JEAN JACOD
ALBERT N. SHIRYAEV

Volume 288

Grundlehren
der mathematischen
Wissenschaften

A Series of
Comprehensive Studies
in Mathematics

LIMIT THEOREMS FOR STOCHASTIC PROCESSES

SECOND EDITION

Springer

Grundlehren der mathematischen Wissenschaften 288

A Series of Comprehensive Studies in Mathematics

Springer-Verlag Berlin Heidelberg GmbH

Jean Jacod
Albert N. Shiryaev

Limit Theorems for Stochastic Processes

Second edition

Springer

Jean Jacod

Université Paris VI
Laboratoire de Probabilités
Tour 56, 3ème etage
4, Place Jussieu
75252 Paris Cedex 05
France

e-mail: jj@ccr.jussieu.fr

Albert N. Shiryaev

Russian Academy of Sciences
Steklov Mathematical Institute
Gubkina 8
119991 Moscow

e-mail: shiryaev@mi.ras.ru

Library of Congress Cataloging-in-Publication Data
Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Jacod, Jean:
Limit theorems for stochastic processes / Jean Jacod ; Albert N. Shiryaev. -
2. ed. - Berlin ; Heidelberg ; New York ; Hong Kong ; London ; Milan ; Paris ;
Tokyo : Springer, 2002
(Grundlehren der mathematischen Wissenschaften ; 288)

Mathematics Subject Classification (2000): 60F17

ISSN 0072-7830

http://www.springer.de

ISBN 978-3-642-07876-7 ISBN 978-3-662-05265-5 (eBook)
DOI 10.1007/978-3-662-05265-5

Originally published by Springer-Verlag Berlin Heidelberg New York in 2003.
Softcover reprint of the hardcover 2nd edition 2003

Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN 10999707 41/3111db - 5 4 3 2 1

To our sons
Olivier, Vincent and Andrei

Limit theorems ...
(by courtesy of Professor A. T. Fomenko of the Moscow State University)

Preface to the Second Edition

Apart from correcting a number of printing mistakes, and some mathematical inaccuracies as well, this second edition contains some new material: indeed, during the fifteen years elapsed since the first edition came out, a large number of new results concerning limit theorems have of course been proved by many authors, and more generally mathematical life has been going on. This gave us the feeling that some of the material in the first edition was perhaps not as important as we thought at the time, while there were some neglected topics which have in fact proved to be very useful in various applications.

So perhaps a totally new book would have been a good thing to write. Our natural laziness prevented us to do that, but we have felt compelled to fill in the most evident holes in this book. This has been done in the most painless way for us, and also for the reader acquainted with the first edition (at least we hope so ...). That is all new material has been added at the end of preexisting chapters.

There are essentially three "new" topics covered below. One concerns stochastic calculus *per se* with a view towards financial mathematics: Section 8 added to Chapter II, which contains a more thorough study of "stochastic exponentials" (or Doléans-Dade exponential) of semimartingales and its inverse, the "stochastic logarithm"; Sections 6 and 7 added to Chapter III, about stochastic integrals of vector-valued integrands with respect to a vector-valued semimartingale (in the finite-dimensional case only). The second topic concerns the so-called "predictable" uniform tightness condition for a sequence of semimartingales, a topic which has reached maturity by now: Section 6 of Chapter VI has been totally rewritten, and Section 6 has been added to Chapter IX and is concerned with the stability results for stochastic differential equations in the light of this uniform tightness condition. Finally, in order to deal with limit theorems for discretized semimartingales on a fixed interval we have added Section 7 to Chapter II (for the prerequisites on this topic) and Section 7 to Chapter IX, where the main and hopefully useful results are presented.

Paris and Moscow
May 2002

Jean Jacod
Albert N. Shiryaev

Introduction

The limit theorems in this book belong to the theory of weak convergence of probability measures on metric spaces.

More precisely, our main aim is to give a systematic exposition of the theory of convergence in law for those stochastic processes that are semimartingales.

The choice of the class of semimartingales as our chief object of study has two reasons. One is that this class is broad enough to accomodate most common processes: discrete-time processes, diffusions, many Markov processes, point processes, solutions of stochastic differential equations, ... Our second reason is that we have in our hands a very powerful tool for studying these processes, namely the stochastic calculus. Since the theory of semimartingales, and related topics as random measures, are not usually associated with limit theorems, we decided to write a rather complete account of that theory, which is covered in the first two chapters. In particular, we devote much space to a careful and detailed exposition of the notion of characteristics of a semimartingale, which extends the well-known "Lévy-Khintchine triplet" for processes with independent increments (drift term, variance of the Gaussian part, Lévy measure), and which plays a particularly important rôle in limit theorems.

The meaning of $X^n \xrightarrow{\mathscr{L}} X$ (that is, the sequence (X^n) of processes converges in law to the process X) is not completely straightforward. The first idea would be to use "finite-dimensional convergence", which says that for any choice $t_1, \ldots, t_p$ of times, then $(X^n_{t_1}, \ldots, X^n_{t_p})$ goes in law to $(X_{t_1}, \ldots, X_{t_p})$. This is usually unsatisfactory because it does not ensure convergence in law of such simple functionals as $\inf(t\colon X^n_t > a)$ or $\sup_{s \leq 1} X^n_s$, etc... In fact, since the famous paper [199] of Prokhorov, the traditional mode of convergence is weak convergence of the laws of the processes, considered as random elements of some functional space. Because semimartingales are right-continuous and have left-hand limits, here the fundamental functional space will always be the "Skorokhod space" $\mathbb{D}$ introduced by Skorokhod in [223]: this space can be endowed with a complete separable metric topology, and $X^n \xrightarrow{\mathscr{L}} X$ will always mean weak convergence of the laws, relative to that topology.

How does one prove that $X^n \xrightarrow{\mathscr{L}} X$?, and in which terms is it suitable to express the conditions? The method proposed by Prokhorov goes as follows:

$$\overset{\text{(i)}}{\left|\begin{array}{l}\text{Tightness of the}\\ \text{sequence } (X^n)\end{array}\right|} + \overset{\text{(ii)}}{\left|\begin{array}{l}\text{Convergence of finite-}\\ \text{dimensional distributions}\end{array}\right|}$$

$$+ \overset{\text{(iii)}}{\left|\begin{array}{l}\text{Characterization of } (X) \text{ by}\\ \text{finite-dimensional distributions}\end{array}\right|} \Rightarrow X^n \xrightarrow{\mathscr{L}} X$$

(as a matter of fact, this is even an equivalence; and of course (iii) is essentially trivial). Sometimes, we will make use of this method. However, it should be emphazised that very often step (ii) is a very difficult (or simply impossible) task to accomplish (with a notable exception concerning the case where the limiting process has independent increments). This fact has led to the development of other strategies; let us mention, for example, the method based upon the "embedding theorem" of Skorokhod, or the "approximation and σ-topological spaces methods" of Borovkov, which allows one to prove weak convergence for large classes of functionals and which are partly based upon (ii). Here we expound the strategy called "martingale method", initiated by Stroock and Varadhan, and which goes as follows:

$$\text{(i)} + \overset{\text{(ii}')}{\left|\begin{array}{l}\text{Convergence of triplets}\\ \text{of characteristics}\end{array}\right|} + \overset{\text{(iii}')}{\left|\begin{array}{l}\text{Characterization of } (X) \text{ by the}\\ \text{triplet of characteristics}\end{array}\right|} \Rightarrow X^n \xrightarrow{\mathscr{L}} X.$$

Here the difficult step is (iii′): we devote a large part of Chapter III to the explicit statement of the problem (called "martingale problem") and to some partial answers.

In both cases, we need step (i): in Chapter VI we develop several tightness criteria especially suited to semimartingales; we also use this opportunity to expose elementary—and less elementary—facts about the Skorokhod topology, in particular for processes indexed by the entire half-line $\mathbb{R}_+$.

The limit theorems themselves are presented in Chapters VII, VIII and IX (the reader can consult [166] for slightly different aspects of the same theory). Conditions insuring convergence always have a similar form, for simple situations (as convergence of processes with independent increments) as well as for more complicated ones (convergence of semimartingales to a semimartingale). Roughly speaking, they say that the triplets of characteristics of X^n converge to the triplet of characteristics of X. As a matter of fact, these conditions are more extensions of two sets of results that are apparently very far apart: those concerning convergence of rowwise independent triangular arrays, as in the book [65] of Gnedenko and Kolmogorov; and those concerning convergence of Markov processes (and especially of diffusion processes, in terms of their coefficients), as in the book [233] of Stroock and Varadhan.

Beside limit theorems, the reader will find apparently disconnected results, which concern absolute continuity for a pair of measures given on a filtered space, and contiguity of sequences of such pairs. In fact, one of our motivations for including such material is that we wanted to give some statistically-oriented applications of our limit theorems (a second motivation is that we indeed find this subject interesting on its own): that is done in Chapter X, where we study convergence of likelihood ratio processes (in particular asymptotic normality) and the so-called "statistical invariance principle" which gives limit theorems under contiguous alternatives.

In order to prepare for these results, we need a rather deep study of contiguity: this is done in Chapter V, in which Hellinger integrals and what we call Hellinger processes are widely used. Hellinger processes are introduced in Chapter IV, which also contains necessary and sufficient conditions for absolute continuity and singularity in terms of the behaviour of those Hellinger processes. Finally, let us mention that some material about convergence in variation is also included in Chapter V.

Within each chapter, the numbering is as follows: 3.4 means statement number 4 in Section 3. When referring to a statement in a previous chapter, say Chapter II, we write II.3.4.

In addition to the usual indexes (Index of Symbols; Index of Terminology), the reader will find in the Index of Topics a reference to all the places in this book where we write about a specific subject: for example, a reader interested only in point processes should consult the Index of Topics first. Finally, all the conditions on the triplets of characteristics which appear in our limit theorems are listed in the Index of Conditions for Limit Theorems.

Parts of this work were performed while one or other author was enjoying the hospitality of the Steklov Mathematical Institute or the Université Pierre et Marie Curie, Paris VI. We are grateful for having had these opportunities.

Paris and Moscow,
June 1987

Jean Jacod
Albert N. Shiryaev

Basic Notation

$\mathbb{R} = (-\infty, +\infty) =$ the set of real numbers, $\mathbb{R}_+ = [0, \infty)$, $\bar{\mathbb{R}} = [-\infty, +\infty]$
$\bar{\mathbb{R}}_+ = [0, \infty]$
$\mathbb{Q} =$ the set of rational numbers, $\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+$
$\mathbb{N} = \{0, 1, 2, \ldots\} =$ the set of integers, $\mathbb{N}^* = \{1, 2, 3, \ldots\}$
$\mathbb{C} =$ the set of complex numbers
$\mathbb{R}^d =$ the Euclidian d-dimensional space
$|x| =$ the Euclidian norm of $x \in \mathbb{R}^d$, or the modulus of $x \in \mathbb{C}$
$x \cdot y =$ the scalar product of $x \in \mathbb{R}^d$ with $y \in \mathbb{R}^d$
$a \vee b = \sup(a, b)$, $a \wedge b = \inf(a, b)$
$x^+ = x \vee 0$, $x^- = (-x) \vee 0$ for $x \in \mathbb{R}$
$1_A =$ the indicator function of the set A
$A^c =$ the complement of the set A
$\varepsilon_a =$ the Dirac measure sitting at point a
a. s. = almost surely
$\lim_{s \uparrow t} = \lim_{s \to t, s \leq t}$, $\lim_{s \uparrow\uparrow t} = \lim_{s \to t, s < t}$
$\lim_{s \downarrow t} = \lim_{s \to t, s \geq t}$, $\lim_{s \downarrow\downarrow t} = \lim_{s \to t, s > t}$
$\otimes =$ tensor product (of spaces, of σ-fields)
$[x] =$ the integer part of $x \in \mathbb{R}_+$
$\mathrm{Re}(x)$, $\mathrm{Im}(x) =$ real and imaginary parts of $x \in \mathbb{C}$
$\ll$ absolute continuity between measures
$\sim$ equivalence between measures
$\perp$ singularity between measures
$\{\cdots\}$ denotes a set

Table of Contents

Chapter I. The General Theory of Stochastic Processes, Semimartingales and Stochastic Integrals

The "General Theory of Stochastic Processes", in spite of its name, encompasses the rather restrictive subject of stochastic processes indexed by $\mathbb{R}_+$. But, within this framework, it expounds deep properties related to the order structure of $\mathbb{R}_+$, and martingales play a central rôle.

By now, there exist several books that give more or less complete accounts on the theory: the basic book [33] of Dellacherie (which however does not deal with stochastic integrals at all), the very complete book [36] of Dellacherie and Meyer, or the book [180] of Métivier... But those may appear as a gigantic investment, for somebody who is not acquainted with the theory beforehand, as might presumably be many of the potential readers of this book. This is why we feel necessary to present a sort of "résumé" that brings out all the needed facts for limit theorems, along the quickest and (hopefully) most painless possible way (although this way is somehow old-fashioned, especially for the presentation of semimartingales and stochastic integrals).

As we wished this book to be as much self-contained as possible, we have provided below all the proofs, with a few exceptions concerning the theory of martingales (regularity of paths, Doob's inequality, Doob's optional theorem), and also two difficult but reasonably well-known results: the Doob-Meyer decomposition of submartingales, and the section theorem (for which we refer to [33] or [36]).

However, despite the fact that all proofs do appear, this chapter is written in the spirit of a résumé, not of a beginner's course: for instance there are almost no examples. So we rather advise the reader to go quickly through the statements (to refresh his mind about notation and definitions) and then to proceed directly to the next chapter.

1. Stochastic Basis, Stopping Times, Optional σ-Field, Martingales

Here are some standard notations to be used in all the book. If $(\Omega, \mathscr{F}, P)$ is a probability space, we denote by $E(X)$ the expectation of any integrable random variable X; if there is some ambiguity as to the measure P, we write $E_P(X)$.

$L^p = L^p(\Omega, \mathscr{F}, P)$, for $p \in [1, \infty)$, is the space of all real-valued random variables X such that $|X|^p$ is integrable, with the usual identification of any two a.s. (= almost surely) equal random variables. Similarly $L^\infty(\Omega, \mathscr{F}, P)$ is the set of all P-essentially bounded real-valued random variables. The corresponding norms are denoted by $\|X\|_{L^p}$.

If $\mathscr{G}$ is a sub-σ-field of $\mathscr{F}$, the conditional expectation of the variable X is well-defined whenever X is integrable or nonnegative or nonpositive, and we denote by $E(X|\mathscr{G})$ any version of it. As a matter of fact, it is also very convenient to use the notion of *generalized conditional expectation*, which is defined for *all* random variables by

1.1 $$E(X|\mathscr{G}) = \begin{cases} E(X^+|\mathscr{G}) - E(X^-|\mathscr{G}) & \text{on the set where } E(|X||\mathscr{G}) < \infty \\ +\infty & \text{elsewhere.} \end{cases}$$

In most cases, $X = Y$ (or $X \leq Y$, etc...) stands for: "$X = Y$ a.s. (almost surely)" (or $X \leq Y$ a.s., etc...).

§ 1a. Stochastic Basis

The reader will immediately notice that our main concern lies in stochastic processes indexed by $\mathbb{R}_+$, or perhaps an interval of $\mathbb{R}_+$. In this case, the theory is built upon what is commonly known as a "stochastic basis", to be recalled below. However, we will occasionally deal with discrete-time processes, that are indexed by $\mathbb{N}$. To help the reader to make the connexion between the two settings, at the end of every section of this chapter we provide an autonomous treatment for the "discrete time": for instance, § 1f of this section provides for the discrete version of what follows.

1.2 **Definition.** A *stochastic basis* is a probability space $(\Omega, \mathscr{F}, P)$ equipped with a *filtration* $\mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}$; here, *filtration* means *increasing* and *right-continuous* family of sub-σ-fields of $\mathscr{F}$ (in other words, $\mathscr{F}_s \subset \mathscr{F}_t$ for $s \leq t$ and $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s$).

By convention, we set: $\mathscr{F}_\infty = \mathscr{F}$ and $\mathscr{F}_{\infty-} = \bigvee_{s \in \mathbb{R}_+} \mathscr{F}_s$. □

The stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$ is also called a *filtered probability space.* In many cases (but not always, as we shall see) it is possible to assume a further property, namely

1.3 **Definition.** The stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is called *complete*, or equivalently is said to *satisfy the usual conditions* if the σ-field $\mathscr{F}$ is P-complete and if every $\mathscr{F}_t$ contains all P-null sets of $\mathscr{F}$. □

It is always possible to "complete" a given stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ as follows:

1.4 $\mathscr{F}^P$ denotes the P-completion of the σ-field $\mathscr{F}$; $\mathscr{N}^P$ denotes the set of all P-null sets of $\mathscr{F}^P$; $\mathscr{F}_t^P$ is the smallest σ-field that contains $\mathscr{F}_t$ and $\mathscr{N}^P$. It is very easy to check that $(\Omega, \mathscr{F}^P, \mathbf{F}^P = (\mathscr{F}_t^P)_{t \in \mathbb{R}_+}, P)$ is a new stochastic basis, called the *completion* of $(\Omega, \mathscr{F}, \mathbf{F}, P)$. □

Let us fix some terminology:

1.5 A *random set* is a subset of $\Omega \times \mathbb{R}_+$. □

1.6 A *process* (or, an *E-valued process*) is a family $X = (X_t)_{t \in \mathbb{R}_+}$ of mappings from Ω into some set E. Unless otherwise stated, E will be $\mathbb{R}^d$ for some $d \in \mathbb{N}^*$. □

A process may, and often will, be considered as a mapping from $\Omega \times \mathbb{R}_+$ into E, via

1.7 $$(\omega, t) \rightsquigarrow X(\omega, t) = X_t(\omega).$$

We shall say indifferently: the process "X", or "(X_t)", or "$(X_t)_{t \in \mathbb{R}_+}$". Each mapping: $t \rightsquigarrow X_t(\omega)$, for a fixed $\omega \in \Omega$, is called a *path*, or a *trajectory*, of the process X.

For example, the indicator function 1_A of a random set A is a process; its paths are the indicator functions of the $\mathbb{R}_+$-sections $\{t: (\omega, t) \in A\}$ of A.

A process X is called *càd* (resp. *càg*, resp. *càdlàg*), for "*c*ontinu *à d*roite" (resp. "*c*ontinu *à g*auche", resp. *c*ontinu *à d*roite avec des *l*imites *à g*auche") in French, if all its paths are right-continuous (resp. are left-continuous, resp. are right-continuous and admit left-hand limits). When X is càdlàg we define two other processes $X_- = (X_{t-})_{t \in \mathbb{R}_+}$ and $\Delta X = (\Delta X_t)_{t \in \mathbb{R}_+}$ by

1.8 $$\begin{cases} X_{0-} = X_0, \quad X_t = \lim_{s \uparrow\uparrow t} X_s \quad \text{for } t > 0 \\ \Delta X_t = X_t - X_{t-} \end{cases}$$

(hence $\Delta X_0 = 0$, which differs from a convention that is sometimes used, as in [183]).

If X is a process and if T is a mapping: $\Omega \to \overline{\mathbb{R}}_+$, we define the "*process stopped at time* T", denoted by X^T, by

1.9 $$X_t^T = X_{T \wedge t}.$$

1.10 A random set A is called *evanescent* if the set $\{\omega: \exists t \in \mathbb{R}_+ \text{ with } (\omega, t) \in A\}$ is P-null; two E-valued processes X and Y are called *indistinguishable* if the random set $\{X \neq Y\} = \{(\omega, t): X_t(\omega) \neq Y_t(\omega)\}$ is evanescent, i.e. if almost all paths of X and Y are the same. □

Note that if X and Y are indistinguishable, one has $X_t = Y_t$ a. s. for all $t \in \mathbb{R}_+$, but the converse is not true. This converse is true, however, when both X and Y are càd, or are càg.

As for random variables, in most cases $X = Y$ (or $X \leq Y$, etc...) for stochastic processes means "up to an evanescent set".

§ 1b. Stopping Times

Let $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis.

1.11 **Definitions.** a) A *stopping time* is a mapping $T: \Omega \to \bar{\mathbb{R}}_+$ such that $\{T \le t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$.

b) If T is a stopping time, we denote by $\mathscr{F}_T$ the collection of all sets $A \in \mathscr{F}$ such that $A \cap \{T \le t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$.

c) If T is a stopping time, we denote by $\mathscr{F}_{T-}$ the σ-field generated by $\mathscr{F}_0$ and all the sets of the form $A \cap \{t < T\}$, where $t \in \mathbb{R}_+$ and $A \in \mathscr{F}_t$. □

One readily checks that $\mathscr{F}_T$ is a σ-field. If $t \in \bar{\mathbb{R}}_+$ and $T(\omega) \equiv t$, then T is a stopping time and $\mathscr{F}_T = \mathscr{F}_t$ (recall that $\mathscr{F}_\infty = \mathscr{F}$ by 1.2); hence the notation $\mathscr{F}_T$ is not ambiguous. Similarly, for $T \equiv t$, one has $\mathscr{F}_{T-} = \mathscr{F}_0$ if $t = 0$, and $\mathscr{F}_{T-} = \bigvee_{s<t} \mathscr{F}_s$ is $t > 0$: hence the notation

1.12 $$\mathscr{F}_{t-} = \begin{cases} \mathscr{F}_0 & \text{if } t = 0 \\ \bigvee_{s<t} \mathscr{F}_s & \text{if } t \in (0, \infty] \text{ (recall 1.2 again for } \mathscr{F}_{\infty-}). \end{cases}$$

The σ-field $\mathscr{F}_t$ is usually interpreted as the set of events that occur before or at time t; then a stopping time is a random time T such that at each time t one may decide whether $T \le t$ or $T > t$ from what one knows up to time t; and $\mathscr{F}_T$ (resp. $\mathscr{F}_{T-}$) is interpreted as the set of events that occur before or at time T (resp. strictly before T).

Now we give a list of well-known and very useful properties of stopping times. All the proofs can be easily provided for by the reader, or may be found in any standard text-book.

1.13 If T is a stopping time and $t \in \mathbb{R}_+$, then $T + t$ is a stopping time. □

1.14 If T is a stopping time, then $\mathscr{F}_{T-} \subset \mathscr{F}_T$ and T is $\mathscr{F}_{T-}$-measurable. □

1.15 If T is a stopping time and if $A \in \mathscr{F}_T$, then

$$T_A(\omega) = \begin{cases} T(\omega) & \text{if } \omega \in A \\ +\infty & \text{if } \omega \notin A \end{cases}$$

is also a stopping time. □

1.16 A mapping $T: \Omega \to \bar{\mathbb{R}}_+$ is a stopping time if and only if $\{T < t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$; in this case, a set $A \in \mathscr{F}$ belongs to $\mathscr{F}_T$ if and only if $A \cap \{T < t\} \in \mathscr{F}_t$ for all $t \in \mathbb{R}_+$ (the right-continuity of the filtration $\mathbf{F}$ is essential for this property). □

1.17 If S, T are two stopping times and if $A \in \mathscr{F}_S$, then $A \cap \{S \le T\} \in \mathscr{F}_T$, $A \cap \{S = T\} \in \mathscr{F}_T$, and $A \cap \{S < T\} \in \mathscr{F}_{T-}$. □

1.18 If (T_n) is a sequence of stopping times, then $S = \bigwedge T_n$ and $T = \bigvee T_n$ are two stopping times, and $\mathscr{F}_S = \bigcap \mathscr{F}_{T_n}$. □

1.19 **Lemma.** *Any stopping time T on the completed stochastic basis $(\Omega, \mathscr{F}^P, \mathbf{F}^P, P)$ is a.s. equal to a stopping time on $(\Omega, \mathscr{F}, \mathbf{F}, P)$.*

Proof. For each $t \in \mathbb{R}_+$ there exists $A_t \in \mathscr{F}_t$ such that $A_t = \{T < t\}$ a.s. (see 1.4). Then $T'(\omega) = \inf(s \in \mathbb{Q}_+ : \omega \in A_s)$ is an $\mathbf{F}$-stopping time (because $\{T' < t\} = \bigcup_{s \in \mathbb{Q}_+, s<t} A_s$ is in $\mathscr{F}_t$) and $T' = T$ a.s. (because $\{T < t\} = \bigcup_{s \in \mathbb{Q}_+, s<t} \{T < s\}$ is a.s. equal to $\{T' < t\}$, for all $t \in \mathbb{R}_+$). □

§ 1c. The Optional σ-Field

Here again, the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is fixed.

1.20 **Definition.** a) A process X is *adapted to the filtration* $\mathbf{F}$ (or, in short, *adapted*) if X_t is $\mathscr{F}_t$-measurable for every $t \in \mathbb{R}_+$.

b) The *optional σ-field* is the σ-field $\mathscr{O}$ on $\Omega \times \mathbb{R}_+$ that is generated by all càdlàg adapted processes (considered as mappings on $\Omega \times \mathbb{R}_+$). □

A process or a random set that is $\mathscr{O}$-measurable is called *optional.*

1.21 **Proposition.** *Let X be an optional process. When considered as a mapping on $\Omega \times \mathbb{R}_+$, it is $\mathscr{F} \otimes \mathscr{R}_+$-measurable. Moreover, if T is a stopping time, then*

a) $X_T 1_{\{T<\infty\}}$ *is $\mathscr{F}_T$-measurable* (hence, X is adapted).

b) *the stopped process X^T is also optional.*

Proof. The set of all processes that are $\mathscr{F} \otimes \mathscr{R}_+$-measurable and meet (a) and (b) for all stopping times is obviously a vector lattice and is stable under pointwise convergence. Thus, by Definition 1.20 and a monotone class argument, it is enough to prove that every càdlàg adapted process X satisfies the claimed properties.

If $n \in \mathbb{N}^*$ we define a new process X^n by putting $X^n_t = X_{k/2^n}$ for $t \in [(k-1)/2^n, k/2^n)$, where $k \in \mathbb{N}^*$. Since

$$\{X^n \in B\} = \bigcup_{k \in \mathbb{N}^*} \left[\{\omega : X_{k/2^n}(\omega) \in B\} \times \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \right],$$

we have $\{X^n \in B\} \in \mathscr{F} \otimes \mathscr{R}_+$ for all Borel sets B, hence X^n is $\mathscr{F} \otimes \mathscr{R}_+$-measurable. Since X is càd, the sequence (X^n) converges pointwise to X, which therefore is also $\mathscr{F} \otimes \mathscr{R}_+$-measurable.

Let T be a stopping time, and put $T_n = \infty$ on the set $\{T = \infty\}$ and $T_n = k/2^n$ on the set $\{(k-1)/2^n \le T < k/2^n\}$. Each T_n is obviously a stopping time, and the sequence (T_n) decreases to T. Since

$$\{X_{T_n} \in B\} \cap \{T_n \le t\} = \bigcup_{k \in \mathbb{N}^*, k/2^n \le t} [\{X_{k/2^n} \in B\} \cap \{T_n = k/2^n\}]$$

is in $\mathscr{F}_t$, we obtain that $X_{T_n} 1_{\{T_n < \infty\}}$ is $\mathscr{F}_{T_n}$-measurable. Since X is càd, $X_{T_n} 1_{\{T_n < \infty\}}$ converges to $X_T 1_{\{T < \infty\}}$. Thus it follows from 1.18 that $X_T 1_{\{T < \infty\}}$ is $\mathscr{F}_T$-measurable and we have (a). Finally X^T is also càdlàg by construction, and since $X_t^T = X_t 1_{\{t<T\}} + X_T 1_{\{T \le t\}}$ it follows from what precedes that X^T is adapted: therefore it is optional, and we have (b). □

There exists a characterization of the optional σ-field that differs from the definition 1.20 and gives some insight for this notion. To this effect, let us first introduce what a stochastic interval is: if S, T are two stopping times, one may define four kinds of *stochastic intervals*, that are the following four random sets:

1.22
$$\begin{cases} [\![S, T]\!] = \{(\omega, t): t \in \mathbb{R}_+, S(\omega) \le t \le T(\omega)\} \\ [\![S, T[\![= \{(\omega, t): t \in \mathbb{R}_+, S(\omega) \le t < T(\omega)\} \\ ]\!]S, T]\!] = \{(\omega, t): t \in \mathbb{R}_+, S(\omega) < t \le T(\omega)\} \\ ]\!]S, T[\![= \{(\omega, t): t \in \mathbb{R}_+, S(\omega) < t < T(\omega)\}. \end{cases}$$

Instead of $[\![T, T]\!]$, we write $[\![T]\!]$: that is, $[\![T]\!]$ is the restriction of the graph of the mapping $T: \Omega \to \overline{\mathbb{R}}_+$, to the set $\Omega \times \mathbb{R}_+$, and we abuse the terminology by calling $[\![T]\!]$ *the graph of the stopping time T.*

The process $1_{[\![0,T[\![}$ is càdlàg, and is obviously adapted if and only if T is a stopping time; then by 1.20 we have $[\![0, T[\![\in \mathscr{O}$ for each stopping time T. More generally:

1.23 **Proposition.** *If S, T are two stopping times and if Y is an $\mathscr{F}_S$-measurable random variable, the four processes $Y 1_{[\![S,T]\!]}$, $Y 1_{[\![S,T[\![}$, $Y 1_{]\!]S,T]\!]}$, $Y 1_{]\!]S,T[\![}$ are optional.*

Proof. It is enough to prove the result when Y is the indicator function of a set $A \in \mathscr{F}_S$. Let us consider for example $X = 1_A 1_{]\!]S,T]\!]}$. Then X is the pointwise limit of $X^n = 1_A 1_{[\![S_n, T_n[\![}$ where $S_n = S + 1/n$ and $T_n = T + 1/n$. X^n is càdlàg by construction and, using 1.17 and the fact that $A \in \mathscr{F}_S \subset \mathscr{F}_{S_n}$, we check that X^n is adapted: hence X^n is optional, and thus so is X. The proof for the other kinds of stochastic intervals is the same. □

1.24 **Proposition.** *Every process X that is càg and adapted is optional.*

Proof. For each $n \in \mathbb{N}^*$ define a new process X^n by

$$X^n = \sum_{k \in \mathbb{N}} X_{k/2^n} 1_{[\![k/2^n, (k+1)/2^n[\![}.$$

Proposition 1.23 yields that X^n is optional. Since X is càg, the sequence (X^n) converges pointwise to X, hence X is optional. □

If X is a càdlàg adapted process, it is obvious that X_- is also adapted; hence it follows from 1.24 that:

1.25 **Corollary.** *If X is a càdlàg adapted process, the two processes X_- and ΔX are optional* (recall that $\Delta X = X - X_-$).

1.26 **Remark.** One may also prove the following, stronger, results, which will not be used in this book:

(a) Any càd adapted process is optional;

(b) the σ-field $\mathcal{O}$ is generated by the stochastic intervals $[\![0, T[\![$, where T is any stopping time. □

Next we study hitting times. Firstly, we have a fairly general (and difficult) result, due to Hunt; although it will not be used in this book, we recall it (without proof: see e.g. [33]) because of its theoretical importance.

1.27 **Theorem.** *If A is an optional random set, its début $T(\omega) = \inf(t\colon (\omega, t) \in A)$ is a stopping time relative to the completed filtration $\mathbf{F}^P$ introduced in* 1.4 (or, equivalently, is a. s. equal to a stopping time of the original filtration $\mathbf{F}$, by Lemma 1.19).

In particular if X is an $\mathbb{R}^d$-valued optional process and if B is a Borel subset of $\mathbb{R}^d$, then $T = \inf(t\colon X_t \in B)$ is a stopping time of the completed filtration $\mathbf{F}^P$ (apply 1.27 to the optional random set $A = \{X \in B\}$).

As for us, instead of using the full force of this result, we shall only use a very particular and easy case, namely

1.28 **Proposition.** a) *If X is an $\mathbb{R}^d$-valued adapted càd process and if B is an open subset of $\mathbb{R}^d$, then $T = \inf(t\colon X_t \in B)$ is a stopping time.*

b) *If X is an $\mathbb{R}$-valued adapted càd process with nondecreasing paths and if $a \in \bar{\mathbb{R}}$, then $T = \inf(t\colon X_t \geq a)$ is a stopping time.*

(Here there is no need to complete the filtration, unlike in 1.27).

Proof. a) Since B is open and X is càd, we have

$$\{T < t\} = \bigcup_{s \in \mathbb{Q}_+, s < t} \{X_s \in B\}$$

and since X is adapted, the right-hand side above is in $\mathscr{F}_t$, so the result follows from 1.16.

b) When X is non-decreasing and càd, then $\{T \leq t\} = \{X_t \geq a\}$, which belongs to $\mathscr{F}_t$ because X is adapted: hence the result. □

We end this paragraph by some easy results on the structure of the jumps of a càdlàg adapted process.

1.30 **Definition.** A random set A is called *thin* if it is of the form $A = \bigcup [\![T_n]\!]$, where (T_n) is a sequence of stopping times; if moreover the sequence (T_n) satisfies $[\![T_n]\!] \cap [\![T_m]\!] = \emptyset$ for all $n \neq m$, it is called an *exhausting sequence* for A. □

Of course a thin set is optional and all its sections $\{t: (\omega, t) \in A\}$ are at most countable; conversely one may prove that any optional set whose sections are at most countable is thin in the sense of 1.30: this is a difficult result, that will not be used here (see [33]).

1.31 **Lemma.** *Any thin random set admits an exhausting sequence of stopping times.*

Proof. Let $A = \bigcup [\![T_n]\!]$, where $(T_n)_{n \in \mathbb{N}}$ is a sequence of stopping times. The set $C_n = \bigcap_{0 \le m \le n-1} \{T_m \neq T_n\}$ is in $\mathscr{F}_{T_n}$ by 1.17, hence 1.15 implies that $S_n = (T_n)_{C_n}$ is a stopping time: the sequence (S_n) is thus an exhausting sequence for A. □

1.32 **Proposition.** *If X is a càdlàg adapted process, the random set $\{\Delta X \neq 0\}$ is thin*; an exhausting sequence for this set $\{\Delta X \neq 0\}$ is called a sequence that *exhausts the jumps* of X.

Proof. Let $n \in \mathbb{N}^*$. Put $S(n, 0) = 0$ and define by induction

$$S(n, p + 1) = \inf(t > S(n, p): |X_t - X_{S(n,p)}| > 2^{-n}).$$

Then for n, p fixed we have $S(n, p + 1) = \inf(t: |Y_t| > 2^{-n})$, where

$$Y = (X - X_{S(n,p)}) 1_{[\![S(n,p),\infty[\![}$$

and Y is a càdlàg adapted process (use 1.23). Hence we deduce from 1.28 that $S(n, p)$ is a stopping time. Moreover 1.21 and 1.25 yield that $A(n, p) = \{S(n, p) < \infty, \Delta X_{S(n,p)} \neq 0\}$ is in $\mathscr{F}_{S(n,p)}$, hence by 1.15 each $T(n, p) = S(n, p)_{A(n,p)}$ is also a stopping time. Now, since X is càdlàg we obviously have: $\lim_{p\uparrow\infty} \uparrow S(n, p) = \infty$ and it easily follows that $\{\Delta X \neq 0\} = \bigcup_{n,p \in \mathbb{N}^*} [\![T(n, p)]\!]$, hence the result. □

§ 1d. The Localization Procedure

In this short subsection, we describe a procedure that is used over and over.

1.33 **Definition.** If $\mathscr{C}$ is a class of processes, we denote by $\mathscr{C}_{\text{loc}}$ the *localized class*, defined as such: a process X belongs to $\mathscr{C}_{\text{loc}}$ if and only if there exists an increasing sequence (T_n) of stopping times (depending on X) such that $\lim_{(n)} T_n = \infty$ a. s. and that each stopped process X^{T_n} belongs to $\mathscr{C}$. The sequence (T_n) is called a *localizing sequence* for X (relative to $\mathscr{C}$). □

For instance, if $\mathscr{C}$ is the class of all *bounded processes*, $\mathscr{C}_{\mathrm{loc}}$ is the class of the so-called *locally bounded processes*. If we may anticipate on the next paragraph, if $\mathscr{C}$ is the class of all submartingales, $\mathscr{C}_{\mathrm{loc}}$ will be the class of the so-called *local* submartingales.

Of course, $\mathscr{C} \subset \mathscr{C}_{\mathrm{loc}}$. The localization is most useful for the classes that satisfy the following property (*all* classes of processes encountered in this book will satisfy the next property!)

1.34 **Definition.** A class $\mathscr{C}$ of processes is called *stable under stopping* if for any $X \in \mathscr{C}$ and any stopping time T, the stopped proced X^T belongs to $\mathscr{C}$. □

1.35 **Lemma.** *Let $\mathscr{C}$ and $\mathscr{C}'$ be two classes of processes, that are stable under stopping. Then*

(a) *$\mathscr{C}_{\mathrm{loc}}$ is stable under stopping, and $(\mathscr{C}_{\mathrm{loc}})_{\mathrm{loc}} = \mathscr{C}_{\mathrm{loc}}$.*

(b) *$(\mathscr{C} \cap \mathscr{C}')_{\mathrm{loc}} = \mathscr{C}_{\mathrm{loc}} \cap \mathscr{C}'_{\mathrm{loc}}$.*

Proof. (a) That $\mathscr{C}_{\mathrm{loc}}$ is stable under stopping is trivial. Let $X \in (\mathscr{C}_{\mathrm{loc}})_{\mathrm{loc}}$, and (T_n) be a localizing sequence such that $X^{T_n} \in \mathscr{C}_{\mathrm{loc}}$. For each $n \in \mathbb{N}$ there exists a localizing sequence $(T(n,p))_{p \in \mathbb{N}}$ such that $(X^{T_n})^{T(n,p)} \in \mathscr{C}$, and there exists an integer p_n such that $P(T(n,p_n) < T_n \wedge n) \leq 2^{-n}$.

Put $S_n = T_n \wedge [\bigwedge_{m \geq n} T(m,p_n)]$. Each S_n is a stopping time and since the sequence (T_n) is increasing, then so is the sequence (S_n). One has:

$$P(S_n < T_n \wedge n) \leq \sum_{m \geq n} P(T(m,p_m) < T_n \wedge n)$$
$$\leq \sum_{m \geq n} P(T(m,p_m) < T_m \wedge m) \leq \sum_{m \geq n} 2^{-m} = 2^{-(n-1)}.$$

Because $\lim_{(n)} T_n = \infty$ a. s., it follows that $\lim_{(n)} S_n = \infty$ a. s., and (S_n) is a localizing sequence. Now,

$$X^{S_n} = ((X^{T_n})^{T(n,p_n)})^{S_n}$$

and, since $\mathscr{C}$ is stable under stopping, it follows that $X^{S_n} \in \mathscr{C}$. Hence $X \in \mathscr{C}_{\mathrm{loc}}$.

(b) The inclusion $(\mathscr{C} \cap \mathscr{C}')_{\mathrm{loc}} \subset \mathscr{C}_{\mathrm{loc}} \cap \mathscr{C}'_{\mathrm{loc}}$ is trivial. Conversely, let $X \in \mathscr{C}_{\mathrm{loc}} \cap \mathscr{C}'_{\mathrm{loc}}$, and let (T_n) and (T'_n) be two localizing sequences such that $X^{T_n} \in \mathscr{C}$ and that $X^{T'_n} \in \mathscr{C}'$. Put $S_n = T_n \wedge T'_n$. The sequence (S_n) is increasing and $\lim_{(n)} S_n = \infty$ a. s.; since $\mathscr{C}$ and $\mathscr{C}'$ are stable under stopping, $X^{S_n} = (X^{T_n})^{T'_n} \in \mathscr{C}$ and similarly $X^{S_n} \in \mathscr{C}'$. Therefore $X \in (\mathscr{C} \cap \mathscr{C}')_{\mathrm{loc}}$. □

The first property above means that one cannot iterate the localization procedure and obtain larger and larger classes of processes. Typically, the previous lemma is set to work in the following sort of situation:

"Theorem". Let $\mathscr{C}, \mathscr{C}', \mathscr{C}''$ be three classes that are stable under stopping; associate to each $X \in \mathscr{C}_{\mathrm{loc}} \cap \mathscr{C}'_{\mathrm{loc}}$ another process $Y = \alpha(X)$, with the property that $\alpha(X^T) =$

$(\alpha(X))^T$ for every stopping time. Then if $\alpha(X) \in \mathscr{C}''_{\text{loc}}$ for all $X \in \mathscr{C} \cap \mathscr{C}'$, we also have $\alpha(X) \in \mathscr{C}''_{\text{loc}}$ for all $X \in \mathscr{C}_{\text{loc}} \cap \mathscr{C}'_{\text{loc}}$.

"Method of Proof": Apply 1.35. In a "real" proof, when we encounter a situation of this type we write the ritual sentence: *by localization, we may assume that* $X \in \mathscr{C} \cap \mathscr{C}'$.

§ 1e. Martingales

In this subsection we review a number of properties of martingales, submartingales and supermartingales, that are essentially due to Doob. They are stated without proof (except for the last two properties), the proofs may be found in most standard books (see e.g. [33], [43]).

1.36 **Definition.** A *martingale* (resp. *submartingale*, resp. *supermartingale*) is an *adapted* process X on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, whose P-almost all paths are *càdlàg*, such that every X_t is integrable, and that for $s \le t$:

$$X_s = E(X_t | \mathscr{F}_s) \quad (\text{resp. } X_s \le E(X_t | \mathscr{F}_s), \text{ resp. } X_s \ge E(X_t | \mathscr{F}_s)). \qquad \square$$

1.37 **Remark.** We somehow depart from standard conventions in this definition; namely, the stochastic basis is *not* assumed to be complete. Nevertheless, the subsequent properties are true, as the reader will check by himself (it is very easy), thanks to the following: if X is a submartingale on the complete basis $(\Omega, \mathscr{F}^P, \mathbf{F}^P, P)$, there exists a process X', P-indistinguishable from X, adapted to the (uncomplete) filtration $\mathbf{F}$, and also an $\mathbf{F}$-stopping time T, such that for all ω the path $X_.(\omega)$ is càd everywhere and làd everywhere except at $T(\omega)$, and moreover $P(T < \infty) = 0$. □

1.38 We say that a process X admits a *terminal variable* X_∞ if X_t converges a. s. to a limit X_∞ as $t \uparrow \infty$; in such a case, the variable X_T is (a. s.) well defined for any stopping time T, with $X_T = X_\infty$ on $\{T = \infty\}$. □

1.39 **Theorem.** *Let X be a supermartingale such that there exists an integrable random variable Y with $X_t \ge E(Y | \mathscr{F}_t)$ for all $t \in \mathbb{R}_+$. Then*

a) (Doob's limit Theorem) *X_t converges a. s. to a finite limit X_∞.*

b) (Doob's stopping Theorem) *If S, T are two stopping times, the random variables X_S and X_T are integrable, and $X_S \ge E(X_T | \mathscr{F}_S)$ on the set $\{S \le T\}$. In particular, X^T is again a supermartingale.*

Now we introduce the two following classes of martingales:

1.40 **Definition.** We denote by $\mathscr{M}$ the class of all *uniformly integrable martingales*, that is of all martingales X such that the family of random variables $(X_t)_{t \in \mathbb{R}_+}$ is uniformly integrable. □

1.41 **Definition.** We denote by $\mathscr{H}^2$ the class of all *square-integrable martingales*, that is of all martingales X such that $\sup_{t \in \mathbb{R}_+} E(X_t^2) < \infty$. □

We obviously have $\mathscr{H}^2 \subset \mathscr{M}$. The following theorem will imply that both $\mathscr{M}$ and $\mathscr{H}^2$ are *stable under stopping*.

1.42 **Theorem.** a) *If X is a uniformly integrable martingale, then X_t converges a.s. and in L^1 to a terminal variable X_∞, and $X_T = E(X_\infty | \mathscr{F}_T)$ for all stopping times T. Moreover, X is square-integrable if and only if X_∞ is square-integrable, in which case the convergence $X_t \to X_\infty$ also takes place in L^2.*

b) *If Y is an integrable random variable, there exists a uniformly integrable martingale X, and only one up to an evanescent set, such that $X_t = E(Y | \mathscr{F}_t)$ for all $t \in \mathbb{R}_+$; moreover, $X_\infty = E(Y | \mathscr{F}_{\infty-})$.*

(Observe that no completion of the filtration is needed here).

1.43 **Theorem** (Doob's inequality). *If X is a square-integrable martingale,*

$$E\left(\sup_{t \in \mathbb{R}_+} X_t^2\right) \leq 4 \sup_{t \in \mathbb{R}_+} E(X_t^2) = 4E(X_\infty^2).$$

Here is another, very useful, characterization of the elements of $\mathscr{M}$:

1.44 **Lemma.** *Let X be an adapted càdlàg process, with a terminal random variable X_∞. Then X is a uniformly integrable martingale if and only if for each stopping time T, the variable X_T is integrable and satisfies $E(X_T) = E(X_0)$.*

Proof. The necessary condition comes immediately from 1.42. To prove the sufficient condition, we remark first that X_∞ is integrable by hypothesis. Then if $t \in \mathbb{R}_+$ and $A \in \mathscr{F}_t$, we define the stopping time T by $T = t$ on A and $T = \infty$ on the complement A^c. We have $E(X_T) = E(X_t 1_A) + E(X_\infty 1_{A^c})$ and $E(X_\infty) = E(X_\infty 1_A) + E(X_\infty 1_{A^c})$. Our assumption implies that $E(X_T) = E(X_\infty)$, hence $E(X_t 1_A) = E(X_\infty 1_A)$ by difference. This being true for all $A \in \mathscr{F}_t$, it follows that $X_t = E(X_\infty | \mathscr{F}_t)$. Then one easily deduce that $X \in \mathscr{M}$ from 1.42. □

1.45 **Definition.** A *local martingale* (resp. a *locally square-integrable martingale*) is a process that belongs to the localized class $\mathscr{M}_{\text{loc}}$ (resp. $\mathscr{H}^2_{\text{loc}}$) constructed from $\mathscr{M}$ (resp. $\mathscr{H}^2$) via 1.33. □

1.46 **Definition.** A process X is *of class* (D) if the set of random variables $\{X_T$: T finite-valued stopping time$\}$ is uniformly integrable. □

1.47 **Proposition.** a) *Each martingale is a local martingale* (hence, $\mathscr{M}_{\text{loc}}$ is also the localized class obtained via 1.33 from the class of martingales).

b) *Each uniformly integrable martingale is a process of class* (D).

c) *A local martingale is a uniformly integrable martingale if and only if it is a process of class* (D).

Proof. a) Let X be a martingale, and put $T_n = n$. Then $X_t^{T_n} = E(X_n|\mathscr{F}_t)$ for all $t \in \mathbb{R}_+$ and 1.42 implies that $X^{T_n} \in \mathscr{M}$.

b) The statement follows from 1.42 and from the well-known fact that if $Y \in L^1$, the set of random variables $\{E(Y|\mathscr{G}): \mathscr{G}$ any sub-σ-field of $\mathscr{F}\}$ is uniformly integrable.

c) Only the sufficient condition remains to be proved. Let $X \in \mathscr{M}_{\text{loc}}$ be of class (D), and let (T_n) be a localizing sequence for X. If $s \leq t$,

$$(1) \qquad X_{s \wedge T_n} = X_s^{T_n} = E(X_t^{T_n}|\mathscr{F}_s) = E(X_{t \wedge T_n}|\mathscr{F}_s).$$

The two sequences $(X_{s \wedge T_n})_{n \in \mathbb{N}}$ and $(X_{t \wedge T_n})_{n \in \mathbb{N}}$ are uniformly integrable because X is of class (D), and they converge to X_s and X_t a.s. respectively because $\lim_{(n)} T_n = \infty$ a.s. Hence the convergence is also in L^1, and so it passes through the conditional expectation in (1), which thus yields $X_s = E(X_t|\mathscr{F}_s)$ and X is a martingale. At last, since X is of class (D), it is *a*-fortiori uniformly integrable. □

We end this paragraph by showing, through two examples, the differences between a uniformly integrable martingale, a martingale, and a local martingale.

1.48 **Example.** Let $(Z_n)_{n \in \mathbb{N}^*}$ be a sequence of i.i.d. random variables with $P(Z_n = 1) = P(Z_n = -1) = 1/2$. Put $\mathscr{F}_t = \sigma(Z_p: p \in \mathbb{N}^*, p \leq t)$ and $X_t = \sum_{1 \leq p \leq [t]} Z_p$, where $[t]$ denotes the integer part of $t \in \mathbb{R}_+$. Then X is trivially a martingale, but by the central limit theorem X_t does not converge a.s. as $t \uparrow \infty$: hence X is not uniformly integrable. □

1.49 **Example.** Let $(A_n)_{n \in \mathbb{N}^*}$ be a measurable partition with $P(A_n) = 2^{-n}$ and $(Z_n)_{n \in \mathbb{N}^*}$ be a sequence of random variables that are independent of the A_n's and with $P(Z_n = 2^n) = P(Z_n = -2^n) = 1/2$. Put $\mathscr{F}_t = \sigma(A_n: n \in \mathbb{N}^*)$ if $t \in [0, 1)$ and $\mathscr{F}_t = \sigma(A_n, Z_n: n \in \mathbb{N}^*)$ if $t \in [1, \infty)$. Put

$$Y_n = \sum_{1 \leq p \leq n} Z_p 1_{A_p}$$

$$X_t = \begin{cases} 0 & \text{if } t \in [0, 1) \\ Y_\infty & \text{if } t \in [1, \infty) \end{cases}$$

$$T_n = \begin{cases} +\infty & \text{on the set } \bigcup_{1 \leq p \leq n} A_p \\ 0 & \text{elsewhere.} \end{cases}$$

(T_n) is clearly a sequence of stopping times that increases to $+\infty$. The process X^{T_n} is equal to 0 (resp. Y_n) on $[\![0, 1[\![$ (resp. $[\![1, \infty[\![$) and Y_n is bounded and independent from the σ-field $\mathscr{F}_{1-}$: hence $X^{T_n} \in \mathscr{M}$ and X is a local martingale. However, it is not a martingale, because $X_1 = Y_\infty$ is not integrable. □

§ 1f. The Discrete Case

When the time-set is not $\mathbb{R}_+$ but $\mathbb{N}$, we have a theory that is similar to the previous one, although much simpler. We will very briefly sketch this theory and show how it is connected to the "continuous-time" one.

1. Let us first define what a stochastic basis is, in this setting.

1.50 **Definition.** A *discrete stochastic basis* is a *probability space* $(\Omega, \mathscr{F}, P)$ equipped with a *filtration* $\mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}}$; here, a *filtration* means an *increasing* family of sub-σ-fields of $\mathscr{F}$ (i.e. $\mathscr{F}_n \subset \mathscr{F}_m$ if $n \leq m$). Notice that the right-continuity has no meaning here. □

A *random set* is a subset of $\Omega \times \mathbb{N}$. A *process* is a family $X = (X_n)_{n \in \mathbb{N}}$ of mappings from Ω into some set E, and it can also be viewed as a mapping from $\Omega \times \mathbb{N}$ into E, via:

$$(\omega, n) \rightsquigarrow X(\omega, n) = X_n(\omega).$$

The notions of càd, càg, or càdlàg processes, have no signification here. However, analoguously to 1.8, to each process X we associate two other processes $X_- = (X_{n-})$ and $\Delta X = (\Delta X_n)$ by

1.51
$$\begin{cases} X_{0-} = X_0, \quad X_{n-} = X_{n-1} \quad \text{if } n \geq 1 \\ \Delta X_n = X_n - X_{n-} \end{cases}$$

If X is a process, and T a mapping: $\Omega \to \overline{\mathbb{N}}$, we define the process X^T "stopped at time T" by $X_n^T = X_{T \wedge n}$.

A *stopping time* T and its associated σ-fields $\mathscr{F}_T$ and $\mathscr{F}_{T-}$ are defined exactly like in 1.11, except that here T is a mapping: $\Omega \to \overline{\mathbb{N}}$ and that $\mathbb{R}_+$ is replaced by $\mathbb{N}$. The properties 1.13 to 1.18 are of course valid. Moreover, we have the obvious and useful additional property:

1.52 A mapping $T: \Omega \to \overline{\mathbb{N}}$ is a stopping time if and only if $\{T = n\} \in \mathscr{F}_n$ for each $n \in \mathbb{N}$; in this case, a set $A \in \mathscr{F}$ belongs to $\mathscr{F}_T$ if and only if $A \cap \{T = n\} \in \mathscr{F}_n$ for each $n \in \mathbb{N}$. □

The notion of optionality is rather trivial here:

1.53 **Definition.** The *optional σ-field* is the σ-field $\mathscr{O}$ on $\Omega \times \mathbb{N}$ that is generated by all *adapted* processes, that is all processes X such that X_n is $\mathscr{F}_n$-measurable for each $n \in \mathbb{N}$. □

Most of the results of § c have no interesting counterpart in the discrete case. Let us however mention 1.21 (which is much easier to prove than in the continuous case), 1.25 that is trivial, and 1.27 that is very easy to prove here:

1.54 Theorem. *If A is an optional random set, its début $T(\omega) = \inf(n \in \mathbb{N}: (\omega, n) \in A)$ is a stopping time* (no completion required here).

Proof. The hypothesis means that the process $X = 1_A$ is adapted; hence the result follows from the equality:

$$\{T \leq n\} = \bigcup_{0 \leq p \leq n} \{X_p = 1\}. \qquad \square$$

Finally, the notion of localization, and all the definitions and theorems of § e about martingales are valid without changes (except that $\mathbb{R}_+$ and $\bar{\mathbb{R}}_+$ are everywhere replaced by $\mathbb{N}$ and $\bar{\mathbb{N}}$, and of course we can drop the "càdlàg" assumption in Definition 1.36). Note that even in the discrete case we have the three notions of a uniformly integrable martingale, of a martingale, and of a local martingale, and the examples 1.48 and 1.49 may easily be translated into the discrete case to show that a martingale may not be uniformly integrable, or that a local martingale may not be a martingale (see also 1.64 below).

2. Now we wish to show that the discrete case actually reduces to a particular case of the general one. To this effect, we consider a *discrete stochastic basis* $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}}, P)$.

We associate to $\mathscr{B}$ a "continuous" stochastic basis $\mathscr{B}'$ as follows:

1.55 $$\mathscr{B}' = (\Omega, \mathscr{F}, \mathbf{F}' = (\mathscr{F}'_t)_{t \in \mathbb{R}_+}, P), \quad \text{with } \mathscr{F}'_t = \mathscr{F}_n \quad \text{for } t \in [n, n+1).$$

In particular, we have:

1.56 $$\mathscr{F}'_n = \mathscr{F}_n \quad \text{if } n \in \mathbb{N}, \quad \mathscr{F}'_{n-} = \mathscr{F}'_{n-1} = \mathscr{F}_{n-1} \quad \text{if } n \in \mathbb{N}^*.$$

1.57 Lemma. *Any $\mathscr{B}$-stopping time T is also a $\mathscr{B}'$-stopping time, and we have $\mathscr{F}'_T = \mathscr{F}_T$ and $\mathscr{F}'_{T-} = \mathscr{F}_{T-}$.*

Proof. For each $A \in \mathscr{F}$ we have $A \cap \{T \leq t\} = A \cap \{T \leq n\}$ when $t \in [n, n+1)$. Hence by 1.56, $A \cap \{T \leq t\} \in \mathscr{F}'_t$ for all $t \in \mathbb{R}_+$ if and only if $A \cap \{T \leq n\} \in \mathscr{F}_n$ for all $n \in \mathbb{N}$: this proves that a $\mathscr{B}$*-stopping time* T is a $\mathscr{B}'$-stopping time and that $\mathscr{F}'_T = \mathscr{F}_T$.

If $A \in \mathscr{F}'_t$ and $t \in [n, n+1)$, we have $A \in \mathscr{F}_n$ and $A \cap \{t < T\} = A \cap \{n < T\}$; since $\mathscr{F}'_0 = \mathscr{F}_0$, it proves that $\mathscr{F}'_{T-} \subset \mathscr{F}_{T-}$. The converse inclusion is proved similarly. $\square$

1.58 Lemma. *Let T' be a $\mathscr{B}'$-stopping time, and put $T = n$ if $n \leq T' < n+1$, $T = \infty$ if $T' = \infty$. Then T is a $\mathscr{B}$-stopping time and $\mathscr{F}'_{T'} = \mathscr{F}_T$, and $\mathscr{F}'_{T'-} \supset \mathscr{F}_{T-}$* (note that in general we do not have $\mathscr{F}'_{T'-} = \mathscr{F}_{T-}$; the proof, which is similar to that of 1.57, is left to the reader).

Now, let X be a process on $\mathscr{B}$. We associate to it a process X' on $\mathscr{B}'$ as follows:

1.59 $$X'_t = X_n \quad \text{if } t \in [n, n+1).$$

Note that X' is càdlàg, and the following statements are obvious:

1.60 X is $\mathscr{B}$-adapted if and only if X' is $\mathscr{B}'$-adapted.

1.61 If T and T' are like in 1.58, we have $(X^T)' = X'^{T'}$.

1.62 The process X is of class (D) on $\mathscr{B}$ if and only if X' is of class (D) on $\mathscr{B}'$.

1.63 The process X is a martingale (resp. a supermartingale, resp. a uniformly integrable martingale, resp. a local martingale) on $\mathscr{B}$ if and only if X' is a martingale (resp. a supermartingale, resp. a uniformly integrable martingale, resp. a local martingale) on $\mathscr{B}'$.

These facts show why the discrete case is indeed "included" within the continuous one. For instance, 1.60 and the property of the process X' given by 1.59 to be automatically càdlàg explain why "optional" and "adapted" mean the same thing on the basis $\mathscr{B}'$.

3. Here is an exception to what we just wrote above: the following does not easily reduce to a property in continuous time.

1.64 **Proposition.** *Let X be an adapted process on $\mathscr{B}$. Then X is a local martingale, if and only if:*

(i) *X_0 is integrable, and*

(ii) *for all $n \in \mathbb{N}^*$, $E(|X_n| \,|\, \mathscr{F}_{n-1}) < \infty$ a.s. and $E(X_n | \mathscr{F}_{n-1}) = X_{n-1}$.*

Remember that $E(\cdot | \mathscr{F}_{n-1})$ is the "extended" conditional expectation. So in (ii) the fact that $E(X_n|\mathscr{F}_{n-1}) = X_{n-1}$ automatically implies that $E(|X_n|\,|\mathscr{F}_{n-1}) < \infty$ (for clarity, we prefer to explicitly state the two conditions). But $E(X_n|\mathscr{F}_{n-1}) = X_{n-1}$ does *not* imply that X_n is integrable; indeed, integrability for all X_n plus $E(X_n|\mathscr{F}_{n-1}) = X_{n-1}$ for $n \in \mathbb{N}^*$ are necessary and sufficient for X to be a *martingale.*

Proof. a) *Necessary condition*: Let (T_n) be a localizing sequence of stopping times for the local martingale X. Then $X_0 = X_0^{T_n}$ is integrable, and $E(X_p^{T_n}|\mathscr{F}_{p-1}) = X_{p-1}^{T_n}$ for all $p \in \mathbb{N}^*$: therefore $E(X_p|\mathscr{F}_{p-1}) = X_{p-1}$ on the $\mathscr{F}_{p-1}$-measurable set $\{T_n > p-1\}$, and since $\bigcup_n \{T_n > p-1\} = \Omega$ we obtain (ii).

b) *Sufficient condition*: Assume (i) and (ii) and set $T_n = \inf(p: \sum_{1 \le k \le p+1} E(|X_k|\,|\mathscr{F}_{n-1}) \ge n)$. Then $\{T_n \ge p\}$ clearly belongs to $\mathscr{F}_{p-1}$, hence T_n is a stopping time. Moreover,

$$E(|X_p^{T_n}|) = E(|X_{T_n \wedge p}|) \le E(|X_0|) + n < \infty$$

and (ii) yields $E(X_p^{T_n}|\mathscr{F}_{p-1}) = X_{p-1}^{T_n}$ because $\{T_n \ge p\} \in \mathscr{F}_{p-1}$. Hence X^{T_n} is a martingale, and X is a local martingale because $T_n \uparrow \infty$ as $n \uparrow \infty$ by (ii) again. $\square$

2. Predictable σ-Field, Predictable Times

§ 2a. The Predictable σ-Field

The meaning of the following notion of "predictable σ-field" is perhaps not immediately apparent, but we will see very clearly what it means in the discrete case in § 2e below. For the time being, we start with a continuous time stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

2.1 **Definition.** The *predictable σ-field* is the σ-field $\mathscr{P}$ on $\Omega \times \mathbb{R}_+$ that is generated by all càg adapted processes (considered as mappings on $\Omega \times \mathbb{R}_+$). □

Proposition 1.24 yields that $\mathscr{P} \subset \mathscr{O}$. A process or a random set that is $\mathscr{P}$-measurable is called *predictable.*

2.2 **Theorem.** *The predictable σ-field is also generated by any one of the following collections of random sets:*

(i) $A \times \{0\}$ *where* $A \in \mathscr{F}_0$, *and* $[\![0, T]\!]$ *where T is any stopping time;*
(ii) $A \times \{0\}$ *where* $A \in \mathscr{F}_0$, *and* $A \times (s, t]$, *where* $s < t$, $A \in \mathscr{F}_s$;

Proof. Let $\mathscr{P}'$ and $\mathscr{P}''$ be the σ-fields respectively generated by the sets in (i) and (ii).

Since the indicator functions of the sets showing in (i) are adapted and càg processes, we have $\mathscr{P}' \subset \mathscr{P}$.

If $A \in \mathscr{F}_s$ and $s < t$, we have $A \times (s, t] =]\!]s_A, t_A]\!]$ with the notation 1.15, which implies that s_A and t_A are two stopping times; hence $]\!]s_A, t_A]\!] = [\![0, t_A]\!] \backslash [\![0, s_A]\!] \in \mathscr{P}'$, and it follows that $\mathscr{P}'' \subset \mathscr{P}'$.

Let X be a càg adapted process, and set for $n \in \mathbb{N}^*$:

$$X^n = X_0 1_{[\![0]\!]} + \sum_{k \in \mathbb{N}} X_{k/2^n} 1_{]\!]k/2^n, (k+1)/2^n]\!]}$$

It is obvious that X^n is a process that is $\mathscr{P}''$-measurable, and the sequence (X^n) converges pointwise to X because X is càg; hence X is $\mathscr{P}''$-measurable, and it follows that $\mathscr{P} \subset \mathscr{P}''$. □

2.3 **Remark.** One might also prove that $\mathscr{P}$ is generated by all adapted processes that have continuous paths (that will not be used here). □

2.4 **Proposition.** *If X is a predictable process and if T is a stopping time,* a) $X_T 1_{\{T<\infty\}}$ *is* $\mathscr{F}_{T-}$*-measurable,* b) *the stopped process X^T is also predictable.*

Proof. The set of all processes that satisfy (a) and (b) is a vector lattice and is stable under pointwise convergence. On the other hand, the collection of all random sets appearing in 2.2(ii) is a Boolean algebra. Hence by 2.2 and a monotone class argument, it is enough to prove that if X is the indicator function of any one of the sets of 2.2(ii), it meets (a) and (b), and this is trivial. □

2.5 **Proposition.** *If S, T are two stopping times and if Y is an $\mathscr{F}_S$-measurable random variable, the process $Y 1_{]\!]S,T]\!]}$ is predictable* (immediate from the definition of predictability, since this process is adapted and càg).

2.6 **Proposition.** *If X is a càdlàg adapted process, then X_- is a predictable process; if moreover X is predictable, then ΔX is predictable* (again immediate from the definition, since X_- is adapted and càg).

§ 2b. Predictable Times

2.7 **Definition.** A *predictable time* is a mapping $T: \Omega \to \overline{\mathbb{R}}_+$ such that the stochastic interval $[\![0, T[\![$ is predictable. □

Every predictable time T is a stopping time: indeed $[\![T, \infty[\![\in \mathscr{P} \subset \mathscr{O}$, so the càdlàg process $X = 1_{[\![T,\infty[\![}$ is adapted, while $\{T \leq t\} = \{X_t = 1\}$. Notice also that if T is a predictable time, then $[\![T]\!] \in \mathscr{P}$ (use $[\![T]\!] = [\![0, T]\!] \backslash [\![0, T[\![$ and 2.2); moreover, if T is a stopping time and if $[\![T]\!] \in \mathscr{P}$, then T is predictable (use $[\![0, T[\![= [\![0, T]\!] \backslash [\![T]\!]$ and 2.2 again).

Here is a list of properties of predictable times, which should be compared to properties 1.13 to 1.18 of stopping times.

2.8 If T is a stopping time and $t > 0$, then $T + t$ is a predictable time, because $[\![0, T + t[\![= \bigcup_{(n)} \left[\!\!\left[0, T + \frac{n-1}{n} t\right]\!\!\right]$ is in $\mathscr{P}$ (recall that $T - t$ is not even a stopping time, in general). □

2.9 **Proposition.** *Let (T_n) be a sequence of predictable times.*

a) *$T = \bigvee T_n$ is a predictable time.*

b) *If $S = \bigwedge T_n$ and $\bigcup_n \{S = T_n\} = \Omega$, then S is a predictable time.*

Proof. (a) We have $[\![0, T[\![= \bigcup_n [\![0, T_n[\![$, which is predictable by hypothesis, hence T is a predictable time.

(b) The hypothesis $\bigcup_n \{S = T_n\} = \Omega$ implies $[\![0, S[\![= \bigcap_n [\![0, T_n[\![$, which again is predictable, hence S is a predictable time. □

In 2.9(b), the property $S = \bigwedge T_n$ is not enough for S to be predictable. For example, let S be a stopping time that is not predictable (we will see later that such an S actually exists!); then $T_n = S + 1/n$ is predictable by 2.8, and $S = \bigwedge T_n$.

2.10 **Proposition.** *Let T be a predictable time and $A \in \mathscr{F}_{T-}$. The time T_A defined in 1.15 is predictable.*

Proof. Note that $T_{A\cup B} = T_A \wedge T_B$ and that $T_{A\cap B} = T_A \vee T_B$; therefore the collection $\mathscr{A} = \{A \in \mathscr{F}\colon T_A$ is a predictable time$\}$ is stable under countable union and countable intersection, by 2.9. If $A \in \mathscr{A}$, then $[\![T, T_A[\![\in \mathscr{P}$; since $[\![0, T_{A^c}[\![= [\![0, \infty[\![\backslash [\![T, T_A[\![$ we obtain that T_{A^c} also is predictable. Hence $\mathscr{A}$ is a σ-field.

It is thus enough to prove that $\mathscr{A}$ contains $\mathscr{F}_0$ and all sets of the form $A = B \cap \{t < T\}$, with $B \in \mathscr{F}_t$. If $A \in \mathscr{F}_0$, then $[\![0, T_A[\![= [\![0, T[\![\cup (A^c \times \mathbb{R}_+)$ which belongs to $\mathscr{P}$, hence $A \in \mathscr{A}$. If $A = B \cap \{t < T\}$ with $B \in \mathscr{F}_t$, we have $A \in \mathscr{F}_t$ and $[\![0, T_{A^c}[\![= [\![0, T[\![\cup (A \times (t, \infty))$, which belongs to $\mathscr{P}$ as well, hence $A^c \in \mathscr{A}$, hence $A \in \mathscr{A}$ and the proof is finished. □

2.11 **Proposition.** *Let S be a predictable time, and $A \in \mathscr{F}_{S-}$, and T be a stopping time. Then $A \cap \{S \le T\} \in \mathscr{F}_{T-}$.*

Proof. We have $A \cap \{S \le T\} = \{S_A \le T < \infty\} \cup (A \cap \{T = \infty\})$. By 2.10 the process $X = 1_{[\![S_A, \infty[\![}$ is predictable, hence $\{S_A \le T < \infty\} = \{X_T 1_{\{T<\infty\}} = 1\}$ is in $\mathscr{F}_{T-}$ by 2.4.

It remains to prove that $A \cap \{T = \infty\} \in \mathscr{F}_{T-}$. Since $A \in \mathscr{F}_{\infty-}$ it is enough to prove this inclusion when $A \in \mathscr{F}_t$ for some $t \in \mathbb{R}_+$. In that case, $A \cap \{T = \infty\} = A \cap \{t < T\} \cap \{T = \infty\}$, which belongs to $\mathscr{F}_{T-}$ by 1.11 and 1.14. □

As a simple corollary of the previous results, we obtain the following complement to Proposition 2.5:

2.12 **Proposition.** *Let S, T be two stopping times, and Y be a random variable. Then*

a) *If T is predictable and Y is $\mathscr{F}_S$-measurable, $Y 1_{]\!]S, T[\![}$ is predictable.*

b) *If S is predictable and Y is $\mathscr{F}_{S-}$-measurable, $Y 1_{[\![S, T]\!]}$ is predictable.*

c) *If S, T are predictable and Y is $\mathscr{F}_{S-}$-measurable, $Y 1_{[\![S, T[\![}$ is predictable.*

Proof. (a) follows from the fact that $Y 1_{]\!]S, T[\![} = (Y 1_{]\!]S, T]\!]}) 1_{[\![0, T[\![}$ and from 2.5. Similarly, (c) follows from (b).

It remains to prove (b). It is enough to prove the claim when Y is the indicator function of a set $A \in \mathscr{F}_{S-}$. In this case, $Y 1_{[\![S, T]\!]} = 1_{[\![S_A, T]\!]}$ and $[\![S_A, T]\!] = [\![0, T]\!] \backslash [\![0, S_A[\![$ is predictable by 2.10. □

2.13 **Proposition.** *Let T be a stopping time, which is the début $T(\omega) = \inf(t\colon (\omega, t) \in A)$ of a predictable random set A. If $[\![T]\!] \subset A$, then T is a predictable time.*

Proof. The hypothesis $[\![T]\!] \subset A$ implies that $[\![T]\!] = A \cap [\![0, T]\!]$, which is predictable because $A \in \mathscr{P}$ and T is a stopping time. Hence T is a predictable time (see after 2.7). □

This should be compared to 1.27 or 1.28; the claim is usually wrong if $[\![T]\!] \not\subset A$: for example any stopping time T is the début of the predictable random set

$]\!]T, \infty[\![$, without necessarily being predictable itself! When the stochastic basis is *complete*, the début of A is automatically a stopping time (see 1.27).

Now we state two difficult and closely related results, which may be found in [33].

2.14 **Theorem** (Predictable Section Theorem). *Suppose that the stochastic basis is complete. Let A be a predictable random set. For every $\varepsilon > 0$ there exists a predictable time T such that $[\![T]\!] \subset A$ and that*

$$P(\omega\colon T(\omega) = \infty \text{ and there exists } t \in \mathbb{R}_+ \text{ with } (\omega, t) \in A) \leq \varepsilon.$$

2.15 **Theorem.** a) *If (T_n) is a sequence of stopping times increasing to T, such that $T_n < T$ on $\{T > 0\}$, then T is a predictable time* (this part is trivial, because our assumptions imply $]\!]0, T[\![= \bigcup_n]\!]0, T_n]\!]$).

b) *Assume that the stochastic basis is complete. If T is a predictable time*, there exists a sequence (T_n) *of stopping times increasing to T, and such that $T_n < T$ on* $\{T > 0\}$.

The claim (b) is not true when the basis is not complete. However, due to 1.19, we deduce that in all cases (complete or not):

2.16 If T is a predictable time, there is an increasing sequence (T_n) of stopping times, such that $T_n < T$ a.s. on $\{T > 0\}$ and $\lim_n T_n = T$ a.s. The sequence (T_n) is called an *announcing sequence* for T. □

This leads us to investigate more closely the relationships between a stochastic basis and its completion (see 1.4). The following complements 1.19:

2.17 **Lemma.** a) *Any predictable time T on the completed stochastic basis $(\Omega, \mathscr{F}^P, \mathbf{F}^P, P)$ is a.s. equal to a predictable time on the original basis.*

b) *Any predictable process X relative to $\mathbf{F}^P$ is indistinguishable from a predictable process relative to* $\mathbf{F}$.

Proof. (a) We reproduce the proof of [36]. Let (T_n) be a sequence of $\mathbf{F}^P$-stopping times, increasing to T and with $T_n < T$ on $\{T > 0\}$ (see 2.15). Let T_n' be an $\mathbf{F}$-stopping time with $T_n' = T_n$ a.s. (by 1.19), and $T' = \sup_n T_n'$, and $A_n = \{0 < T' \neq T_n'\}$, and $T_n'' = (T_n')_{A_n}$, and $S_n = n \wedge (\sup_{m \leq n} T_m'')$. By 1.15, 1.17 and 1.18, S_n is an $\mathbf{F}$-stopping time. The sequence (S_n) clearly increases to a limit S, and $S_n < S$ on $\{S > 0\}$ by construction, hence S is $\mathbf{F}$-predictable. Moreover $T_n < T$ on $\{T > 0\}$, hence $P(A_n) = 0$, hence $T_n'' = T_n$ a.s., hence $S_n = n \wedge T_n$ a.s. and finally we deduce that $S = T$ a.s.

(b) It suffices to prove the result when $X = 1_B$ is the indicator function of a predictable random set. If the random set B is of the form 2.2(ii), the result is trivial (because any set in $\mathscr{F}_t^P$ is a.s. equal to some set in $\mathscr{F}_t$) A monotone class argument then yields the result. □

The section theorem is typically used as such:

2.18 **Proposition.** a) *Let A be a predictable random set. If every predictable time T such that $[\![T]\!] \subset A$ is a.s. infinite, then A is evanescent.*

b) *If X and Y are two predictable processes that satisfy $X_T = Y_T$ a.s. on $\{T < \infty\}$ for all predictable times T, then X and Y are indistinguishable.*

Proof. (b) follows from (a) applied to $A = \{X \neq Y\}$. To prove (a), suppose that A is not evanescent. Thus $\varepsilon := P(\exists t \text{ with } (\omega, t) \in A)$ is positive, and 2.14 yields an $\mathbf{F}^P$-predictable time T with $[\![T]\!] \subset A$ and $P(T < \infty) \geq \varepsilon/2$. By 2.17, $T = S$ a.s. for some predictable time S, relative to $\mathbf{F}$. Let $B = \{\omega: (\omega, S(\omega)) \in A\}$, which belongs to $\mathscr{F}_{S-}$ by 2.4, so $S' = S_B$ is again $\mathbf{F}$-predictable by 2.10, while $[\![S']\!] \subset A$ by construction, so $S' = \infty$ a.s. Moreover, $S = T$ a.s. yields $P(B^c) = 0$, hence $S' = T$ a.s., hence $P(S' < \infty) \geq \varepsilon/2$, which brings a contradiction. □

2.19 **Remark.** As a matter of fact, Theorem 2.14 is true even if the stochastic basis is not complete (see [36]). □

§ 2c. Totally Inaccessible Stopping Times

Now we introduce a class of stopping times that are "orthogonal" in a sense to all predictable times.

2.20 **Definition.** A stopping time T is called *totally inaccessible* if $P(T = S < \infty) = 0$ for all predictable times S. □

2.21 If T is a totally inaccessible stopping time and if S is a stopping time such that $[\![S]\!] \subset [\![T]\!]$ (i.e. $S = T$ on the set $\{S < \infty\}$), then S is also totally inaccessible (a trivial property). □

2.22 **Theorem.** *Let T be a stopping time. There exist a sequence (S_n) of predictable times and a unique* (up to a P-null set) *$\mathscr{F}_T$-measurable subset A of $\{T < \infty\}$, such that the stopping time T_A is totally inaccessible, and that the stopping time T_{A^c} satisfies $[\![T_{A^c}]\!] \subset \bigcup [\![S_n]\!]$.*

T_A is called the *totally inaccessible* part of T, and T_{A^c} its *accessible part.* They are uniquely determined, up to a P-null set. The sequence (S_n) above is of course not unique, and it may be so chosen (as we will see later) that the graphs of the S_n's are pairwise disjoint. If T is totally inaccessible (resp. predictable), its totally inaccessible part is T (resp. $+\infty$) and its accessible part is $+\infty$ (resp. T).

Proof. For any finite family $\{S_i\}$ of predictable times, we set $B(\{S_i\}) = \bigcup_i \{T = S_i < \infty\}$. Then $B(\{S_i\}) \in \mathscr{F}_T$ and the class $\mathscr{B}$ of all sets of the form $B(\{S_i\})$ is stable under finite unions and finite intersections. Let us denote by B a version of the essential supremum of the class $\mathscr{B}$, and $A = \{T < \infty\} \backslash B$.

As is well known, B is the union of a countable family of elements of $\mathscr{B}$: hence there exists a double sequence $(\{S(n,i)\}_{i \le p_n})_{n \in \mathbb{N}}$ of predictable times such that $B = \bigcup_{n \in \mathbb{N}} \bigcup_{i \le p_n} \{T = S(n,i) < \infty\}$, and $[\![T_{A^c}]\!] = [\![T_B]\!] \subset \bigcup_{n \in \mathbb{N}} \bigcup_{i \le p_n} [\![S(n,i)]\!]$.

It remains to prove that T_A is totally inaccessible. If it were not, there would exist a predictable time S with $P(T_A = S < \infty) > 0$. But $P(B^c \cap B(\{S\})) = P(A \cap \{T = S < \infty\}) = P(T_A = S < \infty) > 0$, which contradicts the fact that B is the essential supremum of $\mathscr{B}$.

Finally, the uniqueness of A is obvious. □

As a first application of this theorem, we will deduce some properties of thin sets (see 1.30) that are predictable.

2.23 **Lemma.** a) *If A is a predictable thin set, there is a sequence (T_n) of predictable times whose graphs are pairwise disjoint, such that $[\![T_n]\!] \subset A$ and that $A \backslash \bigcup_n [\![T_n]\!]$ is evanescent.*

b) *If moreover the stochastic basis is complete, we can choose the T_n's so that $A = \bigcup_n [\![T_n]\!]$ (in other words, A admits an exhausting sequence of predictable times).*

Proof. (i) Let (T_n) be a sequence of stopping times exhausting the thin set A, and denote by T'_n and T''_n, respectively, the accessible and the totally inaccessible parts of T_n. 2.22 yields a sequence $(S(n,p))_{p \in \mathbb{N}}$ of predictable times such that $[\![T'_n]\!] \subset \bigcup_p [\![S(n,p)]\!]$. We set $A' = A \cap (\bigcup_{n,p} [\![S(n,p)]\!])$, which belongs to $\mathscr{P}$.

Let S be a predictable time with $[\![S]\!] \subset A \backslash A'$, hence $[\![S]\!] \subset \bigcup_n [\![T''_n]\!]$. Then Definition 2.20 yields that $S = \infty$ a.s. Thus the section Theorem 2.18 implies that $A \backslash A'$ is evanescent.

(ii) Now we re-arrange the "double" sequence $\{S(n,p)\}_{n,p}$ into a sequence $(R_n)_{n \ge 1}$. We set $C_n = \bigcap_{1 \le m \le n-1} \{R_m \ne R_n\}$ and $D_n = C_n \cap \{\omega: (\omega, R_n(\omega)) \in A\}$, so $D_n \in \mathscr{F}_{R_n-}$ by 2.4 and 2.11, and $R'_n = (R_n)_{D_n}$ is also a predictable time. Then clearly $A' = \bigcup_{n \ge 1} [\![R'_n]\!]$ and the $[\![R'_n]\!]$ are pairwise disjoint, so we have proved (a).

(iii) In order to prove (b), it suffices to show that a stopping time T which is a.s. infinite is indeed a predictable time, provided the basis is complete: indeed since $T''_n = \infty$ a.s. by (i), each T''_n will be predictable, and the sequence (T''_n) exhausts the set $A \backslash A'$: combining with (ii) will yield the result. Then, put $S_n = n \wedge (T - 1/n)^+$, and observe that $\{S_n \le t\}$ belongs to $\mathscr{F}_0 = \mathscr{F}_0^P$ for all t, so S_n is a stopping time, which clearly announces T: therefore T is a predictable time. □

2.24 **Proposition.** *If X is a càdlàg predictable process, there is a sequence of predictable times that exhausts the jumps of X. Furthermore, $\Delta X_T = 0$ a.s. on $\{T < \infty\}$ for all totally inaccessible time T.*

Proof. Let (T_n) be a sequence of stopping times that exhausts the jumps of X (see 1.32). For each $m \in \mathbb{N}$ set $B_m = \left[\frac{1}{m-1}, \frac{1}{m}\right)$ $\left(\text{with } \frac{1}{0} = +\infty\right)$, and $C_{n,m} = \{|\Delta X_{T_n}| \in B_m\}$. Then $C_{n,m} \in \mathscr{F}_{T_n}$, hence $S(n,m) = (T_n)_{C_{n,m}}$ is a stopping time and $Y(m) = \sum_n 1_{[\![S(n,m),\infty[\![}$ is an adapted càdlàg process (taking values in $\mathbb{N}$, because X is càdlàg). Thus 1.30 yields that $R(m,q) := \inf(t: Y(m)_t \geq q)$ (for $q \in \mathbb{N}^*$) is a stopping time; moreover, $R(m,q)$ is the début of the predictable set $]\!]R(m,q-1),\infty[\![\cap \{|\Delta X| \in B_m\}$ (with $R(m,0) = 0\}$ and clearly $[\![R(m,q)]\!]$ is included into this set: hence 2.13 yields that $R(m,q)$ is a predictable time. Obviously the sequence $\{R(m,q)\}_{m \in \mathbb{N}, q \in \mathbb{N}^*}$ exhausts the jumps of X.

Finally, $\{\Delta X_T \neq 0, T < \infty\} = \bigcup_{m,q} \{T = R(m,q) < \infty\}$, hence the last result follows from Definition 2.20. □

2.25 **Definition.** A càdlàg process X is called *quasi-left continuous* if $\Delta X_T = 0$ a.s. on the set $\{T < \infty\}$ for every predictable time T. □

2.26 **Proposition.** *Let X be a càdlàg adapted process. There is equivalence between the conditions:*

(a) *X is quasi-left-continuous;*

(b) *there exists a sequence of totally inaccessible stopping times that exhausts the jumps of X;*

(c) *for any increasing sequence (T_n) of stopping times with limit T, we have* $\lim X_{T_n} = X_T$ *a.s. on the set $\{T < \infty\}$.*

Proof. Since there is a sequence of stopping times that exhausts the jumps of X, the equivalence (a)$\Leftrightarrow$(b) immediately follows from 2.22 and 2.25.

Suppose that X is not quasi-left continuous: there exists a predictable time T, which is announced by a sequence (T_n), and such that $P(\Delta X_T \neq 0, T < \infty) > 0$. But $\lim X_{T_n} = X_{T-}$ a.s. on $\{0 < T < \infty\}$, hence $\lim X_{T_n} \neq X_T$ a.s. on the set $\{\Delta X_T \neq 0, T < \infty\}$ and this contradicts (c). Hence we have the implication (c)$\Rightarrow$(a).

Conversely, suppose that (c) fails for some sequence (T_n), and put $S_n = (T_n)_{\{T_n < T\}}$, $S = T_A$ where $A = \bigcap \{T_n < T\}$. Then S is a predictable time announced by the sequence (S_n). Moreover, we obviously have $\{\lim X_{T_n} \neq X_T, T < \infty\} = \{\Delta X_S \neq 0, S < \infty\}$. Hence our assumption implies that $P(\Delta X_S \neq 0, S < \infty) > 0$, and this contradicts (a). Thus, we have the implication (a)$\Rightarrow$(c). □

§ 2d. Predictable Projection

In all this subsection we suppose again that a stochastic basis is given.

2.27 **Lemma.** *If X is a local martingale, then $E(X_T|\mathscr{F}_{T-}) = X_{T-}$ on the set $\{T < \infty\}$ for all predictable times T* (X_T is not necessarily integrable, but here we use the generalized conditional expectation of 1.1).

Proof. Let (T_n) be a localizing sequence for X. Since $\{T \le T_n\} \in \mathscr{F}_{T-}$ by 1.17, we have $E(X_T|\mathscr{F}_{T-}) = E((X^{T_n})_T|\mathscr{F}_{T-})$ and $X_{T-} = (X^{T_n})_{T-}$ on $\{T \le T_n\}$. Hence it is enough to prove the result for each stopped process X^{T_n}.

In other words, we can and will assume that $X \in \mathscr{M}$. The predictable time T is announced by a sequence (S_n) of stopping times. We deduce from 1.39 or 1.42 that $X_{S_n} = E(X_T|\mathscr{F}_{S_n})$ (because $S_n \le T$). The convergence theorem 1.42a applied to the discrete-time uniformly integrable martingale $(X_{S_n}, \mathscr{F}_{S_n})$ shows that $X_{S_n} \to E(X_T|\bigvee_{(n)} \mathscr{F}_{S_n})$ a. s.

$S_n < T$ on the set $\{T > 0\}$, so 1.17 yields $\mathscr{F}_{S_n} \subset \mathscr{F}_{T-}$. On the other hand, the definition of $\mathscr{F}_{T-}$ and the equality $A \cap \{t < T\} = \bigcup_n [A \cap \{t \le S_n\}]$ a.s. on $\{T > 0\}$ yields that $\mathscr{F}_{T-} \subset \bigvee_n \mathscr{F}_{S_n}$ up to the P-null sets. Hence $\mathscr{F}_{T-} = \bigvee_n \mathscr{F}_{S_n}$ up to the P-null sets. Thus we have proved that $X_{S_n} \to E(X_T|\mathscr{F}_{T-})$ a. s. But since (S_n) announces T, we have $X_{S_n} \to X_{T-}$ a.s. on $\{T < \infty\}$, and so we deduce the claim. □

2.28 **Theorem.** a) *Let X be an $\overline{\mathbb{R}}$-valued and $\mathscr{F} \otimes \mathscr{R}_+$-measurable process. There exists a $(-\infty, \infty]$-valued process, called the predictable projection of X and denoted by ${}^pX$, that is determined uniquely up to an evanescent set by the following two conditions*:

(i) *it is predictable,*

(ii) $({}^pX)_T = E(X_T|\mathscr{F}_{T-})$ *on* $\{T < \infty\}$ *for all predictable times T.*

b) *Moreover, if T is any stopping time,*

2.29 $$ {}^p(X^T) = ({}^pX)1_{[\![0,T]\!]} + X_T 1_{]\!]T,\infty[\![}. $$

c) *Moreover, if ${}^pX$ is finite-valued and if X' is a $(-\infty, \infty]$-valued predictable process,*

2.30 $$ {}^p(XX') = X'\,{}^p(X). $$

As a matter of fact, the predictable projection is usually defined only for measurable processes that are bounded, or nonnegative (see e.g. [33]). The notion introduced above is indeed an *extended predictable projection*, which bears the same relation relative to the ordinary predictable projection than the extended conditional expectation does relative to the ordinary conditional expectation.

Proof. 1) The uniqueness, up to an evanescent set, immediately follows from Proposition 2.18.

2) For the existence, we consider first the case of bounded processes. Let $\mathscr{H}$ be the collection of all $\mathscr{F} \otimes \mathscr{R}_+$-measurable bounded processes X for which an associated process ${}^pX$ meeting (i) and (ii) exists. $\mathscr{H}$ is clearly a vector space, stable under pointwise convergence of uniformly bounded processes: indeed if $X(n)$ is such a sequence, then ${}^pX = \limsup_n {}^pX(n)$ meets (i) and (ii) relatively to the limit process $X = \lim_n X(n)$. Hence by a monotone class argument, it is enough to

prove the result for every process X of the form $X_t(\omega) = 1_A(\omega)1_{[u,v)}(t)$, $A \in \mathscr{F}$, $0 \leq u < v$.

Let X be such a process, and consider the bounded martingale M which meets $M_t = E(1_A|\mathscr{F}_t)$. Then ${}^pX := M_- 1_{]\!]u,v]\!]}$ is predictable, and by 2.27:

$$\begin{aligned}({}^pX)_T &= M_{T-}1_{\{u \leq T < v\}} = E(M_T|\mathscr{F}_{T-})1_{\{u \leq T < v\}} \\ &= E(1_A|\mathscr{F}_{T-})1_{\{u \leq T < v\}} = E(X_T|\mathscr{F}_{T-})\end{aligned}$$

on $\{T < \infty\}$ if T is a predictable time. Therefore X belongs to $\mathscr{H}$, and we deduce that $\mathscr{H}$ is indeed the set of all bounded measurable processes.

Finally, if $X \in \mathscr{H}$, the claims (b) and (c) are obvious.

3) Suppose now that X is $\mathscr{F} \otimes \mathscr{R}_+$-measurable and nonnegative, and put $X(n) = X \wedge n$. By 2), ${}^pX(n)$ exists and satisfies (i) and (ii). Moreover $X(n) \leq X(n+1)$, so it follows from the uniqueness that ${}^pX(n) \vee {}^pX(n+1)$ is again a version of the predictable projection of $X(n+1)$. In other words, we may choose versions of ${}^pX(n)$ such that ${}^pX(n) \leq {}^pX(n+1)$. If we put ${}^pX = \lim_n \uparrow {}^pX(n)$, it is immediate to check that ${}^pX$ meets (i) and (ii), as well as (b) and (c).

4) Consider now the general case of an $\overline{\mathbb{R}}$-valued process X. Then ${}^pX$, as defined below, obviously meets (i), (ii), (b), (c):

$${}^pX = \begin{cases} {}^p(X^+) - {}^p(X^-) & \text{on the random set } \{{}^p|X| < \infty\} \\ +\infty & \text{elsewhere.} \end{cases}$$ □

Since X_- is predictable whenever X is càdlàg and adapted, we can rephrase 2.27 as such:

2.31 **Corollary.** *If X is a local martingale, then ${}^p(X) = X_-$ and ${}^p(\Delta X) = 0$.*

We end this subsection by a last notion, of secondary importance.

2.32 **Definition.** The *predictable support* of a measurable random set A is the predictable set $A' = \{{}^p(1_A) > 0\}$, which is defined up to an evanescent set. □

2.33 From 2.28, A' is characterized as such: this is the only (up to an evanescent set) predictable set such that, for every predictable time T, $A \cap [\![T]\!]$ is evanescent if and only if $A' \cap [\![T]\!]$ is evanescent. □

For example, consider the situation of Theorem 2.22, and put $B = [\![T]\!]$. Then the predictable support B' of B is the smallest predictable set that satisfies $[\![T_{A^c}]\!] \subset B'$. More generally:

2.34 **Proposition.** *The predictable support of a thin optional random set is again a thin random set.*

Proof. Let $A = \bigcup [\![T_n]\!]$, where (T_n) is a sequence of stopping times. Let T_n' be the accessible part of T_n, and $(S(n,p))_{p \in \mathbb{N}}$ a sequence of predictable times such that

$[\![T_n']\!] \subset \bigcup_{p\in\mathbb{N}} [\![S(n,p)]\!]$. Finally, let A' be a version of the predictable support of A, and set $B(n,p) = \{\omega: (\omega, S(n,p)(\omega)) \in A'\}$ (so $B(n,p) \in \mathscr{F}_{S(n,p)-}$ by 2.4) and $S'(n,p) = S(n,p)_{B(n,p)}$, which is a predictable time by 2.10. Finally, set $A'' = \bigcup_{n,p} [\![S'(n,p)]\!]$, which is a predictable thin set.

$A'' \subset A'$ by construction. Conversely, let S be a predictable time with $[\![S]\!] \subset A' \backslash A''$. Then

$$P(S < \infty) = P(({}^p 1_A)_S > 0, S < \infty) = P((1_A)_S > 0, S < \infty) \leq \sum_{n\in\mathbb{N}} P(S = T_n < \infty)$$

$$= \sum_{n\in\mathbb{N}} P(S = T_n' < \infty) \leq \sum_{n,p\in\mathbb{N}} P(S = S(n,p) < \infty) = 0$$

(the last equality comes from $[\![S]\!] \subset A' \backslash A''$). Then by 2.18 this implies that $A'' \backslash A'$ is evanescent, that is A'' is also a version of the predictable support of A. □

2.35 **Proposition.** *Let X be a càdlàg adapted process. Then X is quasi-left-continuous if and only if the predictable support of the random set $\{\Delta X \neq 0\}$ is evanescent. In this case, ${}^p X = X_-$* (all this immediately follows from the definition 2.25).

§ 2e. The Discrete Case

We will now examine how the notion of predictability translates in the discrete-time setting of § 1f.

1. We start with a discrete stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n\in\mathbb{N}}, P)$. Of course the notion of càg processes is empty, so Definition 2.1 does not specialize easily to the discrete case; however, characterization 2.2(ii) has an immediate translation:

2.36 **Definition.** The *predictable σ-field* is the σ-field $\mathscr{P}$ on $\Omega \times \mathbb{N}$ that is generated by all processes X such that X_0 is $\mathscr{F}_0$-measurable and that X_n is $\mathscr{F}_{n-1}$-measurable for all $n \in \mathbb{N}^*$. □

Proposition 2.4 is easy to prove, and Proposition 2.6 is completely trivial and reads as follows: if X is an adapted process, then X_- (defined by 1.51) is predictable; if moreover X is predictable, so is ΔX.

Let us turn to the predictable times. We could start with Definition 2.7, but it is easier to remark that if T is a random time and if $X = 1_{[\![0,T[\![}$, then $X_n = 1_{\{T>n\}}$: hence we can give the following definition, which is equivalent to 2.7 in the discrete setting:

2.37 **Definition.** A *predictable time* is a mapping $T: \Omega \to \bar{\mathbb{N}}$ such that $\{T = 0\} \in \mathscr{F}_0$ and that $\{T \leq n\} \in \mathscr{F}_{n-1}$ for all $n \in \mathbb{N}^*$. □

It is clear that a predictable time is a stopping time. The properties 2.8 to 2.11 are of course valid (and easy to prove here); note that in 2.9(b) the assumption $\bigcup_{(n)} \{S = T_n\} = \Omega$ is always satisfied, hence the infimum of any sequence of predictable times is predictable.

It may also be interesting to notice that T is a predictable time if and only if it is a stopping time, for a new filtration $\mathbf{F}^-$ defined by $\mathscr{F}_0^- = \mathscr{F}_0$ and $\mathscr{F}_n^- = \mathscr{F}_{n-1}$ for $n \in \mathbb{N}^*$.

The following statement generalizes at the same time Theorem 2.13 and the section Theorem 2.14:

2.38 **Theorem.** *Let A be a predictable random set. Then its début $T(\omega) = \inf(n \in \mathbb{N}: (\omega, n) \in A)$ is a predictable time and satisfies $[\![T]\!] \subset A$* (that is, T satisfies the conditions of 2.14 with $\varepsilon = 0$).

Proof. That $[\![T]\!] \subset A$ is evident. That T is predictable follows from the equality $\{T \leq n\} = \bigcup_{0 \leq p \leq n} \{(1_A)_p = 1\}$. □

Theorem 2.15 is *not* true in the discrete case, because if (T_n) increases to T and if T and the T_n's are $\bar{\mathbb{N}}$-valued, then $T_n = T$ for n big enough on the set $\{T < \infty\}$. The natural translation of 2.15 in the discrete case is the following statement, which is evident:

2.39 **Theorem.** *A random time T is predictable if and only if $\{T = 0\} \in \mathscr{F}_0$ and if $S = (T - 1)^+$ is a stopping time.* The stopping time S plays the same rôle here than the announcing sequence (T_n) in 2.15.

The notions introduced in § c have no interest here. For instance if we define a totally inaccessible time by 2.20, we have:

2.40 The only totally inaccessible stopping time is $T \equiv \infty$.

Similarly, if a quasi-left-continuous process is defined by 2.25,

2.41 A process X is quasi-left-continuous if and only if $X_n = X_0$ a. s. for every $n \in \mathbb{N}$.

The notion of predictable projection is very simple here. Actually, Theorem 2.28 is valid, and its proof is trivial, once noticed the following property:

2.42 Let X be an adapted process, and define ${}^pX$ by:

$$({}^pX)_0 = X_0, \, ({}^pX)_n = E(X_n | \mathscr{F}_{n-1}) \quad \text{for } n \in \mathbb{N}^*$$

(with the extended conditional expectation introduced in 1.1). Then ${}^pX$ satisfies conditions 2.28(i), (ii). □

Corollary 2.31 is valid, with again a trivial proof (using 2.42). Finally, the notion of the predictable support of a random set has no interest.

2. Now we wish to show how the discrete case actually reduces to a particular case of the general one. To the discrete stochastic basis $\mathscr{B}$ above we associate the "continuous" stochastic basis $\mathscr{B}'$ by 1.55. Then

2.43 A $\mathscr{B}$-stopping time T is predictable for $\mathscr{B}$ if and only if it is predictable for $\mathscr{B}'$, with the announcing sequence $T_n = \left(T - \frac{1}{n}\right)^+$ if $T \le n$ and $T_n = n$ if $T > n$ (compare to 1.57).

2.44 A process X on $\mathscr{B}$ is predictable if and only if the associated process X' (by 1.59) is predictable on $\mathscr{B}'$. For every process X on $\mathscr{B}$, one has $({}^pX)' = {}^p(X')$.

Using these statements, the reader will be able to check that the results concerning the basis $\mathscr{B}$, that have been quickly reviewed above, are corollaries of the corresponding results concerning the basis $\mathscr{B}'$.

3. Increasing Processes

§ 3a. Basic Properties

We first introduce a few notation. The stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is fixed throughout.

3.1 **Definition.** We denote by $\mathscr{V}^+$ (resp. $\mathscr{V}$) the set of all real-valued processes A that are *càdlàg, adapted, with* $A_0 = 0$, and whose each path $t \rightsquigarrow A_t(\omega)$ is *non-decreasing* (resp. has a *finite variation* over each finite interval $[0, t]$). □

We abbreviate by calling a process belonging to $\mathscr{V}^+$ (resp. $\mathscr{V}$) an *adapted increasing process* (resp. an *adapted process with finite-variation*). Observe that if $A \in \mathscr{V}^+$ it admits a terminal variable A_∞ (see 1.38) that takes its values in $\overline{\mathbb{R}}_+$:

3.2 $$A_\infty = \lim_{t \uparrow \infty} A_t.$$

Let $A \in \mathscr{V}$. We denote by "Var(A)" the *variation process* of A, that is the process such that $\mathrm{Var}(A)_t(\omega)$ is the total variation of the function $s \rightsquigarrow A_s(\omega)$ on the interval $[0, t]$. Of course, $\mathrm{Var}(A) = A$ if $A \in \mathscr{V}^+$.

3.3 **Proposition.** *Let $A \in \mathscr{V}$. There exists a unique pair (B, C) of adapted increasing processes such that $A = B - C$ and* $\mathrm{Var}(A) = B + C$ (hence, $\mathrm{Var}(A) \in \mathscr{V}^+$, and $\mathscr{V} = \mathscr{V}^+ \ominus \mathscr{V}^+$). *Moreover if A is predictable, then B, C and* $\mathrm{Var}(A)$ *are also predictable.*

Proof. By a pathwise argument, we obtain the existence of a unique pair (B, C) of processes which are càdlàg, with $B_0 = C_0 = 0$, with non-decreasing paths, and such that $A = B - C$ and $\operatorname{Var}(A) = B + C$: namely, those processes are $B = (A + \operatorname{Var}(A))/2$ and $C = B - A$. Thus it remains only to prove that $\operatorname{Var}(A)$ is adapted (resp. predictable when A is so).

By definition of $\operatorname{Var}(A)$, we have:

$$\operatorname{Var}(A)_t(\omega) = \lim_{(n)} \sum_{1 \le k \le n} |A_{tk/n}(\omega) - A_{t((k-1)/n)}(\omega)|,$$

which clearly is $\mathscr{F}_t$-measurable. Thus $\operatorname{Var}(A)$ is adapted.

Suppose now that A is predictable. $\operatorname{Var}(A)_-$ is càg and adapted, hence predictable, as well as the jump process $\Delta[\operatorname{Var}(A)]$ which is equal to $|\Delta A|$. Thus $\operatorname{Var}(A) = \operatorname{Var}(A)_- + \Delta[\operatorname{Var}(A)]$ is predictable. $\square$

Let again $A \in \mathscr{V}$. For each $\omega \in \Omega$, the path: $t \rightsquigarrow A_t(\omega)$ is the distribution function of a *signed measure* (a positive measure if A is increasing) on $\mathbb{R}_+$, that is finite on each interval $[0, t]$, and that is finite on $\mathbb{R}_+$ if and only if $\operatorname{Var}(A)_\infty(\omega) < \infty$. We denote this measure by $dA_t(\omega)$.

We say that $dA \ll dB$, where $A, B \in \mathscr{V}$, if the measure $dA_t(\omega)$ is *absolutely continuous* with respect to the measure $dB_t(\omega)$ for almost all $\omega \in \Omega$.

Let $A \in \mathscr{V}$ and let H be an optional process. By 1.21, $t \rightsquigarrow H_t(\omega)$ is Borel. Hence we can define the *integral process*, denoted by $H \cdot A$ or by $\int_0^\cdot H_s \, dA_s$, as follows:

3.4 $$H \cdot A_t(\omega) = \begin{cases} \displaystyle\int_0^t H_s(\omega)\, dA_s(\omega) & \text{if } \displaystyle\int_0^t |H_s(\omega)|\, d[\operatorname{Var}(A)]_s(\omega) < \infty \\ +\infty & \text{otherwise.} \end{cases}$$

3.5 **Proposition.** *Let $A \in \mathscr{V}$ (resp. $\mathscr{V}^+$) and let H be an optional process (resp. nonnegative), such that the process $B = H \cdot A$ is finite-valued. Then $B \in \mathscr{V}$ (resp. $\mathscr{V}^+$) and $dB \ll dA$. If moreover A and H are predictable, then B is predictable.*

Proof. If B is finite-valued, it is clearly càdlàg, with $B_0 = 0$, and its paths have finite variation over finite intervals. It remains to prove that B is adapted. Fix $t \in \mathbb{R}_+$; if $\mu(\omega, ds) = dA_s(\omega) 1_{\{s \le t\}}$, for each ω we have a measure $\mu(\omega, \cdot)$ on $[0, t]$, such that $\mu(\cdot, I)$ is $\mathscr{F}_t$-measurable for every interval I; on the other hand, $(\omega, s) \rightsquigarrow H_s(\omega)$ is $\mathscr{F}_t \otimes \mathscr{B}([0, t])$-measurable on $\Omega \times [0, t]$: hence an application of Fubini's Theorem for transition measures shows that B_t is $\mathscr{F}_t$-measurable.

It is obvious that $dB \ll dA$, and that $B \in \mathscr{V}^+$ when A is increasing and when H is nonnegative. Finally, assume that A and H are predictable. Then $\Delta B = H \Delta A$ is predictable, and $B = B_- + \Delta B$ is also predictable. $\square$

This proposition admits a converse which, because of the measurability requirements, is surprisingly difficult to prove; we defer this converse to the end of the subsection (see Proposition 3.13).

3.6 $\mathscr{A}^+$ is the set of all $A \in \mathscr{V}^+$ that are *integrable*: $E(A_\infty) < \infty$.

3.7 $\mathscr{A}$ is the set of all $A \in \mathscr{V}$ that have *integrable variation*: $E(\mathrm{Var}(A)_\infty) < \infty$. By 3.3 we have $\mathscr{A} = \mathscr{A}^+ \ominus \mathscr{A}^+$.

3.8 $\mathscr{A}^+_{\mathrm{loc}}$ and $\mathscr{A}_{\mathrm{loc}}$ are the localized classes constructed from $\mathscr{A}^+$ and $\mathscr{A}$ (recall 1.33); a process in $\mathscr{A}^+_{\mathrm{loc}}$ (resp. $\mathscr{A}_{\mathrm{loc}}$) is called a *locally integrable adapted increasing process* (resp. an *adapted process with locally integrable variation*).

Note that the classes $\mathscr{V}$, $\mathscr{V}^+$, $\mathscr{A}$, $\mathscr{A}^+$ are stable under stopping, and that the localized classes $\mathscr{V}_{\mathrm{loc}}$ and $\mathscr{V}^+_{\mathrm{loc}}$ are just $\mathscr{V}$ and $\mathscr{V}^+$ themselves. We also have the following obvious inclusions:

3.9 $$\mathscr{A}_{\mathrm{loc}} = \mathscr{A}^+_{\mathrm{loc}} \ominus \mathscr{A}^+_{\mathrm{loc}}, \quad \mathscr{A}^+ \subset \mathscr{A}^+_{\mathrm{loc}} \subset \mathscr{V}^+, \quad \mathscr{A} \subset \mathscr{A}_{\mathrm{loc}} \subset \mathscr{V}.$$

3.10 **Lemma.** *Let A be a predictable process that belongs to $\mathscr{V}$. Then there exists a localizing sequence (S_n) of stopping times such that $\mathrm{Var}(A)_{S_n} \le n$ a.s., and in particular $A \in \mathscr{A}_{\mathrm{loc}}$.*

Proof. Let $B = \mathrm{Var}(A)$, which is predictable by 3.3. Set $T_n = \inf(t\colon B_t \ge n)$, which is a predictable time (see 2.13). Let $\{S(n,p)\}_{p \in \mathbb{N}}$ be a sequence of stopping times announcing T_n (see 2.16). There exists an integer p_n such that $P(S(n,p_n) < T_n - 1) \le 2^{-n}$. Finally, set $S_n = \sup_{m \le n} S(m, p_m)$. Then $S_n < \sup_{m \le n} T_m = T_n$ a.s., hence $B_{S_n} \le n$ a.s. Moreover $\lim_n T_n = \infty$, and one checks easily that the sequence (S_n), which is increasing, tends a.s. to $+\infty$. □

3.11 **Lemma.** *Any local martingale that belongs to $\mathscr{V}$ also belongs to $\mathscr{A}_{\mathrm{loc}}$.*

Proof. By localization, it is enough to prove that any $X \in \mathscr{M} \cap \mathscr{V}$ belongs to $\mathscr{A}_{\mathrm{loc}}$. Put $T_n = \inf(t\colon \mathrm{Var}(X)_t > n)$, so the stopping times T_n increase to $+\infty$. For each $t \in \mathbb{R}_+$, $|X_{t-}| \le \mathrm{Var}(X)_{t-}$ and $\Delta[\mathrm{Var}(X)]_t = |\Delta X_t| \le |X_{t-}| + |X_t|$; so $\mathrm{Var}(X)_{T_n} \le 2n + |X_{T_n}|$ and since X_{T_n} is integrable by 1.42, $\mathrm{Var}(X)_{T_n}$ is also integrable. It follows that $\mathrm{Var}(X) \in \mathscr{A}^+_{\mathrm{loc}}$, hence $X \in \mathscr{A}_{\mathrm{loc}}$. □

3.12 **Lemma.** *Let $A \in \mathscr{A}$, and let M be a bounded martingale, and let T be a stopping time. Then $E(M_T A_T) = E(M \cdot A_T)$. If moreover A is predictable, then $E(M_T A_T) = E(M_- \cdot A_T)$.*

Proof. By 3.3 we have $A = B - C$ with $B, C \in \mathscr{A}^+$ (and predictable if A is so); and if the result is true for B and C, it is also evidently true for A: that is, we may and will suppose that $A \in \mathscr{A}^+$.

Let $C_t = \inf(s\colon A_s \ge t)$. Then by 1.28, C_t is a stopping time. Furthermore, it is a predictable time when A is predictable (see 2.13). Moreover, $\{C_t \le s\} = \{A_s \ge t\}$, and thus

(1) $$\int_0^\infty H_s(\omega)\, dA_s(\omega) = \int_0^\infty H_{C_s}(\omega) 1_{\{C_s(\omega) < \infty\}}\, ds$$

for each function $s \rightsquigarrow H_s(\omega)$ of the form $H_s(\omega) = 1_{[0,t]}(s)$. A monotone class argument yields (1) for every bounded measurable function $s \rightsquigarrow H_s(\omega)$, or equivalently for every bounded measurable process H. Then

$$E(M_T A_T) = E\left(\int_0^\infty M_T 1_{\{s \leq T\}}\, dA_s\right)$$

$$= E\left(\int_0^\infty M_T 1_{\{C_s < \infty\}} 1_{\{C_s \leq T\}}\, ds\right) \qquad \text{(apply (1) to } H_s = M_T 1_{\{s\leq T\}}\text{)}$$

$$= \int_0^\infty ds E(M_T 1_{\{C_s<\infty\}} 1_{\{C_s \leq T\}}) \qquad \text{(apply Fubini's Theorem)}$$

$$\text{(2)} \qquad = \int_0^\infty ds E(M_{C_s} 1_{\{C_s<\infty\}} 1_{\{C_s \leq T\}}) \qquad (T \text{ is a stopping time, } M \in \mathscr{M})$$

$$= E\left(\int_0^\infty M_{C_s} 1_{\{C_s < \infty\}} 1_{\{C_s \leq T\}}\, ds\right) \qquad \text{(apply Fubini's Theorem)}$$

$$\text{(3)} \qquad = E\left(\int_0^\infty M_s 1_{\{s \leq T\}}\, dA_s\right) = E(M \cdot A_T) \qquad \text{(apply (1) to } H_s = M_s 1_{\{s\leq T\}}\text{)}.$$

Hence we have obtained the result when A is optional. Suppose now that A is predictable. Then C_s is predictable, and 2.27 implies that in (2) above we may replace M_{C_s} by $M_{(C_s)-}$. Then instead of (3), and if we apply (1) to $H_s = M_{s-} 1_{\{s\leq T\}}$, we obtain

$$= E\left(\int_0^\infty M_{s-} 1_{\{s \leq T\}}\, dA_s\right) = E(M_- \cdot A_T). \qquad \square$$

3.13 **Proposition.** *Let A, $B \in \mathscr{V}$ (resp. $\mathscr{V}^+$) be such that $dB \ll dA$. Then there exists an optional process (resp. nonnegative) H such that $B = H \cdot A$ up to an evanescent set. If moreover A and B are predictable, one may choose H to be predictable.*

Proof. (i) We first prove the result when A and B belong to $\mathscr{A}^+$. For each bounded optional process H, we put

$$\text{(1)} \qquad m(H) = E(H \cdot A_\infty), \quad m'(H) = E(H \cdot B_\infty),$$

thus defining two positive finite measures m and m' on $(\Omega \times \mathbb{R}_+, \mathscr{O})$. If $D \in \mathscr{O}$ satisfies $m(D) = 0$, we have $1_D \cdot A_\infty = 0$ a. s., hence $1_D \cdot B_\infty = 0$ a. s. because $dB \ll dA$, hence $m'(D) = 0$. Thus m' is absolutely continuous with respect to m, and we denote by H a nonnegative optional process that is a version of the Radon-Nikodym derivative of m' with respect to m. Put $B' = H \cdot A$. We will now prove that B and B' are indistinguishable, which is the desired result in the optional case.

For this, it is sufficient to prove that $B'_t = B_t$ a.s. for every $t \in \mathbb{R}_+$ (because B and B' are càdlàg), or equivalently that $E(B'_t 1_F) = E(B_t 1_F)$ for every $t \in \mathbb{R}_+, F \in \mathscr{F}_t$, because B and B' are adapted. Let M be the element of $\mathscr{M}$ that satisfies $M_s = E(1_F|\mathscr{F}_s)$ for every $s \in \mathbb{R}_+$. M is bounded, and we have

$$
\begin{aligned}
(2) \qquad E(B_t 1_F) = E(M_t B_t) &= E(M \cdot B_t) && \text{(apply 3.12)} \\
&= E((HM) \cdot A_t) && \text{(by definition of } H\text{)} \\
&= E(M \cdot B'_t) = E(M_t B'_t) && \text{(again 3.12)} \\
&= E(B'_t 1_F),
\end{aligned}
$$

which yields the result.

(ii) Suppose now that $A, B \in \mathscr{A}^+$ are predictable. We can reproduce the previous proof, with the following changes: define m and m' by (1) for all bounded predictable processes, thus defining two measures m and m' on $(\Omega \times \mathbb{R}_+, \mathscr{P})$. Then take a Radon-Nikodym derivative H that is predictable, thus $B' = H \cdot A$ is predictable. Then in (2) we have $E(M_- \cdot B_t)$ by the preceding lemma; by definition of H this is equal to $E((HM_-) \cdot A_t) = E(M_- \cdot B'_t)$, and this is equal to $E(M_t B'_t)$ again by the lemma: hence the result.

(iii) Suppose now that $A, B \in \mathscr{V}^+$, and put $T_n = \inf(t\colon A_t + B_t \geq n)$ and

$$A(n) = 1_{[\![T_{n-1}, T_n[\![} \cdot A, \; B(n) = 1_{[\![T_{n-1}, T_n[\![} \cdot B.$$

T_n is a stopping time, and is predictable when A and B are so. Then $A(n)$, $B(n) \in \mathscr{A}^+$, and they are predictable when A and B are so, and obviously $dB(n) \ll dA(n)$. Then there an optional nonnegative process $H(n)$ which is predictable when A and B are so, such that $H(n) \cdot A(n)$ and $B(n)$ are indistinguishable. Then $H = \sum H(n) 1_{[\![T_{n-1}, T_n[\![}$ is nonnegative, optional (resp. predictable), and $B = H \cdot A$ up to an evanescent set.

(iv) Finally, suppose that $A, B \in \mathscr{V}$. Let A', A'', B', $B'' \in \mathscr{V}^+$ be such that $A = A' - A''$, $\mathrm{Var}(A) = A' + A''$, $B = B' - B''$, $\mathrm{Var}(B) = B' + B''$, and be predictable if A, B are so. Then dA', dA'', dB', dB'' are absolutely continuous with respect to $d[\mathrm{Var}(A)]$ and by (iii) there exist nonnegative finite-valued processes H', H'', K', K'', predictable if A, B are so, such that $A' = H' \cdot \mathrm{Var}(A)$, $A'' = H'' \cdot \mathrm{Var}(A)$, $B' = K' \cdot \mathrm{Var}(A)$ and $B'' = K'' \cdot \mathrm{Var}(A)$, up to an evanescent set; moreover we may choose versions of H' and H'' such that $H' - H''$ takes only the values -1 and $+1$. Then $H = (K' - K'')/(H' - H'')$ satisfies the required conditions. □

3.14 **Proposition.** a) *Let $A \in \mathscr{A}_{\mathrm{loc}}$ and let M be a local martingale that is locally bounded. Then the process $MA - M \cdot A$ is a local martingale.*

b) *If moreover A is predictable, the process $MA - M_- \cdot A$ is also a local martingale.*

Proof. By localization, we may assume that $A \in \mathscr{A}$ and $M \in \mathscr{M}$, and that M is bounded. We then deduce the claim from 1.44 and 3.12. □

§ 3b. Doob-Meyer Decomposition and Compensators of Increasing Processes

1. We first recall, without proof (see [33, 36]) the following result which is due to Meyer, and is known under the name "Doob-Meyer decomposition of submartingales".

3.15 Theorem. *If X is a submartingale of class* (D) (see 1.46), *there exists a unique (up to indistinguishability) increasing integrable predictable process A with $A_0 = 0$ and such that $X - A$ is a uniformly integrable martingale.*

3.16 Corollary. *Any predictable local martingale which belongs to $\mathscr{V}$ is equal to* 0 (up to an evanescent set).

Proof. By localization, it is enough to prove that if $X \in \mathscr{M} \cap \mathscr{A}$ is predictable, then $X = 0$ up to an evanescent set (recall that $\mathscr{M}_{\text{loc}} \cap \mathscr{V} \subset \mathscr{A}_{\text{loc}}$ by 3.11). We deduce from 3.3 that $X = A - B$, where $A, B \in \mathscr{A}^+$ are predictable. Since $A \in \mathscr{A}^+$ is a submartingale of class (D), 3.15 implies the existence of a unique predictable $A' \in \mathscr{A}^+$ such that $A - A' \in \mathscr{M}$, up to an evanescent set. But both processes $A' = A$ and $A' = B$ meet those requirements, thus $X = A - B$ is equal to 0, up to an evanescent set. □

2. Now we state another very important corollary.

3.17 Theorem. *Let $A \in \mathscr{A}^+_{\text{loc}}$. There is a process, called the compensator of A and denoted by A^p, which is unique up to an evanescent set, and which is characterized by being a predictable process in $\mathscr{A}^+_{\text{loc}}$ meeting any one of the following three equivalent statements*:

(i) *$A - A^p$ is a local martingale*;

(ii) *$E(A^p_T) = E(A_T)$ for all stopping times T*;

(iii) *$E[(H \cdot A^p)_\infty] = E[(H \cdot A)_\infty]$ for all nonnegative predictable processes H.*

Sometimes, A^p is called "*predictable compensator*" of A, or also "*dual predictable projection*" of A.

Proof. We have (iii) $\Rightarrow$ (ii): take $H = 1_{[\![0,T]\!]}$, and (ii) $\Rightarrow$ (iii) (by a monotone class argument, using 2.2 and $A_0 = A^p_0 = 0$). To prove the equivalence (i) $\Leftrightarrow$ (ii), consider a localizing sequence (T_n) such that A^{T_n} and $(A^p)^{T_n}$ are in $\mathscr{A}^+$. Then (i) is equivalent to: $(A - A^p)^{T_n} \in \mathscr{M}$ for each n, which by 1.44 is equivalent to: $E(A^p_{T \wedge T_n}) = E(A_{T \wedge T_n})$ for each n and each stopping time T, and this in turn is equivalent to (ii) because $\lim_{(n)} \uparrow E(A_{T \wedge T_n}) = E(A_T)$ and $\lim_{(n)} \uparrow E(A^p_{T \wedge T_n}) = E(A^p_T)$.

If there were two predictable processes in $\mathscr{A}^+_{\text{loc}}$ satisfying (i), their difference would be a predictable local martingale in $\mathscr{V}$: then 3.16 implies the uniqueness.

It remains to proves the existence of a predictable $A^p \in \mathscr{A}^+_{\text{loc}}$ satisfying (i). Let (T_n) be a localizing sequence such that $A^{T_n} \in \mathscr{A}^+$; then A^{T_n} is a submartingale of class (D). 3.15 yields the existence of a predictable process $B(n) \in \mathscr{A}^+$ such that $A^{T_n} - B(n) \in \mathscr{M}$, and the uniqueness implies that $B(n+1)^{T_n} = B(n)$. Therefore, the process $A^p = \sum_{(n)} B(n) 1_{]\!]T_{n-1}, T_n]\!]}$ is predictable and, because it satisfies $(A^p)^{T_n} = B(n)$, it is in $\mathscr{A}^+_{\text{loc}}$ and $A - A^p \in \mathscr{M}_{\text{loc}}$. □

Here is a trivial, but less classical, extension of this theorem:

3.18 **Theorem.** *Let $A \in \mathscr{A}_{\text{loc}}$. There exists a process, called again the compensator of A and denoted by A^p, which is unique up to an evanescent set, and which is characterized by being a predictable process of $\mathscr{A}_{\text{loc}}$ such that $A - A^p$ is a local martingale.*

Moreover, for each predictable process H such that $H \cdot A \in \mathscr{A}_{\text{loc}}$, then $H \cdot A^p \in \mathscr{A}_{\text{loc}}$ and $H \cdot A^p = (H \cdot A)^p$, and in particular $H \cdot A - H \cdot A^p$ is a local martingale.

Proof. By 3.3 we have $A = B - C$ with $B, C \in \mathscr{A}^+_{\text{loc}}$ (because $B + C = \text{Var}(A) \in \mathscr{A}^+_{\text{loc}}$ by hypothesis). Then $A^p = B^p - C^p$ is a predictable process in $\mathscr{A}_{\text{loc}}$ and $A - A^p \in \mathscr{M}_{\text{loc}}$. If there were two processes satisfying the required properties, their difference would be a predictable local martingale belonging to $\mathscr{A}_{\text{loc}}$, and hence would be equal to 0 up to an evanescent set by 3.16.

Let H be a predictable process such that $H \cdot A \in \mathscr{A}_{\text{loc}}$. Then $H^+ \cdot B \in \mathscr{A}^+_{\text{loc}}$ and by 3.5, $H^+ \cdot B^p$ is a predictable process in $\mathscr{A}^+_{\text{loc}}$. For any other nonnegative predictable process K, $E[(KH^+) \cdot B^p_\infty] = E[(KH^+) \cdot B_\infty]$ by 3.17(iii). But $(KH^+) \cdot B = K \cdot (H^+ \cdot B)$, and similarly for B^p: hence the characterization 3.17(iii) implies that $H^+ \cdot B^p = (H^+ \cdot B)^p$. Similarly we have $H^- \cdot B^p = (H^- \cdot B)^p$ and $(H^\pm \cdot C)^p = H^\pm \cdot C^p$, and all these processes are in $\mathscr{A}^+_{\text{loc}}$. Therefore, the second assertion of the theorem follows from:

$$H \cdot A = H^+ \cdot B + H^- \cdot C - H^- \cdot B - H^+ \cdot C$$
$$H \cdot A^p = H^+ \cdot B^p + H^- \cdot C^p - H^- \cdot B^p - H^+ \cdot C^p. \qquad \square$$

Here are some other, trivial, properties of the compensator:

3.19 If $A \in \mathscr{A}_{\text{loc}}$ is predictable, then $A^p = A$.

3.20 If $A \in \mathscr{A}_{\text{loc}}$ and if T is a stopping time, $(A^T)^p = (A^p)^T$.

3.21 If $A \in \mathscr{A}_{\text{loc}}$, ${}^p(\Delta A) = \Delta(A^p)$ (use 2.31 and $A - A^p \in \mathscr{M}_{\text{loc}}$).

3.22 If $A \in \mathscr{A}_{\text{loc}}$, then A is a local martingale if and only if $A^p = 0$.

3.23 If $A \in \mathscr{M}_{\text{loc}} \cap \mathscr{V}$ and if H is a predictable process such that $H \cdot A \in \mathscr{A}_{\text{loc}}$, then $H \cdot A$ is a local martingale (because $(H \cdot A)^p = H \cdot A^p$ and $A^p = 0$ by 3.22). □

Let us insist on the fact that the compensator A^p of an $A \in \mathscr{A}_{\text{loc}}$ usually *differs* from its predictable projection ${}^pA$ (see an example after the proof of 3.27).

3. A fundamental example: point processes and the Poisson process. By definition, an *adapted point process* is a process $N \in \mathscr{V}^+$ that takes its values in $\mathbb{N}$, and whose jumps are equal to 1 (i.e. the jump process ΔN takes only the values 0 and 1). Sometimes, such a process is called a "simple" point process, by reference to the fact that ΔN takes the values 0 and 1 only: if N_t is the number of "events" occuring in the interval $(0, t]$, this assumption means that two or more events cannot occur exactly at the same time.

We can associate the following sequence of stopping times to the point process N:

3.24 $$T_n = \inf(t: N_t = n)$$

Note that $T_0 = 0$, that $T_n < T_{n+1}$ on the set $\{T_n < \infty\}$, and that $\lim_{(n)} T_n = \infty$. Conversely, the sequence (T_n) completely characterizes the process N, since we have

3.25 $$N = \sum_{n \in \mathbb{N}^*} 1_{[\![T_n, \infty[\![}.$$

Finally, let us remark that any adapted point process is *locally integrable*, and even *locally bounded*, because $N_{T_n} \leq n$.

3.26 **Definitions.** a) An *extended Poisson process on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (or, *relative to* $\mathbf{F}$) is an adapted point process N such that

(i) $E(N_t) < \infty$ for each $t \in \mathbb{R}_+$;

(ii) $N_t - N_s$ is independent of the σ-field $\mathscr{F}_s$ for all $0 \leq s < t$.

b) The function $a(t) = E(N_t)$ is called the *intensity* of N. If this function is continuous, we say that N is a *Poisson process*; if this function is $a(t) = t$, we say that N is a *standard Poisson process.* □

We shall prove later that if N is a Poisson process, the distribution of the variable $N_t - N_s$ is a Poisson distribution with mean $a(t) - a(s)$. For the time being, we content ourselves with the computation of the compensator.

3.27 **Proposition.** *Let N be an extended Poisson process on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ *with intensity* $a(\cdot)$.

(a) *The compensator of N is $N_t^p = a(t)$.*

(b) *N is quasi-left-continuous if and only if it is a Poisson process.* (we shall prove later a converse to (a)).

Proof. The definition of an extended Poisson process immediately yields $E(N_t - N_s | \mathscr{F}_s) = a(t) - a(s)$ for $s \leq t$. Hence $X_t = N_t - a(t)$ is a martingale and $A_t = a(t)$ is a predictable (because deterministic) process in $\mathscr{A}_{\text{loc}}^+$, thus we have (a).

Because of 3.21, we have for all predictable times T:

3.28 $$P(\Delta N_T \neq 0, T < \infty) = E(\Delta N_T 1_{\{T<\infty\}}) = E\{E(\Delta N_T | \mathscr{F}_{T-}) 1_{\{T<\infty\}}\}$$
$$= E(\Delta a(T) 1_{\{T<\infty\}}).$$

Then if a is continuous, 3.28 implies $P(\Delta N_T \neq 0, T < \infty) = 0$ for all predictable times and so N is quasi-left-continuous. Conversely if a has a discontinuity at time t, 3.28 yields $P(\Delta N_t \neq 0) = \Delta a(t) > 0$, and so N is not quasi-left-continuous: hence we have shown (b). □

Let us remark that another application of 2.35 shows that ${}^pN = N_-$ if N is a Poisson process, hence giving an example where ${}^pN \neq N^p$.

§ 3c. Lenglart Domination Property

We will now introduce a relation of "domination" between processes, which is due to Lenglart and which gives rise to some very useful inequalities.

3.29 **Definition.** Let X and Y be two optional processes. We say that X is *L-dominated by* Y if $E(|X_T|) \leq E(|Y_T|)$ for every bounded stopping time T. □

3.30 **Lemma.** *Let X be a càdlàg adapted process which is L-dominated by an increasing process A. For all stopping times T and all $\varepsilon, \eta > 0$,*

a) *if A is predictable,*

3.31 $$P\left(\sup_{s\leq T} |X_s| \geq \varepsilon\right) \leq \frac{\eta}{\varepsilon} + P(A_T \geq \eta);$$

b) *If A is adapted,*

3.32 $$P\left(\sup_{s\leq T} |X_s| \geq \varepsilon\right) \leq \frac{1}{\varepsilon}\left[\eta + E\left(\sup_{s\leq T} \Delta A_s\right)\right] + P(A_T \geq \eta).$$

Proof. We have $P(\sup_{s\leq T}|X_s| \geq \varepsilon) = \lim_n P(\sup_{s\leq T}|X_s| > \varepsilon - 1/n)$, so it suffices to prove 3.31 or 3.32 with the left-hand side replaced by $P(\sup_{s\leq T}|X_s| > \varepsilon)$.

Let $T_n = T \wedge n$. Then $\lim_n \uparrow P(\sup_{s\leq T_n}|X_s| > \varepsilon) = P(\sup_{s\leq T}|X_s| > \varepsilon)$ and $\lim_n \uparrow P(A_{T_n} \geq \eta) \leq P(A_T \geq \eta)$ and $\lim_n \uparrow E(\sup_{s\leq T_n} \Delta A_s) = E(\sup_{s\leq T} \Delta A_s)$, so it suffices to prove the result for each T_n. In other words, we can assume that the stopping time T is bounded.

Set $R = \inf(s\colon |X_s| > \varepsilon)$ and $S = \inf(s\colon A_s \geq \eta)$. Then R is a stopping time, while S is a stopping time (resp. a predictable time) in case (b) (resp. (a)), by 1.28 and 2.13. Moreover $\{\sup_{s\leq T}|X_s| > \varepsilon\} \subset \{A_T \geq \eta\} \cup \{R \leq T < S\}$, hence

3.33 $$P\left(\sup_{s\leq T} |X_s| > \varepsilon\right) \leq P(R \leq T < S) + P(A_T \geq \eta).$$

a) Assume A predictable. Then S is announced by a sequence (S_n) of stopping times, and

$$P(R \leq T < S) \leq \lim_n P(R \leq T < S_n) \leq \lim_n P(|X_{R \wedge T \wedge S_n}| \geq \varepsilon)$$

3.34 $$\leq \frac{1}{\varepsilon} \lim_n E(A_{R \wedge T \wedge S_n})$$

by 3.29. Moreover $S_n < S$ a.s. on $\{S > 0\}$, hence $A_{R \wedge T \wedge S_n} \leq A_{S_n} \leq \eta$ a.s. and 3.31 follows from 3.33 and 3.34.

b) Assume A adapted. Then, as above,

$$P(R \leq T < S) \leq P(|X_{R \wedge T \wedge S}| \geq \varepsilon) \leq \frac{1}{\varepsilon} E(A_{R \wedge T \wedge S})$$

and $A_{R \wedge T \wedge S} \leq \eta + \sup_{s \leq T} \Delta A_s$, so 3.33 yields the result. □

§ 3d. The Discrete Case

Let us now start with a discrete stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}}, P)$ and examine how the preceding notions translate into that simple setting.

1. First of all, the notion of processes with finite-variation is very simple indeed: every discrete-time process has finite-variation! More precisely, Definition 3.1 translates as:

3.35 **Definitions.** a) The set $\mathscr{V}^+$ of *adapted increasing processes* is the set of all adapted processes A with $A_0 = 0$ and $A_n \leq A_{n+1}$ for all $n \in \mathbb{N}$.

b) The set $\mathscr{V}$ of *adapted processes with finite-variation* is the set of all adapted processes A such that $A_0 = 0$. □

If $A \in \mathscr{V}$, its *variation process* $\mathrm{Var}(A)$ is

3.36 $$\mathrm{Var}(A)_n = \sum_{0 \leq p \leq n} |\Delta A_p| = \sum_{1 \leq p \leq n} |A_p - A_{p-1}|.$$

Proposition 3.3 is trivial: take $B_n = \sum_{0 \leq p \leq n} (\Delta A_p)^+$ and $C_n = \sum_{0 \leq p \leq n} (\Delta A_p)^-$.

We say that $dB \ll dA$ if $\Delta B_n = 0$ a.s. on the set $\{\Delta A_n = 0\}$, for each $n \in \mathbb{N}$. We define the *integral process* $H \cdot A$ by

3.37 $$H \cdot A_n = \sum_{0 \leq p \leq n} H_p \Delta A_p = \sum_{1 \leq p \leq n} H_p (A_p - A_{p-1})$$

Proposition 3.5 is also trivial, while for 3.13 it is immediate to check that the process

$$H_n = \frac{B_n - B_{n-1}}{A_n - A_{n-1}} 1_{\{n \geq 1\}} 1_{\{A_n \neq A_{n-1}\}}$$

satisfies the required properties.

The definitions of $\mathscr{A}$, $\mathscr{A}_{\mathrm{loc}}$, $\mathscr{A}^+$, $\mathscr{A}^+_{\mathrm{loc}}$ are similar to those in § 3a, and 3.10, 3.11 and 3.14 hold true (the proofs of 3.10 and 3.11 are basically the same than in the general case; the proof of 3.14 is trivial, up to a localization).

2. Let us turn now to the Doob decomposition Theorem, which holds also in the discrete case, and which we will prove here.

3.38 **Theorem.** *If X is a submartingale of class* (D), *there exists a predictable increasing integrable process A such that $X - A \in \mathscr{M}$; if A' is another such process, then $A'_n = A_n$ a. s. for every $n \in \mathbb{N}$.*

Proof. Suppose first that a process A satisfies all the conditions above. Then, since it is predictable,

$$0 = E[(X - A)_{n+1} - (X - A)_n|\mathscr{F}_n] = E(X_{n+1} - X_n|\mathscr{F}_n) - (A_{n+1} - A_n)$$

and $A_{n+1} - A_n = E(X_{n+1} - X_n|\mathscr{F}_n)$ is uniquely determined, up to a P-null set. Since $A_0 = 0$, this proves the uniqueness. For the existence, we define a predictable increasing process A by $A_0 = 0$ and

$$A_n = \sum_{0 \leq p \leq n-1} E(X_{p+1} - X_p|\mathscr{F}_p)$$

and the same computation as above shows that $X - A$ is a martingale. Moreover $E(A_n) = E(X_n) - E(X_0)$ and since X is of class (D) we obtain that $E(A_\infty) = \sup E(A_n) < \infty$, hence $A \in \mathscr{A}^+$. The fact that $X - A$ is uniformly integrable is then trivial. □

Corollary 3.16 is trivial, and may be stated as follows:

3.39 **Corollary.** *Any local martingale X that is predictable satisfies $X_n = X_0$ a. s. for each $n \in \mathbb{N}$.*

Theorems 3.17 and 3.18 and properties 3.19 to 3.23, are valid. They all are very easy, once noticed that the following formula gives an explicit form for the compensator of $A \in \mathscr{A}_{\mathrm{loc}}$:

3.40 $$(A^p)_n = \sum_{1 \leq p \leq n} [E(A_p|\mathscr{F}_{p-1}) - A_{p-1}]$$

(with the extended conditional expectation).

3. As usual, we associate to the discrete basis $\mathscr{B}$ a continuous basis $\mathscr{B}'$ by 1.55, and to each process X on $\mathscr{B}$ a process X' on $\mathscr{B}'$ by 1.59. Then

3.41 X is a submartingale of class (D) on $\mathscr{B}$ if and only if X' is so on $\mathscr{B}'$; then A is associated to X by 3.38 if and only if A' is associated to X' by 3.15. □

3.42 We have $A \in \mathscr{A}_{\text{loc}}(\mathscr{B})$ is and only if $A' \in \mathscr{A}_{\text{loc}}(\mathscr{B}')$ (with some obvious notation), and then $(A')^p = (A^p)'$. □

4. Semimartingales and Stochastic Integrals

Now we are ready to expound most of the "classical" results on semimartingales. We also introduce stochastic integrals of locally bounded processes: they will be used from time to time in this book; they also are an essential tool to obtain the quadratic variation of semimartingales and to play with their "characteristics".

We go through the shortest possible route, thus leaving undiscovered a large number of interesting properties (the main one being the characterization of semimartingales as the most general "stochastic integrators"). We also emphazise that several presentations are available for the same material and here we follow rather closely the book [36] of Dellacherie and Meyer (including for the proofs).

§ 4a. Locally Square-Integrable Martingales

The stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is fixed throughout the whole section. Recall that $\mathscr{H}^2$ denotes the set of all square-integrable martingales and that $\mathscr{H}^2_{\text{loc}}$ is its localized class (the so-called "locally square-integrable martingales). The following is obvious:

4.1 A local martingale M such that $M_0 \in L^2$ and such that the process ΔM is locally bounded belongs to $\mathscr{H}^2_{\text{loc}}$.

4.2 **Theorem.** *To each pair (M, N) of locally square-integrable martingales one associates a predictable process $\langle M, N \rangle \in \mathscr{V}$, unique up to an evanescent set, such that $MN - \langle M, N \rangle$ is a local martingale. Moreover,*

4.3 $$\langle M, N \rangle = \tfrac{1}{4}(\langle M + N, M + N \rangle - \langle M - N, M - N \rangle)$$

and if M, $N \in \mathscr{H}^2$ then $\langle M, N \rangle \in \mathscr{A}$ and $MN - \langle M, N \rangle \in \mathscr{M}$. Furthermore $\langle M, M \rangle$ is non-decreasing, and it admits a continuous version if and only if M is quasi-left-continuous.

The process $\langle M, N \rangle$ is called the *predictable quadratic covariation* (for a reason to be seen in § 4c), or the *quadratic characteristic*, or also the *angle bracket*, of the pair (M, N). Note that

4.4 $$\langle M, N \rangle = \langle M - M_0, N - N_0 \rangle.$$

Proof. The uniqueness immediately follows from the nullity of any predictable local martingale that belongs to $\mathscr{V}$ (3.16). Then 4.3 follows from the uniqueness and from $MN = [(M + N)^2 - (M - N)^2]/4$.

Because of 4.3 it suffices to prove the existence when $N = M$, and by localization we may assume that $M \in \mathscr{H}^2$. Then the variable $\sup_{(s)} M_s^2$ is integrable by 1.43, and it follows that the process $X = M^2$ is of class (D). Moreover, Jensen inequality implies that X is a submartingale. Then we apply Doob-Meyer decomposition Theorem 3.15 to X, to obtain the existence of a process $\langle M, M \rangle \in \mathscr{A}^+$ which is predictable and satisfies $M^2 - \langle M, M \rangle \in \mathscr{M}$.

It remains to prove the result concerning the quasi-left-continuity of M. A simple computation shows that $\Delta(M^2) = (\Delta M)^2 - 2M_{-}\Delta M$. Then 2.31 and the fact that $M^2 - \langle M, M \rangle \in \mathscr{M}_{\text{loc}}$ imply that $\Delta(\langle M, M \rangle) = {}^p[\Delta(M^2)] = {}^p[(\Delta M)^2]$. It is immediate from the definition of the predictable projection that the sets $\{{}^p(1_{\{\Delta M \neq 0\}}) > 0\}$ and $\{{}^p[(\Delta M)^2] > 0\}$ are equal, up to an evanescent set. Then the result follows from 2.35. □

There is a bijective correspondance between the elements M of $\mathscr{H}^2$ and their terminal variables M_∞. So it is natural to endow $\mathscr{H}^2$ with an Hilbertian structure, as follows: if M, $N \in \mathscr{H}^2$ the scalar product and the norm are

4.5 $$(M, N)_{H^2} = E(M_\infty N_\infty), \quad \|M\|_{H^2} = \|M_\infty\|_{L^2}.$$

With these, $\mathscr{H}^2$ is indeed a Hilbert space: if (M^n) is a Cauchy sequence for $\|\cdot\|_{H^2}$, then the sequence (M_∞^n) is Cauchy in $L^2(\Omega, \mathscr{F}_{\infty-}, P)$ and so goes to a limit M_∞ is this space; then if M is the (unique) martingale with terminal variable M_∞, it belongs to $\mathscr{H}^2$ and $\|M^n - M\|_{H^2} \to 0$.

In view of the previous theorem, we have that

4.6 $$(M, N)_{H^2} = E(\langle M, N \rangle_\infty) + E(M_0 N_0).$$

4.7 **Lemma.** *If M^n converges to M in $\mathscr{H}^2$, then* $\sup_{s \in \mathbb{R}_+} |M_s^n - M_s| \to 0$ *in* L^2.

Proof. That immediately follows from Doob's inequality 1.43. □

4.8 **Corollary.** *The set of all continuous elements of $\mathscr{H}^2$ is a closed subspace of the Hilbert space $\mathscr{H}^2$.*

A fundamental example: the Wiener process.

4.9 **Definitions.** a) A *Wiener process on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (or, *relative to* $\mathbf{F}$) is a continuous adapted process W such that $W_0 = 0$ and

(i) $E(W_t^2) < \infty$ for each $t \in \mathbb{R}_+$, and $E(W_t) = 0$ for each $t \in \mathbb{R}_+$;

(ii) $W_t - W_s$ is independent of the σ-field $\mathscr{F}_s$ for all $0 \leq s \leq t$.

b) The function $\sigma^2(t) = E(W_t^2)$ is called the *variance function* of W. If $\sigma^2(t) = t$, we say that W is a *standard Wiener process*. (Another usual term, instead of "Wiener process", is "Brownian motion".) □

This definition should be compared to the definition of Poisson processes (3.26): the reader will observe that we have not defined an "extended" Wiener process. It is of course well known (and also proved later on) that a Wiener process is Gaussian, and we study all Gaussian martingales (including discontinuous ones) later. Note that the previous definition is just one possibility, among many others, for defining a Wiener process (for instance, we could define W as being a Gaussian process satisfying (ii) with a continuous function $\sigma^2(\cdot)$: that would imply the existence of a version with continuous paths).

For the moment, we just prove the following proposition, which will receive a converse later.

4.10 **Proposition.** *A Wiener process W is a continuous martingale, and its angle bracket $\langle W, W\rangle$ is $\langle W, W\rangle_t(\omega) = \sigma^2(t)$, where $\sigma^2(\cdot)$ is its variance function.*

Proof. The first assertion is trivial (use 4.9ii and $E(W_t) = 0$). Note that σ^2 is continuous, null at 0, and increasing, because by Jensen inequality W^2 is a submartingale. Then it remains to prove that the process $X_t = W_t^2 - \sigma^2(t)$ is a martingale. A simple computation shows that

$$X_t - X_s = (W_t - W_s)^2 - [\sigma^2(t) - \sigma^2(s)] + 2W_s(W_t - W_s);$$

then by 4.9ii and the definition of σ^2, we obtain $E(X_t - X_s|\mathscr{F}_s) = 0$ if $s \leq t$, hence the result. □

§ 4b. Decompositions of a Local Martingale

1. Let us introduce new definitions.

4.11 **Definitions.** a) Two local martingales M and N are called *orthogonal* if their product MN is a local martingale (the terminology will be explained below).

b) A local martingale X is called a *purely discontinuous local martingale* if $X_0 = 0$ and if it is orthogonal to all continuous local martingales. □

4.12 **Comments.** The reader should not be misled by the terminology: "purely discontinuous" is a term that is sort of "orthogonal" to "continuous". However a purely discontinuous local martingale X is (usually) not the sum of its jumps: first of all, the series $\sum_{s\leq t} \Delta X_s$ usually diverges; and even if it converges, its sum usually differs from X_t.

For example, we shall see later that $M_t = N_t - a(t)$ is a purely discontinuous local martingale if N is a Poisson process with intensity function $a(\cdot)$ (recall that a is continuous); indeed in that case $\sum_{s\leq t} \Delta M_s = N_t \neq M_t$. This sort of martingale is the prototype of all purely discontinuous local martingales, a fact that explains why in many places those are also called *compensated sums of jumps.* □

4.13 **Lemma.** a) *A local martingale M is orthogonal to itself if and only if M_0 is square integrable and $M = M_0$ up to an evanescent set.*

b) *A purely discontinuous local martingale which is continuous is a. s. equal to* 0.

c) *Let M, N be two orthogonal local martingales. For all stopping times S, T, the stopped local martingales M^S and N^T are orthogonal.*

Proof. a) Sufficiency is obvious. Conversely, assume that M is orthogonal to itself. That $M_0 \in L^2$ is then trivial. By localization we may assume that M and M^2 belong to $\mathscr{M}$, so $M \in \mathscr{H}^2$. Thus $E(M_t) = E(M_0)$ and $E(M_t^2) = E(M_0^2)$, and these facts imply $M_t = M_0$ a. s. (recall that $M_0 = E(M_t|\mathscr{F}_0)$), and we are finished.

b) This is a trivial consequence of (a).

c) It clearly suffices to prove the claim when $T = \infty$, an hypothesis which we assume further. Let (T_n) be a localizing sequence such that M^{T_n}, N^{T_n}, $(MN)^{T_n}$ belong to $\mathscr{M}$. Then for every stopping time R,

$$\begin{aligned}E((M^S N)_R^{T_n}) &= E((MN)_{R\wedge S}^{T_n}) + E(M_S^{T_n}(N_R^{T_n} - N_S^{T_n})1_{\{S<R\}})\\ &= E(M_0 N_0) + E[M_S^{T_n}1_{\{S<R\}}E(N_R^{T_n} - N_S^{T_n}|\mathscr{F}_S)] = E(M_0 N_0)\end{aligned}$$

by the stopping theorem. Thus Lemma 1.44 gives that $(M^S N)^{T_n}$ belongs to $\mathscr{M}$ for all n, and the result follows. □

4.14 **Lemma.** a) *A local martingale M with $M_0 = 0$ is purely discontinuous if and only if it is orthogonal to all continuous bounded martingales N with $N_0 = 0$.*

b) *A local martingale that belongs to $\mathscr{V}$ is purely discontinuous.*

Proof. a) Only the sufficient condition needs proving. Since $M_0 = 0$, M is orthogonal to N if and only if it is orthogonal to $N - N_0$. Moreover if N is continuous, $N - N_0$ is locally bounded, so the claim follows from a localization.

b) Let $M \in \mathscr{V} \cap \mathscr{M}_{\text{loc}}$. We know that $M \in \mathscr{A}_{\text{loc}}$ (3.11). If N is a continuous bounded martigale, 3.14 yields that $MN - N \cdot M$ is a local martingale; moreover $N \cdot M$ is also a local martingale by 3.23: then MN itself is a local martingale, and in view of (a) that proves the claim. □

Finally, the following explains the terminology "orthogonal":

4.15 **Proposition.** *Let $M, N \in \mathscr{H}^2$. There is equivalence between*:

a) *M and N are orthogonal in the sense of* 4.11.

b) $\langle M, N\rangle = 0$.

c) *For all stopping times T, M^T and $N - N_0$ are orthogonal in the Hilbert space $\mathscr{H}^2$.*

Proof. (a)⇔(b) follows from Theorem 4.2. If M and N are orthogonal in the sense of 4.11, then so are M^T and $N - N_0$ (T arbitrary stopping time), hence $E(M_\infty^T(N_\infty - N_0)) = E(M_0^T(N_0 - N_0)) = 0$ and M^T is orthogonal to $N - N_0$ in

$\mathscr{H}^2$. Conversely, if (c) holds and if T is a stopping time, the stopping theorem yields

$$E(M_T N_T) = E(M_T N_\infty) = E(M_\infty^T(N_\infty - N_0)) + E(M_T N_0) = 0 + E(M_0 N_0).$$

Hence Lemma 1.44 yields that $MN \in \mathscr{M}$, so (a) holds. □

4.16 **Corollary.** *The set $\mathscr{H}^{2,d}$ of all purely discontinuous martingale in $\mathscr{H}^2$ is the orthogonal subspace, in the Hilbert space $\mathscr{H}^2$, of the set $\mathscr{H}^{2,c}$ of all continuous elements of $\mathscr{H}^2$* (note that $\mathscr{H}^{2,c}$ is a closed subspace of $\mathscr{H}^2$ by 4.8).

Proof. Let $M \in \mathscr{H}^{2,c}$ and $N \in \mathscr{H}^{2,d}$ (so $N_0 = 0$). Applying Definition 4.11b and (a) $\Rightarrow$ (c) in 4.15 (for $T = \infty$) yields that M and N are orthogonal in $\mathscr{H}^2$.

Conversely, assume that N is orthogonal to $\mathscr{H}^{2,c}$ in $\mathscr{H}^2$. Firstly if $Y \in L^2(\Omega, \mathscr{F}_0, P)$ we can consider the (trivial) continuous martingale $M_t = Y$ for all t; then $E(N_0 Y) = E(N_0 M_\infty) = E(N_\infty M_\infty) = 0$ by hypothesis, and that being true for all such Y we deduce that $N_0 = 0$ a.s. (recall that N_0 is $\mathscr{F}_0$-measurable). Secondly if $M \in \mathscr{H}^{2,c}$ and if T is a stopping time, then $M^T \in \mathscr{H}^{2,c}$ and so M^T is orthogonal to $N = N - N_0$ in $\mathscr{H}^2$. Therefore 4.15 yields that M and N are orthogonal in the sense of 4.11, and 4.11b yields that $M \in \mathscr{H}^{2,d}$. □

2. Now we turn to the decompositions of a local martingale. The first decomposition is elementary.

4.17 **Proposition.** *Let $a > 0$. Any local martingale M admits a* (non-unique) *decomposition $M = M_0 + M' + M''$, where M' and M'' are local martingales with $M'_0 = M''_0 = 0$, M' has finite variation and $|\Delta M''| \leq a$* (hence $M'' \in \mathscr{H}^2_{\text{loc}}$).

Proof. By localization, it is enough to prove the result when $M \in \mathscr{M}$. Let $b = a/2$. Since M is càdlàg, it has only finitely many jumps of size bigger than b on each finite interval, hence the process $A = \sum_{s \leq \cdot} \Delta M_s 1_{\{|\Delta M_s| > b\}}$ is well-defined, and belongs to $\mathscr{V}$.

Put $T_n = \inf(t\colon \mathrm{Var}(A)_t > n$ or $|M_t| > n)$. We have $\mathrm{Var}(A)_{T_n} \leq n + |\Delta M_{T_n}|$, thus $\mathrm{Var}(A)_{T_n} \leq 2n + |M_{T_n}|$ and since $M \in \mathscr{M}$ it follows that $\mathrm{Var}(A)_{T_n}$ is integrable. Therefore $A \in \mathscr{A}_{\text{loc}}$, with compensator A^p.

It remains to prove that the processes $M' = A - A^p$ and $M'' = M - M' - M_0$ satisfy the required conditions. We have $M' \in \mathscr{M}_{\text{loc}} \cap \mathscr{V}$ by construction. Let $X = \Delta M 1_{\{|\Delta M| \leq b\}}$. Then $\Delta A = \Delta M - X$, hence 3.21 and 2.31 yield $\Delta(A^p) = {}^p(\Delta A) = -{}^pX$; a simple computation shows that $\Delta M'' = \Delta M - \Delta A + \Delta(A^p) = X - {}^pX$; since $|X| \leq b$, we have $|{}^pX| \leq b$ and $|\Delta M''| \leq 2b = a$. □

Our second decomposition is much deeper.

4.18 **Theorem.** *Any local martingale M admits a unique* (up to indistinguishability) *decomposition*

$$M = M_0 + M^c + M^d$$

where $M_0^c = M_0^d = 0$, M^c *is a continuous local martingale, and* M^d *is a purely discontinuous local martingale.*

M^c is called the *continuous part* of M, and M^d its *purely discontinuous part.*

Proof. The uniqueness follows from 4.13b. For the existence, if we decompose M according to 4.17 and use 4.14b and localize, we see that it is enough to consider the case when $M \in \mathscr{H}^2$.

But 4.16 implies that $\mathscr{H}^2$ is the direct sum of $\mathscr{H}^{2,c}$ and $\mathscr{H}^{2,d}$. Then $M = M^c + M^d$ with $M^c \in \mathscr{H}^{2,c}$ and $M^d \in \mathscr{H}^{2,d}$, hence the result. □

4.19 **Corollary.** *Let* M *and* N *be two purely discontinuous local martingales having the same jumps* $\Delta M = \Delta N$ (up to an evanescent set). *Then* M *and* N *are indistinguishable* (apply 4.18 to $M - N$).

§ 4c. Semimartingales

1. We begin with new notation and definitions.

4.20 $\mathscr{L}$ denotes the set of all local martingales M such that $M_0 = 0$.

4.21 **Definitions.** a) A *semimartingale* is a process X of the form $X = X_0 + M + A$ where X_0 is finite-valued and $\mathscr{F}_0$-measurable, where $M \in \mathscr{L}$, and where $A \in \mathscr{V}$. We *denote by* $\mathscr{S}$ the space of all semimartingales.

b) A *special semimartingale* is a semimartingale X which admits a decomposition $X = X_0 + M + A$ as above, with a process A that is *predictable.* We *denote by* $\mathscr{S}_p$ the set of all special semimartingales. □

It is clear that $\mathscr{M}_{\text{loc}} \subset \mathscr{S}_p$ and that $\mathscr{V} \subset \mathscr{S}$. All semimartingales are càdlàg and adapted. The decomposition $X = X_0 + M + A$ in 4.21a is of course not unique. However, because of 3.16, there is at most one such decomposition with A being predictable (up to an evanescent set), and we set:

4.22 **Definition.** If X is a special semimartingale, the unique decomposition $X = X_0 + M + A$ such that $M \in \mathscr{L}$ and that A is a predictable element of $\mathscr{V}$ is called the *canonical decompositon of* X. □

Although it may not be quite apparent from the definition above, the space of semimartingales is a very pleasant space: it stays stable under a large variety of transformations: under stopping (this is evident), under localization (we will see that later), under "change of time", under "absolutely continuous change of probability measure", under "changes of filtration"; and its main property: it is the largest possible class of processes with respect to which one may "reasonnably" integrate all bounded predictable processes: see [35, 10, 36, 98]. Unfortunately, we do not have enough room here to do a thorough study of the class of semimartingales in all details.

4.23 **Proposition.** *Let X be a semimartingale. There is equivalence between:*
(i) *X is a special semimartingale;*
(ii) *there exists a decomposition $X = X_0 + M + A$ where $A \in \mathscr{A}_{\mathrm{loc}}$;*
(iii) *all decompositions $X = X_0 + M + A$ satisfy $A \in \mathscr{A}_{\mathrm{loc}}$;*
(iv) *the process $Y_t = \sup_{s \le t} |X_s - X_0|$ belongs to $\mathscr{A}^+_{\mathrm{loc}}$.*

Proof. That (iii) $\Rightarrow$ (ii) is obvious. Assume (ii): $X = X_0 + M + A$ with $A \in \mathscr{A}_{\mathrm{loc}}$. If $A' = A^p$ and $M' = M + A - A'$, then $M' \in \mathscr{L}$ and A' is predictable and belongs to $\mathscr{V}$, hence $X = X_0 + M' + A'$ is special: thus (ii) $\Rightarrow$ (i).

If $A \in \mathscr{A}_{\mathrm{loc}}$, the increasing càd process $\sup_{s \le .} |A_s|$ belongs to $\mathscr{A}_{\mathrm{loc}}$, because it is smaller than $\mathrm{Var}(A)$. To prove the implication (i) $\Rightarrow$ (iv) it is then enough to show that if $M \in \mathscr{L}$, then $M^*_t = \sup_{s \le t} |M_s|$ is in $\mathscr{A}_{\mathrm{loc}}$. Let (T_n) be a localizing sequence such that $M^{T_n} \in \mathscr{M}$, and set $S_n = \inf(t\colon t \ge T_n \text{ or } |M_t| > n)$. Then $S_n \uparrow \infty$ and $M^*_{S_n} \le n + |M_{S_n}|$, which is integrable because $M^{T_n} \in \mathscr{M}$: therefore $M^* \in \mathscr{A}_{\mathrm{loc}}$.

Finally, assume (iv). Let $X = X_0 + M + A$ be any decomposition 4.21. By hypothesis, $Y \in \mathscr{A}_{\mathrm{loc}}$, and we have just seen above that $M^* \in \mathscr{A}_{\mathrm{loc}}$, hence the process $A^*_t = \sup_{s \le t} |A_s|$ also belongs to $\mathscr{A}_{\mathrm{loc}}$. Since $\mathrm{Var}(A) \le \mathrm{Var}(A)_- + 2A^*$ and since $\mathrm{Var}(A)_-$ is locally bounded (because it is càg and finite-valued and increasing), then $\mathrm{Var}(A) \in \mathscr{A}_{\mathrm{loc}}$: therefore (iv) $\Rightarrow$ (iii). □

Here is a lemma, closely related to 4.17:

4.24 **Lemma.** *If a semimartingale X satisfies $|\Delta X| \le a$, it is special and its canonical decomposition $X = X_0 + M + A$ satisfies $|\Delta A| \le a$ and $|\Delta M| \le 2a$* (in particular if X is continuous, then M and A are continuous).

Proof. If $S_n = \inf(t\colon |X_t - X_0| > n)$, then $S_n \uparrow \infty$ as $n \uparrow \infty$ and $\sup_{s \le S_n} |X_s - X_0| \le n + a$: hence X meets 4.23(iv) and thus is special. If $X = X_0 + M + A$ denotes its canonical decomposition, we deduce from 2.31 that ${}^p(\Delta X) = {}^p(\Delta M) + {}^p(\Delta A) = \Delta A$; therefore $|\Delta X| \le a$ yields $|\Delta A| \le {}^p(|\Delta X|) \le a$, and $|\Delta M| \le 2a$ by difference. □

4.25 **Proposition.** a) *The spaces $\mathscr{S}$ and $\mathscr{S}_p$ are stable under stopping.*
b) *We have $\mathscr{S}_{\mathrm{loc}} = \mathscr{S}$ and $(\mathscr{S}_p)_{\mathrm{loc}} = \mathscr{S}_p$.*
c) *For X to belong to $\mathscr{S}$ it is sufficient that there exists a localizing sequence (T_n) of stopping times, and a sequence $(Y(n))$ of semimartingales, such that $X = Y(n)$ on each interval $[\![0, T_n[\![$.*

Proof. a) is trivial.

c) Since (T_n) increases to $+\infty$, the process X is càdlàg and adapted. Put

$$Z(n) = Y(n)^{T_n} + (X_{T_n} - Y(n)_{T_n})1_{[\![T_n,\infty[\![}.$$

As a sum of a stopped semimartingale and a process in $\mathscr{V}$, $Z(n) \in \mathscr{S}$, and we consider a decomposition $Z(n) = X_0 + M(n) + A(n)$ of type 4.21: note that $X^{T_n} = Z(n)$. Then $X = X_0 + M + A$, where $M = \sum M(n)1_{]\!]T_{n-1},T_n]\!]}$ and $A = X - X_0 - M$. Since

$$M^{T_n} = \sum_{1 \leq p \leq n} [M(p)^{T_p} - M(p)^{T_{p-1}}]$$

(with the convention $T_0 = 0$) we see that $M \in \mathscr{L}$, and we prove similarly that $A \in \mathscr{V}$: hence $X \in \mathscr{S}$.

b) The first assertion follows from (c). Using this first assertion, and the fact that $(\mathscr{A}_{\text{loc}})_{\text{loc}} = \mathscr{A}_{\text{loc}}$, and 4.23(iv), we immediately obtain that $(\mathscr{S}_p)_{\text{loc}} = \mathscr{S}_p$. □

2. Since $\mathscr{V}$ and $\mathscr{M}_{\text{loc}}$ are in $\mathscr{S}$, there are many examples of semimartingales. The Doob-Meyer decomposition Theorem shows that any sub (super) martingale of class (D) is a special semimartingale. Using a localization, on may in fact prove (see [98]) that:

4.26 An adapted process X is a special semimartingale if and only if $X - X_0$ is the difference of two local submartingales (or, supermartingales). This result will not be used in the sequel. □

Next, using the decomposition in Theorem 4.18 (in particular the uniqueness) and statement 4.14b, we readily get the

4.27 **Proposition.** *Let X be a semimartingale. There is a unique* (up to indistinguishability) *continuous local martingale X^c with $X^c_0 = 0$, such that any decomposition $X = X_0 + M + A$ of type* 4.21 *meets $M^c = X^c$* (up to indistinguishability again).

X^c is called the *continuous martingale part* of X.

It is also interesting, and useful, to recognize all "deterministic" processes that are semimartingales:

4.28 **Proposition.** *Let f be a real-valued function on $\mathbb{R}_+$. For the process $X_t(\omega) = f(t)$ to be a semimartingale, it is necessary and sufficient that f be càdlàg, with finite-variation over each finite interval.*

Proof. The sufficient condition is trivial. Conversely, assume that $X \in \mathscr{S}$. Then necessarily f is càdlàg, hence locally bounded, hence X satisfies 4.23(iv) and is special. We consider a decomposition $X = f(0) + M + A$, and a localizing sequence (T_n) such that $M^{T_n} \in \mathscr{M}$ and $A^{T_n} \in \mathscr{A}$ (because $X \in \mathscr{S}_p$). Denote also the distribution of T_n by $F_n(dx)$, which is a probability measure on $\bar{\mathbb{R}}_+$. We have

$$\begin{aligned} g_n(t) := \int F_n(ds) f(s \wedge t) = E(X_{t \wedge T_n}) \\ &= f(0) + E(M_t^{T_n}) + E(A_t^{T_n}) \\ &= f(0) + E(A_t^{T_n}). \end{aligned}$$

Thus g_n is a function with finite variation. But we also have

$$F_n((t, \infty]) f(t) = g_n(t) - \int_{[0,t]} f(s) F_n(ds)$$

and the last term in the right-hand side is also a function (in t) with finite-variation. Thus f has finite-variation on each finite interval $[0, t]$, provided $F_n((t, \infty]) > 0$. But, since $\lim_{(n)} T_n = \infty$, for every $t \in \mathbb{R}_+$ there exists an n big enough so that $F_n((t, \infty]) > 0$, and the result is proved. □

§ 4d. Construction of the Stochastic Integral

In this subsection, we proceed to constructing the stochastic integral of locally bounded predictable processes with respect to a semimartingale.

1. If $X \in \mathscr{V}$ and if H is a bounded process, we have defined in 3.4 an integral process $H \cdot X_t = \int_0^t H_s \, dX_s$. The problem here is to define an integral process when X does not belong to $\mathscr{V}$, but is only a semimartingale: hence the path $X_.(\omega)$ does not define a measure $dX_s(\omega)$ on $\mathbb{R}_+$ (for instance if X is a Wiener process, then almost all paths $t \rightsquigarrow X_t(\omega)$ have infinite variation over each finite interval).

When H is simple enough, it is very easy. More precisely, we denote by $\mathscr{E}$ the set of all processes of the form:

4.29 $$\begin{cases} \text{either } H = Y 1_{[\![0]\!]}, \ Y \text{ is bounded } \mathscr{F}_0\text{-measurable,} \\ \text{or} \quad H = Y 1_{]\!]r,s]\!]}, \ r < s, \ Y \text{ is bounded } \mathscr{F}_r\text{-measurable.} \end{cases}$$

For such an H, the integral process $H \cdot X_t = \int_0^t H_s \, dX_s = \int_{(0,t]} H_s \, dX_s$ has only one "natural" definition (even if dX_s does not make sense), namely:

4.30 $$H \cdot X_t = \begin{cases} 0 & \text{if } H = Y 1_{[\![0]\!]} \\ Y(X_{s \wedge t} - X_{r \wedge t}) & \text{if } H = Y 1_{]\!]r,s]\!]}. \end{cases}$$

4.31 **Theorem.** *Let X be a semimartingale. The map $H \rightsquigarrow H \cdot X$ defined on $\mathscr{E}$ by 4.30 has an extension, still denoted by $H \rightsquigarrow H \cdot X$* (and $H \cdot X$ is called the *stochastic integral* of H with respect to X) *to the space of all locally bounded predictable processes H, with the following properties*:

(i) *$H \cdot X$ is a càdlàg adapted process*;

(ii) *$H \rightsquigarrow H \cdot X$ is linear, up to evanescence* (i.e., $(aH + K) \cdot X$ and $aH \cdot X + K \cdot X$ are indistinguishable);

(iii) *if a sequence (H^n) of predictable processes converges pointwise to a limit H, and if $|H^n| \leq K$ where K is a locally bounded predictable process, then $H^n \cdot X_t \to H \cdot X_t$ in measure for all $t \in \mathbb{R}_+$.*

Moreover this extension is unique, up to evanescence (i.e., if $H \rightsquigarrow \alpha(H)$ is another extension with the same properties, then $\alpha(H)$ and $H \cdot X$ are indistinguishable), *and in* (iii) *above $H^n \cdot X$ converges to $H \cdot X$ in measure, uniformly on finite intervals*: $\sup_{s \leq t} |H^n \cdot X_s - H \cdot X_s| \xrightarrow{P} 0$.

4.32 **Remark.** One could define the integral process by 4.30 for all H of the form 4.29 but without the measurability conditions on Y, that is for simple processes that are not predictable. But the extension is essentially possible for *predictable* processes only.

Similarly, formula 4.30 makes sense for every process X, semimartingale or not. But the extension is possible only when X is a *semimartingale*: as mentioned before, it is a fundamental result by Bichteler, Dellacherie and Mokobodzki, which explains why the space of semimartingales is so important. □

The proof of this theorem is given below, in part *2* of this subsection. Before, we state various properties of the stochastic integrals; in these, X is always a semimartingale, and H, K are locally bounded predictable processes. All equalities (or other statements) are *up to evanescence.*

4.33 $X \rightsquigarrow H \cdot X$ is linear.

4.34 (a) $H \cdot X$ is a semimartingale;

(b) if X is a local martingale, then so is $H \cdot X$;

(c) if $X \in \mathscr{V}$ then $H \cdot X \in \mathscr{V}$ and $H \cdot X$ coincides with the process defined in 3.4 (Stieltjes integral process). □

4.35 $(H \cdot X)_0 = 0$ and $H \cdot X = H \cdot (X - X_0)$.

4.36 $\Delta(H \cdot X) = H \Delta X$.

4.37 $X^T = X_0 + 1_{[\![0,T]\!]} \cdot X$ and $(H \cdot X)^T = (H 1_{[\![0,T]\!]}) \cdot X$ for all stopping times T; more generally, $K \cdot (H \cdot X) = (KH) \cdot X$. □

4.38 If T is a predictable time and Y is an $\mathbb{R}$-valued $\mathscr{F}_{T-}$-measurable random variable, then $(Y 1_{[\![T]\!]}) \cdot X = Y \Delta X_T 1_{[\![T,\infty[\![}$ (observe that $Y 1_{[\![T]\!]}$ necessarily is a locally bounded predictable process). □

We could still enlarge the class of integrands to some non locally bounded processes. That is rather difficult for a semimartingale; but when $X \in \mathscr{H}^2_{\text{loc}}$ it is simple enough, as we shall presently see. Firstly, to all $X \in \mathscr{H}^2_{\mathrm{loc}}$ we associate the following classes of processes:

4.39 $L^2(X)$ (resp. $L^2_{\mathrm{loc}}(X)$) is the set of all predictable processes H such that the process $H^2 \cdot \langle X, X \rangle$ is integrable (resp. locally integrable). □

Note that all locally bounded predictable processes belong to $L^2_{\mathrm{loc}}(X)$, because we know that $\langle X, X \rangle \in \mathscr{A}^+_{\mathrm{loc}}$.

4.40 **Theorem.** *Let $X \in \mathscr{H}^2_{\mathrm{loc}}$. The map $H \rightsquigarrow H \cdot X$* (defined either on $\mathscr{E}$ by 4.30 or for all locally bounded predictable H by 4.31) *has a further extension to the set $L^2_{\mathrm{loc}}(X)$, still denoted by $H \rightsquigarrow H \cdot X$, which meets* 4.31 (i, ii) *and*

(iii′) *if a sequence (H^n) of predictable processes converges pointwise to a limit H and $|H^n| \leq K$ for some $K \in L^2_{\mathrm{loc}}(X)$, then $\sup_{s \leq t} |H^n \cdot X_s - H \cdot X_s| \to 0$ in measure for all $t \in \mathbb{R}_+$.*

Moreover this extension is unique (up to evanescence), *and we have*:

a) $H \cdot X \in \mathscr{H}^2_{\mathrm{loc}}$.

b) $H \cdot X \in \mathscr{H}^2$ *if and only if* $H \in L^2(X)$.

c) *Properties 4.33, 4.35, 4.36, 4.37* (for $H \in L^2_{\mathrm{loc}}(X)$ and $K \in L^2_{\mathrm{loc}}(H \cdot X)$) *hold.*

d) *If $X, Y \in \mathscr{H}^2_{\mathrm{loc}}$ and $H \in L^2_{\mathrm{loc}}(X)$ and $K \in L^2_{\mathrm{loc}}(Y)$, then*

4.41 $$\langle H \cdot X, K \cdot Y \rangle = (HK) \cdot \langle X, Y \rangle.$$

2. Now we proceed to the proof of the results claimed above. We follow Dellacherie and Meyer [36] rather closely, and of course this part may be skipped.

(a) Firstly, we assume that $X \in \mathscr{V}$, and for all locally bounded predictable H we define $H \cdot X$ by 3.4 (Stieltjes integral): this clearly is the unique extension of 4.30 which meets (i, ii, iii) of 4.31 (Lebesgue theorem) and it readily satisfies the last claim in 4.31, as well as properties 4.34 to 4.37 (the only non-evident property, namely 4.34b, has been proved in 3.23).

(b) Next we assume that $X \in \mathscr{H}^2$. We denote by m the positive finite measure on $(\Omega \times \mathbb{R}_+, \mathscr{P})$ defined by $m(B) = E(1_B \cdot \langle X, X \rangle_\infty)$, so that $L^2(X)$ is exactly the Hilbert space $L^2(\Omega \times \mathbb{R}_+, \mathscr{P}, m)$.

Call $\mathscr{E}'$ the vector space spanned by $\mathscr{E}$. An $H \in \mathscr{E}'$ may have several representations as sums of elements of $\mathscr{E}$, say for instance $H = \sum_{(i)} K^i = \sum_{(j)} H^j$ (finite sums), but we readily deduce from 4.30 that $\sum_{(i)} K^i \cdot X = \sum_{(j)} H^j \cdot X$; then $H \cdot X := \sum_{(i)} K^i \cdot X$ is a linear extension of $H \rightsquigarrow H \cdot X$ from $\mathscr{E}$ to $\mathscr{E}'$.

Let $H \in \mathscr{E}'$. Among its various representations, it has one of the form

$$H = Y_0 1_{[\![0]\!]} + \sum_{1 \leq i \leq n} Y_i 1_{]\!]t_i, t_{i+1}]\!]}$$

where $0 = t_0 < \cdots < t_{n+1} < \infty$ and Y_i is bounded and $\mathscr{F}_{t_i}$-measurable. Then

(1) $$H \cdot X = \sum_{1 \leq i \leq n} Y_i (X^{t_{i+1}} - X^{t_i})$$

belongs to $\mathscr{H}^2$ (recall $(x^t)_s = x_{s\wedge t}$), and

$$(H\cdot X)^2 - H^2\cdot\langle X,X\rangle = 2\sum_{1\le i<j\le n} Y_i Y_j (X^{t_{i+1}} - X^{t_i})(X^{t_{j+1}} - X^{t_j})$$
$$+ \sum_{1\le i\le n} Y_i^2[(X^{t_{i+1}})^2 - \langle X,X\rangle^{t_{i+1}} - (X^{t_i})^2$$
$$+ \langle X,X\rangle^{t_i} - 2X_{t_i}(X^{t_{i+1}} - X^{t_i})]$$

obviously is a martingale. Since $H^2\cdot\langle X,X\rangle$ is predictable, the uniqueness of the angle bracket yields

(2) $$\langle H\cdot X, H\cdot X\rangle = H^2\cdot\langle X,X\rangle.$$

By construction $(H\cdot X)_0 = 0$. So, recalling 4.6, we deduce from (2) that $\|H\cdot X\|_{H^2} = E(H^2\cdot\langle X,X\rangle_\infty) = m(H^2)$. Therefore the map $H \rightsquigarrow H\cdot X$ is an isometry from the subspace $\mathscr{E}'$ of the Hilbert space $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$ into the Hilbert space $\mathscr{H}^2$. Since $\mathscr{E}'$ is dense in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m) = L^2(X)$ (see 2.2), this map admits a unique continuous extension (still an isometry) from $L^2(X)$ into $\mathscr{H}^2$, and this extension is again denoted by $H \rightsquigarrow H\cdot X$. Using 4.7, we deduce:

(3) If $H^n \to H$ in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$ (e.g. if $H^n \to H$ pointwise and $|H^n| \le K\in L^2(X)$), then $H^n\cdot X \to H\cdot X$ in $\mathscr{H}^2$ and in particular $\sup_{(s)}|H^n\cdot X_s - H\cdot X_s| \to 0$ in L^2.

(c) Now we prove that the extension in (b) meets (2), 4.35, 4.36, 4.37. That 4.35 holds is evident. If $H\in L^2(X)$ there are $H^n\in\mathscr{E}'$ going to H in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$, so $(H^n\cdot X)_t^2 \to (H\cdot X)_t^2$ in L^1 by (3), whereas $(H^n)^2\cdot\langle X,X\rangle_t \to (H)^2\cdot\langle X,X\rangle_t$ in L^1 by hypothesis; therefore, since by (2) the processes $(H^n\cdot X)^2 - (H^n)^2\cdot\langle X,X\rangle$ are martingales, we deduce that $(H\cdot X)^2 - (H)^2\cdot\langle X,X\rangle$ is also a martingale: thus the characterization of the bracket yields that (2) holds indeed for all $H\in L^2(X)$.

Let $\mathscr{H}$ be the set of all bounded predictable H for which 4.36 holds: $\mathscr{H}$ obviously is a linear space which contains the vector lattice $\mathscr{E}'$, and is stable under pointwise convergence of uniformly bounded sequences by (3). So a monotone class argument shows that $\mathscr{H}$ is the set of all bounded predictable processes. Next, if $H\in L^2(X)$, set $H^n = H 1_{\{|H|\le n\}}$: so $H^n \to H$ pointwise (hence $H^n\Delta X \to H\Delta X$) and in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$ as well, hence $\Delta(H^n\cdot X) \to (H\Delta X)$ uniformly on finite intervals in L^2 by (3): thus $\Delta(H\cdot X)$ and $H\Delta X$ are indistinguishable, and 4.36 holds.

If $K, H\in\mathscr{E}'$ then $K\cdot(H\cdot X) = (KH)\cdot X$ is trivial. Let $K\in\mathscr{E}'$ and $H\in L^2(X)$ and $H^n \to H$ in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$ with $H^n\in\mathscr{E}'$. Then $H^n\cdot X \to H\cdot X$ in $\mathscr{H}^2$, and as an easy consequence of the "explicit form" (1) of the integrals of $K\in\mathscr{E}'$ with respect to any element of $\mathscr{H}^2$ we easily deduce that $K\cdot(H^n\cdot X)_t \to K\cdot(H\cdot X)_t$ in L^2; moreover $KH^n \to KH$ in $L^2(\Omega\times\mathbb{R}_+,\mathscr{P},m)$ as well, so $(KH^n)\cdot X \to (KH)\cdot X$ in $\mathscr{H}^2$ by (3) again: therefore we deduce that $K\cdot(H\cdot X) = (KH)\cdot X$. Finally, the same approximation argument, applied to K this time, shows that $K\cdot(H\cdot X) = (KH)\cdot X$ holds for all $H\in L^2(X)$ and $K\in L^2(H\cdot X)$.

Let T be a stopping time, and set $T_n = k/2^n$ if $(k-1)/2^n \le T < k/2^n$ and $k \le n2^n$, $T_n = n2^n$ otherwise. Then T_n is a stopping time and $1_{[\![0,T_n]\!]} \in \mathscr{E}'$, and $X^{T_n} = X_0 + 1_{[\![0,T_n]\!]} \cdot X$ is obvious. Letting $n \uparrow \infty$ and using (3) again, we get $X^T = X_0 + 1_{[\![0,T]\!]} \cdot X$. Finally, the second claim in 4.37 follows from applying the third one to $K = 1_{[\![0,T]\!]}$ and from the first one.

(d) Now we localize the previous results. Let $X \in \mathscr{H}^2_{\text{loc}}$ and $H \in L^2_{\text{loc}}(X)$. There is a localizing sequence (T_n) such that $X^{T_n} \in \mathscr{H}^2$ and that $H \in L^2(X^{T_n})$. Then 4.37 (proved above) yields $(H \cdot X^{T_{n+1}})^{T_n} = H 1_{[\![0,T_n]\!]} \cdot X^{T_{n+1}} = H \cdot X^{T_n}$. Therefore there is a unique process $H \cdot X \in \mathscr{H}^2_{\text{loc}}$ such that

$$(H \cdot X)^{T_n} = H \cdot X^{T_n} \quad \text{for all } n \in \mathbb{N}^*.$$

Moreover, properties (i, ii, iii') of 4.40, and 4.35, 4.36, 4.37 and (2) are still valid by localization (with $K \in L^2_{\text{loc}}(H \cdot X)$ in 4.37).

At this stage, we have proved the existence of the desired extension in 4.40, and (a), (b), (d) of 4.40 (4.41 following from (2) by polarization). Uniqueness of the extension is also trivial (by localization again). To finish with 4.40 it remains to prove 4.40b. If $H \cdot X \in \mathscr{H}^2$ then by 4.41 and 4.2 we have $H^2 \cdot \langle X, X \rangle \in \mathscr{A}^+$, so $H \in L^2(X)$. Conversely assume that $H \in L^2(X)$. Then Doob's inequality (1.43) yields, if (T_n) is a localizing sequence for $H \cdot X \in \mathscr{H}^2_{\text{loc}}$:

$$\begin{aligned} E\Big(\sup_{(s)} |H \cdot X_s|^2\Big) &= \lim_n \uparrow E\Big(\sup_{s \le T_n} |H \cdot X_s|^2\Big) \le 4 \lim_n \uparrow E[(H \cdot X_{T_n})^2] \\ &= 4 \lim_n \uparrow E(H^2 \cdot \langle X, X \rangle_{T_n}) = 4E(H^2 \cdot \langle X, X \rangle_\infty) < \infty \end{aligned}$$

and thus $H \cdot X$ belongs to $\mathscr{H}^2$.

(e) In our last step we prove Theorem 4.31 and properties 4.33 to 4.38. Let $X \in \mathscr{S}$. According to 4.17 and 4.21,

$$X = X_0 + M + A, \quad M \in \mathscr{H}^2_{\text{loc}}, \quad A \in \mathscr{V}. \tag{4}$$

Then for all locally bounded predictable processes H we set

$$H \cdot X = H \cdot M + H \cdot A.$$

Due to the uniqueness of the extension in (a) and (d) above, this process $H \cdot X$ does not depend upon the decomposition (4) (up to indistinguishability, of course), and it clearly meets 4.35, 4.36, 4.37. Moreover this is an extension of the map $H \rightsquigarrow H \cdot X$ defined by 4.30 which meets (i, ii, iii) and the last claim of 4.31. Furthermore, owing to the uniqueness in (a) and (d) again, this extension is clearly unique, up to evanescence.

Since for all $H \in \mathscr{E}'$ the map $X \rightsquigarrow H \cdot X$ is obviously linear, Property 4.33 is a trivial consequence of the uniqueness proved above.

Next, 4.34a is trivial, 4.34c follows from (a) above, and 4.34b goes as such: let $X \in \mathscr{M}_{\text{loc}}$ with a decomposition (4); then $H \cdot M \in \mathscr{H}^2_{\text{loc}}$ (see (d)) and $A \in \mathscr{V} \cap \mathscr{L}$, so $H \cdot A \in \mathscr{L}$ by (a), hence $H \cdot X \in \mathscr{L}$.

Finally it remains to prove 4.38. Firstly, the increasing process $|Y| 1_{[\![T,\infty[\![}$ is predictable and finite-valued, so it is locally bounded (see 3.10), and $H = Y 1_{[\![T]\!]}$ *a*-fortiori is locally bounded. Using once more a monotone class argument based upon 4.31(iii) we obtain that it is enough to prove the result when Y takes only finitely many values, or even when $Y = 1_A$ for some $A \in \mathscr{F}_{T-}$. Replacing T by T_A ($= T$ on A, $= \infty$ on A^c) and noting that T_A is still predictable (see 2.10), that amounts to prove that $1_{[\![T]\!]} \cdot X = \Delta X_T 1_{[\![T,\infty[\![}$ for all predictable times T. But then, if (T_n) is an a.s. announcing sequence for T (see 2.16), $1_{]\!]T_n,T]\!]} \to 1_{[\![T]\!]}$ pointwise, outside an evanescent set, and $1_{]\!]T_n,T]\!]} \cdot X = X^T - X^{T_n}$ by 4.37. Therefore the last claim in 4.31 yields that $X^T - X^{T_n}$ converges toward $1_{[\![T]\!]} \cdot X$ uniformly on all finite intervals in measure as $n \uparrow \infty$, whereas it also goes to $\Delta X_T 1_{[\![T,\infty[\![}$ pointwise outside an evanescent set (by definition of (T_n)): hence we deduce 4.38.

3. We end this subsection by showing that the stochastic integral of a predictable process that is *càg* may be approximated by Riemann sums. We will give a somewhat sophisticated version of that result, to be used later.

We call an *adapted subdivision* any sequence $\tau = (T_n)_{n \in \mathbb{N}}$ of stopping times with $T_0 = 0$, and $\sup_n T_n < \infty$, and $T_n < T_{n+1}$ on the set $\{T_n < \infty\}$ (a *deterministic* subdivision if all T_n's are constant). The *τ-Riemann approximant* of $H \cdot X$ is the process $\tau(H \cdot X)$ defined by

4.42 $$\tau(H \cdot X)_t = \sum_{n \in \mathbb{N}} H_{T_n}(X_{T_{n+1} \wedge t} - X_{T_n \wedge t}).$$

4.43 A sequence $(\tau_n = (T(n,m))_{m \in \mathbb{N}})_{n \in \mathbb{N}}$ of adapted subdivisions is called a *Riemann sequence* if $\sup_{m \in \mathbb{N}} [T(n,m+1) \wedge t - T(n,m) \wedge t] \to 0$ for all $t \in \mathbb{R}_+$ (that is, if the mesh of the restriction of the subdivisions τ_n to each interval $[0,t]$ tends to 0).

4.44 **Proposition.** *Let X be a semimartingale, H be a càg adapted process, and (τ_n) a Riemann sequence of adapted subdivisions. Then the τ_n-Riemann approximants $\tau_n(H \cdot X)$ converge to $H \cdot X$, in measure uniformly on each compact interval.*

Proof. If $\tau_n = (T(n,m))_{m \in \mathbb{N}}$, define H^n by
$$H^n = \sum_{m \in \mathbb{N}} H_{T(n,m)} 1_{]\!]T(n,m),\, T(n,m+1)]\!]}.$$
Then H^n is predictable, converges pointwise to H because H is càg, and if $K_t = \sup_{s \le t} |H_s|$ then K is adapted, càg, locally bounded, and $|H^n| \le K$. Then the result follows from 4.31 and from the easily checked property that $\tau_n(H \cdot X) = H^n \cdot X$. □

§ 4e. Quadratic Variation of a Semimartingale and Ito's Formula

1. We first define the quadratic variation.

4.45 **Definition.** The *quadratic co-variation* of the two semimartingales X and Y (the *quadratic variation* of X, when $Y = X$) is the following process:

$$[X, Y] = XY - X_0 Y_0 - X_- \cdot Y - Y_- \cdot X$$

(it is defined uniquely, up to an evanescent set). □

The next theorem explains the terminology "quadratic variation". Before, we state the following obvious properties:

4.46
$$\begin{cases} [X, Y]_0 = 0, \quad [X, Y] = [X - X_0, Y - Y_0] \\ [X, Y] = \frac{1}{4}([X + Y, X + Y] - [X - Y, X - Y]). \end{cases}$$

4.47 **Theorem.** *Let X and Y be two semimartingales.*

a) *For any Riemann sequence $\{\tau_n = (T(n, m))_{m \in \mathbb{N}}\}_{n \in \mathbb{N}}$ of adapted subdivisions* (see 4.43), *the processes $S_{\tau_n}(X, Y)$ defined by*

4.48
$$S_{\tau_n}(X, Y)_t = \sum_{m \geq 1} (X_{T(n,m+1)\wedge t} - X_{T(n,m)\wedge t})(Y_{T(n,m+1)\wedge t} - Y_{T(n,m)\wedge t})$$

converge to the process $[X, Y]$, in measure, uniformly on every compact interval.

b) *$[X, Y] \in \mathscr{V}$ and $[X, X] \in \mathscr{V}^+$.*

c) *$\Delta[X, Y] = \Delta X \Delta Y$.*

Proof. In view of 4.45 and of the obvious polarization properties $S_\tau(X, Y) = \frac{1}{4}[S_\tau(X + Y, X + Y) - S_\tau(X - Y, X - Y)]$ and $\Delta X \Delta Y = \frac{1}{4}[(\Delta X + \Delta Y)^2 - (\Delta X - \Delta Y)^2]$ it suffices to prove the claims when $Y = X$.

Since $(x - y)^2 = x^2 - y^2 - 2y(x - y)$, one deduces from 4.42 and 4.48 that $S_{\tau_n}(X, X) = X^2 - X_0^2 - 2\tau_n(X_- \cdot X)$, thus (a) follows from 4.44.

$S_{\tau_n}(X, X)$ is increasing, so we deduce from (a) that $[X, X]$ also is increasing; since it is càdlàg adapted with $[X, X]_0 = 0$, we obtain $[X, X] \in \mathscr{V}^+$, and (b) is proved. It follows from the definition 4.45 and from 4.36 that $\Delta[X, X] = \Delta(X^2) - 2X_- \Delta X$, and since $\Delta(X^2) = (\Delta X)^2 + 2X_- \Delta X$ we deduce that $\Delta[X, X] = (\Delta X)^2$ and (c) is proved. □

We shall give later (in 4.52) a sort of "explicit" formula that complements (c) above. In the meantime we examine various useful properties of the quadratic variation.

4.49 **Proposition.** *Let $X \in \mathscr{S}$ and $Y \in \mathscr{V}$.*

a) *$[X, Y] = \Delta X \cdot Y$ and $XY = Y_- \cdot X + X \cdot Y$.*

b) *If Y is predictable, then $[X, Y] = \Delta Y \cdot X$ and $XY = Y \cdot X + X_- \cdot Y$.*

c) *If Y is predictable and X is a local martingale, then $[X, Y]$ is a local martingale.*

d) *If Y or X is continuous, then $[X, Y] = 0$.*

Proof. a) We first construct a Riemann sequence τ_n as such:

$$T(n, 0) = 0,$$

$$T(n, m + 1) = \inf\left(t > T(n, m): |X_t - X_{T(n,m)}| > \frac{1}{n} \quad \text{or} \quad t > T(n, m) + \frac{1}{n}\right).$$

Then $\sup_{T(n,m)<s<T(n,m+1)}|X_s - X_{T(n,m)}| \le 1/n$, and so for all $t \in \mathbb{R}_+$ such that $|\Delta X_t| > 2/n$ there exists $m \in \mathbb{N}^*$ with $t = T(n,m)$ (m depends on ω, of course), while $|X_{T(n,m+1)} - X_{T(n,m)}| \le 3/n$ whenever $|\Delta X_{T(n,m+1)}| \le 2/n$. Hence

$$|X_{T(n,m+1)} - X_{T(n,m)} - \Delta X_{T(n,m+1)} 1_{\{|\Delta X_{T(n,m+1)}| > 2/n\}}| \le \frac{3}{n}.$$

Therefore the process $A^n = (\Delta X 1_{\{|\Delta X| > 2/n\}}) \cdot Y$, which is in $\mathscr{V}$, has:

$$|S_{\tau_n}(X, Y)_t - A_t^n| \le \frac{3}{n} \operatorname{Var}(Y)_t.$$

Letting $n \uparrow \infty$, we get that $A_t^n \to \Delta X \cdot Y_t$ by Lebesgue convergence theorem; then 4.47a implies $[X, Y] = \Delta X \cdot Y$, which in turn, together with the definition 4.45, implies $XY = Y_- \cdot X + X \cdot Y$.

b) Y is locally bounded (see 3.10), so $Y \cdot X$ and $\Delta Y \cdot X$ are well defined. Let (T_n) be a sequence of predictable times that exhausts the jumps of Y (see 2.24). Then 4.38 yields

$$\left\{\Delta Y \sum_{p \le n} 1_{[\![T_p]\!]}\right\} \cdot X = \sum_{p \le n} \Delta Y_{T_p} \Delta X_{T_p} 1_{[\![T_p, \infty[\![} = \left\{\Delta X \sum_{p \le n} 1_{[\![T_p]\!]}\right\} \cdot Y.$$

Letting $n \uparrow \infty$, the left-hand side above goes to $\Delta Y \cdot X$ and the right-hand side goes to $\Delta X 1_{\{\Delta Y \neq 0\}} \cdot Y$, which equals $\Delta X \cdot Y$ (a trivial property of Stieltjes integrals: the set $\{\Delta X \neq 0\}$ is thin, so $\Delta X \cdot Y = \sum_{s \le .} \Delta X_s \Delta Y_s$). Hence $\Delta Y \cdot X = \Delta X \cdot Y$ and the first claim follows from (a). Moreover, $X \cdot Y = X_- \cdot Y + \Delta X \cdot Y$ (because $Y \in \mathscr{V}$) and $Y_- \cdot X = Y \cdot X - \Delta Y \cdot X$ (because Y is predictable), so the second claim follows from (a) again and from $\Delta Y \cdot X = \Delta X \cdot Y$.

c) This is a trivial consequence of (b) and 4.34b.

d) This follows from (b). □

4.50 **Proposition.** *Let X and Y be two local martingales.*

a) *$XY - X_0 Y_0 - [X, Y]$ is a local martingale.*

b) *If $X, Y \in \mathscr{H}^2_{\mathrm{loc}}$ then $[X, Y]$ belongs to $\mathscr{A}_{\mathrm{loc}}$ and its compensator is $\langle X, Y \rangle$; if moreover $X, Y \in \mathscr{H}^2$, then $XY - [X, Y]$ belongs to $\mathscr{M}$.*

c) *X belongs to $\mathscr{H}^2$ (resp. $\mathscr{H}^2_{\mathrm{loc}}$) if and only if $[X, X]$ belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{\mathrm{loc}}$) and X_0 is square-integrable.*

d) *$X = X_0$ a.s. if and only if $[X, X] = 0$.*

Proof. (a) readily follows from Definition 4.45 of $[X, Y]$ and from 4.34b. We prove (b) and (c) together. By polarization, we may assume that $Y = X$ in (b). By localization it suffices to consider the case $X \in \mathscr{H}^2$, or $[X, X] \in \mathscr{A}$.

Assume first that $X \in \mathscr{H}^2$. Then $X_0 \in L^2$ is obvious. Furthermore, both processes $X^2 - X_0^2 - [X, X]$ by (a) and $X^2 - X_0^2 - \langle X, X \rangle$ by 4.2 belong to $\mathscr{L}$, hence $[X, X] - \langle X, X \rangle \in \mathscr{L} \cap \mathscr{V}$ and thus $[X, X] \in \mathscr{A}_{\mathrm{loc}}$ and $\langle X, X \rangle$ is the compensator of $[X, X]$. We also deduce that $E([X, X]_\infty) = E(\langle X, X \rangle_\infty)$, so indeed $[X, X] \in \mathscr{A}$. Finally, the latter actually yields that $[X, X] - \langle X, X \rangle \in \mathscr{M}$, so in view of 4.2 we also have that $X^2 - [X, X] \in \mathscr{M}$.

Conversely assume that $[X,X] \in \mathscr{A}$ and $X_0 \in L^2$, and let (T_n) be a localizing sequence for the local martingale $X^2 - X_0^2 - [X,X]$. Then

$$\sup_t E(X_t^2) = \sup_t \lim_n \uparrow E[(X_{t \wedge T_n})^2] = \sup_t \lim_n \uparrow E(X_0^2 + [X,X]_{T_n \wedge t}) < \infty$$

and thus $X \in \mathscr{H}^2$.

As for (d), the necessary part is obvious; the sufficient part follows from 4.13a, once noticed that $[X,X] = 0$ implies $X^2 - X_0^2 \in \mathscr{L}$ by (a), so $X - X_0$ is orthogonal to itself. □

4.51 **Lemma.** *Let X be a purely discontinuous martingale* (recall 4.11) *belonging to $\mathscr{H}^2$. Then $[X,X]_t = \sum_{s \leq t} (\Delta X_s)^2$* (we will see later that the same result is true when X is any purely discontinuous local martingale).

Proof. Set $H = \Delta X$. In view of 4.47c, $\sum_{(s)} (H_s)^2 \leq [X,X]_\infty$, which is integrable by 4.50c. Moreover by 2.31 the predictable projection ${}^pH$ of H equals 0, and $H_0 = 0$.

a) For further reference, we will forget for a moment that H is the jump process of X. Define the stopping times $R(n,m)$ by $R(n,0) = 0$, $R(n,m+1) = \inf(t > R(n,m): |H_t| \geq 1/n)$. Since $\sum_{(s)} (H_s)^2 < \infty$ a.s., the thin optional set $D = \bigcup_{n,m} [\![R(n,m)]\!]$ (see 1.30) is equal to the set $\{H \neq 0\}$, up to an evanescent set.

We call J the predictable support of D: it is a thin predictable set by 2.34. Then, due to 1.31 and 2.23, there are stopping times $(S_n)_{n \geq 1}$ and predictable times $(T_n)_{n \geq 1}$ such that $D \backslash J = \bigcup_n [\![S_n]\!]$ and $J = \bigcup_n [\![T_n]\!]$, up to an evanescent set, and the graphs $([\![S_n]\!], [\![T_m]\!])_{n,m \geq 1}$ are pairwise disjoint: so in particular $\{H \neq 0\} \subset (\bigcup_n [\![S_n]\!]) \cup (\bigcup_n [\![T_n]\!])$ up to an evanescent set, and by definition of J all S_n are totally inaccessible.

Set $A^n = H_{S_n} 1_{[\![S_n, \infty[\![}$, which belongs to $\mathscr{A}$. If $A^{n,p}$ is its compensator, then $M^n = A^n - A^{n,p}$ is a martingale. Moreover by 3.21, $\Delta A^{n,p} = {}^p(\Delta A^n) = {}^p(H_{S_n} 1_{[\![S_n]\!]}) = 0$ because S_n is totally inaccessible, so $\Delta M^n = \Delta A^n$.

Set $N^n = H_{T_n} 1_{[\![T_n, \infty[\![}$. For all stopping times T,

$$E(N_T^n) = E(H_{T_n} 1_{\{T_n < \infty, T_n \leq T\}}) = E(({}^pH)_{T_n} 1_{\{T_n < \infty, T_n \leq T\}}) = 0$$

(the second equality comes from $\{T_n < \infty, T_n \leq T\} \in \mathscr{F}_{T_n -}$ and from the definition of ${}^pH$). Hence N^n is a (square-integrable) martingale by 1.44.

Now we set $B_n = \bigcup_{1 \leq m \leq n} ([\![S_m]\!] \cup [\![T_m]\!])$ and $Y^n = \sum_{1 \leq m \leq n} (M^m + N^m)$. Then $Y^n \in \mathscr{L} \cap \mathscr{V}$ and $\Delta Y^n = H 1_{B_n}$ by what precedes. Therefore 4.49a yields for $p \leq n$:

$$[Y^n - Y^p, Y^n - Y^p]_t = \sum_{s \leq t} (H_s)^2 1_{B_n \backslash B_p}(s)$$

which belongs to $\mathscr{A}$, including when $p = 0$ (by convention $Y^0 = 0$ and $B_0 = \emptyset$). Hence 4.50b,c gives $Y^n \in \mathscr{H}^2$ and

$$\|Y^n - Y^p\|_{H^2}^2 = E[(Y_\infty^n - Y_\infty^p)^2] = E([Y^n - Y^p, Y^n - Y^p]_\infty)$$

$$= E\left(\sum_{(s)} (H_s)^2 1_{B_n \backslash B_p}(s)\right).$$

The above tends to 0 as $n, p \uparrow \infty$ (because $E(\sum_{(s)}(H_s)^2) < \infty$). Therefore Y^n goes to a martingale Y in the Hilbert space $\mathscr{H}^2$ as $n \uparrow \infty$, and in particular $\Delta Y^n \to \Delta Y$ uniformly on finite intervals, in L^2 (use 4.7). Thus, since $\{H \neq 0\} \subset \bigcup_n B_n$, we deduce that $\Delta Y = H$.

b) In addition to what precedes, the processes $(Y^n)^2 - [Y^n, Y^n]$ are martingales and converge in L^1, for each time t, to $Y_t^2 - \sum_{s \leq t}(H_s)^2$: hence $Y^2 - \sum_{s \leq \cdot}(H_s)^2$ itself is a martingale. Furthermore each Y^n is a purely discontinuous martingale by 4.14b, so Y also is purely discontinuous by 4.16.

Now we return to our problem. X and Y are two purely discontinuous martingales, and $\Delta X = \Delta Y = H$ up to an evanescent set: hence $X = Y$ up to an evanescent set by 4.19. In particular, $X^2 - \sum_{s \leq \cdot}(\Delta X_s)^2 \in \mathscr{L}$. On the other hand, $X^2 - [X,X] \in \mathscr{L}$ by 4.50a, so $C = [X,X] - \sum_{s \leq \cdot}(\Delta X_s)^2 \in \mathscr{L}$. But in view of 4.47c, C is a continuous element of $\mathscr{V}$, hence it equals 0 (up to an evanescent set) by 3.16, and we are done. □

4.52 **Theorem.** *If X, Y are semimartingales, and if X^c, Y^c denote their continuous martingale parts, then*

4.53
$$[X, Y]_t = \langle X^c, Y^c \rangle_t + \sum_{s \leq t} \Delta X_s \Delta Y_s.$$

Proof. By polarization it is enough to consider the case $Y = X$. We have a decomposition $X = X_0 + X^c + M + A$ where $A \in \mathscr{V}$ and $M \in \mathscr{H}^2_{\text{loc}}$, and by localization we may assume further that $M \in \mathscr{H}^2$. Then

(1) $$[X,X] = [X^c,X^c] + 2[X^c,M] + 2[X^c,A] + [M,M] + 2[M,A] + [A,A].$$

First of all, $[X^c, X^c]$ is continuous by 4.47c, so $[X^c, X^c] = \langle X^c, X^c \rangle$ is deduced from 4.50b. Secondly the previous lemma yields $[M,M] = \sum_{s \leq \cdot}(\Delta M_s)^2$, while $[A,A] = \sum_{s \leq \cdot}(\Delta A_s)^2$ and $[M,A] = \sum_{s \leq \cdot} \Delta M_s \Delta A_s$ by 4.49a: so the sum of the last three terms in (1) equals $\sum_{s \leq \cdot}(\Delta X_s)^2$. It remains to prove that $[X^c, A] = 0$ (which follows from 4.49a again) and that $[X^c, M] = 0$: but X^c and M are orthogonal, since by definition of X^c the local martingale M is purely discontinuous, so $\langle X^c, M \rangle = 0$ (see 4.2 or 4.15). Furthermore $[X^c, M]$, being continuous by 4.47c, equals $\langle X^c, M \rangle$ by 4.50b: hence the claim. □

Finally we state some easy consequences:

4.54 If X, Y are semimartingales and H is a locally bounded predictable process, then $[H \cdot X, Y] = H \cdot [X, Y]$ (follows from 4.53 and 4.36 and 4.41). □

4.55 **Corollary.** *Let X, Y be local martingales.*

a) *The process $[X,X]^{1/2}$ belongs to $\mathscr{A}_{\text{loc}}$.*

b) *$[X, Y] = 0$ whenever X is continuous and Y purely discontinuous.*

c) *$[X, Y] = \langle X, Y \rangle = 0$ whenever X and Y are continuous and orthogonal.*

d) *If X is continuous (resp. purely discontinuous) and if H is a locally bounded predictable process, then $H \cdot X$ is a continuous (resp. purely discontinuous) local martingale.*

Proof. a) Consider a decomposition $X = X_0 + X' + X''$ of type 4.17, so $X'' \in \mathscr{H}^2_{\text{loc}}$ and $X' \in \mathscr{A}_{\text{loc}}$, and let (T_n) be a localizing sequence for both X' and X'' (so $X''^{T_n} \in \mathscr{H}^2$ and $X'^{T_n} \in \mathscr{A}$). Set $S_n = \inf(t : t \geq T_n \text{ or } [X,X]_t \geq n)$, so $S_n \uparrow \infty$ as $n \uparrow \infty$. Then

$$[X,X]^{1/2}_{S_n} \leq \sqrt{n} + |\Delta X_{S_n}| \leq \sqrt{n} + |\Delta X'^{T_n}_{S_n}| + |\Delta X''^{T_n}_{S_n}|$$

and both $\Delta X'^{T_n}_{S_n}$ and $\Delta X''^{T_n}_{S_n}$ are integrable by construction: so $[X,X]^{1/2}_{S_n}$ is integrable.

(b) obviously follows from 4.53 (because $Y^c = 0$ and $X = 0$), and (c) follows from 4.53 and 4.15. The first claim in (d) is obvious (see 4.36), and the second claim comes from 4.53 and 4.54 (if X is purely discontinuous, $\langle (H \cdot X)^c, (H \cdot X)^c \rangle = 0$, which implies $(H \cdot X)^c = 0$). □

2. *Jump process of a local martingale.* For further reference, we characterize below the structure of the "jump process" ΔX of a local martingale.

4.56 **Theorem.** *Let H be an optional process with $H_0 = 0$.*

a) *There is a square-integrable martingale (resp. locally square-integrable martingale) X such that ΔX and H are indistinguishable, if and only if the predictable projection ${}^pH$ is 0 and the increasing process $\sum_{s \leq \cdot} (H_s)^2$ is integrable (resp. locally integrable).*

b) *There exists a local martingale $X \in \mathscr{A}_{\text{loc}}$ (resp. $\mathscr{A}$) such that ΔX and H are indistinguishable, if and only if ${}^pH = 0$ and $\sum_{s \leq \cdot} |H_s| \in \mathscr{A}^+_{\text{loc}}$ (resp. $\mathscr{A}^+$).*

c) *There exists a local martingale X such that ΔX and H are indistinguishable, if and only if ${}^pH = 0$ and $[\sum_{s \leq \cdot} (H_s)^2]^{1/2} \in \mathscr{A}^+_{\text{loc}}$.*

Proof. a) Let $X \in \mathscr{H}^2$ (resp. $\mathscr{H}^2_{\text{loc}}$) and $H = \Delta X$. Then ${}^pH = 0$ by 2.31 and $\sum_{s \leq t} (H_s)^2 \leq [X,X]_t$ by 4.53, so $\sum_{s \leq t} (H_s)^2$ is in $\mathscr{A}$ (resp. $\mathscr{A}_{\text{loc}}$) by 4.50c.

Conversely assume that ${}^pH = 0$ and $\sum_{s \leq \cdot} (H_s)^2 \in \mathscr{A}$. Then part (a) of the proof of Lemma 4.51 shows the existence of $Y \in \mathscr{H}^2$ such that $H = \Delta Y$ up to an evanescent set. The case $\sum_{s \leq \cdot} (H_s)^2 \in \mathscr{A}_{\text{loc}}$ follows by localization.

b) If $X \in \mathscr{L} \cap \mathscr{A}$ (resp. $\mathscr{L} \cap \mathscr{A}_{\text{loc}}$) and $H = \Delta X$, then ${}^pH = 0$ by 2.31 and $\sum_{s \leq \cdot} |H_s| \leq \text{Var}(X)$ belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{\text{loc}}$). Conversely, assume that ${}^pH = 0$ and $\sum_{s \leq \cdot} |H_s| \in \mathscr{A}$ (resp. $\mathscr{A}_{\text{loc}}$); set $A_t = \sum_{s \leq t} H_s$, which belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{\text{loc}}$) as well, and $X = A - A^p$ where A^p is the compensator of A: then $X \in \mathscr{L} \cap \mathscr{A}$ (resp. $\mathscr{L} \cap \mathscr{A}_{\text{loc}}$), and $\Delta X = \Delta A - \Delta A^p = H - {}^pH = H$ by 3.21.

c) Let $X \in \mathscr{M}_{\text{loc}}$ and $H = \Delta X$. Then ${}^pH = 0$ by 2.31 again, and $\sum_{s \leq t} (H_s)^2 \leq [X,X]_t$ by 4.53, so 4.55a yields $(\sum_{s \leq \cdot} (H_s)^2)^{1/2} \in \mathscr{A}_{\text{loc}}$.

Conversely, let H be an optional process with ${}^pH = 0$ and $A^{1/2} \in \mathscr{A}^+_{\text{loc}}$, where $A_t = \sum_{s \leq t} (H_s)^2$. Set $K = H 1_{\{|H| > 1\}}$, $H'' = K - {}^pK$ and $H' = H - H''$, so ${}^pH' =$

${}^pH'' = 0$. Set also $B_t = \sum_{s\le t} |K_s|$, which clearly belongs to $\mathscr{V}^+$. Since $\Delta B \le |\Delta A|^{1/2}$ we deduce from the property $A^{1/2} \in \mathscr{A}^+_{\text{loc}}$ that $B \in \mathscr{A}^+_{\text{loc}}$, and $\sum_{s\le \cdot} |{}^pK_s| \le B^p$ (by 3.21), which also belongs to $\mathscr{A}^+_{\text{loc}}$. Thus $\sum_{s\le \cdot} |H''_s| \in \mathscr{A}^+_{\text{loc}}$ and by (b) there is $X'' \in \mathscr{M}_{\text{loc}}$ with $\Delta X'' = H''$.

Since $|H'|^2 \le 2|H|^2 + 2|H''|^2$, we get $C_t := \sum_{s\le t} |H'_s|^2 \le 2A_t + 2\sum_{s\le t} |H''_s|^2$, so $C_t < \infty$ for $t \in \mathbb{R}_+$. Moreover, since ${}^pH = 0$ we have ${}^pK = -{}^p(H 1_{\{|H|\le 1\}})$, so $|{}^pK| \le 1$, and $|H'| \le 2$ by construction: therefore $\Delta C_t \le 4$, and we deduce that $C \in \mathscr{A}^+_{\text{loc}}$. Then (a) yields a local martingale X' with $\Delta X' = H'$. Hence $X = X' + X''$ meets $\Delta X = H$. □

3. Now we turn to *Ito's formula*. In the following, $D_i f$ and $D_{ij} f$ denote the partial derivatives $\partial f/\partial x^i$ and $\partial^2 f/\partial x^i \partial x^j$.

4.57 **Theorem.** *Let $X = (X^1, \ldots, X^d)$ be a d-dimensional semimartingale, and f a class C^2 function on $\mathbb{R}^d$. Then $f(X)$ is a semimartingale and we have:*

4.58
$$\begin{aligned} f(X_t) = f(X_0) &+ \sum_{i\le d} D_i f(X_-) \cdot X^i + \frac{1}{2} \sum_{i,j\le d} D_{ij} f(X_-) \cdot \langle X^{i,c}, X^{j,c} \rangle \\ &+ \sum_{s\le t} \left[f(X_s) - f(X_{s-}) - \sum_{i\le d} D_i f(X_{s-}) \Delta X^i_s \right] \end{aligned}$$

Of course, this formula implicitely means that all terms are well-defined. In particular the last two terms are processes with finite variation (the first one is continuous, the second one is "purely discontinuous").

Formula 4.58 is also valid when f is complex-valued: take the real and purely imaginary parts separately.

Proof. To simplify notation somewhat, with any C^2 function f on $\mathbb{R}^d$ we associate the C^1 function $\hat{f}$ on $\mathbb{R}^d \times \mathbb{R}^d$ defined by

$$\hat{f}(x,y) = f(x) - f(y) - \sum_{j\le d} D_j f(y)(x^j - y^j),$$

where x^j denotes the j^{th} component of x.

(i) We first prove the result when f is a polynomial on $\mathbb{R}^d$. It suffices to consider the case of monomials and, by induction on the degree and since the result is trivially true for constant functions, it suffices to prove the following: let g be a function meeting $g(X) \in \mathscr{S}$ and 4.58, then $f(x) = x^k g(x)$ also satisfies $f(X) \in \mathscr{S}$ and 4.58.

Since $g(X) \in \mathscr{S}$ and $X^k \in \mathscr{S}$, we have $f(X) \in \mathscr{S}$ by 4.47b and 4.45, and we also have $f(X) = f(X_0) + X^k_- \cdot g(X) + g(X_-) \cdot X^k + [X^k, g(X)]$. Now g satisfies 4.58, hence (using several times 4.36 and 4.37) we obtain:

$$\begin{aligned} (1)\quad f(X) = f(X_0) &+ \sum_{i\le d} (X^k_- D_i g(X_-)) \cdot X^i + \frac{1}{2} \sum_{i,j\le d} (X^k_- D_{ij} g(X_-)) \cdot \langle X^{i,c}, X^{j,c} \rangle \\ &+ \sum_{s\le \cdot} X^k_{s-} \hat{g}(X_s, X_{s-}) + g(X_-) \cdot X^k + [X^k, g(X)]. \end{aligned}$$

Now, if $A \in \mathscr{V}$, we have $[X^k, A] = \sum_{s \leq \cdot} \Delta X^k_s \Delta A_s$ by 4.49a, and the two last terms in 4.58 for g are processes in $\mathscr{V}$; then, using 4.49 and 4.54,

$$(2) \qquad [X^k, g(X)] = \sum_{i \leq d} D_i g(X_-) \cdot [X^k, X^i] + \sum_{s \leq \cdot} \Delta X^k_s \hat{g}(X_s, X_{s-})$$
$$= \sum_{i \leq d} D_i g(X_-) \cdot \langle X^{k,c}, X^{i,c} \rangle + \sum_{s \leq \cdot} \Delta X^k_s (g(X_s) - g(X_{s-})).$$

Then, putting together (1) and (2), we obtain that f satisfies 4.58, once noticed that $\hat{f}(x,y) = y^k \hat{g}(x,y) + (x^k - y^k)(g(x) - g(y))$, and that for $i \neq k$ and $j \neq k$ one has $D_i f = x^k D_i g$, $D_k f = g + x^k D_k g$, $D_{ij} f = x^k D_{ij} g$, $D_{ik} f = D_i g + x^k D_{ik} g$, and $D_{kk} f = 2 D_k g + x^k D_{kk} g$.

(ii) We turn now to the general case. Let $T_n = \inf(t \colon |X_t| > n)$. For each $n \in \mathbb{N}$ we can find a sequence $(g_{nm})_{m \in \mathbb{N}}$ of polynomials, that converge as well as their partial derivatives of first and second order, to f and its partial derivatives of first and second order, uniformly on the ball $\{x \colon |x| \leq n\}$. There exists a constant K_n such that:

$$|x|, |y| \leq n \Rightarrow |\hat{f}(x,y)| \leq K_n |x - y|^2 \quad \text{and} \quad |\hat{g}_{nm}(x,y)| \leq K_n |x - y|^2.$$

Then these inequalities, and the fact that $\sum_{s \leq t} |\Delta X_s|^2 < \infty$ for all t by 4.47, and the fact that $\hat{g}_{nm}(x,y) \to \hat{f}(x,y)$ as $m \uparrow \infty$ for $|x|$, $|y| \leq n$ and an application of Lebesgue convergence theorem, show that

$$(3) \qquad \text{if } t < T_n, \quad \text{then } \sum_{s \leq t} |\hat{f}(X_s, X_{s-})| < \infty \quad \text{and}$$
$$\sum_{s \leq t} \hat{g}_{nm}(X_s, X_{s-}) \to \sum_{s \leq t} \hat{f}(X_s, X_{s-}) \quad \text{as } m \uparrow \infty.$$

Similarly,

$$(4) \qquad g_{nm}(X_t) \to f(X_t) \quad \text{as } m \uparrow \infty, \quad \text{for } t < T_n.$$

(5) $D_i g_{nm}(X_-) \cdot X^i_t \to D_i f(X_-) \cdot X^i_t$ in measure as $m \uparrow \infty$ on the set $\{t < T_n\}$ (apply 4.31 to $H^m = D_i g_{nm}(X_-) 1_{[\![0,T_n]\!]}$ which goes to $D_i f(X_-) 1_{[\![0,T_n]\!]}$).

(6) $D_{ij} g_{nm}(X_-) \cdot \langle X^{i,c}, X^{j,c} \rangle_t \to D_{ij} f(X_-) \cdot \langle X^{i,c}, X^{j,c} \rangle_t$ as $m \uparrow \infty$ on the set $\{t < T_n\}$ (apply Lebesgue convergence theorem).

Recall that $\lim_n T_n = \infty$. Then (3) implies that the process $\sum_{s \leq \cdot} \hat{f}(X_s, X_{s-})$ belongs to $\mathscr{V}$. The other terms in the right-hand side of 4.58 being well defined and being semimartingales, we deduce that the right-hand side of 4.58 defines a process that is a semimartingale. Finally, by (i) we know that each g_{nm} meets 4.58 hence (3), (4), (5), (6) show that f also meets 4.58 for all $t < T_n$, for all $n \in \mathbb{N}$. Thus f meets 4.58 everywhere, and the theorem is proved. □

§ 4f. Doléans-Dade Exponential Formula

In this subsection, we present a first application of Ito's formula. We consider the equation

4.59 $$Y = 1 + Y_- \cdot X \quad (\text{or: } dY = Y_-\, dX \text{ and } Y_0 = 1)$$

where X is a *given* semimartingale, and Y is an unknown càdlàg adapted process. By analogy with the ordinary differential equation $\frac{dy}{dx} = y$, we will call the solution Y the *exponential* of X. We shall encounter this equation in two different settings:

1) when X is a local martingale (real-valued)

2) when X is a complex-valued process with finite variation, in which case Equation 4.59 is solved pathwise: that is, for each ω we consider the deterministic equation $Y_t(\omega) = 1 + \int_0^t Y_{s-}(\omega)\, dX_s(\omega)$.

In order to unify the treatment, we consider at once the case where X is a complex-valued semimartingale, that is $X = X' + iX''$ with $X'$and X'' two real-valued semimartingales. Then 4.59 must be read as a system of two "real-valued" equations, namely

4.60 $$\begin{cases} Y' = 1 + Y'_- \cdot X' - Y''_- \cdot X'' \\ Y'' = \phantom{1 +{}} Y''_- \cdot X' + Y'_- \cdot X'' \end{cases}$$

and $Y = Y' + iY''$.

4.61 **Theorem.** *Let $X = X' + iX''$ be a complex-valued semimartingale. Then Equation 4.59 has one and only one (up to indistinguishability) càdlàg adapted solution. This solution is a semimartingale, is denoted by $\mathscr{E}(X)$, and is given by*

4.62 $$\mathscr{E}(X)_t = \{\exp(X_t - X_0 - \tfrac{1}{2}\langle X'^c, X'^c\rangle_t + \tfrac{1}{2}\langle X''^c, X''^c\rangle_t - i\langle X'^c, X''^c\rangle_t)\} \times \prod_{s \le t} [(1 + \Delta X_s)e^{-\Delta X_s}]$$

where the (possibly) infinite product is absolutely convergent. Furthermore,

a) *If X has finite variation, then so has $\mathscr{E}(X)$.*

b) *If X is a local martingale, then so is $\mathscr{E}(X)$.*

c) *Let $T = \inf(t: \Delta X_t = -1)$. Then $\mathscr{E}(X) \neq 0$ on the interval $[\![0, T[\![$, and $\mathscr{E}(X)_- \neq 0$ on the interval $[\![0, T]\!]$, and $\mathscr{E}(X) = 0$ on the interval $[\![T, \infty[\![$.*

In particular, 4.53 yields that when X has finite variation,

4.63 $$\mathscr{E}(X)_t = e^{X_t - X_0} \prod_{s \le t} (1 + \Delta X_s)e^{-\Delta X_s}.$$

When X is a *real-valued* semimartingale, then

4.64 $$\mathscr{E}(X)_t = e^{X_t - X_0 - 1/2\langle X^c, X^c\rangle_t} \prod_{s \le t} (1 + \Delta X_s)e^{-\Delta X_s}.$$

Proof. (i) $\sum_{s \le t} |\Delta X_s|^2 < \infty$ a. s. for $t \in \mathbb{R}_+$ is deduced from 4.47c applied to X' and X''. Hence there are only finitely many times s inside $[0, t]$ such that $|\Delta X_s| > 1/2$ and since $|\mathrm{Log}(1 + x) - x| \le C|x|^2$ for all $x \in \mathbb{C}$ with $|x| \le 1/2$ (C is some constant) it is obvious that the (possibly) infinite product

(1)
$$V_t = \prod_{s \leq t} (1 + \Delta X_s) e^{-\Delta X_s}$$

is absolutely convergent, a.s. Moreover (1) defines a càdlàg adapted process V with finite variation and $V_0 = 1$.

(ii) Set $Z = X - X_0 - \frac{1}{2}\langle X'^c, X'^c\rangle + \frac{1}{2}\langle X''^c, X''^c\rangle - i\langle X'^c, X''^c\rangle$ and $U = Ve^Z$ (that is, $U = \mathscr{E}(X)$ as given by 4.62). Write $Z = Z' + iZ''$, $V = V' + iV''$, so that $U = f(Z', Z'', V', V'')$ where f is the following function on $\mathbb{R}^4$: $f(x, y, u, v) = e^{x+iy}(u + iv)$. Hence Theorem 4.57 implies that U is a semimartingale, and that

(2)
$$\begin{aligned} U_t = 1 &+ U_- \cdot Z'_t + iU_- \cdot Z''_t + e^{Z_-} \cdot V'_t + ie^{Z_-} \cdot V''_t + \tfrac{1}{2} U_- \cdot \langle Z'^c, Z'^c\rangle_t \\ &- \tfrac{1}{2} U_- \cdot \langle Z''^c, Z''^c\rangle_t + iU_- \cdot \langle Z'^c, Z''^c\rangle_t \\ &+ \sum_{s \leq t} [\Delta U_s - U_{s-}(\Delta Z'_s + i\Delta Z''_s) - e^{Z_{s-}}(\Delta V'_s + i\Delta V''_s)] \end{aligned}$$

because $V'^c = V''^c = 0$. Moreover by definition of Z we have $Z'^c = X'^c$ and $Z''^c = X''^c$, so

(3)
$$Z' + iZ'' + \tfrac{1}{2}\langle Z'^c, Z'^c\rangle - \tfrac{1}{2}\langle Z''^c, Z''^c\rangle + i\langle Z'^c, Z''^c\rangle = X - X_0$$

From (1) it is clear that

(4)
$$e^{Z_-} \cdot V'_t + ie^{Z_-} \cdot V''_t = \sum_{s \leq t} e^{Z_{s-}} \Delta V_s.$$

Moreover,

(5)
$$\Delta U_s = e^{Z_{s-} + \Delta Z_s} V_{s-}(1 + \Delta X_s) e^{-\Delta X_s} - e^{Z_{s-}} V_{s-} = U_{s-} \Delta Z_s$$

because $\Delta Z = \Delta X$. Therefore, injecting (3), (4), (5) into (2) yields

$$U = 1 + U_- \cdot (X - X_0) = 1 + U_- \cdot X.$$

(iii) Conversely, let $Y = Y' + iY''$ be a càdlàg adapted solution to 4.59 (or equivalently 4.60). By 4.34 it is a semimartingale. Let $W = Ye^{-Z}$ and apply Ito's formula to the following function on $\mathbb{R}^4$: $f(x, y, u, v) = e^{-(x+iy)}(u + iv)$:

(6)
$$\begin{aligned} W = 1 &- W_- \cdot Z' - iW_- \cdot Z'' + e^{-Z_-} \cdot Y' + ie^{-Z_-} \cdot Y'' + \tfrac{1}{2} W_- \cdot \langle Z'^c, Z'^c\rangle \\ &- \tfrac{1}{2} W_- \cdot \langle Z''^c, Z''^c\rangle + iW_- \cdot \langle Z'^c, Z''^c\rangle - e^{-Z_-} \cdot \langle Z'^c, Y'^c\rangle \\ &- ie^{-Z_-} \cdot \langle Z'^c, Y''^c\rangle - ie^{-Z_-} \cdot \langle Z''^c, Y'^c\rangle + e^{-Z_-} \cdot \langle Z''^c, Y''^c\rangle \\ &+ \sum_{s \leq \cdot} [\Delta W_s + W_{s-}(\Delta Z'_s + i\Delta Z''_s) - e^{-Z_{s-}}(\Delta Y'_s + i\Delta Y''_s)]. \end{aligned}$$

Since Y is a solution to 4.59, we have $\Delta Y = Y_- \Delta X$ and

(7)
$$\Delta W_s = e^{-Z_{s-} - \Delta Z_s}(Y_{s-} + \Delta Y_s) - e^{-Z_{s-}} Y_{s-} = W_{s-}[e^{-\Delta X_s}(1 + \Delta X_s) - 1]$$

because $\Delta Z = \Delta X$. Moreover, 4.60 and 4.41 yield

$$\langle Z'^c, Y'^c\rangle = Y'_- \cdot \langle Z'^c, X'^c\rangle - Y''_- \cdot \langle Z'^c, X''^c\rangle = Y'_- \cdot \langle Z'^c, Z'^c\rangle - Y''_- \cdot \langle Z'^c, Z''^c\rangle.$$

Similarly, we obtain:

$$\langle Z'^c, Y''^c\rangle = Y'_- \cdot \langle Z'^c, Z''^c\rangle + Y''_- \cdot \langle Z'^c, Z'^c\rangle$$
$$\langle Z''^c, Y'^c\rangle = Y'_- \cdot \langle Z''^c, Z'^c\rangle - Y''_- \cdot \langle Z''^c, Z''^c\rangle$$
$$\langle Z''^c, Y''^c\rangle = Y'_- \cdot \langle Z''^c, Z''^c\rangle + Y'' \cdot \langle Z''^c, Z'^c\rangle$$

and thus

$$\begin{aligned}(8)\qquad &\langle Z'^c, Y'^c\rangle + i\langle Z'^c, Y''^c\rangle + i\langle Z''^c, Y'^c\rangle - \langle Z''^c, Y''^c\rangle \\ &= Y_- \cdot (\langle Z'^c, Z'^c\rangle - \langle Z''^c, Z''^c\rangle + 2i\langle Z'^c, Z''^c\rangle).\end{aligned}$$

Therefore plugging (7), (8) and (3) and $\Delta Y = Y_-\Delta X = Y_-\Delta Z$ into (6) yields

$$\begin{aligned} W &= 1 - W_- \cdot X + e^{-Z_-} \cdot Y + \sum_{s\leq \cdot} W_{s-}[e^{-\Delta X_s}(1 + \Delta X_s) - 1 + \Delta Z_s - \Delta Z_s] \\ &= 1 - W_- \cdot X + W_- \cdot X + \sum_{s\leq \cdot} W_{s-}[e^{-\Delta X_s}(1 + \Delta X_s) - 1] \\ (9)\qquad &= 1 + W_- \cdot A \end{aligned}$$

where $A_t = \sum_{s\leq t}[e^{-\Delta X_s}(1 + \Delta X_s) - 1]$ is a complex-valued process with finite variation (because $\sum_{s\leq t}|\Delta X_s|^2 < \infty$ and $|e^{-x}(1 + x) - 1| \leq C|x|^2$ for all $x \in \mathbb{C}$ such that $|x| \leq 1/2$, for some constant C).

Part (ii) implies that $Y = U$ is a solution of 4.59, and in this case $W = e^{-Z}U = V$. Hence $V = 1 + V_- \cdot A$. Thus if Y is another solution of 4.59 and $W = e^{-Z}Y$ and $\tilde{W} = W - V$, we have by (9):

$$(10)\qquad \tilde{W}_t = \int_0^t \tilde{W}_{s-}\, dA_s.$$

Let $S = \inf(t\colon \tilde{W}_t \neq 0)$. Then (10) implies that $\tilde{W}_S = 0$ on $\{S < \infty\}$. Moreover, there exists $S' \geq S$ with $\{S < \infty\} \subset \{S' > S\}$ and $\int_{(S,S']}|dA_s| \leq 1/2$. Since (10) yields for $t > S$:

$$\tilde{W}_t = \tilde{W}_S + \int_{(S,t]} \tilde{W}_{s-}\, dA_s = \int_{(S,t]} \tilde{W}_{s-}\, dA_s,$$

we deduce that

$$\sup_{t\leq S'} |\tilde{W}_t| \leq \frac{1}{2} \sup_{t\leq S'} |\tilde{W}_t|$$

by definition of S'. Therefore $\sup_{t\leq S'}|\tilde{W}_t| = 0$ and since $S' > S$ on $\{S < \infty\}$, it follows that $S = +\infty$. In other words, $\tilde{W}_t = 0$ for all t, that is $W = V$, and $Y = e^Z W = e^Z V = U$. So we have proved that $U = \mathscr{E}(X)$ is the unique solution of 4.59.

(iv) The claims (a) and (b) immediately follow from 4.59 and properties 4.34. Finally it is obvious from (1) that $V \neq 0$ on $[\![0, T[\![$, $V_- \neq 0$ on $[\![0, T]\!]$ and $V = 0$ on $[\![T, \infty[\![$. Since $\mathscr{E}(X) = Ve^Z$, (c) follows. $\square$

§ 4g. The Discrete Case

1. There is very little to say, indeed, about discrete-time semimartingales. Start with a discrete time basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}}, P)$. We have seen in § 3d that every adapted process X with $X_0 = 0$ is in $\mathscr{V}$. Hence the "natural" class of semimartingales is as such: *a process is a semimartingale if and only if it is adapted.*

For special semimartingales, the easiest way is to use characterization 4.23(iii), which reads as follows: a process X is a *special semimartingale* if and only if X_0 is $\mathscr{F}_0$-measurable, and $X - X_0$ belongs to $\mathscr{A}_{\text{loc}}$ (see § 3d*1*). Then the *canonical decomposition* $X = X_0 + M + A$ is given by:

4.65
$$\begin{cases} A_n = \sum_{1 \leq p \leq n} E(X_p - X_{p-1} | \mathscr{F}_{p-1}) \\ M_n = \sum_{1 \leq p \leq n} [X_p - X_{p-1} - E(X_p - X_{p-1} | \mathscr{F}_{p-1})]. \end{cases}$$

If X and Y are two (locally) square-integrable martingales, their angle bracket is given by

4.66
$$\langle X, Y \rangle_n = \sum_{1 \leq p \leq n} E[(X_p - X_{p-1})(Y_p - Y_{p-1}) | \mathscr{F}_{p-1}]$$

(Theorem 4.2 is then trivial).

The notion of stochastic integral is elementary, since all processes are "simple" in the sense of 4.29, and they also are with finite variation! so if H is any predictable process, $H \cdot X$ is defined by

4.67
$$H \cdot X_n = \sum_{1 \leq p \leq n} H_p(X_p - X_{p-1}) = \sum_{1 \leq p \leq n} H_p \Delta X_p$$

and all properties 4.33 to 4.38 are obvious (we can of course define $H \cdot X$ for any two processes H and X, without measurability conditions; however, property 4.34 for instance needs the predictability for H).

The quadratic variation is naturally defined by (compare to 4.47a):

4.68
$$[X, Y]_n = \sum_{1 \leq p \leq n} (X_p - X_{p-1})(Y_p - Y_{p-1}) = \sum_{1 \leq p \leq n} \Delta X_p \Delta Y_p$$

and all claims of § 4e*1* are obvious or elementary (we have $X^c = Y^c = 0$). Finally, Ito's formula 4.58 reads as

$$f(X_n) = f(X_0) + \sum_{1 \leq p \leq n} \sum_{i \leq d} D_i f(X_{p-1})(X_p^i - X_{p-1}^i)$$
$$+ \sum_{1 \leq p \leq n} \left[f(X_p) - f(X_{p-1}) - \sum_{i \leq d} D_i f(X_{p-1})(X_p^i - X_{p-1}^i) \right],$$

a trivial identity!

2. As usual, we associate the continuous-time basis $\mathscr{B}'$ to $\mathscr{B}$ by 1.55, and to each process X on $\mathscr{B}$ a process X' on $\mathscr{B}'$ by 1.59. Then:

—If X is adapted, then X' is a semimartingale. X is a special semimartingale on $\mathscr{B}$ if and only if X' is such on $\mathscr{B}'$, and in that case the canonical decompositions of X and X' correspond to each other via 1.59.

—With obvious notation, $[X, Y]' = [X', Y']$, and also $\langle X, Y\rangle' = \langle X', Y'\rangle$ when X, Y are locally square-integrable martingales, and also $(H \cdot X)' = H' \cdot X'$.

Chapter II. Characteristics of Semimartingales and Processes with Independent Increments

We continue across our project of expounding the general theory of processes. However, here we touch upon a slightly different aspect of the theory, which at the same time is much less widely known than what was in the first chapter. This is also the aspect which will be most directly useful for limit theorems.

In a sense, this whole chapter is centered around processes with independent increments, although these explicitely appear in Sections 4-6 only. Consider for example a process with stationary independent increments: as is well known, the law of this process is completely characterized by three quantities, namely the "drift", the "diffusion coefficient", the "Lévy measure"; and also the convergence of such processes is entirely determined by the convergence of these corresponding quantities, in a suitable sense.

Our aim is to generalize these notions to semimartingales. After a preliminary section on random measures, this task is performed in Section 2, and examples are provided in Section 3.

Then we start studying processes with independent increments (in short: PII). The main facts are collected in Section 4, where PII which are semimartingales are studied, and one might as well stop reading this chapter at the end of that section. Section 5 deals with PII which are not semimartingales: it is rather difficult, especially considering that at the end we will prove that the most general PII is equal to a PII-semimartingale, plus a deterministic function! Finally, Section 6 deals with "conditional PII", a kind of processes that is a very simple extension of PII's, but that is often encountered in applications (like Cox processes, etc...)

1. Random Measures

The concept of a random measure is again one of those basic notions that could very well have figured in Chapter I, although it is not so well known than semimartingales for instance. It will prove rather essential to our purpose, in the sense that it allows for a very tractable description of the jumps of a càdlàg process.

We describe here the theory of random measures with all details. Nevertheless, it should be kept in mind that the content of this section is nothing else than a straightforward extension of the notions of increasing processes and their compensators.

§ 1a. General Random Measures

1. Like in Chapter I, we start with a (continuous-time) stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}, P)$. Remember that we do not ask the σ-fields to be complete.

Let us also consider an auxiliary measurable space $(E, \mathscr{E})$ which we assume to be a *Blackwell space* (see [36, 59]). The reader who does not know about Blackwell spaces needs not worry: in all the sequel, E will actually be $\bar{\mathbb{R}}^d$ or $\mathbb{R}^d$, or at most a Polish space with its Borel σ-field. The only two properties of Blackwell spaces that will be used are the following:

1.1 First, the σ-field $\mathscr{E}$ is *separable*, i.e. it is generated by a countable algebra. □

For the second property, we need to recall that a *transition kernel* $\alpha(a, db)$ of a measurable space $(A, \mathscr{A})$ into another measurable space $(B, \mathscr{B})$ is a family $(\alpha(a, \cdot)\colon a \in A)$ of positive measures on $(B, \mathscr{B})$, such that $\alpha(\cdot, C)$ is $\mathscr{A}$-measurable for each $C \in \mathscr{B}$.

1.2 Let $(G, \mathscr{G})$ be any measurable space. If m is a positive finite measure on $(G \times E, \mathscr{G} \otimes \mathscr{E})$ with G-marginal $\hat{m}$: $\hat{m}(A) = m(A \times E)$, then there exists a transition kernel $\alpha(g, dx)$ from $(G, \mathscr{G})$ into $(E, \mathscr{E})$ such that $m(B) = \int \hat{m}(dg) \int \alpha(g, dx) 1_B(g, x)$ for all $B \in \mathscr{G} \otimes \mathscr{E}$; we shall also write $m(dg, dx) = \hat{m}(dg)\alpha(g, dx)$. Note that this *disintegration property* is also equivalent to the following property: if Z is any $(E, \mathscr{E})$-valued random variable on any probability space $(\Omega, \mathscr{F}, P)$ and if $\mathscr{F}'$ is a sub-σ-field of $\mathscr{F}$, then Z admits a *regular conditional distribution* with respect to $\mathscr{F}'$. □

We now give the definition of an random measure (we shall only consider nonnegative random measures).

1.3 **Definition.** A *random measure* on $\mathbb{R}_+ \times E$ is a family $\mu = (\mu(\omega; dt, dx)\colon \omega \in \Omega)$ of nonnegative measures on $(\mathbb{R}_+ \times E, \mathscr{R}_+ \otimes \mathscr{E})$ satisfying $\mu(\omega; \{0\} \times E) = 0$ identically. □

We put:

1.4 $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$, with the σ-fields $\tilde{\mathscr{O}} = \mathscr{O} \otimes \mathscr{E}$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$.

A function W on $\tilde{\Omega}$ that is $\tilde{\mathcal{O}}$-measurable (resp. $\tilde{\mathcal{P}}$-measurable) is called an *optional* (resp. a *predictable*) function. If W is a function on $\tilde{\Omega}$ and if H is a process, we write WH or HW for the function $(\omega, t, x) \rightsquigarrow H(\omega, t) W(\omega, t, x)$.

Let μ be a random measure and W an optional function on $\tilde{\Omega}$. Since $(t, x) \rightsquigarrow W(\omega, t, x)$ is $\mathcal{R}_+ \otimes \mathcal{E}$-measurable for each $\omega \in \Omega$, we can define the *integral process* $W * \mu$ by

1.5 $W * \mu_t(\omega)$

$$= \begin{cases} \displaystyle\int_{[0,t] \times E} W(\omega, s, x) \mu(\omega; ds, dx) & \text{if } \displaystyle\int_{[0,t] \times E} |W(\omega, s, x)| \mu(\omega; ds, dx) \text{ is finite,} \\ +\infty & \text{otherwise.} \end{cases}$$

1.6 **Definitions.** a) A random measure μ is called *optional* (resp. *predictable*) if the process $W * \mu$ is optional (resp. predictable) for every optional (resp. predictable) function W.

b) An optional measure μ is called *integrable* if the random variable $1 * \mu_\infty = \mu(\cdot, \mathbb{R}_+ \times E)$ is integrable (or equivalently, if $1 * \mu \in \mathcal{A}^+$).

c) An optional random measure μ is called *$\tilde{\mathcal{P}}$-σ-finite* if there exists a strictly positive predictable function V on $\tilde{\Omega}$ such that the random variable $V * \mu_\infty$ is integrable (or equivalently, $V * \mu \in \mathcal{A}^+$); this property is equivalent to the existence of a $\tilde{\mathcal{P}}$-measurable partition (A_n) of $\tilde{\Omega}$ such that each $(1_{A_n} * \mu)_\infty$ is integrable. □

1.7 **Example.** Let $A \in \mathcal{V}^+$; one associates to A a random measure μ on $\mathbb{R}_+ \times \{1\}$ by $\mu(\omega; dt \times \{1\}) = dA_t(\omega)$. Then

(i) μ is optional; μ is predictable if and only if A is so;

(ii) μ is integrable if and only if A is so (i.e. $A \in \mathcal{A}^+$);

(iii) if $A \in \mathcal{A}^+_{\text{loc}}$ then μ is $\tilde{\mathcal{P}}$-σ-finite (take for (T_n) a localizing sequence for A, so that $A^{T_n} \in \mathcal{A}$, and the predictable partition $A_0 = [\![0]\!] \times \{1\}$, $A_n =]\!]T_{n-1}, T_n]\!] \times \{1\}$ of $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \{1\}$).

However, it is worthwhile to notice that if μ is a random measure on $\mathbb{R}_+ \times \{1\}$ (even a $\tilde{\mathcal{P}}$-σ-finite one), it is not always possible to associate to it a process in $\mathcal{V}^+$ (or $\mathcal{A}^+_{\text{loc}}$), because $\mu(\omega; [0, t] \times \{1\})$ may be infinite for all $t > 0$. □

2. Compensator of a random measure. The main result of this section is the following generalization of Theorem I.3.18:

1.8 **Theorem.** *Let μ be an optional $\tilde{\mathcal{P}}$-σ-finite random measure. There exists a random measure, called the compensator of μ and denoted by μ^p, which is unique up to a P-null set, and which is characterized as being a predictable random measure satisfying either one of the two following equivalent properties:*

(i) *$E(W * \mu^p_\infty) = E(W * \mu_\infty)$ for every nonnegative $\tilde{\mathcal{P}}$-measurable function W on $\tilde{\Omega}$.*

(ii) *For every* $\tilde{\mathscr{P}}$*-measurable function* W *on* $\tilde{\Omega}$ *such that* $|W| * \mu \in \mathscr{A}^+_{\mathrm{loc}}$, *then* $|W| * \mu^p$ *belongs to* $\mathscr{A}^+_{\mathrm{loc}}$, *and* $W * \mu^p$ *is the compensator of the process* $W * \mu$ *(or equivalently,* $W * \mu - W * \mu^p$ *is a local martingale).*

Moreover, there exists a predictable $A \in \mathscr{A}^+$ *and a kernel* $K(\omega, t; dx)$ *from* $(\Omega \times \mathbb{R}_+, \mathscr{P})$ *into* $(E, \mathscr{E})$ *such that*

1.9 $$\mu^p(\omega; dt, dx) = dA_t(\omega)K(\omega, t; dx).$$

Sometimes μ^p is also called *predictable compensator*, or *dual predictable projection*, of μ. Of course, the disintegration 1.9 is by no means unique.

Proof. Let V be a strictly positive predictable function on $\tilde{\Omega}$ such that $V * \mu \in \mathscr{A}^+$. Note that either one of (i) and (ii) implies $V * \mu^p \in \mathscr{A}^+$.

a) We prove the implication (i) $\Rightarrow$ (ii). Let W be a predictable function with $|W| * \mu \in \mathscr{A}^+_{\mathrm{loc}}$, and (T_n) a localizing sequence with $(|W| * \mu)^{T_n} \in \mathscr{A}^+$. Appling (i) to each $|W| 1_{[\![0, T_n]\!]}$, we see that $|W| * \mu^p \in \mathscr{A}_{\mathrm{loc}}$. If T is a stopping time, applying (i) to $W^+ 1_{[\![0, T \wedge T_n]\!]}$ and to $W^- 1_{[\![0, T \wedge T_n]\!]}$ yields that $E(W * \mu^p_{T \wedge T_n}) = (W * \mu_{T \wedge T_n})$. Then I.1.44 implies that each $(W * \mu - W * \mu^p)^{T_n}$ belongs to $\mathscr{M}$, hence $W * \mu - W * \mu^p$ is a local martingale.

(b) We prove the implication (ii) $\Rightarrow$ (i). If $0 \leq W \leq nV$ and if W is predictable, one has $W * \mu \in \mathscr{A}^+$ and (ii) together with I.3.17 imply that $E(W * \mu^p_\infty) = E(W * \mu_\infty)$. If W is any predictable nonnegative function we apply what precedes to each $W(n) = W 1_{\{W \leq nV\}}$ and then let $n \uparrow \infty$ to obtain (i).

(c) We prove the uniqueness. Let $\mathscr{E}_0$ be a countable algebra that generates $\mathscr{E}$, and let μ^p and $\hat{\mu}^p$ satisfy (ii). Then for each $A \in \mathscr{E}_0$ the two processes $(V 1_A) * \mu^p$ and $(V 1_A) * \hat{\mu}^p$ are indistinguishable. Therefore the set

$$N = \bigcup_{A \in \mathscr{E}_0} \{\omega \colon \exists t \text{ with } (V 1_A) * \mu^p_t(\omega) \neq (V 1_A) * \hat{\mu}^p_t(\omega)\}$$

is P-null, while on its complement N^c we have $\mu^p(\omega; \cdot) = \hat{\mu}^p(\omega; \cdot)$.

(d) Finally we prove the existence and 1.9. Put $A = (V * \mu)^p$. For each bounded $\tilde{\mathscr{P}}$-measurable function W, set $m(W) = E[(VW) * \mu_\infty]$, which defines a positive finite measure m on $(\tilde{\Omega}, \tilde{\mathscr{P}})$. We also consider the positive finite measure $\hat{m}$ on $(\Omega \times \mathbb{R}_+, \mathscr{P})$ defined by $\hat{m}(d\omega, dt) = P(d\omega) dA_t(\omega)$, that is $\hat{m}(B) = E(1_B \cdot A_\infty)$ for all $B \in \mathscr{P}$. If $B \in \mathscr{P}$ we have $(V 1_B) * \mu = 1_B \cdot (V * \mu)$ (trivial), hence $[(V 1_B) * \mu]^p = 1_B \cdot A$ by I.3.18 hence $\hat{m}(B) = E[(V 1_B) * \mu_\infty] = m(B \times E)$.

Therefore we can apply property 1.2, thus obtaining a transition kernel α from $(\Omega \times \mathbb{R}_+, \mathscr{P})$ into $(E, \mathscr{E})$, with $m(d\omega, dt, dx) = \hat{m}(d\omega, dt)\alpha(\omega, t; dx)$. We put $K(\omega, t; dx) = \alpha(\omega, t; dx) \dfrac{1}{V(\omega, t, x)}$, that is K is the transition kernel from $(\Omega \times \mathbb{R}_+, \mathscr{P})$ into $(E, \mathscr{E})$ defined by $K(\omega, t; B) = \int_B \alpha(\omega, t; dx) \dfrac{1}{V(\omega, t, x)}$ for all $B \in \mathscr{E}$. Then if W is a nonnegative $\tilde{\mathscr{P}}$-measurable function, the process $(KW)_t(\omega) = \int_E K(\omega, t; dx) W(\omega, t, x)$ is obviously predictable.

Now let us define the random measure μ^p by 1.9. If W is nonnegative and $\tilde{\mathscr{P}}$-measurable, $W * \mu^p = (KW) \cdot A$ is predictable, hence μ^p is a predictable random measure. Moreover, again for W nonnegative $\tilde{\mathscr{P}}$-measurable,

$$E(W * \mu_\infty^p) = \int m(d\omega, dt, dx) \frac{W(\omega, t, x)}{V(\omega, t, x)} = E(W * \mu_\infty),$$

and we have (i). □

Here are some easy properties:

1.10 If μ is a predictable $\tilde{\mathscr{P}}$-σ-finite measure, then $\mu^p = \mu$.

1.11 Let μ be an optional $\tilde{\mathscr{P}}$-σ-finite measure, and let W be a nonnegative predictable function on $\tilde{\Omega}$. Then for each predictable time T,

$$\int_E \mu^p(\{T\} \times dx) W(T, x) = E\left(\left(\int_E \mu(\{T\} \times dx) W(T, x) | \mathscr{F}_{T-}\right) \quad \text{on } \{T < \infty\}$$

($W(T, x)$ stands for $W(\omega, T(\omega), x)$; this property comes from I.3.21 applied to $A = (W 1_{[\![T]\!]}) * \mu$, for which $A^p = (W 1_{[\![T]\!]}) * \mu^p$). □

1.12 **Example.** Let us come back to Example 1.7, assuming that $A \in \mathscr{A}_{\mathrm{loc}}^+$. Then the dual predictable projection μ^p of μ is $\mu^p(\omega; dt \times dx) = dA_t^p(\omega) \otimes \varepsilon_1(dx)$, where A^p is the dual predictable projection of A. □

§ 1b. Integer-Valued Random Measures

1.13 **Definition.** An *integer-valued random measure* is a random measure that satisfies:

(i) $\mu(\omega; \{t\} \times E) \leq 1$ identically;

(ii) for each $A \in \mathscr{R}_+ \otimes \mathscr{E}$, $\mu(\cdot, A)$ takes its values in $\bar{\mathbb{N}}$;

(iii) μ is optional and $\tilde{\mathscr{P}}$-σ-finite. □

Of course, (iii) above is just an ad-hoc property, which we add to the definition in order to avoid repeating it in all the following statements.

1.14 **Proposition.** *If μ is an integer-valued random measure, there exists a thin random set D* (recall I.1.30 for the definition of a thin set) *and an E-valued optional process β such that*

$$\mu(\omega; dt, dx) = \sum_{s \geq 0} 1_D(\omega, s) \varepsilon_{(s, \beta_s(\omega))}(dt, dx),$$

where ε_a denotes the Dirac measure at point a. Note that for any nonnegative optional function W, if (T_n) is a sequence of stopping times that exhausts the thin set D, we have

1.15
$$W * \mu_t = \begin{cases} \sum_{(n)} W(T_n, \beta_{T_n}) 1_{\{T_n \le t\}} \\ \sum_{0<s\le t} W(s, \beta_s) 1_D(s). \end{cases}$$

Note also that conversely, if μ is defined as in 1.14, where D is a thin set and β is an E-valued optional process, then μ is an integer-valued random measure, *provided it is* $\tilde{\mathscr{P}}$-σ-*finite* (the optionality of μ follows from 1.15, I.1.21 and I.1.23).

Proof of 1.14. If we set $D = \{(\omega, t): \mu(\omega; \{t\} \times E) = 1\}$, the definition 1.13 implies that μ has the claimed form for some E-valued process β. The only things to prove are that β is optional, and that D is thin.

Let (A_n) be a $\tilde{\mathscr{P}}$-measurable partition of $\tilde{\Omega}$ such that $1_{A_n} * \mu \in \mathscr{A}^+$ and $(S(n, p))_{p\in\mathbb{N}}$ be the successive jump times of the process $1_{A_n} * \mu$, which by 1.13 is a point process in the sense of I.3.25. So $D = \bigcup_{(n,p)} [\![S(n, p)]\!]$ is a thin set.

Let (T_n) be a sequence of stopping times that exhausts D. For each $t \in \mathbb{R}_+$ and each $C \in \mathscr{E}$, the variable $1_C(\beta_{T_n}) 1_{\{T_n \le t\}} = (1_{[\![T_n]\!] \times C}) * \mu_t$ must be $\mathscr{F}_t$-measurable, because μ is optional. We deduce that β_{T_n} is $\mathscr{F}_{T_n}$-measurable, and since we may choose β as to be equal to an arbitrary fixed point, say a, outside D, we thus obtain an optional process β. □

We also remark that an integer-valued random measure may be considered as the "counting measure" associated to a random point process in $\mathbb{R}_+ \times E$, whose points are the (T_n, β_{T_n}) appearing in 1.15; but this point process is rather particular, because two points can never occur at the same time.

The most useful example of integer-valued measure is the following:

1.16 **Proposition.** *Let X be an adapted càdlàg $\mathbb{R}^d$-valued process. Then*

$$\mu^X(\omega; dt, dx) = \sum_s 1_{\{\Delta X_s(\omega) \neq 0\}} \varepsilon_{(s, \Delta X_s(\omega))}(dt, dx)$$

defines an integer-valued random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ (and in the representation 1.14 we have $D = \{\Delta X \neq 0\}$ and $\beta = \Delta X$).

Proof. Knowing Proposition 1.14 and the comments that follow this proposition, we see that the only thing which remains to be proved is that μ^X is $\tilde{\mathscr{P}}$-σ-finite. Define the stopping times $S(n, p)$ by $S(n, 0) = 0$ and

$$S(n, p + 1) = \inf(t > S(n, p): |X_t - X_{S(n,p)}| > 2^{-n-1}),$$

and put $A(n, p) = [\![0, S(n, p)]\!] \times \{x \in \mathbb{R}^d: |x| > 2^{-n}\}$ and $A(0) = \Omega \times \mathbb{R}_+ \times \{0\}$, and $V = 1_{A(0)} + \sum_{n,p\in\mathbb{N}^*} 1_{A(n,p)} 2^{-n-p}$. Then V is $\tilde{\mathscr{P}}$-measurable and strictly positive (because $S(n, p) \uparrow \infty$ as $p \uparrow \infty$). Moreover $1_{A(0)} * \mu^X_\infty = 0$ and $1_{A(n,p)} * \mu^X_\infty \le p$ by construction of μ^X, and because any jump of X of size $> 2^{-n}$ occurs at one of the times $S(n, p)$. Hence $V * \mu^X_\infty \le \sum_{n,p\in\mathbb{N}^*} p 2^{-n-p}$ and we have the result. □

1.17 Proposition. *Let μ be an integer-valued random measure, $\nu = \mu^p$ its compensator, and $J = \{(\omega, t): \nu(\omega; \{t\} \times E) > 0\}$.*

a) *J is the predictable support of the set D showing in 1.14, and for all predictable times T and nonnegative predictable W:*

1.18 $$\int_E W(T,x)\nu(\{T\} \times dx) = E[W(T,\beta_T)1_D(T)|\mathscr{F}_{T-}] \quad \text{on } \{T < \infty\}.$$

b) *There is a version of ν such that $\nu(\omega, \{t\} \times E) \leq 1$ identically and that the thin set J is exhausted by a sequence of predictable times.*

Proof. a) 1.18 is just 1.11. In particular, $a_t = \nu(\{t\} \times E)$ is the predictable projection of the process 1_D, so the first claim follows from the definition of the predictable support of D.

b) I.2.23 yields a sequence (T_n) of predictable times, whose graphs are pairwise disjoint, with $J' \subset J$ and $J \setminus J'$ evanescent, if $J' = \bigcup [\![T_n]\!]$. Moreover $a_{T_n} \leq 1$ a.s. by (a), while a is predictable. Hence if $A_n = \{a_{T_n} \leq 1\}$ and $T'_n = (T_n)_{A_n}$ and $J'' = \bigcup [\![T'_n]\!]$, then each T'_n is predictable and $J'' \subset J$ and $J \setminus J''$ is evanescent: thus the measure $\nu''(\omega; dt \times dx) = \nu(\omega, dt \times dx)1_{(J\setminus J'')^c}(\omega, t)$ is a.s. equal to ν and is predictable. Therefore it is a version of μ^p, that meets the requirements of (b). □

Then, if we use I.2.35 we obtain the:

1.19 Corollary. *Let X be an adapted càdlàg process and μ^X be the measure associated to its jumps by 1.16. Then X is quasi-left-continuous if and only if there exists a version of $(\mu^X)^p$ that satisfies identically $(\mu^X)^p(\omega; \{t\} \times E) = 0$.*

§ 1c. A Fundamental Example: Poisson Measures

By extension of Definition I.3.26, we put

1.20 Definitions. a) An *extended Poisson measure* on $\mathbb{R}_+ \times E$, *relative to the filtration* $\mathbf{F}$, is an integer-valued random measure μ such that

(i) the positive measure m on $\mathbb{R}_+ \times E$ defined by $m(A) = E[\mu(A)]$ is σ-finite;

(ii) for every $s \in \mathbb{R}_+$ and every $A \in \mathscr{R}_+ \otimes \mathscr{E}$ such that $A \subset (s, \infty) \times E$ and that $m(A) < \infty$, the variable $\mu(\cdot, A)$ is independent of the σ-field $\mathscr{F}_s$.

b) The measure m is called the *intensity measure* of μ.

c) If m satisfies $m(\{t\} \times E) = 0$ for each $t \in \mathbb{R}_+$, then μ is called a *Poisson measure*; if m has the form $m(dt, dx) = dt \times F(dx)$, where F is a positive σ-finite measure on $(E, \mathscr{E})$, then μ is called a *homogeneous Poisson measure.* □

For instance, let N be a point process, and let μ be the measure associated to it like in 1.7 (with $E = \{1\}$); then μ is an extended Poisson (resp. Poisson, resp. homogeneous Poisson) measure if and only if N is an extended Poisson (resp. Poisson, resp. standard Poisson) process.

We shall prove later that a Poisson measure is the counting measure of an "ordinary" Poisson point process on $\mathbb{R}_+ \times E$: that is, if (A_i) is a sequence of pairwise disjoint measurable subsets of $\mathbb{R}_+ \times E$ with $m(A_i) < \infty$ the random variables $(\mu(A_i))$ are independent, and $\mu(A_i)$ has a Poisson distribution with mean $m(A_i)$. At this stage, we content ourselves by computing the compensator of a Poisson measure.

1.21 **Proposition.** *Let μ be an extended Poisson measure on $\mathbb{R}_+ \times E$, relative to the filtration* $\mathbf{F}$, *with intensity measure m. Then its compensator is $\mu^p(\omega; \cdot) = m(\cdot)$.*

(We shall prove later a converse to that statement: see 4.8).

Proof. We define $\mu^p(\omega, \cdot) = m(\cdot)$, which is a "random" measure that is predictable (since it is deterministic). We need to prove 1.7(i). By a monotone class argument, using I.2.2, we see that it suffices to prove 1.7(i) for all W's of the form $W = 1_A 1_{B \times (s,t] \times C}$, where $0 \leq s < t$, $B \in \mathscr{F}_s$, $C \in \mathscr{E}$ and $A \in \mathscr{R}_+ \otimes \mathscr{E}$ satisfies $m(A) < \infty$. By hypothesis, the variables 1_B and $\mu(\cdot, A \cap ((s,t] \times C))$ are independent and integrable, thus

$$E(W * \mu_\infty) = W(1_B \mu(A \cap ((s,t] \times C))) = P(B) m(A \cap ((s,t] \times C)) = E(W * \mu_\infty^p).$$

□

§ 1d. Stochastic Integral with Respect to a Random Measure

This subsection presents the construction of stochastic integrals with respect to a "compensated" integer-valued random measure. At this juncture, we should observe that in Chapter I we have constructed integrals with respect to a semimartingale X, but only for (locally) bounded integrands; the "most general" possible integrands have been reached only when $X \in \mathscr{H}^2_{\text{loc}}$ (and only for integrands such that the integral belong to $\mathscr{H}^2_{\text{loc}}$ itself). But of course there also exists a stochastic integral for "reasonable" non-locally bounded integrands: see [98, 180].

Here we give at once the most general integral, partly because the notion of "locally bounded" has no meaning here.

We start with an integer-valued random measure μ on $\mathbb{R}_+ \times E$, where $(E, \mathscr{E})$ is a "Blackwell space": see § 1a; for all practical purposes, $E = \mathbb{R}^d$. By Proposition 1.14, μ is of the form

1.22
$$\mu(\omega; dt, dx) = \sum_{s \geq 0} 1_D(\omega, s) \varepsilon_{(s, \beta_s(\omega))}(dt, dx)$$

where D is an optional thin set with $(\omega, 0) \notin D$ and β is an E-valued optional process.

We call ν a "good" version of the dual predictable projection of μ, as constructed in 1.17, and we set:

1.23 $$\begin{cases} a_t(\omega) = \nu(\omega;\{t\} \times E) \\ J = \{a > 0\}, \text{ exhausted by the sequence } (T_n) \text{ of predictable times,} \\ \nu^c(\omega; dt, dx) = \nu(\omega; dt, dx) 1_{J^c}(\omega, t). \end{cases}$$

Recall that $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$. With any measurable function W on $\tilde{\Omega}$ we associate the process

1.24 $$\hat{W}_t(\omega) = \begin{cases} \int_E W(\omega,t,x)\nu(\omega;\{t\} \times dx) & \text{if } \int_E |W(\omega,t,x)|\nu(\omega;\{t\} \times dx) < \infty \\ +\infty & \text{otherwise.} \end{cases}$$

1.25 **Lemma.** *If W is $\tilde{\mathscr{P}}$-measurable, then $\hat{W}$ is predictable, and it is a version of the predictable projection of the process $(\omega, t) \rightsquigarrow W(\omega, t, \beta_t(\omega)) 1_D(\omega, t)$.*

In particular, for all predictable times T,

1.26 $$\hat{W}_T = E[W(T, \beta_T) 1_D(T) | \mathscr{F}_{T-}] \quad \text{on } \{T < \infty\}.$$

Proof. Due to 1.24 and to the definition of extended conditional expectation I.1.1 and extended predictable projection I.2.28, it suffices to consider separately W^+ and W^-, or in other words to assume $W \geq 0$. Then $\hat{W}_{T_n} = Y^n_{T_n}$ where $Y^n = (1_{[\![T_n]\!]} W) * \nu$ is predictable. Hence $\hat{W}_{T_n}$ is $\mathscr{F}_{T_n-}$-measurable, and $\hat{W} = \sum_n \hat{W}_{T_n} 1_{[\![T_n]\!]}$ is predictable by I.2.12. Moreover 1.26 coincides with 1.11 and it implies our second claim. □

1.27 **Definitions.** a) We denote by $G_{\mathrm{loc}}(\mu)$ the set of all $\tilde{\mathscr{P}}$-measurable real-valued functions W on $\tilde{\Omega}$ such that the process $\tilde{W}_t(\omega) = W(\omega, t, \beta_t(\omega)) 1_D(\omega, t) - \hat{W}_t(\omega)$ satisfies $[\sum_{s \leq \cdot} (\tilde{W}_s)^2]^{1/2} \in \mathscr{A}^+_{\mathrm{loc}}$.

b) If $W \in G_{\mathrm{loc}}(\mu)$ we call *stochastic integral of W with respect to $\mu - \nu$ and we denote by $W * (\mu - \nu)$* any purely discontinuous local martingale such that ΔX and $\tilde{W}$ are indistinguishable. □

To justify this definition, we observe that if $W \in G_{\mathrm{loc}}(\mu)$, then 1.25 yields ${}^p(\tilde{W}) = 0$, so by I.4.56 there is a local martingale M such that ΔM and $\tilde{W}$ are indistinguishable; then the purely discontinuous part $X = M^d$ (see I.4.18) is a version of $W * (\mu - \nu)$, and by I.4.19 any other version is indistinguishable from X.

It is clear that $G_{\mathrm{loc}}(\mu)$ is a linear space, and $W \rightsquigarrow W * (\mu - \nu)$ is linear (up to indistinguishability) on $G_{\mathrm{loc}}(\mu)$.

The qualifier "stochastic integral with respect to $\mu - \nu$" for the process $W * (\mu - \nu)$ is substantiated by the following:

1.28 **Proposition.** *Let W be a predictable function on $\tilde{\Omega}$, such that $|W| * \mu \in \mathscr{A}^+_{\mathrm{loc}}$* (or equivalently: $|W| * \nu \in \mathscr{A}^+_{\mathrm{loc}}$). *Then $W \in G_{\mathrm{loc}}(\mu)$ and*

1.29 $$W * (\mu - \nu) = W * \mu - W * \nu.$$

Proof. We know that $W * \mu - W * \nu$ is in $\mathscr{L} \cap \mathscr{V}$, hence it is a purely discontinuous local martingale by I.4.14, and a simple computation shows that $\Delta X = \widetilde{W}$. □

1.30 **Proposition.** a) *Let T be a stopping time and $W \in G_{\mathrm{loc}}(\mu)$. Then $W 1_{[\![0,T]\!]}$ belongs to $G_{\mathrm{loc}}(\mu)$ and $(W 1_{[\![0,T]\!]}) * (\mu - \nu) = \{W * (\mu - \nu)\}^T$.*

b) *Let H be a locally bounded predictable process and $W \in G_{\mathrm{loc}}(\mu)$. Then HW belongs to $G_{\mathrm{loc}}(\mu)$ and $(HW) * (\mu - \nu) = H \cdot \{W * (\mu - \nu)\}$.*

Proof. (a) is a particular case of (b): take $H = 1_{[\![0,T]\!]}$. For (b), we observe that $W' = HW$ is $\tilde{\mathscr{P}}$-measurable and $\hat{W}' = H\hat{W}$ and $\widetilde{W}' = H\widetilde{W}$: hence $W' \in G_{\mathrm{loc}}(\mu)$. Therefore $(HW) * (\mu - \nu)$ and $H \cdot [W * (\mu - \nu)]$ are two purely discontinuous local martingales (use I.4.55d for the latter) with the same jumps. □

For further reference, we characterize the property $W \in G_{\mathrm{loc}}(\mu)$ by the integrability of a suitable increasing *predictable* process. With any predictable function W on $\tilde{\Omega}$ we associate two increasing (possibly infinite) predictable processes as such:

1.31 $$C(W)_t = (W - \hat{W})^2 * \nu_t + \sum_{s \le t} (1 - a_s)(\hat{W}_s)^2$$

1.32 $$\bar{C}(W)_t = |W - \hat{W}| * \nu_t + \sum_{s \le t} (1 - a_s)|\hat{W}_s|.$$

1.33 **Theorem.** *Let W be a predictable function on $\tilde{\Omega}$.*

a) *W belongs to $G_{\mathrm{loc}}(\mu)$ and $W * (\mu - \nu)$ belongs to $\mathscr{H}^2$ (resp. $\mathscr{H}^2_{\mathrm{loc}}$) if and only if $C(W)$ belongs to $\mathscr{A}^+$ (resp. $\mathscr{A}^+_{\mathrm{loc}}$), in which case*

1.34 $$\langle W * (\mu - \nu), W * (\mu - \nu) \rangle = C(W).$$

b) *W belongs to $G_{\mathrm{loc}}(\mu)$ and $W * (\mu - \nu)$ belongs to $\mathscr{A}$ (resp. $\mathscr{A}_{\mathrm{loc}}$) if and only if $\bar{C}(W)$ belongs to $\mathscr{A}^+$ (resp. $\mathscr{A}^+_{\mathrm{loc}}$).*

c) *W belongs to $G_{\mathrm{loc}}(\mu)$ if and only if $C(W') + \bar{C}(W'')$ belongs to $\mathscr{A}_{\mathrm{loc}}$, where*

1.35 $$\begin{cases} W' = (W - \hat{W}) 1_{\{|W - \hat{W}| \le 1\}} + \hat{W} 1_{\{|\hat{W}| \le 1\}} \\ W'' = (W - \hat{W}) 1_{\{|W - \hat{W}| > 1\}} + \hat{W} 1_{\{|\hat{W}| > 1\}}. \end{cases}$$

d) *Assume in addition that $\widetilde{W} \ge -1$ identically. Then $\hat{W} \le 1$ on $\{a < 1\}$ up to an evanescent set, and W belongs to $G_{\mathrm{loc}}(\mu)$ if and only if the increasing predictable process $C'(W)$ below belongs to $\mathscr{A}^+_{\mathrm{loc}}$:*

1.36 $$C'(W)_t = (1 - \sqrt{1 + W - \hat{W}})^2 * \nu_t + \sum_{s \le t} (1 - a_s)(1 - \sqrt{1 - \hat{W}_s})^2.$$

Proof. a) By definition of $\widetilde{W}$, the increasing process $A_t = \sum_{s \le t} (\widetilde{W}_s)^2$ is:

$$A = (W - \hat{W})^2 * \mu + \sum_n 1_{D^c}(T_n)(\hat{W}_{T_n})^2 1_{[\![T_n, \infty[\![}$$

(recall 1.23). Since $1 - a$ is the predictable projection of 1_{D^c} (see 1.25) 1.8 yields for all stopping times S:

$$E(A_S) = E[(W - \hat{W})^2 * \mu_S] + \sum_n E[(\hat{W}_{T_n})^2 E(1_{D^c}(T_n)|\mathscr{F}_{T_n-})1_{\{T_n \le S\}}]$$

1.37 $$= E[(W - \hat{W})^2 * \nu_S] + \sum_n E[(\hat{W}_{T_n})^2(1 - a_{T_n})1_{\{T_n \le S\}}] = E[C(W)_S]$$

So $A \in \mathscr{A}^+$ (resp. $\mathscr{A}^+_{\text{loc}}$) if and only if $C(W) \in \mathscr{A}^+$ (resp. $\mathscr{A}^+_{\text{loc}}$), and the necessary and sufficient condition follows from I.4.56a. Moreover if $C(W) \in \mathscr{A}^+_{\text{loc}}$, $X = W * (\mu - \nu)$ satisfies $[X, X] = A$, and 1.37 implies that the compensator of A is $C(W)$: hence 1.34 follows from I.4.50.

b) The increasing process $\bar{A}_t = \sum_{s \le t} |\tilde{W}_s|$ is

$$\bar{A} = |W - \hat{W}| * \mu + \sum_n 1_{D^c}(T_n)|\hat{W}_{T_n}| 1_{[\![T_n, \infty[\![}.$$

So, exactly as in (a), we obtain for all stopping times S:

1.38 $$E(\bar{A}_S) = E(\bar{C}(W)_S).$$

So $\bar{A} \in \mathscr{A}^+$ (resp. $\mathscr{A}^+_{\text{loc}}$) if and only if $\bar{C}(W) \in \mathscr{A}^+$ (resp. $\mathscr{A}^+_{\text{loc}}$), and the result follows from I.4.56b.

c) Since $W = W' + W''$, the sufficient part follows from (a) and (b). Conversely, assume $W \in G_{\text{loc}}(\mu)$. Set $M = W * (\mu - \nu)$ and $A = \sum_{s \le \cdot} \Delta M_s 1_{\{|\Delta M_s| > 1\}}$. Then $A \in \mathscr{V}$, and I.4.23 yields $A \in \mathscr{A}_{\text{loc}}$. A simple computation shows

$$A = W'' * \mu - \sum_n \hat{W}_{T_n} 1_{\{|\hat{W}_{T_n}| > 1\}} 1_{[\![T_n, \infty[\![}.$$

In view of 1.25, the compensator A^p has $\Delta A^p = \hat{W}'' - \hat{W} 1_{\{|\hat{W}| > 1\}}$, hence $M'' = A - A^p$ meets $\Delta M'' = \tilde{W}''$. So I.4.56 yields $W'' \in G_{\text{loc}}(\mu)$ and, since $M'' \in \mathscr{A}_{\text{loc}}$, (b) yields $\bar{C}(W'') \in \mathscr{A}^+_{\text{loc}}$. Moreover $W' = W - W''$ also belongs to $G_{\text{loc}}(\mu)$, and since obviously $|\tilde{W}'| \le 4$ (because $|W'| \le 2$), we have $W' * (\mu - \nu) \in \mathscr{H}^+_{\text{loc}}$. Hence (a) yields $C(W') \in \mathscr{A}^+_{\text{loc}}$.

d) By 1.27, W belongs to $G_{\text{loc}}(\mu)$ if and only if $(\sum_{s \le \cdot} (\tilde{W}_s)^2)^{1/2}$ belongs to $\mathscr{A}^+_{\text{loc}}$, which is clearly equivalent to $A' \in \mathscr{A}^+_{\text{loc}}$, where

1.39 $$A' = \sum_{s \le \cdot} (\tilde{W}_s)^2 1_{\{|\tilde{W}_s| \le 1\}} + \sum_{s \le \cdot} |\tilde{W}_s| 1_{\{|\tilde{W}_s| > 1\}}.$$

Since $\tilde{W} \ge -1$ we have $\hat{W}_{T_n} 1_{D^c}(T_n) \le 1_{D^c}(T_n)$ on $\{T_n < \infty\}$. Taking the predictable projections gives $\hat{W}_{T_n}(1 - a_{T_n}) \le 1 - a_{T_n}$ on $\{T_n < \infty\}$. Hence $\hat{W} \le 1$ on $\{a < 1\}$ up to an evanescent set. Moreover, if $x \ge -1$.

$$(1 - \sqrt{1 + x})^2 \le x^2 1_{\{|x| \le 1\}} + |x| 1_{\{|x| > 1\}} \le \frac{1}{(1 - \sqrt{2})^2}(1 - \sqrt{1 + x})^2.$$

Hence if $\tilde{W} \ge -1$ identically, we have $A'' \le A' \le A''/(1 - \sqrt{2})^2$, where

$$A'' = (1 - \sqrt{1 + W - \hat{W}})^2 * \mu + \sum_n 1_{D^c}(T_n)(1 - \sqrt{1 - \hat{W}_{T_n}})^2 1_{[\![T_n, \infty[\![},$$

and A' is defined in 1.39. Then $A'' \in \mathscr{A}^+_{\text{loc}} \Leftrightarrow A' \in \mathscr{A}^+_{\text{loc}} \Leftrightarrow W \in G_{\text{loc}}(\mu)$. But exactly as in (a), $E(A''_S) = E(C'(W)_S)$ for all stopping times S, so that $A'' \in \mathscr{A}^+_{\text{loc}} \Leftrightarrow C'(W) \in \mathscr{A}^+_{\text{loc}}$, and the result follows. $\square$

2. Characteristics of Semimartingales

The notion of "characteristics" of a semimartingale is designed to replace (or, rather, to extend) the three terms: drift, variance of the Gaussian part, Lévy measure, that characterize the distribution of a process with independent increments.

Although we shall devote Sections 4 and 5 to studying processes with independent increments, perhaps we may slightly anticipate on this subject, in order to provide some insight for the notion of characteristics.

Let then X be a process (say, real-valued) with independent increments, with $X_0 = 0$, and without fixed times of discontinuity. It is well known that X_t has a distribution that is infinitely divisible, and its characteristic function is of the form $E(\exp iuX_t) = \exp \psi_t(u)$, with

2.1 $$\psi_t(u) = iub_t - \frac{u^2}{2}c_t + \int (e^{iux} - 1 - iuh(x))F_t(dx)$$

(Lévy-Khintchine formula), where $b_t \in \mathbb{R}$, $c_t \in \mathbb{R}_+$, F_t is a positive measure which integrates $1 \wedge x^2$, and h is any bounded Borel function with compact support which "behaves like x" near the origin. Moreover, the property of independent increments immediately yields:

2.2 $$\exp(iuX_t)/\exp \psi_t(u) \text{ is a martingale.}$$

Then, if X is a semimartingale, the idea is the following: find two processes (B_t) and (C_t) and a random measure ν such that if we define the *process* $\psi_t(u)$ by 2.1, with b_t (resp. c_t, resp. $F_t(dx)$) replaced by B_t (resp. C_t, resp. $\nu([0,t] \times dx)$), then 2.2 holds. Of course we cannot find a triplet (B, C, ν) that is deterministic (unless X is a semimartingale with independent increments!) but one can find one triplet, and *only one*, that satisfies the required property and that is *predictable*. This triplet is called the *characteristics* of X.

§ 2a. Definition of the Characteristics

In this subsection we consider a d-dimensional semimartingale $X = (X^1, \ldots, X^d)$ defined on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$: we write $X \in \mathscr{S}^d$.

In 2.1 we have seen a (partially) arbitrary function h appear. Let us formalize this fact

2.3 **Definition.** We call $\mathscr{C}_t^d$ (for truncation function) the class of all functions $h : \mathbb{R}^d \to \mathbb{R}^d$ which are bounded and satisfy $h(x) = x$ in a neighbourhood of 0. □

Let $h \in \mathscr{C}_t^d$. Then $\Delta X_s - h(\Delta X_s) \neq 0$ only if $|\Delta X_s| > b$ for some $b > 0$ and the following formulae:

2.4
$$\begin{cases} \check{X}(h)_t = \sum_{s \le t} [\Delta X_s - h(\Delta X_s)] \\ X(h) = X - \check{X}(h) \end{cases}$$

define a d-dimensional process $\check{X}(h)$ in $\mathscr{V}^d$ (i.e. its components are in $\mathscr{V}$) and a d-dimensional semimartingale $X(h)$. Moreover $\Delta X(h) = h(\Delta X)$, which is bounded, hence by I.4.24 $X(h)$ is a special semimartingale (i.e. its components are in $\mathscr{S}_p$) and we consider its canonical decomposition:

2.5 $\quad X(h) = X_0 + M(h) + B(h), \quad M(h) \in \mathscr{L}^d, \quad B(h)$ predictable in $\mathscr{V}^d$.

2.6 **Definition.** Let $h \in \mathscr{C}_t^d$ be fixed. We call *characteristics of* X (or: characteristics *associated with* h, if there may be an ambiguity on h) the triplet (B, C, ν) consisting in:

(i) $B = (B^i)_{i \le d}$ is a predictable process in $\mathscr{V}^d$, namely the process $B = B(h)$ appearing in 2.5.

(ii) $C = (C^{ij})_{i,j \le d}$ is a continuous process in $\mathscr{V}^{d \times d}$, namely

$$C^{ij} = \langle X^{i,c}, X^{j,c} \rangle$$

(X^c is the continuous martingale part of X: see I.4.18).

(iii) ν is a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, namely the compensator of the random measure μ^X associated to the jumps of X by 1.16. □

2.7 **Remark.** We see that C and ν do not depend on the choice of the function h, while $B = B(h)$ does. This corresponds to the following well-known fact: in formula 2.1, if we replace h by another function, then b_t is modified, but neither c_t nor F_t are.

In the sequel, the function h will be arbitrary in $\mathscr{C}_t^d$, but fixed (unless otherwise stated). Originally, h was taken to be $h(x) = x 1_{\{|x| \le 1\}}$, which is sort of canonical (see [106]). But in this book we are interested in limit theorems, so for technical reasons it is wiser to choose h to be continuous. □

2.8 **Remark.** From Definition 2.6, the characteristics are unique up to a P-null set (because the decomposition 2.5, as well as X^c and the bracket $\langle X^{i,c}, X^{j,c} \rangle$ and the compensator of μ^X themselves are unique up to a null set only).

So it is sometimes convenient to also call *characteristics* any triplet (B', C', ν') consisting in an $\mathbb{R}^d$-valued process B', an $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process C', and a random measure ν', such that

$$(B'_{\cdot}(\omega), C'_{\cdot}(\omega), \nu'(\omega; \cdot)) = (B_{\cdot}(\omega), C_{\cdot}(\omega), \nu(\omega; \cdot)) \quad \text{for all } \omega \notin N,$$

where N is a P-null set.

This condition does not imply that (B', C', ν') are predictable (unless the stochastic basis is complete), and does not imply either that, for example, B' has finite variation, or is càdlàg everywhere. Such a (mild!) extension of the notion of characteristics will prove useful in the next chapter, for martingale problems.

□

In all the sequel, we use the notation $*$ introduced in 1.4. We have $E = \mathbb{R}^d$; if f is a function on $\mathbb{R}^d$, a notation like $f * \nu$ means that we integrate $W(\omega, t, x) = f(x)$ with respect to ν; if f is multidimensional, we integrate componentwise and the result is a multidimensional process; if f has an analytic expression, say for instance $f(x) = |x|^2 \wedge 1$, we also write $(|x|^2 \wedge 1) * \nu$.

The non-uniqueness refered to in Remark 2.8 allows for a good version of the characteristics:

2.9 **Proposition.** *One can find a version of the characteristics (B, C, ν) of X which is of the form:*

2.10
$$\begin{cases} B^i = b^i \cdot A \\ C^{ij} = c^{ij} \cdot A \\ \nu(\omega; dt, dx) = dA_t(\omega) K_{\omega,t}(dx) \end{cases}$$

where:

(i) *A is a predictable process in $\mathscr{A}^+_{\mathrm{loc}}$, which may be chosen continuous if and only if X is quasi-left-continuous;*

(ii) *$b = (b^i)_{i \le d}$ is a d-dimensional predictable process;*

(iii) *$c = (c^{ij})_{i,j \le d}$ is a predictable process with values in the set of all symmetric nonegative $d \times d$ matrices;*

(iv) *$K_{\omega,t}(dx)$ is a transition kernel from $(\Omega \times \mathbb{R}_+, \mathscr{P})$ into $(\mathbb{R}^d, \mathscr{R}^d)$ which satisfies:*

2.11
$$\begin{cases} K_{\omega,t}(\{0\}) = 0, \quad \displaystyle\int K_{\omega,t}(dx)(|x|^2 \wedge 1) \le 1 \\ \Delta A_t(\omega) > 0 \Rightarrow b_t(\omega) = \displaystyle\int K_{\omega,t}(dx) h(x) \\ \Delta A_t(\omega) K_{\omega,t}(\mathbb{R}^d) \le 1. \end{cases}$$

It also follows from (iii) and 2.11 above that this "good" version of (B, C, ν) satisfies identically:

2.12 $s \le t \Rightarrow (C^{ij}_t - C^{ij}_s)_{i,j \le d}$ is a symmetric nonnegative matrix;

2.13 $(|x|^2 \wedge 1) * \nu \in \mathscr{A}_{\mathrm{loc}}$ and $\nu(\{t\} \times \mathbb{R}^d) \le 1$;

2.14 $\Delta B_t = \displaystyle\int h(x) \nu(\{t\} \times dx)$.

Proof. (a) Theorem I.4.47 implies that the process $\sum_{s \le \cdot} |\Delta X_s|^2$ is in $\mathscr{V}$. Then the process $(|x|^2 \wedge 1) * \mu^X = \sum_{s \le \cdot} (|\Delta X_s|^2 \wedge 1)$ is locally integrable, because it has bounded jumps, and by definition of ν we deduce that there is a version of ν satisfying 2.13. Moreover, because of 1.17 and because $\mu^X(\mathbb{R}_+ \times \{0\}) = 0$ by construction, one may choose a version with:

2.15 $$\nu(\{t\} \times \mathbb{R}^d) \leq 1, \quad \nu(\mathbb{R}_+ \times \{0\}) = 0.$$

(b) The following formula, where B and C are any versions of the two first characteristics and ν is as above, defines a predictable process A in $\mathscr{A}^+_{\text{loc}}$:

$$A = \sum_{i \leq d} \operatorname{Var}(B^i) + \sum_{i,j \leq d} \operatorname{Var}(C^{ij}) + (|x|^2 \wedge 1) * \nu.$$

Then $dB^i \ll dA$ and $dC^{ij} \ll dA$, Hence I.3.13 implies the existence of predictable processes b^i and c^{ij} such that $B^i = b^i \cdot A$ and $C^{ij} = c^{ij} \cdot A$ up to an evanescent set.

If $A' = (|x|^2 \wedge 1) * \nu$ then $A' = V * \nu$, where $V(\omega, t, x) = |x|^2 \wedge 1 + 1_{\{0\}}(x)$ is a strictly positive $\tilde{\mathscr{P}}$-measurable function on $\tilde{\Omega}$. Hence by 1.8 there is a transition kernel $K'(\omega, t; dx)$ from $(\Omega, \mathbb{R}_+, \mathscr{P})$ into $(\mathbb{R}^d, \mathscr{R}^d)$ such that $\nu(\omega; dt, dx) = dA'_t(\omega) K'(\omega, t; dx)$. Moreover $dA' \ll dA$, hence there is a predictable process H with $A' = H \cdot A$ up to an evanescent set. Then $K_{\omega,t}(dx) = H_t(\omega) K'(\omega, t; dx)$ is again a predictable transition kernel, which meets $\nu(\omega; dt, dx) = dA_t(\omega) K_{\omega,t}(dx)$ P-a. s.

Let $a_t(\omega) = \int K_{\omega,t}(dx) |x|^2 \wedge 1$. From what precedes it follows that $A' = a \cdot A$. But $A - A'$ is increasing by construction, hence $(1 - a) \cdot A$ is increasing, hence the predictable set $F = \{a > 1\}$ meets $1_F \cdot A = 0$. Thus if we replace $K_{\omega,t}(\cdot)$ by 0 if $(\omega, t) \in F$, we still have $\nu(\omega; dt, dx) = dA_t(\omega) K_{\omega,t}(dx)$ a. s., and also $a_t(\omega) \leq 1$ identically. The same type of reasons allows to deduce from 2.15 that one may choose a kernel K that satisfies $K_{\omega,t}(\{0\}) = 0$ and $\Delta A_t(\omega) K_{\omega,t}(\mathbb{R}^d) \leq 1$.

(c) For $u \in \mathbb{Q}^d$, we define the predictable process $\alpha^u = \sum_{i,j \leq d} u^i c^{ij} u^j$, the predictable set $F(u) = \{\alpha^u < 0\}$, and the semimartingale $Y = \sum u^i X^i$. Since $C^{ij} = c^{ij} \cdot A$ and $Y^c = \sum u^i X^{i,c}$, I.4.41 implies

$$1_{F(u)} \cdot \langle Y^c, Y^c \rangle = 1_{F(u)} \alpha^u \cdot A$$

up to an evanescent set. The right-hand side above is nonpositive, and the left-hand side is nonnegative, which is possible only if $1_{F(u)} \cdot A = 0$ up to an evanescent set. Therefore, if we replace each c^{ij} by 0 on the set $F = \bigcup_{u \in \mathbb{Q}^d} F(u)$, we still have $C^{ij} = c^{ij} \cdot A$ a. s., and the new sets $F(u)$'s are empty for each $u \in \mathbb{Q}^d$. Since $C^{ij} = C^{ji}$, it follows that all matrices $(c^{ij})_{i,j \leq d}$ are symmetric nonnegative.

(d) If we consider the decomposition 2.5, we have $\Delta X(h) = \Delta M(h) + \Delta B$, hence I.2.31 yields $\Delta B = {}^p(\Delta X(h))$. On the other hand $\Delta X(h) = h(\Delta X)$, hence 1.18 implies that ${}^p(\Delta X(h))$ is the process $(\int h(x) \nu(\{t\} \times dx))_{t \geq 0}$ which is thus indistinguishable from ΔB. Therefore, if we replace b^i_t by $\int h^i(x) \nu(\{t\} \times dx)$ for each t where $\Delta A_t > 0$, and let b^i unchanged elsewhere, the new process $(b^i \cdot A)_{i \leq d}$ is still a version of the first characteristics of X.

(e) The only statement still unproved is that A may be chosen continuous if and only if X is quasi-left-continuous. This is an immediate consequence of 1.19 and of the last formula in 2.10 and of 2.14, which has just been proved in (d) above. □

In order to establish the grounds for our limit theorems, we set:

2.16 **Definition.** Let $h \in \mathscr{C}_t^d$. We call *modified second characteristic* of X (associated to h) the predictable process $\tilde{C}$ of $\mathscr{V}^{d \times d}$ defined by

$$\tilde{C}^{ij} = \langle M(h)^i, M(h)^j \rangle$$

where $M(h)$ is defined in 2.5 (note that $\Delta X(h) = h(\Delta X)$ and h is bounded, so $\Delta M(h)$ also is bounded: hence each component $M(h)^i$ is a locally square integrable martingale and $\tilde{C}^{ij}$ above is well defined). The triplet $(B, \tilde{C}, \nu)$ is called *modified characteristics* of X. □

Of course, $\tilde{C}$ depends on h, and we sometimes write $\tilde{C}(h)$ to emphazise the dependence. Remark 2.8 also applies to $\tilde{C}$. Here is an explicit computation of $\tilde{C}$ in terms of the characteristics (B, C, ν): formula 2.18 below clearly shows that one can compute $(B, \tilde{C}, \nu)$ in terms of (B, C, ν), and vice-versa.

2.17 **Proposition.** a) *Up to an evanescent set,*

2.18 $$\tilde{C}^{ij} = C^{ij} + (h^i h^j) * \nu - \sum_{s \le \cdot} \left(\int h^i(x) \nu(\{s\} \times dx)\right)\left(\int h^j(x) \nu(\{s\} \times dx)\right)$$
$$= C^{ij} + (h^i h^j) * \nu - \sum_{s \le \cdot} \Delta B_s^i \Delta B_s^j.$$

b) *If (B, C, ν) is the "good" version of the characteristics in 2.9, then one may take $\tilde{C}^{ij} = \tilde{c}^{ij} \cdot A$, where $\tilde{c} = (\tilde{c}^{ij})_{i,j \le d}$ is the predictable process with values in the set of all symmetric nonnegative $d \times d$ matrices given by*

2.19 $$\tilde{c}_t^{ij}(\omega) = c_t^{ij}(\omega) + \int K_{\omega,t}(dx)(h^i h^j)(x)$$
$$- \Delta A_t(\omega)\left[\int K_{\omega,t}(dx) h^i(x)\right]\left[\int K_{\omega,t}(dx) h^j(x)\right].$$

Proof. (b) immediately follows from (a) and 2.10, while the equality between the two right-hand sides of 2.18 follows from 2.14. From I.4.49 we have $\sum_{s \le \cdot} \Delta B_s^i \Delta B_s^j = [B^i, B^j]$, while there exists a constant k such that $|h^i h^j(x)| \le k(|x|^2 \wedge 1)$. Hence by 2.13 we obtain that the right-hand sides of 2.18 make sense and that $(h^i h^j) * \mu^X \in \mathscr{A}_{\mathrm{loc}}$.

Since $M := M(h) = X(h) - X_0 - B$, a simple computation shows

$$[M^i, M^j] = [X(h)^i, X(h)^j] - [B^i, B^j] - [M^i, B^j] - [M^j, B^i].$$

$[M^j, B^i] \in \mathscr{L}$ by I.4.49c, and $[B^i, B^j]$ is predictable, and $\langle M^i, M^j \rangle$ is the compensator of $[M^i, M^j]$. Thus

$$\langle M^i, M^j \rangle = ([X(h)^i, X(h)^j])^p - [B^i, B^j].$$

For obtaining the result, it remains to notice that

$$[X(h)^i, X(h)^j] = C^{ij} + \sum_{s \leq \cdot} \Delta X(h)^i_s \Delta X(h)^j_s = C^{ij} + (h^i h^j) * \mu^X$$

by I.4.53 and the definition of μ^X, and to use the definition of ν and the fact that $(h^i h^j) * \mu^X \in \mathscr{A}_{\text{loc}}$. □

It is clear from Definition 2.6 that (B, C, ν) are characterized in terms of martingales. However, for the sake of later use for convergence theorems, we explicitely state this property in a more tractable way.

On one hand we start with a càdlàg adapted process $X = (X^i)_{i \leq d}$. On the other hand, we are given a triplet (B, C, ν) which is a candidate to being the characteristics of X, relative to some fixed $h \in \mathscr{C}_t^d$: that is, we choose a predictable d-dimensional process B in $\mathscr{V}^d$, a continuous matrix-valued process $C \in \mathscr{V}^{d \times d}$, and a predictable random measure ν on $\mathbb{R}_+ \times \mathbb{R}^d$, such that *2.12, 2.13 and 2.14 are in force.* We also define $\tilde{C}$ by 2.18.

Finally we introduce a set of functions as such:

2.20 $\mathscr{C}^+(\mathbb{R}^d)$ is any family of bounded Borel functions on $\mathbb{R}^d$, vanishing inside a neighbourhood of 0, with the following property: if any two positive measures η and η' on $\mathbb{R}^d$ with $\eta(\{0\}) = \eta'(\{0\}) = 0$ and $\eta(x\colon |x| > \varepsilon) < \infty$ and $\eta'(x\colon |x| > \varepsilon) < \infty$ for all $\varepsilon > 0$ are such that $\eta(f) = \eta'(f)$ for all $f \in \mathscr{C}^+(\mathbb{R}^d)$, then $\eta = \eta'$ (there are many such families; there are such families that are countable and contain only continuous, or even C^∞, functions). □

2.21 **Theorem.** *With the above notation,* there is equivalence between:

a) *X is a semimartingale with characteristics* (B, C, ν).

b) *The following processes are local martingales*:

(i) $M(h) = X(h) - B - X_0$ ($X(h)$ is defined by 2.4);

(ii) $M(h)^i M(h)^j - \tilde{C}^{ij}$ *for* $i, j \leq d$;

(iii) $g * \mu^X - g * \nu$ *for* $g \in \mathscr{C}^+(\mathbb{R}^d)$.

Proof. (a) ⇒ (b): (i) is the definition of the first characteristic B, while (ii) follows from 2.16 and 2.17. Moreover, $g * \mu^X$ is an increasing locally bounded process (because g is bounded, and X has only finitely many jumps in each interval $[0, t]$ and lying in the support of g), so (iii) follows from the definition of the third characteristic ν.

(b) ⇒ (a): By (i), $X = X_0 + \check{X}(h) + M(h) + B$ is clearly a semimartingale, with B as its first characteristic. (iii) shows that, if (T_n) is a localizing sequence for the locally integable processes $g * \mu^X$ and $g * \nu$ (where $g \in \mathscr{C}^+(\mathbb{R}^d)$) and T is any stopping time,

2.22 $$E(g * \mu^X_{T_n \wedge T}) = E(g * \nu_{T_n \wedge T}).$$

Letting $n \uparrow \infty$, we get

2.23 $$E(W * \mu^X_\infty) = E(W * \nu_\infty)$$

for all W of the form $W(\omega, t, x) = g(x)1_{[0,T(\omega)]}(t)$, $g \in \mathscr{C}^+(\mathbb{R}^d)$, T stopping time. Then 2.20 plus a monotone class argument easily shows that 2.23 holds for all predictable nonnegative functions on $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}^d$ (recall I.2.2, and $\mu^X(\{0\} \times \mathbb{R}^d) = \nu(\{0\} \times \mathbb{R}^d) = 0$). So ν is the third characteristic of X.

Then (ii) tells us that $\tilde{C}$ is the second modified characteristic, and comparing to 2.17 yields that C is the second characteristic. □

We end this subsection by showing how $B = B(h)$ and $\tilde{C} = \tilde{C}(h)$ do indeed depend on h.

2.24 **Proposition.** *Let $h, h' \in \mathscr{C}_t^d$. Then, up to an evanescent set,*

2.25 $$B(h) - B(h') = (h - h') * \nu$$

2.26 $$\begin{aligned}\tilde{C}(h)^{ij} - \tilde{C}(h')^{ij} &= (h^i h^j - h'^i h'^j) * \nu \\ &\quad - \sum_{s \le \cdot} \left[\int h^i(x)\nu(\{s\} \times dx) \int h^j(x)\nu(\{s\} \times dx) \right. \\ &\quad \left. - \int h'^i(x)\nu(\{s\} \times dx) \int h'^j(x)\nu(\{s\} \times dx) \right]\end{aligned}$$

Proof. 2.26 is immediate from 2.18 and 2.25. To prove 2.25, we remark first that the process $X(h) - X(h')$ (see 2.4) is in $\mathscr{S}^d$ and has bounded jumps, so its components are in $\mathscr{A}_{\mathrm{loc}}$, and by definition of μ^x it has the form

$$X(h) - X(h') = (h - h') * \mu^X.$$

Hence Theorem 1.8 implies that $(h - h') * \nu$ belongs to $\mathscr{V}^d$ and that $X(h) - X(h') - (h - h') * \nu$ is a d-dimensional local martingale. Then the uniqueness of the canonical decomposition of the special semimartingale $X(h) - X(h')$ gives the result. □

§ 2b. Integrability and Characteristics

This subsection is essentially devoted to proving that one can actually read on the characteristics of the semimartingale X whether it is special, or (locally) square-integrable. It can be skipped at first reading. We begin with a definition:

2.27 **Definition.** The (d-dimensional) semimartingale X is called *locally square-integrable* if it is a special semimartingale whose canonical decomposition $X = X_0 + N + A$ satisfies: N is a locally square-integrable (d-dimensional) local martingale (i.e., its components are in $\mathscr{H}^2_{\mathrm{loc}}$). □

We slightly abuse the terminology in 2.27, as seen in the following:

2.28 **Lemma.** *The semimartingale X is locally square-integrable* (in the sense of 2.27) *if and only if the increasing process $Y_t = \sup_{s \le t} |X_s - X_0|^2$ is locally integrable.*

Proof. Assume first that X is locally square-integrable, with canonical decomposition $X = X_0 + N + A$. Then

$$Y_t \le 2 \sup_{s \le t} |N_s|^2 + 2 \sup_{s \le t} |A_s|^2.$$

There is a localizing sequence (T_n) such that $N^{T_n} \in \mathscr{H}^2$ and $\mathrm{Var}(A)_{T_n}$ is bounded (apply I.3.10). Then Y_{T_n} is integrable.

Conversely, assume that Y is locally integrable. Then I.4.23 implies that X is special, and we again call $X = X_0 + N + A$ its canonical decomposition. Let (T_n) be a localizing sequence with $E(Y_{T_n}) < \infty$ and $\mathrm{Var}(A)_{T_n} \le n$. Then

$$\sup_s |N_s^{T_n}|^2 = \sup_{s \le T_n} |N_s|^2 \le 2\, Y_{T_n} + 2\,\mathrm{Var}(A)_{T_n}$$

is integrable, and so $N^{T_n} \in \mathscr{H}^2$. □

2.29 **Proposition.** *Let X be a semimartingale with characteristics $(B(h), C, \nu)$ relative to $h \in \mathscr{C}_t^d$.*

a) *X is a special semimartingale if and only if $(|x|^2 \wedge |x|) * \nu \in \mathscr{A}_{\mathrm{loc}}$. In this case, the canonical decomposition $X = X_0 + N + A$ satisfies:*

2.30
$$\begin{cases} A = B(h) + (x - h(x)) * \nu \\ \Delta A_t = \displaystyle\int x \nu(\{t\} \times dx). \end{cases}$$

b) *X is a locally square-integrable semimartingale if and only if $|x|^2 * \nu \in \mathscr{A}_{\mathrm{loc}}$. In this case its canonical decomposition $X = X_0 + N + A$ satisfies 2.30 and*

2.31
$$\langle N^i, N^j \rangle = \begin{cases} C^{ij} + (x^i x^j) * \nu - \displaystyle\sum_{s \le \cdot} \int x^i \nu(\{s\} \times dx) \int x^j \nu(\{s\} \times dx) \\ C^{ij} + (x^i x^j) * \nu - \displaystyle\sum_{s \le \cdot} \Delta A_s^i \Delta A_s^j. \end{cases}$$

Proof. a) We use the notation 2.4 and 2.5. Since $X(h)$ is special, then X is special if and only if $\check{X}(h)$ is so; since $\check{X} \in (h) \in \mathscr{V}^d$, from I.4.23 this is the case if and only if $\check{X}(h) \in (\mathscr{A}_{\mathrm{loc}})^d$. Now $\check{X}(h) = (x - h(x)) * \mu^X$ by definition, hence by I.3.20 this is the case if and only if $|x - h(x)| * \nu$ belongs to $\mathscr{A}_{\mathrm{loc}}$. Finally there are constants $C' > C > 0$ such that $|x - h(x)| \le C(|x|^2 \wedge |x|) \le C'((|x|^2 \wedge 1) + |x - h(x)|)$, so in view of 2.13 we obtain the claimed necessary and sufficient condition.

By 2.4 and 2.5 we have $\check{X}(h) = N + A - M(h) - B(h)$ if $X = X_0 + N + A$ is the canonical decomposition of X. This gives that $A - B(h)$ is the compensator of $\check{X}(h) = (x - h(x)) * \mu^X$, thus yielding the first equality in 2.30. The second one is deduced from 2.14.

b) Assume first that $|x|^2 * \nu \in \mathscr{A}_{\text{loc}}$. Then X is special by (a) and call $X = X_0 + N + A$ its canonical decomposition. The assumption clearly implies that the right-hand sides of 2.31 are meaningful, and they are equal by 2.30. The same computation than in the proof of 2.17 (N and A replacing $M(h)$ and $B(h)$) shows that:

(1) $$[N^i, N^j] = C^{ij} + (x^i x^j) * \mu^X - [A^i, A^j] - [A^i, N^j] - [A^j, N^i].$$

We have $C^{ij} \in \mathscr{A}_{\text{loc}}$, $[A^i, A^j] \in \mathscr{A}_{\text{loc}}$ (by I.3.10), $(x^i x^j) * \mu^X \in \mathscr{A}_{\text{loc}}$ because of the assumption, $[A^i, N^j]$ and $[A^j, N^i]$ belong to $\mathscr{A}_{\text{loc}} \cap \mathscr{L}$ (see I.3.11 and I.4.49c). Hence $[N^i, N^j] \in \mathscr{A}_{\text{loc}}$ and I.4.50c implies that $N^i \in \mathscr{H}^2_{\text{loc}}$, hence 2.27 holds. Moreover

$$\langle N^i, N^j \rangle = ([N^i, N^j])^p = C^{ij} + (x^i x^j) * \nu - [A^i, A^j],$$

thus yielding 2.31.

Assume conversely that X is locally square-integrable. Then $[N^i, N^i] \in \mathscr{A}_{\text{loc}}$ and (1) will imply that $(x^i)^2 * \mu^X \in \mathscr{A}_{\text{loc}}$, hence $(x^i)^2 * \nu \in \mathscr{A}_{\text{loc}}$ (see I.3.20), hence $|x|^2 * \nu \in \mathscr{A}_{\text{loc}}$. □

The next result is an exercise about characteristics.

2.32 **Proposition.** *Let X be a semimartingale with characteristics (B, C, ν) relative to a truncation function h. Then the semimartingale $X' = X - B$ admits the following characteristics (B', C', ν'), relative to the same truncation function:*

2.33 $$\begin{cases} B'_t = \sum_{s \le t} \left[\int_{\mathbb{R}^d} \nu(\{s\} \times dx) h(x - \Delta B_s) + h(-\Delta B_s)(1 - \nu(\{s\} \times \mathbb{R}^d)) \right] \\ C' = C \\ \nu'([0,t] \times A) = \int_0^t \int_{\mathbb{R}^d} 1_A(x - \Delta B_s) 1_{\{x \ne \Delta B_s\}} \nu(ds, dx) \\ \qquad\qquad + \sum_{s \le t} [1 - \nu(\{s\} \times \mathbb{R}^d)] 1_{\{\Delta B_s \ne 0\}} 1_A(-\Delta B_s) \end{cases}$$

Proof. That $C' = C$ follows from $X'^c = X^c$ (see I.4.27). Call μ^X and $\mu^{X'}$ the measures associated with X and X' by 1.16, and let W be $\mathscr{P} \otimes \mathscr{R}^d$-measurable and nonnegative. Put $W'(\omega, t, x) = W(\omega, t, x - \Delta B_t(\omega)) 1_{\{x \ne \Delta B_t(\omega)\}}$. Since $\Delta X' = \Delta X - \Delta B$ we have

$$W * \mu^{X'}_\infty = W' * \mu^X_\infty + \sum_{s > 0} W(s, -\Delta B_s) 1_{\{\Delta B_s \ne 0, \Delta X_s = 0\}}.$$

By definition of ν, we have $E(W' * \mu^X_\infty) = E(W' * \nu_\infty)$. Let $D = \{\Delta X \ne 0\}$ and $J = \{(\omega, t) \colon \nu(\omega; \{t\} \times \mathbb{R}^d) > 0\}$ and $a_t(\omega) = \nu(\omega; \{t\} \times \mathbb{R}^d)$. By 1.17, J is the predictable support of the thin optional set D, hence by I.2.23 it admits a sequence (T_p) of exhausting predictable times, up to a null set. Then

$$E(W * \mu_\infty^{X'}) = E(W' * \nu_\infty) + \sum_{p \geq 1} E[W(T_p, -\Delta B_{T_p}) 1_{\{\Delta B_{T_p} \neq 0\}} 1_{D^c}(T_p) 1_{\{T_p < \infty\}}]$$
$$= E(W' * \nu_\infty) + \sum_{p \geq 1} E[W(T_p, -\Delta B_{T_p}) 1_{\{\Delta B_{T_p} \neq 0\}} (1 - a_{T_p}) 1_{\{T_p < \infty\}}]$$

by 1.17a, and if ν' is defined by 2.33 this is equal to $E(W * \nu'_\infty)$. We deduce that ν' is the third characteristic of X'.

We use the notation 2.4 and 2.5. We have

$$X'(h)_t = X'_t - \sum_{s \leq t} [\Delta X'_s - h(\Delta X'_s)]$$
$$= X'_0 + M(h)_t + \sum_{s \leq t} [\Delta X_s - h(\Delta X_s) - \Delta X'_s + h(\Delta X'_s)]$$

Let A_t be the last sum above. The components of the process A are in $\mathscr{V}$ and, since $X'(h)$ is a special semimartingale, we deduce from I.4.23 that they are even in $\mathscr{A}_{\text{loc}}$. From the definition 2.5 of $B' = B'(h)$ we deduce that $B' = A^p$, the compensator of A. But A is the sum of its jumps, and $\Delta A_t = 0$ whenever $\Delta B_t = 0$ (because then $\Delta X_t = \Delta X'_t$), so the process $1_{J^c} \cdot A$ is identically 0. Hence $1_{J^c} \cdot A^p = 0$ as well and we deduce that $B' = A^p$ is the sum of its jumps. Then the first formula in 2.33 follows from the last one and from 2.14. □

§ 2c. A Canonical Representation for Semimartingales

In the present subsection we exhibit a *canonical representation* for multidimensional semimartingales, which is based upon the stochastic integrals defined in § 1d, and which is not to be confounded with the canonical *decomposition* I.4.22 of a special semimartingale.

2.34 **Theorem.** *Let X be a d-dimensional semimartingale, with characteristics (B, C, ν) relative to a truncation function $h \in \mathscr{C}_t^d$, and with the measure μ^X associated to its jumps by 1.16. Then $W^i(\omega, t, x) = h^i(x)$ belongs to $G_{\text{loc}}(\mu^X)$ for all $i \leq d$, and the following representation holds:*

2.35 $$X = X_0 + X^c + h * (\mu^X - \nu) + (x - h(x)) * \mu^X + B.$$

This formula 2.35 is called the *canonical representation* of X.

The stochastic integral $h * (\mu^X - \nu)$ is d-dimensional, and should be read componentwise. Compare this to 2.4 and 2.5: we have $\check{X}(h) = (x - h(x)) * \mu^X$ by definition of $\check{X}(h)$ and μ^X. Then

2.36 $$M(h) = X^c + h * (\mu^X - \nu),$$

so in other words the purely discontinuous part of the d-dimensional local martingale $M(h)$ is $M(h)^d = h * (\mu^X - \nu)$.

Proof. We need only show $M(h)^d = h * (\mu^X - \nu)$. We have $\Delta M(h) = h(\Delta X) - \Delta B$, while if $W^i(\omega, t, x) = h^i(x)$, 2.14 yields $\Delta B^i = \widehat{W^i}$ (with notation 1.24). Thus if we recall that $D = \{\Delta X \neq 0\}$ and $\beta = \Delta X$ for $\mu = \mu^X$, then $\Delta M(h)^i = \widetilde{W^i}$ (see 1.27). In view of I.4.56 we deduce that $W^i \in G_{\text{loc}}(\mu^X)$ and the two purely discontinuous local martingales $M(h)^i$ and $W^i * (\mu^X - \nu)$, having the same jumps, are indistinguishable. □

2.37 **Remark.** The previous result would provide another proof for formula 2.18: just use 1.34 (the proofs of 1.33a and 2.17 are indeed very alike). □

The following corollary complements 2.29a:

2.38 **Corollary.** *Let X be a d-dimensional special semimartingale with characteristics (B, C, ν) and μ^X the measure associated to its jumps by 1.16. Then $W^i(\omega, t, x) = x^i$ belongs to $G_{\text{loc}}(\mu^X)$, and if $X = X_0 + N + A$ is its canonical decomposition, then*

2.39
$$X = X_0 + X^c + x * (\mu^X - \nu) + A$$

Proof. That $|x - h(x)| * \nu \in \mathscr{A}_{\text{loc}}$ follows from 2.29a, so $x^i - h^i(x) \in G_{\text{loc}}(\mu^X)$ by 1.28 and we deduce from 2.34 that $x^i \in G_{\text{loc}}(\mu^X)$. Then 2.39 readily follows from 2.35 and 2.30. □

§ 2d. Characteristics and Exponential Formula

1. In this subsection, we make the various comments written in the introduction, and especially around 2.2, more precise.

Exactly as before Theorem 2.21, we start on one side with a d-dimensional adapted process $X = (X^i)_{i \leq d}$.

On the other side, we are given a triplet (B, C, ν) which is a candidate to being the characteristics of X (associated to some fixed $h \in \mathscr{C}_t^d$). That is, we choose a predictable d-dimensional process $B = (B^i)_{i \leq d}$ in $\mathscr{V}^d$, a continuous matrix-valued process $C = (C^{ij})_{i,j \leq d}$ in $\mathscr{V}^{d \times d}$, and a predictable random measure ν on $\mathbb{R}_+ \times \mathbb{R}^d$. And we suppose that *2.12, 2.13, 2.14 hold* (or equivalently that B, C, ν are given by 2.10 with A, b, c, K meeting all the conditions stated in 2.9).

Now we associate with the triplet (B, C, ν) a process that plays the rôle of the function ψ_t in 2.1. To this effect, we introduce first the usual scalar product notation: if $u, x \in \mathbb{R}^d$, then $u \cdot x = \sum_{i \leq d} u^i x^i$ and similarly for the processes $u \cdot B$ or $u \cdot X$. Analoguously if $z = (z^{ij})_{i,j \leq d}$ is a matrix, we denote by $u \cdot z \cdot u$ the number $u \cdot z \cdot u = \sum_{i,j \leq d} u^i z^{ij} u^j$, and similarly for the process $u \cdot C \cdot u$.

We have $|e^{iu \cdot x} - 1 - iu \cdot h(x)| \leq \alpha(|x|^2 \wedge 1)$ for some constant α, hence since $(|x|^2 \wedge 1) * \nu \in \mathscr{V}$ by hypothesis we may set

2.40 $$A(u)_t = iu\cdot B_t - \frac{1}{2}u\cdot C_t\cdot u + \int (e^{iu\cdot x} - 1 - iu\cdot h(x))\nu([0,t]\times dx).$$

The complex-valued process $A(u)$ such defined is predictable with finite variation (i.e., its real and purely imaginary parts are in $\mathscr{V}$): this is evident for the first and the third terms, and the second one is continuous and increasing by 2.12. Notice that $A(u)$ does not depend on h, by 2.25.

Equivalently, with the notation 2.10,

2.41 $$\begin{cases} A(u) = a(u)\cdot A \\ a(u) = iu\cdot b - \frac{1}{2}u\cdot c\cdot u + \int (e^{iu\cdot x} - 1 - iu\cdot h(x))K(dx). \end{cases}$$

2.42 **Theorem.** *With the preceding notation, there is equivalence between:*

a) *X is a semimartingale, and it admits the characteristics (B, C, ν).*

b) *For each $u\in\mathbb{R}^d$, the process $e^{iu\cdot X} - (e^{iu\cdot X_-})\cdot A(u)$ is a complex-valued local martingale.*

c) *For each bounded function f of class C^2 on $\mathbb{R}^d$, the process*

2.43 $$f(X) - f(X_0) - \sum_{j\le d} D_j f(X_-)\cdot B^j - \frac{1}{2}\sum_{j,k\le d} D_{jk}f(X_-)\cdot C^{jk} - \left\{f(X_- + x) - f(X_-) - \sum_{j\le d} D_j f(X_-)h^j(x)\right\} * \nu$$

is a local martingale.

Let us begin with a lemma (see Gnedenko and Kolmogorov [65]).

2.44 **Lemma.** *Let $b\in\mathbb{R}^d$, let c be a symmetric nonnegative $d\times d$ matrix, and let F be a positive measure on $\mathbb{R}^d$ which integrates $(|x|^2\wedge 1)$ and satisfies $F(\{0\}) = 0$. Then the function*

2.45 $$\psi(u) = iu\cdot b - \frac{1}{2}u\cdot c\cdot u + \int (e^{iu\cdot x} - 1 - iu\cdot h(x))F(dx)$$

admits a unique representation of the form 2.45, i.e. if ψ satisfies 2.45 with (b', c', F') also, then $b' = b$ and $c' = c$ and $F' = F$.

Proof. Let $w\in\mathbb{R}^d\setminus\{0\}$ and define the function $\varphi_w(u) = \psi(u) - \frac{1}{2}\int_{-1}^{1}\psi(u+sw)\,ds$. A simple computation shows that

$$\varphi_w(u) = \frac{1}{6}w\cdot c\cdot w + \int \left(1 - \frac{\sin w\cdot x}{w\cdot x}\right)e^{iu\cdot x}F(dx).$$

Hence the function φ_w is the characteristic function of the measure

$$G_w(dx) = \frac{1}{6}w\cdot c\cdot w\varepsilon_0(dx) + \left(1 - \frac{\sin w\cdot x}{w\cdot x}\right)F(dx).$$

Therefore each measure G_w is uniquely determined by the function φ_w, or by the function ψ. On the other hand, the set of measures $(G_w: w \in \mathbb{R}^d \setminus \{0\})$ completely determines the matrix c and the measure F (recall that $F(\{0\}) = 0$). Finally, b is easily computed from c, F, and the function ψ. □

Proof of Theorem 2.42. (a) ⇒ (c): We consider the decompositions 2.4 and 2.5, with $B = B(h)$, and to simplify we put $M = M(h)$. Then, an application of Ito's formula yields:

$$\begin{aligned} f(X) - f(X_0) = & \sum_{j \le d} D_j f(X_-) \cdot M^j + \sum_{j \le d} D_j f(X_-) \cdot B^j \\ & + \sum_{s \le \cdot} \sum_{j \le d} D_j f(X_{s-})(\Delta X_s^j - h^j(\Delta X_s)) \\ & + \frac{1}{2} \sum_{j,k \le d} D_{jk} f(X_-) \cdot \langle X^{j,c}, X^{k,c} \rangle \\ & + \sum_{s \le \cdot} \left[f(X_s) - f(X_{s-}) - \sum_{j \le d} D_j f(X_{s-}) \Delta X_s^j \right] \\ = & \sum_{j \le d} D_j f(X_-) \cdot M^j + \sum_{j \le d} D_j f(X_-) \cdot B^j \\ & + \frac{1}{2} \sum_{j,k \le d} D_{jk} f(X_-) \cdot C^{jk} + W * \mu^X, \end{aligned}$$

where $W(t,x) = f(X_{t-} + x) - f(X_{t-}) - \sum_{j \le d} D_j f(X_{t-}) h^j(x)$. Let us consider the right-hand side above: the first term is in $\mathscr{L}$; the second and third terms are predictable processes in $\mathscr{V}$; the last term is also in $\mathscr{V}$. But f being bounded, $f(X)$ is a special semimartingale, and I.4.23 implies that this last term is actually in $\mathscr{A}_{\mathrm{loc}}$. Therefore we have $W * \nu - W * \mu^X \in \mathscr{L}$, and the result follows.

(c) ⇒ (b): The only thing to do is to apply (c) to the function $f(x) = e^{iu \cdot x}$ and to notice that $D_j f = iu^j f$, $D_{jk} f = -u^j u^k f$, and

$$f(X_- + x) - f(X_-) - \sum_{j \le d} D_j f(X_-) h^j(x) = e^{iu \cdot X_-}(e^{iu \cdot x} - 1 - iu \cdot h(x)).$$

(b) ⇒ (a): By hypothesis, $e^{iu \cdot X}$ is a complex-valued semimartingale for each $u \in \mathbb{R}^d$. Then, $\sin aX^j$ is in $\mathscr{S}$ for all $a \in \mathbb{R}$. There exists a function f of class C^2 on $\mathbb{R}$, such that $f(\sin x) = x$ if $|x| \le 1/2$. Hence, if $T_n = \inf(t: |X_t^j| > n/2)$, the process X^j coincide with the semimartingale $nf(\sin(X^j/n))$ on the stochastic interval $[\![0, T_n[\![$, and I.4.25 yields $X^j \in \mathscr{S}$.

Let (B', C', ν') be a good version of the characteristics of X, which satisfies all the properties of 2.9. For each $u \in \mathbb{R}^d$ we associate to (B', C', ν') a process $A'(u)$ by 2.40. We have proved the implication (a) ⇒ (b), thus $e^{iu \cdot X} - (e^{iu \cdot X_-}) \cdot A'(u)$ is a local martingale. Then the hypothesis and the uniqueness of the canonical decomposition of the special semimartingale $e^{iu \cdot X}$ show that $(e^{iu \cdot X_-}) \cdot A(u) = (e^{iu \cdot X_-}) \cdot A'(u)$ up to an evanescent set. Integrating the process $e^{-iu \cdot X_-}$, we obtain that $A(u)$ and $A'(u)$ are indistinguishable. Therefore, the set N of all ω for which there exists $u \in \mathbb{Q}^d$ and $t \in \mathbb{Q}_+$ with $A(u)_t(\omega) \neq A'(u)_t(\omega)$ is P-null.

Now we remark that the function ψ defined by 2.45 is continuous, and hence completely characterized by its values on $\mathbb{Q}^d$. Then if we apply Lemma 2.44 we obtain that outside N, we have $B'_t = B_t$, $C'_t = C_t$ and $\nu'([0,t] \times \cdot) = \nu([0,t] \times \cdot)$ for all $t \in \mathbb{Q}_+$, then also for all $t \in \mathbb{R}_+$ because of the right-continuity. Therefore, (B, C, ν) is also a version of the local characteristics of X. □

2. What comes now, and until the end of § 2d, may be skipped at first reading.

2.46 Definition. If T is a predictable time announced by a sequence (T_n), a *local martingale on* $[\![0, T[\![$ is a process M such that each stopped process M^{T_n} is a local martingale. □

It is obvious that this notion does not depend on the announcing sequence (T_n), and that one may find one for which each M^{T_n} is uniformly integrable. Also remark that the values of M outside $[\![0, T[\![$ do not matter, so that it is sufficient that M be defined on the interval $[\![0, T[\![$.

The following theorem is the promised generalization of 2.2:

2.47 Theorem. *Let X be a d-dimensional semimartingale, with characteristics (B, C, ν), and $A(u)$ defined by 2.40. Let $T(u) = \inf(t\colon \Delta A(u)_t = -1)$ and $G(u) = \mathscr{E}[A(u)]$* (defined by I.4.63).

a) *$T(u)$ is a predictable time, $T(u) = \inf(t\colon G(u)_t = 0)$, and the process $(e^{iu\cdot X}/G(u))1_{[\![0,T(u)[\![}$ is a local martingale on $[\![0, T(u)[\![$.*

b) *If G' is another complex-valued predictable process with finite variation, such that $G'_0 = 1$ and that $(e^{iu\cdot X}/G')1_{[\![0,T'[\![}$ is a local martingale on $[\![0, T'[\![$, where $T' = \inf(t\colon G'_t = 0)$, then $T' \leq T(u)$ a.s. and $G' = G(u)$ on $[\![0, T'[\![$ up to an evanescent set.*

Proof. $u \in \mathbb{R}^d$ being fixed, we set $A = A(u)$, $G = G(u)$, $T = T(u)$, $Y = e^{iu\cdot X}$ and $Z = (Y/G)1_{[\![0,T[\![}$.

a) We show exactly as in the proof of I.2.24 that T is a stopping time (even if the basis is not complete): with the notation I.2.24, take $X = A$, $B_m = \{-1\}$, so $T = R(m, 1)$. Moreover $[\![T]\!]$ is included into the predictable set $\{\Delta A = -1\}$, so by I.2.13 it is a predictable time.

Next I.4.63 gives an explicit form for G, and taking the modulus yields:

$$|G_t| = \exp\left[-\frac{1}{2}u\cdot C_t\cdot u - (1 - \cos u\cdot x) * \nu_t\right] \prod_{s\leq t} |1 + \Delta A_s|$$

(recall that $A = A(u)$ is given by 2.40); moreover 2.40 easily yields $|1 + \Delta A_s| \leq 1$: hence the predictable process $|G|$ is decreasing. Therefore $R_n = \inf(t\colon |G_t| \leq 1/n)$ is a predictable time (apply I.1.28 and I.2.13). Since $R_n > 0$, there exists a stopping time S_n such that $S_n < R_n$ a.s. and $P(S_n < R_n - 1/n) \leq 2^{-n}$. Finally, set $T_n = \sup_{m \leq n} S_m$. By I.4.61c we have $\lim_n \uparrow R_n = T = \inf(t: G_t = 0)$. Then one checks easily that $\lim_n \uparrow T_n = T$ a.s. and that $|G| \geq 1/n$ on the interval $[\![0, T_n]\!]$.

We have to prove that for each n, the process Z^{T_n} is a (complex-valued) local martingale. Since $A(u)^{T_n}$ is associated to X^{T_n} just as $A(u)$ is associated to X, and since $G^{T_n} = \mathscr{E}(A^{T_n})$, if we replace X by X^{T_n} we have to prove that Z is a local martingale, knowing that $|G| \geq 1/n$ everywhere.

So from now on we assume that $|G| \geq 1/n$. We choose a function f of class C^2 on $\mathbb{R}^4$, with $f(x, y, u, v) = \dfrac{x + iy}{u + iv}$ whenever $|u + iv| \geq 1/n$. Ito's formula yields

$$(1) \qquad Z = 1 + \frac{1}{G_-} \cdot Y - \frac{Z_-}{G_-} \cdot G + \sum_{s \leq \cdot} \left(\Delta Z_s - \frac{\Delta Y_s}{G_{s-}} + \frac{Z_{s-}}{G_{s-}} \Delta G_s \right).$$

Recall that $N = Y - Y_- \cdot A$ is a local martingale (2.42), while $G = 1 + G_- \cdot A$ by definition of G. Moreover a tedious but simple calculation shows that $\Delta Z - \Delta Y/G_- + Z_- \Delta G/G_- = -\Delta N \Delta A/G_-(1 + \Delta A)$. Hence (1) gives

$$\begin{aligned} Z &= 1 + \frac{1}{G_-} \cdot (N + Y_- \cdot A) - \frac{Z_-}{G_-} \cdot (G_- \cdot A) - \sum_{s \leq \cdot} \frac{\Delta N_s \Delta A_s}{G_{s-}(1 + \Delta A_s)} \\ (2) \qquad &= 1 + \frac{1}{G_-} \cdot N - \frac{1}{G_-(1 + \Delta A)} \cdot \sum_{s \leq \cdot} \Delta N_s \Delta A_s. \end{aligned}$$

Now, N and $\sum_{s \leq \cdot} \Delta N_s \Delta A_s$ are local martingales (use I.4.49c for the latter, after writing $N = N' + iN''$ and $A = A' + iA''$, with N', N'', A', A'' real-valued). Moreover, the processes $1/G_-$ and $1/(1 + \Delta A)$ are locally bounded because $|G| \geq 1/n$, and so $T = \infty$, and so $1 + \Delta A \neq 0$ everywhere. Therefore (2) shows that Z is also a local martingale, and the claim is proved.

b) Let G' satisfy the assumptions of (b), and $Z' = (Y/G') 1_{[\![0, T'[\![}$. Exactly like in (a), we may find an announcing sequence (T'_n) for T' such that $|G'| \geq 1/n$ on $[\![0, T'_n]\!]$. We put $A'(n) = \dfrac{1}{G'_-} \cdot G'^{T'_n}$. Since $G'_0 = 1$ we have $G'^{T'_n} = 1 + G'_- \cdot A'(n)$, hence $G'^{T'_n} = \mathscr{E}[A'(n)]$.

We have $Y^{T'_n} = Z'^{T'_n} G'^{T'_n}$, and $Z'^{T'_n}$ is a local martingale by hypothesis. Applying I.4.49c (to the real and imaginary parts separately!) we obtain that $Y^{T'_n} - Z'_- \cdot G'^{T'_n} = Y^{T'_n} - Y_- \cdot A'(n)$ is a local martingale. By 2.42, $Y^{T'_n} - Y_- \cdot A^{T'_n}$ is also a local martingale, and the uniqueness of the canonical decomposition of the special semimartingale $Y^{T'_n}$ shows that $Y_- \cdot A'(n) = Y_- \cdot A^{T'_n}$; hence $A'(n) = A^{T'_n}$ and $G'^{T'_n} = \mathscr{E}[A'(n)] = G^{T'_n}$ by definition of G. Therefore $G' = G$ on $[\![0, T'[\![$ and from the definition of T and T', this in turn implies that $T' \leq T$. □

2.48 **Corollary.** *Under the assumptions of Theorem 2.42, and under the additional assumption that* $\Delta A(u) \neq -1$ *identically* (or, equivalently, that $G(u) = \mathscr{E}[A(u)]$ never vanishes) *for all* $u \in \mathbb{R}^d$, *there is equivalence between*:

(a) *X is a semimartingale with characteristics (B, C, ν);*
(b) *For each $u \in \mathbb{R}^d$, $e^{iu \cdot X}/G(u)$ is a local martingale.*

Proof. (a) $\Rightarrow$ (b) by 2.47. Conversely, assume (b), and let $M(u) = e^{iu \cdot X}/G(u)$. Then since $G(u)$ is a process with finite variation, $e^{iu \cdot X} = G(u)M(u)$ is a semimartingale and I.4.49b yields

$$\begin{aligned} e^{iu \cdot X} &= e^{iu \cdot X_0} + G(u) \cdot M(u) + M(u)_- \cdot G(u) \\ &= e^{iu \cdot X_0} + G(u) \cdot M(u) + [M(u)_- G(u)_-] \cdot A(u) \end{aligned}$$

because of the definition of $G(u)$. Then $e^{iu \cdot X} - (e^{iu \cdot X_-}) \cdot A(u)$ is a local martingale, and (a) follows from 2.42. □

When $G(u)$ vanishes, the property 2.47(a) cannot characterize (B, C, ν) in general, because it does not say anything on what happens on the interval $]\!]T(u), \infty[\![$ and it may occur that too many $T(u)$'s are too small!

Nevertheless, we can obtain a characterization of this type (and even two, slightly different, characterizations): they extend Theorem 2.47, but are much less important and will not be used in the sequel; henceforth we state the next theorem *without proof* (the proof is essentially the same than for 2.47). We start with a triplet (B, C, ν) and the associated processes $A(u)$ by 2.40. Set

$$\begin{aligned} &T_0(u) = 0, \quad T_{n+1}(u) = \inf(t > T_n(u) \colon \Delta A(u)_t = -1) \\ &K(u) = \{\Delta A(u) = -1\} = \bigcup_{n \geq 1} [\![T_n(u)]\!] \\ &X'(u) = X - \sum_{n \geq 1} \Delta X_{T_n(u)} 1_{[\![T_n(u), \infty[\![} \\ &A'(u) = A(u) - \sum_{n \geq 1} \Delta A(u)_{T_n(u)} 1_{[\![T_n(u), \infty[\![} = A(u) - 1_{K(u)} \cdot A(u) \\ &T(u,t) = \inf(s > t \colon \Delta A(u)_s = -1). \end{aligned}$$

All times $T_n(u)$ and $T(u,t)$ are predictable. We have $T(u,0) = T_1(u)$. We have $\Delta A'(u) \neq -1$ everywhere by construction, hence $\mathscr{E}[A'(u)]$ never vanishes. Also I.4.63 yields that $\mathscr{E}[A(u) - A(u)^t]$ is equal to 1 on $[\![0, t]\!]$, and does not vanish on $[\![0, T(u,t)[\![$. Thus all terms in the next theorem make sense.

2.49 **Theorem.** *With the above notation, there is equivalence between:*
a) *X is a semimartingale with characteristics (B, C, ν).*
b) *For all $u \in \mathbb{R}^d$, $n \in \mathbb{N}^*$,*

2.50 $$e^{iu \cdot X'(u)}/\mathscr{E}[A'(u)] \text{ is a local martingale}$$

2.51 $$E(\exp iu \cdot \Delta X_{T_n(u)} | \mathscr{F}_{T_n(u)-}) = 0 \text{ on } \{T_n(u) < \infty\}.$$

c) *For all $u \in \mathbb{R}^d$ and all t belonging to a dense subset D of $\mathbb{R}_+$ containing 0,*

2.52 $$\{e^{iu \cdot (X - X^t)}/\mathscr{E}[A(u) - A(u)^t]\} 1_{[\![0, T(u,t)[\![} \text{ is a local martingale on } [\![0, T(u,t)[\![$$

2.53 $$E(\exp iu\cdot \Delta X_{T(u,t)}|\mathscr{F}_{T(u,t)-}) = 0 \text{ on } \{T(u,t) < \infty\}.$$

Furthermore, those conditions imply that 2.52 and 2.53 hold for all $t \in \mathbb{R}_+$.

The *a*-priori surprising conditions 2.51 and 2.53 come from the following lemma, of independent interest.

2.54 **Lemma.** *If X is a semimartingale with characteristics (B, C, ν) and if T is a predictable time, then*

$$E(\exp iu\cdot \Delta X_T|\mathscr{F}_{T-}) = 1 + \Delta A(u)_T \quad \text{on } \{T < \infty\}.$$

Proof. 1.18 applied to $W(\omega, t, x) = e^{iu\cdot x} - 1$ yields

$$E(e^{iu\cdot \Delta X_T} - 1|\mathscr{F}_{T-}) = \int (e^{iu\cdot x} - 1)\nu(\{T\} \times dx)$$

on $\{T < \infty\}$, and 2.14 implies that the right-hand side above is $\Delta A(u)_T$. □

3. Some Examples

The archetype of semimartingales—the processes with independent increments—are studied in the next section. Other, somewhat richer, examples (like diffusion processes, etc...) will be studied in Chapter III. However, it seems suitable to give some "elementary" (and some not so elementary) examples as soon as possible.

The first example (§ 3a) concerns the translation of Section 2 into the discrete-time setting: it is very simple, and much in the spirit of the "discrete case" subsections of Chapter I. § 3b below also concerns the discrete case; it is much more important and difficult. Then we deal with the "one-point" point process, mainly to compute the characteristics of the empirical process associated to a sequence of i.i.d. variables.

§ 3a. The Discrete Case

Here we start with a discrete-time basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n\in\mathbb{N}}, P)$. We have seen in § I.4g that a semimartingale is just an adapted process.

So we start with a d-dimensional adapted process $X = (X^i)_{i\leq d}$. The process C and the continuous part $B^c = B - \sum_{s\leq \cdot} \Delta B_s$ of B are "continuous", which in the discrete case means independent of time; since $B_0 = 0$ and $C_0 = 0$, we obtain here that $B^c = 0$ and $C = 0$. On the other hand, 2.14 shows that $B^d = B - B^c = B$ is entirely described by the measure ν. Therefore, we are led to put:

3.1 **Definition.** The *discrete characteristic* of the semimartingale X is the predictable random measure ν on $\mathbb{N} \times \mathbb{R}^d$ defined by:

(i) $\nu(\{0\} \times \cdot) = 0$,

(ii) for $n \geq 1$, $\nu(\cdot; \{n\} \times dx)$ is a version of the regular conditional distribution of the variable $\Delta X_n = X_n - X_{n-1}$ with respect to $\mathscr{F}_{n-1}$. □

Note that 2.13 is replaced by: each $\nu(\{n\} \times \cdot)$, for $n \in \mathbb{N}^*$, is a probability measure.

The equivalent of the process $A(u)$ defined by 2.40 is here:

3.2 $$A(u)_0 = 0, \quad A(u)_n = \sum_{1 \leq p \leq n} \int (e^{iu \cdot x} - 1)\nu(\{p\} \times dx).$$

Theorem 2.42 is replaced by the following trivial statement:

3.3 **Theorem.** *For a predictable random measure ν on $\mathbb{N} \times \mathbb{R}^d$ to be the discrete characteristic of X, it is necessary and sufficient that 3.1(i) holds and that for all $u \in \mathbb{R}^d$ the process $M(u) = e^{iu \cdot X} - e^{iu \cdot X_-} \cdot A(u)$ is a martingale.* Here $A(u)$ is defined by 3.2 and the notation "$\cdot$" is defined by I.3.37, that is we have:

$$M(u)_n = e^{iu \cdot X_0} + \sum_{1 \leq p \leq n} [e^{iu \cdot X_p} - e^{iu \cdot X_{p-1}} - e^{iu \cdot X_{p-1}}(A(u)_p - A(u)_{p-1})].$$

Now we translate part (a) of Theorem 2.47. We suppose that ν is the discrete characteristic of X, and we put:

$$\begin{cases} G(u)_0 = 1 \\ G(u)_n = \prod_{1 \leq p \leq n} \int e^{iu \cdot x}\nu(\{p\} \times dx) \quad \text{for } n \geq 1 \\ R_p = \inf\left(n: |G(u)_n| \leq \frac{1}{p}\right) \end{cases}$$

By I.2.37, each R_p is a predictable time, and I.2.39 implies that $R'_p = (R_p - 1)^+$ is a stopping time. Then, we easily obtain:

3.4 **Theorem.** *With the above assumptions, for each $p \in \mathbb{N}^*$ the process* $([\exp iu \cdot X_{n \wedge R'_p}]/G(u)_{n \wedge R'_p})_{n \in \mathbb{N}}$ *is a martingale.*

We let it to the reader, to formalize part (b) of Theorem 2.47 as well as Theorem 2.49 in the discrete case.

Now, more important than what precedes is how to reduce the discrete case to the continuous one. Instead of associating to $\mathscr{B}$ and X the "usual" continuous time basis $\mathscr{B}'$ by I.1.55 and the process X' by I.1.59, we shall introduce a more general embedding in the next paragraph.

§ 3b. More on the Discrete Case

Let us start again with a discrete basis $\mathscr{B} = (\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\in\mathbb{N}}, P)$, and consider an adapted process $X = (X_n)_{n\in\mathbb{N}}$ on it, with the increments $U_n = \Delta X_n = X_n - X_{n-1}$ (and $U_0 = X_0$).

In limit theorems, it is quite often that one needs "normalizing" the process X in time. That is, one considers the continuous-time process Y defined by

$$Y_t = \sum_{0\le k\le \sigma_t} U_k = X_{\sigma_t}$$

where σ_t is some suitable increasing process taking values in $\mathbb{N}$ or in $\bar{\mathbb{N}}$ (in the latter case, there of course may be a convergence problem in the above formula). Naturally σ_t must satisfy some nice properties, summarized in:

3.5 **Definition.** A *change of time* on the basis $\mathscr{B}$ is a family $\sigma = (\sigma_t)_{t\ge 0}$ such that:
(i) each σ_t is a stopping time on $\mathscr{B}$;
(ii) $\sigma_0 = 0$;
(iii) each path: $t \rightsquigarrow \sigma_t$ is increasing, càd, with jumps equal to 1. □

We associate to it the following continuous-time basis:

3.6 $$\tilde{\mathscr{B}} = (\Omega, \mathscr{F}, \mathbf{G} = (\mathscr{G}_t)_{t\ge 0}, P), \quad \text{with } \mathscr{G}_t = \mathscr{F}_{\sigma_t}.$$

Due to 3.5(iii) and I.1.18, $\mathbf{G}$ is a filtration in the sense of I.1.2. Note that σ is an adapted point process on $\tilde{\mathscr{B}}$, in the sense of § I.3b except that it may take the value $+\infty$. Note also that if $\sigma(t) = [t]$, the integer part of t, we are back to the basis $\mathscr{B}'$ defined by I.1.55.

We set

3.7 $$\tau_k = \inf(t\colon \sigma_t \ge k), \quad k\in\mathbb{N},$$

and we begin with a technical lemma.

3.8 **Lemma.** a) $\tau_0 = 0$ *and* $\mathscr{G}_0 = \mathscr{F}_0$.
b) *Each* τ_k *is a predictable time on* $\tilde{\mathscr{B}}$.
c) *On the set* $\{\tau_k < \infty\}$, *the traces of the* σ*-fields* $\mathscr{F}_k$ *and* $\mathscr{G}_{\tau_k}$ *are equal, and the traces of the* σ*-fields* $\mathscr{F}_{k-1}$ *and* $\mathscr{G}_{(\tau_k)-}$ *are also equal.*

Proof. (a) is trivial. Let $k \ge 1$. If $A \in \mathscr{F}_k$ we have by 3.7:

$$[A \cap \{\tau_k \le t\}] \cap \{\sigma_t = m\} = \begin{cases} \varnothing & \text{if } m < k \\ A \cap \{\sigma_t = m\} & \text{if } m \ge k, \end{cases}$$

which in both cases belongs to $\mathscr{F}_m$. Hence $A \cap \{\tau_k \le t\} \in \mathscr{G}_t$ by I.1.52 and the definition of $\mathscr{G}_t$. We deduce that τ_k is a stopping time on $\tilde{\mathscr{B}}$ (take $A = \Omega$), and that $\mathscr{F}_k \subset \mathscr{G}_{\tau_k}$.

Next, let $A \in \mathscr{G}_{\tau_k}$. Then $A_t := A \cap \{\tau_k \le t < \tau_{k+1}\}$ is in $\mathscr{G}_t$, hence $A_t \cap \{\sigma_t = k\} = A_t$ implies that $A_t \in \mathscr{F}_k$. Since $A \cap \{\tau_k < \infty\} = \bigcup_{t \in \mathbb{Q}_+} A_t$ it follows that $A \cap \{\tau_k < \infty\} \in \mathscr{F}_k$. Therefore the traces of $\mathscr{F}_k$ and of $\mathscr{G}_{\tau_k}$ on $\{\tau_k < \infty\}$ coincide.

We have $\{\tau_k > t\} = \{\sigma_t \le k-1\} \in \mathscr{F}_{k-1} \subset \mathscr{G}_{\tau_{k-1}}$. Hence τ_k is $\mathscr{G}_{\tau_{k-1}}$-measurable and $\tau_k > \tau_{k-1}$ if $\tau_{k-1} < \infty$. We easily deduce that τ_k is a predictable time, announced by the G-stopping times $[\tau_{k-1} + (\tau_k - \tau_{k-1} - 1/n)^+] \wedge n$. If $A \in \mathscr{F}_{k-1}$ we have $A \cap \{\tau_k < \infty\} \in \mathscr{F}_{k-1} \subset \mathscr{G}_{\tau_{k-1}}$, hence $A \cap \{\tau_k < \infty\} \in \mathscr{G}_{(\tau_k)-}$ by I.1.17. Finally if $A \in \mathscr{G}_t$, then $A \cap \{t < \tau_k\} = A \cap \{\sigma_t \le k-1\} \in \mathscr{F}_{k-1}$. This and (a) and the definition of $\mathscr{G}_{(\tau_k)-}$ imply that $\mathscr{F}_k \subset \mathscr{G}_{(\tau_k)-}$, and this finishes the proof. □

Now we come back to the adapted "increment process" $(U_n)_{n \ge 0}$. If $\sigma_t < \infty$ for all $t \in \mathbb{R}_+$ we set

3.9 $$Y_t = \sum_{1 \le k \le \sigma_t} U_k = \sum_{k \ge 1} U_k 1_{\{\tau_k \le t\}}.$$

This is clearly an adapted process on $\tilde{\mathscr{B}}$ with finite variation. Now if σ_t takes on values in $\bar{\mathbb{N}}$ we put

3.10 $$\begin{cases} Y_t^n = \sum_{1 \le k \le \sigma_t \wedge n} U_k = \sum_{1 \le k \le n} U_k 1_{\{\tau_k \le t\}} \\ Y_t = P\text{-}\lim_{(n)} Y_t^n \quad \text{(limit in measure)}, \end{cases}$$

if the latter exists.

3.11 **Theorem.** *Let $h \in \mathscr{C}_t^d$ be any truncation function.*

a) *Formula 3.10 defines a semimartingale Y on $\tilde{\mathscr{B}}$ if and only if for all $t \in \mathbb{R}_+$,*

3.12 $$\sum_{1 \le k \le \sigma_t} |E(h(U_k)|\mathscr{F}_{k-1})| < \infty \quad \text{a.s.}$$

3.13 $$\sum_{1 \le k \le \sigma_t} E(|U_k|^2 \wedge 1|\mathscr{F}_{k-1}) < \infty \quad \text{a.s.}$$

(These conditions, as well as the semimartingale property, are obviously true when $\sigma_t < \infty$ a.s. for all $t \in \mathbb{R}_+$).

b) *In this case, the characteristics (B, C, ν) of Y relative to h are:*

3.14 $$\begin{cases} B_t = \sum_{1 \le k \le \sigma_t} E(h(U_k)|\mathscr{F}_{k-1}) \\ C_t = 0 \\ \nu([0,t] \times g) = \sum_{1 \le k \le \sigma_t} E(g(U_k) 1_{\{U_k \ne 0\}}|\mathscr{F}_{k-1}), \text{ for } g \ge 0 \text{ Borel.} \end{cases}$$

Note that this last formula completely determines the measure ν; it would even suffices to verify this formula for much smaller classes of functions g, for instance the class of C^∞ functions on $\mathbb{R}^d$, bounded, null in a neigbourhood of 0. On the opposite, for each nonnegative $\tilde{\mathscr{P}}$-measurable W on $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}^d$ we have by a monotone class argument:

3.15 $$W * \nu_t = \sum_{1 \leq k \leq \sigma_t} E(W(\tau_k, U_k) 1_{\{U_k \neq 0\}} | \mathscr{F}_{k-1}):$$

This formula also says that, on the set $\{\tau_k < \infty\}$, the measure

3.16 $$\nu(\omega; \{\tau_k(\omega)\} \times dx) + [1 - \nu(\omega; \{\tau_k(\omega)\} \times \mathbb{R}^d)] \varepsilon_0(dx)$$

is a regular version of the conditional distribution of U_k with respect to $\mathscr{F}_{k-1}$.

Proof. Suppose first that Y is a semimartingale, with characteristics (B, C, ν). Let $J = \bigcup_{k \geq 1} [\![\tau_k]\!]$, which is a thin predictable random set on $\tilde{\mathscr{B}}$. From 2.6 it is obvious that the two first characteristics of the semimartingale $1_J \cdot Y$ are $1_J \cdot B$ and $1_J \cdot C$; now, $Y = 1_J \cdot Y$ by construction, hence $B = 1_J \cdot B$ and $C = 1_J \cdot C$. The latter obviously yields $C = 0$, and the former shows that the first formula in 3.14 follows from the last one and from 2.14.

In order to show the last formula in 3.14, consider a nonnegative Borel function g such that $g * \mu^Y$ is in $\mathscr{A}^+_{\text{loc}}$. We have

$$g * \mu^Y_t = \sum_{k \geq 1} g(U_k) 1_{\{U_k \neq 0\}} 1_{\{\tau_k \leq t\}}$$

by 3.10, while 3.8 implies that the increasing process

$$A^g_t = \sum_{k \geq 1} E(g(U_k) 1_{\{U_k \neq 0\}} | \mathscr{F}_{k-1}) 1_{\{\tau_k \leq t\}}$$

is predictable on $\tilde{\mathscr{B}}$. Now for each stopping time T we obtain from I.2.11:

$$\begin{aligned} E(A^g_T) &= \sum_{k \geq 1} E(1_{\{\tau_k \leq T\}} E(g(U_k) 1_{\{U_k \neq 0\}} | \mathscr{F}_{k-1})) \\ &= \sum_{k \geq 1} E(1_{\{\tau_k \leq T\}} g(U_k) 1_{\{U_k \neq 0\}}) = E(g * \mu^Y_T). \end{aligned}$$

Then Theorem I.3.17 gives that $A^g_t = \nu([0, t] \times g)$ a.s.

Finally, 3.12 follows from 3.14, and 3.13 follows from 2.13 and 3.14 again.

Conversely, assume that conditions 3.12 and 3.13 are satisfied. In order to prove sufficiency in (a), it is enough to prove that 3.10 defines a semimartingale, separately for each of the following sequences:

$$U'_k = E(h(U_k) | \mathscr{F}_{k-1}), \quad U''_k = U_k - h(U_k),$$
$$V_k = h(U_k) - E(h(U_k) | \mathscr{F}_{k-1}).$$

This is trivial for (U'_k) (use 3.12). For the other sequences, we first notice that the left-hand side of 3.13 defines a predictable increasing process, so by localization (a procedure that does not affect 3.10) we may assume that

$$E\left(\sum_{k \geq 1} |U_k|^2 \wedge 1\right) = E\left(\sum_{k \geq 1} E(|U_k|^2 \wedge 1 | \mathscr{F}_{k-1})\right) < \infty.$$

Therefore $\sum_{k \geq 1} 1_{\{|U_k| > a\}} < \infty$ a.s. for each $a > 0$ and since $U''_k = 0$ whenever $|U_k| \geq a$, for some a small enough, it follows that 3.10 defines a semimartingale for the sequence (U_k''). It also follows that

$$E\left(\sum_{k\geq 1} |V_k|^2\right) \leq E\left(\sum_{k\geq 1} |h(U_k)|^2\right) < \infty$$

because $|h(x)|^2 \leq c(|x|^2 \wedge 1)$ for some constant c. Due to 3.8, V_k is $\mathscr{F}_{\tau_k}$-measurable and $E(V_k|\mathscr{F}_{\tau_k-}) = 0$. Hence part (a) of the proof of I.4.51 yields that Y^n (defined by 3.10 with (V_k)) are martingales on $\tilde{\mathscr{B}}$, converging in $\mathscr{H}^2$ to a square-integrable martingale Y. So 3.10 for the sequence (V_k) defines a martingale on $\tilde{\mathscr{B}}$, and the sufficient condition in (a) is thus proved. □

3.17 **Remark.** At this stage, it is worthwhile to compare 3.11 with Kolmogorov's three series Theorem. Assume that the U_k's are 1-dimensional and independent, so in 3.12 and 3.13 we have expectations instead of conditional expectations. Assume also that $\sigma_1 \equiv \infty$, for instance, and take $h(x) = x1_{\{|x|\leq 1\}}$. Then 3.12 and 3.13 are clearly equivalent to:

$$(1) \qquad \begin{cases} \sum_k |E(U_k 1_{\{|U_k|\leq 1\}})| < \infty \\ \sum_k E(U_k^2 1_{\{|U_k|\leq 1\}}) < \infty \\ \sum_k P(|U_k| > 1) < \infty. \end{cases}$$

Now, the three series theorem says that the series $\sum U_k$ converges a. s. if and only if

$$(2) \qquad \begin{cases} \sum_k E(U_k 1_{\{|U_k|\leq 1\}}) \quad \text{converges} \\ \sum_k [E(U_k^2 1_{\{|U_k|\leq 1\}}) - E(U_k 1_{\{|U_k|\leq 1\}})^2] < \infty \\ \sum_k P(|U_k| > 1) < \infty. \end{cases}$$

(1) is of course more stringent than (2). In other words, the process Y can be well-defined by 3.10 for all t, and nevertheless not be a semimartingale! □

Assume now that 3.12 and 3.13 are met. Due to 3.14 and 2.18, the modified second characteristic of Y is

$$3.18 \qquad \tilde{C}_t^{ij} = \sum_{1\leq k\leq \sigma_t} [E(h^i h^j(U_k)|\mathscr{F}_{k-1}) - E(h^i(U_k)|\mathscr{F}_{k-1})E(h^j(U_k)|\mathscr{F}_{k-1})]$$

while the processes $A(u)$ and $G(u)$ associated to Y by 2.40 and 2.47 are

$$3.19 \qquad \begin{cases} A(u)_t = \sum_{1\leq k\leq \sigma_t} E(e^{iu\cdot U_k} - 1|\mathscr{F}_{k-1}) \\ G(u)_t = \prod_{1\leq k\leq \sigma_t} E(e^{iu\cdot U_k}|\mathscr{F}_{k-1}). \end{cases}$$

Example: *normalized i.i.d. random variables*. Here we consider a sequence $(Z_k)_{k\geq 1}$ of i.i.d. 1-dimensional random variables. We suppose that $\sigma_t = [nt]$, and we set

$$Y_t^n = a_n \sum_{k\leq [nt]} Z_k.$$

This corresponds to 3.10 with $U_k = a_n Z_k$. Of course 3.12 and 3.13 are met, and the characteristics of Y^n are

3.20
$$\begin{cases} B_t^n = [nt] E(h(a_n Z)) \\ C_t^n = 0, \quad \tilde{C}_t^{n,ij} = [nt]\{E(h^i h^j(a_n Z)) - E(h^i(a_n Z))E(h^j(a_n Z))\} \\ \nu^n([0,t]\times g) = [nt] E(g(a_n Z)). \end{cases}$$

In particular, if the Z_k's are centered with finite variance and $a_n = 1\sqrt{n}$, then Y^n is a locally square-integrable martingale, with

3.21
$$\langle Y^n, Y^n\rangle_t = \frac{[nt]}{n} E(Z^2)$$

(apply 2.29b). Of course, a direct proof of these properties is available and is much simpler than using 3.11!

§ 3c The "One-Point" Point Process and Empirical Processes

1. We begin with an elementary remark. Suppose that $X = N$ is a point process (in the sense of § I.3b) on some stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with A its compensator.

Then X is obviously a semimartingale. Moreover, its characteristics are easily computed through A. Namely, let $h \in \mathscr{C}_t^1$ be a truncation function. Then 2.4 yields $X(h) = h(1)X$ because ΔX is either 0 or 1; similarly, for any function g on $\mathbb{R}$, $g * \mu^X = g(1)X$. Then the characteristics $(B(h), C, \nu)$ of X are:

3.22
$$\begin{cases} B(h) = h(1)A \\ C = 0 \\ \nu(dt,dx) = dA_t \otimes \varepsilon_1(dx) \quad (\varepsilon_1 = \text{Dirac measure at point 1}). \end{cases}$$

Of course it is wise to choose h so that $h(1) = 0$, in which case $B(h) = 0$.

2. Now we consider the one-point process:

3.23
$$N = 1_{[\![T,\infty[\![}$$

where T is an $(0,\infty]$-valued random variable on a probability space $(\Omega, \mathscr{F}, P)$.

3.24 **Lemma.** *The filtration* $\mathbf{F} = (\mathscr{F}_t)_{t\in\mathbb{R}_+}$ *generated by* N (i.e. the smallest filtration to which N is adapted, or equivalently the smallest filtration for which T is a stopping time) *is defined as such: for all* t, $\mathscr{F}_t$ *is the class of all sets of the form* $A = \{T \in B\}$, *where* B *is a Borel subset of* $(0, \infty]$ *such that either* $(t, \infty] \subset B$ *or* $(t, \infty] \cap B = \emptyset$.

Proof. Define $\mathscr{F}_t$ as above. Then $\mathscr{F}_t$ is obviously a σ-field, and $\mathscr{F}_s \subset \mathscr{F}_t$ for $s \le t$.

Next, we show that **F** is right-continuous. Let $A \in \bigcap_{s>t} \mathscr{F}_s$. For $s > t$ there is a Borel set B_s such that $A = \{T \in B_s\}$ and that either $(s, \infty] \subset B_s$ or $(s, \infty] \cap B_s = \emptyset$. Put $B = \liminf_{s \downarrow\downarrow t, s \in \mathbb{Q}} B_s$. Then B is Borel and that $A = \{T \in B\}$ is obvious. Assume that $(t, \infty] \cap B \neq \emptyset$: then $(t, \infty] \cap B$ contains a point $u > t$, and for all rationals s in some interval $(t, v]$ with $v < u$ we must have $u \in B_s$ (by definition of B); then $(s, \infty] \cap B_s \neq \emptyset$ for those s, which implies $(s, \infty] \subset B_s$; applying again the definition of B, we see that $(t, \infty] \subset B$: in other words, $A \in \mathscr{F}_t$, and the claim is proved.

Next, $\{T \le t\} \in \mathscr{F}_t$ by definition of $\mathscr{F}_t$, so T is a stopping time. Finally, let $\mathbf{G} = (\mathscr{G}_t)_{t \in \mathbb{R}_+}$ be any other filtration for which T is a stopping time. Let $A = \{T \in B\} \in \mathscr{F}_t$. If $(t, \infty] \cap B = \emptyset$, then $A = \{T \in B\} \cap \{T \le t\}$ belongs to $\mathscr{G}_t$ because T is a **G**-stopping time; if $(t, \infty] \subset B$, then $A = [\{T \in B\} \cap \{T \le t\}] \cup \{T > t\}$ also belongs to $\mathscr{G}_t$: hence $\mathscr{F}_t \subset \mathscr{G}_t$. $\square$

Now we proceed to computing the compensator of N. We denote by F the distribution of the random variable T: this is a probability measure on $(0, \infty]$. We also define an increasing càdlàg function α by:

3.25 $$\alpha(t) = \int_{(0,t]} \frac{F(ds)}{F([s, \infty])}$$

If $\rho = \sup(s: F([s, \infty]) > 0)$, then $\alpha(t) < \infty$ for $t < \rho$, and $\alpha(t) = \alpha(t \wedge \rho)$, and $\alpha(\rho) < \infty$ if $F(\{\rho\}) > 0$, and $\alpha(\rho-) = \infty$ if $F(\{\rho\}) = 0$.

3.26 **Theorem.** *With the previous notation, and if* **F** *is the filtration generated by* N (see 3.24), *the compensator of* N *is*

3.27 $$A_t = \alpha(t \wedge T).$$

Proof. A is increasing, predictable, càdlàg, with $A_0 = 0$. It is also a. s. finite-valued, because $T \le \rho$ a. s., and $T < \rho$ a. s. if $F(\{\rho\}) = 0$. It remains to prove that $N - A$ is a martingale, and for this it is enough to prove that for all $s \le t$, $B \in \mathscr{F}_s$,

(1) $$E(1_B(N_t - N_s - A_t + A_s)) = 0.$$

If $T \le s$, then $N_t - N_s - A_t + A_s = 0$, so the left-hand side of (1) equals

(2) $$E(1_B 1_{\{T>s\}}(N_t - N_s - A_t + A_s)).$$

Then 3.24 implies that $B \cap \{T > s\}$ is either $\emptyset$ or $\{T > s\}$. In the former case, (2) is obviously 0; in the latter case, (2) equals

$$\begin{aligned}
E(1_{\{T>s\}}(N_t - N_s - A_t + A_s)) &= E(1_{\{s<T\le t\}} - 1_{\{T>s\}}(\alpha(t \wedge T) - \alpha(s))) \\
&= F((s,t]) - \int_{(s,t]} F(du) \int_{(s,u]} \frac{F(dv)}{F([v,\infty])} \\
&\quad - \int_{(t,\infty]} F(du) \int_{(s,t]} \frac{F(dv)}{F([v,\infty])} \\
&= F((s,t]) - \int_{(s,t]} \frac{F(dv)}{F([v,\infty])} \int_{[v,t]} F(du) \\
&\quad - \int_{(s,t]} \frac{F(dv)}{F([v,\infty])} \int_{(t,\infty]} F(du) \\
&= F((s,t] - \int_{(s,t]} \frac{F(dv)}{F([v,\infty])} F([v,\infty]) = 0,
\end{aligned}$$

and we are done. □

3.28 **Example.** If T is exponential with parameter 1, then $A_t = t \wedge T$.

3.29 **Example.** If T is uniform on $(0,1]$, than $A_t = -\mathrm{Log}(1 - t \wedge T)$.

3. *Empirical process.* A very important class of limit theorems concerns empirical processes, with various normalizations. To prepare for that, we compute the characteristics of some of these.

We start with a sequence $(Z_n)_{n\ge 1}$ of i.i.d. random variables, with common distribution G. We assume that *G is supported by* $(0,\infty)$ *and admits the density g* (the second assumption is not essential but it allows to simplify the notation; what is essential, though, is that G has no atom).

In the sequel, we consider a number of different processes. To avoid lengthy repetitions, we introduce the following notation:

3.30 If X is a càdlàg process, $\mathbf{F}^X$ denotes the smallest filtration to which X is adapted. □

Firstly, *the empirical process of size n* is the process

3.31 $$X_t^n = \frac{1}{n} \sum_{1\le i\le n} 1_{\{Z_i \le t\}}.$$

The process nX^n is a. s. a point process in the sense of § I.3b, because $P(Z_i = Z_j) = 0$ for all $i \ne j$, and:

3.32 **Proposition.** *The compensator of the point process $Y^n = nX^n$, relative to the filtration $\mathbf{F}^{Y^n}$, is given by:*

3.33 $$A_t^n = \int_0^t (n - Y_s^n) \frac{g(s)}{G([s,\infty])} ds$$

Proof. Set $N^i = 1_{[\![Z_i,\infty[\![}$. Since $1_{\{s \le t \wedge Z_i\}} = (1 - N^i_{s-})1_{\{s \le t\}}$, Theorem 3.26 yields that the compensator $\hat{A}^i$ of the process N^i, relative to the filtration $\mathbf{F}^{N^i}$, is

(1) $$\hat{A}^i_t = \int_0^t (1 - N^i_s)\frac{g(s)}{G([s,\infty])}ds.$$

Let $\mathbf{F}$ be the smallest filtration that contains all $\mathbf{F}^{N^i}$ for $i \le n$: it is obvious that $\mathscr{F}_t = \bigvee_{1 \le i \le n} \mathscr{F}^{N^i}_t$. Moreover for all t the σ-fields $\mathscr{F}^{N^i}_\infty$ are mutually independent (because the Z_i's are so). Therefore if $s \le t$ and $B_i \in \mathscr{F}^{N^i}_s$ and $B = \bigcap_{1 \le j \le n} B_j$, then

(2) $$\begin{aligned} &E(1_B(N^i_t - N^i_s - \hat{A}^i_t + \hat{A}^i_s)) \\ &= \left[\prod_{j \le n, j \ne i} P(B_j)\right] E(1_{B_i}(N^i_t - N^i_s - \hat{A}^i_t + \hat{A}^i_s)) = 0. \end{aligned}$$

A monotone class argument yields that the left-hand side of (2) is 0 for all $B \in \mathscr{F}_s$, and we deduce that $N^i - \hat{A}^i$ is an $\mathbf{F}$-martingale.

Now $Y^n = \sum_{i \le n} N^i$, and 3.33 and (1) show that $A^n = \sum_{i \le n} \hat{A}^i$. Thus $Y^n - A^n$ is an $\mathbf{F}$-martingale. Moreover A^n is clearly predictable for $\mathbf{F}^{Y^n}$, while $Y^n - A^n$ is adapted to $\mathbf{F}^{Y^n}$: then $Y^n - A^n$ is an $\mathbf{F}^{Y^n}$-martingale, and this proves the result. □

3.34 **Corollary.** *The compensator of the point process $\bar{Y}^n_t = nX^n_{t/n}$, relative to the filtration $\mathbf{F}^{\bar{Y}^n}$, given by*:

$$\bar{A}^n_t = \int_0^t \left(1 - \frac{\bar{Y}^n_s}{n}\right)\frac{g\left(\frac{s}{n}\right)}{G\left(\left[\frac{s}{n},\infty\right]\right)}ds$$

Proof. With the notation of 3.32 we have $\mathscr{F}^{\bar{Y}^n}_t = \mathscr{F}^{Y^n}_{t/n}$, so $A^n_{t/n}$ is the $\mathbf{F}^{\bar{Y}^n}$-compensator of $Y^n_{t/n}$. A change of variables in 3.33 shows that $A^n_{t/n} = \bar{A}^n_t$. □

As we shall see later, the process $\bar{Y}^n$ of 3.34 converges to a Poisson process as n $\uparrow \infty$. Another well-known normalization leads to convergence to a Brownian bridge, and to prepare for this we consider the process

3.35 $$V^n_t = \sqrt{n}(X^n_t - G((0,t])).$$

3.36 **Proposition.** *The characteristic (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$ of the process V^n, relative to the filtration $\mathbf{F}^{V^n}$ and to any truncation function $h \in \mathscr{C}^1_t$ such that $h(x) = x$ for $|x| \le 1$, are given by*:

3.37 $$\begin{cases} B^n_t = -\displaystyle\int_0^t V^n_s \frac{g(s)}{G([s,\infty])}ds \\ C^n = 0, \quad \tilde{C}^n_t = \displaystyle\int_0^t \left[1 - \frac{V^n_s}{\sqrt{n}G([s,\infty])}\right]g(s)\,ds \\ \nu^n(\omega; dt, dx) = \left[n - \sqrt{n}\dfrac{V^n_t(\omega)}{G([t,\infty])}\right]g(t)\,dt \otimes \varepsilon_{1/\sqrt{n}}(dx). \end{cases}$$

Proof. First notice that, with the notation of 3.32,

$$(1)\qquad V_t^n = \sqrt{n}\left(\frac{Y_t^n}{n} - G((0,t])\right), \quad Y_t^n = \sqrt{n}\,V_t^n + nG((0,t]).$$

Hence $\mathbf{F}^{V^n} = \mathbf{F}^{Y^n}$. Secondly, V^n has finite variation, so $C^n = 0$.

Next, $|\Delta V^n| \leq 1/\sqrt{n}$ (recall that $t \rightsquigarrow G((0,t])$ is continuous). Then with the notation 2.4 we have $V^n(h) = V^n$ because $h(x) = x$ for $|x| \leq 1$. Moreover 3.32 yields that $Y^n - A^n$ is a martingale, so by (1), $V_t^n - A_t^n/\sqrt{n} + \sqrt{n}G((0,t])$ is also a martingale. We deduce that

$$(2)\qquad B_t^n = A_t^n/\sqrt{n} - \sqrt{n}G((0,t])$$

is a version of the first characteristic of V^n. A simple computation, using (1), shows that (2) is exactly the first process in 3.37.

Let f be any bounded Borel function on $\mathbb{R}$ with $f(0) = 0$. Since the jumps of Y^n have size 1 (except perhaps on a null set), we deduce from (1) that a.s.,

$$f * \mu^{V^n} = \sum_{\sigma \leq \cdot} f\left(\frac{\Delta Y_s^n}{\sqrt{n}}\right) = \sum_{\sigma \leq \cdot} f\left(\frac{1}{\sqrt{n}}\right) 1_{\{\Delta Y_s^n = 1\}} = f\left(\frac{1}{\sqrt{n}}\right) Y^n.$$

So the compensator of the process $f * \nu^{V^n}$ is $f(1/\sqrt{n})A^n$. Now, if ν^n is defined by 3.37, using (1) once more shows that $f(1/\sqrt{n})A^n = f * \nu^n$, and therefore ν^n is a version of the third characteristic of V^n.

Finally, 2.18 gives $\tilde{C}^n$, as stated in 3.37. □

4. Semimartingales with Independent Increments

We devote the next two sections to studying the fundamental example of processes with independent increments.

4.1 **Definitions.** a) A *process with independent increments* (in short: PII) *on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (or, *relative to the filtration* $\mathbf{F}$) is a càdlàg adapted $\mathbb{R}^d$-valued process X such that $X_0 = 0$ and that for all $0 \leq s \leq t$ the variable $X_t - X_s$ is independent from the σ-field $\mathscr{F}_s$.

b) A *process with stationary independent increments* (in short: PIIS) *on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is a PII such that the distribution of the variables $X_t - X_s$ only depends on the difference $t - s$.

c) A time $t \in \mathbb{R}_+$ is called a *fixed time of discontinuity* for X if $P(\Delta X_t \neq 0) > 0$. □

An extended Poisson process and a Wiener process (see I.3.26 and I.4.9) are PII. A standard Poisson process and a standard Wiener process are PIIS.

Note that if (T_n) is a sequence of stopping times that exhausts the jumps of X, the set D_n of all t's for which $P(T_n = t) > 0$ is at most countable, and the set of fixed times of discontinuity is $\bigcup D_n$. Then:

4.2 The set of fixed times of discontinuity is at most countable (a property which is true for all càdlàg processes). $\square$

Now if X is a PIIS, the distribution of the variables $\Delta X_t = \lim_{s\uparrow\uparrow t}(X_t - X_s)$ does not depend on t. Then it follows from 4.2 that:

4.3 A PIIS has no fixed time of discontinuity. $\square$

Before going through the general case of PII's, we consider separately two examples: Wiener processes and Poisson processes, because they are at the same time much simpler to examine than general PII's, and much more commonly encountered. The first two subsections below may thus serve as an introduction for § 4c. On the other hand, their results are corollaries of the results of § 4c, hence a reader interested only in the general case may as well skip § 4a and § 4b.

§ 4a. Wiener Processes

The next theorem, due to Lévy, complements Proposition I.4.10.

4.4 **Theorem.** *A continuous local martingale W with $W_0 = 0$ is a Wiener process if and only if its angle bracket $\langle W, W\rangle$ is deterministic, say $\langle W, W\rangle_t = \sigma^2(t)$ for some increasing continuous function $\sigma^2(\cdot)$. Then this function is the variance function of W, and for all $0 \le s \le t$ the variable $W_t - W_s$ is Gaussian, centered, with variance $\sigma^2(t) - \sigma^2(s)$.*

Proof. The necessary condition is exactly Proposition I.4.10. Conversely, assume that W is a continuous local martingale with $W_0 = 0$ and $\langle W, W\rangle_t = \sigma^2(t)$. Put $Y_t = \exp\left(iuW_t + \frac{u^2}{2}\sigma^2(t)\right)$. Ito's formula applied to the function $f(x, y) = \exp(iux - y)$ yields:

$$Y = 1 + iuY_- \cdot W,$$

hence Y is a local martingale, and even a (complex-valued) uniformly integrable martingale because $\sup_{s\le t}|Y_s| \le \exp(u^2\sigma^2(t)/2)$. Then if $s \le t$ and $A \in \mathscr{F}_s$, we have

$$\begin{aligned} E(1_A \exp iu(W_t - W_s)) &= E\left[1_A \frac{Y_t}{Y_s} \exp -\frac{1}{2}u^2(\sigma^2(t) - \sigma^2(s))\right] \\ &= P(A) \exp -\frac{1}{2}u^2(\sigma^2(t) - \sigma^2(s)), \end{aligned}$$

which shows at the same time condition I.4.9ii and that $W_t - W_s$ is Gaussian, centered, with variance $\sigma^2(t) - \sigma^2(s)$. Then condition I.4.9i is trivially satisfied, and the proof is finished. □

§ 4b. Poisson Processes and Poisson Random Measures

We first consider Poisson processes: see Definition I.3.26. If $a\colon \mathbb{R}_+ \to \mathbb{R}$ is a càdlàg increasing function, we denote by a^c its *continuous part*:

$$a^c(t) = a(t) - \sum_{0<s\leq t} \Delta a(s).$$

4.5 **Theorem.** a) *A point process N is an extended Poisson process on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ if and only if its compensator N^p is deterministic, say $N_t^p = a(t)$ for some increasing function a.*

b) *In this case, a time t is a fixed time of discontinuity if and only if $\Delta a(t) \neq 0$, and for all $u \in \mathbb{R}$, $0 \leq s \leq t$, we have*:

4.6 $$E(e^{iu(N_t - N_s)}) = [\exp(e^{iu} - 1)(a^c(t) - a^c(s))]\left[\prod_{s<r\leq t} (1 + (e^{iu} - 1)\Delta a(r))\right].$$

This result proves all the statement left unproved after Definition I.3.26. It is basically due to S. Watanabe. In particular

—When N is Poisson, or equivalently when a is continuous, then the distribution of $N_t - N_s$ is Poissonian with mean $a(t) - a(s)$.

—Due to I.3.27b, we see that "quasi-left-continuous" and "without fixed times of discontinuity" are synonymous for an extended Poisson process. This property will be true for all PII's.

One may go further in the structure of an extended Poisson process. Let $J = \{t\colon \Delta a(t) > 0\}$ and put

$$N_t^d = \int_0^t 1_J(s)\, dN_s, \quad N^c = N - N^d.$$

Then N^c and N^d are two independent extended Poisson processes (immediate consequence of I.3.26(ii) and of 4.5). Their compensators are respectively $a^c(t)$ and $a(t) - a^c(t)$. Then N^c is actually a Poisson process, and by 4.6 we have

$$N_t^d = \sum_{s\leq t, s\in J} Y_s$$

where $(Y_s)_{s\in J}$ is a sequence of independent random variable taking the two values 0 and 1, and such that $P(Y_s = 1) = \Delta a(s)$.

Proof of Theorem 4.5. The necessary condition in (a) follows from I.3.27. Conversely, assume that $N_t^p = a(t)$ for some increasing function a. We have $E(N_t) = a(t)$, hence I.3.26i.

Let $u \in \mathbb{R}$. An elementary computation shows that

$$e^{iuN} = 1 + (e^{iu} - 1)e^{iuN_-} \cdot N.$$

Then, by definition of $N^p = a$, the process

$$M_t = e^{iuN_t} - 1 - \int_0^t (e^{iu} - 1)e^{iuN_{s-}}\, da(s)$$

is a (complex-valued) local martingale, and even a martingale because M_t is bounded by $2(1 + a(t))$. Let $A \in \mathscr{F}_s$ with $P(A) > 0$. From the above formula, we obtain for $t \geq s$:

$$1_A e^{iu(N_t - N_s)} = 1_A + 1_A e^{-iuN_s}(M_t - M_s) + (e^{iu} - 1) \int_s^t e^{iu(N_{r-} - N_s)}\, da(r).$$

Let $f(t) = E(1_A \exp iu(N_t - N_s))/P(A)$. Taking the expectation above and using Fubini's Theorem and that M is a martingale, yields

$$f(t) = 1 + \int_s^t f(r-)(e^{iu} - 1)\, da(r), \quad t \geq s.$$

Then I.4.61 implies $f(t) = \mathscr{E}(a')_t$, where a' is the function $a'(r) = (e^{iu} - 1)[a(r) - a(r \wedge s)]$. Now, I.4.63 and the definition of a^c show that $f(t)$ is equal to the right-hand side of 4.6, say $g_{s,t}(u)$, for $s \leq t$:

$$E(1_A \exp iu(N_t - N_s)) = P(A) g_{s,t}(u).$$

This formula implies condition I.3.26i and formula 4.6.

Finally, 4.6 yields $E(\exp iu\Delta N_t) = 1 + (e^{iu} - 1)\Delta a(t)$ (let $s \uparrow\uparrow t$ in 4.6), hence t is a fixed time of discontinuity if and only if $\Delta a(t) \neq 0$. □

Now we examine Poisson random measures. This is slightly outside our main subject, and may be omitted without consequences. We use the notation of § 1b; in particular E is an auxiliary Blackwell space.

If m is a positive σ-finite measure on $(\mathbb{R}_+ \times E, \mathscr{R}_+ \otimes \mathscr{E})$, we put

4.7 $$\begin{cases} J = \{t: m(\{t\} \times E) > 0\} \\ m^d(dt, dx) = m(dt, dx) 1_J(t, x), \quad m^c = m - m^d. \end{cases}$$

J is at most countable. If E is reduced to one point, say 1, and if $m(dt, \{1\}) = da(t)$ for some increasing function a, then m^c corresponds to the continuous part a^c.

4.8 **Theorem.** a) *An integer-valued random measure μ is an extended Poisson random measure on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ if and only if its compensator μ^p is deterministic, say $\mu^p(\omega; \cdot) = m(\cdot)$, for some positive σ-finite measure m on $(\mathbb{R}_+ \times E, \mathscr{R}_+ \otimes \mathscr{E})$ satisfying $m(\{t\} \times E) \leq 1$.*

b) *In this case, for each family $(A_j)_{j \leq d}$ of pairwise disjoint measurable sets in $(\mathbb{R}_+ \times E, \mathscr{R}_+ \otimes \mathscr{E})$ such that $m(A_j) < \infty$, we have:*

4.9
$$\begin{cases} E(\exp \sum_{j\le d} iu_j\mu(A_j)) = \left[\exp \sum_{j\le d} (e^{iu_j} - 1)m^c(A_j)\right] \\ \qquad\qquad \times \prod_{s>0} \left[1 + \sum_{j\le d} (e^{iu_j} - 1)m(\{s\} \times E \cap A_j)\right]. \end{cases}$$

If μ is a *Poisson random measure*, that is if $m = m^c$, the infinite product above disappears and 4.9 implies:

4.10 The (A_j)'s are independent random variables, the distribution of $\mu(A_j)$ is Poisson, with mean $m(A_j)$.

More generally, we associate to the extended Poisson measure μ two new measures:

$$\mu^d(\omega; dt, dx) = \mu(\omega; dt, dx)1_J(t), \quad \mu^c = \mu - \mu^d$$

(or equivalently: if D and β are associated with μ by 1.14, then $D \cap J$ (resp. $D \cap J^c$) and β are associated with μ^d (resp. μ^c)). Then 4.9 implies that μ^c and μ^d are independent, that μ^c is a Poisson random measure with intensity m^c, and that the random variables $(\beta_s 1_D(S))_{s\in J}$ which completely describe μ^d by 1.14 are independent and their distribution is: $P(s \in D, \beta_s \in A) = m(\{s\} \times A)$ for all $s \in J$, $A \in \mathscr{E}$. All these facts exactly generalize the comments made after Theorem 4.5.

Proof of Theorem 4.8. The necessary condition in (a) follows from 1.21. Conversely, assume that $\mu^p = m$ for some measure m. We have $E(\mu(A)) = m(A)$, hence 1.20(i). We also have $\mu(\{0\} \times E) = 0$ (this is in the definition of any random measure). Then to obtain the sufficient condition in (a) and property (b), it suffices to consider $s \in \mathbb{R}_+$, $B \in \mathscr{F}_s$, and measurable subsets $(A_j)_{j\le d}$ of $(s, \infty) \times E$, pairwise disjoints, with $m(A_j) < \infty$, and to prove that

4.11
$$\begin{aligned} &E\left(1_B \exp \sum_{j\le d} iu_j 1_{A_j} * \mu_t\right) \\ &= P(B)\left[\exp \sum_{j\le d} (e^{iu_j} - 1)m^c(A_j \cap (s,t] \times E)\right] \\ &\quad \times \prod_{s<r\le t} \left[1 + \sum_{j\le d} (e^{iu_j} - 1)m(\{r\} \times E \cap A_j)\right] \end{aligned}$$

for every $t \ge s$.

Put $Y = \exp i\sum_{j\le d} u_j 1_{A_j} * \mu$ and $W = \sum_{j\le d} Y_- 1_{A_j}(e^{iu_j} - 1)$. The two processes Y and $1 + W * \mu$ have finite variation, are purely discontinuous, and take the value 1 for $t = 0$; moreover an elementary computation shows that $\Delta Y = \Delta(W * \mu)$. Hence we have $Y = 1 + W * \mu$.

Then by definition of μ^p, the process $M = Y - W * \mu^p$ is a complex-valued local martingale, and even a martingale because it is bounded by $1 + 2\sum m(A_j)$ (recall that $|Y| = 1$ and that $\mu^p = m$). Then, since $Y_s = 1$ and $W * \mu_s^p = 0$ by definition of W and the assumption that $A_j \subset (s, \infty) \times E$, we obtain:

4.12 $$1_B Y_t = 1_B + 1_B(M_t - M_s) + 1_B \int_s^t Y_{r-} \int_E \sum_{j \le d} (e^{iu_j} - 1) 1_{A_j}(r, x) m(dr, dx)$$

Put $f(t) = E(1_B Y_t)/P(B)$ and

$$a(t) = \int_{s \wedge t}^t \int_E \sum_{j \le d} (e^{iu_j} - 1) 1_{A_j}(r, x) m(dr, dx).$$

Then taking the expectation in 4.12, using Fubini's Theorem and the fact that M is a martingale, yield for $t \ge s$:

$$f(t) = 1 + \int_s^t f(r-)\, da(r).$$

Then by I.4.61, we have $f(t) = \mathscr{E}(a)_t$ for $t \ge s$. Therefore, if we notice that $\mathscr{E}(a)_t P(B)$ is equal to the right-hand side of 4.11, we obtain that 4.11 holds, and the theorem is proved. □

§ 4c. Processes with Independent Increments and Semimartingales

Observe first that a PII is not necessarily a semimartingale: for example any deterministic càdlàg process null at 0 is a PII, but it is not a semimartingale unless it has finite variation (I.4.28). We shall describe all PII and their characteristics in Section 5. In this subsection we are only interested in PII's that are semimartingales. Introduce first the notation:

4.13 $$g(u)_t = E(\exp iu \cdot X_t), \quad t \in \mathbb{R}_+,\ u \in \mathbb{R}^d.$$

It is obvious that $g(u)_0 = 1$ and that all functions: $t \rightsquigarrow g(u)_t$ are càdlàg.

4.14 **Theorem.** *Let X be a d-dimensional PII. Then X is also a semimartingale if and only if, for each $u \in \mathbb{R}^d$, the function: $t \rightsquigarrow g(u)_t$ has finite variation over finite intervals.*

We shall prove this theorem simultaneously with the next one, which describes the characteristics of a PII-semimartingale.

4.15 **Theorem.** *Let X be a d-dimensional semimartingale with $X_0 = 0$. Then it is a PII if and only if there is a version (B, C, ν) of its characteristics that is deterministic.*

Furthermore, in this case, the set of all fixed times of discontinuity is $J = \{t: \nu(\{t\} \times \mathbb{R}^d) > 0\}$, and for all $s \le t$, $u \in \mathbb{R}^d$ we have:

4.16
$$E(\exp iu\cdot(X_t - X_s)) = \exp\Bigg[iu\cdot(B_t - B_s) - \frac{1}{2}u\cdot(C_t - C_s)\cdot u + \int_s^t \int_{\mathbb{R}^d} (e^{iu\cdot x} - 1 - iu\cdot h(x))1_{J^c}(r)\nu(dr, dx)\Bigg] \times \prod_{s<r\leq t} \left\{e^{-iu\cdot \Delta B_r}\left[1 + \int (e^{iu\cdot x} - 1)\nu(\{r\}\times dx)\right]\right\}.$$

In particular, $g(u)_t$ is equal to the right-hand side of 4.16 for $s = 0$, and the law of the random variable ΔX_t is

4.17
$$\nu(\{t\}\times dx) + [1 - \nu(\{t\}\times \mathbb{R}^d)]\varepsilon_0(dx).$$

Formula 4.16 and the fact that X has independent increments allows to compute easily the distribution of any family $(X_{t_1},\dots,X_{t_n})$: hence the distribution of the process X is completely characterized by the (deterministic) characteristics of X: we will formalize this property at the end of this subsection (Theorem 4.25).

Note that Theorem 4.4 is a corollary of this one (with $C_t = \sigma^2(t)$ and $B = 0$, $\nu = 0$). Theorem 4.5 is also a corollary (with $C = 0$, $\nu(dt, dx) = da(t)\varepsilon_1(dx)$, $B_t = h(1)a(t)$).

If we combine the previous theorem with 1.19, we obtain:

4.18 **Corollary.** *A PII-semimartingale is quasi-left-continuous if and only if it has no fixed time of discontinuity* (we will see later that the same holds even when the PII is not a semimartingale).

Before proving Theorems 4.14 and 4.15, we state and prove the particular and very important case of PIIS.

4.19 **Corollary.** *A d-dimensional process X is a PIIS if and only if it is a semimartingale admitting a version (B, C, ν) of its characteristics that has the form:*

4.20
$$B_t(\omega) = bt, \quad C_t(\omega) = ct, \quad \nu(\omega; dt, dx) = dtK(dx)$$

where $b\in\mathbb{R}^d$, c is a symmetric nonnegative $d\times d$ matrix, K is a positive measure on $\mathbb{R}^d$ that integrates $(|x|^2\wedge 1)$ and satisfies $K(\{0\}) = 0$. Moreover, for all $t\in\mathbb{R}_+$, $u\in\mathbb{R}^d$ we have:

4.21
$$E(e^{iu\cdot X_t}) = \exp t\left[iu\cdot b - \frac{1}{2}u\cdot c\cdot u + \int (e^{iu\cdot x} - 1 - iu\cdot h(x))K(dx)\right].$$

Thus we have proved Lévy-Khintchine formula (although not in the most elementary way, admittedly!)

Proof of Corollary 4.19. The sufficient condition follows from 4.15; in particular, 4.16 implies that since (B, C, ν) are given by 4.20, then $E(\exp iu \cdot (X_t - X_s))$ depends only on u and $t - s$, thus yielding the stationarity of the increments of X.

Assume conversely that X is a PIIS. It is then trivial that $g(u)_{t+s} = g(u)_t g(u)_s$, hence $g(u)_t$ is of the form $g(u)_t = \exp t\psi(u)$ and in particular $t \rightsquigarrow g(u)_t$ has finite variation. Then 4.14 implies that X is a semimartingale. Hence, according to 4.15, we may choose a version (B, C, ν) of the characteristics that is deterministic: by 4.3 we have $J = \emptyset$, and 4.16 yields

$$t\psi(u) = iu \cdot B_t - \frac{1}{2} u \cdot C_t \cdot u + \int (e^{iu \cdot x} - 1 - iu \cdot h(x)) \nu([0, t] \times dx).$$

Then if we put $b = B_1$, $c = C_1$, $K(dx) = \nu([0, 1] \times dx)$, the uniqueness lemma 2.44 implies 4.20. □

Proof of Theorems 4.14 and 4.15. a) Let us first prove some auxiliary results. Put

$$S(u) = \inf(t\colon g(u)_t = 0)$$
$$h(u)_{s,t} = E(\exp iu \cdot (X_t - X_s)) \quad \text{for } s \leq t.$$

Each $h(u)_{s,t}$ is càdlàg in s and t. We will prove that if X is a PII, then

4.22 $t \rightsquigarrow |g(u)_t|$ is decreasing, and $g(u)_{S(u)-} \neq 0$ if $S(u) < \infty$ (hence $g(u)_t = 0$ if $t \geq S(u)$).

The independence of the increments immediately implies that for $s \leq t$, $g(u)_t = g(u)_s h(u)_{s,t}$. Since $|h(u)_{s,t}| \leq 1$ we obtain the first part of 4.22, and $g(u)_t = 0$ if $t \geq S(u)$. We also deduce that $g(u)_{t-} = g(u)_s h(u)_{s,t-}$ if $s < t$; thus if $S(u) < \infty$ and $g(u)_{S(u)-} = 0$, we should have $h(u)_{s,S(u)-} = 0$ for all $s < S(u)$, a fact which contradicts the property that $h(u)_{s,t-} = E(\exp iu \cdot (X_{t-} - X_s)) \to 1$ when $s \uparrow\uparrow t$. Then we must have 4.22.

We also put

$$Z(u)_t = \begin{cases} \exp iu \cdot X_t / g(u)_t & \text{if } t < S(u) \\ \exp iu \cdot X_{S(u)-} / g(u)_{S(u)-} & \text{if } t \geq S(u) \end{cases}$$

which by 4.22 defines a bounded process. If we consider the process $X' = X 1_{[\![0, S(u)[\![} + X_{S(u)-} 1_{[\![S(u), \infty[\![}$, then X' is obviously another PII and the function associated to it by 4.13 is $g'(u)_t = g(u)_t$ for $t < S(u)$, $g'(u)_t = g(u)_{S(u)-}$ for $t \geq S(u)$. By definition of $Z(u)$ we have for $s \leq t$

$$Z(u)_t = Z(u)_s \frac{g'(u)_s}{g'(u)_t} e^{iu \cdot (X'_t - X'_s)}$$

and the independence of the increments of X' immediately yields that:

4.23 $Z(u)$ is a martingale, and $a \leq |Z(u)| \leq 1$ for some $a > 0$ (this last property is obvious from 4.22 and the definition of $Z(u)$).

b) Let us prove the necessary condition of 4.14. We assume that X is a PII and a semimartingale. Then if we apply Ito's formula to a function f of class C^2 on $\mathbb{R}^{d+2}$ which satisfies $f(x,y,z) = e^{-iu\cdot x}/(y+iz)$ for $|y+iz| \geq a$ (where a appears in 4.23), we obtain that $e^{-iu\cdot X}/Z(u)$ is a semimartingale. Since $g(u)_t = [e^{-iu\cdot X_t}/Z(u)_t] 1_{\{t<S(u)\}}$, it follows that $t \rightsquigarrow g(u)_t$ is a semimartingale, and thus has finite variation by I.4.28.

c) Let us prove the sufficient condition of 4.14. We assume that each $t \rightsquigarrow g(u)_t$ has finite variation. Let $t \in \mathbb{R}_+$. We have $g(u)_t \to 1$ when $u \to 0$, hence there exists $b > 0$ such that $|u| \leq b$ implies $|g(u)_t| > 0$, and therefore if $|u| \leq b$ we have $\exp iu\cdot X_s = Z(u)_s g(u)_s$ for $s \leq t$, and both $Z(u)$ and $(g(u)_s)_{s\geq 0}$ are semimartingales (by 4.23 and the hypothesis). Then I.4.57 yields that $\exp iu\cdot X^t$ is a semimartingale. Now if $|u| > b$ there exists an $n \in \mathbb{N}$ such that $\left|\frac{u}{n}\right| \leq b$, while $\exp iu\cdot X = \left[\exp i\frac{u}{n}\cdot X\right]^n$, hence I.4.57 again implies that $\exp iu\cdot X^t$ is a semimartingale.

Now this is true for all $t \in \mathbb{R}_+$. Therefore by I.4.25 we see that $\exp iu\cdot X$ is a semimartingale, for all $u \in \mathbb{R}^d$, and the proof of 2.42 yields that X is a semimartingale.

d) Let us prove the necessary condition of 4.15. We assume that X is a PII and a semimartingale, and we call (B,C,ν) its characteristics, and $A(u)$ the processes associated by 2.40. By part (b) of this proof, $t \rightsquigarrow g(u)_t$ has finite variation, hence if we apply the uniqueness part of Theorem 2.47 we obtain that $\mathscr{E}[A(u)]$ is indistinguishable from the process $g(u)$ on $[\![0,S(u)[\![$.

More generally, let $s \in \mathbb{R}_+$: then $X - X^s$ is a PII, to which the function "$g(u)_t$" associated is $t \rightsquigarrow h(u)_{s,t}$, and also a semimartingale to which the process "$A(u)$" associated is $A(u) - A(u)^s$. Then the same proof as before shows that $\mathscr{E}[A(u) - A(u)^s]_t = h(u)_{s,t}$ a.s. for all $t \geq s$, $t < \inf(r > s\colon h(u)_{s,r} = 0)$.

Now, $\mathscr{E}[A(u) - A(u)^s]_t(\omega)$ is continuous in u and right continuous in s,t for all $\omega \in \Omega$. Then, by changing (B,C,ν) and $A(u)$ on a P-null set, we may assume that $\mathscr{E}[A(u) - A(u)^s]_t = h(u)_{s,t}$ identically for all $u \in \mathbb{R}^d$, $s \leq t < \inf(r > s\colon h(u)_{s,r} = 0)$.

If $b = \mathscr{E}(a)$ for a complex-valued function with finite variation and $a(0) = 0$, then $a(t) = \int_0^t b(s-)^{-1} db(s)$ whenever $b(t) \neq 0$. Thus we obtain from what precedes that each $A(u)_t - A(u)_{t\wedge s}$ is deterministic for all $t < \inf(r > s\colon \Delta A(u)_r = -1)$. Then it easily follows that $A(u)_t(\omega) = a(u)_t$ identically for some function $a(u)$, and the uniqueness lemma 2.44 yields that B, C, ν do not depend either on ω.

e) Finally, we suppose that X is a semimartingale with $X_0 = 0$ and with deterministic characteristics (B,C,ν). The processes $A(u)$ associated by 2.40 are also deterministic. Let $M(u) = e^{iu\cdot X} - e^{iu\cdot X_-}\cdot A(u)$ be the local martingale obtained in 2.42(b); in fact, $|M(u)_t| \leq 1 + \mathrm{Var}[A(u)]_t$, hence $M(u)$ is a martingale.

Let $s \in \mathbb{R}_+$ and $F \in \mathscr{F}_s$ with $P(F) > 0$ be fixed. For $t \geq s$ we have

$$1_F e^{iu\cdot(X_t - X_s)} = 1_F + 1_F e^{-iu\cdot X_s}(M(u)_t - M(u)_s) + \int_s^t 1_F e^{iu\cdot(X_{r-} - X_s)} dA(u)_r.$$

Hence if $f(t) = E(1_F \exp iu\cdot(X_t - X_s))/P(F)$ for $t \geq s$, and $f(t) = 1$ for $t < s$, we obtain by taking the expectation above and by applying Fubini's Theorem (recall once more that $A(u)$ is deterministic):

$$f(t) = 1 + \int_0^t f(r-)d[A(u) - A(u)^s](r).$$

Thus I.4.61 implies $f(t) = \mathscr{E}[A(u) - A(u)^s]_t$, that is for $t \geq s$:

$$E(1_F e^{iu\cdot(X_t - X_s)}) = P(F)\mathscr{E}[A(u) - A(u)^s]_t.$$

This, being true for all $F \in \mathscr{F}_s$, implies that $X_t - X_s$ is independent from $\mathscr{F}_s$. By I.4.63 and 2.40, $\mathscr{E}[A(u) - A(u)^s]_t$ is equal to the right-hand side of 4.16, hence what precedes also shows formula 4.16.

Moreover, 4.16 yields (make $s \uparrow\uparrow t$):

$$E(\exp iu\cdot\Delta X_t) = 1 + \int (e^{iu\cdot x} - 1)\nu(\{t\} \times dx),$$

which is equal to 1 for all $u \in \mathbb{R}^d$ if and only if $\nu(\{t\} \times \mathbb{R}^d) = 0$. Then the set $J = \{t: \nu(\{t\} \times \mathbb{R}^d) > 0\}$ is the set of fixed times of discontinuity for X, and we also deduce 4.17 from the above formula: that finishes the proof of 4.16. $\square$

Now we explain precisely what we meant in the comment made after Theorem 4.15: that the distribution of a semimartingale-PII is characterized by its characteristics.

To this end, we suppose that X is a given d-dimensional càdlàg process defined on the space Ω. We consider a σ-field $\mathscr{H}$, and the filtration:

4.24
$$\begin{cases} \mathscr{G}_t = \bigcap_{s>t} \mathscr{G}_s^0 \quad \text{and} \quad \mathscr{G} = \bigvee_{(t)} \mathscr{G}_t, \\ \text{where } \mathscr{G}_s^0 = \mathscr{H} \vee \sigma(X_r: r \leq s) \end{cases}$$

4.25 **Theorem.** *Let P and Q be two probability measures on $(\Omega, \mathscr{G})$, such that:*

(i) *P and Q coincide on the σ-field $\mathscr{H}$;*

(ii) *$X_0 = 0$ P-a.s. and Q-a.s.;*

(iii) *X is a semimartingale with the same deterministic characteristics on the two stochastic bases $(\Omega, \mathscr{G}, \mathbf{G}, P)$ and $(\Omega, \mathscr{G}, \mathbf{G}, Q)$;*

Then $P = Q$.

(Note that in (iii) one may replace any one of the stochastic bases by its completion (I.1.4) without altering the result).

Proof. By 4.15, X has independent increments on the two bases. Since the characteristics are the same, 4.16 shows that the expectation of $1_F \exp.i \sum_{j\leq n} u^j\cdot(X_{t_j} - X_{t_{j-1}})$, where $0 = t_0 < t_1 < \cdots < t_n$ and $F \in \mathscr{H}$, is the same for P and for

Q. Since $\mathscr{G}$ is generated by $\mathscr{H}$ and by the variables $X_t - X_s$ and X_0, we obtain $P = Q$. □

We end this subsection with a well-known and very useful result. We state it with PII-semimartingales, but it stays true for all PII, as we shall see later (Proposition 5.29); $\nu(\{t\} \times g)$ stands for $\int \nu(\{t\} \times dx)g(x)$.

4.26 **Proposition.** *Let X be a PII-semimartingale with characteristics (B, C, ν). Let g be a Borel nonnegative function on $\mathbb{R}^d$, vanishing on a neighbourhood of* 0. *Then $X'_t = g * \mu^X_t = \sum_{s \leq t} g(\Delta X_s)$ is a PII, and*

$$E(e^{-X'_t}) = \exp\{-(1 - e^{-g})1_{J^c} * \nu_t + \sum_{s \leq t} \operatorname{Log}(1 - \nu(\{s\} \times (1 - e^{-g})))\}.$$

Proof. $X'_t - X'_s = \sum_{s<r\leq t} g(\Delta X_r)$ is clearly independent from $\mathscr{F}_s$, hence the first claim. If $Z = \exp - X'$, Ito's formula gives

$$\begin{aligned} Z &= 1 - Z_- \cdot (g * \mu^X) + \sum_{s \leq \cdot} Z_{s-}[e^{-g(\Delta X_s)} - 1 + g(\Delta X_s)] \\ &= 1 - Z_-(1 - e^{-g}) * \mu^X \end{aligned}$$

4.27 $$= 1 - M - Z_-(1 - e^{-g}) * \nu$$

where M is the local martingale $M = Z_-(1 - e^{-g}) * \mu^X - Z_-(1 - e^{-g}) * \nu$. Since $0 \leq Z \leq 1$,

$$E[Z_-(1 - e^{-g}) * \mu^X_t] = E[Z_-(1 - e^{-g}) * \nu_t] \leq (1 - e^{-g}) * \nu_t < \infty,$$

and so M is indeed a martingale. Then if $z(t) = E(Z_t)$, 4.27 and Fubini's theorem (because ν is deterministic) yield

$$z(t) = 1 - \int_0^t z(s-)\, dA_s,$$

where $A = (1 - e^{-g}) * \nu$. Thus $z(t) = \mathscr{E}(-A)_t$ is given by I.4.63, which easily yields the claimed formula. □

§ 4d. Gaussian Martingales

As a first application of Theorem 4.15, we elucidate the structure of all Gaussian martingales.

4.28 **Definition.** A *Gaussian martingale* is an $\mathbb{R}^d$-valued martingale X such that:

(i) $X_0 = 0$;

(ii) the distribution of any finite family $(X_{t_1}, \ldots, X_{t_n})$ is Gaussian. □

We put:

4.29 $$\hat{c}^{ij}(t) = E(X_t^i X_t^j), \quad \hat{c}(t) = (\hat{c}^{ij}(t))_{i,j \le d}$$

4.30 **Proposition.** *Let X be a Gaussian martingale. For all $s \le t$, $\hat{c}(t) - \hat{c}(s)$ is a symmetric nonnegative matrix, and for all $u \in \mathbb{R}^d$,*

4.31 $$E(e^{iu\cdot(X_t - X_s)}) = \exp[-\tfrac{1}{2}u\cdot(\hat{c}(t) - \hat{c}(s))\cdot u]$$

Proof. The symmetry is evident. Let $u \in \mathbb{R}^d$ and $a(t) = u\cdot\hat{c}(t)\cdot u$ and $Y = u\cdot X$. Then we have $a(t) = E(Y_t^2)$ and for $s \le t$,

$$Y_t^2 - Y_s^2 = (Y_t - Y_s)^2 + 2Y_s(Y_t - Y_s).$$

Taking the expectation above, and since X, hence Y, are martingales, we obtain $a(t) - a(s) = E((Y_t - Y_s)^2) \ge 0$; this shows that $\hat{c}(t) - \hat{c}(s)$ is nonnegative. Moreover it proves that $Y_t - Y_s$, which is a centered Gaussian variable, admits $a(t) - a(s)$ for its variance, hence 4.31. □

In fact, we have more: a Gaussian martingale X is a Gaussian centered process, whose covariance function is

$$E(X_s^j X_t^k) = \hat{c}^{jk}(s) \quad \text{for } s \le t$$

(immediate from the martingale property). Hence *the distribution of X is entirely characterized by the function $\hat{c}$.* One can also see that otherwise: the martingale property implies that $X_t - X_s$ is uncorrelated with the variables $(X_r, r \le s)$; since the process X is Gaussian, this implies that $X_t - X_s$ is independent from the X_r's for $r \le s$. Hence by 4.31, for all $0 = t_0 \le t_1 \le \cdots \le t_n$ and $u_j \in \mathbb{R}^d$ we have

4.32 $$E\left(\exp \sum_{j \le n} iu_j\cdot(X_{t_j} - X_{t_{j-1}})\right) = \exp\left[-\frac{1}{2}\sum_{j \le n} u_j\cdot(\hat{c}(t_j) - \hat{c}(t_{j-1}))\cdot u_j\right].$$

This formula also shows that $\hat{c}$ completely determines the distribution of the process X.

Now, it may happen that the σ-field $\mathscr{F}_s$ contains non-trivial sets that are not measurable with respect to $\sigma(X_r: r \le s)$; hence $X_t - X_s$ may be non-independent from $\mathscr{F}_s$, and X is not a PII on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ in the sense of definition 4.1. Therefore, let us consider a new filtration:

4.33 $$\mathscr{G}_t = \bigcap_{s>t} \mathscr{G}_s^0, \quad \text{where } \mathscr{G}_s^0 = \sigma(X_r: r \le s).$$

Introduce also some more notation: first, we set

4.34 $$J = \{t > 0: \hat{c}(t) \ne \hat{c}(t-)\}.$$

If $t \in J$, the matrix $\Delta\hat{c}(t) = \hat{c}(t) - \hat{c}(t-)$ is symmetric nonnegative by 4.30, and $\Delta\hat{c}(t) \ne 0$. Then there exists a unique (Gaussian) probability measure K_t on $\mathbb{R}^d$

such that

4.35 $$\int e^{iu\cdot x} K_t(dx) = \exp -\frac{1}{2} u\cdot \Delta\hat{c}(t)\cdot u.$$

4.36 **Theorem.** *Let X be a Gaussian martingale, to which are associated $\hat{c}$, $\mathbf{G} = (\mathscr{G}_t)_{t\geq 0}$, J, K_t as above.*

a) *X is a PII on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{G}, P)$, with characteristics:*

4.37 $$\begin{cases} B_t = \sum_{s\leq t, s\in J} \int h(x) K_s(dx) \\ C_t = \hat{c}(t) - \sum_{s\leq t} \Delta\hat{c}(s) \\ \nu(dt, dx) = \sum_{s\in J} \varepsilon_s(dt) K_s(dx). \end{cases}$$

b) *J is the set of fixed times of discontinuity of X; if $t\in J$, $P(\Delta X_t \neq 0) = 1$ and K_t is the distribution of ΔX_t; and almost all paths are continuous outside J.*

c) *X is a Wiener process* (in the sense of I.4.9) *if and only if $J = \varnothing$.*

Proof. (a) Let $s < r \leq t$. We have seen before (see for instance 4.32) that $X_t - X_r$ is independent from $\mathscr{G}_r^0$. Thus $X_t - X_s = \lim_{r\downarrow\downarrow s}(X_t - X_r)$ is independent from $\bigcap_{r>s} \mathscr{G}_r^0 = \mathscr{G}_s$: therefore X is a PII on $(\Omega, \mathscr{F}, \mathbf{G}, P)$.

Define (B, C, ν) by 4.37. A trivial computation shows that the right-hand sides of 4.16 and of 4.31 coincide. Thus it suffices to prove that the right-hand sides of 4.16, when $u\in\mathbb{R}^d$ and $s\leq t$ take all possible values, completely determine (B, C, ν). From 2.44, these right-hand sides characterize B^c, C, and $\nu(ds, dx)1_{J^c}(s, x)$, and also $\int (e^{iu\cdot x} - 1)\nu(\{r\}\times dx)$ for all $r\in J$; hence they characterize the measures $\nu(\{r\}\times dx)$ for $r\in J$, hence they determine the measure ν, and finally they also determine B because we have $B_t = B_t^c + \sum_{s\leq t} \int h(x)\nu(\{s\}\times dx)$ by 2.14.

(b) The two first assertions come from the fact that $E(\exp iu\cdot \Delta X_t) = \exp[-\frac{1}{2} u\cdot \Delta\hat{c}(t)\cdot u]$, which in turn follows immediately from 4.31. Moreover, 4.37 implies that $1_{J^c} * \nu = 0$, while we have $E(1_{J^c} * \mu_\infty^X) = E(1_{J^c} * \nu_\infty)$ because ν is the compensator of μ^X. Then the last claim follows from the fact that on the set where $1_{J^c} * \mu_\infty^X = 0$ the process X is continuous at each point $t\in J^c$.

c) is an immediate consequence of (b). $\square$

As a trivial consequence, the continuous and purely discontinuous martingale parts of X are:

4.38 $$X^c = X - X^d, \quad X^d = \sum_{s\leq t, s\in J} \Delta X_s$$

where the series defining X_t^d converges in L^2; X^c and X^d are two independent Gaussian martingales, and X^c is a Wiener process.

5. Processes with Independent Increments Which Are Not Semimartingales

As seen before, there are PII's which are not semimartingales. However, they cannot be far apart from being semimartingales, as we will prove below. Beside their theoretical interest, the forthcoming results are useful in providing a simple proof for the sufficient condition under which a sequence of PII's converges (see Chapter VII).

Nevertheless, this section is of secondary importance, and in any case we advise the reader to look only at § 5a, where all the results are stated, and perhaps at the last proposition of this section (Proposition 5.29).

§ 5a. The Results

Let $X = (X^i)_{i \le d}$ be a d-dimensional càdlàg process with $X_0 = 0$, on a given stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

5.1 **Theorem.** *X is a PII if and only if it has the form $X = Y + A$, where*
(i) *Y is a PII and a semimartingale;*
(ii) *$A = (A_t)_{t \ge 0}$ is a deterministic càdlàg function: $\mathbb{R}_+ \to \mathbb{R}^d$, with $A_0 = 0$.*

There are two ways of defining the characteristics of a PII-semimartingale: either via the martingale properties that come up with Definition 2.6 (or Theorem 2.21), or via Formula 4.16. Below we introduce the characteristics of a general PII, first through 4.16.

5.2 **Theorem.** *Let $h \in \mathscr{C}_t^d$ be a truncation function.*
a) *Assume that X is a PII on $(\Omega, \mathscr{F}, \mathbf{F}, P)$. There exists a triplet (B, C, ν) with B depending upon h, and only one, where*

5.3 $B = (B^i)_{i \le d}$ is a càdlàg function: $\mathbb{R}_+ \to \mathbb{R}^d$ with $B_0 = 0$ (B has not necessarily finite variation over finite intervals).

5.4 $C = (C^{jk})_{j,k \le d}$ is a continuous function: $\mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^d$ with $C_0 = 0$, and $C_t - C_s$ a symmetric nonnegative matrix for all $s \le t$.

5.5 ν is a positive measure on $\mathbb{R}_+ \times \mathbb{R}^d$ with:
(i) $\nu(\{0\} \times \mathbb{R}^d) = 0$, $\nu(\mathbb{R}_+ \times \{0\}) = 0$, $\nu([0,t] \times \{x: |x| > \varepsilon\}) < \infty$ for all $\varepsilon > 0$;
(ii) $\nu(\{t\} \times \mathbb{R}^d) \le 1$;
(iii) $\int_0^t \int_{\mathbb{R}^d} |h(x - \Delta B_s)|^2 \nu(ds \times dx) + \sum_{s \le t} (1 - \nu(\{s\} \times \mathbb{R}^d)) |h(-\Delta B_s)|^2 < \infty$;
(iv) $\sum_{s \le t} |\int h(x - \Delta B_s) \nu(\{s\} \times dx) + (1 - \nu(\{s\} \times \mathbb{R}^d)) h(-\Delta B_s)| < \infty$;
(v) $\Delta B_t = \nu(\{t\} \times h) \quad (= \int h(x) \nu(\{t\} \times dx))$;

and such that, if $J = \{t: \nu(\{t\} \times \mathbb{R}^d) > 0\}$, Formula 4.16 holds for all $s \le t$, $u \in \mathbb{R}^d$.

Furthermore, for all $t \in \mathbb{R}_+$ the law of ΔX_t is given by 4.17, and J is the set of fixed times of discontinuity of X.

b) *Conversely, if (B, C, ν) satisfies 5.3, 5.4, 5.5, the right-hand side of 4.16 makes sense, and there is a PII X* (by which we mean in particular a *càdlàg* process) *on some space $(\Omega, \mathscr{F}, P)$ which meets 4.16, and moreover the law $\mathscr{L}(X)$ is entirely determined by (B, C, ν).*

In spite of the fact that B has not necessarily finite variation, it is natural to call (B, C, ν) the *characteristics* of X, writing $B(h)$ if one wishes to emphazise the dependence on h. Observe that (iii) and (iv) in 5.5 do not depend on h, provided ΔB is given by 5.5v (this is not immediately obvious, but comes as a part of the theorem).

Next, we state a characterization of (B, C, ν) analogous to 2.21. For that we need to introduce the "modified second characteristic". However, Formula 2.18 does not necessarily make sense, because $|h|^2 * \nu_t$ may be infinite. But there exists $\varepsilon > 0$ such that $h(x) = x$ for $|x| \leq 2\varepsilon$, hence $h(x - \Delta B) = h(x) - \Delta B$ if $|x| \leq \varepsilon$ and $|\Delta B| \leq \varepsilon$; moreover there are only finitely many $s \leq t$ such that $|\Delta B_s| > \varepsilon$; therefore we readily deduce from 5.5i,iii,iv that

5.6 $$\int_0^t \int_{\mathbb{R}^d} |h(x) - \Delta B_s|^2 \nu(ds, dx) + \sum_{s \leq t} [1 - \nu(\{s\} \times \mathbb{R}^d)] |\Delta B_s|^2 < \infty.$$

Henceforth we can define the following functions:

5.7 $$\tilde{C}_t^{ij} = C_t^{ij} + \int_0^t \int_{\mathbb{R}^d} [h^i(x) - \Delta B_s^i][h^j(x) - \Delta B_s^j] \nu(ds, dx) + \sum_{s \leq t} [1 - \nu(\{s\} \times \mathbb{R}^d)] \Delta B_s^i \Delta B_s^j.$$

Of course, if B has finite variation and if $(|x|^2 \wedge 1) * \nu_t < \infty$, then 5.7 reduces to 2.18 (as a matter of fact, one could prove that $(|x|^2 \wedge 1) * \nu_t < \infty$ is implied by the fact that B has finite variation). If X has no fixed times of discontinuity, in which case B is continuous and 5.5(iii) reads $|h(x)|^2 * \nu_t < \infty$, then 5.7 further reduces to

5.8 $$\tilde{C}^{ij} = C^{ij} + (h^i h^j) * \nu.$$

In all cases we have

5.9 $$\Delta \tilde{C}_t^{ij} = \nu(\{t\} \times (h^i h^j)) + \Delta B_t^i \Delta B_t^j,$$

and $\tilde{C}_t - \tilde{C}_s$ is a symmetric nonnegative matrix for $s \leq t$.

5.10 **Theorem.** *Let (B, C, ν) satisfy 5.3, 5.4, 5.5 relative to some function $h \in \mathscr{C}_t^d$, and define $\tilde{C}$ by 5.7, and $X(h)$ by 2.4, and μ^X by 1.16. There is equivalence between:*

a) *X is a PII with characteristics (B, C, ν).*

b) *Each of the following processes is a local martingale:*

(i) $M(h) = X(h) - B$;
(ii) $M(h)^i M(h)^j - \tilde{C}^{ij}$ $(i,j \leq d)$;
(iii) $g * \mu^X - g * \nu$ ($g \in \mathscr{C}^+(\mathbb{R}^d)$, *where* $\mathscr{C}^+(\mathbb{R}^d)$ *is any family meeting* 2.20).
c) *All the processes in* (i), (ii), (iii) *above are martingales.*
Moreover, (b) *or* (c) *completely determine* (B, C, ν), *and they imply that* ν *is the compensator of* μ^X, *and that* $C^{ij} = \langle M(h)^{i,c}, M(h)^{j,c} \rangle$.

Since $X = X(h) + \check{X}(h)$ and since $\check{X}(h)$ is always a semimartingale, we deduce from 5.10 and I.4.28 that:

5.11 **Corollary.** *A PII is a semimartingale if and only if its first characteristic has finite variation over finite intervals.*

Combining the fact that J is the set of fixed times of discontinuity and that ν is the compensator of the measure μ^X, we then deduce from 1.19 the same result than in 4.18:

5.12 **Corollary.** *A PII is quasi-left continuous if and only if it has no fixed time of discontinuity.*

Finally, using the characterization 5.10, the same proof than in 2.24 yields:

5.13 **Corollary.** *If X is a PII and if h and h' are two truncation functions, the corresponding first characteristics $B(h)$ and $B(h')$ are related by 2.25.*

§ 5b. The Proofs

We break the proof of 5.1, 5.2 and 5.10 into several steps. The function $h \in \mathscr{C}_t^d$ is fixed throughout.

5.14 **Lemma.** *Let X be of the form $X = Y + A$, where Y and A satisfy 5.1(i, ii). Then X is a PII which satisfies all the properties stated in 5.2a, and which also satisfies 5.10b.*

Proof. a) That X is a PII is trivial (recall that A is deterministic).
b) In this part we compute the compensator of the measure μ^X. Call (B', C', ν') the characteristics of Y, and $a'_t = \nu'(\{t\} \times \mathbb{R}^d)$. If $g \in \mathscr{C}^+(\mathbb{R}^d)$,

$$g * \mu_t^X = \sum_{s \leq t} g(\Delta Y_s + \Delta A_s) = g(x + \Delta A) * \mu_t^Y + \sum_{s \leq t} g(\Delta A_s) 1_{\{\Delta Y_s = 0\}}.$$

We know that $[g(x + \Delta A) * \mu^Y]^p = g(x + \Delta A) * \nu'$. As for $F = \sum_{s \leq \cdot} g(\Delta A_s) 1_{\{\Delta Y_s = 0\}}$, we denote by (s_n) a sequence of times that exhausts the jumps of A. Then for all stopping times T,

$$\begin{aligned} E(F_T) &= \sum_n g(\Delta A_{s_n}) P(s_n \leq T, \Delta Y_{s_n} = 0) \\ &= \sum_n g(\Delta A_{s_n}) E[1_{\{s_n \leq T\}} P(\Delta Y_{s_n} = 0 | \mathscr{F}_{(s_n)-})] \\ &= \sum_n g(\Delta A_{s_n}) E(1_{\{s_n \leq T\}}(1 - a'_{s_n})) = E\Big(\sum_{s \leq T} g(\Delta A_s)(1 - a'_s)\Big) \end{aligned}$$

(use 1.17 for the third equality). Hence $F^p = \sum_{s \leq \cdot} g(\Delta A_s)(1 - a'_s)$, and so

$$(g * \mu^X)^p = g(x + \Delta A) * \nu' + \sum_{s \leq \cdot} g(\Delta A_s)(1 - a'_s).$$

Now we define a positive measure ν on $\mathbb{R}_+ \times \mathbb{R}^d$ by

5.15
$$\begin{aligned} \nu([0,t] \times H) = &\int_0^t \int_{\mathbb{R}^d} 1_{H \setminus \{0\}}(x + \Delta A_s) \nu'(ds \times dx) \\ &+ \sum_{s \leq t} (1 - a'_s) 1_{H \setminus \{0\}}(\Delta A_s). \end{aligned}$$

It follows from what precedes that $(g * \mu^X)^p = g * \nu$. This proves 5.10b(iii) and therefore ν is the compensator of μ^X (see the proof of 2.21). In the sequel, we set $a_t = \nu(\{t\} \times \mathbb{R}^d)$.

c) Now we construct the first characteristic B of X, and we prove 5.10b(i). We know that $M'(h) = Y(h) - B'$ is a local martingale, while

$$Y = M'(h) + B' + \sum_{s \leq \cdot} [\Delta Y_s - h(\Delta Y_s)]$$

5.16
$$\begin{aligned} X(h) &= Y + A - \sum_{s \leq \cdot} [\Delta Y_s + \Delta A_s - h(\Delta Y_s + \Delta A_s)] \\ &= M'(h) + A + B' + \sum_{s \leq \cdot} [h(\Delta Y_s + \Delta A_s) - \Delta A_s - h(\Delta Y_s)] \\ &= M'(h) + A + B' + [h(x + \Delta A) - \Delta A - h(x)] * \mu^Y \\ &\quad + \sum_{s \leq \cdot} [h(\Delta A_s) - \Delta A_s] 1_{\{\Delta Y_s = 0\}} \end{aligned}$$

Since $h(x) = x$ for $|x|$ small enough, and h is bounded, and ΔA is locally bounded and equal to $h(\Delta A)$ except on a locally finite set, it is clear that the two processes $F = [h(x + \Delta A) - \Delta A - h(x)] * \mu^Y$ and $G = \sum_{s \leq \cdot} [h(\Delta A_s) - \Delta A_s] 1_{\{\Delta Y_s = 0\}}$ have locally bounded variation, with deterministic compensators given by:

5.17
$$\begin{cases} F^p = [h(x + \Delta A) - \Delta A - h(x)] * \nu' \\ G^p = \sum_{s \leq \cdot} [h(\Delta A_s) - \Delta A_s](1 - a'_s) \end{cases}$$

(for the latter, the proof goes as in (b) above). Then comparing to 5.16 shows that $X(h) = M(h) + B$, where $M(h)$ is a local martingale and

5.18
$$\begin{cases} M(h) = M'(h) + F - F^p + G - G^p \\ B = A + B' + F^p + G^p. \end{cases}$$

This proves 5.10b(i).

d) Now we prove that the condition 5.5 is fulfilled. This is obvious for (i) and (ii). Recalling that $\Delta B'_t = \nu'(\{t\} \times h)$, we get from 5.17 and 5.18:

$$\Delta B_t = \Delta A_t + \Delta B'_t + \int \nu'(\{t\} \times dx)h(x + \Delta A_t) - \Delta A_t a'_t - \Delta B'_t$$

$$+ [h(\Delta A_t) - \Delta A_t](1 - a'_t)$$

$$= \int \nu'(\{t\} \times dx)h(x + \Delta A_t) + (1 - a'_t)h(\Delta A_t) = \nu(\{t\} \times h)$$

by 5.15. Hence 5.5(v) holds.

Let $Z = X - B$. Then $Z = Y - B' - F^p - G^p$ is a PII-semimartingale, with characteristics (B'', C'', ν''). We can reproduce the proof of (b) to get that

5.19
$$\nu''([0,t] \times H) = \int_0^t \int_{\mathbb{R}^d} 1_{H\backslash\{0\}}(x - \Delta B_s)\nu(ds \times dx) + \sum_{s \leq t} (1 - a_s) 1_{H\backslash\{0\}}(-\Delta B_s).$$

The number $|h|^2 * \nu''_t$ is finite (because Z is a semimartingale), and due to 5.19 it is equal to the left-hand side of 5.5(iii), so this condition is met. Finally,

$$\Delta B''_t = \nu''(\{t\} \times h) = \int_{\mathbb{R}^d} h(x - \Delta B_t)\nu(\{t\} \times dx) + (1 - a_t)h(-\Delta B_t),$$

and since $\sum_{s \leq t} |\Delta B''_s| < \infty$ (because B'' has finite variation over $[0,t]$) we deduce that 5.5(iv) is met.

e) Now we show what remains to be proved of 5.2a. As usual, we set $\nu^c = 1_{J^c} \cdot \nu$, and also $\nu'^c = 1_{J'^c} \cdot \nu'$ where $J' = \{t : a'_t > 0\}$. Then 5.15 clearly yields $\nu^c = \nu'^c$ and

$$1 + \nu(\{r\} \times (e^{iu \cdot x} - 1)) = 1 + \nu'(\{r\} \times (e^{iu \cdot x + iu \cdot \Delta A_r} - 1))$$

$$+ (1 - a'_r)(e^{iu \cdot \Delta A_r} - 1)$$

5.20
$$= e^{iu \cdot \Delta A_r}[1 + \nu'(\{r\} \times (e^{iu \cdot x} - 1))].$$

Therefore if $g_u(x) = e^{iu \cdot x} - 1 - iu \cdot h(x)$, formula 4.16 for Y yields

$$E(e^{iu \cdot (X_t - X_s)}) = \exp\{iu \cdot (A_t - A_s) + iu \cdot (B'_t - B'_s)$$

$$- \frac{1}{2} u \cdot (C'_t - C'_s) \cdot u + \nu'^c((s,t] \times g_u)\}$$

$$\times \prod_{s < r \leq t} e^{-iu \cdot \Delta B'_r}[1 + \nu'(\{r\} \times (e^{iu \cdot x} - 1))]$$

5.21
$$= \exp\{iu \cdot (A_t - A_s) + iu \cdot (B'_t - B'_s)$$

$$- \frac{1}{2} u \cdot (C'_t - C'_s) \cdot u + \nu^c((s,t] \times g_u)\}$$

$$\times \prod_{s < r \leq t} e^{-iu \cdot (\Delta A_r + \Delta B'_r)}[1 + \nu(\{r\} \times (e^{iu \cdot x} - 1))].$$

The processes F^p and G^p (see 5.17) are purely discontinuous, with jumps at the times s_n of discontinuity of A, and have finite variation; then

$$\exp[iu\cdot(F_t^p - F_s^p + G_t^p - G_s^p)] \prod_{s<r\leq t} \exp - iu\cdot(\Delta F_r^p + \Delta G_r^p) = 1.$$

Plugging into 5.21 and recalling that $B = A + B' + F^p + G^p$ yield

$$E(e^{iu\cdot(X_t - X_s)}) = \exp\{iu\cdot(B_t - B_s) - \frac{1}{2}u\cdot(C'_t - C'_s)\cdot u + \nu^c((s,t] \times g_u)\}$$
$$\times \prod_{s<r\leq t} e^{-iu\cdot\Delta B_r}[1 + \nu(\{r\} \times (e^{iu\cdot x} - 1))].$$

In other words, we have proved 4.16 for X, provided we set $C = C'$.

Now, we prove that B, C, ν are uniquely determined by 4.16. We first observe that letting $s \uparrow\uparrow t$ in 4.16 gives the law of ΔX_t, namely

5.22 $$E(e^{iu\cdot\Delta X_t}) = 1 + \nu(\{t\} \times (e^{iu\cdot x} - 1)).$$

Since $\nu(\mathbb{R}_+ \times \{0\}) = 0$, this uniquely determines $\nu(\{t\} \times \cdot)$ for all t. Since 5.5(v) holds, each term of the infinite product in 4.16 is uniquely determined, and thus so is "exp ..." in 4.16. Then the uniqueness lemma 2.44, plus $\nu^c(\mathbb{R}_+ \times \{0\}) = 0$ show that $B_t - B_s$ and $C_t - C_s$ and $\nu^c((s,t] \times \cdot)$ are also uniquely determined. Hence our claim is proved.

Finally, 5.22 clearly implies 4.17, and we trivially deduce that J is the set of all fixed times of discontinuity for X.

f) At this stage, it remains to prove 5.10b(ii). But 5.18 yields $M(h)^{i,c} = M'(h)^{i,c}$, so $\langle M(h)^{i,c}, M(h)^{j,c}\rangle = C^{ij}$ $(= C'^{ij})$. Moreover, $\Delta M(h) = h(\Delta X) - \Delta B$ and 5.5(v) holds, so exactly as in the proof of 2.34 we obtain that $h^i \in G_{\text{loc}}(\mu^X)$ and $M(h)^d = h * (\mu^X - \nu)$, so we deduce from 1.33a, by polarization, that

5.23 $$\langle M(h)^{i,d}, M(h)^{j,d}\rangle = (h^i - \Delta B^i)(h^j - \Delta B^j) * \nu + \sum_{s\leq\cdot} (1 - a_s)\Delta B_s^i \Delta B_s^j.$$

Since $\langle N, N\rangle = \langle N^c, N^c\rangle + \langle N^d, N^d\rangle$ for all $N \in \mathscr{H}^2_{\text{loc}}$ (see I.4.15) we deduce that $\tilde{C}^{ij}$, as defined by 5.7, is a version of $\langle M(h)^i, M(h)^j\rangle$ and we are done. □

Proof of Theorem 5.1. The only thing left to be shown is that if X is a PII, then it has the form $X = Y + A$ with Y, A meeting 5.1(i, ii).

a) The basic idea of the proof goes as follows. Let X' be an independent copy of X, and $\tilde{X} = X - X'$. It is obvious that X is again a PII, and that

$$E(e^{iu\cdot\tilde{X}_t}) = |E(e^{iu\cdot X_t})|^2.$$

Now we have seen (in the proof of 4.14 for instance) that $t \rightsquigarrow |E(e^{iu\cdot X_t})|$ is non-increasing, and so is a function with finite variation. Then Theorem 4.14 yields that $\tilde{X}$ is a semimartingale, to which we can apply 4.15 for example. Now, it remains to "fill in" the technical details.

b) Call ν the compensator of the measure μ^X. For each Borel nonnegative bounded function g which equals 0 in a neighbourhood of 0,

$$g * \mu_t^X - g * \mu_s^X = \sum_{s<r\le t} g(\Delta X_r)$$

is clearly independent of the σ-field $\mathscr{F}_s$, so $g * \mu^X$ is an increasing locally bounded PII: therefore by 4.15 its compensator $g * \nu$ should be deterministic. Since ν is completely determined by the processes $g * \nu$ when g ranges through $\mathscr{C}^+(\mathbb{R}^d)$ (see 2.20), it follows that ν is a deterministic measure.

Moreover if $A \in \mathscr{R}^d$, then 1.17 yields

$$\nu(\{t\} \times A) = P(\Delta X_t \in A \setminus \{0\} | \mathscr{F}_{t-}) = P(\Delta X_t \in A \setminus \{0\}),$$

so 4.17 holds, and $J = \{t: \nu(\{t\} \times \mathbb{R}^d) > 0\}$ is the set of fixed times of discontinuity for X.

c) Let (T_n) be a sequence of stopping times that exhausts the jumps of X. For each n one has

$$P(\Delta X'_{T_n} \neq 0, T_n \notin J) = \int P(d\omega) 1_{J^c}(T_n(\omega)) P(\Delta X'_{T_n(\omega)}(\cdot) \neq 0) = 0$$

because X' has the same fixed times of discontinuity than X. Therefore X and X' have a.s. no common jump outside J, and so $|h(x)|^2 1_{J^c} * \mu^X \le |h|^2 * \mu^{\tilde{X}}$. Now, if $\tilde{\nu}$ is the third characteristic of $\tilde{X}$ we have $E(|h|^2 * \mu_t^{\tilde{X}}) = |h|^2 * \tilde{\nu}_t < \infty$ for all $t < \infty$, and it follows that

5.24 $$|h|^2 1_{J^c} * \nu_t = E(|h|^2 1_{J^c} * \mu_t^X) \le |h|^2 * \tilde{\nu}_t < \infty.$$

Set $Y^n = h(x) 1_{J^c} 1_{\{|x|>1/n\}} * \mu^X$, which is a PII with finite variation admitting the following compensator: $Y^{n,p} = h(x) 1_{J^c} 1_{\{|x|>1/n\}} * \nu$, which is deterministic and continuous (the latter by definition of J). So if $M^n = Y^n - Y^{n,p}$, we get $\Delta M^n = h(\Delta X) 1_{J^c} 1_{\{|\Delta X|>1/n\}}$ and I.4.53 yields for $m \ge n$:

$$[M^{m,i} - M^{n,i}, M^{m,i} - M^{n,i}] = |h^i(x)|^2 1_{J^c} 1_{\{1/m<|x|\le 1/n\}} * \mu^X.$$

Then Doob's inequality (I.1.43) and I.4.50b yield

$$\begin{aligned} E(\sup_{s\le t} |M_s^m - M_s^n|^2) &\le 4 \sum_{i\le d} E[(M_t^{m,i} - M_t^{n,i})^2] \\ &= 4 \sum_{i\le d} E([M^{m,i} - M^{n,i}, M^{m,i} - M^{n,i}]_t) \\ &= 4 |h|^2 1_{J^c} 1_{\{1/m<|x|\le 1/n\}} * \nu_t, \end{aligned}$$

which goes to 0 as $m, n \uparrow \infty$ by 5.24. Therefore (M^n) converges to a limit M, in L^2, uniformly on each interval $[0, t]$. Hence M is a martingale, and is càdlàg with $\Delta M = h(\Delta X) 1_{J^c}$; moreover with the notation $X(h)$ and $\check{X}(h)$ of 2.4, we prove exactly as in (b) that $X(h) - Y^n$ and so $X(h) - M^n$ are also PII's. Hence,

5.25 $$\begin{cases} Z = X(h) - M \text{ is a PII} \\ X - Z = \check{X}(h) + M \text{ is a semimartingale} \\ \Delta Z = h(\Delta X) 1_J \quad (\text{so } |\Delta Z| \text{ is bounded}). \end{cases}$$

d) Repeat the procedure (a): take an independent copy Z' of Z, and set $\tilde{Z} = Z - Z'$; then $\tilde{Z}$ is a PII-semimartingale, and if $\tilde{v}$ now denotes its third characteristic, $|x|^2 * \tilde{v}_t < \infty$ (because $|x|^2 \wedge 1 * \tilde{v}_t < \infty$ and $\tilde{v}([0,t] \times \cdot)$ only charges a bounded subset of $\mathbb{R}^d$). Then 4.17 applied to $\tilde{Z}$ yields $\sum_{s \le t} E(|\Delta \tilde{Z}_s|^2) \le |x|^2 * \tilde{v}_t < \infty$.

Let $U_s = \Delta Z_s - E(\Delta Z_s)$. Then $E(|\Delta \tilde{Z}_s|^2) = 2E(|U_s|^2)$ by construction of $\tilde{Z}$, hence

5.26 $$\sum_{s \le t} E(|U_s|^2) < \infty.$$

Now we set $N_t^n = \sum_{s \le t, s \in J_n} U_s$, where (J_n) is an increasing sequence of finite sets with $\bigcup J_n = J$. Since Z is a PII, N^n is obviously a martingale, and for $m \ge n$,

$$[N^{m,i} - N^{n,i}, N^{m,i} - N^{n,i}]_t = \sum_{s \le t, s \in J_m \backslash J_n} (U_s^i)^2.$$

Hence exactly like in (c) we obtain

$$E\left(\sup_{s \le t} |N_s^m - N_s^n|^2\right) \le 4 \sum_{s \le t, s \in J_m \backslash J_n} E(|U_s|^2),$$

which goes to 0 as $m, n \uparrow \infty$ by 5.26. Hence, again as in (c), (N^n) converges to a limit N in L^2, uniformly on each interval $[0,t]$, and N is a martingale, and $\Delta N_s = U_s$ if $s \in J$, and $Z - N$ is a PII. In view of 5.25, we obtain:

5.27 $$\begin{cases} W = Z - N \text{ is a PII} \\ X - W = \check{X}(h) + M + N \text{ is a semimartingale} \\ \Delta W_t = 1_J(t) E[h(\Delta X_t)]. \end{cases}$$

e) Repeat (a) once more: take an independent copy W' of W, and set $\tilde{W} = W - W'$; then $\tilde{W}$ is a PII-semimartingale, and it is continuous because $\Delta \tilde{W}_t = \Delta W_t - \Delta W_t' = 0$ by 5.27. Therefore we deduce from 4.16 applied to $\tilde{W}$ (and in which $v = 0$ because $\tilde{W}$ is continuous) that each $\tilde{W}_t$ is a Gaussian variable. Hence $\tilde{W}_t = W_t - W_t'$ is square-integrable, thus so is W_t. Set $F_t = E(W_t)$. Then the process $\hat{N} = W - F$ has again independent increments, and $E(\hat{N}_t - \hat{N}_s | \mathscr{F}_s) = 0$ if $s \le t$: hence that process $\hat{N}$ admits a càdlàg version. Since W is càdlàg, then F is indeed a càdlàg function. Finally, due to 5.27, we obtain:

$$\begin{cases} X - F = \hat{N} + \check{X}(h) + M + N \text{ is a semimartingale} \\ F \text{ is a càdlàg function with } F_0 = 0. \end{cases}$$

Then $Y = X - F$ is also certainly a PII, and we are finished. □

Proof of Theorem 5.2. In view of 5.1 and 5.14, part (a) has already been shown. In order to prove (b), let us start with (B, C, v) fulfilling 5.3, 5.4 and 5.5.

Firstly, if $J = \{t: v(\{t\} \times \mathbb{R}^d) > 0\}$ then 5.5(i, iii) implies that $(|x|^2 \wedge 1) 1_{J^c} * v_t < \infty$ for all t. Therefore if $g_u(x) = e^{iu \cdot x} - 1 - iu \cdot h(x)$, we have $|g_u| 1_{J^c} * v_t < \infty$ and

$$h(u)_{s,t} = \exp\{iu\cdot(B_t - B_s) - \frac{1}{2}u\cdot(C_t - C_s)\cdot u + \int_s^t \int_{\mathbb{R}^d} g_u(x) 1_{J^c}(s) \nu(ds \times dx)\}$$

is well-defined, and is the characteristic function of an infinitely divisible distribution on $\mathbb{R}^d$ (Lévy-Khintchine formula). Secondly, we define ν'' by 5.19: then ν'' is related to ν as ν is to ν' in the proof of 5.14, with $-\Delta B$ instead of ΔA, and so 5.20 reads:

$$\begin{aligned}\alpha(u,r) &:= e^{-iu\cdot\Delta B_r}[1 + \nu(\{r\} \times (e^{iu\cdot x} - 1))] = 1 + \nu''(\{r\} \times (e^{iu\cdot x} - 1)) \\ &= 1 + \nu''(\{r\} \times g_u) - iu\cdot\nu''(\{r\} \times h)\end{aligned}$$

and $u \rightsquigarrow \alpha(u,r)$ is the characteristic function of a random variable, and thus so is any finite product $u \rightsquigarrow \prod_{i\le n} \alpha(u,r_i)$.

Moreover, 5.19 and 5.5(i, iii, iv) imply $(|x|^2 \wedge 1) * \nu''_t < \infty$ and $\sum_{s\le t} |\nu''(\{s\} \times h)| < \infty$, hence for each $\theta > 0$ there is a constant C_θ with $\sum_{r\le t} |\alpha(u,r) - 1| \le C_\theta$ for all u with $|u| \le \theta$. Therefore the (possibly) infinite product

$$h'(u)_{s,t} = \prod_{s<r\le t} \alpha(u,r) = \prod_{s<r\le t, r\in J} \alpha(u,r)$$

converges and defines a continuous function $h'(\cdot)_{s,t}$, which thus is again a characteristic function. We deduce that the right-hand side of 4.16, namely $g(u)_{s,t} = h(u)_{s,t} h'(u)_{s,t}$, is well defined and is a characteristic function. Furthermore $g(u)_{s,t}$ is trivially càdlàg in s and t for $s \le t$.

Now, Kolmogorov's extension theorem yields a process X' on some space $(\Omega, \mathscr{F}, P)$, such that $X'_0 = 0$ and that for $0 = t_0 < \cdots < t_p$, $u_j \in \mathbb{R}^d$,

$$E\left[\exp i \sum_{1\le j\le p} u_j\cdot(X'_{t_j} - X'_{t_{j-1}})\right] = \prod_{1\le j\le p} g(u_j)_{t_{j-1},t_j}. \tag{1}$$

Note that X' has independent increments by construction. We set $\mathscr{F}_t^0 = \sigma(X'_s\colon s \le t)$, $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0$. If $T(u) = \inf(t\colon g(u)_{0,t} = 0)$ then $g(u)_{0,s} \ne 0$ for all $s < T(u)$. Hence we easily deduce from (1) and from a monotone class argument that $Z'(u)_t = [\exp iu\cdot X'_t]/g(u)_{0,t}$, defined for $t < T(u)$, satisfies $E(Z'(u)_t|\mathscr{F}_s^0) = Z'(u)_s$ for $s \le t < T(u)$. The path-regularity properties of martingales then imply that there exists a càdlàg martingale $(Z(u)_t)_{0\le t<T(u)}$, relative to the completed filtration $\mathbf{F}^P$, with $Z(u)_t = Z'(u)_t$ a.s.

On the other hand, for each $N \in \mathbb{N}^*$ there is $\theta_N > 0$ such that $|u| \le \theta_N \Rightarrow g(u)_{0,N} \ne 0$, hence $|u| \le \theta_N \Rightarrow T(u) > N$. Therefore if $\Omega_0 = \bigcap_{N\in\mathbb{N}^*} \{\omega\colon Z'(u)_t(\omega) = Z(u)_t(\omega)$ for all $t \in \mathbb{Q} \cap [0,N]$, all $u \in \mathbb{Q}^d$ with $|u| \le \theta_N\}$, then $P(\Omega_0) = 1$.

If $\omega \notin \Omega_0$ we set $X_t(\omega) = 0$ for all t. If $\omega \in \Omega_0$, and since $t \rightsquigarrow g(u)_{0,t}$ is càdlàg, there obviously exists a càdlàg function $X_\cdot(\omega)$ such that $\exp iu\cdot X_t(\omega) = g_{0,t}(u) Z(u)_t(\omega)$ for all $t < N$, $u \in \mathbb{Q}^d$ with $|u| \le \theta_N$, and $X_t(\omega) = X'_t(\omega)$ for all $t \in \mathbb{Q}_+$. Then we have defined a càdlàg process X, adapted to $\mathbf{F}^P$, with $X_0 = 0$, and such that for all $t \in \mathbb{Q}_+$ it satisfies $X_t = X'_t$ a.s.

In particular, (1) holds with X instead of X', if $t_j \in \mathbb{Q}_+$. If $t_j \in \mathbb{R}_+$, there are $t_j^n \in \mathbb{Q}_+$ with $t_j^n \downarrow t_j$ as $n \uparrow \infty$. Using the right-continuity of X_t and $g(u)_{s,t}$ in s, t

allows to pass to the limit in (1) and so to prove that (1) holds for all $t_j \in \mathbb{R}_+$, with X instead of X': but then, that means that X is a PII in the sense of 4.1, which meets 4.16, and we are done. □

Proof of Theorem 5.10. In view of 5.1 and 5.14, we know that (a) ⇒ (b). That (b) completely determines (B, C, ν) follows from the fact that a deterministic martingale is constant in time, so $B, \tilde{C}, g * \nu$ are characterized by (b); then ν is uniquely determined (see the proof of 2.21) and in particular is the compensator of μ^X; finally, if we know $B, \tilde{C}, \nu$, we deduce C from 5.7.

Next we show that (b) ⇒ (a). Firstly, (b.i) implies that $X' = X - B$ is a semimartingale, with characteristics denoted by (B', C', ν'). If we closely examine the proof of 2.32 we see that in order to compute B' and ν' we do not use the fact that X is a semimartingale, but only that 2.21(i, iii) hold; but here, this is exactly (b)(i, iii), so the same computation goes through and B' and ν' are given by 2.33 in terms of B and ν (this has also been proved in parts (b) and (c) of the proof of 5.14, with $(X, Y, -A)$ instead of (X', X, B)).

We will now prove that $C' = C$. With notation 2.4 and 2.5, and similarly to part (c) of the proof of 5.14, we get:

$$\begin{aligned} M'(h) - M(h) &= X'(h) - X(h) - B' + B \\ &= X - B - \sum_{s \le \cdot} [\Delta X_s - \Delta B_s - h(\Delta X_s - \Delta B_s)] \\ &\quad - X + \sum_{s \le \cdot} [\Delta X_s - h(\Delta X_s)] - B' + B \\ &= -B' + \sum_{s \le \cdot} [h(\Delta X_s - \Delta B_s) + \Delta B_s - \Delta X_s], \end{aligned}$$

which has finite variation: hence $M'(h)^c = M(h)^c$. Moreover $\Delta M(h) = h(\Delta X) - \Delta B$ and b(iii) yields that ν is the compensator of μ^X, so exactly as in (f) of the proof of 5.14, $M(h)^d = h * (\mu^X - \nu)$ and 5.23 holds. Comparing to the definition 5.7 of $\tilde{C}$, we obtain:

$$\langle M(h)^{i,c}, M(h)^{j,c} \rangle = \langle M(h)^i, M(h)^j \rangle - \langle M(h)^{i,d}, M(h)^{j,d} \rangle = C^{ij}$$

because $\tilde{C}^{ij} = \langle M(h)^i, M(h)^j \rangle$ by (b.ii). Since $M'(h)^c = M(h)^c$, we deduce $C' = C$ from the definition of C'.

In other words, (B', C', ν') are given by 2.33, and in particular they are deterministic. Thus Theorem 4.15 yields that X' is a PII, and so is X. Finally, because of the already shown implication (a) ⇒ (b), the characteristics of X have to be (B, C, ν).

It remains to prove the equivalence (b) ⇔ (c). That (c) ⇒ (b) is trivial. Conversely, assume (b). Then $\tilde{C}^{ii}$ is the compensator of $[M(h)^i, M(h)^i]$ by (b.ii), and therefore $E([M(h)^i, M(h)^i]_t) = \tilde{C}_t^{ii} < \infty$ for all t: by I.4.50c that implies that each stopped local martingale $(M(h)^i)^t$ is indeed in $\mathscr{H}^2$; then it is a martingale, so (c.i) holds, and $M(h)^i M(h)^j - \tilde{C}^{ij}$ is also a martingale, so (c.ii) holds. Similarly if

$g \in \mathscr{C}^+(\mathbb{R}^d)$, (b.iii) implies that $g * \nu$ is the compensator of $g * \mu^X$, so $E(g * \mu_t^X) = g * \nu_t < \infty$ and thus $\sup_{s \leq t} |g * \mu_s^X|$ is integrable: we readily deduce (c.iii). $\square$

We end this section with two additional results, for further reference. The first one has already been shown in the course of the proof of 5.10, but presents some interest in itself.

5.28 **Proposition.** *Let X be a PII with characteristics (B, C, ν); then $X' = X - B$ is a PII-semimartingale whose characteristics (B', C', ν') are given by 2.33.*

5.29 **Proposition.** *Let X be a PII with characteristics (B, C, ν). Let g be a Borel nonnegative function on $\mathbb{R}^d$, vanishing on a neighbourhood of* 0. *Then $X' = g * \mu^X$ is a PII and*

$$E(e^{-X'_t}) = \exp\left\{-(1 - e^{-g})1_{J^c} * \nu_t + \sum_{s \leq t} \operatorname{Log}(1 - \nu(\{s\} \times (1 - e^{-g})))\right\}.$$

Proof. It is exactly the same than for 4.26, in which the only fact that was used is the property: $\nu = (\mu^X)^p$. $\square$

6. Processes with Conditionally Independent Increments

In this section we introduce a class of processes that is slightly more general than PII's. They will be used only when we study weak convergence of random variables (resp. processes) toward a *mixture* of infinitely divisible variables (resp. of PII's). Hence this whole section can be skipped without harm for most of the sequel.

As usual in this chapter, we start with a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. Recall first that a variable Y and a σ-field $\mathscr{H}$ are said *conditionally independent with respect to the σ-field $\mathscr{G}$* if

6.1 $$E(f(Y)Z|\mathscr{G}) = E(f(Y)|\mathscr{G})E(Z|\mathscr{G})$$

for all bounded measurable functions f and all bounded $\mathscr{H}$-measurable variable Z (as a matter of fact, it is enough to check 6.1 when f and Z are indicator functions).

6.2 **Definition.** Let $\mathscr{H}$ be a sub-σ-field of $\mathscr{F}_0$. A *process with $\mathscr{H}$-conditionally independent increments* (in short: $\mathscr{H}$-PII) is a càdlàg adapted $\mathbb{R}^d$-valued process X such that $X_0 = 0$ and that for all $0 \leq s \leq t$ the variable $X_t - X_s$ and the σ-field $\mathscr{F}_s$ are conditionally independent with respect to $\mathscr{H}$. $\square$

Remark that a PII is a $\mathscr{H}$-PII for any sub σ-field $\mathscr{H}$ of $\mathscr{F}_0$.

All that has been said in § 4c for PII's stays valid for $\mathscr{H}$-PII, provided we replace everywhere "deterministic" by "$\mathscr{H}$-measurable". More precisely, we now state the generalizations of Theorems 4.14 and 4.15. For this, we introduce the processes:

6.3 $$G(u)_t = E(\exp iu \cdot X_t | \mathscr{H}), \quad u \in \mathbb{R}^d, t \in \mathbb{R}_+.$$

We also introduce the:

6.4 *Hypothesis*: (i) There exists a transition probability $R = R(\omega, d\omega')$ from $(\Omega, \mathscr{H})$ into $(\Omega, \mathscr{F})$, which is a *regular version of the conditional probability relative to* $\mathscr{H}$.

(ii) There is an increasing family $(\mathscr{G}_t^0)_{t \in \mathbb{R}_+}$ of *separable* σ-fields, such that $\mathscr{F}_t = \bigcap_{s>t} (\mathscr{H} \vee \mathscr{G}_s^0)$. □

6.5 **Theorem.** *Assume 6.4. Let X be a d-dimensional $\mathscr{H}$-PII. Then X is a semimartingale if and only if for each $u \in \mathbb{R}^d$ one may find a version of the process $G(u)$ that is càdlàg with finite variation over compact sets.*

6.6 **Theorem.** *Assume 6.4. Let X be a d-dimensional semimartingale with $X_0 = 0$, and let $h \in \mathscr{C}_t^d$ be a truncation function. Then X is a $\mathscr{H}$-PII if and only if there exists a version (B, C, ν) of its characteristics that is $\mathscr{H}$-measurable.*

Moreover, in that case we have for all $u \in \mathbb{R}^d$, $s \leq t$:

6.7 $$\begin{aligned} &E(\exp iu \cdot (X_t - X_s) | \mathscr{H}) \\ &= \exp\left[iu \cdot (B_t - B_s) - \frac{1}{2} u \cdot (C_t - C_s) \cdot u \right. \\ &\quad \left. + \int_s^t \int_{\mathbb{R}^d} (e^{iu \cdot x} - 1 - iu \cdot h(x)) 1_{J^c}(r) \nu(dr, dx) \right] \\ &\quad \times \prod_{s<r\leq t} \left\{ e^{-iu \cdot \Delta B_r} \left[1 + \int (e^{iu \cdot x} - 1) \nu(\{r\} \times dx) \right] \right\} \end{aligned}$$

where $J = \{(\omega, t): \nu(\omega; \{t\} \times \mathbb{R}^d) > 0\}$. In particular, $G(u)_t$ equals the right-hand side of 6.7 for $s = 0$.

Note that Theorems 4.14 and 4.15 are indeed particular cases of these two theorems, at least under 6.4(ii): take $\mathscr{H}$ to be the trivial σ-field.

6.8 **Remark.** We have limited ourselves to conditional PII that are semimartingales. Needless to say, Theorems 5.1, 5.2 and 5.10 have similar extensions for conditional PII's. □

6.9 **Remark.** *Theorem 6.5 and 6.6 are still valid without 6.4.* But in that case it is necessary to entirely rework the proofs of 4.14 and 4.15, instead of just using the results of these two theorems. □

Despite the previous remark, we decided to prove these theorems only under 6.4, since the next lemma shows that it really is a weak assumption.

6.10 **Lemma.** *Let Y be a càdlàg n-dimensional process. If* **F** *is the smallest filtration to which Y is adapted, and such that $\mathscr{H} \subset \mathscr{F}_0$, and if $\mathscr{F} = \mathscr{F}_{\infty-}$, then 6.4 holds.*

Proof. 6.4(ii) is obvious, with $\mathscr{G}_t^0 = \sigma(Y_s\colon s \leq t)$. Now, we anticipate on Chapter VI: There we will show that the space $\mathbb{D}(\mathbb{R}^n)$ of all càdlàg functions $y\colon \mathbb{R}_+ \to \mathbb{R}^n$ is endowed with a topology (the Skorokhod topology) under which it is a Polish space, whose Borel σ-field $\mathscr{D}(\mathbb{R}^n)$ is the σ-field generated by all the evaluation maps: $y \rightsquigarrow y(t)$, $t \in \mathbb{R}_+$.

Then Y maps Ω into $\mathbb{D}(\mathbb{R}^n)$, and $\mathscr{F} = \mathscr{H} \vee \mathscr{G}_{\infty-}$ where $\mathscr{G}_{\infty-}$ is the pre-image of $\mathscr{D}(\mathbb{R}^n)$ by the map Y. Then $(\Omega, \mathscr{F})$ is measure-isomorphic to a subset of the product space $(\Omega, \mathscr{H}) \otimes (\mathbb{D}(\mathbb{R}^n), \mathscr{D}(\mathbb{R}^n))$. Since $\mathbb{D}(\mathbb{R}^n)$ is in particular a Blackwell space (see 1.2), we immediately deduce 6.4i. □

Before proving Theorems 6.5 and 6.6, we give some applications. Firstly, note that if X is a $\mathscr{H}$-PII, then by 6.7 the distribution of X is completely determined by the restriction of P to $\mathscr{H}$ and by (B, C, ν). Hence the following result, with the same proof than for 4.25 (note that the filtration **G** below has the structure described in Lemma 6.10, so that 6.4 holds for P and Q):

6.11 **Theorem.** *We use the notation 4.24, and we consider two probability measures P and Q on $(\Omega, \mathscr{G})$ such that:*

(i) *P and Q coincide on the σ-field $\mathscr{H}$;*

(ii) *$X_0 = 0$ P-a.s. and Q-a.s.;*

(iii) *X is a semimartingale with the same $\mathscr{H}$-measurable characteristics on the two stochastic bases $(\Omega, G, \mathbf{G}, P)$ and $(\Omega, G, \mathbf{G}, Q)$;*

Then $P = Q$.

6.12 **Example.** *Cox processes.* The classical definition of a Cox process is the following one: it is a point process $N = (N_t)_{t \geq 0}$ defined on $(\Omega, \mathscr{F}, P)$, and there exists a continuous increasing process A (usually of the form $A_t = \int_0^t \lambda_s \, ds$ or even $A_t = t\Lambda$) such that if $\mathscr{H}$ denotes the σ-field generated by the variables A_t's, then "conditionally with respect to $\mathscr{H}$" the point process N is a Poisson process with intensity $dA_t(\omega)$.

Then, if we compare this definition with 4.5 and 6.5, we immediately obtain that a point process is a Cox process if and only if it is a $\mathscr{H}$-PII for some sub σ-field $\mathscr{H} \subset \mathscr{F}_0$, and is quasi-left-continuous. □

Now we proceed to the proofs of 6.5 and 6.6. The basic tools are the following two lemmas.

6.13 **Lemma.** *Assume 6.4. Then X is a $\mathscr{H}$-PII if and only if for P-almost all ω, X is a PII under the measure $R(\omega, \cdot)$.*

Proof. Let $\hat{\mathscr{G}}_t^0$ be a countable algebra generating $\mathscr{G}_t^0$. Let also $\mathscr{C}^+(\mathbb{R}^d)$ be a countable family of functions, satisfying 2.20. Denote by Ω_0 the set of all ω such that

$$(1) \quad \int R(\omega, d\omega') 1_A(\omega') f[(X_t - X_s)(\omega')] = R(\omega, A) \int R(\omega, d\omega') f[(X_t - X_s)(\omega')]$$

for all $f \in \mathscr{C}^+(\mathbb{R}^d)$, $A \in \hat{\mathscr{G}}_s^0$, $s \leq t$, $s, t \in \mathbb{Q}_+$.

a) We first prove that X is a PII for $R(\omega, \cdot)$ if and only if $\omega \in \Omega_0$. The condition is trivially necessary. Conversely assume that $\omega \in \Omega_0$. By a montone class argument, (1) holds for all $s \leq t$, $s, t \in \mathbb{Q}_+$, $f \in \mathscr{C}^+(\mathbb{R}^d)$ and $A \in \mathscr{G}_s^0$. Since $\mathscr{H}$ is $R(\omega, \cdot)$-trivial, it also holds for all $A \in \mathscr{H} \vee \mathscr{G}_s^0$. Then if we take $s, t \in \mathbb{R}_+$, $s \leq t$, and if we write (1) for $s', t' \in \mathbb{Q}_+$ with $s' > s$ and $t' > t$, and then let $s' \downarrow s$ and $t' \downarrow t$, we see that (1) holds for all $A \in \bigcap_{s' > s} (\mathscr{H} \vee \mathscr{G}_{s'}^0) = \mathscr{F}_s$. Another monotone class argument finally yields (1) for all $s \leq t$, $A \in \mathscr{F}_s$ and f bounded Borel (recall 2.20): hence $X_t - X_s$ is independent of $\mathscr{F}_s$ under $R(\omega, \cdot)$.

b) It remains to prove that X is a $\mathscr{H}$-PII is and only if $P(\Omega_0) = 1$. For this we observe that:

$$(2) \quad \begin{cases} \text{the left-hand side of (1) is a version of } E[1_A f(X_t - X_s) | \mathscr{H}], \\ \text{the right-hand side of (1) is a version of } P(A|\mathscr{H}) E(f(X_t - X_s)|\mathscr{H}). \end{cases}$$

Then if X is a $\mathscr{H}$-PII, for all $A, f, s \leq t$ we have (1) P-a.s. Since Ω_0 is a countable intersection of sets on which (1) holds, we deduce $P(\Omega_0) = 1$.

Conversely, assume $P(\Omega_0) = 1$. We have seen in (a) that if $\omega \in \Omega_0$, then (1) holds for all $s \leq t$, f bounded Borel, $A \in \mathscr{F}_s$: then (2) gives that $\mathscr{F}_s$ and $X_t - X_s$ are $\mathscr{H}$-conditionally independent, and we are finished. □

6.14 **Lemma.** *Assume 6.4. Let M be a càdlàg adapted bounded process. Then M is a martingale for P if and only if for P-almost all ω it is a martingale for $R(\omega, \cdot)$.*

Proof. The proof is similar to that of 6.13. We use the same notation $\hat{\mathscr{G}}_t^0$, and Ω_0 is the set of all ω such that

$$(1) \qquad \int R(\omega, d\omega') 1_A(\omega') (M_t - M_s)(\omega') = 0$$

for all $A \in \hat{\mathscr{G}}_s^0$, $s \leq t$, $s, t \in \mathbb{Q}_+$.

a) M is a martingale for $R(\omega, \cdot)$ if and only if $\omega \in \Omega_0$: that is trivially necessary. Conversely if $\omega \in \Omega_0$, one shows like in 6.13 that (1) holds for all $s \leq t$, $A \in \mathscr{F}_s$, hence the claim.

b) The left-hand side of (1) is a version of $E(1_A(M_t - M_s)|\mathscr{H})$, so if M is a martingale for P then (1) holds P-a. s. Since Ω_0 is defined through a countable number of relations (1), we deduce $P(\Omega_0) = 1$. Conversely if $P(\Omega_0) = 1$, and since $\omega \in \Omega_0$ implies that (1) holds in fact for all $s \le t$, $A \in \mathscr{F}_s$, then taking expectations shows that $E(1_A(M_t - M_s)) = 0$, and so M is a martingale. □

6.15 **Corollary.** *Assume 6.4. The càdlàg adapted process X is a semimartingale with characteristics (B, C, ν) for P if and only if for P-almost all ω, it is a semimartingale with characteristics (B, C, ν) for $R(\omega, \cdot)$.*

Proof. That X is a semimartingale with characteristics (B, C, ν) is equivalent to conditions (i, ii, iii) of 2.21, with for $\mathscr{C}^+(\mathbb{R}^d)$ some countable family. Moreover, the local martingales showing in 2.21 are all locally bounded by construction. Therefore X is a semimartingale with characteristics (B, C, ν) if and only if a given countable family of processes consists only in locally bounded local martingales.

Now, by localization, 6.14 stays valid for locally bounded local martingales: hence the claim immediately follows from this extension of 6.14. □

Proofs of 6.5 and 6.6. We remark that a version of $G(u)_t$ is

$$G(u)_t(\omega) = \int R(\omega, d\omega') \exp iu \cdot X_t(\omega').$$

Hence 6.5 readily follows from 4.14 and 6.13 and 6.15. Similarly, we deduce 6.6 from 4.15 and 6.13 and 6.15, one noticed that the characteristics are $\mathscr{H}$-measurable if and only if they are $R(\omega, \cdot)$-a. s. deterministic, for P-almost all ω. □

7. Progressive Conditional Continuous PIIs

In the previous section we introduced processes with conditionally independent increments, the conditioning being w.r.t. a σ-field $\mathscr{H}$ included into the initial σ-field $\mathscr{F}_0$. Here we no longer assume that $\mathscr{H} \subset \mathscr{F}_0$, that is we make a sort of "progressive" conditioning as time increases. On the other hand we specialise the setting by considering only *continuous* processes, which in addition are local martingales.

Processes with the structure described below will be encountered as limiting processes in Section 7 of Chapter IX: they are especially useful when one deals with numerical computations of solutions of stochastic differential equations (through the Euler scheme, for example), and for statistical analysis of processes like diffusion processes when they are observed at discrete times only.

To keep things simple, we will assume a special structure: we start with a given stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, \mathbf{P})$, and we "extend" it in the following way:

7.1 **Definition.** We call *extension* of $\mathscr{B}$ another stochastic basis $\widetilde{\mathscr{B}} = (\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbf{F}} = (\widetilde{\mathscr{F}}_t)_{t\geq 0}, \widetilde{P})$ constructed as follows: we have an auxiliary filtered space $(\Omega', \mathscr{F}', \mathbf{F}' = (\mathscr{F}_t')_{t\geq 0})$ such that each σ–field $\mathscr{F}_{t-}'$ is separable, and a transition probability $Q_\omega(d\omega')$ from $(\Omega, \mathscr{F})$ into $(\Omega', \mathscr{F}')$, and we set

7.2
$$\begin{gathered}\widetilde{\Omega} = \Omega \times \Omega', \quad \widetilde{\mathscr{F}} = \widetilde{\mathscr{F}} \otimes \mathscr{F}', \quad \widetilde{\mathscr{F}}_t = \cap_{s>t} \mathscr{F}_s \otimes \mathscr{F}_s', \\ \widetilde{P}(d\omega, d\omega') = P(d\omega) Q_\omega(d\omega').\end{gathered}$$

Further, the extension $\widetilde{\mathscr{B}}$ is called *very good* if, for all $t \geq 0$ and all $A' \in \mathscr{F}_t'$, the map $\omega \mapsto Q_\omega(A')$ is P–a.s. equal to an $\mathscr{F}_t$–measurable random variable. □

We denote by $\widetilde{E}$ the expectation w.r.t. $\widetilde{P}$. Any variable or process defined on either Ω or Ω' can be considered as a variable or process on $\widetilde{\Omega}$, by setting for example $X(\omega, \omega') = X(\omega)$ or $X(\omega, \omega') = X(\omega')$. The terminology "very good extension", as well as the following lemma, come from [109]:

7.3 **Lemma.** *The extension $\widetilde{\mathscr{B}}$ is very good if and only if all martingales on the basis $\mathscr{B}$ are also martingales on $\widetilde{\mathscr{B}}$.*

Proof. Suppose first that $\widetilde{\mathscr{B}}$ is a very good extension. Let M be a martingale on $\mathscr{B}$. For all $s \geq t \geq 0$ and $A \in \mathscr{F}_t$, $A' \in \mathscr{F}_t'$, we have by 7.2:

$$\widetilde{E}(1_{A\times A'} M_s) = E(1_A M_s Q(A')) = E(1_A M_t Q(A')) = \widetilde{E}(1_{A\times A'} M_t).$$

Therefore $M_t = \widetilde{E}(M_s | \mathscr{F}_t \otimes \mathscr{F}_t')$. Applying this to a sequence t_n decreasing strictly to r, and using the right-continuity of M, we deduce that $M_r = \widetilde{E}(M_s | \widetilde{\mathscr{F}}_r)$ for all $r < s$, so M is a martingale on $\widetilde{\mathscr{B}}$.

Conversely suppose that any martingale M on $\mathscr{B}$ is also a martingale on $\widetilde{\mathscr{B}}$. Let $A' \in \mathscr{F}_t'$ and $A \in \mathscr{F}$, and call Z and M the martingales on $\mathscr{B}$ such that $Z_s = E(Q(A') | \mathscr{F}_s)$ and $M_s = P(A | \mathscr{F}_s)$. We have

$$E(1_A Q(A')) = \widetilde{E}(1_A\, 1_{A'}) = \widetilde{E}(M_t 1_{A'})$$

because $\Omega \times A' \in \widetilde{\mathscr{F}}_t$ and M is a martingale on $\widetilde{\mathscr{B}}$ by hypothesis. We deduce that

$$E(1_A Q(A')) = E(M_t Q(A')) = E(M_t Z_t) = E(1_A Z_t).$$

This, being true for all $A \in \mathscr{F}$, yields that $Z_t = Q(., A')$ P–a.s.: so the extension $\widetilde{\mathscr{B}}$ is very good. □

7.4 **Definition.** A process Z on the extension $\widetilde{\mathscr{B}}$ is called an *$\mathscr{F}$–progressive conditional martingale* (resp. an *$\mathscr{F}$–progressive conditional PII*) if it is adapted to $\widetilde{\mathbf{F}}$ and if for P–almost all ω the process $Z(\omega, .)$ is a martingale (resp. a PII) on the basis $\mathscr{B}_\omega = (\Omega', \mathscr{F}', \mathbf{F}', Q_\omega)$. $\square$

Note that if $(\Omega, \mathscr{F}, \mathbf{F}, \mathbf{P})$ and $\mathscr{H}$ satisfy Hypothesis 6.4, any $\mathscr{H}$–PII is a progressive conditional PII in the above sense: set $\mathscr{H}_t = \mathscr{H}$ for all t and $\mathscr{B} = (\Omega, \mathscr{F}, (\mathscr{H}_t), P)$ and $(\Omega', \mathscr{F}', \mathbf{F}') = (\Omega, \mathscr{F}, \mathbf{F})$ and $Q_\omega(d\omega') = R(\omega, d\omega')$. But of course if we start with a basis $\mathscr{B}$ having a genuine filtration instead of a constant one, then a progressive conditional PII is not an $\mathscr{H}$–PII for any given σ–field $\mathscr{H}$, in general.

Let us come back to the general situation of this Section. For notational simplicity let $\mathscr{M}_b$ denote the set of all bounded martingales on $\widetilde{\mathscr{B}}$. For any q–dimensional locally square-integrable martingale Z on $\widetilde{\mathscr{B}}$, the $q \times q$–dimensional process $(\langle Z^i, Z^j\rangle)_{1\le i,j\le d}$ is denoted by $\langle Z, Z\rangle$; this a process taking its values in the set $\mathscr{S}_q$ of all $q \times q$ symmetric nonnegative matrices, and it is non-decreasing for the strong order of this set.

7.5 **Proposition.** *Let $\underset{\sim}{Z}$ be a continuous adapted q–dimensional process on the very good extension $\widetilde{\mathscr{B}}$, with $Z_0 = \underset{\sim}{0}$. The following statements are equivalent:*

(i) *Z is a local martingale on $\widetilde{\mathscr{B}}$, orthogonal to all elements of $\mathscr{M}_b$, and the bracket $\langle Z, Z\rangle$ is adapted to* $\mathbf{F}$.

(ii) *Z is an $\mathscr{F}$–progressive conditional martingale and PII.*

In this case, the $\mathscr{F}$–conditional law of Z is characterized by the process $\langle Z, Z\rangle$ (i.e., for P–almost all ω, the law of $Z(\omega, .)$ under Q_ω depends only on the function $t \rightsquigarrow \langle Z, Z\rangle_t(\omega)$).

Proof. a) We first prove that, if each Z_t is $\widetilde{P}$–integrable, then Z is an $\mathscr{F}$–progressive conditional martingale iff it is a martingale on $\widetilde{\mathscr{B}}$ which is orthogonal to all elements of $\mathscr{M}_b$ (which by Lemma 7.3 are also martingales on $\widetilde{\mathscr{B}}$). For this, we can and will assume that Z is 1-dimensional.

Let $t \le s$ and let U, U' be bounded measurable function on $(\Omega, \mathscr{F}_t)$ and $(\Omega', \mathscr{F}_t')$ respectively, and $M \in \mathscr{M}_b$. Then

$$7.6 \quad \widetilde{E}(UU'M_sZ_s) = \int P(d\omega)U(\omega)M_s(\omega) \int Q_\omega(d\omega')U'(\omega')Z_s(\omega, \omega'),$$

$$7.7 \quad \widetilde{E}(UU'M_tZ_t) = \int P(d\omega)U(\omega)M_t(\omega) \int Q_\omega(d\omega')U'(\omega')Z_t(\omega, \omega').$$

Assume first that Z is an $\mathscr{F}$–progressively conditional martingale. For P–almost all ω we have

$$\int Q_\omega(d\omega')U'(\omega')Z_s(\omega, \omega') = \int Q_\omega(d\omega')U'(\omega')Z_t(\omega, \omega'),$$

and the latter is $\mathscr{F}_t$–measurable as a function of ω because the extension is very good. Since $M \in \mathscr{M}_b$, we deduce that 7.6 and 7.7 are equal: thus MZ is a martingale on $\widetilde{\mathscr{B}}$, so Z is a martingale (take $M \equiv 1$), orthogonal to all elements of $\mathscr{M}_b$.

Next we prove the sufficient condition. Take V bounded $\mathscr{F}_s$–measurable, and consider the martingale $M_r = E(V|\mathscr{F}_r)$. With the notation above we have equality between 7.6 and 7.7, and further in 7.7 we can replace $M_t(\omega)$ by $M_s(\omega) = V(\omega)$ because the last integral is $\mathscr{F}_t$–measurable in ω. Taking $U = 1$, we get

$$\int P(d\omega)V(\omega) \int Q_\omega(d\omega')U'(\omega')Z_s(\omega, \omega')$$
$$= \int P(d\omega)V(\omega) \int Q_\omega(d\omega')U'(\omega')Z_t(\omega, \omega').$$

Hence for P–almost ω, $Q_\omega(U'Z_s(\omega, .)) = Q_\omega(U'Z_t(\omega, .))$. Using the separability of the σ–field $\mathscr{F}'_{t-}$ and the continuity of Z, we have this relation P–almost surely in ω, simultaneously for all $t \leq s$ and all $\mathscr{F}'_{t-}$–measurable variable U': this gives the $\mathscr{F}$–progressive conditional martingality for Z.

b) Assume that (i) holds. If $Y^{ij} = Z^iZ^j - \langle Z^i, Z^j\rangle$ and $Y = (Y^{ij})_{1\leq i,j\leq q}$, a simple application of Ito's formula and the fact that Z is continuous show that, since Z is orthogonal to all $M \in \mathscr{M}_b$, the same holds for Y. Each $T_n = \inf(t : \sum_{i,j} |\langle Z^i, Z^j\rangle_t| > n)$ is a stopping time on $\mathscr{B}$, and $T_n \uparrow \infty$ as $n \to \infty$. Then $Z(n)_t = Z_{t\wedge T_n}$ and $Y(n)_t = Y_{t\wedge T_n}$ are continuous $\mathbf{F}$–martingale, orthogonal to all $M \in \mathscr{M}_b$, and obviously $|Z(n)_t|$ and $|Y(n)_t|$ are integrable: by (a), and by letting $n \uparrow \infty$, we deduce that for P–almost all ω, under Q_ω the process $Z(n)(\omega, .)$ is a continous martingale with deterministic bracket $\langle Z, Z\rangle(\omega)$: thus by Theorem 4.15 it is, under Q_ω, a PII whose law is entirely determined by the function $t \mapsto \langle Z, Z^*\rangle_t(\omega)$: hence (ii) and the last claim.

c) Assume now (ii). There is a P–full set $A \in \mathscr{F}$ such that for all $\omega \in A$, the process $Z(\omega, .)$ is both a PII and a martingale on $\mathscr{B}_\omega$. With $F_t^{ij}(\omega) = \int Q_\omega(d\omega')Z_t^i(\omega, \omega')Z_t^j(\omega, \omega')$, the process $(Z^iZ^j)(\omega, .) - F^{ij}(\omega)$ is a martingale on $\mathscr{B}_\omega$ for $\omega \in A$: that is, $Z^iZ^j - F^{ij}$ is an $\mathscr{F}$–progressively conditional martingale. By localizing at the $\mathbf{F}$–stopping times $T_n = \inf(t : \sum_{i,j} |F_t^{ij}| > n)$ and by (a), we deduce that Z and $Z^iZ^j - F^{ij}$ are local martingales on $\widetilde{\mathscr{B}}$, orthogonal to all $M \in \mathscr{M}_b$. Since $F = (F^{ij})_{i,j}$ is continuous, adapted to $\mathbf{F}$, and of bounded variation (since it is non-decreasing for the strong order in $\mathscr{S}_q$), it follows that it is a version of $\langle Z, Z\rangle$, hence we have (i). □

Let now M be a continous d–dimensional local martingale, and let $\mathscr{M}_b(M^\perp)$ be the class of all elements of $\mathscr{M}_b$ which are orthogonal to M (i.e., to all components of M). If $d \geq 2$, for the stochastic integral $u \cdot M$ below we anticipate on §4a of Chapter III and use freely Theorem III.4.5: no circular argument involved!

7.8 **Definition.** A q–dimensional process Z on the extension is called an *M–biased $\mathscr{F}$–progressive conditional martingale-PII* if it can be written as $Z = Z' + u \cdot M$, where Z' is an $\mathscr{F}$–progressive conditional martingale and PII, and u is a predictable $\mathbb{R}^q \otimes \mathbb{R}^d$–valued process on $\mathscr{B}$.

7.9 **Proposition.** *Let $\widetilde{Z}$ be a continuous adapted q–dimensional process on the very good extension $\widetilde{\mathscr{B}}$, with $Z_0 = 0$. The following statements are equivalent:*

(i) Z is a local martingale on the extension, orthogonal to all elements of $\mathscr{M}_b(M^\perp)$, and the brackets $\langle Z, Z\rangle$ and $\langle Z, M\rangle$ are adapted to $\mathbf{F}$.

(ii) Z is an M-biased $\mathscr{F}$–progressive conditional martingale-PII.

Then the $\mathscr{F}$–conditional law of Z is characterized by the processes M, $\langle Z, Z\rangle$ and $\langle Z, M\rangle$.

Proof. Under either (i) or (ii), Z and M are continous local martingales on $\widetilde{\mathscr{B}}$ (use the fact that the extension is very good, and use Proposition 7.5 and Definition 7.8 under (ii)). We write $F = \langle Z, Z\rangle$, $G = \langle Z, M\rangle$ and $H = \langle M, M\rangle$, with components F^{ij}, G^{ik} and H^{kl} for $1 \le i, j \le q$ and $1 \le k, l \le d$.

If (ii) holds, Definition 7.8 and Proposition 7.5 yield for all $N \in \mathscr{M}_b$:

$$G^{ik} = \sum_{l=1}^{d} u^{il} \cdot H^{lk}, \qquad F^{ij} = \langle Z'^i, Z'^j\rangle + \sum_{k,l=1}^{d} u^{ik}u^{jl} \cdot H^{kl},$$

$$\langle Z^i, N\rangle = \sum_{k=1}^{d} u^{ik} \cdot \langle M^k, N\rangle.$$

Then (i) follows, and further u and $\langle Z', Z'\rangle$ are determined by F, G and H. Since $u \cdot M$ is $\mathscr{F}$–measurable, the last claim follows from Proposition 7.5 again.

Conversely, assume (i). There are a continuous increasing process A and $\mathbf{F}$–predictable processes f, g, h with values in $\mathbb{R}^q \otimes \mathbb{R}^q$, $\mathbb{R}^q \otimes \mathbb{R}^d$ and $\mathbb{R}^d \otimes \mathbb{R}^d$ respectively, such that $F_t = \int_0^t f_s dA_s$, $G_t = \int_0^t g_s dA_s$ and $H_t = \int_0^t h_s dA_s$ with matrix notation: the transpose of a matrix f is $f^\star$. The process (M, Z) is a continuous local martingale on the extension, with bracket $K = k \cdot A$, where $k = \begin{pmatrix} h & g^* \\ g & f \end{pmatrix}$. By triangularization we may write $k = zz^*$, where

7.10 $$z = \begin{pmatrix} v & 0 \\ uv & w \end{pmatrix},$$

for suitable matrix-valued $\mathbf{F}$–predictable processess u, v and w, and $h = vv^*$, $g = uvv^*$ and $f = uvv^*u^* + ww^*$. Set $Y = u \cdot M$ and $Z' = Z - Y$. Since $\widetilde{\mathscr{B}}$ is a very good extension, Z' is a local martingale on $\widetilde{\mathscr{B}}$, while a simple calculation shows that $\langle Z', Z'^*\rangle = (ww^*) \cdot A$, which is adapted to $\mathbf{F}$. Further, $\langle Z', N\rangle = \langle Z, N\rangle - u \cdot \langle M, N\rangle$: first this implies that $\langle Z', N\rangle = 0$ if $N \in \mathscr{M}_b(M^\perp)$ (since then $\langle Z, N\rangle = 0$ by hypothesis), second this implies that when $N = \alpha \cdot M$ we have $\langle Z', N\rangle = (g\alpha^* - uvv^*\alpha) \cdot A = 0$. Thus Z' is orthogonal to all $N \in \mathscr{M}_b$, and it is an $\mathscr{F}$–progressive conditional martingale-PII by Proposition 7.5. Hence (ii) holds. □

We end this section with a construction of $\mathscr{F}$–progressive conditional martingale-PII or M–biased conditional martingale-PPI with given brackets.

In Proposition 7.5, the process $\langle Z, Z\rangle$ is a continuous adapted non-decreasing $\mathcal{S}_q$–valued process on $\mathcal{B}$, null at 0. In Proposition 7.9, the bracket of (M, Z) is a continuous adapted non-decreasing $\mathcal{S}_{d+q}$–valued process on $\mathcal{B}$, null at 0. Conversely we have:

7.11 **Proposition.** a) *Let F be a continuous adapted nondecreasing $\mathcal{S}_q$–valued process, with $F_0 = 0$, on $\mathcal{B}$. There exists a continuous $\mathcal{F}$–progressive conditional martingale and PII Z on a very good extension, such that $\langle Z, Z\rangle = F$.*

b) *Let K be a continuous adapted nondecreasing $\mathcal{S}_{d+q}$–valued process with $K_0 = 0$, and let M be a continuous d–dimensional local martingale with $\langle M^i, M^j\rangle = K^{ij}$ for $1 \leq i, j \leq d$, on the basis $\mathcal{B}$. There exists a continuous M–biased $\mathcal{F}$–progressive conditional martingale-PII Z on a very good extension, such that $\langle Z^i, M^j\rangle = K^{d+i,j}$ for $1 \leq i \leq q$, $1 \leq j \leq d$, and $\langle Z^i, Z^j\rangle = K^{d+i,d+j}$ for $1 \leq i, j \leq q$.*

Of course (a) is a particular case of (b) (take $M = 0$), but in the proof below (b) is obtained as a consequence of (a).

Proof. a) Let $(\Omega', \mathcal{F}', \mathbf{F}')$ be the canonical space of all $\mathbb{R}^d$–valued continuous functions on $\mathbb{R}_+$, with the canonical process $Z_t(\omega') = \omega'(t)$ and the filtration generated by Z. For each ω, denote by Q_ω the unique probability measure on $(\Omega', \mathcal{F}')$ under which Z is a centered Gaussian process with covariance $\int Z_t^i Z_s^j dQ_\omega = F_{s\wedge t}^{ij}(\omega)$. Then Z is a PII and a martingale under each Q_ω: Defining $\mathcal{B}$ by 7.1 gives the result.

b) As in the previous proof, we can write $K = k \cdot A$ for some continuous adapted increasing process A and some predictable process $k = zz^*$ with z as in 7.10. By (a) we have a continuous $\mathcal{F}$–progressive conditional martingale-PII Z' on a very good extension, with $\langle Z', Z'\rangle_t = (ww^*) \cdot A$. Then $Z = Z' + u \cdot M$ satisfies our requirements. □

We even have a more "concrete" way of constructing Z above, when K is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}_+$. Let $(\Omega', \mathcal{F}', \mathbf{F}')$ be the canonical space of all $\mathbb{R}^q$–valued continuous functions on $\mathbb{R}_+$, with the canonical process $W_t(\omega') = \omega'(t)$ and the filtration generated by W, and call P' the unique probability measure on $(\Omega', \mathcal{F}')$ under which the process W is a standard q–dimensional Wiener process. Then we construct the "canonical Wiener extension" $\widetilde{\mathcal{B}}$ of $\mathcal{B}$ as in 7.1, with $Q_\omega = P'$ for all $\omega \in \Omega$. This is clearly a very good extension of $\mathcal{B}$.

7.12 **Proposition.** *Let K and M be as in Proposition 7.11, and assume that $K_t = \int_0^t k_s ds$ with k predictable $\mathcal{S}_{d+q}$–valued. Then we can choose a version of k of the form $k = zz^*$ with $z = \begin{pmatrix} v & 0 \\ uv & w \end{pmatrix}$, and on the canonical q–dimensional Wiener extension $\widetilde{\mathcal{B}}$ of $\mathcal{B}$ the process*

7.13 $$Z = u \cdot M + w \cdot W$$

is a continuous M–biased $\mathscr{F}$–progressive conditional martingale-PII, such that $\langle Z^i, M^j \rangle = K^{d+i,j}$ *for* $1 \le i \le q$ *and* $1 \le j \le d$, *and* $\langle Z^i, Z^j \rangle = K^{d+i,d+j}$ *for* $1 \le i, j \le q$.

Proof. The first claim has already been proved. 7.13 defines a continuous q–dimensional local martingale on the canonical Wiener extension, and a simple computation shows that it has the required brackets. □

8. Semimartingales, Stochastic Exponential and Stochastic Logarithm

§ 8a. More About Stochastic Exponential and Stochastic Logarithm

The *Stochastic* or *Doléans-Dade exponential* of a (real-valued or complex-valued) semimartingale X was introduced in § I.4f, as a solution of the (integral) equation

8.1 $$Y = 1 + Y_- \cdot X \quad \text{(or: } dY = Y_- dX \text{ and } Y_0 = 1).$$

Further, Theorem I.4.61 states that, in the case of a real-valued process X, Equation 8.1 has a unique (up to evanescence) *càdlàg* adapted solution $Y = \mathscr{E}(X)$ (in the class of semimartingales) given by

8.2 $$\mathscr{E}(X)_t = e^{X_t - X_0 - \frac{1}{2}\langle X^c, X^c\rangle_t} \prod_{0<s\le t} (1 + \Delta X_s) e^{-\Delta X_s}.$$

Also, if $\Delta X \neq -1$ identically, then both processes $\mathscr{E}(X)$ and $\mathscr{E}(X)_-$ are non-vanishing, and in this case, the mapping $X \mapsto \mathscr{E}(X)$ can be inverted: if $Y = \mathscr{E}(X)$ then in analogy to real calculus ($Y = e^X \Leftrightarrow X = \log Y$) we will call its inverse X the *stochastic logarithm* of Y, and write it as $X = \mathscr{L}\text{og}\,(Y)$. This is the object of the following:

8.3 **Theorem.** *Let Y be a real-valued semimartingale such that the two processes Y_- and Y do not vanish. Then the process*

8.4 $$X = \frac{1}{Y_-} \cdot Y \qquad \text{(or: } dX = \tfrac{1}{Y_-}\, dY \text{ and } X_0 = 0),$$

also denoted by $\mathscr{L}\text{og}\,(Y)$ and called the stochastic logarithm of Y, is the unique (up to evanescence) semimartingale X such that

8.5 $$Y = Y_0\, \mathscr{E}(X), \qquad X_0 = 0.$$

Moreover it satisfies $\Delta X \neq -1$ identically, and we also have

8.6 $$\mathscr{L}\text{og}\, Y_t = \log\left|\frac{Y_t}{Y_0}\right| + \frac{1}{2Y_-^2}\cdot\langle Y^c, Y^c\rangle_t - \sum_{s\leq t}\left(\log\left|1+\frac{\Delta Y_s}{Y_{s-}}\right| - \frac{\Delta Y_s}{Y_{s-}}\right).$$

It is worth mentionning that the process Y needs not be positive for $\mathscr{L}\text{og}\,(Y)$ to exist, in accordance with the fact that the stochastic exponential $\mathscr{E}(X)$ may take negative values.

Proof. The non-vanishing assumptions for Y_- and Y imply that $T_n = \inf(t : |Y_{t-}| \leq 1/n)$ increases to $+\infty$, hence the process $1/Y_-$ is locally bounded and the stochastic integral in 8.4 makes sense. Let $Y' = Y/Y_0$: then $Y'_0 = 1$, and X defined by 8.4 is also $X = (1/Y'_-)\cdot Y'$. Hence, since Y'_- does not vanish,

$$1 + Y'_-\cdot X = 1 + Y'_-\cdot\left(\frac{1}{Y'_-}\cdot Y'\right) = 1 + \left(Y'_-\frac{1}{Y'_-}\right)\cdot Y' = Y'.$$

Thus $Y' = \mathscr{E}(X)$ and 8.5 holds. Further, $\Delta X = \Delta Y/Y_-$ is clearly nowhere equal to -1.

Let X' be any other semimartingale satisfying 8.5. With $Y' = Y/Y_0$ again, 8.5 yields $Y' = \mathscr{E}(X')$, hence $Y' = 1 + Y'_-\cdot X'$, hence $Y = Y_0 + Y_-\cdot X'$. Then since $X'_0 = X_0 = 0$,

$$X' = \frac{Y_-}{Y_-}\cdot X' = \frac{1}{Y_-}\cdot(Y - Y_0) = \frac{1}{Y_-}\cdot Y = X,$$

and we have the uniqueness.

To prove the representation 8.6 we wish to apply Ito's formula I.4.57 to $\log|Y|$. Since the "log" explodes at 0 we have to be careful: for each n, we consider a C^2 function f_n on $\mathbb{R}$, with $f_n(x) = \log|x|$ when $|x| \geq 1/n$. Then for all n and $t < T_n$ we have

$$\log|Y_t| = \log|Y_0| + \frac{1}{Y_-}\cdot Y_t - \frac{1}{2Y_-^2}\cdot\langle Y^c, Y^c\rangle_t + \sum_{s\leq t}\left[\log|Y_s| - \log|Y_{s-}| - \frac{\Delta Y_s}{Y_{s-}}\right].$$

This together with 8.4 yields 8.6 for $t < T_n$ and, since $\lim_n T_n = \infty$ by hypothesis, we obtain 8.6 everywhere. □

Another way for saying essentially the same thing is that $\mathscr{E}$ is a one-to-one map from the set of (equivalences classes for indistinguishability of) semimartingales X with $X_0 = 0$ and $\Delta X \neq 1$ onto the set of (equivalences classes for indistinguishability of) semimartingales Y with both Y and Y_- non-vanishing and $Y_0 = 1$, and the reciprocal map is $\mathscr{L}\text{og}$. When $X_0 = 0$ or $Y_0 = 1$ fail, we have the following trivial result:

8.7 **Corollary.** a) *If X is a semimartingale satisfying $\Delta X \neq -1$ identically, then $\mathscr{L}\text{og}\,(\mathscr{E}(X)) = X - X_0$.*

b) *If Y is a semimartingale such that Y and Y_- do not vanish, then* $\mathscr{E}(\mathscr{L}\mathrm{og}\,(Y)) = Y/Y_0$.

Another useful set of formulas consists in rewriting 8.2 and 8.3 by the way of the jump measures μ^X and μ^Y associated with X and Y. This readily gives:

8.8 if $\Delta X > -1$ then
$$\mathscr{E}(X) = \exp\left\{X - X_0 - \tfrac{1}{2}\langle X^c, X^c\rangle + \left(\log(1+x) - x\right) * \mu^X\right\},$$

8.9
$$\mathscr{L}\mathrm{og}\,Y = \log\left|\frac{Y}{Y_0}\right| + \frac{1}{2Y_-^2}\cdot\langle Y^c, Y^c\rangle - \left(\log\left|1 + \frac{y}{Y_-}\right| - \frac{y}{Y_-}\right) * \mu^Y.$$

8.10 **Theorem.** *Let X and $\bar{X}$ be two real-valued semimartingales. Then we have $e^X = \mathscr{E}(\bar{X})$ if and only if $\bar{X} = \mathscr{L}\mathrm{og}\,(e^X)$, and in this case we also have $X_0 = 0$ and $\Delta\bar{X} > -1$ identically, and*

8.11
$$\bar{X} = \bar{X}_0 + X + \frac{1}{2}\langle X^c, X^c\rangle + (e^x - 1 - x) * \mu^X,$$

8.12
$$X = \bar{X} - \bar{X}_0 - \frac{1}{2}\langle \bar{X}^c, \bar{X}^c\rangle + (\log(1+x) - x) * \mu^{\bar{X}},$$

8.13
$$\Delta\bar{X} = e^{\Delta X} - 1, \qquad \Delta X = \log(1 + \Delta\bar{X}).$$

Furthermore the characteristics (B, C, ν^X) and $(\bar{B}, \bar{C}, \nu^{\bar{X}})$ of X and $\bar{X}$, w.r.t. some truncation functions h and $\bar{h}$ respectively, can be computed ones from the others through the formulas:

8.14
$$\begin{cases} \bar{B} = B + \frac{C}{2} + [\bar{h}(e^x - 1) - h(x)] \star \nu^X, \\ \bar{C} = C, \\ 1_G(x) \star \nu^{\bar{X}} = 1_G(e^x - 1) \star \nu^X, \quad G \in \mathscr{R}, \end{cases}$$

8.15
$$\begin{cases} B = \bar{B} - \frac{\bar{C}}{2} + [h(\log(1+x)) - \bar{h}(x)] \star \nu^{\bar{X}}, \\ C = \bar{C}, \\ 1_G(x) \star \nu^X = 1_G(\log(1+x)) \star \nu^{\bar{X}}, \quad G \in \mathscr{R}. \end{cases}$$

Proof. Let $Y = e^X$, which satisfies $Y > 0$ and $Y_- > 0$. Then by Theorem 8.3 we have $\bar{X} = \mathscr{L}\mathrm{og}\,Y$ iff $Y = \mathscr{E}(\bar{X})$, hence the equivalence. If this holds, then $Y_0 = 1$ which gives $X_0 = 0$, and $\mathscr{E}(\bar{X}) > 0$ which gives $\Delta\bar{X} > -1$ by 8.2. Next, 8.12 follows from 8.8; this yields the second formula 8.13, hence the first one; next 8.11 follows by inverting 8.12, after observing that 8.12 yields $X^c = \bar{X}^c$.

As said already, we have $X^c = \bar{X}^c$, hence $\bar{C} = C$. The third formulas in 8.14 and 8.15 readily follow from 8.13 when ν^X and $\nu^{\bar{X}}$ are substituted with μ^X and $\mu^{\bar{X}}$, and these relations pass to the predictable compensators. Finally we write the decomposition 2.35 for X and plug it in 8.11 and compare with the decomposition 2.35 written for $\bar{X}$; this gives after deleting $X^c = \bar{X}^c$:

$$B+\frac{C}{2}+h\star(\mu^X-\nu^X)+(e^x-1-h(x))\star\mu^X = \bar{B}+\bar{h}\star(\mu^{\bar{X}}-\nu^{\bar{X}})+(x-\bar{h}(x))\star\mu^{\bar{X}}.$$

In view of the relation between the jump measures seen before, the last term on the right above equals the last term on the left, plus $A = (h(x) - \bar{h}(e^x - 1)) \star \mu^X$. Note that the function $k(x) = h(x) - \bar{h}(e^x - 1)$ is bounded and is equivalent to $x^2/2$ near 0. Therefore A has locally finite variation and a predictable compensator $A' = k \star \nu^X$. Thus

$$\bar{B} + A' - B - \frac{C}{2} = (h + k) \star (\mu^X - \nu^X) + \bar{h} \star (\mu^{\bar{X}} - \nu^{\bar{X}}).$$

The left side is predictable with locally finite variation, and the right side is a local martingale: then Corollary I.3.16 yields that both sides vanish, and this gives the first formula in 8.14. The first formula in 8.15 is easily deduced from 8.14. □

Combining Theorems 4.15 and 8.10, we immediately deduce:

8.16 **Corollary.** *If X is a real-valued PIIS then $\bar{X} = \mathscr{L}\mathrm{og}\,(e^X)$ is also a PIIS with $\Delta\bar{X} > -1$. If conversely $\bar{X}$ is a PIIS with $\Delta\bar{X} > -1$, then $X = \log\mathscr{E}(\bar{X})$ is also a PIIS.*

Furthermore in this case the characteristics of X and $\bar{X}$ are given by 4.20 for some triplet (b, c, K) and $(\bar{b}, \bar{c}, \bar{K})$ (relative to some truncation functions h and $\bar{h}$), and we have

8.17
$$\begin{cases} \bar{b} = b + \frac{c}{2} + \int[\bar{h}(e^x - 1) - h(x)]\, K(dx), \\ \bar{c} = c, \\ \bar{K}(G) = \int I_G(e^x - 1)\, K(dx), \quad G \in \mathscr{R}. \end{cases}$$

8.18
$$\begin{cases} b = \bar{b} - \frac{\bar{c}}{2} + \int[h(\log(1 + x)) - \bar{h}(x)]\, \bar{K}(dx), \\ c = \bar{c}, \\ K(G) = \int I_G(\log(1 + x))\, \bar{K}(dx), \quad G \in \mathscr{R}. \end{cases}$$

A last thing is in order here: the stochastic exponential is a generalization of the ordinary exponential in the sense that both of them solve the linear differential equation. What about the usual property that the exponential of a sum is the product of the exponential ? A routine use of Ito's formula gives that for any two semimartingales X and X', we have

8.19 $$\mathscr{E}(X)\mathscr{E}(X') = \mathscr{E}(X + X' + [X, X']).$$

This equation may be inverted, to give

8.20 $$\mathscr{L}\mathrm{og}\,(YY') = \mathscr{L}\mathrm{og}\,(Y) + \mathscr{L}\mathrm{og}\,(Y') + [\mathscr{L}\mathrm{og}\,(Y), \mathscr{L}\mathrm{og}\,(Y')]$$

for any two semimartingales Y and Y' which do not vanish and whose left limits do not vanish either.

§ 8b. Multiplicative Decompositions and Exponentially Special Semimartingales

1. As we know every special semimartingale X has an *additive decomposition* $X = X_0 + M + A$ with $M \in \mathscr{M}_{\mathrm{loc}}$ and A predictable in $\mathscr{V}$. This has a "multiplicative" counterpart which is given below, and which for simplicity is stated only for positive semimartingales with $X_0 = 1$.

8.21 **Theorem.** *Let X be a semimartingale with $X_0 = 1$, such that X and X_- take their values in* $(0, \infty)$. *Then X admits a multiplicative decomposition $X = LD$, where L is a positive local martingale and D a positive predictable process with locally finite variation and $L_0 = D_0 = 1$, if and only if X is a special semimartingale.*

In this case, the multiplicative decomposition is unique (up to evanescence), and is given as follows, where $X = 1 + M + A$ is the canonical (additive) decomposition of X and $\frac{1}{X_- + \Delta A}$ is necessarily locally bounded and positive:

8.22 $$L = \mathscr{E}\left(\frac{1}{X_- + \Delta A} \cdot M\right), \qquad D = \frac{1}{\mathscr{E}\left(-\frac{1}{X_- + \Delta A} \cdot A\right)}.$$

Proof. a) If a multiplicative decomposition $X = LD$ exists, by I.4.49(b) we have

8.23 $$X = LD = 1 + D \cdot L + L_- \cdot D.$$

But $D \cdot L \in \mathscr{L}$ and $L_- \cdot D$ is predictable and in $\mathscr{V}$, thus X is a special semimartingale.

Suppose that we have two multiplicative decompositions $X = LD = L'D'$, for which we can write 8.23. The uniqueness of the canonical decomposition yields $L_- \cdot D = L'_- \cdot D'$ (up to evanescence), or $L_- dD = L'_- dD'$ in "differential form". Since L, L, D and D' do not vanish and $L/L' = D'/D$, we deduce that $(1/D_-)dD = (1/D'_-)dD'$. Then if $U = \frac{1}{D'_-} \cdot D'$, we also have $\frac{1}{D_-} \cdot D = U$. In other words (see e.g. Theorem 8.3) we have $D = D' = \mathscr{E}(U)$, and this proves the uniqueness.

b) It remains to prove the following: if $X = 1 + M + A$ is the canonical decomposition of the special semimartingale X (with $X > 0$ and $X_- > 0$), and if $K = X_- + \Delta A$ and $H = 1/K$ and $N = H \cdot M$ and $B = H \cdot A$, then the predictable process H is locally bounded and positive and further $X = LD$, where $L = \mathscr{E}(N)$ and $D = 1/\mathscr{E}(-B)$ (that is, L and D are given by 8.22: note that N is a local martingale, hence L also, and that B is predictable with finite variation, hence D also).

c) First, we observe that $K = {}^pX$, the predictable projection of X: indeed ${}^pA = A = A_- + \Delta A$ is obvious and ${}^pM = M_-$ by Corollary I.2.31. Then for any predictable time T, and since $X > 0$, we get $K_T = E(X_T|\mathscr{F}_{T-}) > 0$ a.s. on $\{T < \infty\}$. Then I.2.18 yields that $K > 0$ outside an evanscent set, and we deduce $0 < H < \infty$.

This is not enough, though, to have H locally bounded. But $H = 1/X_-V$, where $V = 1+\Delta A/X_-$, and the process $1/X_-$ is positive locally bounded, so it is enough to prove that the positive finite process $1/V$ is locally bounded. The process $U_t = \sum_{s \leq t} \Delta A_s/X_{s-}$ is in $\mathscr{V}$, so the successive times where $\Delta U = V - 1 \leq 1/2$ form a strictly increasing sequence (T_n) going to $+\infty$. Then $S_n = \inf(t : V_t \leq 1/n)$ satisfies $[\![S_n]\!] \subset \cup_p [\![T_p]\!]$ for $n \geq 2$ and since $V > 0$ everywhere the sequence S_n increases to $+\infty$ as well, and further $[\![S_n]\!] \subset \{V \leq 1/n\}$: so Proposition I.2.13 implies that S_n is a predictable time, and by I.2.16 for each n there is a stopping time $S'_n < S_n$ such that $P(S'_n < S_n - 1) \leq 2^{-n}$. Therefore the sequence of stopping times $R_n = \sup_{1 \leq p \leq n} S'_p$ increases a.s. to $+\infty$, and $V \geq 1/n$ on $[\![0, R_n]\!]$: this proves that $1/V$ is locally bounded.

d) Since we know that H is locally bounded we can define L and D as above, and it remains to prove that $X = LD = \mathscr{E}(N)/\mathscr{E}(-B)$. Observe that $\Delta B = \frac{\Delta A}{X_- + \Delta A} < 1$ identically, hence $\mathscr{E}(-B) > 0$ and $\mathscr{E}(-B)_- > 0$. Then we apply Ito's formula to the function $f(x, y) = \frac{x}{y}$, or rather to a C^2 function coinciding with f for y outside an arbitrarily small neighbourhood of 0 (exactly as we did in the proof of Theorem 8.3), to obtain with $D' = 1/D$ and $X' = LD$:

$$X'_t = 1 + \frac{1}{D'_-} \cdot L_t - \frac{L_-}{D'^{\,2}_-} \cdot D'_t + \sum_{s \leq t} \left(\frac{L_s}{D'_s} - \frac{L_{s-}}{D'_{s-}} - \frac{\Delta L_s}{D'_{s-}} + \frac{L_{s-} \Delta D'_s}{D'^{\,2}_{s-}} \right).$$

Since $L = 1 + L_- \cdot N$ and $D' = 1 - D'_- \cdot B$, we get

$$\begin{aligned}
X'_t &= 1 + X'_- \cdot N_t + X'_- \cdot B_t + \sum_{s \leq t} X'_{s-} \left(\frac{1 + \Delta N_s}{1 - \Delta B_s} - 1 - \Delta N_s - \Delta B_s \right) \\
&= 1 + X'_- \cdot N_t + X'_- \cdot B_t + \sum_{s \leq t} X'_{s-} \frac{\Delta N_s + \Delta B_s}{1 - \Delta B_s} \Delta B_s \\
&= 1 + X'_- \left(1 + \frac{\Delta B}{1 - \Delta B} \right) \cdot N_t + X'_- \left(1 + \frac{\Delta B}{1 - \Delta B} \right) \cdot B_t \\
&= 1 + \frac{X'_- H}{1 - \Delta B} \cdot X_t = 1 + \frac{X'_-}{X_-} \cdot X_t
\end{aligned}$$

because $\Delta B = H\Delta A$ and $1 - H\Delta A = HX_-$, hence $\frac{H}{1-\Delta B} = \frac{1}{X_-}$. Then $\frac{1}{X'}\cdot X' = \frac{1}{X_-}\cdot X$ and, exactly as in Part (a) of the proof, we deduce that $X' = X$. □

2. Let us recall that any semimartingale X satisfies $(|x|^2 \wedge 1) * \nu \in \mathscr{V}$ (ν being the third characteristic of X), while it is special iff $(|x|^2 \wedge |x|) * \nu \in \mathscr{V}$ (by Proposition 2.29; recall that a predictable process is in $\mathscr{V}$ iff it is in $\mathscr{A}_{\text{loc}}$). Since $|x|1_{\{|x|>1\}} \leq x^2 \wedge |x| \leq x^2 + |x|1_{\{|x|>1\}})$, it also follows that the semimartingale X is special iff we have

8.24 $$|x|1_{\{|x|>1\}} * \nu \in \mathscr{V}.$$

8.25 **Definition.** The real-valued semimartingale X *exponentially special* if $\exp(X - X_0)$ is a special semimartingale.

By analogy with 2.29 or 8.24, we can express this property in terms of ν.

8.26 **Proposition.** *Let X be a real-valued semimartingale with characteristics (B, C, ν) relative to $h \in \mathscr{C}_t^1$. The the following three properties are equivalent:*
a) *X is exponentially special,*
b) *the process $(e^x - 1 - h(x)) * \nu$ is in $\mathscr{V}$,*
c) *the process $e^x 1_{\{x>1\}} * \nu$ is in $\mathscr{V}$.*

Proof. Set $\bar{X} = \mathscr{L}\text{og}\, Y$ with $Y = e^{X-X_0}$. We have $\bar{X} = (1/Y_-)\cdot Y$ and $Y = 1 + Y_-\cdot\bar{X}$ (see Theorem 8.3). Since the stochastic integral process of a locally bounded predictable process w.r.t. a special semimartingale is special, we see that (a) is equivalent to the fact that $\bar{X}$ is special, which in turn is equivalent by 8.24 to the fact that $|x|1_{\{|x|>1\}} \star \nu^{\bar{X}}$ belongs to $\mathscr{V}$ (where $\nu^{\bar{X}}$ is the third characteristic of $\bar{X}$). Further, $\bar{X}$ and $X - X_0$ are connected one with the other like in Theorem 8.10, and of course $\nu^{X-X_0} = \nu$. Since $|e^x - 1| > 1$ iff $x > \log 2$, we deduce from 8.14 that (a) is equivalent to the property

8.27 $$(e^x - 1)1_{\{x>\log 2\}} \star \nu \in \mathscr{V}.$$

Since we know that $(x^2 \wedge |x|) \star \nu \in \mathscr{V}$, the implications 8.27 $\Rightarrow$ (c) $\Rightarrow$ (b) $\Rightarrow$ 8.27 follow respectively from the next three inequalities, where C is a constant depending on h:

$$\begin{aligned} e^x 1_{\{|x|>1\}} &\leq (e^x - 1)1_{\{x>\log 2\}} + x^2 \wedge |x|, \\ |e^x - 1 - h(x)| &\leq C(e^x 1_{\{x>1\}} + x^2 \wedge |x|), \\ (e^x - 1)1_{\{|x|>\log 2\}} &\leq C(|e^x - 1 - h(x)| + x^2 \wedge |x|). \end{aligned}$$

□

If X is a special semimartingale with canonical decomposition $X = X_0 + M + A$, then A is sometimes called the "compensator" of X, and it is the unique predictable process in $\mathscr{V}$ such that $X - X_0 - A \in \mathscr{M}_{\text{loc}}$. By analogy, we set:

8.28 **Definition.** Let X be a real-valued semimartingale. A predictable process $V \in \mathscr{V}$ is called an *exponential compensator* of X if $\exp(X - X_0 - V) \in \mathscr{M}_{\mathrm{loc}}$.

In other words, we have the *multiplicative decomposition* of Theorem 8.21 for e^{X-X_0}, namely $e^{X-X_0} = LD$ with $D = e^V$; note that D is predictable with locally finite variation and $D_0 = 1$. In view of Theorem 8.21, the following is obvious:

8.29 **Proposition.** *A real-valued semimartingale X has an exponential compensator if and only if it is exponentially special. In this case the exponential compensator is unique, up to evanescence.*

Chapter III. Martingale Problems and Changes of Measures

In limit theorems, one needs to characterize the law (or distribution) of various processes, in particular of the limiting process. As is well known, the law of a process is indeed characterized by the family of its "finite-dimensional" distributions. However, one is very rarely able to explicitely compute these finite-dimensional distributions, except for PII. On the other hand, many usual processes are semimartingales; and a natural tool has emerged in Chapter II for studying them, namely their characteristics: at least, they are often easy to compute.

The first question addressed to in the present chapter can loosely be stated this way: to which point do the characteristics of a semimartingale indeed characterize the law of the semimartingale? the answer may be positive, as for PIIs for example (we have seen that in Theorem II.4.15), or "reasonably often" for diffusion processes; in general, though, the answer is negative.

The characteristics can be defined in terms of a number of processes being local martingales (Theorem II.2.21), so the question may be stated as such: what are all the probability measures on a given filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ under which all members of a given family $\mathscr{X}$ of processes are local martingales? such a problem is called a *martingale problem.* Martingale problems pertaining to semimartingales and their characteristics are introduced in Section 2, where many examples are described. In Section 1 we introduce "general" martingale problems, and also the problems related to point processes and multivariate point processes and random measures: it is a good introduction to Section 2 because it is simpler; and it does not (formally) reduce to martingale problems associated with semimartingales, whereas it has practical interest of its own. However, someone interested only in semimartingales can proceed directly to Section 2.

Sections 3 and 5 are devoted to studying the first steps of a problem that is very important for statistics of processes (the next steps are considered in Chapters IV and V). More specifically, we study what happens to a martingale or a semimartingale or a random measure when one replaces the original measure P by another probability measure P' which is absolutely continuous with respect to P. Part of the story consists in various versions of "Girsanov's Theorem", which all hinge upon considering the basic "density processes" of P' with respect to P, and this is done in Section 3. Another part, examined in Section 5, consists in actually computing this density process.

In the middle of it (Section 4), we study a problem related to Girsanov's Theorem and also to martingale problems: can we represent all local martingales on a stochastic basis as stochastic integrals with respect to some "basic" local martingales or random measures. The statements of the various results are independent from Girsanov's Theorems (and from Section 3), but the proofs are not! on the other hand, these representation results are crucial for Section 5.

1. Martingale Problems and Point Processes

§ 1a. General Martingale Problems

This subsection is sort of a general introduction to this chapter: we describe what a "general" martingale problem is. The ingredients are as follows:

1.1 $(\Omega, \mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}, \mathscr{F})$ is a *filtered space* (no measure on it, yet); and $\mathscr{H}$ is a sub-σ-field of $\mathscr{F}_0$, called the *initial σ-field.*

1.2 $\mathscr{X}$ is a family of optional $\overline{\mathbb{R}}$-valued processes on $(\Omega, \mathbf{F}, \mathscr{F})$.

1.3 Definition. Let P_H be a probability measure on $(\Omega, \mathscr{H})$ (called the *initial condition*). Then a *solution to the martingale problem* associated with $\mathscr{X}$ and P_H is a probability measure P on $(\Omega, \mathscr{F})$ such that:

(i) the restriction $P_{|\mathscr{H}}$ of P to $\mathscr{H}$ equals P_H;

(ii) each process $X \in \mathscr{X}$ is a local martingale on the stochastic basis $(\Omega, \mathbf{F}, \mathscr{F}, P)$. □

This formulation explains why in Chapter I we have carefully dealt with non-complete stochastic bases: here the filtration $\mathbf{F}$ cannot be complete since we do not know the measure P beforehand, and a martingale problem may have many solutions.

We have already encountered several examples:

1.4 Example. *The standard Wiener process.* Let W be a continuous adapted process defined on $(\Omega, \mathbf{F}, \mathscr{F})$. Then it follows from Theorem II.4.4 that W is a standard Wiener process under a measure P if and only if P is a solution to the martingale problem associated with

$$\begin{cases} -\mathscr{H} = \sigma(W_0),\ P_H \text{ is the measure such that } P_H(W_0 = 0) = 1, \\ -\mathscr{X} = \{W, Y\}, \text{ with } Y_t = W_t^2 - t. \end{cases}$$ □

1.5 **Example.** *The standard Poisson process.* Let N be an adapted point process on $(\Omega, \mathscr{F}, \mathbf{F})$. Then Theorem II.4.5 yields that N is a standard Poisson process under a measure P is and only if P is a solution to the martingale problem associated with

$$\begin{cases} -\mathscr{H} = \{\Omega, \phi\},\ P_H \text{ is the trivial probability measure on } (\Omega, \mathscr{H}), \\ -\mathscr{X} = \{X\} \text{ with } X_t = N_t - t. \end{cases}$$ □

1.6 **Remarks.** 1) In most cases, all elements of $\mathscr{X}$ are $\mathbb{R}$-valued and càdlàg. However there are instances (encountered sometimes in this book) where it is not possible to assume this hypothesis, which anyway does not simplify the discussion by any mean.

2) On the opposite, one could even relax the adaptedness of the elements of $\mathscr{X}$, but that would notably complicate the matter, with no benefit for us in this book. □

§ 1b. Martingale Problems and Random Measures

Here we describe the simplest of the two classes of martingale problems encountered in this book. We are given the filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ and the initial σ-field $\mathscr{H}$, as in 1.1.

Let also $(E, \mathscr{E})$ be an auxiliary Blackwell space (see § II.1a: for all practical purposes, $E = \mathbb{R}^d$ or $E = \bar{\mathbb{R}}^d$) and we consider an *integer-valued random measure* μ on $\mathbb{R}_+ \times E$, in the sense of II.1.13: that is, μ is an optional random measure, taking its values in $\bar{\mathbb{N}}$, and such that $\mu(\omega; \{t\} \times E) \le 1$ identically; the difference with II.1.13 is that we cannot impose beforehand "$\tilde{\mathscr{P}}$-σ-finiteness" because there is no probability measure. However, we impose the following condition on μ:

1.7 There is a strictly positive $\tilde{\mathscr{P}}$-measurable function on $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$ (recall that $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$) such that $V * \mu_\infty(\omega) < \infty$ for all ω ($V * \mu$ is the Stieltjes integral process of V with respect to μ: see II.1.5). □

It is obviously possible to replace V by $V \wedge 1$ above; then, if $T_n = \inf(t: V * \mu_t \le n)$, $T_n(\omega) \uparrow \infty$ for all ω as $n \uparrow \infty$ and $V * \mu_{T_n} \le n + 1$. Hence $V' = \sum_{n \ge 1} 2^{-n} V 1_{[\![0, T_n]\!] \times E}$ is $\tilde{\mathscr{P}}$-measurable and strictly positive and has $V' * \mu_\infty \le \sum_{n \ge 1} (n + 1) 2^{-n}$. So 1–7 yields:

1.8 There is a strictly positive $\tilde{\mathscr{P}}$-measurable function V on $\tilde{\Omega}$ such that $V * \mu_\infty$ is bounded. □

In particular, we then deduce that μ is $\tilde{\mathscr{P}}$-σ-finite relatively to *all* probability measures on $(\Omega, \mathscr{F})$.

1.9 **Definition.** Assume 1.1 and let μ satisfy 1.7. Let P_H be an initial condition (i.e., a probability measure on $(\Omega, \mathscr{H})$). Let ν be a predictable random measure on $\mathbb{R}_+ \times E$. Then, a *solution to the martingale problem* associated with $(\mathscr{H}, \mu)$ *and* (P_H, ν) is a probability measure P on $(\Omega, \mathscr{F})$ such that:

(i) the restriction $P_{|\mathscr{H}}$ of P to $\mathscr{H}$ equals P_H;

(ii) ν is the compensator of μ on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

We denote by $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ the set of all solutions. □

1.10 **Proposition.** *P belongs to $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ if and only if it is a solution to the martingale problem 1.3, with for $\mathscr{X}$ the family of all processes of the form*

1.11 $$X = (WV) * \mu - (WV) * \nu$$

where W ranges through all nonnegative bounded $\tilde{\mathscr{P}}$-measurable functions, and V is given in 1.8.

Proof. Note that each process 1.11 is well defined (with values in $[-\infty, \infty)$).

Assume first that $P \in \mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$. Then $WV * \mu \in \mathscr{A}^+$ and so by definition of the compensator (see II.1.8(ii)), $WV * \mu - WV * \nu \in \mathscr{M}$.

Conversely, assume that all X above are local martingales on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ for some measure P. Then $E(WV * \mu_\infty) = E(WV * \nu_\infty)$ for all bounded nonnegative $\tilde{\mathscr{P}}$-measurable W (because $WV * \nu$ is then the compensator of $WV * \mu$), and by the monotone limit theorem the same holds for all nonnegative $\tilde{\mathscr{P}}$-measurable W, bounded or not. Then ν is the compensator of μ on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, and we are done. □

1.12 **Remark.** In the definition of the family $\mathscr{X}$ above, it is not necessary to take all those many W's. We could take all W of the form $W(\omega, t, x) = g(x)$ where g ranges through the set of all bounded nonnegative measurable functions on $(E, \mathscr{E})$, or even through a smaller subset as the class $\mathscr{C}^+(\mathbb{R}^d)$ defined in II.2.20 when $E = \mathbb{R}^d$. □

1.13 **Proposition.** *$\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ is a convex set.*

Proof. Let P, $P' \in \mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ and $Q = \alpha P + (1 - \alpha) P'$ where $\alpha \in (0, 1)$. Then $Q_{|\mathscr{H}} = \alpha P_{|\mathscr{H}} + (1 - \alpha) P'_{|\mathscr{H}} = P_H$, and for all nonnegative $\tilde{\mathscr{P}}$-measurable function W on $\tilde{\Omega}$,

$$\begin{aligned} E_Q(W * \mu_\infty) &= \alpha E_P(W * \mu_\infty) + (1 - \alpha) E_{P'}(W * \mu_\infty) \\ &= \alpha E_P(W * \nu_\infty) + (1 - \alpha) E_{P'}(W * \nu_\infty) = E_Q(W * \nu_\infty) \end{aligned}$$

so ν is the Q-compensator of μ. □

As a first example, we have the (extended) Poisson measures (see § II.1c). Theorem II.4.8 can be re-formulated as such:

1.14 **Theorem.** *Assume that in 1.9 the measure ν is deterministic* (i.e., $\nu(\omega, \cdot) = m(\cdot)$ for some measure m).

a) *P belongs to $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ if and only if μ is an extended Poisson measure on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with intensity measure ν and $P_{|\mathscr{H}} = P_H$.*

b) *Assume further that $\mathscr{F} = \mathscr{F}_{\infty-}$ and that $\mathbf{F}$ is the smallest filtration such that μ is optional and that $\mathscr{H} \subset \mathscr{F}_0$. Then $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ contains at most one element.*

Proof. (a) is nothing else than II.4.8a. Moreover, II.4.8 shows that if P is a solution, then $E(1_B \exp \sum_{j \leq d} iu_j \mu(A_j))$, where $B \in \mathscr{H}$, $u_j \in \mathbb{R}$, $A_j \in \mathscr{E}$, depends only upon ν and $P(B) = P_H(B)$. Since $\mathscr{F} = \mathscr{H} \vee \sigma(\mu(A)\colon A \in \mathscr{E})$, we deduce uniqueness for P. □

This theorem deals with uniqueness. There is also an existence result, but for it the space $(\Omega, \mathscr{F})$ must be large enough to accomodate all possible "paths" of the integer-valued random measure. Let us just state the result in a loose way: we assume that Ω is the *canonical space* of all integer-valued random measures on $\mathbb{R}_+ \times E$, and we denote by μ the canonical random measure on Ω, and we call $\mathbf{F}$ the smallest filtration for which μ is optional and $\mathscr{F} = \mathscr{F}_{\infty-}$. Finally, let $\mathscr{H}$ be the trivial σ-field $\{\Omega, \emptyset\}$ with the trivial probability measure P_H. Then:

1.15 If m is any positive σ-finite measure on $\mathbb{R}_+ \times E$ such that $m(\{t\} \times E) \leq 1$ for all t, and if $\nu(\omega, \cdot) = m(\cdot)$, there is one and only one solution in $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$. □

(For a proof, see e.g. [98]). Of course, there also exists a version of 1.15 with a non-trivial initial condition.

§ 1c. Point Processes and Multivariate Point Processes

1. We start again with 1.1, and we assume that $(\Omega, \mathscr{F}, \mathbf{F})$ is endowed with an *adapted point process* N (see § I.3b): recall that if $(T_n)_{n \geq 1}$ denote the successive jump times and $T_0 = 0$, then $T_n < T_{n+1}$ if $T_n < \infty$, and

1.16
$$N = \sum_{n \geq 1} 1_{[\![T_n, \infty[\![}.$$

We also consider an initial condition P_H on $(\Omega, \mathscr{H})$, and an increasing càdlàg predictable process A with $A_0 = 0$. We are interested in those measures P for which A is the compensator of N and $P_{|\mathscr{H}} = P_H$.

That problem can be fitted into the framework of the previous subsection, as follows: set $E = \{1\}$. Then N can be considered as the "distribution function" of an integer-valued random measure μ on $\mathbb{R}_+ \times \{1\}$, namely (see Example II.1.7):

1.17 $$\mu(dt, dx) = \sum_{n\geq 1} 1_{\{T_n<\infty\}}\varepsilon_{(T_n,1)}(dt \times dx)$$

(which obviously meets 1.7: take $V = \sum_{n\geq 1} 2^{-n} 1_{[\![0,T_n]\!]\times\{1\}}$), and similarly A is the distribution function of a predictable random measure ν:

1.18 $$\nu(dt, dx) = dA_t \varepsilon_1(dx)$$

Moreover, we have the following (see II.1.12):

1.19 P belongs to $\mathscr{s}(\mathscr{H}, \mu|P_H, \nu)$ is and only if $P_{|\mathscr{H}} = P_H$ and A is the compensator of N on $(\Omega, \mathscr{F}, \mathbf{F}, P)$. □

Finally, we add one more hypothesis:

1.20 $\mathscr{F} = \mathscr{F}_{\infty-}$ and $\mathbf{F}$ is the smallest filtration to which N is adapted and such that $\mathscr{H} \subset \mathscr{F}_0$ (i.e. $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0$, $\mathscr{F}_s^0 = \mathscr{H} \vee \sigma(N_r: r \leq s)$). □

We wish to prove the following uniqueness result:

1.21 **Theorem.** *With the above assumptions and notation* (in particular 1.20) *there is at most one probability measure P such that $P_{|\mathscr{H}} = P_H$ and that A is the compensator of N on $(\Omega, \mathscr{F}, \mathbf{F}, P)$* (in other words, $\mathscr{s}(\mathscr{H}, \mu|P_H, \nu)$ has at most one solution).

1.22 **Remark.** Here again we have stated a uniqueness result only. As for existence, this is another matter:

1) One may hope for a solution only if Ω is rich enough: for instance, if it is the canonical space of all point processes.

2) Even so, we are not insured of existence, because the point process may be led to "explode" at a finite time by its compensator.

In fact, one can prove *existence and uniqueness* if Ω is the set of all $\bar{\mathbb{N}}$-valued counting functions (i.e. of the form 1.16, but with $\lim_n \uparrow T_n$ being *finite* or infinite): see [94]. □

2. In fact, Theorem 1.21 is a particular case of a similar statement concerning multivariate point processes, which we introduce now. 1.1 is still in force.

1.23 **Definition.** Let $(E, \mathscr{E})$ be a Blackwell space. An *E-valued multivariate point process* is an integer-valued random measure μ on $\mathbb{R}_+ \times E$ such that $\mu(\omega; [0,t] \times E) < \infty$ for all ω, $t \in \mathbb{R}_+$. □

Let us introduce the stopping times $T_n = \inf(t: \mu([0,t] \times E) \geq n)$. Then $T_n < T_{n+1}$ if $T_n < \infty$, and $T_n \uparrow \infty$ as $n \uparrow \infty$, and according to II.1.14 there are $\mathscr{F}_{T_n}$-measurable E-valued random variables Z_n such that

1.24 $$\mu(dt, dx) = \sum_{n \geq 1} 1_{\{T_n < \infty\}} \varepsilon_{(T_n, Z_n)}(dt, dx).$$

Note that μ meets 1.7: take $V = \sum_{n \geq 1} 2^{-n} 1_{[\![0, T_n]\!] \times E}$. We also consider an initial condition P_H on $(\Omega, \mathscr{H})$, and a predictable random measure ν on $\mathbb{R}_+ \times E$. Finally, we assume the following:

1.25 $\mathscr{F} = \mathscr{F}_{\infty-}$ and $\mathbf{F}$ is the smallest filtration for which μ is optional and $\mathscr{H} \subset \mathscr{F}_0$ (i.e. $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0$ and $\mathscr{F}_s^0 = \mathscr{H} \vee \sigma(\mu([0,r] \times B) \colon r \leq s, B \in \mathscr{E})$). □

Since a point process can be viewed as an E-valued multivariate point process with E reduced to one point, in which case 1.25 reduces to 1.20, the following clearly extends 1.21:

1.26 **Theorem.** *With the above assumptions* (in particular μ is a multivariate point process and 1.25 holds), *the set $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ contains at most one point.*

1.27 **Remarks.** 1) 1.26 is usually wrong when μ is an integer-valued random measure, but not a multivariate point process (we have seen, however, that it is correct for Poisson random measures).

2) As for point processes, an existence result may be found in [94]. □

3. The proof of 1.26 is broken into several steps, the two first ones being mere extensions of the results in § II.3c.2. For $n \in \mathbb{N}^*$, set

1.28 $$\mathscr{G}(0) = \mathscr{H}, \quad \mathscr{G}(n) = \mathscr{H} \vee \sigma(T_1, Z_1, T_2, Z_2, \ldots, T_n, Z_n).$$

1.29 **Lemma.** *Assume 1.25.*

a) *A set B belongs to $\mathscr{F}_t$, if and only if for each $n \in \mathbb{N}$ there is a set $B_n \in \mathscr{G}(n)$ such that $B \cap \{t < T_{n+1}\} = B_n \cap \{t < T_{n+1}\}$.*

b) *A process H is predictable, if and only if H_0 is $\mathscr{F}_0$-measurable and if for each $n \in \mathbb{N}$ there is a $\mathscr{G}(n) \otimes \mathscr{R}_+$-measurable process $H(n)$, with*

1.30 $$H = H_0 + \sum_{n \in \mathbb{N}} H(n) 1_{]\!]T_n, T_{n+1}]\!]}.$$

Proof. a) Denote by $\mathscr{K}_t$ the set of all B's having the claimed property: it clearly is a σ-field.

If $B \cap \{t < T_{n+1}\} = B_n \cap \{t < T_{n+1}\}$ with $B_n \in \mathscr{G}(n)$ for all n, then $B = \bigcup_n [B_n \cap \{t < T_{n+1}\} \cap \{T_n \leq t\}]$, and since $\mathscr{G}(n) \subset \mathscr{F}_{T_n}$ (a trivial property) we deduce that $\mathscr{K}_t \subset \mathscr{F}_t$.

Let $s \leq t$, $p \in \mathbb{N}$, $A \in \mathscr{E}$ and $B = \{\mu([0,s] \times A) = p\}$. Then $B_n := \{\sum_{1 \leq i \leq n} 1_A(Z_i) 1_{[0,s]}(T_i) = p\}$ belongs to $\mathscr{G}(n)$ and $B \cap \{t < T_{n+1}\} = B_n \cap \{t < T_{n+1}\}$ for all $n \in \mathbb{N}$. Hence we deduce that $B \in \mathscr{K}_t$. Moreover $\mathscr{H} = \mathscr{G}(0) \subset \mathscr{K}_t$ is obvious: then 1.25 yields $\mathscr{F}_t^0 \subset \mathscr{K}_t$.

Let $B \in \bigcap_{s>t} \mathscr{K}_s$. For all $s > t$, $n \in \mathbb{N}$ there is a set $B_{n,s} \in \mathscr{G}(n)$ with $B \cap \{s < T_{n+1}\} = B_{n,s} \cap \{s < T_{n+1}\}$. Then $B_n = \limsup_{s \in Q, s \downarrow\downarrow t} B_{n,s}$ belongs to $\mathscr{G}(n)$, and $B \cap \{t < T_{n+1}\} = B_n \cap \{t < T_{n+1}\}$, and $B \in \mathscr{K}_t$ follows.

So far, we have seen that $\mathscr{F}_t^0 \subset \mathscr{K}_t \subset \mathscr{F}_t$ and $\bigcap_{s>t} \mathscr{K}_s \subset \mathscr{K}_t$, while $\mathscr{F}_t = \bigcap_{s<t} \mathscr{F}_s^0$ by definition: we deduce that $\mathscr{K}_t = \mathscr{F}_t$.

b) Since $\mathscr{G}(n) \subset \mathscr{F}_{T_n}$, the sufficient condition follows from I.2.12: take first $H(n) = Y(n) 1_{]\!]u,v]\!]}$ where $Y(n)$ is a $\mathscr{G}(n)$-measurable random variable, then conclude by linearity and a monotone class argument. For the converse, a monotone class argument again yields that it suffices to prove the claim when $H = 1_{B \times \{0\}}$, $B \in \mathscr{F}_0$ (then it is obvious) and when $H = 1_{B \times (t,s]}$, $B \in \mathscr{F}_t$: let B_n be associated to B as in (a); then H is of the form 1.30 with $H(n) = 1_{B_n \times (t,s]}$. □

1.31 **Remark.** One could prove more precise results, which will not be used here (see [98, 128]), for example: $\mathscr{F}_t = \mathscr{F}_t^0$ and $\mathscr{F}_{T_n} = \mathscr{G}(n)$. □

The next result has interest of its own. Let P be a probability measure on $(\Omega, \mathscr{F})$, and set

1.32 $G_n(\omega; ds, dx)$ is a regular version of the conditional distribution of (T_{n+1}, Z_{n+1}) with respect to $\mathscr{G}(n)$ (so $G_n(\omega; \cdot)$ is supported by $(T_n(\omega), \infty] \times E$ or by $\{\infty\} \times E$). □

1.33 **Theorem.** *Under 1.25, a version of the compensator of the E-valued multivariate point process μ is*

1.34
$$\nu(dt, dx) = \sum_{n \geq 1} \frac{1}{G_n([t, \infty] \times E)} 1_{\{t \leq T_{n+1}\}} G_n(dt, dx).$$

In particular if $F_n(dt) = G_n(dt \times E)$, the point process $N = \sum_{n \geq 1} 1_{[\![T_n, \infty[\![}$ has the compensator $A_t = \nu([0, t] \times E)$, which reads as

1.35
$$A_t = \sum_{n \geq 1} \int_0^{T_{n+1} \wedge t} \frac{1}{F_n([s, \infty])} F_n(ds),$$

and this generalizes Theorem II.3.26.

Proof. Define ν by 1.34. We need to prove that for all nonnegative $\tilde{\mathscr{P}}$-measurable functions W, then

1) $W * \nu$ is predictable,

2) $E(W * \nu_\infty) = E(W * \mu_\infty)$.

By a monotone class argument, it suffices to prove these facts when $W = H \otimes g$, where H is a predictable nonnegative process and g is a nonnegative measurable function on $(E, \mathscr{E})$. Then H has a representation 1.30, so $W * \nu =$

$\sum_{n\geq 1} K(n) 1_{]\!]T_n, T_{n+1}]\!]}$ with

$$K(n)_t = \sum_{0\leq p\leq n-1} \int_{(T_p, T_{p+1}]\times E} \frac{1}{G_p([s,\infty]\times E)} H(p)_s g(x) G_p(ds,dx)$$
$$+ \int_{(T_n, t]\times E} \frac{1}{G_n([s,\infty]\times E)} H(n)_s g(x) G_n(ds,dx)$$

(recall that $G_p([0,T_p]\times E) = 0$ if $T_p < \infty$). Then $K(n)$ is $\mathscr{G}(n)\otimes\mathscr{R}_+$-measurable by definition of the G_p's, so 1.29 implies that $W * \nu$ is predictable.

Next, we will show that $E(W1_{]\!]T_n, T_{n+1}]\!]} * \nu_\infty) = E(W1_{]\!]T_n, T_{n+1}]\!]} * \mu_\infty)$: then summing up on $n\in\mathbb{N}$ gives (2). We can again assume that $W = H\otimes g$ with H given by 1.30, so $W1_{]\!]T_n, T_{n+1}]\!]} = H(n) g(x) 1_{]\!]T_n, T_{n+1}]\!]}$ where $H(n)$ is $\mathscr{G}(n)\otimes\mathscr{R}_+$-measurable, so

$$
\begin{aligned}
&E(W1_{]\!]T_n, T_{n+1}]\!]} * \nu_\infty)\\
&= E\left(\int_{(T_n, T_{n+1}]\cap(T_n,\infty)} \int_E \frac{H(n)_s g(x)}{G_n([s,\infty]\times E)} G_n(ds,dx)\right)\\
&= E\left(\int_{(T_n,\infty]} G_n(du\times E) \int_{(T_n,u]\cap(T_n,\infty)} \int_E \frac{H(n)_s g(x)}{G_n([s,\infty]\times E)} G_n(ds,dx)\right) && \text{(by 1.32)}\\
&= E\left(\int_{(T_n,\infty)} \frac{H(n)_s g(x)}{G_n([s,\infty]\times E)} G_n(ds,dx) \int_{[s,\infty]} G_n(du\times E)\right) && \text{(Fubini)}\\
&= E\left(\int_{(T_n,\infty)\times E} H(n)_s g(x) G_n(ds,dx)\right)\\
&= E(H(n)_{T_{n+1}} g(Z_{n+1}) 1_{\{T_{n+1}<\infty\}}) && \text{(by 1.32)}\\
&= E(W1_{]\!]T_n, T_{n+1}]\!]} * \mu_\infty) && \text{(by 1.24)}
\end{aligned}
$$

and we are finished. □

Proof of Theorem 1.26. Let P^1 and P^2 belong to $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$, and call $G_n^i(\omega; ds\times dx)$ a regular version of the conditional distribution of (T_{n+1}, Z_{n+1}), with respect to $\mathscr{G}(n)$, for the measure P^i. The idea of the proof is as follows: 1.34 holds P^i-a. s. if we replace G_n by G_n^i; so "$G_n^1 = G_n^2$" and since $P^1 = P^2$ on $\mathscr{G}(0)$, an induction on n shows that $P^1 = P^2$. Unfortunately, the actual proof is somewhat more intricate, because G_n^i is defined up to a P^i-null set only!

Our induction hypothesis is that $P^1 = P^2$ on $\mathscr{G}(n)$: it is met for $n = 0$ because $P^i_{|\mathscr{H}} = P_H$ and $\mathscr{G}(0) = \mathscr{H}$. We will prove that $P^1 = P^2$ on $\mathscr{G}(n+1)$, so the induction yields $P^1 = P^2$ on $\bigvee_n \mathscr{G}(n)$, which equals $\mathscr{F}$ by 1.25.

From now on, $n\in\mathbb{N}$ is fixed. Let $\mathscr{E}_0$ be a countable algebra which generates the σ-field $\mathscr{E}$. For every $A\in\mathscr{E}_0$ the process $X_t^A = \nu((T_n\wedge t, t]\times A)$ is predictable and increasing, so we deduce from 1.29 that there is a $\mathscr{G}(n)\otimes\mathbb{R}_+$-measurable increasing process $\bar{X}^A$ such that $\bar{X}_t^A(\omega) = X_t^A(\omega)$ for all $t\leq T_{n+1}(\omega)$.

For $i = 1, 2$, and for $A \in \mathscr{E}_0$, we set

1.36 $$X_t^{i,A} = \int_{[0,t]\times A} \frac{1}{G_n^i([s,\infty] \times E)} G_n^i(ds, dx),$$

which is also $\mathscr{G}(n) \otimes \mathscr{R}_+$-measurable. Moreover, the explicit form of the previous theorem shows that $X_{t\wedge T_{n+1}}^{i,A} = X_{t\wedge T_{n+1}}^{A}$ P^i-a. s.; hence if $S_A^i = \inf(t: X_t^{i,A} \neq \overline{X}_t^A)$ we deduce that $P^i(S_A^i < T_{n+1}) = 0$. Now, S_A^i is $\mathscr{G}(n)$-measurable, so by definition of G_n^i,

$$E^i(1_{\{S_A^i<\infty\}} G_n^i((S_A^i, \infty] \times E)) = P^i(S_A^i < \infty, S_A^i < T_{n+1}) = 0.$$

Therefore if $I^i(\omega) = \{t \in \mathbb{R}_+ : G_n^i([t,\infty] \times E) > 0\}$, we deduce that

1.37 $$P^i(S_A^i \in I^i) = 0.$$

Now we apply the induction hypothesis. The set $\{\omega: S_A^i(\omega) \in I^i(\omega)\}$ belongs to $\mathscr{G}(n)$, so 1.37 yields $P^j(S_A^i \in I^i) = 0$ for $i, j = 1, 2$. In other words,

1.38 $$P^i(X_t^{1,A} = X_t^{2,A} \text{ for all } t \in I^1 \cap I^2 \text{ and all } A \in \mathscr{E}_0) = 1, \quad i = 1, 2.$$

Let $U_s^i = G_n^i((s,\infty] \times E)$; applying 1.36 for $A = E$ yields $dU_s^i = -U_{s-}^i\, dX_s^{i,E}$, while $U_0^i = 1$. Hence I.4.61 implies that $U^i = \mathscr{E}(-X^{i,E})$ (Doléans-Dade exponential), and 1.38 yields

$$P^i(U_s^1 = U_s^2 \text{ for all } s \in I^1 \cap I^2) = 1, \quad i = 1, 2$$

Now, if $U_s^1 = U_s^2$ for all $s \in I^1 \cap I^2$, we also have $U_s^1 = U_s^2$ for all $s \in \mathbb{R}_+$ (because U^i is decreasing, nonnegative, and $U_s^i = 0$ if s is the right-end point of I^i). Since $I^i = \{s: U_{s-}^i > 0\}$, we deduce:

1.39 $$P^i(I^1 = I^2 \text{ and } G_n^1(\cdot \times E) = G_n^2(\cdot \times E)) = 1.$$

Finally, apply 1.36 once more: we have $G_n^i(ds \times A) = U_{s-}^i 1_{I^i}(s)\, dX_s^{i,A}$. Therefore 1.38 and 1.39 yield

$$P^i(G_n^1(\cdot \times A) = G_n^2(\cdot \times A) \text{ for all } A \in \mathscr{E}_0) = 1$$

and we deduce (because $\mathscr{E}_0$ generates $\mathscr{E}$) that $G_n^1 = G_n^2$ P^i-a. s. for $i = 1, 2$. Since $P^1 = P^2$ on $\mathscr{G}(n)$ and $\mathscr{G}(n+1) = \mathscr{G}(n) \vee \sigma(T_{n+1}, Z_{n+1})$, we easily deduce that $P^1 = P^2$ on $\mathscr{G}(n+1)$ from the definitions of G_n^1 and G_n^2, and we are done. □

2. Martingale Problems and Semimartingales

Now we introduce our second and most important class of martingale problems, namely those related to characteristics of semimartingales. Then we give some examples, and the last subsection presents the notion of "local uniqueness", a technical but very useful concept; however, this last subsection may be skipped at first reading.

§ 2a. Formulation of the Problem

In the whole section, we start with a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ and an initial σ-field $\mathscr{H}$ (see 1.1), and also an initial condition P_H (which is a probability measure on $(\Omega, \mathscr{H})$). Note once more that we do not have a measure on $(\Omega, \mathscr{F})$ yet.

Another basic ingredient is our fundamental process:

2.1 $X = (X^i)_{i \leq d}$ is a d-dimensional càdlàg adapted process on $(\Omega, \mathscr{F}, \mathbf{F})$. X has vocation to being a semimartingale, and so we fix:

2.2 $h \in \mathscr{C}_t^d$, a truncation function (see II.2.3);

2.3 A triplet (B, C, ν) (a candidate for the characteristics of X), where:

(i) $B = (B^i)_{i \leq d}$ is $\mathbf{F}$-predictable, with finite variation over finite intervals, and $B_0 = 0$;

(ii) $C = (C^{ij})_{i,j \leq d}$ is $\mathbf{F}$-predictable, continuous, $C_0 = 0$, and $C_t - C_s$ is a nonnegative symmetric $d \times d$ matrix for $s \leq t$;

(iii) ν is an $\mathbf{F}$-predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, which charges neither $\mathbb{R}_+ \times \{0\}$ nor $\{0\} \times \mathbb{R}^d$, and such that $(|x|^2 \wedge 1) * \nu_t(\omega) < \infty$ and $\int \nu(\omega; \{t\} \times dx) h(x) = \Delta B_t(\omega)$ and $\nu(\omega; \{t\} \times \mathbb{R}^d) \leq 1$ identically. □

(these properties are exactly the properties of the "nice" version of characteristics, as constructed in Proposition II.2.9).

2.4 **Definition.** A *solution to the martingale problem associated with* $(\mathscr{H}, X)$ and $(P_H; B, C, \nu)$ is a *probability measure* P on $(\Omega, \mathscr{F})$ such that:

(i) the restriction $P_{|\mathscr{H}}$ of P to $\mathscr{H}$ equals P_H;

(ii) X is a semimartingale on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with characteristics (B, C, ν) relative to the truncation function h.

We denote by $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ the set of all solutions P. □

Although "semimartingale problem" might seem more appropriate a name, the terminology "martingale problem" is commonly used for the above, for it reduces indeed to a problem of type 1.3, as seen below. Before, we recall the definition of the following càdlàg processes (see II.2.4, II.2.5 and II.2.18):

2.5
$$\begin{cases} \check{X}(h)_t = \sum_{s \leq t} [\Delta X_s - h(\Delta X_s)] \\ X(h) = X - \check{X}(h) \\ M(h) = X(h) - X_0 - B. \end{cases}$$

2.6 $$\tilde{C}^{ij} = C^{ij} + (h^i h^j) * \nu - \sum_{s \leq \cdot} \Delta B_s^i \Delta B_s^j$$

Note that $\tilde{C}$ satisfies 2.3(ii), except that it is càdlàg and not necessarily continuous. As usual, μ^X is the random measure associated with the jumps of X by II.1.16.

The following is just a reformulation of Theorem II.2.21:

2.7 **Theorem.** *A probability measure P belongs to $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ if and only if it is a solution to the martingale problem 1.3 associated with P_H and the family $\mathscr{X}$ of processes consisting in the following:*

(i) $M(h)^i$, $i \leq d$

(ii) $M(h)^i M(h)^j - \tilde{C}^{ij}$, $i, j \leq d$

(iii) $g * \mu^X - g * \nu$, $g \in \mathscr{C}^+(\mathbb{R}^d)$ (see II.2.20).

2.8 **Corollary.** *The set $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ is a convex set.*

Proof. Let P, P' be two solutions, and $Q = bP + (1-b)P'$ a convex combination, $b \in [0,1]$. That $Q_{|\mathscr{H}} = P_H$ is obvious. Let Y be any of the processes in 2.7: then $Y_0 = 0$ and $|\Delta Y|$ is bounded by construction. If $T_n = \inf(t: |Y_t| > n)$ then (T_n) is a sequence of stopping times increasing to $+\infty$, and Y^{T_n} is bounded. Then by I.1.47, Y^{T_n} is a uniformly integrable martingale for P and for P', and for any stopping time S we have $E_P(Y_S^{T_n}) = 0$. Thus

$$E_Q(Y_S^{T_n}) = bE_P(Y_S^{T_n}) + (1-b)E_{P'}(Y_S^{T_n}) = 0$$

and we deduce from I.1.44 that Y^{T_n} is a Q-martingale. Hence Y is a Q-local martingale, and $Q \in \mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ by 2.7. □

2.9 **Remarks.** 1) In some situations (as in § 2c below) the assumptions in 2.1 and 2.3 are too strong, and should be replaced by:

2.10 $$\begin{cases} X & \text{is adapted and } (\overline{\mathbb{R}})^d\text{-valued} \\ B & \text{is predictable and } (\overline{\mathbb{R}})^d\text{-valued} \\ C & \text{is predictable and } (\overline{\mathbb{R}})^{d \times d}\text{-valued} \\ \nu & \text{is a predictable random measure on } \mathbb{R}_+ \times \mathbb{R}^d. \end{cases}$$

Then in Definition 2.4, one should replace (ii) by:

(ii′) X is indistinguishable from a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, whose characteristics are indistinguishable from (B, C, ν).

Of course, if P is a solution, then X is P-a. s. càdlàg and $\mathbb{R}^d$-valued and (B, C, ν) meets 2.3 except on a P-null set. In this situation, 2.7 and 2.8 remain true, as is easily seen, provided in 2.6 we set $\tilde{C}_t^{ij} = +\infty$ whenever the right-hand side is not well defined, and in 2.7 we add:

2.11 B has P-almost surely finite variation over finite intervals.

2) One can even go further, dropping adaptedness or predictability from 2.10. Then if P is a solution, X, B, C, ν are adapted or predictable with respect to the

completed filtration $\mathbf{F}^P$ only: this rends comparison between solutions difficult, and for example 2.8 fails. □

We end this subsection with a description of additional assumptions that naturally complement 1.1 and 2.1. So far $(\Omega, \mathscr{F}, \mathbf{F})$ is arbitrary, up to the fact that it supports the adapted process X. However, except in very specific cases, we cannot hope for *uniqueness* of the solution of $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ unless we have:

2.12 **F** is generated by X and $\mathscr{H}$, by which we mean:

(i) $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0$ and $\mathscr{F}_s^0 = \mathscr{H} \vee \sigma(X_r : r \leq s)$ (in other words, **F** is the smallest filtration such that X is adapted and $\mathscr{H} \subset \mathscr{F}_0$);

(ii) $\mathscr{F} = \mathscr{F}_{\infty-} (= \bigvee_t \mathscr{F}_t)$. □

We already encountered 2.12 in § 1c. As for *existence* of a solution to our martingale problem, we need more structure on Ω, and a typical situation is as such:

2.13 *The canonical setting.* Ω is the "*canonical space*" (also denoted by $\mathbb{D}(\mathbb{R}^d)$) of all càdlàg functions $\omega: \mathbb{R}_+ \to \mathbb{R}^d$; X is the "canonical process" defined by $X_t(\omega) = \omega(t)$; $\mathscr{H} = \sigma(X_0)$; finally **F** is generated by X and $\mathscr{H}$ in the sense of 2.12. □

In the canonical setting, or more generally when $\mathscr{H} = \sigma(X_0)$, we can identify the initial measure P_H with the distribution of X_0 in $\mathbb{R}^d$:

2.14 If $\mathscr{H} = \sigma(X_0)$ and if η is a probability measure on $\mathbb{R}^d$, *we also denote by* η *the measure on* $(\Omega, \mathscr{H})$ defined by $\eta(X_0 \in A) = \eta(A)$. □

§ 2b. Example: Processes with Independent Increments

We have already encountered, and essentially solved in Chapter II, a series of martingale problems related to processes with (conditionally) independent increments. For instance, if P_H is an arbitrary probability measure on $(\Omega, \mathscr{H})$, we can re-formulate Theorem II.4.4:

2.15 **Theorem.** *Let* $d = 1$ *and assume that* $X_0 = 0$. *Let* σ^2 *be a continuous increasing function with* $\sigma^2(0) = 0$. *Then* P *belongs to* $\mathscr{s}(\mathscr{H}, X | P_H; 0, \sigma^2, 0)$ *if and only if* X *is* (indistinguishable from) *a Wiener process with variance function* σ^2 *on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

(this is essentially the same as Example 1.4).

More generally, we can write Theorems II.4.15 and II.4.25 and II.5.2b in our present setting:

2.16 **Theorem.** *Let (B, C, ν) meet 2.3 and be deterministic.*

a) *P belongs to $\mathfrak{s}(\mathscr{H}, X | P_H; B, C, \nu)$ if and only if $P_{|\mathscr{H}} = P_H$ and $X - X_0$ is a PII on $(\Omega, F, \mathbf{F}, P)$ whose distribution is described by II.4.16.*

b) *Assume further that $\mathscr{F}$ is generated by X and $\mathscr{H}$* (see 2.12). *Then $\mathfrak{s}(\mathscr{H}, X | P_H; B, C, \nu)$ contains at most one element P.*

c) *Assume further the canonical setting 2.13. Then for any probability measure η on $\mathbb{R}^d$, $\mathfrak{s}(\mathscr{H}, X | \eta; B, C, \nu)$ contains one and only one solution.*

For processes with conditionally independent increments we have similar results; actually, we state only the analogues to (a) and (b) above (they are re-statements of II.6.6 and II.6.11: note that 2.12 implies Hypothesis II.6.4).

2.17 **Theorem.** *Assume that $\mathscr{F}$ is generated by X and $\mathscr{H}$* (see 2.12), *and let (B, C, ν) meet 2.3 and be $\mathscr{H}$-measurable. Then $\mathfrak{s}(\mathscr{H}, X | P_H; B, C, \nu)$ contains at most one element P, in which case $X - X_0$ is an $\mathscr{H}$-PII on $(\Omega, \mathscr{F}, \mathbf{F}, P)$.*

§ 2c. Diffusion Processes and Diffusion Processes with Jumps

Diffusions and diffusions with jumps rank among first in importance: many Markov processes are diffusions with jumps (also called "Ito processes"); they also are the solutions to a wide class of stochastic differential equations, hence their usefulness in applications; finally they provide most of the non-trivial examples for martingale problems.

Of course it is outside the scope of this book to develop any sizeable portion of the theory of diffusion processes. We simply account here for some of their basic properties, in connection with martingale problems. The reference books are Stroock and Varadhan [232], Dynkin [47], Liptcer and Shiryaev [158] or Gikhman and Skorokhod [61, 62] and also the paper [231] or [98, chapters XIII and XIV] for diffusions with jumps, and the more recent book [283] by Revuz and Yor.

In all this subsection, the *canonical setting* of 2.13 is in force.

2.18 **Definition.** Let P be a probability measure on $(\Omega, \mathscr{F})$. Then X is called a *diffusion with jumps* on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ if it is a semimartingale with the following characteristics (recall that the truncation function h is fixed):

2.19
$$\begin{cases} B_t^i(\omega) = \int_0^t b^i(s, X_s(\omega))\, ds \quad (= +\infty \text{ if the integral diverges}) \\[2ex] C_t^{ij}(\omega) = \int_0^t c^{ij}(s, X_s(\omega))\, ds \quad (= +\infty \text{ if the integral diverges}) \\[2ex] \nu(\omega; dt \times dx) = dt \times K_t(X_t(\omega), dx) \end{cases}$$

where:

$$\begin{cases} b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \quad \text{is Borel} \\ c\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d \quad \text{is Borel, } c(s,x) \text{ is symmetric nonnegative} \\ K_t(x,dy) \quad \text{is a Borel transition kernel from } \mathbb{R}_+ \times \mathbb{R}^d \text{ into } \mathbb{R}^d\text{, with} \\ \qquad K_t(x,\{0\}) = 0. \end{cases}$$

Moreover,

a) if $\nu = 0$, X is called a *diffusion* (it is then a. s. continuous);

b) if $b(s,x)$, $c(s,x)$, $K_s(x,dy)$ do not depend upon s, X is called an *homogeneous* diffusion (with jumps). □

Observe that (B,C,ν) might not meet all conditions in 2.3, but it certainly satisfies the weaker conditions stated in Remark 2.9. So if η is the law $\mathscr{L}(X_0)$ under P, we will also write $P \in \mathscr{s}(\mathscr{H}, X|\eta; B, C, \nu)$ (recall that $\mathscr{H} = \sigma(X_0)$ and that we identify a measure on $\mathbb{R}^d$ with its preimage on $(\Omega, \mathscr{H})$ by X_0).

Next, we proceed to relate the above with stochastic differential equations. For this, let $\mathscr{B}' = (\Omega', \mathscr{F}', \mathbf{F}', P')$ be another stochastic basis, endowed with:

2.20 The *driving terms*:

(i) $W = (W^i)_{i \le m}$, an m-dimensional standard Wiener process (i.e., each W^i is a standard Wiener process, and the W^i's are independent);

(ii) $\not{p}$, a Poisson random measure on $\mathbb{R}_+ \times E$ with intensity measure $q(dt,dx) = dt \otimes F(dx)$; here, $(E,\mathscr{E})$ is an arbitrary Blackwell space (one may take $E = \mathbb{R}$ in *all* cases if one wishes), and F is a positive σ-finite measure on $(E,\mathscr{E})$. □

Let us also be given:

2.21 The *coefficients*:

$$\begin{cases} \beta = (\beta^i)_{i \le d}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \\ \gamma = (\gamma^{ij})_{i \le d, j \le m}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^m \\ \delta = (\delta^i)_{i \le d}, \text{ a Borel function: } \mathbb{R}_+ \times \mathbb{R}^d \times E \to \mathbb{R}^d. \end{cases}$$ □

2.22 The *initial variable*: it is an $\mathscr{F}'_0$-measurable $\mathbb{R}^d$-valued random variable ξ. □

Recall that h is the truncation function, and set $h'(x) = x - h(x)$. The stochastic differential equation is as follows:

2.23
$$\begin{cases} Y_0 = \xi \\ dY_t = \beta(t,Y_t)\,dt + \gamma(t,Y_t)\,dW_t + h \circ \delta(t, Y_{t-}, z)(\not{p}(dt,dz) - q(dt,dz)) \\ \qquad + h' \circ \delta(t, Y_{t-}, z)\not{p}(dt,dz) \end{cases}$$

(hence if $\not{p}$ has a "jump" at point (t,z), then $\Delta Y_t = \delta(t, Y_{t-}, z)$).

2.24 **Definition.** a) A *solution-process* (or: *strong solution*) to 2.23, on the basis $\mathscr{B}'$ and relative to the driving terms $(W, \not p)$, is a càdlàg adapted process Y such that for each $i \leq d$,

2.25 $$Y^i = \xi^i + \beta^i(Y)\cdot t + \sum_{j\leq m} \gamma^{ij}(Y_-)\cdot W^j + h^i\circ\delta(Y_-)*(\not p - q) + h'^i\circ\delta(Y_-)*\not p.$$

Of course, it is implicit that the various integrals make sense: $\beta^i(s, Y_s)$ is Stieltjes-integrable with respect to ds, and $\gamma^{ij}(s, Y_{s-})\in L^2_{\text{loc}}(W^j)$, and $h^i\circ\delta(s, Y_{s-}, z)\in G_{\text{loc}}(\not p)$, and $h'^i\circ\delta(s, Y_{s-}, z)$ is Stieltjes-integrable with respect to $\not p$.

b) A *solution-measure* (or: *weak solution*) to 2.23 *with initial condition* η (a probability measure on $\mathbb{R}^d$) is a probability measure P on $(\Omega, \mathscr{F})$ (the canonical space) with the following property: there exists a stochastic basis $\mathscr{B}'$ with driving terms $(W, \not p)$ meeting 2.20 and with an $\mathscr{F}'_0$-measurable variable ξ meeting $\mathscr{L}(\xi) = \eta$, and with a solution-processs Y on $\mathscr{B}'$, such that P be the law of Y (i.e.: P is the image of P' under the map: $\omega'\in\Omega' \rightsquigarrow Y(\omega')\in\Omega$). □

Now we review some of the main properties (see e.g. [98]).

2.26 **Theorem.** *Let* η *be an initial condition* (a probability on $\mathbb{R}^d$), *and* β, γ, δ *be coefficients* (see 2.21).

The set of all solution-measures to 2.23 with initial condition η *is the set* $\mathscr{s}(\mathscr{H}, X|\eta; B, C, \nu)$ *of all solutions to a martingale problem on the canonical space, where* (B, C, ν) *are given by 2.19, with*

2.27 $$\begin{cases} b = \beta, \quad c = \gamma\gamma^T \quad \left(\text{i.e. } c^{ij} = \sum_{1\leq k\leq m} \gamma^{ik}\gamma^{jk}\right) \\ K_t(y, A) = \int 1_{A\setminus\{0\}}(\delta(t, y, z))F(dz). \end{cases}$$

2.28 **Remarks.** 1) Let Y be a solution-process. Then 2.25 gives the canonical representation II.2.35 of Y, namely

$$Y_0 = \xi, \quad Y^{i,c} = \sum_{j\leq m} \gamma^{ij}(Y_-)\cdot W^j, \quad h*(\mu^Y - \nu) = h\circ\delta(Y_-)*(\not p - q)$$
$$(x - h(x))*\mu^Y = h'\circ\delta(Y_-)*\not p, \quad B = \beta(Y)\cdot t$$

(immediate to check).

2) That a solution-measure belongs to $\mathscr{s}(\mathscr{H}, X|\eta; B, C, \nu)$ is easy: indeed if Y is a solution-process, it follows from 1) above that its characteristics are given by 2.19, with Y in place of X, provided b, c, K are given by 2.27. Then it remains to transport the characteristics onto Ω, which is not very difficult.

3) 2.27 gives (b, c, K) is terms of (β, γ, δ). Quite often one rather starts with a diffusion process with jumps in the sense of 2.18, i.e. a solution to $\mathscr{s}(\mathscr{H}, X|\eta; B, C, \nu)$, where the given data are (b, c, K). Then, are there coefficients (β, γ, δ) such that 2.27 holds? the answer is: yes. That γ exists, with $c = \gamma\gamma^T$, is well known (provided $m \geq d$, or at least $m \geq$ maximal rank of $c(s, y)$). That δ exists is also

"classical", provided the measure F is infinite and has no atom: see [49, 70, 98, 225]. □

2.29 **Remark.** Very often one considers an equation 2.23 with $h(x) = x$, and so $h'(x) = 0$, and the last term disappears. Of course, h is no longer a truncation function. Nevertheless, the set of all solutions-measures of

$$\text{2.30} \qquad \begin{cases} Y_0 = \xi \\ dY_t = \beta(t, Y_t)\,dt + \gamma(t, Y_{t-})\,dW_t + \delta(t, Y_{t-}, z)(\mu(dt, dz) - q(dt, dz)) \end{cases}$$

where ξ is a random variable with a given distribution η is exactly $\mathscr{s}(\mathscr{H}, X|\eta; B, C, \nu)$ with (B, C, ν) given by 2.19, provided c and K are given by 2.27 and

$$\text{2.31} \qquad b(t, y) = \beta(t, y) + \int K_t(y, dy')[h(y') - y']$$

(this comes from II.2.29a). □

The following is a classical result of existence of solution-processes in case of Lipschitz coefficients.

2.32 **Theorem.** *Assume the following two conditions*

(i) *Local Lipschitz coefficients.* For each $n \in \mathbb{N}^*$ there is a constant θ_n and a function $\rho_n\colon E \to \mathbb{R}_+$ with $\int \rho_n(z)^2 F(dz) < \infty$, such that for $t \leq n$, $|y| \leq n$, $|y'| \leq n$:

$$|\beta(s, y) - \beta(s, y')| \leq \theta_n |y - y'|, \quad |\gamma(s, y) - \gamma(s, y')| \leq \theta_n |y - y'|$$
$$|h \circ \delta(s, y, z) - h \circ \delta(s, y', z)| \leq \rho_n(z) |y - y'|$$
$$|h' \circ \delta(s, y, z) - h' \circ \delta(s, y', z)| \leq \rho_n(z)^2 |y - y'|$$

(ii) *Linear growth.* for each $n \in \mathbb{N}^*$ there are θ_n and ρ_n as above, such that for all $t \leq n$ and all $y \in \mathbb{R}^d$:

$$|\beta(s, y)| \leq \theta_n(1 + |y|), \quad |\gamma(s, y)| \leq \theta_n(1 + |y|)$$
$$|h \circ \delta(s, y, z)| \leq \rho_n(z)(1 + |y|), \quad |h' \circ \delta(s, y, z)| \leq [\rho_n(z)^2 \wedge \rho_n(z)^4](1 + |y|).$$

Then 2.23 has a solution-process Y, and only one (up to indistinguishability) *on any stochastic basis $\mathscr{B}'$ supporting driving terms (W, μ) as in 2.20.*

In the above case, we also have uniqueness of the solution-measure, as deduced from the following result:

2.33 **Theorem.** *Suppose that on any stochastic basis $\mathscr{B}'$ supporting driving terms (W, μ) as in 2.20, there is at most one solution-process* (up to indistinguishability). *Then, if there is a solution-measure, with a given initial condition, this solution-measure is unique.*

Finally we state an existence and uniqueness result (we do not give the sharpest possible result).

2.34 **Theorem.** *Assume that b is bounded, that c is bounded and continuous on $\mathbb{R}_+ \times \mathbb{R}^d$ and everywhere invertible, and that the functions $(t, y) \rightsquigarrow \int_A (|z|^2 \wedge 1) K_t(y, dz)$ are bounded and continuous for all $A \in \mathscr{R}^d$.*

Then, there is a transition kernel $P_{x,r}(d\omega)$ from $(\mathbb{R}^d \times \mathbb{R}_+, \mathscr{R}^d \otimes \mathscr{R}_+)$ into $(\Omega, \mathscr{F})$ (the canonical space) *with the following property:*

For every (x, r), $P_{x,r}$ is the unique probability measure under which the canonical process X is a diffusion with jumps, with $P_{x,r}(X_0 = x) = 1$ and with characteristics given by 2.19 where $b(s, y)$, $c(s, y)$, $K_s(y, dy')$ are replaced by $b(r + s, y)$, $c(r + s, y)$, $K_{r+s}(y, dy')$.

Then in particular, $P_{x,0}$ is the unique solution of $\mathscr{s}(\mathscr{H}, X | \varepsilon_x; B, C, \nu)$ where (B, C, ν) are given by 2.19.

The conditions above on (b, c, K) are essentially *much weaker* than the conditions in 2.32 ("essentially" only, because of the uniform boundedness imposed upon the coefficients; those could be relaxed in the same way than in 2.32(iv)), except for a *very important point*, namely that c is nondegenerate (while in 2.32 one could have $\gamma = 0$ and $c = 0$). So the combination of 2.32 and 2.33 gives an existence and uniqueness result that indeed is different from 2.34.

§ 2d. Local Uniqueness

1. For "technical" reasons, useful for limit theorems as well as for studying absolute continuity or singularity questions, we need a form of uniqueness for martingale problems that is stronger than the "ordinary" uniqueness of the solution to $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$.

The setting is as in § 2a, and we assume further that $\mathbf{F}$ is generated by X and $\mathscr{H}$ (2.12). In the sequel the stopping times with respect to the "filtration" $(\mathscr{F}_t^0)_{t \in \mathbb{R}_+}$ play a central rôle: to emphasize the fact that $(\mathscr{F}_t^0)$ is not a filtration in our sense (it is not right-continuous, in general), we give a specific name to its stopping times:

2.35 **Definition.** A *strict stopping time* (or, a stopping time relative to $(\mathscr{F}_t^0)_{t \geq 0}$) is a map $T\colon \Omega \to \overline{\mathbb{R}}_+$ such that $\{T \leq t\} \in \mathscr{F}_t^0$ for all $t \in \mathbb{R}_+$. If T is a strict stopping time, then $\mathscr{F}_T^0$ denotes the σ-field of all $A \in \mathscr{F}$ such that $A \cap \{T \leq t\}$ belongs to $\mathscr{F}_t^0$ for all $t \in \mathbb{R}_+$. □

Since $\mathscr{F}_{t-} \subset \mathscr{F}_t^0 \subset \mathscr{F}_t$ for $t > 0$, we readily obtain:

2.36 —A strict stopping time is a stopping time; $\mathscr{F}_T^0 \subset \mathscr{F}_T$; and $\mathscr{F}_{T-} \subset \mathscr{F}_T^0$ on the set $\{T > 0\}$.

—A predictable time T is a strict stopping time if $\{T = 0\} \in \mathscr{F}_O^0$. □

Let (B, C, ν) be a triplet of type 2.3. If T is a stopping time, we know the stopped processes X^T, B^T, C^T, and the "stopped random measure" ν^T is defined as

$$\nu^T(\omega; ds, dx) = \nu(\omega; ds, dx) 1_{\{s \le T(\omega)\}}$$

(hence $W * \nu^T = (W * \nu)^T$).

2.37 **Definition.** Assume 2.12. We say that *local uniqueness* holds for the martingale problem 2.4 if, for every *strict* stopping time T, any two solutions P and P' of the "stopped" martingale problem $\mathscr{s}(\mathscr{H}, X^T | P_H; B^T, C^T, \nu^T)$ coincide on the σ-field $\mathscr{F}_T^0$.

(Local uniqueness implies uniqueness: take $T \equiv \infty$). □

2. A general result. It is apparent from the very definition that local uniqueness is not going to be easy to check, unless we can prove that it is implied by uniqueness. We cannot hope for that in general. However, it is true when the martingale problem has a sort of "Markovian" type which we now describe.

From now on we suppose that $(\Omega, \mathscr{F}, \mathbf{F})$ is the canonical space, with the canonical process X and $\mathscr{H} = \mathscr{F}_0^0$: see 2.13. For each $t \in \mathbb{R}_+$ there is a mapping (the "shift"):

2.38 $\qquad \theta_t: \Omega \to \Omega$ defined by $X_s \circ \theta_t(\omega) = X_{s+t}(\omega) \quad \forall s \ge 0, \forall \omega \in \Omega.$

We are given a triplet (B, C, ν) of type 2.3. For each $t \in \mathbb{R}_+$ we also have a triplet $(p_t B, p_t C, p_t \nu)$ of type 2.3, such that

2.39 (i) the mappings $(\omega, t) \rightsquigarrow (p_t B)_s(\omega)$, and $(\omega, t) \rightsquigarrow (p_t C)_s(\omega)$ (for $s \in \mathbb{R}_+$), and $(\omega, t) \rightsquigarrow (p_t \nu)(\omega, A)$ (for $A \in \mathscr{R}_+ \otimes \mathscr{R}^d$) are $\mathscr{F} \otimes \mathscr{R}_+$-measurable;

(ii) for all $\omega \in \Omega$, $s \in \mathbb{R}_+$, $A \in \mathscr{R}^d$,

$$(p_t B)_s(\theta_t \omega) = B_{t+s}(\omega) - B_t(\omega)$$

$$(p_t C)_s(\theta_t \omega) = C_{t+s}(\omega) - C_t(\omega)$$

$$(p_t \nu)(\theta_t \omega; (0, s] \times A) = \nu(\omega; (t, t+s] \times A). \qquad \square$$

(In particular, $p_0 B = B$, $p_0 C = C$, $p_0 \nu = \nu$).

2.40 **Theorem.** *In addition to the above assumptions* (namely, 2.13 and 2.39), *suppose that there is a transition kernel $P_{x,t}(d\omega)$ from $(\mathbb{R}^d \times \mathbb{R}_+, \mathscr{R}^d \otimes \mathscr{R}_+)$ into $(\Omega, \mathscr{F})$ such that for every (x, t), $P_{x,t}$ is a solution to the martingale problem $\mathscr{s}(\mathscr{H}, X | \varepsilon_x; p_t B, p_t C, p_t \nu)$.*

If furthermore the problem $\mathscr{s}(\mathscr{H}, X | \varepsilon_x; B, C, \nu)$ has a unique solution (which necessarily is $P_{x,0}$!), *then local uniqueness holds for this problem.*

Before proceeding to the proof, we give two examples.

2.41 Corollary (Diffusions with jumps): *Under the assumptions of 2.34, we have existence and local uniqueness.*

Proof. It suffices to observe that the triplet

$$(p_t B)_s(\omega) = \int_0^s b(t+r, X_r(\omega))\,dr, \quad (p_t C)_s(\omega) = \int_0^s c(t+r, X_r(\omega))\,dr$$
$$(p_t \nu)(\omega; ds, dy) = ds\, K_{t+s}(X_s(\omega), dy)$$

satisfies 2.39. □

2.42 Corollary. *If (B, C, ν) are deterministic, we have existence and local uniqueness.*

Proof. We set

$$(p_t B)_s = B_{t+s} - B_t, \quad (p_t C)_s = C_{t+s} - C_t, \quad p_t \nu([0,s] \times A) = \nu((t, t+s] \times A),$$

so $(p_t B, p_t C, p_t \nu)$ is deterministic and obviously satisfies 2.39. By 2.16 $\mathscr{s}(\mathscr{H}, X | \varepsilon_x; p_t B, p_t C, p_t \nu)$ admits a unique solution $P_{x,t}$, for which $X - X_0$ is a PII. Moreover, if $h(u)_{s,t}$ is the right-hand side of II.4.16 and $E_{x,t}$ denotes the expectation with respect to $P_{x,t}$, then II.4.16 yields for $0 = s_0 < \cdots < s_p$, $u_p \in \mathbb{R}^d$:

$$E_{x,t}\left(\exp i\left\{u_0 \cdot X_0 + \sum_{1 \le j \le p} u_j \cdot (X_{t_j} - X_{t_{j-1}})\right\}\right)$$
$$= [\exp i u_0 \cdot x] \prod_{1 \le j \le p} h(u_j)_{t+t_{j-1}, t+t_j},$$

which is Borel in (x, t); we easily deduce that $P_{x,t}(d\omega)$ is a kernel, and the claim follows from 2.16 and 2.40. □

3. The proof of Theorem 2.40 is broken into several steps, the idea being as follows: let T be a strict stopping time, and P be a solution of the stopped problem. Then we call Q the law of the process X obtained by pasting together at time T the process X^T under P, and the process X under $P_{X_T, T}$, and we prove that Q is the (unique) solution to $\mathscr{s}(\mathscr{H}, X | \varepsilon_x; B, C, \nu)$: so the restriction of P to $\mathscr{F}_{T-}$, which equals the restriction of Q to $\mathscr{F}_{T-}$, is uniquely determined. This is a technique that is well known in the theory of Markov processes; the first two lemmas below have been originally designed for Markov processes.

2.43 Lemma. a) *Let $A \in \mathscr{F}$. Then A belongs to $\mathscr{F}_t^0$ if and only if, for every pair (ω, ω') such that $X_s(\omega) = X_s(\omega')$ $\forall s \le t$, we have $(\omega \in A) \Leftrightarrow (\omega' \in A)$.*

b) *Let T be an $\mathscr{F}$-measurable mapping: $\Omega \to \overline{\mathbb{R}}_+$. Then T is a strict stopping time if and only if, for every pair (ω, ω') such that $X_s(\omega) = X_s(\omega')$ $\forall s \le T(\omega)$, then $T(\omega) = T(\omega')$.*

c) *If T is a strict stopping time and $A \in \mathscr{F}$, then A belongs to $\mathscr{F}_T^0$ if and only if, for every pair (ω, ω') such that $X_s(\omega) = X_s(\omega')\ \forall s \leq T(\omega) (= T(\omega')$ by (b)), then $(\omega \in A) \Leftrightarrow (\omega' \in A)$.*

d) *Let H be a predictable process, and $t > 0$, and (ω, ω') such that $X_s(\omega) = X_s(\omega')\ \forall s < t$. Then $H_t(\omega') = H_t(\omega)$.*

Proof. a) Define $a_t \colon \Omega \to \Omega$ by $X_s \circ a_t(\omega) = X_{s \wedge t}(\omega)$ for all s (the "stopping operator"). For all s, $X_s \circ a_t$ is $\mathscr{F}_t{}^0$-measurable; for $s \leq t$, $X_s = X_s \circ a_t$ is $a_t^{-1}(\mathscr{F})$-measurable: hence $\mathscr{F}_t{}^0 = a_t^{-1}(\mathscr{F})$.

If $A \in \mathscr{F}$ satisfies the stated condition, then $1_A = 1_A \circ a_t$ and so $A \in \mathscr{F}_t{}^0$. Conversely if $A \in \mathscr{F}_t{}^0$ we have $1_A = 1_B \circ a_t$ for some $B \in \mathscr{F}$ by what precedes, and we deduce that A meets the stated condition.

(b, c) Suppose first that T is a strict stopping time, and let (ω, ω') with $X_s(\omega) = X_s(\omega')\ \forall s \leq T(\omega)$. Let $t = T(\omega)$. If $t = \infty$ then $\omega' = \omega$ and $T(\omega') = \infty$. If $t < \infty$, then $\omega \in \{T = t\} \in \mathscr{F}_t{}^0$, hence $\omega' \in \{T = t\}$ by (a) and again $T(\omega') = T(\omega)$: so we have proved the necessary condition of (b). Moreover, let $A \in \mathscr{F}_T^0$ and $\omega \in A$, so $\omega \in A \cap \{T = t\} \in \mathscr{F}_t{}^0$, so $\omega' \in A$ by (a). Conversely let $A \in \mathscr{F}$ satisfy the stated property in (c), and set $A_t = A \cap \{T \leq t\}$. If (ω, ω') are such that $X_s(\omega) = X_s(\omega')\ \forall s \leq t$ and $\omega \in A_t$, then $X_s(\omega) = X_s(\omega')\ \forall s \leq T(\omega)$, hence $\omega' \in A_t$ because $T(\omega') = T(\omega)$ and because of our hypothesis on A: hence (a) yields $A_t \in \mathscr{F}_t{}^0$, thus $A \in \mathscr{F}_T^0$.

It remains to prove the sufficient condition in (b). Let $t \geq 0$ and (ω, ω') with $X_s(\omega) = X_s(\omega')\ \forall s \leq t$. If $\omega \in \{T \leq t\}$, then $X_s(\omega) = X_s(\omega')\ \forall s \leq T(\omega)$, hence $T(\omega') = T(\omega)$ and $\omega' \in \{T \leq t\}$ by hypothesis: thus (a) yields that $\{T \leq t\} \in \mathscr{F}_t{}^0$, and we deduce that T is a strict stopping time.

d) By I.2.2 it suffices to consider two cases: (1) $H = 1_{A \times \{0\}}$ $(A \in \mathscr{F}_0)$, then $H_t \equiv 0$. (2) $H = 1_{A \times (u,v]}$ $(A \in \mathscr{F}_u)$. Then $H_t \equiv 0$ if $t \notin (u, v]$; if $u < t \leq v$, $H_t = 1_A$ and $A \in \mathscr{F}_t{}^0$: so the claim follows from (a). □

2.44 Lemma. *Let T be a strict stopping time.*

a) *X_T is $\mathscr{F}_T^0$-measurable.*

b) *The traces $\{T < \infty\} \cap \mathscr{F}$ and $\{T < \infty\} \cap (\mathscr{F}_T^0 \vee \theta_T^{-1}(\mathscr{F}))$ are equal.*

c) *If S is another strict stopping time, there is an $\mathscr{F}_T^0 \otimes \mathscr{F}$-measurable mapping $V \colon \Omega \times \Omega \to \overline{\mathbb{R}}_+$, such that $V(\omega, \cdot)$ is a strict stopping time for all $\omega \in \Omega$, and $S(\omega) \vee T(\omega) = T(\omega) + V(\omega, \theta_T \omega)$ on $\{T < \infty\}$.*

Proof. a) follows immediately from 2.43c.

b) It obviously suffices to consider the case when $T < \infty$ identically. $\mathscr{F}_T^0 \subset \mathscr{F}$ is obvious. If $A \in \mathscr{R}^d$ then $\theta_T^{-1}(X_s \in A) = \{X_{T+s} \in A\} \in \mathscr{F}_{T+s}^0$ by (a), so $\theta_T^{-1}(\mathscr{F}) \subset \mathscr{F}$.

For the converse, let $A \in \mathscr{R}^d$. Then $\{X_t \in A\} = \{X_t \in A, T \geq t\} \cup \{X_t \in A, T < t\}$. That $\{X_t \in A, T \geq t\} \in \mathscr{F}_T{}^0$ is obvious. The set $B = \{(\omega, \omega') : T(\omega) < t, X_{t - T(\omega)}(\omega') \in A\}$ belongs to $\mathscr{F}_T{}^0 \otimes \mathscr{F}$, and $\{X_t \in A, T < t\} = \{\omega : (\omega, \theta_T \omega) \in B\} \in \mathscr{F}_T{}^0 \vee \theta_T^{-1}(\mathscr{F})$. Thus $\{X_t \in A\} \in \mathscr{F}_T{}^0 \vee \theta_T^{-1}(\mathscr{F})$, and since t is arbitrary we deduce that $\mathscr{F} \subset \mathscr{F}_T{}^0 \vee \theta_T^{-1}(\mathscr{F})$.

c) It suffices to consider the case when $T < \infty$ identically, and $S \geq T$. $S - T$ is $\mathscr{F}$-measurable, hence (b) implies that $S(\omega) - T(\omega) = V(\omega, \theta_T\omega)$ where V is $\mathscr{F}_T \otimes \mathscr{F}$-measurable: $\Omega \times \Omega \to \bar{\mathbb{R}}_+$. By (a), the set $B = \{(\omega, \omega'): X_T(\omega) = X_0(\omega')\}$ belongs to $\mathscr{F}_T^0 \otimes \mathscr{F}$, so we may replace V by $V1_B$ without altering the measurability nor $S(\omega) = T(\omega) + V(\omega, \theta_T\omega)$.

Fix ω, ω'_1, ω'_2 with $X_s(\omega'_1) = X_s(\omega'_2)$ $\forall s \leq V(\omega, \omega'_1)$. In view of 2.43b, it remains to prove that $V(\omega, \omega'_1) = V(\omega, \omega'_2)$.

If $X_0(\omega'_1) \neq X_T(\omega)$, then $X_0(\omega'_2) \neq X_T(\omega)$ and $V(\omega, \omega'_1) = V(\omega, \omega'_2) = 0$ (because $V = V1_B$). Assume now that $X_0(\omega'_i) = X_T(\omega)$ for $i = 1, 2$. Define $\tilde{\omega}_i \in \Omega$ by:

$$X_s(\tilde{\omega}_i) = \begin{cases} X_s(\omega) & \text{if } s \leq T(\omega) \\ X_{s-T(\omega)}(\omega'_i) & \text{if } s > T(\omega). \end{cases}$$

Hence $X_s(\tilde{\omega}_1) = X_s(\tilde{\omega}_2) = X_s(\omega)$ for $s \leq T(\omega)$, and the $\mathscr{F}_T^0 \otimes \mathscr{F}$-measurability of V yields that $V(\tilde{\omega}_1, \cdot) = V(\tilde{\omega}_2, \cdot) = V(\omega, \cdot)$, while 2.43b yields $T(\omega) = T(\tilde{\omega}_1) = T(\tilde{\omega}_2)$. Hence $\theta_T\tilde{\omega}_i = \omega'_i$, and thus

2.45 $$S(\tilde{\omega}_i) = T(\tilde{\omega}_i) + V(\tilde{\omega}_i, \theta_T\tilde{\omega}_i) = T(\omega) + V(\omega, \omega'_i).$$

Therefore the assumptions on ω'_1, ω'_2 give $X_s(\tilde{\omega}_1) = X_s(\tilde{\omega}_2)$ $\forall s \leq S(\tilde{\omega}_1)$. So 2.43b implies $S(\tilde{\omega}_2) = S(\tilde{\omega}_1)$, and 2.45 yields $V(\omega, \omega'_2) = V(\omega, \omega'_1)$, which is the desired result. □

In the following lemma, P is a probability measure on $(\Omega, \mathscr{F})$, and $P_{x,t}(d\omega)$ is a transition probability measure from $(\mathbb{R}^d \times \mathbb{R}_+, \mathscr{R}^d \otimes \mathscr{R}_+)$ into $(\Omega, \mathscr{F})$ such that $P_{x,t}(X_0 = x) = 1$. T is a *strict stopping time*. From 2.44, to each $A \in \mathscr{F}$ one may associate $\bar{A} \in \mathscr{F}_T^0 \otimes \mathscr{F}$ such that $A \cap \{T < \infty\} = \{\omega: T(\omega) < \infty, (\omega, \theta_T\omega) \in \bar{A}\}$ ($\bar{A}$ is not necessarily unique), and we set:

2.46 $$Q(A) = P(A \cap \{T = \infty\}) + \int P(d\omega) P_{X_T(\omega), T(\omega)}(d\omega') 1_{\{T(\omega) < \infty\}} 1_{\bar{A}}(\omega, \omega').$$

2.47 **Lemma.** *2.46 defines a probability measure on* $(\Omega, \mathscr{F})$.

Proof. Since $\bar{Q}(\bar{A}) = \int P(d\omega) P_{X_T(\omega), T(\omega)}(d\omega') 1_{\{T(\omega) < \infty\}} 1_{\bar{A}}(\omega, \omega')$ obviously defines a measure on $(\Omega \times \Omega, \mathscr{F}_T^0 \otimes \mathscr{F})$, with mass $\bar{Q}(\Omega \times \Omega) = P(T < \infty)$, it suffices to prove that 2.46 defines $Q(A)$ unambiguously, or, equivalently, that $\bar{Q}(\bar{A}) = 0$ if $A = \{\omega: T(\omega) < \infty, (\omega, \theta_T\omega) \in \bar{A}\}$ is empty.

Let $\bar{A}$ as above, with $A = \emptyset$. If $\bar{Q}(\bar{A}) > 0$, there are ω, ω' with $(\omega, \omega') \in \bar{A}$, $T(\omega) < \infty$, $X_0(\omega) = X_T(\omega')$ (because $P_{x,t}(X_0 = x) = 1$). Define $\tilde{\omega} \in \Omega$ by

$$X_t(\tilde{\omega}) = \begin{cases} X_t(\omega) & \text{if } t \leq T(\omega) \\ X_{t-T(\omega)}(\omega') & \text{if } t > T(\omega). \end{cases}$$

Exactly as in the proof of 2.44c, we obtain that $T(\tilde{\omega}) = T(\omega)$, and $\theta_T(\tilde{\omega}) = \omega'$ and $(\tilde{\omega}, \omega') \in \bar{A}$, thus $(\tilde{\omega}, \theta_T\tilde{\omega}) \in \bar{A}$ and A contains $\tilde{\omega}$, which contradicts the assumption $A = \emptyset$. □

2.48 Lemma. *In addition to the assumptions in 2.47, we consider a family of càdlàg processes N and $p_t N$ ($t \in \mathbb{R}_+$) such that*

(i) $N_0 = (p_t N)_0 = 0$; N_s and $(p_t N)_s$ are $\mathscr{F}_s^0$-measurable;

(ii) $(\omega, t) \rightsquigarrow (p_t N)_s(\omega)$ is $\mathscr{F} \otimes \mathscr{R}_+$-measurable;

(iii) $|\Delta N| \leq a$, $|\Delta(p_t N)| \leq a$ for some constant a.

Then if N is a P-local martingale and $p_t N$ is a $P_{x,t}$-local martingale for all (x,t), the following defines a Q-local martingale:

2.49 $$\tilde{N}_t(\omega) = \begin{cases} N_t(\omega) & \text{if } t < T(\omega) \\ N_T(\omega) + (p_{T(\omega)}N)_{t-T(\omega)}(\theta_T \omega) & \text{if } t \geq T(\omega). \end{cases}$$

Proof. $\tilde{N}$ is càdlàg and $\tilde{N}_0 = 0$. By 2.44b, $\tilde{N}_t$ is $\mathscr{F}$-measurable. Let ω, ω' be such that $X_s(\omega) = X_s(\omega')$ $\forall s \leq t$. Then (i) yields that $N_s(\omega) = N_s(\omega')$ and $(p_r N)_s(\omega) = (p_r N)_s(\omega')$ for all $s \leq t, r \in \mathbb{R}_+$. Then $N_T(\omega) = N_T(\omega')$ and $T(\omega) = T(\omega')$ if $T(\omega) \leq t$ (thus N_T is $\mathscr{F}_T^0$-measurable) and $\tilde{N}_t(\omega) = \tilde{N}_t(\omega')$, so we deduce that $\tilde{N}_t$ is $\mathscr{F}_t^0$-measurable (apply 2.43).

Set $T_\rho = \inf(t: |\tilde{N}_t| > \rho)$, which is an $\mathbf{F}$-stopping time, and $R_n = \lim_{m \uparrow \infty} T_{n-1/m}$, which also equals: $\inf(t: |\tilde{N}_t| \geq n \text{ or } |\tilde{N}_{t-}| \geq n)$. Thus if $X_s(\omega) = X_s(\omega')$ for all $s \leq R_n(\omega)$, we also have $\tilde{N}_s(\omega) = \tilde{N}_s(\omega')$ for all $s \leq R_n(\omega)$, and so $R_n(\omega) = R_n(\omega')$: therefore R_n is a strict stopping time, and so is $S_n = n \wedge R_n$. We call V_n the mapping associated to S_n by 2.44c.

Let S be another strict stopping time, and V the mapping associated to it by 2.44c. If E_Q and $E_{x,t}$ denote the expectations with respect to Q and $P_{x,t}$, we can write

$$\begin{aligned} E_Q(\tilde{N}_S^{S_n}) &= E_Q(\tilde{N}_{S \wedge S_n}) = E_Q(N_{S \wedge T \wedge S_n}) + E_Q(\tilde{N}_{S \wedge S_n} - \tilde{N}_{S \wedge T \wedge S_n}) \\ &= 0 + E_Q(1_{\{S \wedge S_n > T\}}(p_T N)_{V(\cdot,\theta_T) \wedge V_n(\cdot,\theta_T)} \circ \theta_T) \end{aligned}$$

(because $N^{S_n \wedge T}$ is a bounded P-martingale, and by definition of V and V_n),

$$= \int P(d\omega) 1_{\{S_n(\omega) \wedge S(\omega) > T(\omega)\}} E_{X_{T(\omega)}, T(\omega)}[(p_{T(\omega)}N)_{V(\omega,\cdot) \wedge V_n(\omega,\cdot)}]$$

(because $\{S_n \wedge S > T\} \in \mathscr{F}_T^0$ and by 2.46). Finally, since $(p_t N)^{V_n(\omega,\cdot)}$ is a bounded $P_{x,t}$-martingale for all (x,t) and all $\omega \in \Omega$, we deduce that the above equals 0.

If now S is a stopping time, $S + 1/n$ is a strict stopping time (see 2.36) and since N^{S_n} is bounded and càdlàg we get $E_Q(\tilde{N}_S^{S_n}) = \lim_m E_Q(\tilde{N}_{S+1/m}^{S_n}) = 0$. Hence I.1.44 implies that $\tilde{N}^{S_n}$ is a Q-martingale (recall that $S_n \leq n$). Since $S_n \uparrow \infty$ as $n \uparrow \infty$, we obtain the result. □

Proof of Theorem 2.40. Let P be a solution of the stopped problem $\mathscr{s}(\mathscr{H}, X^T | \varepsilon_x; B^T, C^T, \nu^T)$, where T is a strict stopping time. Let Q be defined by 2.46, where $P_{x,t}$ is given in the assumptions of the theorem. We will prove that $Q \in \mathscr{s}(\mathscr{H}, X | \varepsilon_x; B, C, \nu)$, hence $Q = P_{x,0}$ by the uniqueness assumption. Now, if $A \in \mathscr{F}_T^0$, the set $\bar{A} = A \times \Omega \in \mathscr{F}_T^0 \otimes \mathscr{F}$ satisfies $A = \{\omega: T(\omega) < \infty, (\omega, \theta_T \omega) \in \bar{A}\}$ and 2.46 immediately yields $Q(A) = P(A)$. Therefore $P = P_{x,0}$ on $(\Omega, \mathscr{F}_T^0)$. In other words, we have local uniqueness.

It remains to prove that $Q \in \mathscr{s}(\mathscr{H}, X|\varepsilon_x; B, C, \nu)$. To see that, we first observe that $Q = \varepsilon_x$ on $(\Omega, \mathscr{H})$ by construction. Next, let N be any of the following processes, with the notation of 2.7:

$$N^i = [M(h)^i]^T, \quad N^{ij} = [M(h)^i M(h)^j - \tilde{C}^{ij}]^T, \quad N^f = (f * \mu^X - f * \nu)^T$$

and $p_t N$ is accordingly defined by

$$p_t N^i = X(h)^i - X_0^i - p_t B^i, \quad p_t N^{ij} = (p_t N^i)(p_t N^j) - p_t \tilde{C}^{ij}$$

$$p_t N^f = f * \mu^X - f * (p_t \nu)$$

($p_t \tilde{C}$ is defined by 2.6, from $p_t B, p_t C, p_t \nu$). Then those N and $p_t N$ meet all conditions (i)–(iii) of 2.48 (note that a predictable process, null at 0, is adapted to $(\mathscr{F}_t{}^0)$), by 2.39(i). Moreover, each N is a P-local martingale by our assumption on P being a solution to the stopped problem (see 2.7), and similarly each $p_t N$ is a $P_{x,t}$-local martingale.

To each family $(N, p_t N)$ we associate the process $\tilde{N}$ by 2.49, and $\tilde{N}$ is a Q-local martingale. Now, an easy computation based upon 2.39(ii) and the properties $(X(h)_s - X_0) \circ \theta_t = X(h)_{s+t} - X(h)_t$ and $(f * \mu^X)_s \circ \theta_t = f * \mu^X_{s+t} - f * \mu^X_t$, shows that

$$\tilde{N}^i = M(h)^i, \quad \tilde{N}^{ij} = M(h)^i M(h)^j - \tilde{C}^{ij}, \quad \tilde{N}^f = f * \mu^X - f * \nu.$$

Therefore 2.7 allows to conclude that $Q \in \mathscr{s}(\mathscr{H}, X|\varepsilon_x; B, C, \nu)$. □

3. Absolutely Continuous Changes of Measures

Here we consider two probability measures P and P' on a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$. The basic hypothesis is either that P' is absolutely continuous with respect to P (we write $P' \ll P$), or a slightly weaker assumption: a "local" absolute continuity of P' with respect to P.

Our main aim is to compute the characteristics of a semimartingale X relative to P', from its characteristics relative to P; these computations, and other related computations concerning martingales and random measures, are known as "*Girsanov's Theorems*". The main ingredient which shows up in them is the *density process* of P' relative to P: this is a martingale Z on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ such that for each $t \in \mathbb{R}_+$, Z_t is the Radon-Nikodym derivative $dP'_{|\mathscr{F}_t}/dP_{|\mathscr{F}_t}$ of the restrictions of P' and P to $(\Omega, \mathscr{F}_t)$.

§ 3a. The Density Process

At first we introduce some notation: E and E' denote the expectations with respect to the measures P and P'. For each stopping time T we set

$$P_T = \text{restriction of } P \text{ to } \mathscr{F}_T,$$

3.1

$$P_{T-} = \text{restriction of } P \text{ to } \mathscr{F}_{T-},$$

and similarly for P'_T and P'_{T-}.

3.2 **Definition.** We say that P' is *locally absolutely continuous* with respect to P, and we write $P' \overset{\text{loc}}{\ll} P$, if $P'_t \ll P_t$ for all $t \in \mathbb{R}_+$. □

Usually, a "local" property is localized along a sequence of stopping times (see § I.1d). The present notion indeed satisfies the same rule, as shown in the following:

3.3 **Lemma.** *$P' \overset{\text{loc}}{\ll} P$ if and only if there is an increasing sequence (T_n) of stopping times, such that $P'(\lim_n \uparrow T_n = \infty) = 1$ and that $P'_{T_n} \ll P_{T_n}$ for all $n \in \mathbb{N}^*$.*

Proof. The necessary part is obvious. To prove the sufficient part, let $A \in \mathscr{F}_t$ with $P(A) = 0$. Then

$$P'(A) = \lim_n P'(A \cap \{T_n > t\}) = 0$$

because $A \cap \{T_n > t\} \in \mathscr{F}_{T_n}$ and $P'_{T_n} \ll P_{T_n}$. Hence $P'_t \ll P_t$. □

3.4 **Theorem.** *Assume that $P' \overset{\text{loc}}{\ll} P$. There is a unique* (up to P- and P'-indistinguishability) *P-martingale Z, such that $Z_t = dP'_t/dP_t$* (the Radon-Nikodym derivative) *for all $t \in \mathbb{R}_+$. Moreover*

(i) *one may take $Z \geq 0$ identically.*

(ii) *If T is a stopping time, in restriction to the set $\{T < \infty\}$ we have $P'_T \ll P_T$ and $Z_T = dP'_T/dP_T$.*

(iii) *If T is a predictable time, in restriction to the set $\{T < \infty\}$ we have $P'_{T-} \ll P_{T-}$ and $Z_{T-} = dP'_{T-}/dP_{T-}$.*

The P-martingale Z is called the *density process* of P', relative to P. Observe that $E(Z_t) = 1$ for all $t \in \mathbb{R}_+$.

Proof. Set $U^n = dP'_n/dP_n$, so $U^n \in L^1(\Omega, \mathscr{F}, P)$. Call Y^n the P-martingale such that $Y^n_t = E(U^n|\mathscr{F}_t)$ if $t < n$ and $Y^n_t = U^n$ if $t \geq n$ (see I.1.42): one may clearly take a version such that $Y^n \geq 0$. Set $Z = \sum_{n\geq 1} Y^n 1_{[\![n-1,n[\![}$, which is càdlàg and adapted and $Z \geq 0$. Let T be a stopping time, and $A \in \mathscr{F}_T$. Then

$$\begin{aligned} E(1_A 1_{\{T<\infty\}} Z_T) &= \sum_{n\geq 1} E(1_A 1_{\{n-1\leq T<n\}} Y^n_T) = \sum_{n\geq 1} E(1_A 1_{\{n-1\leq T<n\}} U^n) \\ &= \sum_{n\geq 1} P'(A \cap \{n-1 \leq T < n\}) = P'(A \cap \{T < \infty\}). \end{aligned}$$

We deduce (ii). Taking $A \in \mathscr{F}_t$ and $T \equiv t$ (resp. $T \equiv s > t$) we obtain that $E(1_A Z_t) = P'(A) = E(1_A Z_s)$, hence Z is a P-martingale.

If T is a predictable time, and $A \in \mathscr{F}_{T-}$, then

$$P'(A \cap \{T < \infty\}) = E(1_A 1_{\{T<\infty\}} Z_T) = E(1_A 1_{\{T<\infty\}} Z_{T-})$$

by I.2.27, hence (iii). Finally, the Radon-Nikodym derivative dP'_t/dP_t is P-a. s. and P'-a. s. unique, hence the uniqueness of Z follows. □

Now come some easy properties of the density process.

3.5 **Proposition.** *Assume that $P' \overset{\text{loc}}{\ll} P$ and let Z be the density process.*

a) *We have $P'(\inf_{(t)} Z_t > 0) = 1$.*

b) *There is equivalence between:*

(i) $P'_{\infty-} \ll P_{\infty-}$;

(ii) $P'(\sup_{(t)} Z_t < \infty) = 1$;

(iii) *Z is a P-uniformly integrable martingale.*

Proof. a) Let $T_n = \inf(t: Z_t < 1/n)$. Then 3.4(ii) yields

$$P'(T_n < \infty) = E(Z_{T_n} 1_{\{T_n<\infty\}}) \leq \frac{1}{n},$$

hence $P'(\bigcap_n \{T_n < \infty\}) = 0$ and the claim follows.

b) (i) $\Rightarrow$ (ii): By Doob's limit theorem I.1.39, Z_t converges P-a. s. to a finite limit as $t \uparrow \infty$. Under (i), the same holds P'-a. s., hence (ii).

(ii) $\Rightarrow$ (iii): $E(Z_s 1_{\{Z_s>n\}}) = P'(Z_s > n) \leq P'(\sup_{(t)} Z_t > n)$ goes to 0 as $n \uparrow \infty$, hence the family $(Z_s)_{s\in\mathbb{R}_+}$ is P-uniformly integrable.

(iii) $\Rightarrow$ (i): Under (iii), $Z_t \to Z_\infty$ in $L^1(\Omega, \mathscr{F}, P)$, while for each $A \in \mathscr{F}_t$, $P'(A) = E(1_A Z_t) = E(1_A Z_\infty)$. A monotone class argument yields that $P'(A) = E(1_A Z_\infty)$ for all $A \in \mathscr{F}_{\infty-}$, and we deduce (i). □

3.6 **Lemma.** *Let Z be a nonnegative P-supermartingale* (e.g. the density process when $P' \overset{\text{loc}}{\ll} P$). *Then $T = \inf(t: Z_t = 0$ or $Z_{t-} = 0)$ is a stopping time, and $Z = 0$ P-a. s. on $[\![T, \infty[\![$.*

Proof. Let $T_n = \inf(t: Z_t < 1/n)$, which is a stopping time. We have $T = \lim_n \uparrow T_n$. By Doob's stopping theorem I.1.39, $E(Z_T|\mathscr{F}_{T_n}) \leq Z_{T_n} \leq \frac{1}{n}$ on $\{T_n < \infty\}$, hence $E(Z_T 1_{\{T<\infty\}}) \leq 1/n$ for all n, hence $Z_T = 0$ P-a.s. on $\{T < \infty\}$. Now, if $S_n = \inf(t > T: Z_t > 1/n)$, then S_n is a stopping time and $E(Z_{S_n}|\mathscr{F}_T) \leq Z_T = 0$ on $\{T < \infty\}$, which implies that $S_n = \infty$ P-a.s., and so $Z = 0$ P-a. s. $[\![T, \infty[\![$. □

3.7 **Lemma.** *Assume that $P' \overset{\text{loc}}{\ll} P$ and that Z is the density process. Let X and Y be two predictable processes. Then X is P'-indistinguishable from Y if and only if $X 1_{\{Z_->0\}}$ is P-indistinguishable from $Y 1_{\{Z_->0\}}$.*

In particular, if X and Y are predictable with finite variation over finite intervals, then they are P'-indistinguishable if and only if $1_{\{Z_->0\}} \cdot X$ and

$1_{\{Z_->0\}} \cdot Y$ are P-indistinguishable: this is due to the structure of the set $\{Z_- > 0\}$ as described in 3.6.

Proof. This is a simple corollary of the version I.2.18 of the predictable section theorem, once observed that if S is a predictable time, then $P'(X_S \neq Y_S, S < \infty) = E(1_{\{X_S \neq Y_S, S<\infty\}} Z_{S-})$ by 3.3(iii). □

§ 3b. Girsanov's Theorem for Local Martingales

1. We begin with a general result.

3.8 **Proposition.** *Assume that* $P' \overset{\text{loc}}{\ll} P$ *and let* Z *be the density process. Let* M' *be a càdlàg adapted process.*

a) *$M'Z$ is a P-martingale if and only if M' is a P'-martingale.*

b) *If $M'Z$ is a P-local martingale, then M' is a P'-local martingale.*

c) *If M' is a P'-local martingale with a localizing sequence (T_n) having $P(\lim_n \uparrow T_n = \infty) = 1$, then $M'Z$ is a P-local martingale.*

Observe that (a) is equivalent to the following well-known formula: if Y is bounded (or P'-integrable) and $\mathscr{F}_s$-measurable, and $t \leq s$,

3.9
$$E'(Y|\mathscr{F}_t) = \frac{1}{Z_t} E(YZ_s|\mathscr{F}_t).$$

Proof. a) Let $A \in \mathscr{F}_t$. Then $E'(1_A M'_t) = E(1_A Z_t M'_t)$. Therefore $E'(M'_t - M'_s|\mathscr{F}_s) = 0$ (for $s \leq t$) if and only if $E(Z_t M'_t - Z_s M'_s|\mathscr{F}_s) = 0$, and the equivalence follows.

b) Let (T_n) be a localizing sequence for the P-local martingale $M'Z$ and $T = \lim_n \uparrow T_n$. Then 3.4 yields $P'(T < \infty) = E(Z_T 1_{\{T<\infty\}}) = 0$ (because $T = \infty$ P-a.s.). Moreover $M'^{T_n}Z = (M'Z)^{T_n} + M'_{T_n}(Z - Z_{T_n})1_{[\![T_n,\infty[\![}$ is clearly a P-martingale, so the claim follows from (a).

c) $M'^{T_n}Z$ is a P-martingale by (a), so $(M'Z)^{T_n} = M'^{T_n}Z - M'_{T_n}(Z - Z_{T_n})1_{[\![T_n,\infty[\![}$ is also a P-martingale. □

3.10 **Corollary.** *Under the assumptions of 3.8, if $T_n = \inf(t\colon Z_t < 1/n)$ and if all $(M'Z)^{T_n}$ are P-local martingales, then M' is a P'-local martingale.*

Proof. $M'^{T_n}Z = (M'Z)^{T_n} + M'_{T_n}(Z - Z_{T_n})1_{[\![T_n,\infty[\![}$ is clearly a P-local martingale, so M'^{T_n} is a P'-local martingale by 3.8b, and $T_n \uparrow \infty$ P'-a. s. by 3.5. □

2. Now we arrive to the "classical" Girsanov's Theorem.

3.11 **Theorem.** *Assume that* $P' \overset{\text{loc}}{\ll} P$ *and let Z be the density process. Let M be a P-local martingale such that $M_0 = 0$ and that the P-quadratic covariation $[M, Z]$*

has P-locally integrable variation, and denote by $\langle M, Z\rangle$ its P-compensator (that is coherent with I.4.50). *Then the process*

3.12 $$M' = M - \frac{1}{Z_-} \cdot \langle M, Z\rangle$$

is P'-a.s. well defined, and is a P'-local martingale. Moreover the P-quadratic variation $\langle M^c, M^c\rangle$ of the continuous part M^c (relative to P) *of M is also a version of the P'-quadratic variation of the continuous part* (relative to P') *of M'.*

Proof. a) Let $T_n = \inf(t\colon Z_t < 1/n)$ and $A = (1/Z_-) \cdot \langle M, Z\rangle$; which is obviously well defined on each $[\![0, T_n]\!]$. Since $T_n \uparrow \infty$ P'-a.s. by 3.5, A and M' are P'-a.s. well-defined.

b) Since $MZ = M_- \cdot Z + Z_- \cdot M + [Z, M]$, $(MZ)^{T_n} - \langle M, Z\rangle^{T_n}$ is a P-local martingale. Since A^{T_n} is predictable with finite variation, I.4.49b yields $(AZ)^{T_n} = A \cdot Z^{T_n} + Z_- \cdot A^{T_n}$, so $(AZ)^{T_n} - Z_- \cdot A^{T_n}$ is also a P-local martingale. Moreover, $Z_- \cdot A^{T_n} = \langle M, Z\rangle^{T_n}$ by definition of A. Hence by difference we obtain that $(M'Z)^{T_n}$ is a P-local martingale, thus M' is a P'-local martingale by 3.10.

c) From (b) we deduce that M is a P'-semimartingale (with canonical decomposition $M = M' + A$). Now, the "Riemann sums" $S_{\tau_n}(M, M)$ defined by I.4.48 do not depend upon the probability measure, hence the quadratic variation $[M, M]$ is the same for P and P', by I.4.47a. The last claim follows, because by I.4.53 the quadratic variation of the continuous martingale part of a semimartingale is just the "continuous part" of the quadratic variation of the semimartingale. □

As a corollary, we obtain the

3.13 **Theorem.** *Assume that $P' \overset{\text{loc}}{\ll} P$. Any P-semimartingale X is also a semimartingale relative to P', and the quadratic variation process $[X, X]$ for P is also a version of the quadratic variation for P'.*

For proving this theorem, we need a preliminary lemma.

3.14 **Lemma.** *Let Z and M be two local martingales, such that $|\Delta M| \le a$ for some constant a. Then $[M, Z]$ has locally integrable variation.*

Proof. Due to I.4.55 we may assume by localization that $U = [\sum_s (\Delta Z_s)^2]^{1/2}$ is integrable. Call A the variation process of $[M, Z]$ (which is known to belong to $\mathscr{V}$) and $T_n = \inf(t\colon A_t > n)$. Then $A_{T_n} \le n + \Delta A_{T_n} = n + |\Delta M_{T_n} \Delta Z_{T_n}| \le n + a|\Delta Z_{T_n}| \le n + aU$ on $\{T_n < \infty\}$: thus $E(A_{T_n}) < \infty$. □

Proof of Theorem 3.13. There is a decomposition $X = X_0 + M + A$ where $A \in \mathscr{V}$ and M is a P-local martingale with bounded jumps. A clearly is a P'-semimartingale, and we deduce from 3.11 and 3.14 that M, and thus X also, are P'-semimartingales. Finally the claim concerning the quadratic variation is proved as in (c) of the proof of 3.11. $\square$

§ 3c. Girsanov's Theorem for Random Measures

We consider an auxiliary Blackwell space $(E, \mathscr{E})$ (see § II.1a). To every random measure μ on $\mathbb{R}_+ \times E$ defined on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ we associate the following:

3.15 M_μ^P is the positive measure on $(\tilde{\Omega}, \mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{E})$ defined by $M_\mu^P(W) = E(W * \mu_\infty)$ for all measurable nonnegative functions W. $\square$

Assume furthermore that μ is $\tilde{\mathscr{P}}$-σ-finite on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, which is equivalent to saying that the restriction of the measure M_μ^P to $(\tilde{\Omega}, \tilde{\mathscr{P}})$ is σ-finite. Then, there is a notion of "conditional expectation relative to M_μ^P", with respect to the σ-field $\tilde{\mathscr{P}}$: for every nonnegative measurable function W the "conditional expectation" $W' = M_\mu^P(W|\tilde{\mathscr{P}})$ is the M_μ^P-a.e. unique $\tilde{\mathscr{P}}$-measurable function such that

3.16 $$M_\mu^P(WU) = M_\mu^P(W'U) \text{ for all nonnegative } \tilde{\mathscr{P}}\text{-measurable } U.$$

Furthermore, exactly as in I.1.1, we can next defined a "generalized" conditional M_μ^P-expectation for *all measurable functions W*.

3.17 **Theorem.** *Assume that* $P' \overset{\text{loc}}{\ll} P$ *and let Z be the density process. Let* μ *be an integer-valued random measure on* $\mathbb{R}_+ \times E$ *defined on the stochastic basis* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (this implies in particular that it is $\tilde{\mathscr{P}}$-σ-finite relative to P), *and denote by* ν *its P-compensator.*

a) μ *is also* $\tilde{\mathscr{P}}$-σ-*finite relative to* P'.

b) *Let Y be a* $\tilde{\mathscr{P}}$-*measurable nonnegative function on* $\tilde{\Omega}$. *Let* ν' *be a version of the* P'-*compensator of* μ. *There is equivalence between*:

(i) $\nu' = Y \cdot \nu$ P'-a. s. (where $Y \cdot \nu(\omega; dt, dx) = \nu(\omega; dt, dx) Y(\omega, t, x)$);

(ii) $1_{\{Z_- > 0\}} \cdot \nu' = Y 1_{\{Z_- > 0\}} \cdot \nu$ P-a. s.;

(iii) YZ_- *is a version of the conditional expectation* $M_\mu^P(Z|\tilde{\mathscr{P}})$.

Moreover, any nonnegative version Y of $M_\mu^P\left(\frac{Z}{Z_-} 1_{\{Z_- > 0\}} \middle| \tilde{\mathscr{P}}\right)$ *has the above properties.*

c) *There is a version of* ν' *that meets identically*

3.18 $$\begin{cases} \nu' = Y \cdot \nu & \text{for some } \tilde{\mathscr{P}}\text{-measurable nonnegative function } Y, \\ \nu(\omega; \{t\} \times E) = 1 \Rightarrow \nu'(\omega; \{t\} \times E) = 1. \end{cases}$$

3.19 **Remark.** Except for the last claim in 3.18, all the above remains true for any $\tilde{\mathscr{P}}$-σ-finite random measure μ on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, whether it is integer-valued or not (only the integer-valued case is of interest for us). $\square$

Proof. Let us first show some auxiliary facts. By hypothesis there is a strictly positive $\tilde{\mathscr{P}}$-measurable function V with $V * \mu_\infty \in L^1(\Omega, \mathscr{F}, P)$, and so $V * \nu_\infty \in L^1(\Omega, \mathscr{F}, P)$ as well, and we can always assume that $V \leq 1$ (if not, replace V by $V \wedge 1$). We define the stopping times:

$$T_n = \inf(t: t \geq n, \text{ or } V * \mu_t \geq n, \text{ or } V * \nu_t \geq n, \text{ or } Z_t > n)$$

and $A = \bigcup_n [\![0, T_n]\!]$ and $T = \lim_n \uparrow T_n$. Then $P(T < \infty) = 0$. Moreover,

3.20 $$V * \mu_{T_n} \leq n + 1, \quad V * \nu_{T_n} \leq n + 1.$$

Furthermore, the stopped P-martingale Z^n is uniformly integrable, so $Z_{T_n} = (Z^n)_{T_n}$ (recall $T_n \leq n$) is P-integrable, and so

3.21 $$\sup_{s \leq T_n} Z_s \leq n + Z_{T_n} \in L^1(\Omega, \mathscr{F}, P).$$

Now we can proceed to the proof itself.

a) $P'(T < \infty) = E(Z_T 1_{\{T<\infty\}}) = 0$ because $P(T < \infty) = 0$. Therefore the $\tilde{\mathscr{P}}$-measurable and strictly positive function $V' = 1_{A^c \times E} + \sum_{n \geq 1} 2^{-n} V 1_{[\![0, T_n]\!] \times E}$ has $V' * \mu_\infty = \sum_{n\geq 1} 2^{-n} V * \mu_{T_n} \leq \sum_{n \geq 1} (n+1) 2^{-n}$ P'-a. s., and the claim follows.

b) Firstly, let Y be a nonnegative version of $M^P_\mu\left(\frac{Z}{Z_-} 1_{\{Z_->0\}} \middle| \tilde{\mathscr{P}}\right)$. Then $YZ_- = M^P_\mu(Z 1_{\{Z_->0\}} | \tilde{\mathscr{P}}) = M^P_\mu(Z | \tilde{\mathscr{P}})$ by 3.6, so Y meets (iii). Secondly, we prove the equivalence between (i), (ii), (iii):

(iii) $\Rightarrow$ (i): Assume that Y meets (iii). The random measure $Y \cdot \nu$ is clearly predictable and $W * (Y \cdot \nu) = (WY) * \nu$. Hence if we prove that

3.22 $$E'(W * \mu_\infty) = E'((WY) * \nu_\infty)$$

for all nonnegative $\tilde{\mathscr{P}}$-measurable W, (i) will follow from II.1.8. In fact, owing to the monotone limit theorem, it suffices to prove 3.22 in the following three cases:

(1) $0 \leq W \leq V 1_{\{Y \leq q\}} 1_{[\![0, T_n]\!] \times E}$ for some $q, n \in \mathbb{N}^*$
(2) $0 \leq W \leq V 1_{\{Y = \infty\}} 1_{[\![0, T_n]\!] \times E}$ for some $n \in \mathbb{N}^*$
(3) $0 \leq W \leq 1_{A^c \times E}$.

In case (3), 3.22 is trivial because A^c is P'-evanescent by (a). If W is $\tilde{\mathscr{P}}$-measurable and nonnegative, assumption (iii) for Y yields

3.23 $$\begin{aligned} E[(ZW) * \mu_{T_n}] &= M^P_\mu(ZW 1_{[\![0, T_n]\!] \times E}) = M^P_\mu(Z_- Y W 1_{[\![0, T_n]\!] \times E}) \\ &= E[(Z_- WY) * \mu_{T_n}] = E[(Z_- WY) * \nu_{T_n}] \end{aligned}$$

(the last equality comes from the definition of ν).

Now we apply 3.23 to $W = V 1_{\{Y=\infty\}}$. By 3.20 and 3.21, the left-hand side of 3.23 is smaller than $(n+1) E(\sup_{s \leq T_n} Z_s)$, which is finite. On the other hand, $Z_- V Y 1_{\{Y=\infty\}} * \mu_{T_n}$ is either 0 or ∞, and similarly for ν instead of μ, so these random variables should be P-a. s. 0; since they are $\mathscr{F}_n$-measurable (recall $T_n \leq n$) and $P'_n \ll P_n$, they also are P'-a. s. 0, or in other words μ and ν P'-a. s. do not charge the set $\{Y = \infty\} \cap [\![0, T_n]\!] \cap \{Z_- > 0\}$. Since $P'(\inf_t Z_t > 0) = 1$, we deduce that μ and ν P'-a. s. do not charge $\{Y = \infty\} \cap [\![0, T_n]\!]$. Thus if W is as in (2), both sides of 3.22 equal 0.

Finally let W be as in (1). Let $G = (W * \mu)^{T_n}$ and $B = [(WY) * \nu]^{T_n}$, which are increasing, bounded by 3.20, and predictable for B. Then 3.21 yields that the processes $GZ^{T_n} - Z \cdot G$ and $BZ^{T_n} - Z_- \cdot B$ are of class (D) for P, while they are P-local martingales by I.4.49, so they are indeed P-uniformly integrable martingales. Therefore the stopping theorem yields

$$E'(W * \mu_\infty) = E'(G_{T_n}) = E(Z_{T_n} G_{T_n}) = E(Z \cdot G_{T_n}) = E[(ZW) * \mu_{T_n}]$$

$$E'[(WY) * \nu_\infty] = E'(B_{T_n}) = E(Z_{T_n} B_{T_n}) = E(Z_- \cdot B_{T_n}) = E[(Z_- WY) * \nu_{T_n}].$$

Those are equal by 3.23, so 3.22 holds.

(i) $\Rightarrow$ (ii): That readily follows from 3.7.

(ii) $\Rightarrow$ (iii): Let Y satisfy (ii) and $Y' = M^P_\mu\left(\frac{Z}{Z_-} 1_{\{Z_- > 0\}} | \tilde{\mathscr{P}}\right)$, which is known to satisfy (iii); from what precedes, Y' also satisfies (i) and (ii). So $Y 1_{\{Z_- > 0\}} \cdot \nu = Y' 1_{\{Z_- > 0\}} \cdot \nu$ P-a. s., thus the $\tilde{\mathscr{P}}$-measurable set $A = \{Y \neq Y', Z_- > 0\}$ has $1_A * \nu_\infty = 0$ P-a. s. Therefore $M^P_\mu(A) = E(1_A * \mu_\infty) = E(1_A * \nu_\infty) = 0$, hence $YZ_- = Y'Z_-$ M^P_μ-a. s., and since Y' meets (iii), so does Y.

c) Let $\nu' = Y \cdot \nu$ be a version of the P'-compensator of μ, with Y as above. Set $a_t(\omega) = \nu(\omega; \{t\} \times E)$ and $a'_t(\omega) = \nu'(\omega; \{t\} \times E)$. There is a sequence (S_n) of predictable times such that $\{a = 1\} = \bigcup [\![S_n]\!]$ up to a P-evanescent set. Then if $D = \{(\omega, t): \mu(\omega; \{t\} \times E) = 1\}$, II.1.18 applied to $W = 1$ yields $1 = a_{S_n} = P(S_n \in D | \mathscr{F}_{S_n-})$ on $\{S_n < \infty\}$, so $P(S_n < \infty, S_n \notin D) = 0$. Since $P' \overset{\text{loc}}{\ll} P$ we deduce that $P'(S_n < \infty, S_n \notin D) = 0$ as well, so II.1.18 again yields $a'_{S_n} = P'(S_n \in D | \mathscr{F}_{S_n-}) = 1$ P'-a. s. on $\{S_n < \infty\}$. Therefore if

$$\hat{\nu}' = Y' \cdot \nu \quad \text{with } Y'(\omega, t, z) = \begin{cases} Y(\omega, t, z) & \text{if } a_t(\omega) \neq 1 \text{ or } a_t(\omega) = a'_t(\omega) = 1 \\ 1 & \text{if } a'_t(\omega) \neq 1 = a_t(\omega) \end{cases}$$

we have $\hat{\nu}' = \nu'$ P'-a. s., and 3.18 is met by $\hat{\nu}'$. □

§ 3d. Girsanov's Theorem for Semimartingales

1. Here we consider a d-dimensional *semimartingale* $X = (X^i)_{i \leq d}$ on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with characteristics (B, C, ν) relative to a given truncation function h.

We denote by X^c the continuous martingale part of X, relative to P. Let A be an increasing predictable process such that $C^{ij} = c^{ij} \cdot A$ (see II.2.10).

3.24 **Theorem.** *Assume that $P' \overset{\text{loc}}{\ll} P$, and let X be as above. There exist a $\tilde{\mathscr{P}}$-measurable nonnegative function Y and a predictable process $\beta = (\beta^i)_{i \leq d}$ satisfying*

3.25 $$|h(x)(Y-1)| * \nu_t < \infty \ P'\text{-a.s. for } t \in \mathbb{R}_+$$

3.26 $$\left|\sum_{j\le d} c^{ij}\beta^j\right| \cdot A_t < \infty \text{ and } \left(\sum_{j,k\le d} \beta^j c^{jk}\beta^k\right) \cdot A_t < \infty \ P'\text{-a.s. for } t \in \mathbb{R}_+$$

and such that a version of the characteristics of X relative to P' are

3.27 $$\begin{cases} B'^i = B^i + \left(\sum_{j\le d} c^{ij}\beta^j\right) \cdot A + h^i(x)(Y-1) * \nu \\ C' = C \\ \nu' = Y \cdot \nu \end{cases}$$

Moreover, Y and β meet all the above conditions, if and only if

3.28 $$\begin{cases} YZ_- = M^P_{\mu^X}(Z|\tilde{\mathscr{P}}) \\ \langle Z^c, X^{i,c}\rangle = \left(\sum_{j\le d} c^{ij}\beta^j Z_-\right) \cdot A, \end{cases}$$

(up to a P-null set, of course), *where Z is the density process, Z^c is its continuous martingale part relative to P, and $\langle Z^c, X^{i,c}\rangle$ is the bracket relative to P* (which also equals $[Z, X^{i,c}]$). *Moreover it is even possible to choose Y so that*

3.29 $$\nu(\omega;\{t\}\times\mathbb{R}^d) = 1 \Rightarrow \nu'(\omega;\{t\}\times\mathbb{R}^d) = \int Y(\omega,t,x)\nu(\omega;\{t\}\times dx) = 1$$

Observe that 3.25 and 3.26 imply that the processes in 3.27 are P'-a.s. finite-valued.

We begin with a lemma, of independent interest, and for which the setting is as follows: let $Y = (Y^i)_{i\le d}$ be a continuous local martingale on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. Let A be an increasing process, and $c = (c^{ij})_{i,j\le d}$ a predictable process with values in the set of all $d\times d$ symmetric nonnegative matrices, such that

3.30 $$\langle Y^i, Y^j\rangle = c^{ij} \cdot A.$$

3.31 **Lemma.** *In addition to the above, let U be another continuous local martingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$. Then there is a predictable process $H = (H^i)_{i\le d}$ such that for $i = 1, \dots, d$,*

3.32 $$\langle U, Y^i\rangle = \left(\sum_{j\le d} c^{ij}H^j\right) \cdot A.$$

Moreover, for any such H, the increasing process $(\sum_{i,j\le d} H^i c^{ij} H^j) \cdot A$ is locally integrable on $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

Proof. We consider the $(d+1)$-dimensional continuous local martingale $\bar{Y} = (Y, U)$. Since the claims do not depend upon the choice of the pair (A, c), as long as 3.30 holds, we can choose A so that $\langle \bar{Y}^i, \bar{Y}^j \rangle = \bar{c}^{ij} \cdot A$ for all $i, j \leq d + 1$, with $\bar{c} = (\bar{c}^{ij})_{i,j \leq d+1}$ a $(d + 1) \times (d + 1)$ symmetric nonnegative matrix-valued predictable process. Moreover, we can write $\bar{c} = \begin{pmatrix} c & \alpha \\ \alpha^T & \gamma \end{pmatrix}$ where γ is a predictable $\mathbb{R}$-valued process, and α is a predictable $\mathbb{R}^d$-valued process, and α^T denotes the transpose. Now, it suffices to prove that $\alpha = cH$ for a predictable $\mathbb{R}^d$-valued process (hence 3.32), and that $H^T cH \leq \gamma$ (hence the last claim, because $\gamma \cdot A = \langle U, U \rangle$).

But those facts are elementary properties of symmetric matrices, which can be proved as follows. There are two predictable processes Π and Λ with values respectively in the sets of orthogonal $d \times d$ matrices and diagonal $d \times d$ matrices, such that $c = \Pi^{-1}\Lambda\Pi$. If $\tilde{\Lambda}$ is the diagonal matrix whose entries are $\tilde{\Lambda}^{ii} = (\Lambda^{ii})^{-1}$ (resp. $= 0$) if $\Lambda^{ii} > 0$ (resp. $= 0$), then we set $H = \Pi^{-1}\tilde{\Lambda}\Pi\alpha$, so H is predictable. Moreover, let $\bar{\Pi} = \begin{pmatrix} \Pi & 0 \\ 0 & 1 \end{pmatrix}$ (a $(d + 1) \times (d + 1)$ matrix), so $\bar{\Pi}\bar{c}\bar{\Pi}^{-1} = \begin{pmatrix} \Lambda & \Pi\alpha \\ (\Pi\alpha)^T & \gamma \end{pmatrix}$ is again a symmetric nonnegative matrix, hence $(\Pi\alpha)^i = 0$ whenever $\Lambda^{ii} = 0$. Therefore $cH = c\Pi^{-1}\tilde{\Lambda}\Pi\alpha = \Pi^{-1}\Lambda\tilde{\Lambda}\Pi\alpha$ equals $\Pi^{-1}\Pi\alpha = \alpha$ and we get the first claim.

Finally, let H be such that $cH = \alpha$. Then $\bar{c} = \begin{pmatrix} c & \alpha \\ \alpha^T & \gamma \end{pmatrix}$ being symmetric nonnegative, for all $u \in \mathbb{R}$ we have

$$H^T cH + 2uH^T cH + u^2\gamma = (H^T, u)\bar{c}\begin{pmatrix} H \\ u \end{pmatrix} \geq 0.$$

Therefore $H^T cH \leq \gamma$ and we are done. □

Proof of Theorem 3.24. In the proof, all stochastic integrals, brackets, etc, ... are relative to the measure P.

a) We know by 3.13 that X is a P'-semimartingale, and that the quadratic covariation process $[X^i, X^j]$ is the same for P and P'. Since C^{ij} is the "continuous part" of $[X^i, X^j]$ we deduce that $C' = C$ is also a version of the second characteristic of X for P'.

b) Set $Y = M^P_\mu\left(\frac{Z}{Z_-} 1_{\{Z_- > 0\}} \middle| \tilde{\mathscr{P}}\right)$, where $\mu = \mu^X$ is the measure associated with the jumps of X by II.1.16. Then 3.17 yields that $\nu' = Y \cdot \nu$ is a version of the third characteristic of X under P', and that Y meets the first equality in 3.28.

c) Set $M = h * (\mu - \nu)$, which is a purely discontinuous local martingale. Then the $\mathbb{R}^d$-valued process $[M, Z]$ is $[M, Z] = \sum_{s \leq \cdot} \Delta M_s \Delta Z_s$. Let (T_n) be a sequence of predictable times that exhausts the predictable thin set $J = \{(\omega, t) \colon \nu(\omega; \{t\} \times \mathbb{R}^d) > 0\}$. By II.2.14, ΔB is P-indistinguishable from $\nu(\{t\} \times h)$, hence by Definition II.1.27 of the stochastic integral $h * (\mu - \nu)$,

$$[M, Z] = (h(x)\Delta Z 1_{J^c}) * \mu + \sum_n \Delta Z_{T_n}(h(\Delta X_{T_n}) - \Delta B_{T_n}) 1_{[\![T_n,\infty[\![}$$

3.33 $$= (h(x)\Delta Z) * \mu - [Z, B].$$

$[Z, B]$ is a local martingale with finite variation (see I.4.49), so it has locally integrable variation (for P). Since M has bounded jumps by construction, $[M, Z]$ has locally integrable variation by 3.14, hence we deduce from 3.33 that there is a localizing sequence (S_n) such that

3.34 $$E(|h(x)\Delta Z| * \mu_{S_n}) < \infty.$$

$Z_-(Y-1) = M_\mu^P(\Delta Z|\tilde{\mathscr{P}})$ by definition of Y (recall that $\Delta Z = 0$ if $Z_- = 0$ by 3.6). Then $Z_-|Y-1| \leq M_\mu^P(|\Delta Z||\tilde{\mathscr{P}})$ and

$$\begin{aligned} E(|h(x)Z_-(Y-1)| * \nu_{S_n}) &= E(|h(x)Z_-(Y-1)| * \mu_{S_n}) \\ &= M_\mu^P(|h(x)Z_-(Y-1)1_{[\![0,S_n]\!]}|) \leq M_\mu^P(|h(x)\Delta Z 1_{[\![0,S_n]\!]}|) \\ &= E(|h(x)\Delta Z| * \mu_{S_n}) \end{aligned}$$

is finite by 3.34. Then $Z_-|h(x)(Y-1)| * \nu$ belongs to $\mathscr{A}_{\text{loc}}(P)$; since $1/Z_-$ is P'-locally bounded by 3.5, property 3.25 follows.

Let S be a stopping time with $S \leq S_n$ for some $n \in \mathbb{N}$. The same argument as above leads to

$$\begin{aligned} E[(h(x)Z_-(Y-1)) * \nu_S] &= E[(h(x)Z_-(Y-1)) * \mu_S] = M_\mu^P(h(x)Z_-(Y-1)1_{[\![0,S]\!]}) \\ &= M_\mu^P(h(x)\Delta Z 1_{[\![0,S]\!]}) = E[(h(x)\Delta Z) * \mu_S]. \end{aligned}$$

Therefore $h(x)Z_-(Y-1) * \nu$ is the P-compensator of $h(x)\Delta Z * \mu$. Using again that $[Z, B]$ is a P-local martingale, we deduce from 3.33 that the P-compensator of $[M, Z]$ is $\langle M, Z\rangle = Z_- \cdot (h(x)(Y-1) * \nu)$. Thus 3.11 yields

3.35 $$M - h(x)(Y-1) * \nu \quad \text{is a } P'\text{-local martingale.}$$

d) Lemma 3.31 gives a predictable process $H = (H^i)_{i \leq d}$ such that $\langle Z^c, X^{i,c}\rangle = (\sum_{j \leq d} c^{ij}H^j) \cdot A$ and $(\sum_{i,j \leq d} H^i c^{ij} H^j) \cdot A \in \mathscr{A}_{\text{loc}}(P)$. Then if

$$\beta^i = \frac{H^i}{Z_-} 1_{\{Z_- > 0\}}$$

the process $\beta = (\beta^i)_{i \leq d}$ meets 3.26 because $1/Z_-$ is P'-locally bounded. Moreover $(\sum_{j \leq d} c^{ij}\beta^j) \cdot A = (1/Z_-) \cdot \langle Z^c, X^{i,c}\rangle$ by construction, so 3.11 yields

3.36 $$X^{i,c} - \left(\sum_{j \leq d} c^{ij}\beta^j\right) \cdot A \quad \text{is a } P'\text{-local martingale.}$$

Now, the process $X(h)$ defined by II.2.4 equals $X_0 + X^c + M + B$. It immediately follows from 3.25 and 3.26 that if B' is defined by 3.27, then $X(h) - X_0 - B'$ is a P'-local martingale: that is, B' is a version of the first characteristic of X under P'.

e) Y and β meet 3.28 by construction. Let Y' and β' satisfy 3.25 and 3.26 and 3.27. Firstly, 3.17 yields that $1_{\{Y' \neq Y, Z_- > 0\}} * \nu_\infty = 0$ P-a. s., hence $1_{\{Y' \neq Y, Z_- > 0\}} * \mu_\infty = 0$ P-a. s. and $YZ_- = Y'Z_-$ M^P_μ-a. s.: thus Y' satisfies 3.27.

Secondly, since the first characteristic is P'-a. s. unique and $Y' \cdot \nu = Y \cdot \nu$ P'-a. s., we deduce that $(\sum_{j \le d} c^{ij}\beta^j) \cdot A = (\sum_{j \le d} c^{ij}\beta'^j) \cdot A$ P'-a. s., and 3.7 yields $(\sum_{j \le d} c^{ij}\beta^j Z_-) \cdot A = (\sum_{j \le d} c^{ij}\beta'^j Z_-) \cdot A$ P-a. s.: so β' satisfies 3.27 as well.

Finally, the last assertion follows from 3.17c. □

2. Now we suppose that P is a convex combination of two probability measures P' and P'' on $(\Omega, \mathscr{F})$, namely

3.37 $$P = \alpha' P' + \alpha'' P'', \quad \text{with } \alpha', \alpha'' > 0 \quad \text{and } \alpha' + \alpha'' = 1.$$

3.38 **Lemma.** *We have $P' \ll P$, $P'' \ll P$, and there is a version Z' (resp. Z'') of the density process of P' (resp. P'') with respect to P, such that we have identically:*

3.39 $$\alpha' Z' + \alpha'' Z'' = 1, \quad 0 \le Z' \le 1/\alpha', \quad 0 \le Z'' \le 1/\alpha''.$$

Proof. Let Z' and Z'' be the density processes of P' and P'', which are obviously absolutely continuous with respect to P. If $A \in \mathscr{F}_t$, then

$$P(A) = \alpha' P'(A) + \alpha'' P''(A) = E[1_A(\alpha' Z'_t + \alpha'' Z''_t)].$$

Hence $\alpha' Z'_t + \alpha'' Z''_t = 1$ P-a. s. It follows that we can choose versions of Z' and Z'' such that $\alpha' Z' + \alpha'' Z'' = 1$ and $Z' \ge 0$ and $Z'' \ge 0$ identically, hence 3.39 holds. □

3.40 **Theorem.** *Assume 3.37, and let Z', Z'' be as in 3.38. Let X be a d-dimensional semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P')$ and on $(\Omega, \mathscr{F}, \mathbf{F}, P'')$, with respective characteristics (B', C', ν') and (B'', C'', ν''), relative to the same truncation function h.*

Then X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with the following characteristics relative to h:

3.41 $$\begin{cases} B = \alpha' Z'_- \cdot B' + \alpha'' Z''_- \cdot B'' \\ C = \alpha' Z'_- \cdot C' + \alpha'' Z''_- \cdot C'' \\ \nu = \alpha' Z'_- \cdot \nu' + \alpha'' Z''_- \cdot \nu''. \end{cases}$$

In particular, this shows that if X is a càdlàg adapted process, the set of all probability measures for which it is a semimartingale is a *convex set.*

3.42 **Remark.** In our definition II.2.6 of characteristics (say, under P') the process B' has finite variation over finite intervals. However, any process $\hat{B}'$ that is P'-indistinguishable from B' might serve as the first characteristic, although it may have infinite variation on some finite interval (on a P'-null set, of course, but not necessarily on a P-null set). Nevertheless, by 3.7, $\hat{B}'$ is P-indistinguishable from B' on the random set $\{Z'_- > 0\}$, so $Z_- \cdot \hat{B}'$ is P-a. s. well-defined, and also P-indistinguishable from $Z_- \cdot B'$. The same holds for C', of course.

It follows from this argument that 3.41 makes sense P-a. s. for any version of the characteristics (B', C', ν') and (B'', C'', ν''), even "bad" ones. □

Proof. According to the above argument, we choose for (B', C', ν') and (B'', C'', ν'') "good" versions of the characteristics, as constructed in II.2.9 for example.

We prove first that X is a P-semimartingale. With the notation II.2.4, $\check{X}(h)$ has finite variation, so it certainly is a P-semimartingale. Next, we have $X(h) - X_0 = M' + B' = M'' + B''$, where M' and M'' are càdlàg processes with bounded jumps and are respectively a P'- and P''-local martingale, localized by the same sequence of stopping time $S_n = \inf(t\colon |M_t| > n \text{ or } |M'_t| > n)$, which satisfies $\lim_n S_n(\omega) = \infty$ for all ω. We have

$$\begin{aligned} X(h) - X_0 &= \alpha' Z'(X(h) - X_0) + \alpha'' Z''(X(h) - X_0) \\ &= \alpha' Z' B' + \alpha'' Z'' B'' + \alpha' Z' M' + \alpha'' Z'' M''. \end{aligned}$$

We deduce from 3.8c that $N = \alpha' Z' M' + \alpha'' Z'' M''$ is a P-local martingale. By Ito's formula, $\alpha' Z' B' + \alpha'' Z'' B'' = B + \hat{N}$, where B is defined in 3.41 and $\hat{N} = \alpha' B' \cdot Z' + \alpha'' B'' \cdot Z''$ is a P-local martingale. Then $X(h) - X_0 = B + N + \hat{N}$, which gives that $X(h)$ is a P-semimartingale, so X also a P-semimartingale, and which also gives that B is the first characteristic of X under P.

Call $\hat{C}$ the second characteristic of X under P. By 3.24 we have $\hat{C} = C'$ P'-a. s., hence $\alpha' Z'_- \cdot \hat{C} = \alpha' Z'_- \cdot C'$ P-a. s. by 3.7, and similarly $\alpha'' Z''_- \cdot \hat{C} = \alpha'' Z''_- \cdot C''$ P-a. s.; summing up and using 3.39 yields $\hat{C} = C$ P-a. s., if C is defined by 3.41.

Call $\hat{\nu}$ the third characteristic of X under P, and set $Y' = M^P_{\mu^X}\left(\frac{Z'}{Z'_-} 1_{\{Z'_- > 0\}} \middle| \tilde{\mathscr{P}}\right)$ and $Y'' = M^P_{\mu^X}\left(\frac{Z''}{Z''_-} 1_{\{Z''_- > 0\}} \middle| \tilde{\mathscr{P}}\right)$. We know that $M^P_{\mu^X}(Z' | \tilde{\mathscr{P}}) = Z'_- Y'$ and $M^P_{\mu^X}(Z'' | \tilde{\mathscr{P}}) = Z''_- Y''$, hence $\alpha' Z'_- Y' + \alpha'' Z''_- Y'' = 1$ $M^P_{\mu^X}$-a. s. and, up to a modification of Y' and Y'' on an $M^P_{\mu^X}$-null set we may assume that $\alpha' Z'_- Y' + \alpha'' Z''_- Y'' = 1$ identically.

By 3.17b(ii), $Z'_- \cdot \nu' = Z'_- Y' \cdot \hat{\nu}$ P-a. s. and $Z''_- \cdot \nu'' = Z''_- Y'' \cdot \hat{\nu}$ P-a. s., thus $\nu := \alpha' Z'_- \cdot \nu' + \alpha'' Z''_- \cdot \nu'' = \hat{\nu}$ P-a. s., and we are done. □

§ 3e. The Discrete Case

It may be worth explaining how Girsanov's Theorem looks like when the continuous-time filtration is replaced by a discrete-time one. Furthermore, even though the discrete case reduces to the continuous one (through I.1.55), the proofs in the discrete case are so much more elementary!

Let $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}})$ be a discrete-time filtered space, endowed with two probability measures P and P'. We assume that $P' \overset{\text{loc}}{\ll} P$ (same definition than 3.2: $P'_n \ll P_n$ for all $n \in \mathbb{N}$). Exactly like in 3.4 there is a density process $Z = (Z_n)_{n \in \mathbb{N}}$

which is a P-martingale, and 3.4i, ii, iii hold. In particular

3.43 $$Z_n = \frac{dP'_n}{dP_n},$$

and indeed the verification that the process defined by 3.43 is a P-martingale in trivial here (much easier than in 3.4). 3.6 reads as:

3.44 $$Z_{n+1} = 0 \ P\text{-a.s. on the set } \{Z_n = 0\}$$

and this is also obvious to check: write $Z_n = E(Z_{n+1}|\mathscr{F}_n)$ and recall that $Z \geq 0$.

In 3.12 or 3.15 the process Z/Z_- plays a central rôle. In our present setting, it is replaced by

3.45 $$\alpha_n = \frac{Z_n}{Z_{n-1}} 1_{\{Z_{n-1}>0\}}$$

3.46 **Theorem.** *In addition to the above, let M be a local martingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ such that $M_0 = 0$ and that*

3.47 $$E(|M_p - M_{p-1}|\alpha_p|\mathscr{F}_{p-1}) < \infty \quad \forall p \in \mathbb{N}^*.$$

Then the process

3.48 $$M'_n = M_n - \sum_{1\leq p\leq n} E(\alpha_p(M_p - M_{p-1})|\mathscr{F}_{p-1})$$

in a P'-local martingale.

The reader should recognize Theorem 3.11 here: $[M, Z]$ is the process

$$[M, Z]_n = \sum_{1\leq p\leq n} (M_p - M_{p-1})Z_{p-1}(\alpha_p - 1)$$

and so its P-compensator is

$$\langle M, Z\rangle_n = \sum_{1\leq p\leq n} Z_{p-1}E((M_p - M_{p-1})\alpha_p|\mathscr{F}_{p-1})$$

(observe that $E(|M_p - M_{p-1}||\mathscr{F}_{p-1}) < \infty$ by I.1.64). Hence 3.48 is exactly 3.12.

Proof. $M_0 = 0$ by construction. Moreover, for any $\mathscr{F}_n$-measurable random variable Y, we have by definition of Z and α:

$$E'(Y|\mathscr{F}_{n-1}) = \frac{1}{Z_{n-1}} E(YZ_n|\mathscr{F}_{n-1}) = E(Y\alpha_n|\mathscr{F}_{n-1})$$

(recall 3.44). Hence, using 3.47:

$$\begin{aligned} E'(M_n|\mathscr{F}_{n-1}) &= E(M_n\alpha_n|\mathscr{F}_{n-1}) = E(\alpha_n(M_n - M_{n-1})|\mathscr{F}_{n-1}) + E(M_{n-1}\alpha_n|\mathscr{F}_{n-1}) \\ &= E(\alpha_n(M_n - M_{n-1})|\mathscr{F}_{n-1}) + M_{n-1} \end{aligned}$$

($E(\alpha_n|\mathscr{F}_{n-1}) = 1$ because Z is a P-martingale). In particular, this implies $E'(|M_n||\mathscr{F}_{n-1}) < \infty$. If we compare to 3.48, we deduce $E'(|M'_n||\mathscr{F}_{n-1}) < \infty$ and $E'(M'_n|\mathscr{F}_{n-1}) = M'_{n-1}$: hence the result follows from I.1.64. □

There is also a version of Theorem 3.24 about the characteristic of a semimartingale X on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ (see II.3.11), but it does not seem very useful, and we leave it as an (easy) exercise for the reader.

4. Representation Theorem for Martingales

In this section we address the following problem: let X be a d-dimensional semimartingale on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with characteristics (B, C, ν) and with X^c its continuous martingale part and $\mu = \mu^X$ the measure associated with its jumps by II.1.16. Then, is it the case that every local martingale is the sum of a stochastic integral with respect to X^c and a stochastic integral with respect to $\mu - \nu$? Beside its own interest, this property (when true) will allow us to explicitely compute the density process of any other measure P' such that $P' \overset{\text{loc}}{\ll} P$ with respect to P.

To begin with, we need to expound some complements to Chapter I (§ 4a) and to Chapter II (§ 4b), as for example the meaning of a stochastic integral with respect to the multi-dimensional local martingale X^c.

§ 4a. Stochastic Integrals with Respect to a Multi-Dimensional Continuous Local Martingale

This subsection complements § I.4d, to which we borrow all our notation. The stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is fixed, as well as a *continuous* d-dimensional local martingale $X = (X^i)_{i \leq d}$.

In Theorem I.4.40 we introduced the most general stochastic integrals with respect to each component X^i separately, namely the integrals of those processes in $L^2_{\text{loc}}(X^i)$. So if we wish to integrate with respect to X it seems natural, at first sight, to do as follows: let $H = (H^i)_{i \leq d}$ be a predictable process with $H^i \in L^2_{\text{loc}}(X^i)$ for all $i \leq d$, then set

4.1 $$H \cdot X = \sum_{i \leq d} H^i \cdot X^i.$$

However, we will see that usually this is *not* the most general stochastic integral of d-dimensional processes with respect to X.

In order to extend 4.1 we consider a factorization:

4.2 $\langle X^i, X^j \rangle = c^{ij} \cdot A$, $(c^{ij})_{i,j \leq d}$ predictable process with values in the set of all nonnegative symmetric $d \times d$ matrices, A predictable increasing process. □

(There are many such factorizations: see the proof of II.2.9). For every predictable process $H = (H^i)_{i \leq d}$ we set $H \cdot c \cdot H = \sum_{i,j \leq d} H^i c^{ij} H^j$, and

4.3 $L^2(X)$ (resp. $L^2_{\text{loc}}(X)$) is the set of all predictable processes H such that the increasing process $(H \cdot c \cdot H) \cdot A$ is integrable (resp. locally integrable). □

(Compare to I.4.39; clearly $L^2(X)$ and $L^2_{\text{loc}}(X)$ do not depend on the factorization in 4.2).

For the next theorem, one needs a further observation. Let $Y \in \mathscr{H}^2_{\text{loc}}$. Then $d\langle Y, X^i\rangle_t \ll d\langle X^i, X^i\rangle_t$: indeed if $Z^1 = X^i$ and $Z^2 = Y$, there is a factorization $\langle Z^i, Z^j\rangle = \tilde{c}^{ij} \cdot \tilde{A}$ as in 4.2, and $\tilde{c}^{12} = 0$ whenever $\tilde{c}^{11} = 0$ because of the non-negativeness of $\tilde{c}$: thus $d\langle Z^1, Z^2\rangle_t \ll d\langle Z^1, Z^1\rangle_t$, or equivalently $d\langle Y, X^i\rangle_t \ll c^{ii}_t\, dA_t \ll dA_t$. Therefore, there exists a predictable process c^{Yi} such that:

4.4 $$\langle Y, X^i\rangle = c^{Yi} \cdot A, \quad c^{Yi} = 0 \text{ whenever } c^{ii} = 0.$$

4.5 **Theorem.** *Let $X = (X^i)_{i \le d}$ be a continuous local martingale, and $H \in L^2_{\text{loc}}(X)$.*

a) *If $H(n) = H 1_{\{|H| \le n\}}$ then $H(n) \cdot X$* (defined by 4.1) *converges in measure, uniformly on every compact interval, to a limit denoted $H \cdot X$.*

b) *$H \cdot X$ is also characterized as follows: it is the unique* (up to evanescence) *continuous local martingale, null at* 0, *such that*

4.6 $$\langle H \cdot X, Y\rangle = \left(\sum_{i \le d} H^i c^{Yi}\right) \cdot A$$

for all $Y \in \mathscr{H}^2_{\text{loc}}$.

c) *If $H, K \in L^2_{\text{loc}}(X)$, then*

4.7 $$\langle H \cdot X, K \cdot X\rangle = (H \cdot c \cdot K) \cdot A.$$

d) *$H \cdot X$ belongs to $\mathscr{H}^2$ if and only if $H \in L^2(X)$.*

e) *$H \rightsquigarrow H \cdot X$ is linear* (up to evanescence) *on $L^2_{\text{loc}}(X)$, and*

4.8 for all stopping times T, $H 1_{[\![0,T]\!]} \in L^2_{\text{loc}}(X)$ and $(H 1_{[\![0,T]\!]}) \cdot X = (H \cdot X)^T$.

4.9 if K is a predictable locally bounded $\mathbb{R}$-valued process, then $KH \in L^2_{\text{loc}}(X)$ and $(KH) \cdot X = K \cdot (H \cdot X)$.

Proof. a) Assume first $H \in L^2(X)$. Set $M^n := H(n) \cdot X$, as defined by 4.1; then 4.2 and I.4.41 yield for $m \le n$:

$$\langle M^n - M^m, M^n - M^m\rangle_\infty = (H \cdot c \cdot H) 1_{\{m < |H| \le n\}} \cdot A_\infty$$

which goes to 0 in L^1 as $n, m \uparrow \infty$, whereas by I.4.6 the expected value of the above equals $\|M^n - M^m\|^2_{H^2}$: hence M^n belongs to $\mathscr{H}^2$, and converges to a limit M (which we also denote $H \cdot X$) in the Hilbert space $\mathscr{H}^2$; furthermore $H \cdot X$ is continuous by I.4.8. Since $H(n) 1_{[\![0,T]\!]} \cdot X = (H(n) \cdot X)^T$ for all n and all stopping times T by I.4.37, we obtain that 4.8 holds by passing to the limit as $n \uparrow \infty$.

Next, assume $H \in L^2_{\text{loc}}(X)$ and let (T_n) be a localizing sequence such that $H 1_{[\![0,T_n]\!]} \in L^2(X)$. Then $H 1_{[\![0,T_n]\!]} \cdot X = (H 1_{[\![0,T_{n+1}]\!]} \cdot X)^{T_n}$ from what precedes, and by pasting together we construct a continuous process $H \cdot X \in \mathscr{H}^2_{\text{loc}}$ which meets the requirement, and 4.8 as well.

b) We first prove that $H \cdot X$ satisfies 4.6. For that, and due to 4.8, it suffices by localization to consider the case when $H \in L^2(X)$ and $Y \in \mathscr{H}^2$. Then $M^n = H(n) \cdot X$ has, by I.4.41:

$$\langle M^n, Y \rangle_t = \left(\sum_{i \le d} H^i c^{Yi} \right) 1_{\{|H| \le n\}} \cdot A_t$$

(recall 4.4), which converges in L^1 to $(\sum_{i \le d} H^i c^{Yi}) \cdot A_t$ for all times t; on the other hand, $M_t^n Y_t \to (H \cdot X)_t Y_t$ in L^1 as $n \uparrow \infty$ (because $M^n \to M$ in $\mathscr{H}^2$), while $M^n Y - \langle M^n, Y \rangle$ is a martingale for all n; therefore $(H \cdot X) Y - (\sum_{i \le d} H^i c^{Yi}) \cdot A$ is a martingale, and due to the fact that $(\sum_{i \le d} H^i c^{Yi}) \cdot A$ is continuous and to the characterization of $\langle H \cdot X, Y \rangle$ (see I.4.2), we deduce 4.6.

Now let $M' \in \mathscr{H}^2_{\text{loc}}$ with $M'_0 = 0$ and $\langle M', Y \rangle = (\sum_{i \le d} H^i c^{Yi}) \cdot A$ for all $Y \in \mathscr{H}^2_{\text{loc}}$. By 4.4 we obtain $c^{X^j,k} = c^{jk}$, hence $c^{M'i} = \sum_{j \le d} H^j c^{ji}$, hence $\langle M', M' \rangle = (H \cdot c \cdot H) \cdot A$. In particular, $M = H \cdot X$ has $\langle M, M \rangle = (H \cdot c \cdot H) \cdot A$, and the same proof also yields $\langle M, M' \rangle = (H \cdot c \cdot H) \cdot A$. Thus $\langle M - M', M - M' \rangle = 0$, so $M - M'$ is orthogonal to itself (I.4.15), and we deduce $M' = M$ (I.4.13).

c) We have proved 4.7 for $K = H$ above, and the general case follows by polarization.

d) That follows from (c), via the same proof than for I.4.40b.

e) The first claim is obvious, and 4.8 has already been shown. The first claim in 4.9 is also obvious. Finally, if $Y \in \mathscr{H}^2_{\text{loc}}$ we deduce from 4.6 and I.4.41 that

$$\langle K \cdot (H \cdot X), Y \rangle = K \cdot \langle H \cdot X, Y \rangle = \left(K \sum_{i \le d} H^i c^{Yi} \right) \cdot A = \left(\sum_{i \le d} K H^i c^{Yi} \right) \cdot A$$

and $(KH) \cdot X = K \cdot (H \cdot X)$ then follows from (b). □

4.10 **Example.** This is to show that the integral constructed in 4.5 is more general than the one in 4.1.

Let Y and Z be two independent standard Wiener processes on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, so $\langle Y, Y \rangle_t = \langle Z, Z \rangle_t = t$ and $\langle Y, Z \rangle = 0$. Let K be a predictable process with values in $(0, 1)$, and set $X^1 = Y$ and $X^2 = K \cdot Y + (1 - K) \cdot Z$. So we easily check that 4.2 holds with $A_t = t$, and $c^{11} = 1$, $c^{12} = K$ and $c^{22} = (K)^2 + (1 - K)^2$.

Now, take the 2-dimensional process $H = (H^1, H^2)$ with components $H^1 = -K/(1 - K)$ and $H^2 = 1/(1 - K)$. Then $H \cdot c \cdot H = 1$ and thus $H \in L^2_{\text{loc}}(X)$. However, $H^1 \in L^2_{\text{loc}}(X^1)$ if and only if

$$\int_0^t \left(\frac{K_s}{1 - K_s} \right)^2 ds < \infty \quad \text{a.s. for all } t \in \mathbb{R}_+,$$

and this property certainly fails if $K_s = 1 - s$ for $s \in (0, \frac{1}{2})$. □

Now we "project" an arbitrary local martingale on X, according to the orthogonality defined in I.4.11:

4.11 **Theorem.** *Let $X = (X^i)_{i \le d}$ be a continuous local martingale, and Z be an arbitrary local martingale. We use notation* 4.2.

a) *There is a predictable process $H = (H^i)_{i \le d}$ such that*

4.12
$$[Z, X^i] = \langle Z^c, X^i \rangle = \left(\sum_{j \le d} c^{ij} H^j \right) \cdot A.$$

b) *Any predictable process meeting* 4.12 *belongs to $L^2_{\mathrm{loc}}(X)$, and the stochastic integral $H \cdot X$ does not depend upon the chosen version of H, and $Y = Z - H \cdot X$ is orthogonal to all components of X and*

4.13
$$[Y, X^i] = \langle Y^c, X^i \rangle = 0.$$

Proof. (a) and the first claim in (b) are nothing else than Lemma 3.31, once noticed that $[Z, X^i] = \langle Z^c, X^i \rangle$ (as well as $[Y, X^i] = \langle Y^c, X^i \rangle$ in 4.13) because of the continuity of X^i (see I.4.53).

That $Y = Z - H \cdot X$ satisfies 4.13 immediately follows from 4.6 and 4.12, and the orthogonality of Y and X^i then follows from I.4.15.

Finally let H' be another process satisfying 4.12, and $Y' = Z - H' \cdot X$. Then $Y - Y' = (H' - H) \cdot X$ is orthogonal to all X^i, and with notation 4.4 that implies $c^{Y-Y',i} \cdot A = 0$ for all i; then 2.6 yields $\langle Y - Y', (H' - H) \cdot X \rangle = 0$: so $Y - Y'$ is orthogonal to itself, and $Y_0 = Y'_0 = 0$. Therefore $Y = Y'$ and the integral $H \cdot X$ does not depend on the version chosen for H. □

4.14 **Remark.** This theorem partially explains why $L^2_{\mathrm{loc}}(X)$ is the largest set of integrands with respect to X, since a local martingale that is orthogonal to X should certainly not be a stochastic integral with respect to X! As a matter of fact, one could do exactly the same thing if X were not continuous, but $X^i \in \mathscr{H}^2_{\mathrm{loc}}$ for all i: but doing so would not yield the most general integrals, because an integral $H \cdot X$ can be discontinuous if X is so, and thus there is no reason for $H \cdot X$ to belong to $\mathscr{H}^2_{\mathrm{loc}}$ itself. □

§ 4b. Projection of a Local Martingale on a Random Measure

The setting is exactly as in Section II.1, to which we borrow all notation. We start with an integer-valued random measure μ on $\mathbb{R}_+ \times E$, defined on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, and where $(E, \mathscr{E})$ is some auxiliary Blackwell space. Recall that $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$.

Let us recall some facts. Firstly, μ has the form

4.15
$$\mu(\omega; dt, dx) = \sum_{s \ge 0} 1_D(\omega, s) \varepsilon_{(s, \beta_s(\omega))}(dt, dx)$$

where D is an optional thin set and β is an optional E-valued process. We call ν a "good" version of the compensator of μ, so that if

4.16 $$a_t(\omega) = \nu(\omega; \{t\} \times E),$$

then $a \leq 1$ identically. Moreover, according to II.1.24, for each measurable function W on $\tilde{\Omega}$ we set

4.17 $$\hat{W}_t(\omega) = \begin{cases} \int_E W(\omega, t, x)\nu(\omega; \{t\} \times dx) & \text{if this integral converges} \\ +\infty & \text{otherwise,} \end{cases}$$

and by II.1.25 we have:

4.18 (i) If W is $\tilde{\mathscr{P}}$-measurable nonnegative, then $\hat{W}$ is a version of the predictable projection of the process $1_D(t)W(t, \beta_t)$, or equivalently, $\hat{W}_T = E(W(T, \beta_T)1_D(T)|\mathscr{F}_{T-})$ on $\{T < \infty\}$ if T is predictable;

(ii) In particular, a is the predictable projection of 1_D. □

Finally, recall that we have defined the measure M_μ^P by 3.15, and we have a notion of M_μ^P-conditional expectation with respect to $\tilde{\mathscr{P}}$. Then 4.18 admits the following extension:

4.19 **Lemma.** *Let W be an $\mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{E}$-measurable function on $\tilde{\Omega}$ and let $W' = M_\mu^P(W|\tilde{\mathscr{P}})$. Then for all predictable times T,*

$$\hat{W}'_T = E(W(T, \beta_T)1_D(T)|\mathscr{F}_{T-}) \quad \text{on } \{T < \infty\}.$$

Proof. Due to the definition of generalized conditional expectations, it suffices to consider the case $W \geq 0$. Let $A \in \mathscr{F}_{T-}$. Then

$$\begin{aligned} E(\hat{W}'_T 1_{A \cap \{T<\infty\}}) &= E(1_A W'(T, \beta_T)1_D(T)) && \text{(by 4.18)} \\ &= M_\mu^P(W' 1_{[\![T_A]\!] \times E}) && \text{(definition of } M_\mu^P) \\ &= M_\mu^P(W 1_{[\![T_A]\!] \times E}) && \text{(because } [\![T_A]\!] \times E \in \tilde{\mathscr{P}}) \\ &= E(1_A W(T, \beta_T)1_D(T)) && \text{(definition of } M_\mu^P), \end{aligned}$$

hence the result. □

4.20 **Theorem.** *Let X be a local martingale, and $U = M_\mu^P(\Delta X|\tilde{\mathscr{P}})$* (here, ΔX is considered as defined on $\tilde{\Omega}$ by $\Delta X(\omega, t, x) = \Delta X_t(\omega)$).

a) *There is a version of U such that $\{a = 1\} \subset \{\hat{U} = 0\}$.*

b) *Let $W = U + \dfrac{\hat{U}}{1-a} 1_{\{a<1\}}$. Then $W \in G_{\mathrm{loc}}(\mu)$, and if $Y = W * (\mu - \nu)$ and $Z = X - Y$, then we have $M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = 0$.*

$M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = 0$ may be interpreted as Z being "orthogonal" to μ, so Y is a sort of *projection* of X on μ (or, rather, on the space of all integrals of the form $V * (\mu - \nu)$).

Proof. a) We shall apply the predictable section theorem, in its version I.2.18, to the predictable random set $A = \{a = 1\} \cap \{\hat{U} \neq 0\}$. Indeed, let T be a predictable time such that $[\![T]\!] \subset A$. We deduce from 4.18:

$$1 = a_T = E(1_D(T)|\mathscr{F}_{T-}) \quad \text{on } \{T < \infty\}$$

and so $1_D(T) = 1$ a.s. on $\{T < \infty\}$. Then 4.19 yields on $\{T < \infty\}$:

$$\hat{U}_T = E(\Delta X_T 1_D(T)|\mathscr{F}_{T-}) = E(\Delta X_T|\mathscr{F}_{T-}) = 0$$

(use I.2.27); since $[\![T]\!] \subset \{\hat{U} \neq 0\}$ we deduce that $T = \infty$ a.s.: therefore A is evanescent. So if we replace U by $U 1_{A \times E}$ we do not alter the property $U = M_\mu^P(\Delta X|\tilde{\mathscr{P}})$, while $\hat{U} = 0$ identically on $\{a = 1\}$.

b) Suppose first that the claims are true for some version U, and let U' be another version of $M_\mu^P(\Delta X|\tilde{\mathscr{P}})$, and $W' = U' + [\hat{U}'/(1 - a)] 1_{\{a<1\}}$. Then $U' = U$ M_μ^P-a.e., so $1_{\{U' \neq U\}} * \mu_\infty = 0$ a.s., hence $1_{\{U' \neq U\}} * \nu_\infty = 0$ a.s., hence $\tilde{W}$ and $\tilde{W}'$ are indistinguishable (use notation II.1.27). Then $W' \in G_{\text{loc}}(\mu)$ and $Z' = X - W' * (\mu - \nu)$ is indistinguishable from Z: thus $M_\mu^P(\Delta Z'|\tilde{\mathscr{P}}) = 0$ as well, and U' also meets the claimed properties. Thus, to prove the result we can use the version of U constructed in (a). In this case, with $\frac{0}{0} = 0$ we have

4.21 $$W = U + \frac{\hat{U}}{1 - a}, \quad \hat{W} = \frac{\hat{U}}{1 - a}, \quad \tilde{W}_t = U(t, \beta_t) 1_D(t) - \frac{\hat{U}_t}{1 - a_t} 1_{D^c}(t).$$

Now, the claimed properties are "linear" in X: hence, due to I.4.17 and up to a localization, it suffices to prove the result separately for the two cases $X \in \mathscr{L} \cap \mathscr{A}$ and $X \in \mathscr{H}^2$. Let (T_n) be a sequence of predictable times that exhausts the predictable thin set $\{a > 0\}$.

Suppose first $X \in \mathscr{L} \cap \mathscr{A}$. Then 4.19 and 4.21 and II.1.32 yield:

$$E(\bar{C}(W)_\infty) = E(|U| * \nu_\infty) + \sum_n E(|\hat{U}_{T_n}| 1_{\{T_n < \infty\}}) = M_\mu^P(|U|) + \sum_n E(|\hat{U}_{T_n}| 1_{\{T_n < \infty\}})$$

$$\leq M_\mu^P(|\Delta X|) + \sum_n E(|\Delta X_{T_n}| 1_{\{T_n < \infty\}}) \leq 2E\left(\sum_{s>0} |\Delta X_s|\right) < \infty.$$

Therefore II.1.33b gives $W \in G_{\text{loc}}(\mu)$.

Secondly, suppose that $X \in \mathscr{H}^2$. Due to 4.21 and II.1.31,

$$E[C(W)_\infty]$$

$$= E(U^2 * \nu_\infty) + \sum_n E\left(\frac{1}{1 - a_{T_n}} (\hat{U}_{T_n})^2 1_{\{T_n < \infty\}}\right)$$

$$= E(U^2 * \mu_\infty) + \sum_n E\left[\frac{1}{1 - a_{T_n}} 1_{\{T_n < \infty\}} E(\Delta X_{T_n} 1_D(T_n)|\mathscr{F}_{T_n-})^2\right] \qquad \text{(by 4.19)}$$

$$= M_\mu^P(U^2) + \sum_n E\left(\frac{1}{1 - a_{T_n}} 1_{\{T_n < \infty\}} E(\Delta X_{T_n} 1_{D^c}(T_n)|\mathscr{F}_{T_n-})^2\right)$$

(because $E(\Delta X_{T_n}|\mathscr{F}_{T_n-}) = 0$ on $\{T_n < \infty\}$). Hence

$$E[C(W)_\infty]$$

$$\leq M_\mu^P((\Delta X)^2) + \sum_n E\left(\frac{1}{1-a_{T_n}} 1_{\{T_n<\infty\}} E((\Delta X_{T_n})^2|\mathscr{F}_{T_n-}) E(1_{D^c}(T_n)|\mathscr{F}_{T_n-})\right)$$

$$= M_\mu^P((\Delta X)^2) + \sum_n E(1_{\{T_n<\infty\}} E((\Delta X_{T_n})^2|\mathscr{F}_{T_n-}))$$

$$= E\left[\sum_{s>0} (\Delta X_s)^2 1_D(s) + \sum_n 1_{\{T_n<\infty\}} (\Delta X_{T_n})^2\right] \leq 2E([X,X]_\infty) < \infty.$$

Therefore $C(W) \in \mathscr{A}^+$ and $W \in G_{\mathrm{loc}}(\mu)$ by II.1.33a.

Adding these two results, we obtain that $W \in G_{\mathrm{loc}}(\mu)$ for any local martingale X. Finally, if $Z = X - W*(\mu-\nu)$ we have $\Delta Z = \Delta X - \widetilde{W}$, and $\widetilde{W} = U$ M_μ^P-a.e. (because $M_\mu^P(D^c \times E) = 0$ by definition). Hence $M_\mu^P(\Delta Z|\tilde{\mathscr{P}}) = M_\mu^P(\Delta X|\tilde{\mathscr{P}}) - U$ is equal to 0. □

§ 4c. The Representation Property

Until the end of the present section, the setting is as follows: the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ supports a semimartingale $X = (X^i)_{i\leq d}$ with characteristics (B, C, ν), relative to some truncation function h; X^c is the continuous martingale part of X, and $\mu = \mu^X$ is defined by II.1.16.

4.22 **Definition.** We say that *a local martingale M has the representation property relative to X* if it has the form

4.23 $$M = M_0 + H \cdot X^c + W*(\mu - \nu)$$

where $H = (H^i)_{i\leq d}$ belongs to $L^2_{\mathrm{loc}}(X^c)$ (see 4.3) and $W \in G_{\mathrm{loc}}(\mu)$ (see II.1.27). □

In many respects it would be more natural to suppress the reference to a basic semimartingale X: thus X^c would be replaced by an arbitrary continuous local martingale, and $\mu = \mu^X$ by an arbitrary integer-valued random measure μ with compensator ν. The results presented below would remain true.

4.24 **Lemma.** *Every local martingale M has a decomposition*

4.25 $$M = H \cdot X^c + W*(\mu-\nu) + N,$$

where $H \in L^2_{\mathrm{loc}}(X^c)$, $W \in G_{\mathrm{loc}}(\mu)$, and

4.26 $$\langle N^c, (X^c)^i\rangle = 0 \quad \forall i \leq d, \quad M_\mu^P(\Delta N|\tilde{\mathscr{P}}) = 0.$$

Moreover, this decomposition is unique, up to indistinguishability (although H and W are not necessarily unique!)

Proof. a) The existence of a decomposition 4.25 follows immediately from Theorems 4.11 and 4.20. In particular, $M^c = H \cdot X^c + N^c$ and the uniqueness of this decomposition for M^c also follows from 4.11.

To obtain the uniqueness of 4.25 it remains to prove that any purely discontinuous local martingale M of the form $M = W * (\mu - \nu)$, with $M_\mu^P(\Delta M|\tilde{\mathscr{P}}) = 0$, has $M = 0$ (a. s.). We use the notation $a_t = \nu(\{t\} \times \mathbb{R}^d)$ of 4.16, and we know that a is the predictable projection ${}^p(1_{\{\Delta X \neq 0\}})$ of the indicator process $1_{\{\Delta X \neq 0\}}$, while we have seen in the proof of 4.20 that $\{a = 1\} \subset \{\Delta X \neq 0\}$, up to an evanescent set. Now, by definition of the stochastic integral $W * (\mu - \nu)$, $\Delta M_t = W(t, \Delta X_t) 1_{\{\Delta X_t \neq 0\}} - \hat{W}_t$, hence $\Delta M = W - \hat{W}$ M_μ^P-a. s., and $W - \hat{W} = M_\mu^P(\Delta M|\tilde{\mathscr{P}}) = 0$ M_μ^P-a. s. by hypothesis: thus we remain with $\Delta M = -\hat{W} 1_{\{\Delta X = 0\}}$ (up to an evanescent set), and even with $\Delta M = -\hat{W} 1_{\{\Delta X = 0\}} 1_{\{a < 1\}}$ because $\{a = 1\} \subset \{\Delta X \neq 0\}$. Therefore ${}^p(\Delta M) = -\hat{W} 1_{\{a<1\}}(1 - a)$ and since ${}^p(\Delta M) = 0$ by I.2.31 we obtain that $\hat{W} = 0$ on $\{a < 1\}$, hence $\Delta M = 0$ and so $M = 0$ by I.4.19. □

4.27 **Corollary.** *The following three statements are equivalent*:

(i) *All local martingales have the representation property.*

(ii) *All local martingales satisfying 4.26 are trivial* (a local martingale N is called *trivial* if $N_t = N_0$ a. s. for all $t \in \mathbb{R}_+$).

(iii) *All bounded martingales satisfying 4.26 are trivial.*

Proof. The equivalence (i)$\Leftrightarrow$(ii) readily follows from the previous lemma, and (ii)$\Rightarrow$(iii) is evident. It remains to prove that if (ii) fails, then so does (iii).

Therefore we assume that (ii) fails, so there exists a non-trivial local martingale Y with 4.26. We will exhibit a non-trivial local martingale M_0 with $M_0 = 0$ and bounded jumps and 4.26. Then, if $T_n = \inf(t\colon |M_t| > n)$, M^{T_n} is a bounded martingale and meets 4.26 and is nontrivial for n large enough: so the claim will follow.

If Y^c is non-trivial, $M = Y^c$ obviously has the required properties. Hence we further assume that Y is purely discontinuous, and we single out two cases:

a) $M_\mu^P(\Delta Y \neq 0) > 0$: since $M_\mu^P(\Delta Y|\tilde{\mathscr{P}}) = 0$, this implies that the σ-fields $\mathscr{O} \otimes \mathscr{R}^d$ and $\tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{R}^d$ of $\tilde{\Omega}$ have not the same completion. So there is a set $A \in \mathscr{O} \otimes \mathscr{R}^d$ such that $M_\mu^P(A) < \infty$ and $M_\mu^P(V \neq 0) > 0$, where $V = 1_A - M_\mu^P(1_A|\tilde{\mathscr{P}})$. The process $M = V * \mu$ is càdlàg, adapted, with integrable variation, $M_0 = 0$, and for all stopping times T:

$$E(M_T) = M_\mu^P(V 1_{[\![0,T]\!]}) = 0$$

because $M_\mu^P(V|\tilde{\mathscr{P}}) = 0$ by construction. Hence M is a non-trivial martingale (see I.1.44) with $|\Delta M| \leq 1$, while 4.26 is met by construction.

b) $M_\mu^P(\Delta Y \neq 0) = 0$, that is X and Y have a. s. no common jump. Suppose first that

4.28 $\Delta Y_T \neq 0$ on $\{T < \infty\}$ for a totally inaccessible time T with $P(T < \infty) > 0$

Set $A = 1_{[\![T,\infty[\![}$ and let A^p be its compensator and $M = A - A^p$: then M is a non-trivial purely discontinuous local martingale, and $M_0 = 0$, and $|\Delta M| \leq 1$, and $\Delta M = \Delta A = 0$ on $\{\Delta X \neq 0\}$, hence 4.26 holds for M.

Finally, suppose that 4.28 fails, and recall that $\{\Delta Y \neq 0\} \cap \{\Delta X \neq 0\} = \emptyset$ and $\{a = 1\} \subset \{\Delta X \neq 0\}$, up to an evanescent set ($a_t = \nu(\{t\} \times \mathbb{R}^d)$; see the proof of 4.24). Since Y is not trivial, there is a predictable time T and two constants $\alpha > 0$, $\beta \in (0,1)$ such that $P(A) > 0$, where $A = \{T < \infty, |\Delta Y_T| > \alpha\}$ and that $[\![T]\!] \subset \{a \leq \beta\}$. Set

$$M = U 1_{[\![T,\infty[\![} \quad \text{with} \quad U = 1_A - \frac{P(A|\mathscr{F}_{T-})}{1 - a_T} 1_{\{\Delta X_T = 0\}}.$$

M is bounded because $|U| \leq 1 + \dfrac{1}{1-\beta}$; moreover

$$E(U|\mathscr{F}_{T-}) = P(A|\mathscr{F}_{T-})\left[1 - \frac{P(\Delta X_T = 0|\mathscr{F}_{T-})}{1 - a_T}\right] = 0$$

on $\{T < \infty\}$. We easily deduce that M is a martingale. Finally, $\Delta X_T = 0$ on the set A, so $\Delta M_T = 0$ if $\Delta X_T \neq 0, T < \infty$: then $\Delta M = 0$ M_μ^P-a.s., and 4.26 holds. □

§ 4d. The Fundamental Representation Theorem

1. As we shall presently see, the representation property is closely related to a martingale problem. In order to fit Definition 2.4, we add an *initial condition* to our on-going setting (which is as in § 4c): namely, $\mathscr{H}$ is a sub-σ-field of $\mathscr{F}_0$, and $P_H = P_{|\mathscr{H}}$ is the restriction of P to $\mathscr{H}$.

In particular, our given measure P is a solution to the problem $\mathscr{s}(\mathscr{H}, X|P_H; B, C, \nu)$ introduced in 2.4.

4.29 **Theorem.** *In addition to the above, assume that $\mathscr{F} = \mathscr{F}_{\infty-}$. Then the following four statements are equivalent:*

(i) *All local martingales have the representation property, relative to X; moreover $\mathscr{F}_0$ is contained in the σ-field generated by $\mathscr{H}$ and the P-null sets of $\mathscr{F}$.*

(ii) *P is an extremal point in the convex set $\mathscr{s}(\mathscr{H}, X|P_H; B, C, \nu)$.*

(iii) *If $P' \in \mathscr{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $P' \overset{\text{loc}}{\ll} P$, then $P' = P$.*

(iv) *If $P' \in \mathscr{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $P' \ll P$, then $P' = P$.*

4.30 **Remark.** (i) is a property of the filtration $\mathbf{F}$ and thus only involves the restriction of P to $\mathscr{F}_{\infty-}$. Contrarywise, (ii)–(iv) involve the "full" measure P on $(\Omega, \mathscr{F})$. This is why we assumed $\mathscr{F} = \mathscr{F}_{\infty-}$.

Now, when the inclusion $\mathscr{F}_{\infty-} \subset \mathscr{F}$ is strict, anyone of (ii), (iii), (iv) obviously implies (i). □

Proof. (iii) $\Rightarrow$ (iv): evident.

(iv) $\Rightarrow$ (ii): Suppose that P is a convex combination $P = \alpha P' + (1-\alpha)P''$ of two other solutions P' and P'', with $0 < \alpha < 1$. Then obviously $P' \ll P$ and $P'' \ll P$, hence (iv) yields $P' = P'' = P$: thus P is extremal in the set of all solutions.

(ii) $\Rightarrow$ (i): If the representation property fails, 4.27 gives a non-trivial bounded martingale M with $M_0 = 0$, which meets 4.26, and we may of course assume that $|M| \leq 1$. If the second property of (i) fails, there is a set $A \in \mathscr{F}_0$ such that $P(A|\mathscr{H})$ is not a.s. equal to 1_A; set $M_t = 1_A - P(A|\mathscr{H})$ for all t, so M is a martingale, with $|M| \leq 1$ and which obviously satisfies 4.26.

Hence, in case (i) fails, we have constructed a martingale $Z = 1 + M$ with $0 \leq Z \leq 2$, $E(Z_t) = 1$ for all $t \in \overline{\mathbb{R}}_+$, and 4.26 and $P(Z_\infty = 1) < 1$. So $P'(d\omega) = P(d\omega)Z_\infty(\omega)$ defines a new probability measure P' on $(\Omega, \mathscr{F})$ and 3.28 is met with $Y = 1, \beta = 0$: hence 3.24 yields that X is a P'-semimartingale with characteristics (B, C, ν). Moreover $E(Z_\infty|\mathscr{H}) = E(Z_0|\mathscr{H}) = 1$ by construction, hence the restriction of P' to $\mathscr{H}$ equals P_H and thus $P' \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$. Finally $P' \neq P$ because $P(Z_\infty = 1) < 1$.

Similarly if $Z' = 1 - M$, then $\tilde{P}'(d\omega) = P(d\omega)Z'_\infty(\omega)$ defines a solution $\tilde{P}' \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $\tilde{P}' \neq P$. Since $Z + Z' = 2$, for all $A \in \mathscr{F}$:

$$P'(A) + \tilde{P}'(A) = E(Z_\infty 1_A) + E(Z'_\infty 1_A) = 2P(A)$$

and thus $P = \frac{1}{2}(P' + \tilde{P}')$, hence contradicting (ii).

(i) $\Rightarrow$ (iii): Let $P' \in \mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$ with $P' \overset{\text{loc}}{\ll} P$, and call Z the density process.

P' and P coincide on $\mathscr{H}$, thus $E(Z_0|\mathscr{H}) = 1$. Since Z_0 is P-a.s. equal to an $\mathscr{H}$-measurable variable, it follows that $Z_0 = 1$ P-a.s.

Next, apply Theorem 3.24: by hypothesis we have 3.27 with $\beta = 0$ and $Y = 1$. Comparing to 3.28, we obtain that Z satisfies 4.26. Then, due to the representation property, 4.27 implies that Z is trivial: therefore $Z_t = 1$ P-a.s. for all $t \in \mathbb{R}_+$: we deduce that P' and P coincide on $\mathscr{F}_t$ for all $t \in \mathbb{R}_+$. Since $\mathscr{F} = \mathscr{F}_{\infty-}$ it follows that $P' = P$. $\square$

The second property in (i) is included to connect the representation property with the martingale problem $\mathfrak{s}(\mathscr{H}, X|P_H; B, C, \nu)$, but it is by no way our main concern.

Here is a simple corollary, obtained by applying 4.29 with $\mathscr{H} = \mathscr{F}_0$ (so the second property in (i) is automatically met), and for which we recall that P_0 is the restriction of P to $\mathscr{F}_0$.

4.31 **Corollary.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$. There is equivalence between:*

(i) *All local martingales have the representation property, relative to X.*

(ii) *If $P' \overset{\text{loc}}{\ll} P$ and $P'_0 = P_0$ and X admits (B, C, ν) for P'-characteristics, then $P' = P$.*

(iii) *If $P' \ll P$ and $P'_0 = P_0$ and X admits (B, C, ν) for P'-characteristics, then $P' = P$.*

The second property in 4.29(i) is a sort of "0–1 law" (it is a 0–1 law when $\mathscr{H}$ is the trivial σ-field). In the same vein, we have the following (rather unimportant) property:

4.32 **Proposition.** *Assume that* $\mathbf{F}$ *is generated by* X *and* $\mathscr{H}$ (see 2.12). *Then, if all local martingales have the representation property relative to* X, *the* σ-*field* $\mathscr{F}_t$ *is contained in the* σ-*field generated by* $\mathscr{F}_t^0$ *and by all* P-*null sets of* $\mathscr{F}$, *for all* $t \in \mathbb{R}_+$.

Proof. Let $A \in \mathscr{F}_t$ and call M the bounded martingale with terminal variable $M_\infty = 1_A$ (so $M_s = 1_A$ a.s. for $s \geq t$). By hypothesis, there exist $H \in L^2_{\text{loc}}(X^c)$ and $W \in G_{\text{loc}}(\mu)$ such that $M = M_0 + H \cdot X^c + W * (\mu - \nu)$. Hence $1_A = M_t = M_{t-} + W(t, \Delta X_t) 1_{\{\Delta X_t \neq 0\}} - \hat{W}_t$ a.s. The processes M_- and $\hat{W}$ are predictable, hence $M_{t-} - \hat{W}_t$ is $\mathscr{F}_{t-}$-measurable, and thus $\mathscr{F}_t^0$-measurable. The function W is $\tilde{\mathscr{P}}$-measurable and ΔX_t is $\mathscr{F}_t^0$-measurable, hence $W(t, \Delta X_t) 1_{\{\Delta X_t \neq 0\}}$ is $\mathscr{F}_t^0$-measurable: the claim follows. □

2. Now we give some examples: indeed, each time P is the unique solution of the problem $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ it certainly is extremal in this set! so all cases of uniqueness lead to a representation property.

Firstly, we deduce from 2.15 (or 2.16) and 4.29 the classical representation result for the martingales of a Wiener process.

4.33 **Theorem.** *Assume 2.12 and that* X *is a Wiener process on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$. *Then*

a) *Every local martingale has the form* $M = M_0 + H \cdot X$, *for some* $H \in L^2_{\text{loc}}(X)$ (in particular, every local martingale is continuous).

b) *If* $\mathscr{H}$ *is trivial or if* $\mathscr{H} = \sigma(X_0)$, *then each set of* $\mathscr{F}_0$ *has measure* 0 *or* 1 ("0–1 law"); in particular in (a), M_0 is a.s. constant.

More generally, we deduce from 2.17:

4.34 **Theorem.** *Assume 2.12 and that* X *is an* $\mathscr{H}$-*conditional PII on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$. *Then*

a) *Every local martingale has the form* $M = M_0 + H \cdot X^c + W * (\mu - \nu)$, *for some* $H = (H^i)_{i \leq d} \in L^2_{\text{loc}}(X^c)$, $W \in G_{\text{loc}}(\mu)$.

b) *If* $\mathscr{H}$ *is trivial or if* $\mathscr{H} = \sigma(X_0)$, *then each set of* $\mathscr{F}_0$ *has measure* 0 *or* 1 (in this case, X is a PII).

The reader will write by himself the representation results for diffusions and diffusions with jumps, obtained as corollaries of either 2.34, or 2.32 and 2.33.

3. The same sort of results holds for the martingale problems introduced in § 1b.

More specifically, we forget about the process X, and we suppose that $(\Omega, \mathscr{F}, \mathbf{F}, P)$ is equipped with an integer-valued random measure μ on $\mathbb{R}_+ \times E$ (with $(E, \mathscr{E})$ an auxiliary Blackwell space), with compensator ν. We also consider an initial σ-field $\mathscr{H} \subset \mathscr{F}_0$ and the restriction P_H of P to $\mathscr{H}$. Then all results of §4c and of the present subsection remain true, provided "X" is dropped everywhere. For instance, the representation property goes as follows:

4.35 A local martingale M has the *representation property relative to* μ if it has the form $M = M_0 + W * (\mu - \nu)$ for some $W \in G_{\text{loc}}(\mu)$. □

Then our main theorem 4.29 holds, with "μ" instead of "X" in (i), and $\mathscr{s}(\mathscr{H}, \mu | P_H, \nu)$ in (ii, iii, iv). Of course, each uniqueness result provides a representation result. For instance, 1.26 yields

4.36 If μ is a multivariate point process and 1.25 holds, then all local martingales have the form $M = M_0 + W * (\mu - \nu)$ for some $W \in G_{\text{loc}}(\mu)$. □

A-priori, the integral $W * (\mu - \nu)$ in 4.36 is a stochastic integral, although $\mu - \nu$ is a finite measure in restriction to every set $[0, t] \times E$ (recall that μ is a multivariate point process). However, the following shows that this integral is indeed a Stieltjes integral.

4.37 **Theorem.** *Assume that μ is a multivariate point process and that 1.25 holds. Then all local martingales have the form $M = M_0 + W * \mu - W * \nu$ where W is a $\tilde{\mathscr{P}}$-measurable function such that $|W| * \mu$ is locally integrable.*

This is a remarquable result, which in particular implies that *all local martingales have finite variation* if 1.25 holds and μ is a multivariate point process.

For instance, let N be a point process with compensator A, and assume 1.20. Then all martingales have the form

$$M = M_0 + H \cdot N - H \cdot A \tag{4.38}$$

where H is a predictable process that is Stieltjes-integrable with respect to N and A.

Proof. We use the notation of §4b, in particular a (see 4.16) and $\hat{W}$ (see 4.17). Let $M \in \mathscr{M}_{\text{loc}}$, and apply Theorem 4.20: set $U = M^P_\mu(\Delta M | \tilde{\mathscr{P}})$ and $W = U + \frac{\hat{U}}{1 - a} 1_{\{a<1\}}$, so $W \in G_{\text{loc}}(\mu)$ and $M = M_0 + W * (\mu - \nu) + N$ with $M^P_\mu(\Delta N | \tilde{\mathscr{P}}) = 0$. Then the representation property 4.36, plus 4.27, yield that $N = 0$, and it remains to prove that $|W| * \mu \in \mathscr{A}_{\text{loc}}$.

μ has the form 1.24, so up to a localization we may assume that

$$\mu(\cdot) = \sum_{1 \leq n \leq p} 1_{\{T_n < \infty\}} \varepsilon_{(T_n, Z_n)}(\cdot)$$

for some $p \in \mathbb{N}^*$. Next, the process $\sup_{s \leq \cdot} [(1 - a_s)^{-1}] 1_{\{a_s < 1\}}$ is increasing and predictable and finite-valued, so up to a further localization we may assume that

4.39 $$\beta = \sup_{s,\omega} \frac{1}{1 - a_s(\omega)} 1_{\{a_s(\omega)<1\}} < \infty.$$

Furthermore we have already seen (e.g. in the beginning of the proof of 3.17) that $\sup_{s \leq \cdot} |M_s|$ is locally integrable, so up to a further localization we may finally assume that

4.40 $$\sup_s |M_s| \quad \text{is integrable.}$$

By definition of U we have

4.41 $$M_\mu^P(|U|) \leq M_\mu^P(|\Delta M|) = E\left(\sum_{1 \leq n \leq p} 1_{\{T_n<\infty\}} |\Delta M_{T_n}|\right) < \infty$$

from 4.40. Let (S_q) be a sequence of predictable times that exhausts the thin predictable set $\{a > 0\}$. Then $\{\hat{U} \neq 0\} \subset \bigcup_q [\![S_q]\!]$, and 4.18 yields

$$\begin{aligned}\sum_{1 \leq n \leq p} E(|\hat{U}_{T_n}| 1_{\{T_n<\infty\}}) &\leq p \sum_q E(|\hat{U}_{S_q}| 1_{\{S_q<\infty\}})\\ &\leq p \sum_q E\left(\sum_{1 \leq n \leq p} |U(T_n, Z_n)| 1_{\{S_q=T_n<\infty\}}\right)\end{aligned}$$

4.42 $$\leq pE\left(\sum_{1 \leq n \leq p} |U(T_n, Z_n)| 1_{\{T_n<\infty\}}\right) = pM_\mu^P(|U|).$$

Hence,

$$\begin{aligned}E(|W| * \mu_\infty) &\leq \sum_{1 \leq n \leq p} E\left(1_{\{T_n<\infty\}}\left[|U(T_n, Z_n)| + \frac{|\hat{U}_{T_n}|}{1 - a_{T_n}} 1_{\{a_{T_n}<1\}}\right]\right)\\ &\leq \sum_{1 \leq n \leq p} E(1_{\{T_n<\infty\}} |U(T_n, Z_n)|) + \beta p M_\mu^P(|U|)\end{aligned}$$

(use 4.39 and 4.42), which equals $(1 + \beta p)M^P(|U|)$, which is finite by 4.41, and the result is proved. □

5. Absolutely Continuous Change of Measures: Explicit Computation of the Density Process

The setting is essentially the same as in Section 4: let $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis endowed with a d-dimensional semimartingale $X = (X^i)_{i \leq d}$ with characteristics (B, C, ν), relative to some fixed truncation function h. X^c denotes the continuous martingale part of X, and $\mu = \mu^X$ is the random measure associated with the jumps of X by II.1.16. We use the notation

5.1 $$a_t(\omega) = \nu(\omega; \{t\} \times \mathbb{R}^d)$$

5.2 $$\hat{W}_t(\omega) = \begin{cases} \int \nu(\omega; \{t\} \times dx) W(\omega, t, x) & \text{if this integral converges} \\ +\infty & \text{otherwise} \end{cases}$$

(as in 4.17). We also choose a "good version" of C which has:

5.3 $C^{ij} = c^{ij} \cdot A$: A is a continuous increasing process; $c = (c^{ij})_{i,j \le d}$ is predictable with values in the set of $d \times d$ symmetric nonnegative matrices.

We also consider another measure P' under which X is a semimartingale with characteristics (B', C', ν'), relative to the same truncation function h. Most of the time we will assume $P' \overset{\text{loc}}{\ll} P$, and our objective is to compute the density process Z of P' relative to P, explicitely in terms of (B, C, ν) and (B', C', ν').

§ 5a. All P-Martingales Have the Representation Property Relative to X

Since $P' \overset{\text{loc}}{\ll} P$, Girsanov's Theorem 3.24 implies that one can find:

5.4 $$\begin{cases} \beta = (\beta^i)_{i \le d}, & \text{a predictable process} \\ Y, & \text{a } \tilde{\mathscr{P}}\text{-measurable nonnegative function on } \tilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \\ & \text{(recall that } \tilde{\mathscr{P}} = \mathscr{P} \otimes \mathscr{R}^d), \end{cases}$$

such that

5.5 $$\begin{cases} B'^i = B^i + \left(\sum_{j \le d} c^{ij} \beta^j\right) \cdot A + h^i(x)(Y - 1) * \nu \\ \quad (= +\infty \text{ if one of the integrals diverges}) \\ C' = C \\ \nu' = Y \cdot \nu \end{cases}$$

up to a P'-null set.

We wish to prove that the density process Z is the Doléans-Dade exponential of a local martingale N, which we presently construct in all generality. Set

5.6 $$\sigma = \inf(t\text{: either } \hat{Y}_t > 1, \text{ or } a_t = 1 \text{ and } \hat{Y}_t < 1),$$

which obviously is a predictable time, with $\sigma > 0$ (and $P'(\sigma = \infty) = 1$ by 3.17). Next, set

5.7 $$\begin{aligned} H = (\beta \cdot c \cdot \beta) 1_{[\![0,\sigma[\![} \cdot A + (1 - \sqrt{Y})^2 1_{[\![0,\sigma[\![} * \nu \\ + \sum_{s \le \cdot} (\sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s})^2 1_{\{s < \sigma\}}. \end{aligned}$$

H is predictable and $\overline{\mathbb{R}}_+$-valued with nondecreasing paths and $H_0 = 0$. However it is not necessarily an "increasing process" in the sense of § I.3a because it may take the value $+\infty$ and it may fail to be right-continuous. In view of their forthcoming importance, we give a special name to the processes of that type:

5.8 **Definitions.** (i) A *generalized increasing process* is a process H with the following properties: it is $\overline{\mathbb{R}}_+$-valued and $H_0 = 0$ and its paths are non-decreasing, and if $T = \inf(t: H_t = \infty)$ then H is right-continuous on $[\![0, T[\![$ (and of course on $]\!]T, \infty[\![$); however, it may happen that $H_T < \infty$, while $H_{T+} = \infty$ if $T < \infty$.

(ii) We say that the generalized increasing process H *does not jump to infinity* if $H_{T-} = +\infty$ on $\{T < \infty\}$, in which case all its paths are everywhere right-continuous. □

So the process H defined in 5.7 is a generalized increasing process, because the integrals in 5.7 may start diverging at time T, while they converge on the closed interval $[0, T]$; we set $T = \inf(t: H_t = \infty)$ and

5.9 $$T_n = \inf(t: H_t \geq n), \quad \Delta = [\![0, \sigma[\![\cap \left(\bigcup_n [\![0, T_n]\!] \right).$$

T_n is a stopping time, but may fail to be predictable (one cannot apply I.2.13 because H_{T_n} may be strictly smaller than n, in case $T_n = T$). Δ is a predictable random interval and $[\![0]\!] \subset \Delta$.

5.10 **Proposition.** *There is a process N, unique* (up to P-indistinguishability) *on the set Δ, such that for every stopping time S such that $[\![0, S]\!] \subset \Delta$ the stopped process N^S is the following P-local martingale*:

5.11 $$N^S = (\beta 1_{[\![0,S]\!]}) \cdot X^c + \left(Y - 1 + \frac{\hat{Y} - a}{1 - a} 1_{\{a<1\}} \right) 1_{[\![0,S]\!]} * (\mu - \nu).$$

Extending Definition II.2.46 one might say that N is a "local martingale on the predictable random interval Δ". Of course, 5.11 implicitly assumes that:

5.12 $$\begin{cases} \beta 1_{[\![0,S]\!]} \in L^2_{\mathrm{loc}}(X^c) \\ V 1_{[\![0,S]\!]} \in G_{\mathrm{loc}}(\mu), \quad \text{where } V = \left\{ Y - 1 + \dfrac{\hat{Y} - a}{1 - a} 1_{\{a<1\}} \right\} 1_{[\![0,\sigma[\![}. \end{cases}$$

Proof. Define V by 5.12. With the convention $0/0 = 0$ we have $\hat{V} = \dfrac{\hat{Y} - 1}{a - 1}$ on $[\![0, \sigma[\![$. Therefore

5.13 $$\tilde{V}_t = \left[(Y(t, \Delta X_t) - 1) 1_{\{\Delta X_t \neq 0\}} - \frac{\hat{Y}_t - a_t}{1 - a_t} 1_{\{\Delta X_t = 0\}} \right] 1_{\{t<\sigma\}}$$

satisfies $\tilde{V} \geq -1$ identically. With the notation II.1.36 we recognize that

5.14 $$H = (\beta \cdot c \cdot \beta) 1_{[\![0,\sigma[\![} \cdot A + C'(V).$$

Now, if $[\![0,S]\!] \subset \Delta$ we have $H_S < \infty$, hence the predictable (stopped) process H^S belongs to $\mathscr{A}^+_{\mathrm{loc}}$. Due to II.1.33 and 4.3, we deduce that 5.12 holds, and the right-hand side of 5.11 defines a P-local martingale. If S' is another stopping time with $[\![0,S']\!] \subset \Delta$, the right-hand sides of 5.11 written for S and for S' obviously coincide on $[\![0,S \wedge S']\!]$, up to a P-evanescent set: the existence and uniqueness of the process N with the claimed properties follow, provided we can write $\Delta = \bigcup_n [\![0,S_n]\!]$ P-a.s., for some sequence (S_n) of stopping times.

But the latter is easy: there is a sequence (σ_n) of stopping times that P-a.s. announces σ (which is predictable: see I.2.16). Then $S_n = \sigma_n \wedge T_n$ does the job. □

From now on to the end of this subsection, we assume that $P' \overset{\mathrm{loc}}{\ll} P$, and Z denotes the density process of P' relative to P, and

5.15 $$R_n = \inf(t\colon Z_t < 1/n),$$

so that by 3.6:

5.16 $$\{Z_- > 0\} \cap]\!]0,\infty[\![= \bigcup_n]\!]0,R_n]\!] \quad \text{up to a } P\text{-evanescent set.}$$

5.17 **Lemma.** *Assume $P' \overset{\mathrm{loc}}{\ll} P$, and define Δ by 5.9.*

a) *The density process Z has the form*

5.18 $$\begin{aligned} Z &= Z_0 + (Z_-\beta)\cdot X^c + Z_-\left(Y - 1 + \frac{\hat{Y} - a}{1 - a} 1_{\{a<1\}}\right) * (\mu - \nu) + Z' \\ &= Z_0 + (Z_-\beta)\cdot X^c + (Z_- V) * (\mu - \nu) + Z' \end{aligned}$$

where V is given by 5.12 and Z' is a P-local martingale with $Z'_0 = 0$ and $\langle Z'^c, X^{i,c}\rangle = 0$ for $i = 1, \dots, d$ and $M^P_\mu(\Delta Z'|\tilde{\mathscr{P}}) = 0$.

b) *$\bigcup_n [\![0,R_n]\!] \subset \Delta$ up to a P-evanescent set.*

Proof. In view of 5.5, Theorem 3.24 gives that β and Y meet 3.28, and in particular $M^P_\mu(\Delta Z|\tilde{\mathscr{P}}) = Z_-(Y - 1)$. Then Theorems 4.11 and 4.20 yield the first formula in 5.18.

Moreover, Theorem 3.24 yields $P'(\sigma < \infty) = 0$. Since σ is predictable, $Z_{\sigma-} = dP'_{\sigma-}/dP_{\sigma-}$ on $\{\sigma < \infty\}$ (see 3.4(iii)), so $Z_{\sigma-} = 0$ P-a.s. on $\{\sigma < \infty\}$: therefore $Z_- 1_{[\![0,\sigma[\![} = Z_-$ and we deduce the last equality in 5.18; we also deduce, using 5.16, that $\bigcup_n [\![0,R_n]\!] \subset [\![0,\sigma[\![$ up to a P-evanescent set.

Now $(1/Z_-)1_{[\![0,R_n]\!]} \le n$, so we deduce from $Z_-\beta \in L^2_{\mathrm{loc}}(X^c)$ and $Z_- V \in G_{\mathrm{loc}}(\mu)$ that $\beta 1_{[\![0,R_n]\!]} \in L^2_{\mathrm{loc}}(X^c)$ and $V 1_{[\![0,R_n]\!]} \in G_{\mathrm{loc}}(\mu)$. Then it follows from 4.3 and II.1.33 (recall that $\tilde{V} \ge -1$: see 5.13) that $H_{R_n \wedge t} < \infty$ P-a.s. for all $t < \infty$. This clearly yields $[\![0,R_n]\!] \subset \bigcup_p [\![0,T_p]\!]$ up to a P-evanescent set, and the last claim follows. □

5.19 **Theorem.** *Assume that $P' \overset{\mathrm{loc}}{\ll} P$ and that all P-local martingales have the representation property relative to X. Then the density process Z satisfies*

5.20 $$Z = Z_0 + (Z_-\beta)\cdot X^c + Z_-\left(Y - 1 + \frac{\hat{Y} - a}{1 - a} 1_{\{a<1\}}\right) * (\mu - \nu).$$

Moreover, if Δ and N are given by 5.9 and 5.11,

5.21 $$Z_t = \begin{cases} Z_0 \exp\left(N_t - \frac{1}{2}(\beta\cdot c\cdot\beta)\cdot A_t\right) \prod_{s\leq t} (1 + \Delta N_s)e^{-\Delta N_s} & \text{if } t\in\Delta \\ 0 & \text{if } t\notin\Delta. \end{cases}$$

Proof. a) 5.20 immediately follows from Lemma 5.17a and from the equivalence (i)$\Leftrightarrow$(ii) in 4.27.

b) We already know that $Z_t = 0$ if $t\notin\Delta$ by 5.17b. Moreover, since $[\![0, R_n]\!] \subset \Delta$, comparing 5.20 and 5.11 yields:

$$Z^{R_n} = Z_0 + Z_- \cdot N^{R_n}.$$

Hence $Z(n) = Z^{R_n}/Z_0$ (with $0/0 = 0$) satisfies $Z(n) = 1 + Z(n)_- \cdot N^{R_n}$ and I.4.61 yields that $Z(n) = \mathscr{E}(N^{R_n})$. Hence $Z^{R_n} = Z_0\mathscr{E}(N^{R_n})$. Since

$$\langle (N^{R_n})^c, (N^{R_n})^c\rangle = (\beta\cdot c\cdot\beta)1_{[\![0,R_n]\!]}\cdot A$$

by 4.7, in view of I.4.64 we recognize that $Z_t^{R_n} = Z_t$ is given by 5.21, for $t \leq R_n$.

It remains to prove that if $t\in\Delta$ and $t > R_n$ for all n, then the right-hand side of 5.21, say $\tilde{Z}_t$, equals 0. Set $R = \lim_n \uparrow R_n$ and $S = \inf(t\colon \tilde{Z}_t = 0 \text{ or } \tilde{Z}_{t-} = 0)$. By construction of $\tilde{Z}$, we have $\tilde{Z} = 0$ on $[\![S, \infty[\![$. Then we have two cases:

1) $R_n = R$ for some n: then $\tilde{Z}_R = \tilde{Z}_{R_n} = Z_{R_n} = Z_R = 0$, hence $S = R$ and $\tilde{Z}_t = 0$ because $t \geq S$;

2) $R_n < R$ for all n: then $\tilde{Z}_{R-} = \lim \tilde{Z}_{R_n} = \lim Z_{R_n} = Z_{R-} = 0$, hence $S = R$ and $\tilde{Z}_t = 0$ again. □

5.22 **Corollary.** *In addition to the assumptions of 5.19, assume that one of the following three statements is met:*

(i) *either $Z_t > 0$ P-a.s. for all $t\in\mathbb{R}_+$,*

(ii) *or $T \leq \sigma$ and the generalized increasing process H does not jump to infinity* (see 5.8),

(iii) *or for P-almost all ω there is a t with $(\omega, t)\in\Delta$ and $\Delta N_t(\omega) = -1$.*

Then we have

5.23 $$Z_t = Z_0 \exp\left(N_t - \frac{1}{2}(\beta\cdot c\cdot\beta)\cdot A_t\right) \prod_{s\leq t} (1 + \Delta N_s)e^{-\Delta N_s}.$$

Proof. The claim is obvious under (i) or (iii). Suppose now (ii), that is $H_{T-} = \infty$ on $\{T < \infty\}$ and $T \leq \sigma$. Therefore $T_n < T$ for all n on $\{T < \infty\}$, and $\Delta = [\![0, T[\![$. Since $\bigcup_n [\![0, R_n]\!] \subset \Delta$, we have $R_n < T$ on $\{T < \infty\}$ and so $\lim_{t\uparrow T, t<T} Z_t = 0$ on $\{T < \infty\}$. Therefore 5.23 readily follows from 5.21. □

We give below an interesting consequence. We consider an "initial σ-field" $\mathscr{H}$ (a sub-σ-field of $\mathscr{F}_0$) and P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$.

5.24 Theorem. *Assume that $P' \overset{\text{loc}}{\ll} P$.*

a) *If all P-local martingales have the representation property relative to X, then all P'-local martingales also have the representation property relative to X* (on the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P')$ for the latter, of course!)

b) *If P is extremal in $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ and if $\mathscr{F} = \mathscr{F}_{\infty-}$, then P' is extremal in $\mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$.*

Proof. In view of 4.29, (a) is the particular case of (b) obtained when $\mathscr{H} = \mathscr{F}_0$.

Assume that P is extremal in $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$, and let $\tilde{P}' \in \mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$ with $\tilde{P}' \ll P'$. Then $\tilde{P}' \overset{\text{loc}}{\ll} P$ and we call $\tilde{Z}$ the density process of $\tilde{P}'$ relative to P.

$E_P(\tilde{Z}_0 | \mathscr{H}) = E_P(Z_0 | \mathscr{H})$ is the density dP'_H/dP_H: then we deduce from the second property in 4.29i that $\tilde{Z}_0 = Z_0$ P-a.s. It remains to apply Theorem 5.19: since $\tilde{Z}_0 = Z_0$ P-a.s., $\tilde{Z}_t$ and Z_t are both given P-a.s. by the same formula 5.21. Hence $\tilde{Z}_t = Z_t$ P-a.s., hence $\tilde{P}' = P'$ on each $\mathscr{F}_t$. Since $\mathscr{F} = \mathscr{F}_{\infty-}$ we deduce that $\tilde{P}' = P'$. Therefore P' satisfies 4.29iv, and the extremality of P' in $\mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$ follows. □

5.25 Remark. Lemma 5.17 and Theorem 5.19 contain a very remarquable fact, namely that $Z_{\sigma-} = 0$ on $\{\sigma < \infty\}$. Indeed, if we define Z by 5.21 there is absolutely no reason why $Z_{\sigma-}$ should be equal to 0 for $\sigma < \infty$ (except on the set $\{H_{\sigma-} = \infty\}$): this property comes in fact from the "implicit" assumption on the characteristics (B', C', ν') that they are the characteristics of X for a measure P' with $P' \overset{\text{loc}}{\ll} P$. □

§ 5b. P' Has the Local Uniqueness Property

Here again we begin with some general results. β and Y are as in 5.4, and σ, H, Δ are defined in 5.6, 5.7 and 5.9, and N is the process introduced in 5.10. We set

5.26 $$\tilde{Z}_t = \begin{cases} \tilde{Z}_0 \exp\left(N_t - \frac{1}{2}(\beta \cdot c \cdot \beta) \cdot A_t\right) \prod_{s \le t} (1 + \Delta N_s) e^{-\Delta N_s} & \text{if } t \in \Delta \\ 0 & \text{if } t \notin \Delta \end{cases}$$

where $\tilde{Z}_0$ is an $\mathscr{F}_0$-measurable random variable, with $\tilde{Z}_0 \ge 0$ and $E(\tilde{Z}_0) = 1$.

Since $\Delta N = \tilde{V} \ge -1$ (see 5.13), we have $\tilde{Z} \ge 0$. Moreover if S is a stopping time with $[\![0, S]\!] \subset \Delta$, one shows like in the proof of Theorem 5.19 that $\tilde{Z}^S = \tilde{Z}_0 \mathscr{E}(N^S)$, and thus $\tilde{Z}^S$ is a P-local martingale.

Moreover, we assume that $B'(\omega)$, $C'(\omega)$, $\nu'(\omega)$ are given by 5.5 *for all ω* (and not only P'-a.s.). That is an essential hypothesis for the next lemma, in which we construct another measure $\tilde{P}'$ which has a-priori nothing to do with P'.

5.27 **Lemma.** *Let S be a stopping time such that $[\![0,S]\!] \subset \Delta$, and assume further that $\tilde{Z}^S$ is a P-uniformly integrable martingale. Then, under the probability measure $\tilde{P}'(d\omega) = P(d\omega)\tilde{Z}_S(\omega)$, X^S is a semimartingale with characteristics (B'^S, C'^S, v'^S)* (see before 2.37 for the notation v'^S).

Proof. As seen before, $\tilde{Z}^S = \tilde{Z}_0 \mathscr{E}(N^S)$, and so 5.11 yields

5.28 $$\tilde{Z}^S = \tilde{Z}_0 + (\tilde{Z}^S_- \beta 1_{[\![0,S]\!]}) \cdot X^c + (\tilde{Z}^S_- V 1_{[\![0,S]\!]}) * (\mu - v)$$

with V as in 5.12. Then 4.6 implies

$$\langle (\tilde{Z}^S)^c, X^{i,c} \rangle = \left(\tilde{Z}^S_- \left(\sum_{j \le d} c^{ij} \beta^j \right) 1_{[\![0,S]\!]} \right) \cdot A.$$

Moreover 5.13 gives $\tilde{Z}^S = \tilde{Z}^S_-(Y-1)1_{[\![0,S]\!]}$ M^P_μ-a.s., and the latter is $\tilde{\mathscr{P}}$-measurable. Hence $M^P_\mu(\Delta \tilde{Z}^S | \tilde{\mathscr{P}}) = \tilde{Z}^S_-(Y-1)1_{[\![0,S]\!]}$: it easily follows that 3.28 is fulfilled by X^S and $\tilde{Z}^S$ (relatively to (B^S, C^S, v^S) and (B'^S, C'^S, v'^S)). Since $\tilde{Z}^S$ is the density process of $\tilde{P}'$ relative to P, the result is deduced from Theorem 3.24. □

Ultimately we will assume local uniqueness (see 2.37) for the martingale problem to which P' is a solution. So we place ourselves in the proper setting: let $\mathscr{H} \subset \mathscr{F}_0$ be an initial σ-field, and P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$, and we suppose that **F** is generated by X and $\mathscr{H}$ (see 2.12). Recall that strict stopping times have been defined in 2.35.

5.29 *Hypothesis.* $\Delta = \bigcup_n [\![0,\sigma_n]\!]$ up to a P-evanescent set, where (σ_n) is a sequence of *strict stopping times.* □

Observe that Δ is always of the form $\bigcup_n [\![0,\sigma_n]\!]$ for a sequence (σ_n) of stopping times: take for instance $\sigma_n = \sigma'_n \wedge T_n$, where T_n is like in 5.9 and (σ'_n) is a sequence that announces the positive predictable time σ. However, those σ_n are not necessarily strict stopping times.

Nevertheless, *5.29 is fulfilled* when *H does not jump to infinity* (5.8) *and* $\sigma \ge \lim_n T_n (= \inf(t\colon H_t = \infty))$, because then each T_n is predictable (use I.2.13) and positive, so is strict by 2.36, and $\Delta = \bigcup_n [\![0, T_n]\!]$.

5.30 **Lemma.** *Assume 2.12 and 5.29. There is a sequence (S_n) of strict stopping times, such that $\Delta = \bigcup_n [\![0,S_n]\!]$ P-a.s., and that $\tilde{Z}^{S_n}$ is a P-uniformly integrable martingale for all n.*

Proof. a) We assume first that $\Delta = \Omega \times \mathbb{R}_+$, hence $\sigma \equiv \infty$ and H is finite-valued (and so is an "ordinary" right-continuous increasing process).

V is given by 5.12 and we associate V' and V'' with it by II.1.35. Set

$$K = (\beta \cdot c \cdot \beta) \cdot A + C(V') + \bar{C}(V'')$$
$$K' = (1 + \tilde{Z}_- + (\tilde{Z}_-)^2) \cdot K.$$

In view of 5.14 and of II.1.33, we see that K, and hence K' also, are predictable, and belong to $\mathscr{A}_{\mathrm{loc}}$.

Set $S_n = \inf(t\colon K'_t \geq n)$. Then S_n is a predictable time (by I.2.13), and $S_n > 0$, and $S_n \uparrow \infty$ as $n \uparrow \infty$. It remains to prove that $\tilde{Z}^{S_n}$ is P-uniformly integrable.

For simplicity, set $S = S_n$. By 5.28, $\tilde{Z}^S$ is a P-local martingale, as well as $\tilde{Z}^{S-} := \tilde{Z}_0 + 1_{[\![0,S[\![} \cdot \tilde{Z}^S$, S being predictable. 5.28 yields

5.31
$$\begin{aligned}\tilde{Z}^{S-} &= Z_0 + (\tilde{Z}_- \beta 1_{[\![0,S[\![}) \cdot X^c + (\tilde{Z}_- V' 1_{[\![0,S[\![}) * (\mu - \nu) \\ &\quad + (\tilde{Z}_- V'' 1_{[\![0,S[\![}) * (\mu - \nu)\end{aligned}$$

Moreover, $K'_{S-} \leq n$ by construction, thus the definition of K' yields

$$\begin{aligned}(\tilde{Z}_-^2(\beta \cdot c \cdot \beta) 1_{[\![0,S[\![}) \cdot A_\infty &\leq n \\ C(\tilde{Z}_- V' 1_{[\![0,S[\![})_\infty = \tilde{Z}_-^2 \cdot C(V')_{S-} &\leq n \\ \bar{C}(\tilde{Z}_- V'' 1_{[\![0,S[\![})_\infty = \tilde{Z}_- \cdot \bar{C}(V'')_{S-} &\leq n.\end{aligned}$$

Then 4.5d implies $(\tilde{Z}_- \beta 1_{[\![0,S[\![}) \cdot X^c \in \mathscr{H}^2$, and II.1.33a implies $(\tilde{Z}_- V' 1_{[\![0,S[\![}) * (\mu - \nu) \in \mathscr{H}^2$, and II.1.33b yields $(\tilde{Z}_- V'' 1_{[\![0,S[\![}) * (\mu - \nu) \in \mathscr{A}$. Since $\tilde{Z}_0$ is P-integrable, we deduce from 5.31 that $\tilde{Z}^{S-}$ is P-uniformly integrable.

Now, $\tilde{Z}^S = \tilde{Z}^{S-} + \Delta \tilde{Z}_S 1_{[\![S,\infty[\![}$, so it remains to prove that $\Delta \tilde{Z}_S 1_{\{S<\infty\}}$ is P-integrable. We have $\Delta \tilde{Z}_S = \tilde{Z}_{S-} \Delta N_S$ on $\{S < \infty\}$, and $\tilde{Z}_{S-}$ is P-integrable from what precedes, so it suffices to prove that $E(|\Delta N_S| | \mathscr{F}_{S-}) 1_{\{S<\infty\}}$ is bounded. But, recalling 5.13 for $\hat{V} = \Delta N$, and that $a \leq 1$, $\hat{Y} \leq 1$, we get on $\{S < \infty\}$:

$$|\Delta N_S| \leq 1 + Y(S, \Delta X_S) 1_{\{\Delta X_S \neq 0\}} + \frac{2}{1 - a_S} 1_{\{\Delta X_S = 0\}}$$

and since S is predictable,

$$E(|\Delta N_S| | \mathscr{F}_{S-}) \leq 1 + \hat{Y}_S + \frac{2}{1 - a_S} P(\Delta X_S = 0 \mid \mathscr{F}_{S-}) = \hat{Y}_S + 3 \leq 4.$$

b) We turn now to the general case. Fix $p \in \mathbb{N}^*$, and stop all processes at time σ_p: then (a) applies and gives a sequence $(S_{n,p})_{n\geq 1}$ of positive predictable, and thus strict, stopping times, with $\Omega \times \mathbb{R}_+ = \bigcup_n [\![0, S_{n,p}]\!]$ and $(\tilde{Z}^{\sigma_p})^{S_{n,p}}$ is a P-uniformly integrable martingale.

Then $S'_{n,p} = S_{n,p} \wedge \sigma_p$ is again a strict stopping time, and $\Delta = \bigcup_{n,p} [\![0, S'_{n,p}]\!]$, and $\tilde{Z}^{S'_{n,p}}$ is a P-uniformly integrable martingale. Relabelling the sequence $(S'_{n,p})_{n,p \in \mathbb{N}^*}$ gives the result. $\square$

Now we state our main result, to be compared to Theorem 5.19: here we assume local uniqueness for the problem $\mathfrak{s}(\mathscr{H}, X | P'_H; B', C', \nu')$ whose P' is a solution (and hence the unique solution); in 5.19 we assumed the representation property for all P-local martingales, a property which in practice can only be checked when uniqueness holds for $\mathfrak{s}(\mathscr{H}, X | P_H; B, C, \nu)$.

5.32 **Theorem.** *Assume that* (B', C', ν') *is given everywhere by 5.5, where* (β, Y) *meets 5.4. Assume that* $\mathbf{F}$ *is generated by* X *and* $\mathscr{H}$ (see 2.12) *and that 5.29 holds, and that local uniqueness holds for the martingale problem* $\mathfrak{s}(\mathscr{H}, X | P'_H; B', C', \nu')$, *with* P' *as its unique solution.*

Then if $P' \overset{\text{loc}}{\ll} P$ *the density process* Z *of* P' *relative to* P *is given by 5.21* (or 5.23 if either one of the additional assumptions in 5.22 holds), *where* Z_0 *is* (P-a. s. equal to) *the Radon-Nikodym derivative* $Z_H = dP'_H/dP_H$.

5.33 **Remark.** That (B', C', ν') is everywhere given by 5.5 is a crucial hypothesis (it has been used in Lemma 5.27, which is fundamental for proving 5.32).

Very often the measure P' is given first, and there are usually many versions of the P'-characteristics of X: then one has to choose *a version that satisfies 5.5*, and local uniqueness ought to hold *relatively to this choice of* (B', C', ν').

We illustrate this point with an example. $\Omega = \{1, 2\}$ with $P(1) = P(2) = \frac{1}{2}$ and $P'(1) = 1$, so $P' \ll P$. Let X be

$$X_t(1) = 0, \quad X_t(2) = (t-1)^+$$

and $\mathbf{F}$ be given by 2.12 ($\mathscr{H}$ = trivial σ-field). Then $B = X$, $C = 0$, $\nu = 0$ are the P-characteristics of X, and the density process is

$$Z_t(\omega) = 1 \text{ for } t < 1; \quad Z_t(1) = 2 \text{ and } Z_t(2) = 0 \text{ for } t \geq 1.$$

Hence *Z is not of the form 5.21* (because $N = 0$, since $\mu = 0$ and $X^c = 0$). However a natural set of characteristics for X under P' is $B' = 0$, $C' = 0$, $\nu' = 0$, for which local uniqueness obviously holds! but of course this set of characteristics does not meet 5.5.

Indeed, the only set of characteristics that meets 5.5 is $(B', C', \nu') = (B, C, \nu)$, and *all* probability measures on Ω are solutions of the corresponding martingale problem. $\square$

Proof of 5.32. We define $\tilde{Z}$ by 5.26, with $\tilde{Z}_0 = Z_H := dP'_H/dP_H$, and we consider the sequence (S_n) constructed in Lemma 5.30. Then 5.27 implies that $P'^n(d\omega) = P(d\omega)\tilde{Z}_{S_n}(\omega)$ is a solution of the stopped problem $\mathfrak{s}(\mathscr{H}, X^{S_n} | P'_H; B'^{S_n}, C'^{S_n}, \nu'^{S_n})$ (note that $P_0'^n = Z_H \cdot P_0$, hence $P_H'^n = P'_H$ is obvious).

Now we apply local uniqueness: S_n is a strict stopping time, hence $P'^n = P'$ on $\mathscr{F}^0_{S_n}$, and a-fortiori on $\mathscr{F}_{(S_n)-}$, in restriction to the set $\{S_n > 0\}$. So $P_t'^n = P'_t$ on the set $\{S_n > t\}$, and we deduce that $\tilde{Z} = Z$ on $\bigcup_n [\![0, S_n[\![$. Because of 5.17 and 5.26, we also have $\tilde{Z} = Z = 0$ on Δ^c. It remains to see what happens at time S, on the set $\bigcup_n \{S_n = S < \infty\}$: we have $Z_S = 0$ on this set (because Z is càd), while $\tilde{Z}_S \geq 0$, and so

$$\tilde{Z}^{S_n} - Z^{S_n} = \tilde{Z}_S 1_{\{S = S_n < \infty\}} 1_{[\![S, \infty[\![} \geq 0.$$

Therefore for each $t \in \mathbb{R}_+$ the process $M(n, t) = \tilde{Z}^{S_n \wedge t} - Z^{S_n \wedge t}$, which is a P-martingale, is also nonnegative and satisfies $E_P(M(n, t)_0) = 0$. Hence $M(n, t) =$

0 P-a.s., and so $\tilde{Z}_S = 0$ P-a.s. on $\{S = S_n \leq t\}$, and so $\tilde{Z}_S = 0$ P-a.s. on $\bigcup_n \{S = S_n < \infty\}$. □

Our last result anticipates on the next chapter, but it comes as a natural corollary at this point.

5.34 **Theorem.** *Assume that (B', C', ν') is given everywhere by 5.5, and that $\mathbf{F}$ is generated by X and $\mathscr{H}$, and that 5.29 holds, and finally that local uniqueness holds for the martingale problem $\mathfrak{s}(\mathscr{H}, X | P'_H; B', C', \nu')$, with P' as its unique solution.*

Then if $P'_H \ll P_H$ and if $P'(H_t < \infty) = 1$ for all $t \in \mathbb{R}_+$ we have $P' \overset{\text{loc}}{\ll} P$ and the density process Z of P' relative to P is given by 5.21, with $Z_0 = dP'_H/dP_H$.

Proof. Set $Z_H = dP'_H/dP_H$ and define $\tilde{Z}$ by 5.26 with $\tilde{Z}_0 = Z_H$. We can reproduce the proof of 5.32 up to the point $P'^n = P'$ on $\{0 < S_n\} \cap \mathscr{F}_{(S_n)-}$.

We have $P'(S_n > t) \uparrow 1$ as $n \uparrow \infty$ for all $t \in \mathbb{R}_+$, by hypothesis. If $A \in \mathscr{F}_t$, then $A \cap \{S_n > t\} \in \mathscr{F}_{S_n-}$, and if moreover $P(A) = 0$ we then obtain

$$P'(A) = \lim_n P'(A \cap \{S_n > t\}) = \lim_n P'^n(A \cap \{S_n > t\}) = 0$$

because $P'^n \ll P$. We then deduce that $P'_t \ll P_t$, and so $P' \overset{\text{loc}}{\ll} P$.

Finally, the last claim follows from 5.32. □

§ 5c. Examples

In this subsection we give some examples of "explicit" computations of the density process, and various other applications of what precedes.

Except when explicitely mentionned otherwise, we will place ourselves in the *canonical setting* 2.13.

1. Processes with independent increments. We assume here that X (the canonical process) is a PII under P and under P', with respective characteristics (B, C, ν) and (B', C', ν'). Those are deterministic, so we will assume that 5.5 holds with *deterministic* term β and Y. Then the process H is deterministic, as well as σ and Δ (see 5.6 and 5.9): in particular, Hypothesis 5.29 is obviously met.

5.35 **Theorem.** *In addition to the above hypotheses, assume that $P' \overset{\text{loc}}{\ll} P$. Then the density process Z of P' relative to P is given by 5.21, and the process $N = (N_t)_{t \in \Delta}$ is a process with independent increments on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, indexed by the time interval Δ.*

Proof. The first claim follows either from Theorem 5.19 (using 4.34) or from Theorem 5.32 (using 2.42).

For the second claim, we need to prove that if $s \in \Delta$, the stopped process N^s is a PII. Now, the explicit form 5.11 gives for $r \leq t$:

$$N_t^s - N_r^s = \beta 1_{(r,t]} \cdot X_s^c + \left(Y - 1 + \frac{\hat{Y} - a}{1 - a} 1_{\{a<1\}} \right) 1_{(r,t]} * (\mu - \nu)_s$$

and the integrands are deterministic: so $N_t^s - N_r^s$ is clearly measurable with respect to the σ-field generated by $X_u^c - X_r^c$ for $u \geq r$ and by $\mu((r,u] \times A)$ for $u \geq r$, $A \in \mathbb{R}^d$, which itself is contained in $\sigma(X_u - X_r: u \geq r)$ (actually, it is equal to this σ-field). Then the PII property of X implies that this σ-field is P-independent from $\mathscr{F}_r$, and so N^s is a PII on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. □

2. *Generalized diffusion processes.* For simplicity we consider only the 1-dimensional case ($d = 1$). We assume again the canonical setting, and also the following: under P, X is a standard Wiener process, so $B = 0$, $C_t = t$, $\nu = 0$. Under P', X is a semimartingale with characteristics

5.36 $$B_t' = \int_0^t b_s' \, ds, \quad C_t' = t, \quad \nu' = 0$$

and with $P'(X_0 = 0) = 1$. Note that we can write

5.37 $X = B' + W$, with W a standard Wiener process on $(\Omega, \mathscr{F}, \mathbf{F}, P')$

(because $W = X - B'$ is a continuous local martingale with $W_0 = 0$ and with quadratic variation $\langle W, W \rangle_t = C_t' = t$). So X is called a *generalized diffusion process* under P': it is indeed as a diffusion process (with diffusion coefficient $c = 1$ here), except that $b_s(\omega)$ is arbitrary (predictable) instead of being of the form $b_s(\omega) = b(s, X_{s-}(\omega))$ as for "ordinary" diffusions.

We will see other results similar to the following one, in the next chapter.

5.38 **Theorem.** *Assume all the above, and set* $T = \inf(t: \int_0^t (b_s')^2 \, ds = \infty)$.

a) *If* $P' \overset{\text{loc}}{\ll} P$, *then* $P'(T = \infty) = 1$ *and the density process* Z *of* P' *relative to* P *is*

5.39 $$Z_t = \begin{cases} \exp\left[\int_0^t b_s' \, dX_s - \frac{1}{2} \int_0^t (b_s')^2 \, ds\right] & \text{if } t < T \\ 0 & \text{if } t \geq T. \end{cases}$$

b) *We have* $P \overset{\text{loc}}{\ll} P'$ *if and only if* $P(T = \infty) = 1$, *in which case the density process* Z' *of* P *relative to* P' *is*

5.40 $$Z_t' = \begin{cases} \exp\left[-\int_0^t b_s' \, dW_s - \frac{1}{2} \int_0^t (b_s')^2 \, ds\right] & \\ \quad = \exp\left[-\int_0^t b_s' \, dX_s + \frac{1}{2} \int_0^t (b_s')^2 \, ds\right] & \text{if } t < T \\ 0 & \text{if } t \geq T. \end{cases}$$

c) *If* $P'(T = \infty) = P(T = \infty) = 1$, *then* P *and* P' *are locally equivalent* (i.e. $P \overset{\text{loc}}{\ll} P'$ and $P' \overset{\text{loc}}{\ll} P$).

d) *If $T = \infty$ identically, then local uniqueness* (and of course uniqueness) *holds for the martingale problem* $\mathfrak{s}(\mathscr{H}, X|\varepsilon_0; B', C', \nu')$.

A priori, the stochastic integrals in 5.39 (resp. 5.40) are relative to P (resp. P'); however, one could prove that there always exists a version of those stochastic integrals that is valid for both P and P'. Of course, if the assumption of (c) is in force, we have 5.39 and 5.40, and we recognize that $ZZ' = 1$, as it should be.

As for existence of a solution to $\mathfrak{s}(\mathscr{H}, X|\varepsilon_0; B', C', \nu')$, it is assumed by hypothesis in the previous theorem. If we do not know beforehand that this problem has a solution P', we can prove the following: when $T = \infty$ identically, the problem admits a solution P'_t on each σ-field $\mathscr{F}_t$ and P'_t coincides with P'_s on $\mathscr{F}_s$ if $s \le t$. Using a limit theorem technique (as in the proof of X.1.4 of Chapter X: this explicitely uses the canonical setting) one could show that the family (P'_t) uniquely extends to a probability measure P' on $(\Omega, \mathscr{F})$, which thus is the unique solution to the problem $\mathfrak{s}(\mathscr{H}, X|\varepsilon_0; B', C', \nu')$.

Proof. a) We will apply Theorem 5.19, using the fact that all P-local martingales have the representation property relative to X (see 4.33): we have 5.5 with $\beta = b'$, so $H_t = \int_0^t (b'_s)^2\,ds$. Hence 5.39 reduces to 5.21. Moreover $P'(T \le t) = E_P(Z_t 1_{\{T \le t\}}) = 0$ by 5.39, hence $P'(T < \infty) = 0$.

b) Now we "exchange" P and P', so 5.5 holds with $\beta' = -b'$ and $H'_t = \int_0^t (b'_s)^2\,ds$ again. Moreover local uniqueness holds for the martingale problem $\mathfrak{s}(\mathscr{H}, X|\varepsilon_0; B, C, \nu)$ (see 2.41 or 2.42), so the sufficient condition follows from 5.34. Conversely, assume $P \overset{\text{loc}}{\ll} P'$. Then 5.40 follows from 5.32; moreover $P(T < \infty) = 0$ is then proved as in (a), and this gives the necessary condition.

c) Assume $P'(T = \infty) = P(T = \infty) = 1$. Then $P \overset{\text{loc}}{\ll} P'$ by (b), and Z' is given by 5.40. Since $P'(T < \infty) = 0$, we deduce that $P'(Z'_t = 0) = 0$, so $P' \overset{\text{loc}}{\ll} P$ follows.

d) Let S be a strict stopping time, and let Q be a solution to the stopped martingale problem $\mathfrak{s}(\mathscr{H}, X^S|\varepsilon_0; B'^S, C'^S, \nu'^S)$. Since $T = \infty$ identically, one proves exactly as in Theorem 5.34 that $P_{S \wedge t} \ll Q_{S \wedge t}$ and the density process $\tilde{Z}'^S$ of P relative to Q, on the stochastic interval $[\![0, S]\!]$, is again given by 5.40. Using again that $T = \infty$, we deduce that $Z'^S_t > 0$ Q-a.s. on $\{t \le S\}$: so $P_{S \wedge t} \sim Q_{S \wedge t}$ and the density process of Q with respect to P is thus Z, as given by 5.39, on the interval $[\![0, S]\!]$; therefore we deduce that $Q_S = P'_S$, and the claim is proved. □

3. *Point processes and multivariate point processes.* Exactly as in Section 4, it is obvious that all the results of the present section remain valid, mutatis mutandis, if we replace the basic semimartingale X by a random measure μ on $\mathbb{R}_+ \times E$, with $(E, \mathscr{E})$ an auxiliary Blackwell space: we denote by ν and ν' its compensators with respect to P and P'; then only Y shows up in 5.4, and 5.5 reads simply as $\nu' = Y \cdot \nu$ P'-a.s., and σ is not changed, and H is given by

5.41 $$H = (1 - \sqrt{Y})^2 1_{[\![0,\sigma[\![} * \nu + \sum_{s \le \cdot} (\sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s})^2 1_{\{s < \sigma\}}$$

and Δ is still defined by 5.9, and 5.11 reads

5.42 $$N^S = \left(Y - 1 + \frac{\hat{Y} - a}{1 - a} 1_{\{a<1\}}\right) 1_{[\![0,S]\!]} * (\mu - \nu).$$

Then it readily follows from 5.19 and 4.37 that:

5.43 **Theorem.** *Assume that μ is an E-valued multivariate point process and that 1.25 holds. If $P' \overset{loc}{\ll} P$, the density process Z of P' relative to P is given by 5.21, where N is defined in 5.42.*

In particular, this applies for point processes (when E reduces to one point). So let us assume that X is a point process with compensators A and A' under P and P', and that

5.44 $$A' = y \cdot A, \quad y \text{ nonnegative predictable.}$$

Then the formula giving Z has a particularly pleasant form when A is continuous (i.e., X is quasi-left continuous):

5.45 **Theorem.** *Assume 1.20 and 5.44 and that A is continuous, and let $(T_n)_{n\geq 1}$ be the successive jump times of X. Then if $P' \overset{loc}{\ll} P$ the density process Z of P' relative to P is given by*

5.46 $$Z_t = \begin{cases} e^{(y-1)\cdot A_t} \prod_{n:T_n \leq t} y_{T_n} & \text{if } t < T := \inf(t\colon y \cdot A_t = \infty) \\ 0 & \text{if } t \geq T. \end{cases}$$

6. Integrals of Vector-Valued Processes and σ-martingales

So far we have constructed the stochastic integral of a predictable process H w.r.t. a 1-dimensional semimartingale X, when H is locally bounded: see Chapter I, §4d. Also, in § 4a of the present chapter we constructed stochastic integrals of a d-dimensional integrands H w.r.t. a d-dimensional continuous local martingale X. Here we wish to do the same when X is a d-dimensional semimartingale, in such a way that we obtain the "most general" integrands.

Let us consider a d-dimensional semimartingale $X = (X^i)_{1\leq i\leq d}$ on a stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$. For any d-dimensional predictable process $H = (H^i)_{1\leq i\leq d}$ which is *locally bounded*, we can define the stochastic integral process $H \cdot X$ by

6.1 $$H \cdot X = \sum_{i=1}^{d} H^i \cdot X^i,$$

where obviously each integral $H^i \cdot X^i$ makes sense. When extending this integral to more general integrands H, we encounter two additional difficulties: first, even in the 1–dimensional case, there are problems due to the non–boundedness of H. Second, we may have "interferences" between components: to give a trivial example let $d = 2$ and $X^1 = X^2$; then any predictable process of the form $H = (H^1, -H^1)$ should lead to a null integral process, although $H^1 \cdot X^1$ might not be defined.

Our construction is made in three steps: first for X being a locally square-integrable martingale, second when X is of locally finite variation, third for a general semimartingale X. We also end this Section with a brief account on the notion of σ–martingales, a notion which (slightly) generalizes local martingales.

§ 6a. Stochastic Integrals with Respect to a Multi-Dimensional Locally Square-integrable Martingale

This paragraph is very similar to § 4a. Suppose that X is a d–dimensional locally square-integrale martingale, that is, each component X^i is in $\mathscr{H}^2_{\text{loc}}$. We consider the angle brackets $C^{ij} = \langle X^i, X^j \rangle$ and a factorization 4.2 for them, that is

6.2 $$C^{ij} = c^{ij} \cdot F$$

where $c = (c^{ij})_{1 \le i,j \le d}$ is predictable, taking values in the set of all symmetric nonegative $d \times d$ matrices, and F is an increasing predictable process. By analogy with 4.3, we set

6.3 $L^2(X)$ (resp. $L^2_{\text{loc}}(X)$) is the set of all d–dimensional predictable processes H such that the increasing process $(\sum_{i,j} H^i c^{ij} H^j) \cdot F$ is integrable (resp. locally integrable).

Clearly this set does not depend on the choice of the pair (F, c) satisfying 6.2. Then, similarly to Theorem 4.5, we have:

6.4 **Theorem.** *Let X be a d–dimensional locally square-integrable martingale, and $H \in L^2_{\text{loc}}(X)$.*

a) *If $H(n) = H 1_{\{|H| \le n\}}$, then $H(n) \cdot X$* (defined by 6.1)) *converges in measure, uniformly on every compact inteval, to a limit denoted by $H \cdot X$.*

b) *$H \cdot X$ is also characterized as follows: it is the unique* (up to evanescence) *locally square-integrable martingale, null at* 0, *such that*

6.5 $$\langle H \cdot Y \rangle = \left(\sum_{i \le d} H^i c^{Yi} \right) \cdot F$$

for all $Y \in \mathscr{H}^2_{\text{loc}}$, where c^{Yi} is a predictable process such that $\langle Y, X^i \rangle = c^{Yi} \cdot F$.

c) *If* $H, K \in L^2_{\text{loc}}(X)$, *then*

6.6
$$\langle H \cdot X, K \cdot X\rangle = (H \cdot c \cdot K) \cdot F.$$

d) $H \cdot X$ *always belongs to* $\mathscr{H}^2_{\text{loc}}(X)$, *and it belongs to* $\mathscr{H}^2$ *if and only if* $H \in L^2(X)$.

e) $L^2_{\text{loc}}(X)$ *is stable by localization and is a linear space and* $H \mapsto H \cdot H \cdot X$ *is linear* (up to evanescence), *and*

6.7 for all stopping times T, $H1_{[\![0,T]\!]} \in L^2_{\text{loc}}(X)$ and $(H1_{[\![0,T]\!]}) \cdot X = (H \cdot X)^T$;

6.8 if $K \in L^2_{\text{loc}}(H \cdot X)$, then $KH \in L^2_{\text{loc}}(X)$ and $(KH) \cdot X = K \cdot (H \cdot X)$.

The proof is exactly the same as for Theorem 4.5 and is thus omitted. In the sequel, we need a few more properties, stated in the next proposition

6.9 **Proposition.** a) *Let* X *be a* d*–dimensional locally square-integrable martingale, and* $H \in L^2_{\text{loc}}(X)$. *Then for any sequence* (D_n) *of predictable sets increasing to* $\Omega \times \mathbb{R}_+$, *we have* $H1_{D_n} \in L^2_{\text{loc}}(X)$ *and* $\sup_{s \leq t} |(H1_{D_n}) \cdot X_s - H \cdot X_s| \to 0$ *in measure for all* $t \in \mathbb{R}_+$, *and this convergence also holds in* L^2 *if further* $H \in L^2(X)$. *Moreover we have*

6.10
$$\Delta(H \cdot X) = \sum_{i \leq d} H^i \Delta X^i \quad \text{(written as } H \cdot \Delta X).$$

b) *If* Y *is another* d*–dimensional locally square-integrable martingale and* $a \in \mathbb{R}$ *and* $H \in L^2_{\text{loc}}(X) \cap L^2_{\text{loc}}(Y)$, *then* $H \in L^2_{\text{loc}}(X + aY)$ *and* $H \cdot (X + aY) = H \cdot X + aH \cdot Y$.

Proof. That $H1_{D_n} \in L^2_{\text{loc}}(X)$ is obvious. Set $M^n = (H1_{D_n}) \cdot X - H \cdot X$. We have $M^n \in \mathscr{H}^2_{\text{loc}}$, and 6.6 yields $\langle M^n, M^n\rangle_t = (H \cdot c \cdot H)1_{D \setminus D_n} \cdot F_t$, which goes to 0 in measure, and even in L^1 when $H \in L^2(X)$: so we get the desired convergence of M^n to 0. Finally if $D_n = \{|H| \leq n\}$ and $H^n = H1_{D_n}$ then H^n is locally bounded, hence satisfies 6.10; passing to the limit in n gives 6.10 for H.

For (b), we set $D_n = \{|H| \leq n\}$, which increases to $\Omega \times \mathbb{R}_+$, and $H^n = H1_{D_n}$. We can always choose factorizations 6.2 for X and Y with the same process F, so that $H \in L^2_{\text{loc}}(X + aY)$ is obvious. Further (a) yields that $H^n \cdot X$ and $H^n \cdot Y$ and $H^n \cdot (X + aY)$ converge in measure to $H \cdot X$ and $H \cdot Y$ and $H \cdot (X + aY)$ respectively, while obviously $H^n \cdot X + H^n \cdot Y = H^n \cdot (X + aY)$ because H^n is bounded, hence the last result. $\square$

§ 6b. Integrals with Respect to a Multi-Dimensional Process of Locally Finite Variation

Another, much simpler, situation is when X has components in $\mathscr{V}$. In this case, and similarly to 6.2, we can find an increasing optional process F and an optional $\mathbb{R}^d$–valued process $a = (a^i)_{1\leq i\leq d}$ such that

6.11 $$X^i = a^i \cdot F,$$

and further F and a may be chosen predictable if X is so. Then we set

6.12 $L^0(X)$ is the set of all d–dimensional predictable processes H such that the increasing process $|\sum_i H^i a^i| \cdot F$ is finite-valued. We then put

6.13 $$H \cdot X = (\sum_i H^i a^i) \cdot F.$$

Again, neither the set $L^0(X)$ nor the integral process $H \cdot X$ depend on the choice of the pair (A, a) satisfying 6.11. Moreover, 6.13 and 6.1 coincide when H is locally bounded. Note that even in this simple case we may have $H \in L^0(X)$ without each component H^i belonging to $L^0(X^i)$. The following set of results is trivial:

6.14 **Theorem.** *Let X be a d–dimensional process with components in $\mathscr{V}$. Then if $H \in L^0(X)$, we have:*

a) *For any sequence (D_n) of predictable sets increasing to $\Omega \times \mathbb{R}_+$, we have $H 1_{D_n} \in L^0(X)$ and $\sup_{s\leq t} |(H 1_{D_n}) \cdot X_s - H \cdot X_s| \to 0$ in measure for all $t \in \mathbb{R}_+$.*

b) *$H \cdot X$ belongs to $\mathscr{V}$, and is predictable when X is predictable. Moreover the variation process of $H \cdot X$ is $\mathrm{Var}(H \cdot X) = |\sum_{j\leq d} H^j a^j| \cdot F$.*

c) *$L^0(X)$ is stable by localization and is a linear space, and $H \mapsto H \cdot X$ is linear.*

d) *If Y is another d–dimensional process with components in $\mathscr{V}$ and $a \in \mathbb{R}$ and $H \in L^0(X) \cap L^0(Y)$, then $H \in L^0(X + aY)$ and $H \cdot (X + aY) = H \cdot X + aH \cdot Y$.*

e) *We have 6.7 and 6.8 (with L^0 instead of L^2_{loc}), and 6.10.*

Now it may happen that X has components in both $\mathscr{V}$ and $\mathscr{H}^2_{\mathrm{loc}}$, and that H belongs to $L^0(X)$ and to $L^2_{\mathrm{loc}}(X)$; *a priori* we then have two integrals: the first one, denoted for a moment by $H \overset{m}{\cdot} X$, in the sense of § 6a, and the second one denoted by $H \overset{v}{\cdot} X$, in the sense of 6.13. Happily, these two integrals coincide:

6.15 **Lemma.** *If all components of $X - X_0$ are both in $\mathscr{V}$ and in $\mathscr{H}^2_{\mathrm{loc}}$ and if $H \in L^2_{\mathrm{loc}}(X) \cap L^0(X)$, then $H \overset{m}{\cdot} X = H \overset{v}{\cdot} X$, up to evanescence.*

Proof. Let $H(n) = H1_{\{|H| \leq n\}}$ and $Y(n) = H(n) \overset{m}{\cdot} X$ and $Z(n) = H(n) \overset{v}{\cdot} X$. Since $H(n)$ is bounded, the construction of the stochastic integral in Chapter I (see I.4.34(b)) shows that $Y(n) = Z(n)$. Now, 6.4a and 6.14a yield that $Y(n)$ and $Z(n)$ converge in measure locally uniformly towards $H \overset{m}{\cdot} X$ and $H \overset{v}{\cdot} X$, respectively: hence the result. □

§ 6c. Stochastic Integrals with Respect to a Multi-Dimensional Semimartingale

We are ready now for the general case: X is an arbitrary d–dimensional semimartingale. Recall from (4) in § I.4d (written for each component) that we can write, in many different ways:

6.16 $$X = X_0 + M + A, \quad M^i \in \mathscr{H}^2_{\mathrm{loc}}, \quad A^i \in \mathscr{V},$$

6.17 **Definition.** We say that a d–dimensional *predictable* process H is *integrable w.r.t.* X if there exists a decomposition 6.16 such that $H \in L^2_{\mathrm{loc}}(M) \cap L^0(A)$, and in this case we define the integral process by

6.18 $$H \cdot X = H \cdot M + H \cdot A$$

We denote by $L(X)$ the set of all (predictable) integrable processes H.

In view of Lemma 6.15, the stochastic integral process $H \cdot X$ above does not depend on the decomposition 6.16. Also, any locally bounded predictable process H belongs to $L(X)$, and in this case $H \cdot X$ can also be defined by 6.1 and 6.18 makes sense for *any* decomposition 6.16. The main properties are gathered in the following theorem:

6.19 **Theorem.** *Let X be a d–dimensional semimartingale. Then if $H \in L(X)$, we have:*

a) *For any sequence (D_n) of predictable sets increasing to $\Omega \times \mathbb{R}_+$, we have $H1_{D_n} \in L(X)$ and $\sup_{s \leq t} |(H1_{D_n}) \cdot X_s - H \cdot X_s| \to 0$ in measure for all $t \in \mathbb{R}_+$.*

b) *$H \cdot X$ is a semimartingale.*

c) *$L(X)$ is stable by localization and is a linear space, and $H \mapsto H \cdot X$ is linear, up to evanscence.*

d) *If Y is another d–dimensional semimartingale and $a \in \mathbb{R}$ and $H \in L(X) \cap L(Y)$, then $H \in L(X + aY)$ and $H \cdot (X + aY) = H \cdot X + aH \cdot Y$, up to evanscence.*

e) *We have 6.7 and 6.8 (with L instead of L^2_{loc}), and 6.10.*

Proof. (a) (resp. (b), resp. (d), resp. 6.7 and 6.10 in (e)) readily follow from 6.9a and 6.14a (resp. 6.4d and 6.14b, resp. 6.9b and 6.14d, resp. 6.4e, 6.9a and 6.14e). Note also that 6.8 holds when in addition K is locally bounded. So it remains to prove (c) and 6.8 in (e): this is more difficult, and requires a number of steps.

i) First we prove the following fact: suppose that X admits a (necessarily unique) decomposition 6.16 with A predictable, and let $H \in L(X)$ be such that $H \cdot X$ has bounded jumps: then $H \in L^2_{\text{loc}}(M) \cap L^0(A)$.

Indeed let $Y = H \cdot X$, which by I.4.24 is $Y = N + B$ with $N \in \mathscr{H}^2_{\text{loc}}$ and B predictable and in $\mathscr{V}$. Let $D_n = \{|H| \leq n\}$ and $H(n) = H 1_{D_n}$ and $Y(n) = H(n) \cdot X$. By 6.8 for a bounded K we have $Y(n) = 1_{D_n} \cdot Y$ and $Y(n)$ is a special semimartingale with canonical decomposition $Y(n) = N(n) + B(n)$, where $N(n) = 1_{D_n} \cdot N$ and $B(n) = 1_{D_n} \cdot B$. On the other hand, since $H(n)$ is bounded, we also have $N(n) = H(n) \cdot N$ and $B(n) = H(n) \cdot A$.

Consider a pair (F, c) associated with M by 6.2. Then, using 6.6,

$$\begin{aligned}((H \cdot c \cdot H) 1_{D_n}) \cdot F_t &= (\sum_{i,j} H(n)^i c^{ij} H(n)^j) \cdot F_t \\ &= \langle N(n), N(n) \rangle_t = 1_{D_n} \cdot \langle N, N \rangle_t,\end{aligned}$$

which increases to $(H \cdot c \cdot H) \cdot F_t$ and also to $\langle N, N \rangle_t$: hence $(H \cdot c \cdot H) \cdot F$ is locally integrable and $H \in L^2_{\text{loc}}(M)$. Similarly if the pair (F, a) is associated with A by 6.11 we have, with Var(.) denoting the variation process:

$$(1_{D_n} |\sum_i H^i a^i|) \cdot F_t = |\sum_i H(n)^i a^i| \cdot F_t = \text{Var}(B(n))_t = 1_{D_n} \cdot \text{Var}(B)_t.$$

Then passing to the limit gives $|\sum_i H^i a^i|) \cdot F = \text{Var}(B)$, hence $H \in L^0(A)$.

ii) We can now prove (c). Let $H, H' \in L(X)$, and set $Y = H \cdot X$ and $Y' = H' \cdot X$, and $D = \{|\Delta X| > 1\} \cup \{|\Delta Y| > 1\} \cup \{|\Delta Y'| > 1\}$. The optional set D is *discrete*, meaning that for all $\omega \in \Omega$ and $t > 0$, the set $\{s : s \leq t, (\omega, s) \in D\}$ is finite. Therefore we can set $G_t = X_0 + \sum_{s \leq t} \Delta X_s 1_D(s)$ and $K_t = \sum_{s \leq t} \Delta Y_s 1_D(s)$ and $K'_t = \sum_{s \leq t} \Delta Y'_s 1_D(s)$. Since D is discrete we obviously have $H, H' \in L^0(G)$ and $K = H \cdot G$ and $K' = H' \cdot G$. Then (d) yields that if $\bar{X} = X - G$ and $\bar{Y} = Y - K$ and $\bar{Y}' = Y' - K'$, then $H, H' \in L(\bar{X})$ and $\bar{Y} = H \cdot \bar{X}$ and $\bar{Y}' = H' \cdot \bar{X}$.

Henceforth, after noting the obvious facts that $H + aH' \in L(G)$ and $(H + aH') \cdot G = K + aK'$, and in view of (d) again, it remains to prove that $H + aH' \in L(\bar{X})$ and $(H + aH') \cdot \bar{X} = \bar{Y} + a\bar{Y}'$. But $\bar{X}'$ has bounded jumps, so by Lemma I.4.24 it can be written as $\bar{X} = M + A$ where the components of M are in $\mathscr{H}^2_{\text{loc}}$ and A is predictable and in $\mathscr{V}$. Further $\bar{Y} = H \cdot \bar{X}$ and $\bar{Y}' = H' \cdot \bar{X}$ have bounded jumps, so by (i) the two processes H and H' are in $L^2_{\text{loc}}(M) \cap L^0(A)$: it then suffices to apply 6.4e and 6.14c.

iii) Finally we prove 6.8, that is we start with $H \in L(X)$ and $H' \in L(Y)$, where $Y = H \cdot X$. We want to prove that $HH' \in L(X)$ and that $Y' = H' \cdot Y$ satisfies $Y' = (H'H) \cdot X$.

We proceed like in (ii): define D, G, K, $\bar{X}$ and $\bar{Y}$ as above, and set $K'_t = \sum_{s \leq t} \Delta Y'_s 1_D(s)$ and $\bar{Y}' = Y' - K'$. As above we have $H \in L^0(G)$ and $K =$

$H \cdot G$, and also $H' \in L^0(K)$ and $K' = H' \cdot K$; by 6.14e we get $HH' \in L^0(\bar{G})$ and $K' = (H'H) \cdot \bar{G}$. So it remains to prove that $H'H \in L(\bar{X})$ and that $\bar{Y}' = (H'H) \cdot \bar{X}$, while we know that $H \in L(\bar{X})$ and $H' \in L(\bar{Y})$.

As in (ii), $\bar{X} = M + A$ where the components of M are in $\mathscr{H}^2_{\mathrm{loc}}$ and A is predictable and in $\mathscr{V}$. Further $H \in L^2_{\mathrm{loc}}(M) \cap L^0(A)$ and $N = H \cdot M$ has components in $\mathscr{H}^2_{\mathrm{loc}}$ and $B = H \cdot A$ is predictable and in $\mathscr{V}$ and $\bar{Y} = N + B$. Now $\bar{Z} = K \cdot \bar{Y}$ has bounded jumps, so (i) yields that H' is in $L^2_{\mathrm{loc}}(N) \cap L^0(B)$. Then 6.4e and 6.14c imply that $HH' \in L^2_{\mathrm{loc}}(M) \cap L^0(A)$ and $H' \cdot N = (HH') \cdot M$ and $H' \cdot B = (HH') \cdot A$: the result follows. □

The above stochastic integral has nice properties, but also some others which do not correspond to intuition. For example the following questions are natural, even in the 1–dimensional case:

(1) If $X \in \mathscr{V}$, is it true that $H \cdot X \in \mathscr{V}$?

(2) If $X \in \mathscr{S}$, is it true that $H \cdot X \in \mathscr{S}$?

6.20 *The answer to (1) is Negative:* Consider a sequence $(\xi_n)_{n \geq 1}$ of i.i.d. variables with $P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}$, and the process $X_t = \sum_{n \geq 1} 2^{-n} \xi_n 1_{[q_n, \infty)}(t)$, where $(q_n)_{n \geq 1}$ is an enumeration of the rationals of (0, 1). Then X is in $\mathscr{H}^2$ (relatively to the smallest filtration w.r.t. which it is adapted) and $\langle X, X \rangle_t = \sum_{n \geq 1} 2^{-2n} 1_{[q_n, \infty)}(t) \leq 1$; and X is also in $\mathscr{A}$ because $\mathrm{Var}(X)_t = \sum_{n \geq 1} 2^{-n} |1_{[q_n, \infty)}(t)| \leq 1$. The deterministic process $H_t = \sum_{n \geq 1} \frac{2^n}{n} 1_{[q_n, \infty)}(t)$ belongs to $L^2(X)$ because $H^2 \cdot \langle X, X \rangle_t \leq \sum_{n \geq 1} \frac{1}{n^2} < \infty$, but the integral $Y = H \cdot X$ has $\mathrm{Var}(Y)_1 = \sum_{n \geq 1} \frac{1}{n} = +\infty$ and so $Y \notin \mathscr{V}$.

This counterexample also shows that, even in the 1–dimensional case, when $X \in \mathscr{V}$ the inclusion $L^0(X) \subset L(X)$ may be strict. □

6.21 *The answer to (2) is Negative:* Consider two independent variables: τ which is exponential with parameter 1, and ξ with $P(\xi = 1) = P(\xi = -1) = \frac{1}{2}$, and the process $X_t = \xi 1_{[\tau, \infty)}(t)$. Relatively to the smallest filtration w.r.t. which X is adapted, this process X is in $\mathscr{H}^2$ and also in $\mathscr{A}$. The deterministic process $H_t = \frac{1}{t} 1_{(0, \infty)}(t)$ belongs to $L^0(X)$, and $Y = H \cdot X$ is $Y_t = \frac{\xi}{\tau} 1_{[\tau, \infty)}(t)$. But Y does not belong to $\mathscr{S}$ (and not to $\mathscr{S}_p$ either), because $E(|Y_T|) = +\infty$ for every stopping time T such that $P(T > 0) > 0$. □

On the contrary, we also have "positive" answers to some other natural questions, as shown in the next result. Below, X is a d–dimensional semimartingale, and X^c denotes its continuous martingale part, taken componentwise, i.e. $X^{c,i}$ is the continuous martingale part of X^i.

6.22 **Proposition.** *Let X be a d–dimensional semimartingale and let $H \in L(X)$.*

a) *We have $H \in L^2_{\mathrm{loc}}(X^c)$, and $(H \cdot X)^c = H \cdot X^c$.*

b) *Let D be any discrete optional subset of $\Omega \times \mathbb{R}_+$ such that the processes $\Delta X 1_{D^c}$ and $(H \cdot \Delta X) 1_{D^c}$ are bounded, and let $Z_t = \sum_{s \le t} \Delta X_s 1_D(s)$, and denote by $X - Z = X_0 + M + A$ the canonical decomposition of the special semimartingale $X - Z$. Then H belongs to $L^2_{\text{loc}}(M)$, to $L^0(A)$ and to $L^0(Z)$, and $H \cdot M + H \cdot A$ is the canonical decomposition of the special semimartingale $H \cdot X - \sum_{s \le \cdot} \Delta Y_s 1_D(s)$.*

Proof. a) Consider any decomposition 6.16 such that $H \in L^2_{\text{loc}}(M) \cap L^0(A)$. Note that $X^c = M^c$. If $C^{ij} = \langle M^i, M^j \rangle$ and $C'^{ij} = \langle M^{c,i}, M^{c,j} \rangle$, there are factorizations 6.2 for C^{ij} and C'^{ij} with the same process F, in which case $c - c'$ also takes its values in the set of all symmetric nonnegative matrices. Then the property $H \in L^2_{\text{loc}}(M)$ immediately implies $H \in L^2_{\text{loc}}(M^c)$ and $H \in L^2_{\text{loc}}(M - M^c)$. Further 6.4d and 6.5 and 6.10 give that $H \cdot M^c$ is a continuous local martingale and that $H \cdot (M - M^c)$ is a locally square-integrable martingale orthogonal to all continuous martingales (since $M - M^c$ is so). Hence $(H \cdot M)^c = H \cdot M^c$. Since $H \cdot A \in \mathscr{V}$, we deduce the result from $H \cdot X = H \cdot M + H \cdot A$.

b) The semimartingales $X - Z$ and $H \cdot X - \sum_{s \le \cdot} \Delta Y_s 1_D(s)$ have bounded jumps, hence are special: so the result is just a restatement of part (i) of the proof of Theorem 6.19, plus the obvious facts that $H \in L^0(Z)$ and that $H \cdot Z = \sum_{s \le \cdot} (H \cdot \Delta X_s) 1_D(s)$. □

6.23 **Proposition.** *Let X be a 1–dimensional semimartingale and $H \in L(X)$ and $Y = H \cdot X$. If H never vanishes, then $H' = 1/H$ belongs to $L(Y)$ and $X = X_0 + H' \cdot Y$.*

Proof. Consider a decomposition 6.16 for X with $H \in L^2_{\text{loc}}(M) \cap L^0(A)$. Set $N = H \cdot M$ and $B = H \cdot A$, so $Y = N + B$ is a decomposition 6.16 for Y. On the one hand we trivially have $A = H' \cdot B$. On the other hand $\langle N, N \rangle = H^2 \cdot \langle M, M \rangle$, so $\langle M, M \rangle = H'^2 \cdot \langle N, N \rangle$: it readily follows that $H' \in L^2_{\text{loc}}(N)$ and thus $M = (HH') \cdot M = H' \cdot N$ by 6.19e. □

The last three results will not be explicitely used in the sequel and are given without proof:

6.24 **Proposition.** *Let X be a d–dimensional semimartingale and $H \in L(X)$ and $Y = H \cdot X$. Let P' be another probability measure on $(\Omega, \mathscr{F})$ which is locally absolutely continuous w.r.t. P. Then on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P')$, the process H is integrable w.r.t. X and its stochastic integral process coincide Q–a.s. with Y* (see [98] and apply 6.21 with $D = \{|\Delta X| > 1\} \cup \{|\Delta Y| > 1\}$).

6.25 **Proposition.** *Let X be a d–dimensional semimartingale and $H \in L(X)$ and $Y = H \cdot X$. Let $\mathbf{G}$ be a sub-filtration of $\mathbf{F}$ (i.e., $\mathscr{G}_t \subset \mathscr{F}_t$ for all t). Then if further X is adapted to $\mathbf{G}$ and if H is $\mathbf{G}$–predictable, on the basis*

$(\Omega, \mathscr{F}, \mathbf{G}, P)$, *the process* H *is integrable w.r.t.* X *and its stochastic integral process coincide* P*–a.s. with* Y (see [268]).

For the last result let us recall what is the so-called Emery's topology of semimartingales. For any two semimartingales X, Y we set

$$d(X, Y) = \sup\left(\sum_{n\geq 1} 2^{-n}\, E(|H\cdot(X-Y)_n \bigwedge 1) : H \text{ predictable}, |H| \leq 1\right).$$

This is a metric on the space $\mathscr{S}$, on which we identify two indistinguishable semimartingales. Then we have:

6.26 **Proposition.** *Let* X *be a* d*–dimensional semimartingale. Then the set* $\{H\cdot X : H \in L(X)\}$ *is closed for the Emery topology* (see [281]).

6.27 **Remark.** For defining $H\cdot X$ we started with the integrals of (locally) *bounded* predictable processes, as developed in Chapter I, and then we used the slightly strange definition 6.17. Considering the work already done in Chapter I, this is the shortest route towards the properties in Theorem 6.19, and at the end only those properties are really of use: the definition itself is in fact *ad hoc* for obtaining these properties.

However there is another route, which takes care of the present section and of Section I.4 at the same time: we start (in the 1–dimensional case or the multidimensional case as well) with the formula I.4.30 for "very simple" H as in I.4.29; then one extends it by linearity and continuity as much as possible, in such a way that the resulting process is a semimartingale: we end up with the same notion of stochastic integral as above. This route is not any shorter than the one taken here, but it gives some other interesting insights. □

6.28 **Remark.** At this stage, we can ask about the degree of generality we have reached, even in the 1–dimensional case. Some natural questions are as follows:

1) What are the most general predictable integrands w.r.t. some $X \in \mathscr{L}$? This question has two different answers: one may consider X as a semimartingale and integrate all elements of $L(X)$; but of course $H\cdot X$ is not necessarily in $\mathscr{L}$, as seen in 6.21. One may also ask for the most general H's for which $H\cdot X \in \mathscr{L}$. In view of 6.21, 6.10 and I.4.56, the answer is as follows: if $X \in \mathscr{L}$ the most general set of predictable H's such that $H\cdot X$ makes sense and belongs to $\mathscr{L}$ is the set $L^1_{\text{loc}}(X)$ (contained in $L(X)$) of all predictable H such that $(H^2\cdot[X,X])^{1/2}$ is locally integrable. There is also a multi-dimensional version, left to the reader.

2) The same question arises when $X \in \mathscr{S}$. Then the biggest set of predictable integrands is then $L(X)$: this follows for example from Proposition 6.26, or from a description of the jump processes ΔX for all semimartingales.

3) Is it possible to go beyond *predictable* integrands? If we want continuity properties like Lebesgue convergence theorem or 6.19a, the answer is in general NO. But for special processes the set of integrands may be much bigger: if $X \in \mathscr{V}$, any bounded measurable process has a "stochastic" integral, whether predictable or adapted or even not adapted. If X is a *continuous* local martingale any bounded *progressively measurable* process H has a stochastic integral, which indeed coincides with the integral of the predictable projection H^p of H. If X has jumps and is of locally infinite variation, then the predictability is essentially necessary (although some efforts have been done towards stochastic integrals of optional processes, but with rather unsatisfactory answers). □

§ 6d. Stochastic Integrals: A Predictable Criterion

In the way for a more "concrete" description of the set $L(X)$ of predictable integrands w.r.t. a d–dimensional semimartingale X, it is tempting to try for "predictable" criteria based upon the characteristics of X. This is easily done as a consequence of Proposition 6.22.

We denote by (B, C, ν) the characteristics of X with respect to a truncation function h. We also use a factorization of the form II.2.10:

6.29 $$B^i = b^i \cdot A, \quad C^{ij} = c^{ij} \cdot A, \quad \nu(\omega, dt, dx) = K_{\omega,t}(dx)\, dA_t(\omega),$$

where A is a predictable increasing process and b, c and K are predictable. Recall that if $\Delta A_t > 0$, then $\int K_t(dx)(|x| \wedge 1) < \infty$.

6.30 **Theorem.** *A predictable $\mathbb{R}^d$–valued process H belongs to $L(X)$ if and only if the nonnegative process*

$$\begin{aligned} v(H)_s = {} & H_s \cdot c_s \cdot H_s + \int_{\{|x|\le 1\}} K_s(dx)(1 \wedge |H_s \cdot x|^2) \\ & - \Delta A_s \left| \int_{\{|x|\le 1, |H_s \cdot x| \le 1\}} K_s(dx) H_s \cdot x \right|^2 \\ & + \left| H_s \cdot b_s + \int K_s(dx) \left((H_s \cdot x) 1_{\{|x|\le 1, |H_s\cdot x|\le 1\}} - H_s \cdot h(x) \right) \right| \end{aligned}$$

is in $L^0(A)$, that is if $\int_0^t v(H)_s dA_s < \infty$ a.s. for all $t < \infty$.

Proof. Let H be given and set $D = \{|\Delta X| > 1\} \cup \{|H \cdot \Delta X| > 1\}$. Proposition 6.22 yields that H belongs to $L(X)$ if and only we have (i) and (ii) below:

(i) D is a.s. discrete,

(ii) putting $Z_t = \sum_{s\le t} \Delta X_s 1_D(s)$ and letting $X - Z = X_0 + M + G$ be the canonical decomposition of $X - Z$, then H is in $L^2_{\mathrm{loc}}(M)$ and in $L^0(G)$

(under (i), Z makes sense and $X - Z$ is a special semimartingale).

We set $\widetilde{D} = \{(\omega, t, x) : |x| \leq 1, |H_t(\omega) \cdot x| \leq 1\}$. First, (i) is equivalent to having $1_{\widetilde{D}^c \cap \{|x| \leq 1\}} \star \mu_t < \infty$ a.s. for all t, which in turn amounts to

6.31 $$1_{\widetilde{D}^c \cap \{|x| \leq 1\}} \star \nu_t < \infty, \qquad \text{a.s. for all } t < \infty$$

Next we can compute M and G: by II.2.35 we can write

$$M + G = X^c + h(x) \star (\mu - \nu) + B + (x 1_{\widetilde{D}} - h(x)) \star \mu.$$

The last term above is with finite variation and is also a special semimartingale, hence it has locally integrable variation with compensator $(x 1_{\widetilde{D}} - h(x)) \star \nu$. The uniqueness of the canonical decomposition gives

$$G = B + (x 1_{\widetilde{D}} - h(x)) \star \nu, \qquad M = X^c + (x 1_{\widetilde{D}}) \star (\mu - \nu).$$

So we see that $G = g \cdot A$, where $g_t = b_t + \int (x 1_{\{|x| \leq 1, |H_t \cdot x| \leq 1\}} - h(x)) K_t(dx)$. Finally $\langle M^i, M^j \rangle = \sigma^{ij} \cdot A$, where

$$\begin{aligned} \sigma_t^{ij} &= c^{ij} + \int K_t(dx) x^i x^j 1_{\widetilde{D}}(t, x) \\ &\quad - \Delta A_t \int K_t(dx) x^i 1_{\widetilde{D}}(t, x) \int K_t(dx) x^j 1_{\widetilde{D}}(t, x). \end{aligned}$$

Then (ii) is equivalent to having the process $|H \cdot g| + H \cdot \sigma \cdot H$ in $L^0(A)$: combining this with 6.31 gives the result. □

As an example, we consider the case where X is a PIIS, with characteristics (b, c, F): hence in 6.29 we may take $A_t = t$ and $b_t = b$, $c_t = c$ and $K_t = F$. Then we have $H \in L(X)$ if and only if $\int_0^t v(H)_s ds < \infty$ a.s. for all t, where

6.32 $$\begin{aligned} v(H)_s &= H_s \cdot c \cdot H_s + \int_{\{|x| \leq 1\}} F(dx) (1 \bigwedge |H_s \cdot x|^2) \\ &\quad + \left| H_s \cdot b + \int F(dx) \left((H_s \cdot x) 1_{\{|x| \leq 1, |H_s \cdot x| \leq 1\}} - H_s \cdot h(x) \right) \right|. \end{aligned}$$

Let us now consider the case where X is a 1–dimensional stable process with index $\alpha \in (0, 2)$: we have $c = 0$ and $F(dx) = \left(\frac{a_+}{|x|^{1+\alpha}} 1_{\{x>0\}} + \frac{a_-}{|x|^{1+\alpha}} 1_{\{x<0\}} \right) dx$. We also take the truncation function h to be $h(x) = x 1_{\{|x| \leq 1\}}$. Then if $|H_s| \geq 1$ we have

$$\int_{\{|x| \leq 1\}} F(dx) (1 \bigwedge |H_s \cdot x|^2) = \frac{a_+ + a_-}{\alpha} \left(\frac{2}{2 - \alpha} |H_s|^\alpha - 1 \right).$$

If $|H_s| \geq 1$ again, the last term in 6.32 equals

$$\begin{cases} \left| H_s \tilde{b} - \frac{a_+ - a_-}{1-\alpha} |H_s|^\alpha \operatorname{sign}(H_s) \right|, \quad \text{with } \tilde{b} = b - \frac{a_+ - a_-}{1-\alpha}, & \text{if } \alpha \neq 1 \\ |H_s b - (a_+ - a_-) H_s \log |H_s|| & \text{if } \alpha = 1; \end{cases}$$

Then one readily deduces that the real-valued predictable process H belongs to $L(X)$ if and only if:

- If $\alpha \in (1,2)$, or if $\alpha = 1$ and $a_+ = a_-$, or if $\alpha \in (0,1)$ and $\tilde{b} = 0$, then $\int_0^t |H_s|^\alpha ds < \infty$ a.s. for all t,
- If $\alpha = 1$ and $a_+ \neq a_-$, then $\int_0^t |H_s|(\log^+ |H_s|) ds < \infty$ a.s. for all t,
- If $\alpha \in (0,1)$ and $\tilde{b} \neq 0$, then $\int_0^t |H_s| ds < \infty$ a.s. for all t.

In particular when X is a symmetric stable process of index α then $H \in L(X)$ iff $\int_0^t |H_s|^\alpha ds < \infty$ a.s. for all t: so we find in $L(X)$ the class of integrands which was previously known (see e.g. [288]) to be a "natural" set of integrands w.r.t. a symmetric stable process.

§ 6e. Σ–localization and σ–martingales

In Chapter I we introduced the "localization procedure", which is " used over and over". For example a local martingale (an element of $\mathscr{M}_{\mathrm{loc}}$) is a process X for which there exists a sequence (T_n) of stopping times increasing to infinity, and such that each stopped process X^{T_n} is a uniformly integrable martingale (an element of $\mathscr{M}$). Clearly if X is an adapted process, saying that $X - X_0$ is a local martingale is equivalent to saying that it is a semimartingale such that each process $1_{\Sigma_n} \cdot X$ is a uniformly integrable martingale, where $\Sigma_n = [\![0, T_n]\!]$ and the T_n's are stopping times increasing to infinity.

In the above, the Σ_n's constitute a particular sequence of predictable sets increasing to $\Omega \times \mathbb{R}_+$. A more general "localization procedure" is naturally defined in an analoguous way, by using an *arbitrary* sequence of predictable sets increasing to $\Omega \times \mathbb{R}_+$; it will be called *$\Sigma$–localization*:

6.33 **Definition.** A semimartingale X is called a σ–martingale if there exists an increasing sequence of predictable sets Σ_n such that $\bigcup \Sigma_n = \Omega \times \mathbb{R}_+$, and if for any $n \geq 1$ the process $X^{\Sigma_n} = 1_{\Sigma_n} \cdot X$ is a uniformly integrable martingale. We denote by $\mathscr{M}_\sigma$ the class of all σ–martingales.

A d–dimensional process $X = (X^i)_{i \leq d}$ is called a σ–martingale if each of its components is in $\mathscr{M}_\sigma$: this clearly implies the existence of a sequence Σ_n as above (not depending on i), such that each $1_{\Sigma_n} \cdot X^i$ is a uniformly integrable martingale.

6.34 **Proposition.** *The class $\mathscr{M}_\sigma$ is stable by Σ–localization, and thus by localization (i.e., $(\mathscr{M}_\sigma)_{\mathrm{loc}} = \mathscr{M}_\sigma$). In particular any local martingale is a σ–martingale.*

Proof. First note that $X \in \mathscr{M}_\sigma$ iff there is a predictable countable partition (Σ'_n) of $\Omega \times \mathbb{R}_+$ such that each $1_{\Sigma'_n} \cdot X$ belongs to $\mathscr{M}$ (the relation with the sequence (Σ_n) of Definition 6.33 being obviously $\Sigma_n = \cup_{1\leq i\leq n}\Sigma'_i$, or conversely $\Sigma'_n = \Sigma_n\backslash\Sigma_{n-1}$, with $\Sigma_0 = \emptyset$).

We can now prove our result. Let X be a semimartingale such that there exists a sequence (Σ_n) of predictable sets increasing to $\Omega\times\mathbb{R}_+$, with $\Sigma_0 = \emptyset$, and such that $X(n) = 1_{\Sigma_n}\cdot X$ belongs to $\mathscr{M}_\sigma$ for all n. For each n we thus have a predictable partition $(\Sigma(n,m) : m \geq 1)$ of $\Omega \times \mathbb{R}_+$ such that $X(n,m) = 1_{\Sigma(n,m)}\cdot X(n)$ is in $\mathscr{M}$. Let $\Sigma'(n,m) = \Sigma(n,m) \cap (\Sigma_n\backslash\Sigma_{n-1})$. Obviously $X'(n,m) = 1_{\Sigma'(n,m)}\cdot X = 1_{\Sigma'(n,m)}\cdot X(n)$ is a local martingale, localized by a sequence $(T(n,m,p) : p \geq 1)$ of stopping times, and we set $T(n,m,0) = 0$ and $\Sigma''(n,m,p) = \Sigma'(n,m)\cap]\!]T(n,m,p-1), T(n,m,p)]\!]$ (with a closed interval instead of a semi-open one when $p = 0$). Then $1_{\Sigma''(n,m,p)}\cdot X = X'(n,m)^{T(n,m,p)} - X'(n,m)^{T(n,m,p-1)}$ is in $\mathscr{M}$. Since the $(\Sigma''(n,m,p) : n, m, p \geq 1)$ constitute a countable predictable partition of $\Omega \times \mathbb{R}_+$, we get that $X \in \mathscr{M}_\sigma$. □

In order to study σ-martingales it is convenient to first give a characterization of these processes in terms of the characteristics, along the lines of Proposition II.2.29. Suppose that $X = (X^i)_{i\leq d}$ is a d-dimensional semimartingale and let (B, C, ν) denote its characteristics with respect to a truncation function h. We have the factorization 6.29. Then the process $t \mapsto \int_{\mathbb{R}^d}(|x|^2 \wedge 1)K_t(dx)$ is always in $L^0(A)$. Further II.2.29 says that X is a special semimartingale (resp. is a locally square-integrable semimartingale) if and only if the process $t \mapsto \int_{\mathbb{R}^d}(|x|^2 \wedge |x|)K_t(dx)$ (resp. $t \mapsto \int_{\mathbb{R}^d}|x|^2K_t(dx)$) is in $L^0(A)$ (recall that an increasing finite-valued and predictable process is necessarily locally bounded). Similarly, we have:

6.35 **Proposition.** *Let $X = (X^i)_{i\leq d}$ be a semimartingale with characteristics (B, C, ν) given by 6.29.*

a) *$X - X_0$ is a local martingale if and only if the process $t \mapsto \int_{\mathbb{R}^d}(|x|^2 \wedge |x|)$ $K_t(dx)$ belongs to $L^0(A)$ and if we have the following two properties:*

6.36
$$b(\omega,t) + \int_{\mathbb{R}^d}(x - h(x))\,K_{\omega,t}(dx) = 0 \quad \text{for } P(d\omega)dA_t(\omega) \text{ almost all } (\omega,t)$$

6.37
$$\Delta A_t(\omega) > 0 \implies \int_{\mathbb{R}^d} x\,K_{\omega,t}(dx) = 0 \quad \text{for } P(d\omega)dA_t(\omega) \text{ almost all } (\omega,t)$$

b) *X is a σ-martingale if and only if we have 6.36, 6.37 and*

6.38
$$\int_{\mathbb{R}^d}(|x|^2 \wedge |x|)\,K_{\omega,t}(dx) < \infty \quad \text{for } P(d\omega)dA_t(\omega) \text{ almost all } (\omega,t)$$

Furthermore $X - X_0$ is a local martingale if and only if is it a σ-martingale and a special semimartingale.

Proof. (a) follows from II.2.29 and from the fact that X is a local martingale iff it is a special semimatingale with $B + (x - h(x)) \star \nu = 0$.

(b) First assume $X \in \mathscr{M}_\sigma$, and let (Σ_n) be the sequence of associated predictable sets in Definition 6.33. The characteristics of $X(n) = 1_{\Sigma_n} \cdot X$ are given by 6.29 with the same A, and (b, c, K) replaced by $b(n) = 1_{\Sigma_n} b$, $c(n) = 1_{\Sigma_n} c$ and $K(n)_{\omega,t}(dx) = 1_{\Sigma_n}(\omega, t) K_{\omega,t}(dx)$. Since each $X(n)$ is a local martingale, we deduce from (a) that 6.36, 6.37 and 6.38 hold $dP \times dA_t$–a.s. on Σ_n. This and $\Sigma_n \uparrow \Omega \times \mathbb{R}_+$ yield the necessary condition.

Conversely assume 6.36, 6.37 and 6.38. Let $H_t(\omega)$ denote the left side of 6.38, and set $\Sigma_n = \{H = \infty\} \cup \{H \leq n\}$, which is predictable and increases to $\Omega \times \mathbb{R}_+$. The characteristics of $X(n) = 1_{\Sigma_n} \cdot X$ are as described above, hence they readily satisfy 6.33 and 6.34 and $\int_0^t dA_s 1_{\Sigma_n}(s) \int_{\mathbb{R}^d} (|x|^2 \wedge |x|)\, K_s(dx) < \infty$ a.s.: then (a) implies that $X(n)$ is a local martingale, and we conclude by 6.34.

Finally the last claim readily follows from (a), (b) and II.2.29. □

6.39 **Remark.** This proposition shows the difference between a local martingale and a σ–martingale. When X is a PIIS, in 6.29 we can take $A_t = t$, in which case the terms b, c, K are independent of ω and t. Therefore it readily follows from 6.35 that a PIIS which is a σ–martingale is a local martingale (and in fact even a martingale). The following example shows that it is not necessarily the case when X is a general PII.

6.40 *Example:* Let $(\xi_n)_{n \geq 1}$ be a sequence of independent random variables with

$$P(\xi_n = n) = P(\xi_n = -n) = \frac{1}{2n^2}, \qquad P(\xi_n = 0) = 1 - \frac{1}{n^2}.$$

Take $t_n = 1 - \frac{1}{n}$ and set

$$X_t = \sum_{\{n : t_n \leq t\}} \xi_n, \qquad t \geq 0.$$

With probability one only a finite number of ξ_n's differ from zero, so the process X is well defined. Set $\Sigma_n = \Omega \times ([0, 1 - \frac{1}{n}] \cup [1, \infty))$. For any $n \geq 1$, $1_{\Sigma_n} \cdot X$ is a uniformly integrable martingale, so X is a σ–martingale. Further, X is also obviously a PII, but it is *not* a local martingale.

Although a σ–martingale is not necessarily a local martingale, some rather mild additional properties yield this implication. For example, a σ–martingale which is bounded from below (or from above) by a constant is a local martingale. Hence any locally bounded σ–martingale is a local martingale, and in particular any *continuous* σ–martingale is a local martingale.

A process X which can be written as $Y = H \cdot M$ for some $M \in \mathscr{L}$ and some $H \in L(M)$ is sometimes called a *martingale transform*. In the discrete-time setting the set of all martingales transforms is $\mathscr{L}$. In general the class of martingale transforms is exactly $\mathscr{M}_\sigma$, as shown in the next result.

6.41 **Theorem.** *Let X be a d-dimensional semimartingale. The following three properties are equivalent:*

(i) *The process X is a σ-martingale.*

(ii) *There exist a d-dimensional local martingale M and processes $H^i \in L(M^i)$ such that and $X^i = X_0^i + H^i \cdot M^i$.*

(iii) *There exist a d-dimensional local martingale N with $E([N^i, N^i]_\infty^{1/2}) < \infty$ and a strictly positive process $K \in \cap_{1\le i\le d} L(N^i)$ such that $X^i = X_0^i + K \cdot N^i$.*

Proof. (i) $\Rightarrow$ (iii). Denote by μ the jump measure of X and by (B, C, ν) its characteristics, with the factorization 6.29: since A is locally bounded (since it is predictable), up to replacing A by an "equivalent" process we can even suppose that $A_\infty \le 1$ identically. By assumption we have the properties 6.36, 6.37 and 6.38 and, up to modifying the characteristics on a null set, we can assume that 6.38 holds for all (ω, t).

Set $H_t = \int_{\mathbb{R}^d} (|x|^2 \bigwedge |x|)\, K_t(dx)$ and $\Gamma_0' = \emptyset$ and $\Gamma_n' = \{|H| \le n\} \cap \{|c| \le n\}$ (where c is as in 6.29) and $\Gamma_n = \Gamma_n' \backslash \Gamma_{n-1}'$ for $n \ge 1$. Set also $\widetilde{K} = \sum_{n\ge1} \frac{1}{n^3} 1_{\Gamma_n}$, which is predictable with $0 < \widetilde{K} \le 1$. By 6.38, the Γ_n's constitute a predictable partition of $\Omega \times \mathbb{R}_+$. On the one hand we have

$$\widetilde{K}^2 \cdot C_\infty^{ii} = \sum_{n\ge1} \frac{1}{n^6} (c^{ii} 1_{\Gamma_n}) \cdot A_\infty \le \sum_{n\ge1} \frac{1}{n^5} 1_{\Gamma_n} \cdot A_\infty \le \alpha$$

for some constant α (recall $A_\infty \le 1$). Hence $K \in L^2(X^{i,c})$ and we can define the continuous martingales $N^{i,c} = \widetilde{K} \cdot X^{i,c}$. On the other hand we set $W(\omega, t, x) = \widetilde{K}_t(\omega) x^i$ and consider the associated processes $\widehat{W}$ and $\widetilde{W}$ as in II.1.24 and II.1.27. By 6.37 we have $\widehat{W} = 0$, so $\widetilde{W}_t = \widetilde{K}_t \Delta X_t^i$. Therefore the process $G_t = \left(\sum_s (\widetilde{W}_s)^2\right)^{1/2}$ has, with $(T_p)_{p\ge1}$ denoting the sequence of jump times of X with size > 1:

$$\begin{aligned}
G_\infty &\le \left((\widetilde{K}^2 |x|^2 1_{\{|x|\le1\}} \star \mu_\infty + \sum_{p\ge1} \widetilde{K}_{T_p}^2 |\Delta X_{T_p}|^2 \right)^{1/2} \\
&\le 1 + (\widetilde{K}^2 |x|^2 1_{\{|x|\le1\}} \star \mu_\infty + \sum_{p\ge1} \widetilde{K}_{T_p} |\Delta X_{T_p}| \\
&\le 1 + (\widetilde{K}(|x|^2 \bigwedge |x|)) \star \mu_\infty
\end{aligned}$$

because $(\sum_n a_n^2)^{1/2} \le \sum_n a_n$ for any sequence $a_n \ge 0$ and also $\sqrt{a} \le 1 + a$ if $a \ge 0$, and $\widetilde{K} \le 1$. Then, coming back to the definition of $\widetilde{K}$ and using the fact that ν is the predictable compensator of μ, we get

$$
\begin{aligned}
E(G_\infty) &\le 1 + E\left[(\widetilde{K}(|x|^2 \textstyle\bigwedge |x|)) \star \nu_\infty\right] \\
&= 1 + \sum_{n\ge1} \frac{1}{n^3}\, E\left[((|x|^2 \textstyle\bigwedge |x|)1_{\Gamma_n}) \star \nu_\infty\right] \\
&\le 1 + \sum_{n\ge1} \frac{1}{n^3}\, E\left[(H 1_{\Gamma_n})\cdot A_\infty\right] \;\le\; 1 + \sum_{n\ge1} \frac{1}{n^2} \;=\; \beta
\end{aligned}
$$

(we use the fact that $H \le n$ on Γ_n). Then by II.1.27 we can define a purely discontinuous local martingale by $N^{i,d} = (\widetilde{K}x^i)\star(\mu-\nu)$. Then $N^i = N^{i,c}+N^{i,d}$ has $[N^i, N^i] = K^2\cdot C^{ii} + G^2$, hence $E(|N^i, N^i]_\infty^{1/2}) \le \sqrt{\alpha} + \beta < \infty$, and N^i is a martingale.

Now, the characteristics of each process $X(n) = 1_{\Gamma_n}\cdot X$ are of the form 6.29 with $b(n) = b1_{\Gamma_n}$, $c(n) = c1_{\Gamma_n}$ and $K(n)_t(x) = K_t(dx)1_{\Gamma_n}(t)$, so they satisfy all conditions in 6.35a: hence $X(n) \in \mathscr{L}$, and we have $X(n)^i = 1_{\Gamma_n}\cdot X^{i,c} + (x^i 1_{\Gamma_n}) \star (\mu - \nu)$. Therefore $1_{\Gamma_n}\cdot N^i = K\cdot X(n)^i = (K1_{\Gamma_n})\cdot X^i$, and

$$
(\sum_{n=1}^{N} 1_{\Gamma_n})\cdot N^i = (\widetilde{K}\sum_{n=1}^{N} 1_{\Gamma_n})\cdot X.
$$

Then since $\cup_{1\le n\le N}\Gamma_n$ increases to $\Omega \times \mathbb{R}_+$ as $N \uparrow \infty$, we deduce that $N^i = \widetilde{K}\cdot X^i$ by the dominated convergence theorem for stochastic integrals. It remains to apply Proposition 6.23: if $K = 1/\widetilde{K}$, then $K \in L(N^i)$ and $X^i = X_0^i + K\cdot N^i$.

(iii) $\Rightarrow$ (ii). This implication is obvious.

(ii) $\Rightarrow$ (i). It is enough to treat the 1–dimensional case. Set $\Sigma_n = \{|H| \le n\}$. Note that $1_{\Sigma_n}\cdot X = H(n)\cdot M$, where $H(n) = H1_{\Sigma_n}$. Since $H(n)$ is bounded, the process $H(n)\cdot M$ is a local martingale, so we have stopping times $\tau_{n,m}$ increasing to infinity as $m \uparrow \infty$, and such that $(H(n)\cdot M)^{\tau_{n,m}}$ is a uniformly integrable martingale: then it remains to set $\Sigma_{n,m} = \Sigma_n \cap [\![0, \tau_{n,m}]\!]$ and to observe that $1_{\Sigma_{n,m}}\cdot X = (H(n)\cdot M)^{\tau_{n,m}}$ and that the double sequence $\Sigma_{n,m}$: $n, m \ge 1)$ is a predictable partition of $\Omega \times \mathbb{R}_+$. □

As seen in Example 6.21 the class of all local martingales is not stable by stochastic integration. An interesting feature of σ–martingales is that this class *is* stable by stochastic integration:

6.42 Proposition. *If X is a d–dimensional σ–martingale and if $H \in L(X)$, then $H\cdot X$ is also a σ–martingale.*

Proof. Let Σ_n be associated with X as in Definition 6.33, and set $\Sigma'_n = \{|H| \le n\}$. The sets $\Sigma''_n = \Sigma_n \cap \Sigma'_n$ are predictable and increase to $\Omega \times \mathbb{R}_+$. Moreover if $H(n) = H1_{\Sigma'_n}$ and $X(n) = 1_{\Sigma_n}\cdot X$, then $1_{\Sigma''_n}\cdot(H\cdot X) = H(n)\cdot X(n)$ is a local martingale: hence each $1_{\Sigma''_n}\cdot(H\cdot X)$ is a local martingale and a further localization yields that $H\cdot X$ is a σ–martingale. □

7. Laplace Cumulant Processes and Esscher's Change of Measures

§ 7a. Laplace Cumulant Processes of Exponentially Special Semimartingales

The natural place of this section would rather be at the end of Chapter II, except that we will use the vector stochastic integrals defined just above. Let $X = (X^1, \dots, X^d)$ be d–dimensional semimartingale with jump maesure μ and characteristics (B, C, ν) defined in 2.6 (w.r.t. some truncation function h). As in 6.29 we introduce the following factorization for

7.1 $$B = b \cdot A, \qquad C = c \cdot A, \qquad \nu(\omega; dt, dx) = dA_t(\omega) F_{(\omega,t)}(dx),$$

where A is a predictable process in $\mathscr{V}$ and b, c, F are predictable (we use the notation F instead of K here because the letter K will be used for another purpose below).

Observe that if $u \in \mathbb{R}^d$ we have obtained in Theorem II.2.47 and II.2.49 the multiplicative decomposition (in the sense of II.8.19) for the semimartingale $e^{iu\cdot X}$, at least when the process $A(u)$ of 2.40 has no jump equal to -1: namely the process $G(u) = \mathscr{E}[A(u)]$ plays the rôle of D in the multiplicative decomposition of $e^{iu\cdot X}$, or rather of $e^{iu\cdot(X-X_0)}$ (although these processes are complex-valued, the theory is the same; see II.8.19). And when X is a PII then $G(u)_t$ is also (as a function of u) the characteristic function of X_t: so in general $G(u)$ could be called the *Fourier transform process*, or *Fourier cumulant process*. Now, when a random variable is sufficiently integrable it is convenient sometimes to replace the characteristic function by the Laplace transform; in the setting of multidimensional processes, this is $E(e^{\lambda\cdot X_t})$ for $\lambda \in \mathbb{R}^d$, as soon as the integral exists, and the Laplace cumulant is the logarithm of this integral. This is what this subsection is all about, except that we replace λ by a d–dimensional predictable process.

7.2 **Definition.** a) Let X be a d–dimensional semimartingale and let θ be in $L(X)$ and such that the semimartingale $\theta \cdot X$ is exponentially special. Then the *Laplace cumulant* $\widetilde{K}^X(\theta)$ of X at θ is defined as the *additive compensator* (i.e. the part which is predictable with locally bounded variation in the canonical decomposition) of the real-valued special semimartingale $\mathscr{L}og\,(e^{\theta\cdot X})$.

b) The *modified Laplace cumulant* $K^X(\theta)$ of X at θ is

7.3 $$K^X(\theta) = \log\, \mathscr{E}(\widetilde{K}^X(\theta)).$$

7.4 **Theorem.** *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special. Then*

7.5 $$\widetilde{K}^X(\theta) = \widetilde{\varkappa}(\theta) \cdot A,$$

where (for the product notation, see § II.2d)

7.6 $\quad \widetilde{\varkappa}(\theta)_t = \theta_t \cdot b_t + \frac{1}{2}\,\theta_t \cdot c_t \cdot \theta_t + \int \left(e^{\theta_t \cdot x} - 1 - \theta_t \cdot h(x)\right) F_t(dx),$

which belongs to $L^0(A)$. Further we have

7.7 $$K^X(\theta)_t = \widetilde{K}^X(\theta)_t + \sum_{s \le t} \left[\log\left(1 + \Delta \widetilde{K}^X(\theta)_s\right) - \Delta \widetilde{K}^X(\theta)_s\right],$$

7.8 $$\widetilde{K}^X(\theta)_t = K^X(\theta)_t + \sum_{s \le t} \left[e^{\Delta K^X(\theta)_s} - 1 - \Delta K^X(\theta)_s\right],$$

7.9 $$\Delta \widetilde{K}^X(\theta)_t = \int \left(e^{\theta_t \cdot x} - 1\right) \nu(\{t\} \times dx),$$

7.10 $$\Delta K^X(\theta)_t = \log\left(1 + \Delta \widetilde{K}^X(\theta)_t\right).$$

Finally, if X is quasi-left continuous then $K^X(\theta) = \widetilde{K}^X(\theta)$ and $K^X(\theta)$ is continuous.

Proof. The processes $Y = \theta \cdot X$ and $\bar{Y} = \mathscr{L}\mathrm{og}\,(e^Y)$ are in the same relationship than X and $\bar{X}$ in Theorem II.8.10, and so they satisfy II.8.11, and also $Y_0 = \bar{Y}_0 = 0$.

(i) We first compute the characteristics (B', C', ν') of Y, w.r.t. some truncation function h' on $\mathbb{R}$. 6.6 and 6.22a give $C' = (\theta \cdot c \cdot \theta) \cdot A$. Then 6.10 gives that the jump measure μ' of Y is the image of μ by the map $x \mapsto \theta \cdot x$, and since θ is predictable ν' is the image of ν by the same map.

The computation of B' is more delicate. Set $D = \{|\Delta X| > 1\} \cup \{|\Delta Y| > 1\}$ and $U = \{(\omega, t, x) : |x| > 1 \text{ or } |\theta_t(\omega) \cdot x| > 1\}$ and $Z_t = \sum_{s \le t} \Delta X_s 1_D(s) = (x 1_U) \star \mu_t$. By 6.22 we have $X = X_0 + Z + X^c + M + G$, where the components of M are purely discontinuous elements of $\mathscr{H}^2_{\mathrm{loc}}$ and the components of G are predictable and in $\mathscr{V}$, and further $\theta \in L^2_{\mathrm{loc}}(X^c) \cap L^0(Z) \cap L^2_{\mathrm{loc}}(M) \cap L^0(G)$.

Below we write $V \sim W$ if V and W are two processes such that $V - W$ has components in $\mathscr{L}$. If we use II.2.35 for X we get $Z + G \sim (x - h(x)) \star \mu + B$, that is $G \sim B + (x 1_{U^c} - h(x)) \star \mu$. Since $G - B$ is predictable with components in $\mathscr{V}$ we get that $(x 1_{U^c} - h(x)) \star \mu$ is with locally integrable variation, with compensator $(x 1_{U^c} - h(x)) \star \nu$, and thus

7.11 $$G = B + (x 1_{U^c} - h(x)) \star \nu.$$

On the other hand we have $Y \sim \theta \cdot Z + \theta \cdot G$, and $\theta \cdot Z$ is obviously equal to $((\theta \cdot x) 1_U) \star \mu$. But II.2.35 for Y gives $Y \sim (y - h'(y)) \star \mu' + B' = (\theta \cdot x - h'(\theta \cdot x)) \star \mu + B'$. Hence by difference we get $B' \sim \theta \cdot G + (h'(\theta \cdot x) - (\theta \cdot x) 1_{U^c}) \star \mu$, and as before we deduce

7.12 $$B' = \theta \cdot G + (h'(\theta \cdot x) - (\theta \cdot x) 1_{U^c}) \star \nu.$$

Now we use 7.1 to deduce from 7.11 and 7.12 that

7.13 $\quad B' = b' \cdot A, \quad$ where $\quad b'_t = \theta_t \cdot b_t + \int F_t(dx)(h'(\theta_t \cdot x) - \theta_t \cdot h(x)).$

(ii) Now we apply II.8.11 to the pair $(Y, \bar{Y})$ to obtain

$$\bar{Y} = Y^c + h' \star (\mu' - \nu') + B' + \frac{1}{2}C' + (e^x - 1 - h'(x)) \star \mu'.$$

The two first terms on the right above are local martingales, the third and fourth ones are predictable with components in $\mathscr{V}$, and the last three ones have components in $\mathscr{V}$. Since $\bar{Y}$ is special, it follows from I.4.23(iii) that the last term is indeed of locally integrable variation, with compensator $(e^x - 1 - h'(x)) \star \nu'$: therefore the compensator of $\bar{Y}$ is, by using 7.2:

$$\begin{aligned}\widetilde{K}(\theta) &= B' + \frac{1}{2}C' + (e^x - 1 - h'(x)) \star \nu' \\ &= B' + \frac{1}{2}(\theta \cdot c \cdot \theta) \cdot A + (e^{\theta \cdot x} - 1 - h'(\theta \cdot x)) \star \nu\end{aligned}$$

and, in view of 7.13, we readily get 7.5 and 7.6.

To prove the remaining claims, we note that by 7.5 and II.2.14,

$$\Delta \widetilde{K}^X(\theta)_t = \int \left(e^{\theta_t \cdot x} - 1\right) \nu(\{t\} \times dx).$$

So we have 7.9. Further, the processes $K^X(\theta)$ and $\widetilde{K}^X(\theta)$ are again in the same relationship than X and $\widetilde{X}$ in Theorem II.8.10, and $K^X(\theta)_0 = \widetilde{K}^X(\theta)_0 = 0$, and also these pocesses have a vanishing continuous martingale part: so 7.7, 7.8 and 7.10 follow from II.8.12, II.8.11 and II.8.13 respectively. $\square$

7.14 **Proposition.** *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special. Then the modified Laplace cumulant $K^X(\theta)$ is also the exponential compensator of $\theta \cdot X$.*

Proof. Let $W = e^{\theta \cdot X}$. As seen before Proposition II.8.29, the exponential compensator V of $\theta \cdot X$ is $V = \log D$, where $W = LD$ is the multiplicative decomposition of W. Then if $W = 1 + M + A$ is the canonical decomposition of W, Theorem II.8.21 gives that $D = 1/\mathscr{E}(U)$, where $U = -\frac{1}{W_- + \Delta A} \cdot A$. Then $V = \log D$ and so in view of 7.3 we only have to prove that $D = \mathscr{E}(\widetilde{K}^X(\theta))$.

By definition of $\widetilde{K}^X(\theta)$ we have $\mathscr{L}\text{og}\, W = N' + \widetilde{K}^X(\theta)$ for some $N' \in \mathscr{L}$. Since $\mathscr{L}\text{og}\, W = \frac{1}{W_-} \cdot M + \frac{1}{W_-} \cdot A$, the uniqueness of the canonical decomposition yields $A = W_- \cdot \widetilde{K}^X(\theta)$. Then $\Delta A = W_- \Delta \widetilde{K}^X(\theta)$ and we obtain

$$U = -\frac{1}{1 + \Delta \widetilde{K}^X(\theta)} \cdot \widetilde{K}^X(\theta).$$

Then $\Delta U = -\frac{\Delta \widetilde{K}^X(\theta)}{1+\Delta \widetilde{K}^X(\theta)}$ and $1+\Delta U = \frac{1}{1+\Delta \widetilde{K}^X(\theta)}$. Hence $D = 1/\mathscr{E}(U)$ and II.8.2 yield

$$\begin{aligned} D_t &= \exp\left(\frac{1}{1+\Delta \widetilde{K}^X(\theta)} \cdot \widetilde{K}^X(\theta)_t \right. \\ &\qquad \left. + \sum_{s \le t} \left(\log(1+\Delta \widetilde{K}^X(\theta)_s) - \frac{\Delta \widetilde{K}^X(\theta)_s}{1+\Delta \widetilde{K}^X(\theta)_s}\right)\right) \\ &= \exp\left(\widetilde{K}^X(\theta)_t + \sum_{s \le t} \left(\log(1+\Delta \widetilde{K}^X(\theta)_s) - \Delta \widetilde{K}^X(\theta)_s\right)\right) \\ &= \mathscr{E}(\widetilde{K}^X(\theta))_t, \end{aligned}$$

and we are finished. □

7.15 **Remarks.** 1) Similarly to Theorem II.2.49, one could prove that a process X is a semimartingale with characteristics (B, C, ν) if and only if its cumulant processes $\widetilde{K}^X(\theta)$ are given by 7.5 and 7.6 for the constant processes $\theta \equiv \lambda$, for all λ in an open set U of $\mathbb{R}^d$: this of course supposes that 7.6 makes sense, that is that the processes $e^{\lambda \cdot x} 1_{\{|x|>1\}} \star \nu$ are in $\mathscr{V}$ for all $\lambda \in U$ (in addition to the usual requirements on characteristics of semimartingales for (B, C, ν)).

2) One could also define the cumulant process for integrands θ whose components are complex-valued; if $\theta = iu$ for $u \in \mathbb{R}^d$, one sees that $\widetilde{K}^X(\theta) = A(u)$, as given by II.2.40 (compare II.2.40 with 7.5–7.6).

3) When X is a PII the cumulant processes $\widetilde{K}^X(\theta)$ and $K^X(\theta)$ are deterministic as soon as θ is a deterministic process. The converse holds as well. In this case if $\theta \equiv \lambda$ for some $\lambda \in \mathbb{R}^d$, we have under appropriate integrability conditions (in fact, the property $\lambda \in U$ for U as in (1) above):

$$E(e^{\lambda \cdot X_t}) = \mathscr{E}[\widetilde{K}^X(\lambda)] = e^{K^X(\lambda)}.$$

§ 7b. Esscher Change of Measure

So far we have considered two probability measures P and P' and studied how the characteristics of a semimartingale X under P are modified when we replace P by P'. The point of view here is slightly different: we start with a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ endowed with a d–dimensional semimartingale X with characteristics (B, C, ν); we want to construct another measure P', locally equivalent to P, in such a way that X has some specified properties under P', like being a local martingale.

This problem amounts of course to constructing the density process $Z_t = dP'_t/dP_t$ in such a way that the new characteristics (B', C', ν') of X under P'

have some given properties, using the connection between characteristics and the density process provided by Girsanov's Theorem 3.24: this is in general a difficult "martingale problem" (see, for example, § 4c and § 5a,b).

1) A "natural" way of solving this problem has been introduced by Esscher [264], in connection with some actuarial problems. We first illustrate this method on a simple "random variable" problem, where no filtration is involved: Suppose that X is a real-valued random variable given on a probability space $(\Omega, \mathscr{F}, P)$ and such that $P(X > 0) > 0$ and $P(X < 0) > 0$. The problem is to construct a measure P' equivalent to P and such that $E_{P'}[X] = 0$.

Practically speaking, the idea of Esscher is as follows: we first construct the measure $Q \sim P$ by

$$Q(d\omega) = ce^{-X(\omega)^2} P(d\omega),$$

where c is the norming constant $c = 1/E[e^{-X^2}]$). Then put $\varphi(\theta) = E_Q[e^{\theta X}]$ for $\theta \in \mathbb{R}$, and

7.16
$$Z_\theta(\omega) = \frac{e^{\theta X(\omega)}}{\varphi(\theta)} \quad \Big(= e^{\theta X(\omega) - K(\theta)} \quad \text{with } K(\theta) = \log \varphi(\theta)\Big).$$

($x \mapsto e^{\theta x}/\varphi(\theta)$ is called the Esscher's transform.) It is obvious from that construction that $E_Q[Z_\theta] = 1$ and that $Q \sim P$.

Now we construct the measures P'_θ by

$$P'_\theta(d\omega) = Z_\theta(\omega) Q(d\omega) = \frac{e^{\theta X(\omega)}}{\varphi(\theta)} Q(d\omega).$$

It is clear that

$$E_{P'_\theta}[X] = E_Q\left(\frac{Xe^{\theta X}}{\varphi(\theta)}\right) = \frac{\varphi'(\theta)}{\varphi(\theta)}.$$

where φ' is the derivative of φ. The assumptions $P(X > 0) > 0$ and $P(X < 0) > 0$ yield that the strictly convex function φ reaches its minimum at a unique point θ', so the equation $\varphi'(\theta) = 0$ as θ' for its unique solution. Then we define $P' = P'_{\theta'}$: we have $P' \sim P$ and $E_{P'}(X) = 0$.

2) Now we are going to the Esscher's change of measures in the semimartingale setting. We have a given d–dimensional semimartingale $X = (X^1, \dots, X^d)$ on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with characteristics (B, C, ν) relative to some truncation function h and with jump measure μ.

We will intensively use the Laplace cumulant processes $K^X(\theta)$ and $\tilde{K}^X(\theta)$ defined in Definition 7.2. By analogy with 7.16 and using the interpretation of $K^X(\theta)$ given in Proposition 7.14, we introduce (under the assumptions that $\theta \in L(X)$ and that $\theta \cdot X$ is exponentially special) the following process Z^θ:

7.17
$$Z^\theta = \exp\{\theta \cdot X - K^X(\theta)\}.$$

By definition the process Z^θ is a positive local martingale with $Z^\theta_0 = 1$.

If further Z^θ is a uniformly integrable martingale, we may define the probability measure $P^\theta(d\omega) = P(d\omega)Z^\theta_\infty(\omega)$, and the density process of P^θ w.r.t. P is obviously Z^θ.

Now we want to the processes θ such that the exponential process e^X or the process X itself are local martingales with respect to P^θ.

7.18 **Theorem.** *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is an exponentially special semimartingale and that the process Z^θ defined in 7.17 is a uniformly integrable martingale, and let P^θ be as above. Let also $\theta^{(i)}$ be the process with* jth *component equal to θ^j if $j \neq i$ and $\theta^i + 1$ if $j = i$. Then the process e^{X^i} is a P^θ–local martingale if and only if $\theta^{(i)} \cdot X$ is exponentially special and if we have P–almost surely*

7.19
$$K^X(\theta^{(i)}) = K^X(\theta).$$

Proof. By 3.8, e^{X^i} is a P^θ–local martingale iff $Z^\theta e^{X^i} = e^{\theta^{(i)} \cdot X - K^X(\theta)}$ is a P–local martingale. By Propositions II.8.27 and 7.14, this is the case iff $\theta^{(i)} \cdot X$ is an exponentially special semimartingale and $K^X(\theta^{(i)}) = K^X(\theta)$: hence the result. □

7.20 *Example.* Suppose that X is a 1–dimensional PIIS with characteristics (b, c, F). Consider also a number $\theta \in \mathbb{R}$ (no longer a process, in this example). Under the assumption

7.21
$$\int_{\{|x|>1\}} e^{\theta x}\, F(dx) < \infty,$$

the process θX is exponentially special (see II.8.26), and by Theorem 7.4 we have $K^X(\theta)_t = \varkappa(\theta)t$, where

$$\varkappa(\theta) = \theta b + \frac{1}{2}\theta^2 c + \int (e^{\theta x} - 1 - \theta h(x))F(dx).$$

Suppose that 7.21 holds for all $\theta \in \mathbb{R}$. Then 7.19 writes as $\varkappa(\theta + 1) = \varkappa(\theta)$, i.e.

7.22
$$b + c(\tfrac{1}{2} + \theta) + \int (e^{\theta x}(e^x - 1) - h(x))F(dx) = 0.$$

If this equation has a solution (say, θ^*) then it is unique and from Theorem 7.3 we get that e^X is a local martingale under P^{θ^*}. Further, the process X is still a PIIS under P^{θ^*}: to see this, check (using Theorem 7.23 below) that the characteristics of X under P^{θ^*} are still deterministic. □

The next result is an "explicit" expression of the process Z^θ in terms of the original process X.

7.23 **Theorem.** a) *The process Z^θ above can be written as*

7.24
$$Z^\theta = \mathscr{E}\Big(\theta\cdot X^c + W^\theta * (\mu-\nu)\Big),$$
where

7.25
$$W^\theta(\omega,t,x) = \frac{e^{\theta_t(\omega).x}-1}{1+\Delta\widetilde{K}^X_t(\theta)(\omega)}.$$

b) *If Z^θ is a uniformly integrable martingale, the characteristics $(B^\theta, C^\theta, \nu^\theta)$ of the semimartingale X with respect to the probability $P^\theta(d\omega) = P(d\omega)\,Z^\theta(\omega)$ are given by*

7.26
$$\begin{cases} B^{\theta,i} = B^i + \left(\sum_{j\le d} c^{ij}\theta^j\right)\cdot A + h^i(x)(Y^\theta-1)*\nu, \\ C^\theta = C, \\ \nu^\theta = Y^\theta\cdot\nu, \end{cases}$$

where $Y^\theta = W^\theta + \frac{1}{1+\Delta\tilde{K}^X(\theta)}$ and A and c are as in 7.1.

Proof. a) We have $Z^\theta = e^V$, where $V = \theta\cdot X - K^X(\theta)$. Then if $\bar{V} = \mathscr{L}\mathrm{og}\,(e^V)$, we have $Z^\theta = \mathscr{E}(\bar{V})$. Since Z^θ is a local martingale, then so is $\bar{V}$, and thus $\bar{V}$ is entirely characterized by its continuous martingale part $\bar{V}^c$ and by its jumps $\Delta\bar{V}$. Now II.8.12 and 6.22 yield $\theta\in L(X^c)$ and

7.27
$$\bar{V}^c = V^c = (\theta\cdot X)^c = \theta\cdot X^c.$$

Next, II.8.13 and 7.10 and 6.10 give

$$\Delta\bar{V} = e^{\Delta V}-1 = e^{\Delta(\theta\cdot X)-\Delta K^X(\theta)}-1 = \frac{e^{\theta\cdot\Delta X}}{1+\Delta\widetilde{K}^X(\theta)}-1.$$

But with W^θ as in 7.25 and $\widetilde{W^\theta}$ as in II.1.27 (with the measure μ), we readily check (use 7.9) that $\widetilde{W^\theta} = \Delta\bar{V}$. Since $\bar{V}$ is a local martingale, we deduce that W^θ is integrable w.r.t. $\mu-\nu$ and that $W^\theta \star (\mu-\nu)$ is equal to $\bar{V}-\bar{V}^c$. This and 7.27 give 7.24.

b) We use Theorem 3.24. All we need to check is the validity of 3.28 with $Z = Z^\theta$ and with $Y = Y^\theta$ as in our theorem. Since $Z^\theta = 1 + Z^\theta_-\cdot\bar{V}$, the second formula 3.28 follows from 7.27 and 6.5. We also have $\Delta Z^\theta = Z^\theta_-\Delta\bar{V}$, so the first formula 3.28 can be written as $M^P_\mu(\Delta\bar{V}|\widetilde{\mathscr{P}}) = Y^\theta-1$. Since $\Delta\bar{V} = W^\theta(t,\Delta X_t)1_{\{\Delta X_t\neq 0\}} - \frac{\Delta\tilde{K}^X(\theta)}{1+\Delta\tilde{K}^X(\theta)}$, we obtain $M^P_\mu(\Delta\bar{V}|\widetilde{\mathscr{P}}) = W^\theta - \frac{\Delta\tilde{K}^X(\theta)}{1+\Delta\tilde{K}^X(\theta)}$, which equals $Y^\theta-1$. □

7.28 **Theorem.** *Let $\theta\in L(X)$ be such that $\theta\cdot X$ is exponentially special and such that Z^θ is a uniformly integrable martingale, and consider the probability $P^\theta(d\omega) = P(d\omega)Z^\theta(\omega)$. Then X is P^θ–local martingale if and only if we have the following three properties:*

7.29 $$\left(e^{\theta\cdot x}\,(|x|^2 \bigwedge |x|)\right) \star \nu \in \mathscr{V},$$

and, up to a $P(d\omega)dA_t(\omega)$ null set:

7.30 $$b_t^i + \sum_{j\le d} c_t^{ij}\theta_t^j + \int \left(\frac{e^{\theta_t\cdot x}}{1+\Delta\widetilde{K}^X(\theta)_t}\, x^i - h^i(x)\right) F_t(dx) = 0,$$

7.31 $$\Delta A_t > 0 \quad \Longrightarrow \quad \int x\, e^{\theta_t\cdot x}\, F_t(dx) = 0.$$

Proof. By 7.26 the characteristics $(B^\theta, C^\theta, \nu^\theta)$ of X) with respect to P^θ have the form

$$B^\theta = b^\theta \cdot A, \quad C^\theta = c^\theta \cdot A, \quad \nu^\theta(\omega, dt, dx) = dA_t(\omega) F^\theta_{\omega,t}(dx),$$

where

7.32 $$\begin{cases} b_t^{\theta,i} = b_t^i + \sum_{j\le d} c_t^{ij}\theta_t^j + \int h^i(x)\left(\frac{e^{\theta_t\cdot x}}{1+\Delta\widetilde{K}^X(\theta)_t} - 1\right) F_t(dx), \\ c^\theta = c, \\ F_t^\theta(dx) = \frac{e^{\theta_t\cdot x}}{1+\Delta\widetilde{K}^X(\theta)_t}\, F_t(dx), \end{cases}$$

So, since $P^\theta \sim P$, we deduce from 6.35 that S is a P^θ–local martingale iff we have 7.30 and the following two properties:

7.33 $$\left(\frac{e^{\theta_t\cdot x}}{1+\Delta\widetilde{K}^X(\theta)_t}\,(|x|^2 \bigwedge |x|)\right) \star \nu \in \mathscr{V},$$

and also outside a $P(d\omega)dA_t(\omega)$–null set the following two properties:

7.34 $$\Delta A_t > 0 \quad \Longrightarrow \quad \int x \frac{e^{\theta_t\cdot x}}{1+\Delta\widetilde{K}^X(\theta)_t}\, F_t(dx) = 0.$$

But the process $\frac{1}{1+\Delta\widetilde{K}^X(\theta)}$ is predictable positive and equals 1 except on a discrete set, so it is locally bounded: then 7.33 and 7.34 are clearly equivalent to 7.29 and 7.31 respectively. □

7.35 *Example.* Consider again Example 7.20: X is a 1–dimensional PIIS with characteristics (b, c, F), and θ is a (non-random) real. Then $\widetilde{K}^X(\theta)$ is deterministic and so in Theorem 7.24 we see that Z^θ is the stochastic exponential of a PIIS, and also under P^θ the process X is again a PIIS. The conditions in Theorem 7.28 amount to say that $x \mapsto e^{\theta x}\,|x|1_{\{|x|>1\}}$ is F–integrable and to

7.36 $$b + c\theta + \int (xe^{\theta x} - h(x))F(dx) = 0.$$

So, the PIIS X with characteristics (b, c, F) is P^θ–local martingale if and only if θ is a root of Equation 7.36 and the above integrability condition holds.

Chapter IV. Hellinger Processes, Absolute Continuity and Singularity of Measures

The question of absolute continuity or singularity (ACS) of two probability measures has been investigated a long time ago, both for its theoretical interest and for its applications to mathematical statistics. S. Kakutani in 1948 [125] was the first to solve the ACS problem in the case of two measures P and P' having a (possibly infinite) product form: $P = \mu_1 \otimes \mu_2 \otimes \cdots$ and $P' = \mu'_1 \otimes \mu'_2 \otimes \cdots$, when $\mu_n \sim \mu'_n$ (μ_n and μ'_n are equivalent) for all n; he proved a remarquable result, known as the "Kakutani alternative", which says that either $P \sim P'$, or $P \perp P'$ (P and P' are mutually singular). Ten years later, Hajek [80] and Feldman [53] proved a similar alternative for Gaussian measures, and several authors gave effective criteria in terms of the covariance functions or spectral quantities, for the laws of two Gaussian processes.

With the development of martingale theory and stochastic calculus, the study of ACS problems was made possible for a wide class of stochastic processes: diffusions, point processes, semimartingales, ... The results obtained, in turn, illustrate the power of the general theory of stochastic processes and of stochastic integration theory.

In this chapter, we study the ACS problem for two measures P and P' defined on a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$. We seek for so-called "predictable criteria" (i.e., expressed in terms of predictable processes), especially because when P and P' are laws of semimartingales, these predictable criteria can often be expressed in terms of the characteristics of the two processes (for instance, if the processes are solutions of stochastic differential equations, they are expressed in terms of the coefficients of the equations).

In Section 1 we introduce the tools to be used later. Recall that the ACS problem for two measures can be attacked (in a very elementary way) through the "Kakutani-Hellinger metric" and the "Hellinger integrals". As we consider a filtered space, we introduce a family of predictable processes, the "Hellinger processes", which allow us to visualize how the Hellinger integrals vary in function of time (i.e., when considering the Hellinger integrals of the restrictions of the two measures to every σ-field $\mathscr{F}_t$).

Section 2 is devoted to stating and proving the main "general" results. For these first two sections we heavily use the stochastic calculus as introduced in Chapters I and II, but we use nothing of Chapter III except the very simple §§ III.3a,b.

Then in Sections 3 and 4 we show how it is possible to explicitely compute the Hellinger processes in concrete situations, such as the case of martingale problems introduced in § III.2. We give some examples of such computations and of ACS criteria for point processes, diffusion processes, and processes with independent increments.

1. Hellinger Integrals and Hellinger Processes

§ 1a. Kakutani-Hellinger Distance and Hellinger Integrals

Let $(\Omega, \mathscr{F})$ be a measurable space endowed with two probability measures P and P'; let Q be a third probability measure such that

1.1 $$P \ll Q, \quad P' \ll Q.$$

We consider the Radon-Nikodym derivatives

1.2 $$z = \frac{dP}{dQ}, \quad z' = \frac{dP'}{dQ}.$$

The expectations with respect to P, P', Q are denoted by E_P, $E_{P'}$, E_Q.

The *Kakutani-Hellinger distance $\rho(P, P')$ between P and P'* is the nonnegative number whose square is

1.3 $$\rho^2(P, P') = \frac{1}{2} \int_\Omega (\sqrt{dP} - \sqrt{dP'})^2,$$

which is a (loose) way of writing

1.4 $$\rho^2(P, P') = \frac{1}{2} \int_\Omega \left(\sqrt{\frac{dP}{dQ}} - \sqrt{\frac{dP'}{dQ}} \right)^2 dQ = \frac{1}{2} E_Q[(\sqrt{z} - \sqrt{z'})^2]$$

(1.3 makes clear that $\rho(P, P')$ does not depend on the measure Q satisfying 1.1: we shall see that presently). Therefore

1.5 $$\rho^2(P, P') = 1 - H(P, P'),$$

where $H(P, P')$ is the *Hellinger integral* of P and P', defined by

1.6 $$H(P, P') = E_Q(\sqrt{zz'}),$$

or, with the "notation" 1.3: $H(P, P') = \int_\Omega \sqrt{dP\, dP'}$.

For technical reasons (the first of them being the resolution of "absolute continuity-singularity problems"), we also introduce the *Hellinger integral of order α* by

1.7 $$H(\alpha; P, P') = E_Q(z^\alpha z'^{1-\alpha}),$$

where $\alpha \in (0, 1)$. We then have $H(P, P') = H(\frac{1}{2}; P, P')$.

1.8 **Lemma.** a) *The Hellinger integral of order α, defined by 1.7, does not depend upon the measure Q satisfying 1.1* (hence $\rho(P,P')$ does not depend upon Q either).

b) *The Kakutani-Hellinger distance ρ defined by 1.3* (or 1.4) *is a distance on the set of all probability measures on $(\Omega, \mathscr{F})$.*

We shall see in Chapter V that the topology defined by the distance ρ is the topology of convergence in variation.

Proof. a) It suffices to prove that if Q satisfies 1.1 and $\bar{Q}$ is another probability measure with $Q \ll \bar{Q}$ (so $\bar{Q}$ meets 1.1 as well), then $E_Q(z^\alpha z'^{1-\alpha}) = E_{\bar{Q}}(\bar{z}^\alpha \bar{z}'^{1-\alpha})$, where $\bar{z} = dP/d\bar{Q}$ and $\bar{z}' = dP'/d\bar{Q}$. Let $Z = dQ/d\bar{Q}$, so that $\bar{z} = Zz$ and $\bar{z}' = Zz'$. Hence

$$E_Q(z^\alpha z'^{1-\alpha}) = E_{\bar{Q}}(Zz^\alpha z'^{1-\alpha}) = E_{\bar{Q}}(\bar{z}^\alpha \bar{z}'^{1-\alpha}).$$

b) If $\rho(P,P') = 0$, then $z = z'$ Q-a.s., so $P' = P$. Let P, P', P'' be three probability measures on $(\Omega, \mathscr{F})$, and Q another measure such that $P \ll Q$, $P' \ll Q$, $P'' \ll Q$, and $z = dP/dQ$, $z' = dP'/dQ$, $z'' = dP''/dQ$. The triangular inequality in $L^2(Q)$ yields

$$[E_Q(\sqrt{z} - \sqrt{z''})^2]^{1/2} \leq [E_Q(\sqrt{z} - \sqrt{z'})^2]^{1/2} + [E_Q(\sqrt{z'} - \sqrt{z''})^2]^{1/2}$$

and $\rho(P,P'') \leq \rho(P,P') + \rho(P',P'')$ follows. □

In particular, to compute $H(\alpha; P, P')$ we can use the measure

1.9 $$Q = \tfrac{1}{2}(P + P'),$$

which obviously satisfies 1.1. In this case, there is a version of z and z' such that the following equality holds identically:

1.10 $$z + z' = 2.$$

1.11 **Lemma.** *Assume 1.1. Then*

a) *The following are equivalent:*
- (i) $P' \ll P$
- (ii) $P'(z > 0) = 1$
- (iii) $H(\alpha; P, P') \to 1$ *as* $\alpha \downarrow 0$.

b) *The following are equivalent:*
- (i) $P' \perp P$ (i.e.: P and P' are singular).
- (ii) $P'(z > 0) = 0$
- (iii) $H(\alpha; P, P') \to 0$ *as* $\alpha \downarrow 0$
- (iv) $H(\alpha; P, P') = 0$ *for all* $\alpha \in (0,1)$.
- (v) $H(\alpha; P, P') = 0$ *for some* $\alpha \in (0,1)$.

Proof. From the definitions of z and z',

$$P(z = 0) = E_Q(z 1_{\{z=0\}}) = 0$$

$$P'(A \cap \{z > 0\}) = E_Q(z' 1_{A \cap \{z>0\}}) = E_Q\left(z' \frac{z}{z} 1_{A \cap \{z>0\}}\right) = E_P\left(\frac{z'}{z} 1_{A \cap \{z>0\}}\right), \quad \text{for}$$

$A \in \mathscr{F}$. The equivalences (i)$\Leftrightarrow$(ii) easily follow in both cases a) and b). Moreover, $z^\alpha z'^{1-\alpha} \to z' 1_{\{z>0\}}$ as $\alpha \downarrow 0$, and $0 \le z^\alpha z'^{1-\alpha} \le \alpha z + (1-\alpha)z'$, which is Q-integrable. Hence Lebesgue convergence theorem yields

$$\lim_{\alpha \downarrow 0} H(\alpha; P, P') = E_Q(z' 1_{\{z>0\}}) = P'(z > 0),$$

and so (ii)$\Leftrightarrow$(iii) follows in both cases a) and b).

Finally, $H(\alpha; P, P') = E_{P'}[(z/z')^\alpha 1_{\{z'>0\}}]$ and $P'(z' > 0) = 1$, hence for each $\alpha \in (0,1)$ we have the equivalence: $P'(z > 0) = 0 \Leftrightarrow H(\alpha; P, P') = 0$. Thus (ii)$\Leftrightarrow$(iv)$\Leftrightarrow$(v) in (b) follows. □

§ 1b. Hellinger Processes

1. For the remainder of this section, we consider a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ with $\mathscr{F} = \mathscr{F}_{\infty-}$ and two fixed probability measures P, P' on $(\Omega, \mathscr{F})$.

Let Q be another probability measure on $(\Omega, \mathscr{F})$. Instead of 1.1, and for reasons that will become apparent later on, we only assume (see III.3.2)

1.12 $$P \overset{\text{loc}}{\ll} Q, \quad P' \overset{\text{loc}}{\ll} Q.$$

We call z and z' the *density processes* of P and P', relative to Q (see III.3.4): they are Q-martingales, and

1.13 $$\begin{cases} T \quad \text{stopping time} \Rightarrow z_T = dP_T/dQ_T \text{ and } z'_T = dP'_T/dQ_T \text{ on } \{T < \infty\} \\ T \quad \text{predictable time} \Rightarrow z_{T-} = dP_{T-}/dQ_{T-} \\ \qquad\qquad\qquad\qquad \text{and } z'_{T-} = dP'_{T-}/dQ_{T-} \text{ on } \{T < \infty\} \end{cases}$$

where P_T, P'_T, Q_T (resp. P_{T-}, P'_{T-}, Q_{T-}) are the restrictions of P, P', Q to $\mathscr{F}_T$ (resp. $\mathscr{F}_{T-}$).

Moreover:

1.14 If 1.1 holds, then z and z' are Q-uniformly integrable martingales, and 1.13 holds on Ω and not only on $\{T < \infty\}$ (because $z_\infty = z_{\infty-}$ and $\{T = \infty\} \cap \mathscr{F}_{T-} = \{T = \infty\} \cap \mathscr{F}_T = \{T = \infty\} \cap \mathscr{F}$). □

Set

1.15 $$\begin{cases} R_n = \inf(t: z_t < 1/n), \quad R = \lim_n \uparrow R_n, \quad \Gamma = \bigcup_n [\![0, R_n]\!] \\ R'_n = \inf(t: z'_t < 1/n), \quad R' = \lim_n \uparrow R'_n, \quad \Gamma' = \bigcup_n [\![0, R'_n]\!] \\ S_n = R_n \wedge R'_n, \quad S = R \wedge R' = \lim_n \uparrow S_n, \quad \Gamma'' = \Gamma \cap \Gamma' = \bigcup_n [\![0, S_n]\!], \end{cases}$$

and recall from III.3.6 that, up to a Q-evanescent set,

1.16
$$\Gamma = \{z_- > 0\} \cup [\![0]\!],\ [\![0, R[\![= \{z > 0\};$$
$$\Gamma' = \{z'_- > 0\} \cup [\![0]\!],\ [\![0, R'[\![= \{z' > 0\}.$$

1.17 **Lemma.** *Let $\alpha \in (0,1)$. Then the process $Y(\alpha) = z^\alpha z'^{1-\alpha}$ is a Q-supermartingale, and the stopped process $Y(\alpha)^t$ is of class* (D) *for all $t \in \mathbb{R}_+$. If furthermore 1.1 holds, $Y(\alpha)$ itself is of class* (D).

Proof. $0 \le Y(\alpha) \le \alpha z + (1-\alpha)z'$, while for all $t \in \mathbb{R}_+$ the stopped Q-martingales z^t and z'^t are of class (D) (resp., z and z' are of class (D), under 1.1): hence $Y(\alpha)$ has the same properties.

Moreover, the function $(u,v) \rightsquigarrow u^\alpha v^{1-\alpha}$ being concave on $\mathbb{R}_+^2$, Jensen's inequality yields for $s \le t$ (see [188], II.6.1):
$$E_Q(z_t^\alpha z_t'^{1-\alpha} | \mathscr{F}_s) \le E_Q(z_t | \mathscr{F}_t)^\alpha E_Q(z'_t | \mathscr{F}_s)^{1-\alpha} = z_s^\alpha z_s'^{1-\alpha}. \qquad \square$$

1.18 **Theorem.** *Let $\alpha \in (0,1)$ and $Y(\alpha) = z^\alpha z'^{1-\alpha}$. There exists a predictable increasing $\overline{\mathbb{R}}_+$-valued process $h(\alpha)$, unique up to Q-indistinguishability, which meets $h(\alpha)_0 = 0$ and the following two conditions*:

1.19
$$h(\alpha) = 1_{\Gamma''} \cdot h(\alpha)$$

1.20
$$M(\alpha) = Y(\alpha) + Y(\alpha)_- \cdot h(\alpha) \text{ is a } Q\text{-martingale.}$$

If moreover 1.1 holds, then $M(\alpha)$ is Q-uniformly integrable.

Proof. By I.3.15 and a localization at arbitrarily large fixed times, we get a Q-martingale $M(\alpha)$ and an increasing finite-valued predictable process $A(\alpha)$ with $A(\alpha)_0 = 0$ and $M(\alpha) = Y(\alpha) + A(\alpha)$ (Doob-Meyer decomposition). Moreover, under 1.1, $Y(\alpha)$ and hence $M(\alpha)$ are of class (*D*).

$Y(\alpha) = 0$ on $[\![S, \infty[\![$ and $Y(\alpha)_{S_n} \to 0$ as $n \uparrow \infty$ on $\bigcap_n \{S_n < S\}$: we deduce that $1_{\Gamma''^c} \cdot Y(\alpha)$, which is the limit of $1_{]\!]S_n, \infty]\!]} \cdot Y(\alpha) = Y(\alpha) - Y(\alpha)^{S_n}$, equals 0. Thus $1_{\Gamma''^c} \cdot M(\alpha) = 1_{\Gamma''^c} \cdot A(\alpha)$ and, since a predictable local martingale with finite variation is constant (see I.3.16), we obtain

1.21
$$1_{\Gamma''^c} \cdot A(\alpha) = 0 \quad Q\text{-a.s.}$$

Moreover $Y(\alpha)_- > 0$ on $\Gamma'' \cap]\!]0, \infty[\![$. Then the process $h(\alpha) = \left(\frac{1}{Y(\alpha)_-} 1_{\Gamma''}\right) \cdot A(\alpha)$ meets all the claimed properties (1.20 follows from $A(\alpha) = Y(\alpha)_- \cdot h(\alpha)$, which itself follows from 1.21).

Finally, the uniqueness of $h(\alpha)$ follows from the uniqueness of Doob-Meyer decomposition and from 1.19 and the fact that $Y(\alpha)_- > 0$ on $\Gamma'' \cap]\!]0, \infty[\![$. $\square$

1.22 **Theorem.** *The process $h(\alpha)$ of 1.18 does not depend upon the measure Q satisfying 1.12, in the following sense*: *if $\bar{Q}$ is another measure with $Q \overset{\text{loc}}{\ll} \bar{Q}$* (hence

$\bar{Q}$ meets 1.12 as well) *and if* $h(\alpha)$ *and* $\bar{h}(\alpha)$ *are the processes computed through* Q *and* $\bar{Q}$, *then* $h(\alpha)$ *and* $\bar{h}(\alpha)$ *are* Q-*indistinguishable.*

In particular, no matter which Q is used, $h(\alpha)$ is unique up to a P- and P'-evanescent set (observe, though, that $h(\alpha)$ and $\bar{h}(\alpha)$ above are not necessarily $\bar{Q}$-indistinguishable).

Proof. Let Z be the density process of Q with respect to $\bar{Q}$; let $\bar{z}$ and $\bar{z}'$ be the density processes of P and P' with respect to $\bar{Q}$. Hence $\bar{z} = Zz$, $\bar{z}' = Zz'$, and $\bar{Y}(\alpha) := \bar{z}^{\alpha}\bar{z}'^{1-\alpha} = ZY(\alpha)$. Put $A(\alpha) = Y(\alpha)_- \cdot h(\alpha)$, $\bar{A}(\alpha) = \bar{Y}(\alpha)_- \cdot \bar{h}(\alpha)$ and $M(\alpha) = Y(\alpha) + A(\alpha)$.

$A(\alpha)$ is predictable and increasing, so Ito's formula (I.4.49c) applied for $\bar{Q}$ gives $A(\alpha)Z = A(\alpha) \cdot Z + Z_- \cdot A(\alpha)$, and

$$\bar{Y}(\alpha) = ZM(\alpha) - ZA(\alpha) = ZM(\alpha) - A(\alpha) \cdot Z - Z_- \cdot A(\alpha).$$

Now, $A(\alpha) \cdot Z$ is a $\bar{Q}$-local martingale; $ZM(\alpha)$ is also a $\bar{Q}$-martingale by III.3.8a, and $Z_- \cdot A(\alpha) = \bar{Y}(\alpha)_- \cdot h(\alpha)$. Thus $\bar{Y}(\alpha) + \bar{Y}(\alpha)_- \cdot h(\alpha)$ as well as $\bar{Y}(\alpha) + \bar{Y}(\alpha)_- \cdot \bar{h}(\alpha)$ are $\bar{Q}$-local martingales. The uniqueness of the canonical $\bar{Q}$-decomposition of $\bar{Y}(\alpha)$ yields

1.23 $$\bar{Y}(\alpha)_- \cdot h(\alpha) = \bar{Y}(\alpha)_- \cdot \bar{h}(\alpha) \quad \bar{Q}\text{-a.s.}$$

Moreover $Q(\inf_{s \le t} Z_s > 0) = 1$ for all $t < \infty$, hence $\{\bar{Y}(\alpha)_- > 0\} \supset \Gamma'' \cap]\!]0, \infty[\![$ up to a Q-evanescent set. Integrating $\bar{Y}(\alpha)_-^{-1} 1_{\Gamma''}$ against 1.23 yields $h(\alpha) = \bar{h}(\alpha)$ up to a Q-evanescent set. □

1.24 **Definition.** a) The (P- and P'-unique) increasing predictable process $h(\alpha)$ constructed in 1.18 is called *Hellinger process in the strict sense, of order* α, *between* P *and* P'.

b) A *Hellinger process of order* α *between* P *and* P' is any increasing process $h'(\alpha)$ such that $1_{\Gamma''} h'(\alpha)$ and $1_{\Gamma''} h(\alpha)$ (or equivalently $1_{\Gamma''} \cdot h'(\alpha)$ and $1_{\Gamma''} \cdot h'(\alpha)$) are P- and P'-indistinguishable.

To emphazise the non-symmetric rôle of P and P', we write $h(\alpha; P, P')$ for these processes. □

1.25 **Remark.** We shall see in § 3c that if $(\Omega, \mathscr{F}, \mathbf{F})$ is the canonical space with the canonical d-dimensional càdlàg process X (see III.2.13 for the "canonical setting"), and if P and P' are two measures under which X is a PII, there are *versions of the Hellinger processes* $h(\alpha; P, P')$ *that are deterministic.* In general, this is not an easy result. However we sketch here a simple proof (as a corollary of Theorem 1.22), in the particular case where P and P' are locally mutually equivalent (i.e. $P' \overset{\text{loc}}{\ll} P$ and $P \overset{\text{loc}}{\ll} P'$).

In that case, we call Z the density process of P' relative to P. Then $Z > 0$ everywhere, so Theorem III.5.35 yields that Z is given by III.5.23, where A is deterministic and N is a PII relative to P; it easily follows that Z_t/Z_s is P-

independent from $\mathscr{F}_s$ for $s \le t$. Now, due to 1.22, we can use the measure $Q = P$ for computing $h(\alpha; P, P')$, so $z = 1$ and $z' = Z$ and $Y(\alpha) = Z^{1-\alpha}$ (notation of 1.18). Moreover the Hellinger integral $H(\alpha)_t = H(\alpha; P_t, P'_t)$ equals $E_P(Z_t^{1-\alpha})$, so $t \rightsquigarrow H(\alpha)_t$ is non-increasing, positive and càdlàg, with $H(\alpha)_0 = 1$. Furthermore if $M = Y(\alpha)/H(\alpha)$, we get for $s < t$:

$$\frac{H(\alpha)_t}{H(\alpha)_s} = \frac{1}{H(\alpha)_s} E_P\left[\left(\frac{Z_t}{Z_s}\right)^{1-\alpha} Z_s^{1-\alpha}\right]$$

$$= \frac{1}{H(\alpha)_s} E_P\left[\left(\frac{Z_t}{Z_s}\right)^{1-\alpha}\right] E_P(Z_s^{1-\alpha}) = E_P\left[\left(\frac{Z_t}{Z_s}\right)^{1-\alpha}\right]$$

$$E_P\left(\frac{M_t}{M_s}\middle|\mathscr{F}_s\right) = \frac{H(\alpha)_s}{H(\alpha)_t} E_P\left[\left(\frac{Z_t}{Z_s}\right)^{1-\alpha}\middle|\mathscr{F}_s\right] = \frac{H(\alpha)_s}{H(\alpha)_t} E_P\left[\left(\frac{Z_t}{Z_s}\right)^{1-\alpha}\right] = 1,$$

so M is a P-martingale. Moreover I.4.49c (Ito's formula) yields

$$Y(\alpha) = MH(\alpha) = H(\alpha) \cdot M + M_- \cdot H(\alpha)$$

and, comparing to 1.20 (here $\Gamma'' = \Omega \times \mathbb{R}_+$) we obtain $Y(\alpha)_- \cdot h(\alpha) = -M_- \cdot H(\alpha)$, so

1.26 $$h(\alpha) = -\frac{1}{H(\alpha)_-} \cdot H(\alpha)$$

and the claim follows. □

2. The previous remark, and also the name "Hellinger process", suggest a relationship between $h(\alpha; P, P')$ and the Hellinger integrals:

1.27 **Proposition.** *Let $h(\alpha)$ be any version of $h(\alpha; P, P')$, and $Y(\alpha) = z^\alpha z'^{1-\alpha}$. Then*

1.28 T stopping time $\Rightarrow H(\alpha; P_T, P'_T) = H(\alpha; P_0, P'_0) - E_Q[Y(\alpha)_- \cdot h(\alpha)_T]$

1.29 T predictable time $\Rightarrow H(\alpha; P_{T-}, P'_{T-}) = H(\alpha; P_0, P'_0) - E_Q[Y(\alpha)_- \cdot h(\alpha)_{T-}]$

Proof. The right-hand sides of 1.28 and 1.29 depend only upon the restrictions of $h(\alpha)$ to Γ'', hence we may assume that $h(\alpha)$ is the process of 1.18.

Assume first that T is bounded. Then $H(\alpha; P_T, P'_T) = E_Q[Y(\alpha)_T]$ by 1.13, so we deduce 1.28 from 1.20. Similarly $H(\alpha; P_{T-}, P'_{T-}) = E_Q[Y(\alpha)_{T-}]$, so 1.29 follows from 1.20 and from the property $E_Q[M(\alpha)_{T-}] = E_Q[M(\alpha)_T]$ when T is bounded and predictable.

We turn to the general case. 1.28 is true for $T \wedge n$, and $E_Q[Y(\alpha)_- \cdot h(\alpha)_{T \wedge n}] \uparrow E_Q[Y(\alpha)_- \cdot h(\alpha)_T]$. On the other hand, if we have $Q = \frac{P + P'}{2}$, then

$$H(\alpha; P_{T \wedge n}, P'_{T \wedge n}) = E_Q[Y(\alpha)_{T \wedge n}] \downarrow E_Q[Y(\alpha)_T] = H(\alpha; P_T, P'_T)$$

and thus 1.28 holds for T. 1.29 is proved similarly. □

Finally, we state a technical result.

1.30 **Lemma.** *Let $h(\alpha)$ be a version of $h(\alpha; P, P')$. Then, up to a Q-null set:*

a) $\Delta h(\alpha) \leq 1$ *on* Γ'', $\Delta h(\alpha) < 1$ *on* $[\![0, S[\![$.

b) *If T is a predictable time with $T \geq S$, then $\Delta h(\alpha)_T = 1$ on the set* $\bigcup_n \{T = S_n < \infty\}$.

Proof. We can assume without loss of generality that $h(\alpha)$ is the Hellinger process in the strict sense.

a) Let T be a predictable time. Then I.2.27 and 1.18 yield on $\{T < \infty\}$:

1.31 $$0 = E_Q[\Delta M(\alpha)_T | \mathscr{F}_{T-}] = E_Q[Y(\alpha)_T | \mathscr{F}_{T-}] - Y(\alpha)_{T-}[1 - \Delta h(\alpha)_T].$$

Since $Y(\alpha) \geq 0$, and $Y(\alpha)_- > 0$ on $\Gamma'' \cap]\!]0, \infty[\![$ we get $\Delta h(\alpha)_T \leq 1$ (recall that $\Delta h(\alpha) = 0$ outside Γ''). Hence that $\Delta h(\alpha) \leq 1$ up to a Q-evanescent set follows from the predictable section theorem I.2.18.

Now, $T = \inf(t: \Delta h(\alpha)_t = 1)$ is a predictable time; 1.31 gives that $E_Q[Y(\alpha)_T | \mathscr{F}_{T-}] = 0$ on $\{T < \infty\}$, hence $Y(\alpha)_T = 0$ Q-a.s. on $\{T < \infty\}$ and thus $T \geq S$ Q-a.s.: this finishes the proof of (a).

b) Let T be a predictable time, with $T \geq S$. Then $Y(\alpha)_T = 0$ on $\{T < \infty\}$ and so 1.31 yields $Y(\alpha)_{T-}[1 - \Delta h(\alpha)_T] = 0$ on $\{T < \infty\}$. Since $Y(\alpha)_{T-} > 0$ on the set $\bigcup_n \{T = S_n < \infty\}$, it follows that $\Delta h(\alpha)_T = 1$ (Q-a.s.) on this set. □

This implies in particular that the sets $\{\Delta h(\alpha; P, P') = 1\} \cap \Gamma''$ when α ranges through $(0, 1)$, are Q-a.s. equal (and they equal the "biggest predictable random set" included in $[\![S]\!] \cap \Gamma''$).

§ 1c. Computation of Hellinger Processes in Terms of the Density Processes

Although the properties stated in 1.18 do characterize the Hellinger process $h(\alpha)$, they do not give any "explicit" form for it. In this subsection we provide a way of computing $h(\alpha)$, in terms of the "characteristics" of the density processes z and z' of P and P' with respect to Q, without any kind of assumptions on P and P'. In the next section we make some restrictive assumptions on P and P' (as: they are the unique solutions of some martingale problems), thus allowing for a more explicit form for $h(\alpha)$ even.

The setting is the same than in § 1b. In particular, unless otherwise stated, we assume 1.12 only, regarding Q. We introduce a function $\varphi_\alpha: \mathbb{R}_+^2 \to \mathbb{R}_+$ (where $\alpha \in (0, 1)$), defined by:

1.32 $$\varphi_\alpha(u, v) = \alpha u + (1 - \alpha)v - u^\alpha v^{1-\alpha}.$$

1.33 **Theorem.** *We denote by z^c and z'^c the continuous martingale parts of z and z', relative to Q. We also denote by $\nu^{(z,z')}$ the third Q-characteristic of the bi-dimensional process (z, z').*

a) $\nu^{(z,z')}$ *only charges the* $\mathscr{P} \otimes \mathscr{R}^2$*-measurable set* $A = \{(\omega, t, x, y): t > 0, x \geq -z_{t-}(\omega), x = 0 \text{ if } z_{t-}(\omega) = 0, y \geq -z'_{t-}(\omega), y = 0 \text{ if } z'_{t-}(\omega) = 0\}$.

b) *A version of the Hellinger process* $h(\alpha; P, P')$ *is given by*

1.34
$$h(\alpha) = \frac{\alpha(1-\alpha)}{2}\left\{\frac{1}{z_-^2}\cdot\langle z^c, z^c\rangle - \frac{2}{z_- z'_-}\cdot\langle z^c, z'^c\rangle + \frac{1}{z'^2_-}\cdot\langle z'^c, z'^c\rangle\right\}$$
$$+ \varphi_\alpha\left(1 + \frac{x}{z_-}, 1 + \frac{y}{z'_-}\right) * \nu^{(z,z')}.$$

The last integral above stands for (with $0/0 = 0$):

$$\int_0^t \int_{\mathbb{R}^2} \varphi_\alpha\left(1 + \frac{x}{z_{s-}}, 1 + \frac{y}{z'_{s-}}\right)\nu^{(z,z')}(ds \times (dx, dy)).$$

Proof. a) Since $\Delta z + z_- \geq 0$ and $\Delta z = 0$ if $z_- = 0$ and $\Delta z' + z'_- \geq 0$ and $\Delta z' = 0$ if $z'_- = 0$, the random measure $\mu^{(z,z')}$ associated with the jumps of the bi-dimensional process (z, z') only charges A. Since A is predictable, this property carries over to $\nu^{(z,z')}$.

b) Fix $n \in \mathbb{N}^*$, and let f be a function of class C^2 on $\mathbb{R}^2$, such that $f(x, y) = x^\alpha y^{1-\alpha}$ for $x \geq 1/n$ and $y \geq 1/n$. Hence $Y(\alpha) = f(z, z')$ on $[\![0, S_n[\![$, and Ito's formula yields:

(1) $f(z^{S_n}, z'^{S_n})$
$$= Y(\alpha)_0 + \alpha\frac{Y(\alpha)_-}{z_-}\cdot z^{S_n} + (1-\alpha)\frac{Y(\alpha)_-}{z'_-}\cdot z'^{S_n}$$
$$+ \frac{\alpha(\alpha-1)}{2}\frac{Y(\alpha)_-}{z_-^2}\cdot\langle z^c, z^c\rangle^{S_n} + \alpha(1-\alpha)\frac{Y(\alpha)_-}{z_- z'_-}\cdot\langle z^c, z'^c\rangle^{S_n}$$
$$- \frac{(1-\alpha)\alpha}{2}\frac{Y(\alpha)_-}{z'^2_-}\cdot\langle z'^c, z'^c\rangle^{S_n}$$
$$+ \sum_{s \leq \cdot \wedge S_n}\left\{f(z_s, z'_s) - Y(\alpha)_{s-} - \alpha\frac{Y(\alpha)_{s-}}{z_{s-}}\Delta z_s - (1-\alpha)\frac{Y(\alpha)_{s-}}{z'_{s-}}\Delta z'_s\right\}.$$

If $s \leq S_n$, we also have by 1.32:

$$Y(\alpha)_s - Y(\alpha)_{s-} - \alpha\frac{Y(\alpha)_{s-}}{z_{s-}}\Delta z_s - (1-\alpha)\frac{Y(\alpha)_{s-}}{z'_{s-}}\Delta z'_s$$
$$= Y(\alpha)_{s-}\left\{\left(1 + \frac{\Delta z_s}{z_{s-}}\right)^\alpha\left(1 + \frac{\Delta z'_s}{z'_{s-}}\right)^{1-\alpha} - 1 - \alpha\frac{\Delta z_s}{z_{s-}} - (1-\alpha)\frac{\Delta z'_s}{z'_{s-}}\right\}$$
$$= -Y(\alpha)_{s-}\varphi_\alpha\left(1 + \frac{\Delta z_s}{z_{s-}}, 1 + \frac{\Delta z'_s}{z'_{s-}}\right),$$

and we set for simplicity

$$k(\alpha) = \frac{\alpha(1-\alpha)}{2}\left\{\frac{1}{z_-^2}\cdot\langle z^c, z^c\rangle - \frac{2}{z_- z'_-}\cdot\langle z^c, z'^c\rangle + \frac{1}{z'^2_-}\cdot\langle z'^c, z'^c\rangle\right\}.$$

Then (1) yields for $t < S_n$:

$$(2) \qquad Y(\alpha)_t^{S_n} = Y(\alpha)_0 + \alpha\frac{Y(\alpha)_-}{z_-}\cdot z_t^{S_n} + (1-\alpha)\frac{Y(\alpha)_-}{z'_-}\cdot z_t'^{S_n} - Y(\alpha)_-\cdot k(\alpha)_t^{S_n}$$
$$- Y(\alpha)_- 1_{[\![0,S_n]\!]}\varphi_\alpha\left(1+\frac{x}{z_-}, 1+\frac{y}{z'_-}\right) * \mu_t^{(z,z')}.$$

Now, the computation of $Y(\alpha)_s$ above precisely shows that the jumps of $Y(\alpha)^{S_n}$ and those of the process in the right-hand side of (2) are equal: since the latter equals $Y(\alpha)^{S_n}$ on $[\![0, S_n[\![$, we deduce that (2) holds for all $t \in \mathbb{R}_+$.

Now, consider (2). By 1.18, $Y(\alpha)^{S_n}$ is a special semimartingale. The two first stochastic integrals in (2) are Q-local martingales; $k(\alpha)^{S_n}$ is obviously in $\mathscr{A}^+_{\text{loc}}$ (for Q), and so is $Y(\alpha)_-\cdot k(\alpha)^{S_n}$. Hence the last process in (2), which has finite variation, ought to belong to $\mathscr{A}_{\text{loc}}$ by I.4.23. Therefore II.1.7 implies that its compensator under Q is $Y(\alpha)_- 1_{[\![0,S_n]\!]}\varphi_\alpha\left(1+\frac{x}{z_-}, 1+\frac{y}{z'_-}\right) * \nu^{(z,z')}$. We deduce that

$$Y(\alpha)^{S_n} + Y(\alpha)_-\cdot k(\alpha)^{S_n} + Y(\alpha)_- 1_{[\![0,S_n]\!]}\varphi_\alpha\left(1+\frac{x}{z_-}, 1+\frac{y}{z'_-}\right) * \nu^{(z,z')}$$

is a Q-local martingale. Comparing to 1.20, we deduce that the process $h(\alpha)$ defined by 1.34 coincides with the Hellinger process on each interval $[\![0, S_n]\!]$; since $\Gamma'' = \bigcup_n [\![0, S_n]\!]$, this proves the result. □

1.35 **Corollary.** *Assume that* $Q = \frac{P+P'}{2}$, *and call* ν^z *the third* Q*-characteristic of* z. *Then the Hellinger process of order* α *in the strict sense is*

$$1.36 \qquad h(\alpha) = \frac{\alpha(1-\alpha)}{2}\left(\frac{1}{z_-} + \frac{1}{z'_-}\right)^2\cdot\langle z^c, z^c\rangle + \varphi_\alpha\left(1+\frac{x}{z_-}, 1-\frac{x}{z'_-}\right) * \nu^z,$$

where $z' = 2 - z$.

Proof. Since $z + z' = 2$, we get $z^c + z'^c = 0$, which implies $\langle z'^c, z'^c\rangle = \langle z^c, z^c\rangle = -\langle z^c, z'^c\rangle$. Moreover $\Delta z' = -\Delta z$, thus if μ^z is the random measure associated with the jumps of z, we have

$$\int_0^t\int_{\mathbb{R}^2} W(\omega, s, (x,y))\mu^{(z,z')}(\omega; dt\times(dx,dy)) = \int_0^t\int_{\mathbb{R}} W(\omega, s, (x,-x))\mu^z(\omega; ds, dx)$$

and this relation carries over to the compensators $\nu^{(z,z')}$ and ν^z, whenever W is predictable. Hence 1.34 immediately takes the form 1.36.

To see that $h(\alpha)$ is actually the Hellinger process in the strict sense, we observe that $1_{\Gamma''^c} \cdot z = 0$, so neither $\langle z^c, z^c \rangle$ nor v^z charge Γ''^c: hence $h(\alpha) = 1_{\Gamma''} \cdot h(\alpha)$. □

1.37 **Corollary.** *Assume that $P' \overset{\text{loc}}{\ll} P$, and call Z the density process of P' with respect to P. Denote by Z^c the continuous martingale part of Z, relative to P, and ν^Z the third P-characteristic of Z. Then the Hellinger process of order α in the strict sense is*

1.38
$$h(\alpha) = \frac{\alpha(1-\alpha)}{2} \frac{1}{Z_-^2} \cdot \langle Z^c, Z^c \rangle + \left\{ \alpha + (1-\alpha)\left(1 + \frac{x}{Z_-}\right) - \left(1 + \frac{x}{Z_-}\right)^{1-\alpha} \right\} * \nu^Z$$

and in particular:

1.39
$$h\left(\frac{1}{2}\right) = \frac{1}{8} \frac{1}{Z_-^2} \cdot \langle Z^c, Z^c \rangle + \frac{1}{2} \left\{ 1 - \sqrt{1 + \frac{x}{Z_-}} \right\}^2 * \nu^Z.$$

Proof. We choose $Q = P$, hence $z = 1$ and $z' = Z$. Then $\langle z^c, z^c \rangle = \langle z^c, z'^c \rangle = 0$ and if μ^Z is the random measure associated with the jumps of Z, we have

$$\int_0^t \int_{\mathbb{R}^2} W(\omega, s, (x, y)) \mu^{(z,z')}(\omega; ds \times (dx, dy)) = \int_0^t \int_{\mathbb{R}} W(\omega, s, (0, y)) \mu^Z(\omega; ds, dy)$$

and this relation carries over to $\nu^{(z,z')}$ and ν^Z whenever W is predictable. Hence 1.34 has the form 1.38, and the property of $h(\alpha)$ to be the Hellinger process in the strict sense is proved as in 1.35. □

§ 1d. Some Other Processes of Interest

This subsection is purely technical in nature: we introduce a family of processes "which behave like Hellinger processes", and some of them will be used (like Hellinger processes again) for studying absolute continuity and contiguity, and also in Chapter X.

The setting is the same than in § 1b; in particular 1.12, 1.13, 1.15 hold.

We consider a function ψ: $\mathbb{R}_+ \to \mathbb{R}_+$ with the following property:

1.40 $\dfrac{\psi(x)}{|x-1|^2 \wedge |x-1|}$ is bounded (hence in particular $\psi(1) = 0$ and ψ is locally bounded),

and we set $\psi(+\infty) = 0$. We also use below the conventions $\frac{0}{0} = 0$ and $\frac{a}{0} = +\infty$ for $a \in (0, \infty]$. Then the next formula defines a "generalized increasing process" in the sense of III.5.8:

1.41
$$\jmath(\psi) = \sum_{s \leq \cdot} \frac{z'_s}{z'_{s-}} \psi\left(\frac{z_s/z_{s-}}{z'_s/z'_{s-}}\right).$$

1.42 Proposition. *Let ψ be as above. There exists a predictable increasing $\bar{\mathbb{R}}_+$-valued process $i(\psi)$, unique up to Q-indistinguishability, which has $i(\psi)_0 = 0$ and*

1.43 $$i(\psi) = 1_{\Gamma''} \cdot i(\psi)$$

1.44 $$j(\psi)^{S_n} - i(\psi)^{S_n} \text{ is a } Q\text{-local martingale for all } n \in \mathbb{N}^*.$$

Proof. a) Firstly, we prove that each stopped process $j(\psi)^{S_n}$ is Q-locally integrable. So we fix $n \in \mathbb{N}^*$, and we denote by K the supremum of $\psi(x)/(|x-1|^2 \wedge |x-1|)$. Let $B = \left\{|\Delta z| \leq \frac{1}{2n}\right\} \cap \left\{|\Delta z'| \leq \frac{1}{2n}\right\}$. We have $j(\psi)^{S_n} = F + G$, where

$$F_t = \sum_{s \leq t \wedge S_n} 1_B(s) \frac{z'_s}{z_s} \psi\left(\frac{z_s/z_{s-}}{z'_s/z'_{s-}}\right)$$

$$G_t = \sum_{s \leq t \wedge S_n} 1_{B^c}(s) \frac{z'_s}{z_s} \psi\left(\frac{z_s/z_{s-}}{z'_s/z'_{s-}}\right).$$

On $B \cap [\![0, S_n]\!]$ we have $z'/z'_- \geq 1/2$, hence

$$\frac{z'}{z'_-}\psi\left(\frac{z/z_-}{z'/z'_-}\right) \leq K\frac{z'}{z'_-}\left(\frac{z/z_-}{z'/z'_-} - 1\right)^2 = K\frac{(\Delta z/z_- - \Delta z'/z'_-)^2}{z'/z'_-}$$
$$\leq 4K\left[\left(\frac{\Delta z}{z_-}\right)^2 + \left(\frac{\Delta z'}{z'_-}\right)^2\right] \leq 4Kn^2(\Delta z^2 + \Delta z'^2).$$

Therefore $F_t \leq 4Kn^2(1_B \cdot [z,z]_t + 1_B \cdot [z',z']_t)$. So F is a finite-valued increasing process, and its jumps are bounded by $2K$: hence F is Q-locally integrable.

On the other hand, we easily deduce from I.4.56 that the two increasing processes $A_t = \sup_{s \leq t}|\Delta z_s|$ and $A'_t = \sup_{s \leq t}|\Delta z'_s|$ are Q-locally integrable. 1.40 again yields on $[\![0, S_n]\!]$:

$$\frac{z'}{z'_-}\psi\left(\frac{z/z_-}{z'/z'_-}\right) \leq K\frac{z'}{z'_-}\left|\frac{z/z_-}{z'/z'_-} - 1\right| = K\left|\frac{\Delta z}{z_-} - \frac{\Delta z'}{z'_-}\right| \leq Kn(|\Delta z| + |\Delta z'|).$$

Therefore if $T_0 = 0, \ldots, T_p = \inf(t > T_{p-1}: t \in B^c)$ we deduce that $G_{T_p \wedge t} \leq Knp(A_t + A'_t)$, and so the stopped process G^{T_p} is Q-locally integrable. Now, B^c is "discrete", so $T_p \uparrow \infty$ as $p \uparrow \infty$, and we deduce that G itself is Q-locally integrable. This finishes the proof that $j(\psi)^{S_n}$ is Q-locally integrable.

b) Because of what precedes, $j(\psi)^{S_n}$ admits a Q-compensator, say $i(\psi, n)$, which is obviously constant on $[\![S_n, \infty[\![$. It is also obvious that $i(\psi, n+1) = i(\psi, n)$ Q-a. s. on $[\![0, S_n]\!]$ (uniqueness of the compensator), hence

$$i(\psi) = \sum_{n \geq 1} 1_{]\!]S_{n-1}, S_n]\!]} \cdot i(\psi, n)$$

has the claimed properties. □

The above is the analogue of Theorem 1.18. Similarly, Theorem 1.22 has its counterpart (with essentially the same proof):

1.45 **Proposition.** *The process $i(\psi)$ of 1.42 does not depend upon the measure Q satisfying 1.12, in the following sense: if $\bar{Q}$ is another measure with $Q \overset{\text{loc}}{\ll} \bar{Q}$, and if $i(\psi)$ and $\bar{i}(\psi)$ are the processes computed through Q and $\bar{Q}$, then $i(\psi)$ and $\bar{i}(\psi)$ are Q-indistinguishable.*

Proof. $Z, \bar{z}, \bar{z}'$ denote the density processes of Q, P, P' with respect to $\bar{Q}$, so $\bar{z} = zZ$ and $\bar{z}' = z'Z$. We define $j(\psi)$ and $\bar{j}(\psi)$ by 1.41, and we use the notation 1.15: S_n and Γ'', $\bar{S}_n$ and $\bar{\Gamma}''$. Since $\inf_t Z_t > 0$ Q-a.s., we have $\Gamma'' = \bar{\Gamma}''$ Q-a.s., and so it suffices to prove that $i(\psi)^{S_n \wedge \bar{S}_{n'}} = \bar{i}(\psi)^{S_n \wedge \bar{S}_{n'}}$ Q-a.s. for all $n, n' \in \mathbb{N}$. Hence, upon stopping all processes at time $S_n \wedge \bar{S}_{n'}$ we may and will assume that $S_n \equiv \bar{S}_{n'} \equiv \infty$ for some $n, n' \in \mathbb{N}^*$. This implies in particular that $\bar{Q} \sim Q$.

Set $A = z'_- \cdot j(\psi)$ and $B = z'_- \cdot i(\psi)$, so 1.44 and $S_n \equiv \infty$ yield that $M = A - B$ is a Q-local martingale, and we call (T_p) a localizing sequence for Q. Since $\bar{Q} \sim Q$ we have $T_p \uparrow \infty$ $\bar{Q}$-a.s. as $p \uparrow \infty$, so III.3.8c implies that MZ is a $\bar{Q}$-local martingale. Ito's formula yields

$$ZA = Z \cdot A + A_- \cdot Z, \quad ZB = Z_- \cdot B + B \cdot Z,$$

hence

$$Z \cdot A = ZM - A_- \cdot Z + B \cdot Z + Z_- \cdot B,$$

and $ZM - A_- \cdot Z + B \cdot Z$ is a $\bar{Q}$-local martingale. So $Z_- \cdot B$ is the $\bar{Q}$-dual predictable projection of $Z \cdot A$. Furthermore, $Z \cdot A = \bar{z}'_- \cdot j(\psi)$ by 1.41, and $\bar{S}_{n'} \equiv \infty$, hence the characterization 1.42 of $\bar{i}(\psi)$ yields $\bar{z}'_- \cdot \bar{i}(\psi) = Z_- \cdot B = \bar{z}'_- \cdot i(\psi)$ $\bar{Q}$-a.s. Applying again that $\bar{S}_{n'} \equiv \infty$, we get $\bar{i}(\psi) = i(\psi)$ $\bar{Q}$-a.s., and we are finished. □

1.46 **Definition.** Let ψ be a function satisfying 1.40. We denote by $i(\psi; P, P')$ any increasing process $i'(\psi)$ such that $1_{\Gamma''} i'(\psi)$ and $1_{\Gamma''} i(\psi)$ (or equivalently, $1_{\Gamma''} \cdot i'(\psi)$ and $1_{\Gamma''} \cdot i(\psi)$) are $(P + P')$-indistinguishable, where $i(\psi)$ is the process constructed in 1.42. □

Due to 1.45, this definition makes sense and $1_{\Gamma''} \cdot i(\psi; P, P')$ is defined uniquely, up to a $(P + P')$-evanescent set, regardless of the dominating measure Q. We can compute $i(\psi; P, P')$ by formulae similar to 1.34 or 1.36 or 1.38:

1.47 **Proposition.** *Let ψ satisfy 1.40, and recall that $\psi(\infty) = 0$ by convention.*

a) *If $\nu^{(z,z')}$ denotes the third Q-characteristic of the bi-dimensional process (z, z'), a version of $i(\psi; P, P')$ is given by*

1.48 $$i(\psi) = \left(1 + \frac{y}{z'_-}\right) \psi\left(\frac{1 + x/z_-}{1 + y/z'_-}\right) * \nu^{(z,z')}$$

$$\left(\text{which stands for: } i(\psi)_t = \int_0^t \int_{\mathbb{R}^2} \left(1 + \frac{y}{z'_{s-}}\right) \psi\left(\frac{1 + x/z_{s-}}{1 + y/z'_{s-}}\right) \nu^{(z,z')}(ds \times (dx, dy))\right)$$

b) *If* $Q = \dfrac{P + P'}{2}$ *and if* ν^z *denotes the third* Q*-characteristic of* z, *a version of* $\imath(\psi; P, P')$ *is given by*

1.49
$$\imath(\psi) = \left(1 - \frac{x}{z'_-}\right)\psi\left(\frac{1 + x/z_-}{1 - x/z'_-}\right) * \nu^z$$

c) *If* $P' \overset{\text{loc}}{\ll} P$, *if* Z *is the density process of* P' *relative to* P, *and if* ν^Z *denotes the third* P*-characteristic of* Z, *a version of* $\imath(\psi; P, P')$ *is given by*

1.50
$$\imath(\psi) = \left(1 + \frac{x}{Z_-}\right)\psi\left(\frac{1}{1 + x/Z_-}\right) * \nu^Z.$$

Proof. a) With the notation $\mu^{(z,z')}$ of the proof of 1.33, we deduce from 1.41 that

$$\jmath(\psi) = \left(1 + \frac{y}{z'_-}\right)\psi\left(\frac{1 + x/z_-}{1 + y/z'_-}\right) * \mu^{(z,z')},$$

and 1.48 readily follows from 1.42.

(b) and (c) are deduced from (a) exactly as in 1.35 and 1.37. □

A particularly interesting case concerns the function

1.51
$$\psi(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x > 0. \end{cases}$$

1.52 **Definition.** a) We call *Hellinger process of order* 0 between P and P', and we denote by $h(0; P, P')$, any version of the process $\imath(\psi; P, P')$ with ψ given by 1.51.

b) The *Hellinger process of order* 0 *in the strict sense* is a version $h(0)$ of $h(0; P, P')$ which meets $h(0) = 1_{\Gamma''} \cdot h(0)$ (it is unique, up to a $(P + P')$-evanescent set). □

In other words, $h(0; P, P')$ is any increasing process such that $1_{\Gamma''} \cdot h(0; P, P')$ is the Q-compensator of the process

1.53
$$\sum_{s \leq \cdot} 1_{\Gamma''}(s) \frac{z'_s}{z'_{s-}} 1_{\{z_s = 0\}} = \frac{z'_S}{z'_{S-}} 1_{\{0 < S < \infty, z_S = 0 < z_{S-}\}} 1_{[\![S, \infty[\![}.$$

The reason for which we call $h(0; P, P')$ "Hellinger process of order 0" comes from the following corollary of 1.47, in which we use the function

1.54
$$\varphi_0(u, v) := v 1_{\{u=0\}} = v\psi\left(\frac{u}{v}\right) = \lim_{\alpha \downarrow\downarrow 0} \varphi_\alpha(u, v).$$

1.55 **Corollary.** a) *If* $\nu^{(z,z')}$ *denotes the* Q*-characteristic of the bi-dimensional process* (z, z'), *a version of* $h(0; P, P')$ *is given by*

1.56 $h(0;P,P') = \left(1+\frac{y}{z'_-}\right)1_{\{x+z_-=0\}} * \nu^{(z,z')} = \varphi_0\left(1+\frac{x}{z_-}, 1+\frac{y}{z'_-}\right) * \nu^{(z,z')}$.

b) $Q = \frac{P+P'}{2}$ *and if* ν^z *denotes the third* Q*-characteristic of* z*, a version of* $h(0;P,P')$ *is given by*

1.57 $$h(0;P,P') = \left(1-\frac{x}{z'_-}\right)1_{\{x+z_-=0\}} * \nu^z = \varphi_0\left(1+\frac{x}{z_-}, 1-\frac{x}{z'_-}\right) * \nu^z.$$

c) *If* $P' \overset{\text{loc}}{\ll} P$*, then* $h(0;P,P') = 0$ *is the Hellinger process of order* 0 *in the strict sense.*

1.58 **Remark.** 1.56 shows that formally, $h(0;P,P')$ is the process $h(\alpha;P,P')$ for $\alpha = 0$ (compare to 1.34). Moreover, $Y(0) := z' 1_{\{z>0\}}$ equals $\lim_{\alpha\downarrow\downarrow 0} Y(\alpha)$ (notation of 1.18) and is also a Q-supermartingale. However, *1.20 is not valid for* $\alpha = 0$ in general: indeed, $Y(0)_{S_n} = z'_{S_n} \to z'_{S-}$ on $\bigcap_n \{S_n < S\}$ as $n \uparrow \infty$, hence $1_{(\Gamma'')^c} \cdot Y(0)$ is not always equal to 0. □

Finally, we give a result similar to 1.30, which of secondary importance (observe that the function ψ given by 1.51 meets the requirements of the lemma below).

1.59 **Lemma.** *Let* ψ *be a function meeting 1.40 and such that* $\psi(x) < \psi(0)$ *for all* $x > 0$*. Let* $\imath(\psi)$ *be any version of* $\imath(\psi;P,P')$*. Then, up to a* Q*-null set,*

a) $\Delta\imath(\psi) \le \psi(0)$ *on* Γ''*, and* $\Delta\imath(\psi) < \psi(0)$ *on* $[\![0,S[\![$.

b) *If* T *is a predictable time with* $T \ge S$*, then* $\Delta\imath(\psi)_T = \psi(0)$ Q-a.s. *on the set* $\bigcup_n \{S_n = T < \infty\}$.

Proof. Without loss of generality, we can assume that $\imath(\psi) = 1_{\Gamma''} \cdot \imath(\psi)$ is the process defined in 1.42. Let $\gamma = \psi(0)$.

a) Let T be a predictable time. Then 1.42 yields on $\bigcup_n \{T \le S_n, T < \infty\}$:

1.60 $$\Delta\imath(\psi)_T = E_Q(\Delta\jmath(\psi)_T|\mathscr{F}_{T-}) \le \gamma E_Q(z'_T/z'_{T-}|\mathscr{F}_{T-}) = \gamma$$

because $\psi \le \gamma$, and $E_Q(z'_T|\mathscr{F}_{T-}) = z'_{T-}$. Moreover $\Delta\imath(\psi)_T = 0$ on $\bigcap_n \{S_n < T < \infty\}$: hence $\Delta\imath(\psi) \le \gamma$ up to a Q-evanescent set by the predictable section theorem I.2.18.

Next, $T = \inf(t: \Delta\imath(\psi)_t = \gamma)$ is predictable. We have $\psi(x) < \gamma$ for $x > 0$, while 1.60 implies $\Delta\jmath(\psi)_T = \gamma z'_T/z'_{T-}$ Q-a.s. on $\{T < \infty\}$, hence $\frac{z_T/z_{T-}}{z'_T/z'_{T-}} = 0$ Q-a.s. on $\{T < \infty, z'_T > 0\}$, hence $z_T = 0$ Q-a.s. on $\{T < \infty, z'_T > 0\}$ and this implies that $Q(T \ge S) = 1$.

b) Let T be a predictable time, with $T \ge S$. Then on $\{T < \infty\}$ we have either $z'_T = 0$, or $z_T = 0 < z'_T$, and in the latter case, $\psi\left(\frac{z_T/z_{T-}}{z'_T/z'_{T-}}\right) = \gamma$, hence $\Delta\jmath(\psi)_T =$

$\gamma z'_T/z'_{T-}$ in both cases. Thus $\Delta i(\psi)_T = E_Q(\gamma z'_T/z'_{T-}|\mathscr{F}_{T-}) = \gamma$ on $\bigcup_n \{T = S_n < \infty\}$. □

§ 1e. The Discrete Case

1. Now we consider a discrete-time filtered space $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}})$ endowed with two probability measures P and P'. Let Q meet 1.12, and $z = (z_n)_{n \in \mathbb{N}}$ and $z' = (z'_n)_{n \in \mathbb{N}}$ be the density processes of P and P' with respect to Q (see § III.3e). According to III.3.45, we set:

1.61 $$\beta_n = \frac{z_n}{z_{n-1}}, \quad \beta'_n = \frac{z'_n}{z'_{n-1}}$$

with the convention $0/0 = 0$ (recall that $z_n = 0$ if $z_{n-1} = 0$). For $\alpha \in (0, 1)$ any version of the Hellinger process $h(\alpha; P, P')$, say $h(\alpha)$, is again characterized by 1.20 which, translated into the discrete time setting, reads:

1.62 (i) $n \rightsquigarrow h(\alpha)_n$ is non-decreasing, $h(\alpha)_0 = 0$, $h(\alpha)_n$ is $\mathscr{F}_{n-1}$-measurable;
(ii) $M(\alpha)_n = z_n^\alpha z_n'^{1-\alpha} - \sum_{1 \le p \le n} z_{p-1}^\alpha z_{p-1}'^{1-\alpha}[h(\alpha)_p - h(\alpha)_{p-1}]$ is a Q-martingale. □

This allows to very easily compute $h(\alpha)$ in terms of the two processes (β_n) and (β'_n) defined in 1.61:

1.63 Proposition. *Let $\alpha \in (0, 1)$. The following formulae give two versions of the process $h(\alpha; P, P')$:*

1.64 $$\begin{cases} h(\alpha)_n = \sum_{1 \le p \le n} E_Q(1 - \beta_p^\alpha \beta_p'^{1-\alpha} | \mathscr{F}_{p-1}) \\ h(\alpha)_n = \sum_{1 \le p \le n} E_Q[\varphi_\alpha(\beta_p, \beta'_p) | \mathscr{F}_{p-1}]. \end{cases}$$

These two versions of $h(\alpha; P, P')$ are in general different, and they also differ in general from the Hellinger process in the strict sense.

Proof. We have $E_Q(\beta_p|\mathscr{F}_{p-1}) = 1_{\{\beta_{p-1} > 0\}}$ and $E_Q(\beta'_p|\mathscr{F}_{p-1}) = 1_{\{\beta'_{p-1} > 0\}}$. Then the concavity of the function $(u, v) \rightsquigarrow u^\alpha v^{1-\alpha}$ on $\mathbb{R}_+^2$ yields

$$E_Q(\beta_p^\alpha \beta_p'^{1-\alpha} | \mathscr{F}_{p-1}) \le 1_{\{\beta_{p-1} > 0, \beta'_{p-1} > 0\}} \le 1,$$

and so both processes $h(\alpha)$ defined in 1.64 meet 1.62(i). Moreover, we easily deduce from the definition of $M(\alpha)$ and from 1.61 that $E_Q[M(\alpha)_n - M(\alpha)_{n-1}|\mathscr{F}_{n-1}) = 0$, so 1.62(ii) holds. □

The reader will have recognized 1.32 in the second formula 1.64 (in the discrete case, we have $z^c = z'^c = 0$, and we can use II.3.1). Similarly, if ψ is a function satisfying 1.40, then a version of $i(\psi)$ is given by

1.65 $$\imath(\psi)_n = \sum_{1 \le p \le n} E_Q(\beta'_p \psi(\beta_p/\beta'_p) | \mathscr{F}_{p-1})$$

1.66 $$h(0; P, P')_n = \sum_{1 \le p \le n} E_Q(\beta'_p 1_{\{\beta_p = 0\}} | \mathscr{F}_{p-1})$$

(we recover 1.64 for $\alpha = 0$). Indeed, 1.65 readily follows from the definition of $\jmath(\psi)$, which here reads as (see 1.41):

$$\jmath(\psi)_n = \sum_{1 \le p \le n} \beta'_p \psi(\beta_p/\beta'_p).$$

Moreover, if we remember that for any $\mathscr{F}_p$-measurable nonnegative variable Y,

$$E_{P'}(Y | \mathscr{F}_{p-1}) = \frac{1}{z'_{p-1}} E_Q(Y z'_p | \mathscr{F}_{p-1}) = E_Q(Y \beta'_p | \mathscr{F}_{p-1})$$

on $\{z'_{p-1} > 0\}$, 1.65 has the form

1.67 $$\imath(\psi)_n = \sum_{1 \le p \le n} E_{P'}\left(\psi\left(\frac{\beta_p}{\beta'_p}\right) \middle| \mathscr{F}_{p-1}\right).$$

Finally, suppose that $P' \overset{\text{loc}}{\ll} P$, and call Z the density process of P' relative to P, and

1.68 $$\alpha_n = Z_n / Z_{n-1}$$

(always with $0/0 = 0$). Then 1.64 and 1.65 read as:

1.69 $$\begin{cases} h(\alpha)_n = \sum_{1 \le p \le n} E_P(1 - (\alpha_p)^{1-\alpha} | \mathscr{F}_{p-1}) \\ h(\alpha)_n = \sum_{1 \le p \le n} E_P(\varphi_\alpha(1, \alpha_p) | \mathscr{F}_{p-1}) \qquad \text{for } \alpha \in [0,1) \end{cases}$$

1.70 $$\imath(\psi)_n = \sum_{1 \le p \le n} E_P\left(\alpha_p \psi\left(\frac{1}{\alpha_p}\right) \middle| \mathscr{F}_{p-1}\right)$$

2. The case of independent random variables. Now we specialize even further. Namely, we suppose that P (resp. P') is the law of a sequence of independent random variables. This is formalized as such:

1.71 $$\begin{cases} \Omega = (\mathbb{R}^d)^{\mathbb{N}^*}, \quad \xi_n(\omega) = n^{\text{th}} \text{ coordinate (in } \mathbb{R}^d) \text{ of } \omega \in \Omega \\ \mathscr{F} = (\mathscr{R}^d)^{\mathbb{N}^* \otimes}, \quad \mathscr{F}_0 = \text{trivial } \sigma\text{-field}, \quad \mathscr{F}_n = \sigma(\xi_1, \ldots, \xi_n) \quad \text{for } n \ge 1 \\ P = \bigotimes_{n \in \mathbb{N}^*} \rho_n, \quad P' = \bigotimes_{n \in \mathbb{N}^*} \rho'_n \end{cases}$$

where ρ_n and ρ'_n are probability measures on $\mathbb{R}^d$: then under P (resp. P') the "canonical" random variables ξ_n are independent, and the law of ξ_n is ρ_n (resp. ρ'_n).

For each $n \in \mathbb{N}^*$ we consider a measure $\bar{\rho}_n$ such that $\rho_n \ll \bar{\rho}_n$ and $\rho'_n \ll \bar{\rho}_n$, and we call $\eta_n = d\rho_n / d\bar{\rho}_n$ and $\eta'_n = d\rho'_n / d\bar{\rho}_n$ the Radon-Nikodym derivatives (they are Borel nonnegative functions on $\mathbb{R}^d$). Thus if $Q = \bigotimes_{n \in \mathbb{N}^*} \bar{\rho}_n$, we readily obtain $P \overset{\text{loc}}{\ll} Q$, $P' \overset{\text{loc}}{\ll} Q$, and

1.72 $$z_n = \begin{cases} 1 & \text{if } n = 0 \\ \prod_{1 \le p \le n} \eta_p(\xi_p) & \text{if } n \ge 1 \end{cases}, \quad z'_n = \begin{cases} 1 & \text{if } n = 0 \\ \prod_{1 \le p \le n} \eta'_p(\xi_p) & \text{if } n \ge 1 \end{cases},$$

and in particular $\beta_n = \eta_n(\xi_n) 1_{\{z_{n-1} > 0\}}$ and similarly for β'_n, with the notation 1.61.

1.73 **Proposition.** a) *If* $\alpha \in (0, 1)$, *for all* $n \in \bar{\mathbb{N}}^*$ *the Hellinger integral* $H(\alpha; P_n, P'_n)$, *where* P_n *and* P'_n *are the restrictions of* P *and* P' *to* $\mathscr{F}_n$ (so $P_\infty = P, P'_\infty = P'$), *is*

1.74 $$H(\alpha; P_n, P'_n) = \prod_{p=1}^{n} H(\alpha; \rho_p, \rho'_p).$$

b) *If* $\alpha \in (0, 1)$, *a version of the Hellinger process* $h(\alpha; P, P')$ *is*

1.75 $$h(\alpha)_n = \sum_{1 \le p \le n} [1 - H(\alpha; \rho_p, \rho'_p)].$$

c) *If* ψ *satisfies 1.40, a version of* $\imath(\psi; P, P')$ *is*

1.76 $$\imath(\psi)_n = \sum_{1 \le p \le n} \int \rho'_p(dx) \psi(\eta_p(x)/\eta'_p(x)).$$

In particular, $h(\alpha)$ and $\imath(\psi)$ are deterministic.

Proof. a) By 1.7 and 1.72, and by the independance of the ξ_p's under Q,

$$H(\alpha; P_n, P'_n) = E_Q(z_n^\alpha z_n'^{1-\alpha}) = E_Q\left(\prod_{1 \le p \le n} \eta_p(\xi_p)^\alpha \eta'_p(\xi_p)^{1-\alpha} \right)$$
$$= \prod_{p=1}^{n} \int \bar{\rho}_p(dx) \eta_p(x)^\alpha \eta'_p(x)^{1-\alpha} = \prod_{p=1}^{n} H(\alpha; \rho_p, \rho'_p).$$

b) We use 1.64, the fact that $h(\alpha)_n$ is arbitrary on the set $\{z_{n-1} = 0\} \cup \{z'_{n-1} = 0\}$, and the values of β_n and β'_n given after 1.72, and the independance of the ξ_n's under Q again:

$$h(\alpha)_n = \sum_{1 \le p \le n} E_Q(1 - \eta_p(\xi_p)^\alpha \eta'_p(\xi_p)^{1-\alpha} | \mathscr{F}_{p-1})$$
$$= \sum_{1 \le p \le n} \int \bar{\rho}_p(dx)[1 - \eta_p(x)^\alpha \eta'_p(x)^{1-\alpha}],$$

which equals the right-hand side of 1.75. Finally, 1.76 is proved similarly, using 1.67. □

In particular, if $\psi(u) = 1_{\{u=0\}}$, 1.76 gives:

1.77 $$h(0)_n = \sum_{1 \le p \le n} \rho'_p(\{x : \eta_p(x) = 0\}).$$

2. Predictable Criteria for Absolute Continuity and Singularity

§ 2a. Statement of the Results

In this section P and P' are two probability measures on a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$. To avoid tedious complications, we assume that $\mathscr{F} = \mathscr{F}_{\infty-} = \bigvee_t \mathscr{F}_t$. As usual, for every stopping times T we denote by P_T and P'_T the restrictions of P and P' to $(\Omega, \mathscr{F}_T)$. Our aim is to describe whether $P'_T \ll P_T$ or $P'_T \perp P_T$ in a "predictable way", and more precisely in terms of the Hellinger processes of various order, as introduced in the previous section.

As a matter of fact, there are many different implications or equivalences, and we will not fear redundancy. Our first (and perhaps main) result concerns absolute continuity (all proofs are in § 2b).

2.1 **Theorem.** *Assume that* $\mathscr{F} = \mathscr{F}_{\infty-}$, *and let* T *be a stopping time. For* $\alpha \in [0,1)$ *let* $h(\alpha)$ *be any version of the Hellinger process* $h(\alpha; P, P')$ *of order* α (see 1.24 for $\alpha > 0$ and 1.52 for $\alpha = 0$). *There is equivalence between*:

(i) $P'_T \ll P_T$;

(ii) $P'_0 \ll P_0$ *and* $P'(h(\frac{1}{2})_T < \infty) = 1$ *and* $P'(h(0)_T = 0) = 1$;

(iii) $P'_0 \ll P_0$ *and* $h(\alpha)_T \xrightarrow{P'} 0$ *as* $\alpha \downarrow\downarrow 0$

($\xrightarrow{P'}$ means convergence in measure, relative to P').

Of course, in this theorem as well as in all others in this section, one could replace $h(\frac{1}{2})$ by $h(\beta)$, for any *fixed* $\beta \in (0,1)$.

As it is natural, what happens at time 0 plays a specific (and rather trivial) rôle. In order to make the conditions more homogeneous, we introduce a set G_0, which is $(P + P')$-uniquely determined by the following:

2.2 $$G_0 \in \mathscr{F}_0, \quad \text{such that} \begin{cases} P'_0 \sim P_0 & \text{in restriction to } G_0 \\ P'_0 \perp P_0 & \text{in restriction to } (G_0)^c. \end{cases}$$

G_0 can be constructed as follows: let Q be any measure such that $P \overset{\text{loc}}{\ll} Q$ and $P' \overset{\text{loc}}{\ll} Q$, and denote as usual the density processes of P and P' relative to Q by z and z'; then

2.3 $$\begin{cases} G_0 = \{z_0 > 0, z'_0 > 0\} & (P + P')\text{-a. s.} \\ \text{and then: } \begin{cases} G_0 = \{z_0 > 0\} & P'\text{-a. s. (because } z'_0 > 0 \quad P'\text{-a. s.)} \\ G_0 = \{z'_0 > 0\} & P\text{-a. s.,} \end{cases} \end{cases}$$

and we deduce from Lemma 1.11 that

2.4 $$\begin{cases} P'_0 \ll P_0 \Leftrightarrow P'(G_0) = 1 \Leftrightarrow P'(z_0 > 0) = 1 \\ P'_0 \perp P_0 \Leftrightarrow P'(G_0) = 0 \Leftrightarrow P'(z_0 > 0) = 0. \end{cases}$$

At this stage, we introduce two other sets of interest (T is again a stopping time, and $h(\alpha)$ is a version of $h(\alpha; P, P')$):

2.5
$$\begin{cases} G_T = G_0 \cap \left\{ h\left(\frac{1}{2}\right)_T < \infty \right\} \cap \{h(0)_T = 0\} \\ \tilde{G}_T = G_0 \cap \left\{ \limsup_{\alpha \downarrow\downarrow 0} h(\alpha)_T = 0 \right\} \end{cases}$$

(if $T \equiv 0$, then $G_T = \tilde{G}_T = G_0$, as given by 2.2).

The next theorem includes the equivalence (i)$\Leftrightarrow$(ii) of 2.1; it also features a sort of symmetry between the equivalence problem and the singularity problem.

2.6 **Theorem.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$, let T be a stopping time, and define G_T and $\tilde{G}_T$* by 2.5.

a) *There is equivalence between:* (i) $P'_T \ll P_T$;
(ii) $P'(G_T) = 1$;
(iii) $P'(\tilde{G}_T) = 1$.

b) *$P'_T \perp P_T$ implies $P'(G_T) = 0$ and $P'(\tilde{G}_T) = 0$.*

c) *If either $P'_0 \perp P_0$ ($\Leftrightarrow P'(G_0) = 0$) or $P'(h(\frac{1}{2})_T < \infty) = 0$, then $P'_T \perp P_T$.*

We derive two corollaries, the first one being just a restatement of 2.6 when $T \equiv \infty$.

2.7 **Corollary.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$.*

a) *There is equivalence between:* (i) $P' \ll P$;
(ii) $P'(G_\infty) = 1$;
(iii) $P'(\tilde{G}_\infty) = 1$.

b) *$P' \perp P$ implies $P'(G_\infty) = P'(\tilde{G}_\infty) = 0$.*

c) *If $P'_0 \perp P_0$ or if $P'(h(\frac{1}{2})_\infty < \infty) = 0$, then $P' \perp P$.*

2.8 **Corollary.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$ and that $P' \overset{\text{loc}}{\ll} P$.*

a) *$P' \ll P \Leftrightarrow P'(h(\frac{1}{2})_\infty < \infty) = 1$.*

b) *$P' \perp P \Leftrightarrow P'(h(\frac{1}{2})_\infty < \infty) = 0$.*

Proof. $P'_0 \ll P_0$ is obvious, and yields $P'(G_0) = 1$. Moreover $h(0) = 0$ is a version of $h(0; P, P')$ by 1.55, hence (a) (resp. (b)) follows from 2.7a (resp. 2.7b, c). □

The various criteria derived above for $P'_T \ll P_T$ are called *predictable criteria* because the processes $h(\alpha)$ are (or, may be chosen) predictable. We do not know whether it is possible to derive a *predictable* criterion (necessary and sufficient condition) for $P'_T \perp P_T$, but $P'(G_T) = 0$ or $P'(\tilde{G}_T) = 0$ are not enough for $P'_T \perp P_T$, as shown below:

2.9 **Example.** Let σ and τ be two independent random variables on the space $(\Omega, \mathscr{F}, Q)$, with $Q(\sigma = 1) = Q(\sigma = -1) = 1/2$ and τ being exponential with parameter 1. Set

$$z_t = \begin{cases} 1 & \text{if } t < \tau \text{ or if } t \geq \tau \geq 1 \\ 2 & \text{if } t \geq \tau \text{ and } \tau < 1 \text{ and } \sigma = 1 \\ 0 & \text{if } t \geq \tau \text{ and } \tau < 1 \text{ and } \sigma = -1 \end{cases}$$

and $z' = 2 - z$ and $\mathscr{F}_t = \sigma(z_s \colon s \leq t)$. Then $\mathbf{F} = (\mathscr{F}_t)_{t \geq 0}$ is a filtration, and upon shrinking $\mathscr{F}$ we may assume that $\mathscr{F} = \mathscr{F}_{\infty-}$. Moreover z and z' are bounded nonnegative martingales on $(\Omega, \mathscr{F}, \mathbf{F}, Q)$ with $z_0 = z'_0 = 1$, and so are the density processes of the measures $P = z_\infty \cdot Q$ and $P' = z'_\infty \cdot Q$.

Moreover, with the notation of 1.35,

$$\nu^z(dt, dx) = \tfrac{1}{2} 1_{[\![0, \tau \wedge 1]\!]}(t)\, dt [\varepsilon_1(dx) + \varepsilon_{-1}(dx)]$$

so 1.36 and 1.57 yield $h(\alpha)_t = t \wedge \tau \wedge 1$ for $\alpha \in [0, 1)$. Thus $P'(G_\infty) = P'(\tilde{G}_\infty) = 0$. However, P and P' are *not* mutually singular, since they coincide on the set $\{\tau \geq 1\}$. □

Let $Q = \frac{P+P'}{2}$. The density processes z and z' satisfy $z + z' = 2$, and as above we set $S = \inf(t : z_t = 0 \text{ or } z'_t = 0)$ and $\Gamma'' = \{z_- > 0 \text{ and } z'_- > 0\} \cup [\![0]\!]$. We know that for every stopping time T,

2.10
$$\begin{cases} P'_T \ll P_T \Leftrightarrow P'(z_T > 0) = 1 \\ P'_T \perp P_T \Leftrightarrow P'(z_T > 0) = 0, \end{cases}$$

while $\{z_T > 0\}$ is naturally written as a three-fold intersection:

2.11
$$\begin{cases} \{z_T > 0\} = \{z_0 > 0\} \cap A_T \cap B_T, \text{ where} \\ A_T = \{0 < S \leq T, z_{S-} = 0\}^c,\ B_T = \{0 < S \leq T, S < \infty, z_{S-} > z_S = 0\}^c \end{cases}$$

and $\{z_T = 0\}$ is the disjoint union of $\{z_0 = 0\}$ and $(A_T)^c$ and $(B_T)^c$. In view of our formulation of Theorem 2.6, and comparing 2.10 with 2.5, it is tempting to write the following equalities:

(i) $G_0 = \{z_0 > 0\}$ P'-a.s.

(ii) $\{h(\frac{1}{2})_T < \infty\} = A_T$ P'-a.s.

(iii) $\{h(0)_T = 0\} = B_T$ P'-a.s.

(i) is true (see 2.3). (ii) and (iii) cannot be both true, since if they were one would have an equivalence in 2.6b. However, we do have the following properties, which go in the direction of (ii) and (iii), and which offer a nice interpretation of the previous theorems:

2.12 Lemma. *Assume that* $Q = \dfrac{P + P'}{2}$ *and let* $h'(\alpha) = 1_{\Gamma''} \cdot h(\alpha)$ *be the Hellinger process of order* α *in the strict sense. Then*

a) $A_T \subset \{h'(\frac{1}{2})_T < \infty\}$ *P'-a.s.*

b) $P'(B_T) = 1 \Leftrightarrow P'(h'(0)_T = 0) = 1$.

c) *On the set* $\{h'(\frac{1}{2})_T < \infty\}$ *we have* $h'(0)_T = \lim_{\alpha \downarrow\downarrow 0} h'(\alpha)_T$.

d) $\{h'(\frac{1}{2})_T < \infty\} \cap \{h'(0)_T = 0\} \subset A_T \cap B_T$ *P'-a.s.*

We end this series of results with two other criteria, which are closely related to the previous ones, and sometimes useful. The first one improves upon 2.6a,b.

2.13 **Theorem.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$. Let T be a stopping time and define G_T and $\tilde{G}_T$ by 2.5. Let $F \in \mathscr{F}_T$.*

a) *If $P'(F) = P'(F \cap G_T)$, or if $P'(F) = P'(F \cap \tilde{G}_T)$, then $P'_T \ll P_T$ in restriction to the set F.*

b) *If $P'_T \perp P_T$ in restriction to the set F, then $P'(F \cap G_T) = P'(F \cap \tilde{G}_T) = 0$.*

The last result is a set of two "non-predictable" criteria for $P'_T \ll P_T$. We first introduce two $[0, \infty]$-valued processes (with the convention $2/0 = +\infty$; here again $Q = \frac{P+P'}{2}$):

2.14 $$Z_t = \frac{z'_t}{z_t}, \quad \alpha_t = \begin{cases} Z_t/Z_{t-} & \text{if } 0 < Z_{t-} < \infty \\ 0 & \text{if } Z_{t-} = 0 \\ +\infty & \text{if } Z_{t-} = +\infty. \end{cases}$$

2.15 **Theorem.** *Assume that $\mathscr{F} = \mathscr{F}_{\infty-}$ and let T be a stopping time. There is equivalence between:*

(i) $P'_T \ll P_T$;

(ii) $P'_0 \ll P_0$ *and* $P'(h(\frac{1}{2})_T < \infty) = 1$ *and* $P'(B_T) = 0$ (see 2.11 for B_T);

(iii) $P'_0 \ll P_0$ *and* $P'(h(\frac{1}{2})_T < \infty) = 1$ *and* $P'(\sup_{t \leq T} \alpha_t < \infty) = 1$.

§ 2b. The Proofs

It turns out that the key result is Lemma 2.12, all other statements being rather simple corollaries. So we begin by proving this lemma, through a number of steps.

We always assume $\mathscr{F} = \mathscr{F}_{\infty-}$. Let $Q = \dfrac{P + P'}{2}$, with the usual notation z, z', S_n, S, Γ'' (see § 1a); $h(\alpha)$ is a version of $h(\alpha; P, P')$, and $h'(\alpha) = 1_{\Gamma''} \cdot h(\alpha)$ is the Hellinger process in the strict sense.

2.16 **Lemma.** *We have 2.12a: $A_T \subset \{h'(\frac{1}{2})_T < \infty\}$ P'-a.s.*

Proof. Since 1.1 holds, we deduce from 1.18 that $E_Q(\sqrt{z_- z'_-} \cdot h'(\frac{1}{2})_\infty) < \infty$, hence $\sqrt{z_- z'_-} \cdot h'(\frac{1}{2})_\infty < \infty$ Q-a.s. Since $\sqrt{z_- z'_-} \geq \frac{1}{n}$ on $[\![0, S_n]\!]$ we deduce

2.17 $$h'(\tfrac{1}{2})_{S_n} < \infty \quad Q\text{-a.s.}$$

But on A_T there are only three possibilities (see 2.11):

1) either $T \leq S_n$ for some $n \in \mathbb{N}^*$, in which case $h'(\frac{1}{2})_T \leq h'(\frac{1}{2})_{S_n}$;

2) or $S = 0$, in which case $h'(\frac{1}{2})_T = 0$;

3) or $T > S_n$ for all n and $z_{S-} > 0$; but since $\inf_s z'_s > 0$ P'-a.s. we have $z'_{S-} > 0$

P'-a.s., and so there is P'-a.s. an integer n (depending on ω) such that $S_n = S$, in which case $h'(\frac{1}{2})_T = h'(\frac{1}{2})_{S_n}$.

Then, the claim follows from 2.17. $\square$

2.18 **Lemma.** *We have 2.12b*: $P'(B_T) = 1 \Leftrightarrow P'(h'(0)_T = 0) = 1$.

Proof. Since z' is bounded, I.3.12 yields

$$E_{P'}(h'(0)_{T\wedge S_n}) = E_Q(z'_\infty h'(0)_{T\wedge S_n}) = E_Q(z'_- \cdot h'(0)_{T\wedge S_n}),$$

which by 1.42 equals $E_Q(z'_- \cdot j_{T\wedge S_n})$, where j is the process defined by 1.53. Letting $n \uparrow \infty$, we get $h'(0)_{T\wedge S_n} \uparrow h'(0)_T$ because $1_{\Gamma''} \cdot h'(0) = h'(0)$, while one easily deduces from 1.53 that $z'_- \cdot j_{T\wedge S_n} \uparrow z'_- \cdot j_T$. Thus, using 1.53 again, we obtain

$$E_{P'}(h'(0)_T) = E_Q(z'_- \cdot j_T) = E_Q(z'_T 1_{(B_T)^c}) = P'((B_T)^c)$$

(use $B_T \in \mathscr{F}_T$ and 1.14 for the last equality). The claim follows. $\square$

Next, we need some properties of the functions φ_α defined by 1.32 and 1.54 for $\alpha \in [0,1)$, and which we recall below:

$$\varphi_\alpha(u,v) = \begin{cases} \alpha u + (1-\alpha)v - u^\alpha v^{1-\alpha} & \text{for } \alpha \in (0,1) \\ v 1_{\{u=0\}} & \text{for } \alpha = 0. \end{cases}$$

2.19 **Lemma.** *Let* $\alpha \in (0,1)$.

a) $\varphi_\alpha \leq 8\varphi_{1/2}$.

b) *There is a constant* γ_α *such that* $\varphi_{1/2} \leq \gamma_\alpha \varphi_\alpha$.

c) $\varphi_\alpha(u,v) \leq 8\alpha(1 \vee \operatorname{Log} N)\varphi_{1/2}(u,v)$ *for all* $N \geq 1, 0 \leq v \leq Nu$.

Proof. Since $\varphi_\alpha(0,v) = (1-\alpha)v$, all the claims are obvious for $u = 0$, $v \geq 0$ $\left(\text{provided } \gamma_\alpha \geq \dfrac{1}{2(1-\alpha)} \text{ in (b)}\right)$. If $u > 0$ we have

$$\varphi_\alpha(u,v) = u\varphi_\alpha\left(1, \frac{v}{u}\right) = uf_\alpha\left(\frac{v}{u}\right), \tag{2.20}$$

where $f_\alpha(x) = \alpha + (1-\alpha)x - x^{1-\alpha}$. We have

$$\frac{\partial f_\alpha}{\partial \alpha}(x) = 1 - x + x^{1-\alpha} \operatorname{Log} x \leq 1 - x + x \operatorname{Log} x,$$

and thus $f_\alpha(x) \leq \alpha(1 - x + x \operatorname{Log} x)$. One easily checks that $1 - x + x \operatorname{Log} x \leq 4(1-\sqrt{x})^2$ for $x \in [0,4]$; if $N \geq 4$ and $x \in [4,N]$, the following holds:

$$1 - x + x \operatorname{Log} x \leq 1 - x + x \operatorname{Log} N \leq x \operatorname{Log} N \leq 4(\operatorname{Log} N)(1-\sqrt{x})^2.$$

Putting together these estimates, and noting that $f_{1/2}(x) = \frac{1}{2}(1-\sqrt{x})^2$, we get

$$f_\alpha(x) \leq 8\alpha(1 \vee \operatorname{Log} N) f_{1/2}(x) \quad \text{for } 0 \leq x \leq N, N \geq 1.$$

In view of 2.20, this gives (c). It also gives $\varphi_\alpha(u,v) \le 8\varphi_{1/2}(u,v)$ for $v \le u$ (take $N = 1$). Since $\varphi_\alpha(u,v) = \varphi_{1-\alpha}(v,u)$ we also deduce that $\varphi_\alpha(u,v) \le 8\varphi_{1/2}(u,v)$ for $u \le v$, and thus (a) holds.

It remains to prove (b). We have $f_\alpha(x) > 0$ for all $x \in \mathbb{R}_+ \backslash \{1\}$, so the function $g_\alpha = f_{1/2}/f_\alpha$ is continuous on $\mathbb{R}_+ \backslash \{1\}$. Moreover $f_\alpha(x) \sim \frac{\alpha(1-\alpha)}{2}(x-1)^2$ as $x \to 1$ and $f_\alpha(x) \sim (1-\alpha)x$ as $x \uparrow \infty$; hence $g_\alpha(x) \to \frac{1}{4\alpha(1-\alpha)}$ as $x \to 1$ and $g_\alpha(x) \to \frac{1}{2(1-\alpha)}$ as $x \uparrow \infty$. Therefore g_α is bounded by a constant, say γ_α, and $\varphi_{1/2} \le \gamma_\alpha \varphi_\alpha$ follows again from 2.20. □

2.21 **Corollary.** a) $h'(\alpha) \le 8h'(\frac{1}{2})$ *for all* $\alpha \in (0,1)$.

b) *For every* $\alpha \in (0,1)$ *there is a constant* $\tilde{\gamma}_\alpha$ *such that* $h'(\frac{1}{2}) \le \tilde{\gamma}_\alpha h'(\alpha)$.

c) *2.12c holds: on the set* $\{h'(\frac{1}{2})_T < \infty\}$ *we have* $h'(0)_T = \lim_{\alpha \downarrow\downarrow 0} h'(\alpha)_T$.

Proof. Recall that, with the notation of 1.35,

2.22 $$h'(\alpha) = \frac{\alpha(1-\alpha)}{2}\left(\frac{1}{Z_-} + \frac{1}{Z'_-}\right)^2 \cdot \langle z^c, z^c \rangle + \varphi_\alpha\left(1 + \frac{x}{Z_-}, 1 - \frac{x}{Z'_-}\right) * \nu^z.$$

Then (a) and (b) follow from 2.19 $\left(\text{note that } \frac{\alpha(1-\alpha)}{2} \le \frac{1}{8} \text{ for all } \alpha \text{ and take } \tilde{\gamma}_\alpha = \gamma_\alpha \vee \frac{1}{4\alpha(1-\alpha)}\right)$. Moreover $\varphi_\alpha \to \varphi_0$ pointwise as $\alpha \to 0$, so (c) follows from 2.19 again and from Lebesgue convergence theorem. □

2.23 **Lemma.** *2.12d holds:* $\{h'(\frac{1}{2})_T < \infty\} \cap \{h'(0)_T = 0\} \subset A_T \cap B_T$ P'*-a.s.*

Proof. Set $C_T = \{h'(\frac{1}{2})_T < \infty\} \cap \{h'(0)_T = 0\}$ and, for $N \ge e$,

$$K(N) = 8\varphi_{1/2}\left(1 + \frac{x}{Z_-}, 1 - \frac{x}{Z'_-}\right) 1_{\{N(1+x/Z_-) < 1 - x/Z'_-\}} * \nu^z$$

(same notation as in the previous proof). 2.22 and 2.19 yield, since $\operatorname{Log} N \ge 1$ and $\frac{\alpha(1-\alpha)}{2} \le \alpha$:

2.24 $$h'(\alpha) \le 8\alpha(\operatorname{Log} N) h'(\tfrac{1}{2}) + K(N).$$

We have $\varphi_{1/2}(u,v) 1_{\{Nu < v\}} \to \varphi_{1/2}(u,v) 1_{\{u=0<v\}} = \frac{1}{2}\varphi_0(u,v)$ as $N \uparrow \infty$, so by definition of the set C_T we deduce that $K(N)_T \downarrow 0$ as $N \uparrow \infty$ on C_T (Lebesgue convergence theorem). Hence if $\eta > 0$ is fixed, there exists $N \ge e$ such that

2.25 $$P'(C_T \cap \{h'(\tfrac{1}{2})_T \ge N \quad \text{or} \quad K(N)_T \ge \eta\}) \le \eta.$$

By construction the processes $h'(1/2)$ and $K(N)$ are predictable, increasing and right-continuous (they are "generalized increasing processes that do not jump to infinity", in the sense of III.5.8). So $V = \inf(t: h'(\frac{1}{2})_t \geq N \text{ or } K(N)_t \geq \eta)$ is a predictable time (I.2.13), and 2.25 implies

2.26 $$P'(C_T \cap D) \leq \eta, \quad \text{where } D = \{V \leq T, V < \infty\}.$$

Recall from 1.18 that $Y(\alpha) = z^\alpha z'^{1-\alpha}$ is a Q-supermartingale and that $Y(\alpha) - Y(\alpha)_- \cdot h'(\alpha)$ is a Q-uniformly integrable martingale for $\alpha \in (0,1)$. Since V is predictable and $0 \leq Y(\alpha) \leq 2$,

$$E_Q(Y(\alpha)_0 - Y(\alpha)_{V-}) = E_Q(Y(\alpha)_- \cdot h'(\alpha)_{V-}) \leq 2E_Q(h'(\alpha)_{V-}) \leq 16\alpha(\operatorname{Log} N)N + 2\eta,$$

the last inequality coming from 2.24 and the definition of V. Now, let $\alpha \downarrow\downarrow 0$. We have $Y(\alpha) \to z' 1_{\{z>0\}}$, so the above inequality yields

2.27 $$P'(z_0 > 0) - P'(z_{V-} > 0) = E_Q(z'_0 1_{\{z_0>0\}} - z'_{V-} 1_{\{z_{V-}>0\}}) \leq 2\eta$$

(use 1.14). We can write

$$\begin{aligned} P'(C_T) &= P'(C_T \cap \{z_0 = 0\}) + P'(C_T \cap \{z_0 > 0\}) \\ &\leq P'(C_T \cap \{z_0 = 0\}) + P'(C_T \cap \{z_{V-} > 0\}) + 2\eta \qquad \text{(by 2.27)} \\ &\leq P'(C_T \cap \{z_0 = 0\}) + P'(C_T \cap \{z_{V-} > 0\} \cap D^c) + 3\eta \quad \text{(by 2.26).} \end{aligned}$$

But $D^c \cap \{z_{V-} > 0\} \subset \{z_T > 0\}$ by definition, hence

$$P'(C_T) \leq P'(C_T \cap \{z_0 = 0\}) + P'(C_T \cap \{z_T > 0\}) + 3\eta.$$

Since $\eta > 0$ is arbitrary, we get $C_T = C_T \cap \{z_0 = 0 \text{ or } z_T > 0\}$ P'-a. s. Moreover 2.11 gives $\{z_0 = 0 \text{ or } z_T > 0\} = A_T \cap B_T$ P'-a. s., hence the claim. □

Now we can proceed to the proof of Theorems 2.1, 2.6, 2.13 and 2.15. Recall that G_T and $\tilde{G}_T$ are defined by 2.5, and we similarly define G'_T $\tilde{G}'_T$ with the Hellinger processes in the strict sense $h'(\alpha)$ instead of $h(\alpha)$. We obviously have

2.28 $$G_T \subset G'_T, \quad \tilde{G}_T \subset \tilde{G}'_T.$$

Proof of Theorem 2.13. In view of 2.21b,c we have $G'_T = \tilde{G}'_T$, while 2.5 and 2.11 and 2.12d yield $G'_T \subset \{z_T > 0\}$ P'-a. s.

If $F \in \mathscr{F}_T$ meets $P'(F) = P'(F \cap G_T)$ (resp. $P'(F) = P'(F \cap \tilde{G}_T)$), then $F \subset G_T$ (resp. $F \subset \tilde{G}_T$) P'-a. s., so 2.28 implies $F \subset G'_T$ P'-a. s., or equivalently $F \subset \{z_T > 0\}$ P'-a. s.: this obviously yields (a).

Conversely, assume that $P'_T \perp P_T$ in restriction to F. Then $z_T = 0$ P'-a. s. on F, hence $F \cap G'_T = F \cap \tilde{G}'_T = \varnothing$ P'-a. s., and using again 2.28 we obtain (b). □

Proof of Theorem 2.6. (b) follows from 2.13b applied to $F = \Omega$, and the implications (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (i) follow from 2.13a applied to $F = \Omega$ again.

If $P'_0 \perp P_0$, that $P'_T \perp P_T$ is obvious. Assume now that $P'(h(\frac{1}{2})_T < \infty) = 0$. By construction, we have

2.29
$$\begin{cases} h(\alpha)_T = h'(\alpha)_T & \text{on } \{z_T > 0, z'_T > 0\} \\ P'(z'_T > 0) = 1, \end{cases}$$

so $\{h'(\frac{1}{2})_T < \infty, h(\frac{1}{2})_T = \infty\} \subset \{z_T = 0\}$ P'-a.s. Moreover 2.11 and 2.12a yield $\{z_T > 0\} \subset A_T \subset \{h'(\frac{1}{2})_T < \infty\}$ P'-a.s.: hence $\{h(\frac{1}{2})_T = \infty\} \subset \{z_T = 0\}$ P'-a.s., and so our assumption implies $P'(z_T = 0) = 1$, which in turn gives $P'_T \perp P_T$. Hence we have proved (c).

Finally, assume $P'_T \ll P_T$, so $P'(z_T > 0) = 1$. Then 2.29 shows that $h(\alpha)_T = h'(\alpha)_T$ P'-a.s. for all $\alpha \in [0,1)$. Moreover $P'(G_0) = P'(A_T) = P'(B_T) = 1$, thus 2.12a,b,c yields $P'(h'(\frac{1}{2})_T < \infty) = 1$ and $P'(h'(0)_T = 0) = 1$ and $P'(\limsup_{\alpha \downarrow\downarrow 0} h'(\alpha)_T = 0) = 1$. Since we can replace $h'(\alpha)$ by $h(\alpha)$ above, coming back to the definition 2.5 of G_T and $\tilde{G}_T$ yields $P'(G_T) = P'(\tilde{G}_T) = 1$ so (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) in (a), and we are finished. □

Proof of Theorem 2.1. The implications (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) follow from 2.6a. Assume now (iii), then a-fortiori $h'(\alpha)_T \xrightarrow{P'} 0$ as $\alpha \downarrow\downarrow 0$, so 2.21b yields $h'(\frac{1}{2})_T < \infty$ P'-a.s., and 2.12c then yields $h'(0)_T = 0$ P'-a.s. In other words, $P'(G'_T) = 1$: then 2.6a (applied with the Hellinger processes in the strict sense) gives $P'_T \ll P_T$. □

Proof of Theorem 2.15. The equivalence (i) $\Leftrightarrow$ (ii) follows from 2.12b and from the equivalence (i) $\Leftrightarrow$ (ii) in 2.1.

Assume (i). To obtain (iii) it is enough (in view of 2.1) to prove that $\sup_{t \le T} \alpha_t < \infty$ P'-a.s. To see this, we observe that $\inf_t z_t > 0$ P-a.s., so $P'_T \ll P_T$ implies $\sup_{t \le T} Z_t < \infty$ P'-a.s.; we also have $\inf_t z'_t > 0$ P'-a.s., hence $\inf_t Z_t > 0$ P'-a.s. (recall that $z + z' = 2$); hence clearly $\sup_{t \le T} \alpha_t < \infty$ P'-a.s., and we have obtained that (i) $\Rightarrow$ (iii).

On $(B_T)^c$ we have $S \le T$ and $S < \infty$ and $z_{S-} > z_S = 0$, so $Z_S = \infty$ while $0 < Z_{S-} < \infty$: hence $\sup_{t \le T} \alpha_t = \infty$ on $(B_T)^c$, and therefore $P'(\sup_{t \le T} \alpha_t < \infty) = 1$ yields $P'(B_T) = 1$. Thus (iii) $\Rightarrow$ (ii). □

2.30 **Remark.** The proof proposed above is not the simplest one for proving Theorems 2.1 and 2.6, essentially because we want a unified proof for both the singularity and the absolute continuity; moreover, this proof allows to obtain at the same time Theorem 2.13 and Lemma 2.12, which have an interest in their own. In Chapter V we will see a slightly different, and somewhat simpler, proof in the setting of "contiguity of sequences of measures". □

§ 2c. The Discrete Case

1. Here we translate some of the previous results in the discrete case setting. The situation is as in § 1e: we have a discrete-time filtered space $(\Omega, \mathscr{F}, \mathbf{F} = (\mathscr{F}_n)_{n \in \mathbb{N}})$ endowed with two measures P and P'. Let Q meet 1.12, and $z = (z_n)_{n \in \mathbb{N}}$, $z' = (z'_n)_{n \in \mathbb{N}}$ be the density processes of P and P' with respect to Q. We also set (with $\frac{0}{0} = 0$):

$$\beta_n = z_n/z_{n-1}, \quad \beta'_n = z'_n/z'_{n-1}.$$

We assume that $\mathscr{F} = \mathscr{F}_{\infty-} = \bigvee_n \mathscr{F}_n$, and that T is a stopping time. In view of 1.63 and 1.66, and recalling that $\varphi_{1/2}(u,v) = \frac{1}{2}(\sqrt{u} - \sqrt{v})^2$, Theorem 2.6 yields:

2.31 **Theorem.** a) *We have $P'_T \ll P_T$ if and only if $P'_0 \ll P_0$ and*

2.32 $$\sum_{1 \le n \le T} E_Q[(\sqrt{\beta_n} - \sqrt{\beta'_n})^2 | \mathscr{F}_{n-1}] < \infty \quad P'\text{-a. s.}$$

2.33 $$\sum_{1 \le n \le T} E_Q(\beta'_n 1_{\{\beta_n = 0\}} | \mathscr{F}_{n-1}) = 0 \quad P'\text{-a. s.}$$

b) *If $P'_T \perp P_T$ then with P'-probability 1 we have either $z_0 = 0$ or $\sum_{1 \le n \le T} E_Q[(\sqrt{\beta_n} - \sqrt{\beta'_n})^2 | \mathscr{F}_{n-1}] = \infty$ or $\sum_{1 \le n \le T} E_Q(\beta'_n 1_{\{\beta_n = 0\}} | \mathscr{F}_{n-1}) = 0$.*

c) *If $P'_0 \perp P_0$ or if $\sum_{1 \le n \le T} E_Q[(\sqrt{\beta_n} - \sqrt{\beta'_n})^2 | \mathscr{F}_{n-1}] = \infty$ P'-a. s., then $P'_T \perp P_T$.*

The translation of 2.15 also has some interest. We assume that $Q = \dfrac{P + P'}{2}$ and we set (again with $2/0 = +\infty$):

2.34 $$Z_n = \frac{z'_n}{z_n}, \quad \alpha_n = \begin{cases} Z_n/Z_{n-1} & \text{if } 0 < Z_{n-1} < \infty \\ 0 & \text{if } Z_{n-1} = 0 \\ +\infty & \text{if } Z_{n-1} = +\infty. \end{cases}$$

2.35 **Theorem.** *We have $P'_T \ll P_T$ if and only if $P'_0 \ll P_0$ and 2.32 holds and $P'(\sup_{1 \le n \le T} \alpha_n < \infty) = 1$.*

Finally, assume that $P' \overset{\text{loc}}{\ll} P$, and call Z the density process of P' with respect to P and set $\alpha_n = Z_n/Z_{n-1}$ (with $0/0 = 0$). Observe that Z_n and α_n are exactly the same as in 2.34. Then 2.8 yields

2.36 **Theorem.** *Assume that $P' \overset{\text{loc}}{\ll} P$. Then*

a) *$P'_T \ll P_T$ if and only if $\sum_{1 \le n \le T} E_P[(1 - \sqrt{\alpha_n})^2 | \mathscr{F}_{n-1}] < \infty$ P'-a. s.*

b) *$P'_T \perp P_T$ if and only if $\sum_{1 \le n \le T} E_P[(1 - \sqrt{\alpha_n})^2 | \mathscr{F}_{n-1}] = \infty$ P'-a. s.*

2. Now we specialize further, assuming 1.71: that is P and P' are laws of a sequence of independent random variables.

2.37 **Theorem.** *Assume 1.71.*

a) *$P' \ll P$ if and only if the following two conditions holds:*

(i) $\sum_n [1 - H(\frac{1}{2}; \rho_n, \rho'_n)] < \infty$

(ii) *$\rho'_n \ll \rho_n$ for all $n \in \mathbb{N}^*$.*

b) $P' \perp P$ *if and only if at least one of the following two conditions hold*:
(i′) $\sum_n [1 - H(\frac{1}{2}; \rho_n, \rho'_n)] = \infty$,
(ii′) *there exists* $n \in \mathbb{N}^*$ *with* $\rho'_n \perp \rho_n$.

Proof. a) If $P' \ll P$ then (ii) is obvious, and $h(\frac{1}{2})_n = \sum_{1 \le p \le n} [1 - H(\frac{1}{2}; \rho_p, \rho'_p)]$ is a version of $h(\frac{1}{2}; P, P')$ by 1.73. Thus (i) follows from 2.1.

Conversely assume (i) and (ii). Then (ii) yields $P' \overset{\text{loc}}{\ll} P$, and thus $P' \ll P$ follows from (i) because of 2.8.

b) (i′) $\Rightarrow P' \perp P$ by 2.6c; (ii′) $\Rightarrow P' \perp P$ is trivial.

For the converse, suppose that neither (i′) nor (ii′) is met. Then $H(\frac{1}{2}; \rho_n, \rho'_n) > 0$ for all n by 1.11b, and so the convergence of the series $\sum [1 - H(\frac{1}{2}; \rho_n, \rho'_n)]$ implies the convergence of the infinite product $\prod H(\frac{1}{2}; \rho_n, \rho'_n)$ toward a positive number: hence 1.74 implies $H(\frac{1}{2}; P, P') > 0$, and 1.11b again contradicts $P' \perp P$. □

Then comes a very interesting corollary, known as the *Kakutani alternative*:

2.38 **Corollary.** *Assume 1.71; assume further that* $\rho_n \sim \rho'_n$ *for all n. Then either* $P' \sim P$ *or* $P' \perp P$.

Proof. If $\sum_n [1 - H(\frac{1}{2}; \rho_n, \rho'_n)] < \infty$, 2.37a implies $P' \sim P$. Otherwise, 2.37b implies $P' \perp P$. □

2.39 **Remarks.** 1) In 2.37 one has a criterion for singularity, as well as for absolute continuity: this is in opposition with the general case (see 2.9), and is due to the particular structure of the problem at hand. The same will also hold for the continuous-time analogue, i.e. laws of PII (see 4.33).

2) There is a direct proof of 2.37(a) and of the sufficient part of 2.37(b), using 1.11 and 1.74 (as for the proof of the necessary part in 2.37(b)). □

3. Hellinger Processes for Solutions of Martingale Problems

In this section, we attempt to compute a version of the Hellinger process $h(\alpha; P, P')$ and of the processes $\imath(\psi; P, P')$ of § 1d, when P and P' are solutions of two martingale problems in the sense of § III.2, based upon the same fundamental process X. One wishes to compute them in terms of the characteristics of X under P and P', only.

Our setting and notation are introduced in § 3a. In § 3b we analyse the general problem, and we give an explicit form for $h(\alpha; P, P')$ and $\imath(\psi; P, P')$ when P and P' are dominated by a measure Q under which a martingale representation holds: this assumption is very difficult to check in most cases; however, it is met in some cases, as for example in the case of point processes. In § 3c we compute $h(\alpha; P, P')$

and $\mathscr{i}(\psi; P, P')$ when local uniqueness (see § III.2d) holds for the two martingale problems associated with P and P'.

§ 3a. The General Setting

The filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ is endowed with a càdlàg d-dimensional process $X = (X^i)_{i \le d}$ and a sub-σ-field $\mathscr{H}$, such that (see III.2.12):

3.1 $$\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s^0, \quad \text{where } \mathscr{F}_t^0 = \mathscr{H} \vee \sigma(X_s\colon s \le t), \text{ and } \mathscr{F}_\cdot = \mathscr{F}_{\infty -}$$

We denote by $\mu = \mu^X$ the random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ associated with the jumps of X.

We fix a truncation function $h \in \mathscr{C}_t^d$ (see II.2.3) and *two triplets* (B, C, ν) *and* (B', C', ν') *with the following* (the same as III.2.3):

3.2 (i) B and B' are predictable, d-dimensional, with finite variation over finite intervals, and $B_0 = B'_0 = 0$;

(ii) $C = (C^{ij})_{i,j \le d}$ and $C' = (C'^{ij})_{i,j \le d}$ are two continuous adapted processes with $C_0 = C'_0 = 0$, such that $C_t - C_s$ and $C'_t - C'_s$ are nonnegative symmetric $d \times d$ matrices for all $s \le t$;

(iii) ν is a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$ which charges neither $\{0\} \times \mathbb{R}^d$ nor $\mathbb{R}_+ \times \{0\}$, and such that $|x|^2 \wedge 1 * \nu_t(\omega) < \infty$ and $a_t(\omega) := \nu(\omega; \{t\} \times \mathbb{R}^d) \le 1$ and $\Delta B_t(\omega) = \int \nu(\omega; \{t\} \times dx) h(x)$; ν' has the same properties with B' instead of B and $a'_t(\omega) = \nu'(\omega; \{t\} \times \mathbb{R}^d)$. $\square$

We also consider two probability measures P and P' on $(\Omega, \mathscr{F})$, which are solutions to the martingale problems $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ and $\mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$ respectively: this means that P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$, and that X is a semimartingale with characteristics (B, C, ν) (resp. (B', C', ν')) under P (resp. P').

Next, we associate to those two triplets several random sets and processes.

Firstly, in view of I.3.13 it is easy to find an increasing predictable finite-valued process A and two processes c, c' taking values in the set of nonnegative symmetric $d \times d$ matrices and predictable, such that

3.3 $$C = c \cdot A, \quad C' = c' \cdot A \quad \text{up to a } (P + P')\text{-evanescent set.}$$

Secondly, let λ be a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, such that $|x|^2 \wedge 1 * \lambda_t < \infty$ for all $t < \infty$ and that

3.4 $$\nu \ll \lambda, \quad \nu' \ll \lambda$$

(take e.g. $\lambda = \nu + \nu'$). Using notation III.3.15, we deduce from 3.4 that if $Q = (P + P')/2$, then $M_\nu^Q \ll M_\lambda^Q$ on $(\tilde{\Omega}, \mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{R}^d)$; so there is a $\tilde{\mathscr{P}}$-measurable nonnegative function U such that $U = dM_\nu^Q / dM_\lambda^Q$ in restriction to $(\tilde{\Omega}, \tilde{\mathscr{P}})$. Then for all nonnegative predictable W,

$$E_Q(W * (U \cdot \lambda)_\infty) = E_Q((WU) * \lambda_\infty) = M^Q_\lambda(WU) = M^Q_\nu(W) = E_Q(W * \nu_\infty).$$

Since ν and $U \cdot \lambda$ are predictable measures, it follows from II.1.8 that $U \cdot \lambda = \nu$ Q-a.s. So we have got a nonnegative predictable function U on $\tilde{\Omega}$, and similarly we obtain a function U' with the same properties, such that

3.5 $$\nu = U \cdot \lambda, \quad \nu' = U' \cdot \lambda \quad (P + P')\text{-a.s.}$$

(*Note*: Doob's Theorem on "measurable Radon-Nikodym derivatives" [36] allows to find an $\mathscr{F} \otimes \mathscr{R}_+ \otimes \mathscr{R}^d$-measurable U such that $\nu = U \cdot \lambda$ identically; whether one can find a predictable U with $\nu = U \cdot \lambda$ identically is not clear!)

Thirdly, we define a predictable random set Σ and a predictable process $\tilde{B} = (\tilde{B}^i)_{i \le d}$ on Σ by

3.6 $$\begin{cases} \Sigma = \{(\omega, t) \colon |h(x)(U - U')| * \lambda_t(\omega) < \infty\} \\ \tilde{B}_t = B_t - B'_t - h(x)(U - U') * \lambda_t \quad \text{for } t \in \Sigma. \end{cases}$$

Observe that the process $K = |h(x)(U - U')| * \lambda$ is predictable, and is a generalized increasing process in the sense of III.5.8; moreover, up to a $(P + P')$-evanescent set, we have $a_t = \int U(t,x)\lambda(\{t\} \times dx) \le 1$, and similarly for U', while h is bounded, so $K_t - K_{t-}$ is bounded: therefore if $\sigma_n = \inf(t \colon K_t \ge n)$ we have $K_{\sigma_n} < \infty$ and $\Sigma = \bigcup_n [\![0, \sigma_n]\!]$ $(P + P')$-a.s. Furthermore, $\tilde{B}$ is *continuous* on Σ, again up to a $(P + P')$-evanescent set.

Finally, there is a decomposition:

3.7 $$\tilde{B}^i = \Big(\sum_{j \le d} c^{ij} \tilde{\beta}^j\Big) \cdot A + \tilde{b}^i \cdot A + \tilde{B}'^i \quad (P + P')\text{-a.s. on } \Sigma, \text{ where}$$

(i) $\tilde{\beta}$, $\tilde{b}$ are predictable; $\tilde{B}'$ is continuous predictable with finite variation over the compact subsets of Σ;

(ii) $d\tilde{B}'^i_t$ and dA_t are $(P + P')$-a.s. mutually singular on Σ;

(iii) for all $(\omega, t) \in \Sigma$, the vector $\tilde{b}_t(\omega)$ is orthogonal to the image of $\mathbb{R}^d$ by the linear map associated to the matrix $c_t(\omega)$. □

Note that $\tilde{\beta}$ and $\tilde{b}$ are not unique, but the decomposition 3.7 is unique up to a $(P + P')$-evanescent set on Σ. To see that 3.7 exists, consider first a $(P + P')$-a.s. (pathwise) Lebesgue decomposition on Σ: $d\tilde{B}^i_t = d\tilde{B}'^i_t + \tilde{b}'^i_t\, dA_t$ of $d\tilde{B}^i_t$ relative to dA_t, the predictability of $\tilde{B}'$ and $\tilde{b}'$ being insured by I.3.13; then decompose $\tilde{b}'$ as $\tilde{b}' = c\tilde{\beta} + \tilde{b}$ with $\tilde{b}$, $\tilde{\beta}$ as above.

We also define a stopping time τ by

3.8 $$\tau = \inf(t \colon \text{either } t \notin \Sigma, \text{ or } C_t \ne C'_t, \text{ or } t \in \Sigma \text{ and } \tilde{b} \cdot A_t + \tilde{B}'_t \ne 0).$$

We obviously can, and will, assume that $c \equiv c'$ on $[\![0, \tau]\!]$.

We end this long list of notation with the definition of several processes that have vocation to being versions of $h(\alpha; P, P')$ or $\imath(\psi; P, P')$. Below, $\alpha \in (0,1)$ and ψ is always a function satisfying 1.40.

3.9 $$h^0(\alpha) = \frac{\alpha(1-\alpha)}{2}(\tilde{\beta}\cdot c\cdot\tilde{\beta})1_\Sigma\cdot A + \varphi_\alpha(U,U')*\lambda + \sum_{s\le\cdot}\varphi_\alpha(1-a_s, 1-a'_s)$$

3.10 $$h^0(0) = U'1_{\{U=0\}}*\lambda + \sum_{s\le\cdot}(1-a'_s)1_{\{a_s=1\}}$$

3.11 $$\imath^0(\psi) = U'\psi\left(\frac{U}{U'}\right)*\lambda + \sum_{s\le\cdot}(1-a'_s)\psi\left(\frac{1-a_s}{1-a'_s}\right).$$

(recall that by convention $0/0 = 0$ and $a/0 = \infty$ for $a > 0$). Observe that 3.10 is also 3.9 taken for $\alpha = 0$. For $\alpha = 1/2$ we obtain:

3.12 $$\begin{aligned} h^0\left(\frac{1}{2}\right) &= \frac{1}{8}(\tilde{\beta}\cdot c\cdot\tilde{\beta})1_\Sigma\cdot A + \frac{1}{2}(\sqrt{U}-\sqrt{U'})^2*\lambda \\ &\quad + \frac{1}{2}\sum_{s\le\cdot}(\sqrt{1-a_s}-\sqrt{1-a'_s})^2. \end{aligned}$$

3.13 **Remark.** If $\nu_t(\omega)$ and $\nu'_t(\omega)$ denote the restrictions of the measures $\nu(\omega;\cdot)$ and $\nu'(\omega;\cdot)$ to $[0,t]\times\mathbb{R}^d$, then $\frac{1}{2}(\sqrt{U}-\sqrt{U'})^2*\lambda_t = \rho^2(\nu_t,\nu'_t)$, where ρ is the *Hellinger distance* given by 1.5. □

3.14 **Remark.** It is obvious that changing λ, U, U' modifies Σ and $\tilde{B}$ on a $(P+P')$-evanescent set only (as long as 3.5 remains true); similarly $(\tilde{\beta}\cdot c\cdot\tilde{\beta})1_\Sigma\cdot A$ does not depend upon the choice of $\tilde{\beta}$ in 3.7. Since moreover $\varphi_\alpha(ub,vb) = b\varphi_\alpha(u,v)$, it is then obvious that $h^0(\alpha)$ and $\imath^0(\psi)$ are $(P+P')$-a.s. uniquely determined, no matter which λ, U, U', $\tilde{\beta}$ are chosen. □

§ 3b. The Case Where *P* and *P'* Are Dominated by a Measure Having the Martingale Representation Property

1. In this section, we suppose that Q is another probability measure on $(\Omega,\mathscr{F})$, such that $P \overset{\text{loc}}{\ll} Q$ and $P' \overset{\text{loc}}{\ll} Q$ and X is a Q-semimartingale. We denote by $(\bar{B},\bar{C},\bar{\nu})$ the Q-characteristics of X. Without loss of generality, we can assume that $\bar{C} = \bar{c}\cdot A$ Q-a.s., where $\bar{c} = (\bar{c}^{ij})_{i,j\le d}$ is a predictable matrix-valued process and A is the same than in 3.3 (since in 3.3 we can always replace A by a "bigger" process).

We call z and z' the density processes of P and P' with respect to Q. We have 1.13, and we use the notation Γ, Γ', Γ'', R_n, R'_n, S_n, S of 1.15.

In view of Theorem III.3.24 and of III.3.7 there are two predictable processes $\beta = (\beta^i)_{i\le d}$ and $\beta' = (\beta'^i)_{i\le d}$ and two nonnegative $\tilde{\mathscr{P}}$-measurable functions Y and Y' on $\tilde{\Omega} = \Omega\times\mathbb{R}_+\times\mathbb{R}^d$ (recall that $\tilde{\mathscr{P}} = \mathscr{P}\otimes\mathscr{R}^d$), with

3.15 $$\begin{cases} \langle z^c,(X^i)^c\rangle = \left(z_-\sum_{j\le d}\bar{c}^{ij}\beta^j\right)\cdot A, & \langle z'^c,(X^i)^c\rangle = \left(z'_-\sum_{j\le d}\bar{c}^{ij}\beta'^j\right)\cdot A \\ M^Q_\mu(\Delta z|\tilde{\mathscr{P}}) = z_-Y, & M^Q_\mu(\Delta z'|\tilde{\mathscr{P}}) = z'_-Y' \end{cases}$$

($z^c, z'^c, (X^i)^c$ are the continuous martingale parts relative to Q) and

3.16 $$\begin{cases} |h(x)(Y-1)| * \bar{\nu}_t < \infty & Q\text{-a.s. on } \{t \in \Gamma\} \\ |h(x)(Y'-1)| * \bar{\nu}_t < \infty & Q\text{-a.s. on } \{t \in \Gamma'\}, \end{cases}$$

3.17 $$\begin{cases} \left.\begin{array}{l} B^i = \overline{B}^i + \left(\sum_{j\le d} \bar{c}^{ij}\beta^j\right) \cdot A + h^i(x)(Y-1) * \bar{\nu}, \\ C = \overline{C}, \quad \nu = Y \cdot \bar{\nu} \end{array}\right\} & Q-\text{a.s. on } \Gamma \\ \left.\begin{array}{l} B'^i = \overline{B}^i + \left(\sum_{j\le d} \bar{c}^{ij}\beta'^j\right) \cdot A + h^i(x)(Y'-1) * \bar{\nu}, \\ C' = \overline{C}, \quad \nu' = Y' \cdot \bar{\nu} \end{array}\right\} & Q-\text{a.s. on } \Gamma' \end{cases}$$

3.18 $$\{\bar{a} = 1\} \cap \Gamma \subset \{a = 1\}, \quad \{\bar{a} = 1\} \cap \Gamma' \subset \{a' = 1\} \quad Q\text{-a.s.}$$

where $\bar{a}_t = \bar{\nu}(\{t\} \times \mathbb{R}^d)$. We obviously may, and will, assume that $c \equiv c' \equiv \bar{c}$ on $[\![0,\tau]\!] \cap \Gamma''$, where τ is defined by 3.8.

3.19 **Lemma.** *$\Gamma'' \subset \Sigma \cap [\![0,\tau]\!]$ up to a Q-evanescent set, and a version of $\tilde{\beta}$ on Γ'' is $\tilde{\beta} 1_{\Gamma''} = (\beta - \beta') 1_{\Gamma''}$.*

Proof. We have $Y \cdot \bar{\nu} = U \cdot \lambda$ and $Y' \cdot \bar{\nu} = U' \cdot \lambda$ Q-a.s. on $\Gamma'' \times \mathbb{R}^d$ by 3.17 and 3.5. Hence on $\{t \in \Gamma''\}$ we have Q-a.s.:

$$\begin{aligned} |h(x)(U - U')| * \lambda_t &= |h(x)(Y - Y')| * \bar{\nu}_t \\ &\le |h(x)(Y-1)| * \bar{\nu}_t + |h(x)(Y'-1)| * \bar{\nu}_t < \infty \end{aligned}$$

and so $\Gamma'' \subset \Sigma$ up to a Q-evanescent set. Similarly, 3.17 yields Q-a.s. on $\{t \in \Gamma''\}$:

$$\begin{aligned} \tilde{B}_t &= B_t - B'_t - h(x)(Y - Y') * \bar{\nu}_t \\ &= B_t - h(x)(Y-1) * \bar{\nu}_t - B'_t + h(x)(Y'-1) * \bar{\nu}_t \\ &= \left[\sum_{j\le d} \bar{c}^{\cdot j}(\beta^j - \beta'^j)\right] \cdot A_t, \end{aligned}$$

while 3.17 again yields $C_t = C'_t = \overline{C}_t$ Q-a.s. on $\{t \in \Gamma''\}$. Comparing to 3.7, we thus obtain $\Gamma'' \subset [\![0,\tau]\!]$ up to a Q-evanescent set, and $\beta - \beta'$ is a version of $\tilde{\beta}$ on Γ''. □

3.20 **Theorem.** *Assume that the density processes z and z' have the following representation* (with $0/0 = 0$; recall 3.18, so that the next formulae are nothing else than III.5.20 for z and z'):

3.21 $$\begin{cases} z = z_0 + (z_-\beta) \cdot X^c + z_-\left(Y - 1 + \dfrac{a - \bar{a}}{1 - \bar{a}}\right) * (\mu - \bar{\nu}) \\ z' = z'_0 + (z'_-\beta') \cdot X^c + z'_-\left(Y' - 1 + \dfrac{a' - \bar{a}}{1 - \bar{a}}\right) * (\mu - \bar{\nu}). \end{cases}$$

a) *If* $\alpha \in [0,1)$, *the process* $h^0(\alpha)$ *defined by 3.9 or 3.10 is a version of the Hellinger process* $h(\alpha; P, P')$.

b) *If* ψ *satisfies 1.40, the process* $\imath^0(\psi)$ *defined by 3.11 is a version of the process* $\imath(\psi; P, P')$ *of 1.46.*

We begin with a lemma:

3.22 **Lemma.** *Assume 3.21, and call* $\nu^{(z,z')}$ *the third* Q*-characteristic of the bi-dimensional process* (z, z'). *Then for all* $\mathscr{P} \otimes \mathscr{R}^2$*-measurable nonnegative functions* W *on* $\Omega \times \mathbb{R}_+ \times \mathbb{R}^2$ *with* $W(\omega, t, 0, 0) = 0$,

3.23
$$\begin{aligned} W * \nu^{(z,z')} &= W(\cdot, z_-(Y-1), z'_-(Y'-1)) * \bar{\nu} \\ &\quad + \sum_{s \leq \cdot} (1 - \bar{a}_s) W\left(s, z_{s-} \frac{\bar{a}_s - a_s}{1 - \bar{a}_s}, z'_{s-} \frac{\bar{a}_s - a'_s}{1 - \bar{a}_s}\right). \end{aligned}$$

Proof. The formula 3.23 clearly defines a predictable random measure. Hence by the very definition of $\nu^{(z,z')}$ it suffices to prove that for all W as above, we have:

3.24 $E_Q\left(\sum_{(s)} W(s, \Delta z_s, \Delta z'_s)\right) = Q$-expectation of the right-hand side of 3.23, taken at time $+\infty$.

Now, we easily deduce from 3.21 (as in the proof of III.5.10 for instance!) that outside a Q-evanescent set,

$$\Delta z_t = z_{t-}\left[(Y(t, \Delta X_t) - 1) 1_{\{\Delta X_t \neq 0\}} + \frac{\bar{a}_t - a_t}{1 - \bar{a}_t} 1_{\{\Delta X_t = 0\}}\right]$$

and a similar formula for $\Delta z'_t$. Let (T_n) be a sequence of predictable times, which exhausts Q-a. s. the predictable thin set $\{\bar{a} > 0\}$. Then the left-hand side of 3.24 is

$$\begin{aligned} E_Q\Bigg[& W(\cdot, z_-(Y-1), z'_-(Y'-1)) * \mu_\infty \\ & + \sum_n 1_{\{\Delta X_{T_n} = 0\}} W\left(T_n, z_{T_n-} \frac{\bar{a}_{T_n} - a_{T_n}}{1 - \bar{a}_{T_n}}, z'_{T_n-} \frac{\bar{a}_{T_n} - a'_{T_n}}{1 - \bar{a}_{T_n}}\right)\Bigg]. \end{aligned}$$

Using the definition of $\bar{\nu}$ ($= Q$-compensator of μ) and the property $1 - \bar{a}_{T_n} = Q(\Delta X_{T_n} = 0 | \mathscr{F}_{T_n-})$ and the fact that $W(T_n, \cdot, \cdot)$ above is $\mathscr{F}_{(T_n)-}$-measurable, the left-hand side of 3.24 becomes

$$\begin{aligned} E_Q\Bigg[& W(\cdot, z_-(Y-1), z'_-(Y'-1)) * \bar{\nu}_\infty \\ & + \sum_n (1 - \bar{a}_{T_n}) W\left(T_n, z_{T_n-} \frac{\bar{a}_{T_n} - a_{T_n}}{1 - \bar{a}_{T_n}}, z'_{T_n-} \frac{\bar{a}_{T_n} - a'_{T_n}}{1 - \bar{a}_{T_n}}\right)\Bigg], \end{aligned}$$

which obviously equals the Q-expected value of the right-hand side of 3.23, evaluated at time $+\infty$. Hence we are finished. □

Proof of Theorem 3.20. We have seen in Remark 3.14 that $h^0(\alpha)$ and $\imath^0(\psi)$ do not depend upon λ, U, U', $\tilde{\beta}$, so long as 3.5 and 3.7 holds. So if we replace λ by $1_{\Gamma''}\cdot\bar{\nu} + 1_{\Gamma''^c}\cdot\lambda$, 3.17 shows that 3.5 still holds, provided we replace U and U' by $Y1_{\Gamma''} + U1_{\Gamma''^c}$ and $Y'1_{\Gamma''} + U'1_{\Gamma''^c}$, respectively.

Recall also that $h(\alpha; P, P')$ and $\imath(\psi; P, P')$ are uniquely determined on Γ'' only. Hence, since $\Gamma'' \subset \Sigma$ by 3.19, and if we compare 1.34 and 3.9 (resp. 1.48 and 3.11), we see that it is enough to prove the following (recall that $Y = U$, $Y' = U'$, $\lambda = \bar{\nu}$ on Γ'', and that $h^0(0) = \imath^0(\psi)$ for $\psi(u) = 1_{\{u=0\}}$):

3.25 $$(\tilde{\beta}\cdot c\cdot\tilde{\beta})\cdot A = \frac{1}{z_-^2}\cdot\langle z^c, z^c\rangle - \frac{2}{z_-z'_-}\cdot\langle z^c, z'^c\rangle + \frac{1}{z'^2_-}\cdot\langle z'^c, z'^c\rangle \quad \text{on } \Gamma''$$

3.26 $$\begin{aligned}&\varphi_\alpha(Y, Y') * \bar{\nu} + \sum_{s\le\cdot} \varphi_\alpha(1 - a_s, 1 - a'_s)\\ &= \varphi_\alpha\left(1 + \frac{x}{z_-}, 1 + \frac{y}{z'_-}\right) * \nu^{(z,z')} \quad \text{on } \Gamma''\end{aligned}$$

3.27 $$\begin{aligned}&Y'\psi\left(\frac{Y}{Y'}\right) * \bar{\nu} + \sum_{s\le\cdot} (1 - a'_s)\psi\left(\frac{1 - a_s}{1 - a'_s}\right)\\ &= \left(1 + \frac{y}{z'_-}\right)\psi\left(\frac{1 + x/z_-}{1 + y/z'_-}\right) * \nu^{(z,z')} \quad \text{on } \Gamma''.\end{aligned}$$

Firstly, 3.21 and III.4.5 yield on Γ'':

$$\langle z^c, z^c\rangle = z_-^2(\beta\cdot c\cdot\beta)\cdot A,$$
$$\langle z^c, z'^c\rangle = z_-z'_-(\beta\cdot c\cdot\beta')\cdot A,$$
$$\langle z'^c, z'^c\rangle = z'^2_-(\beta'\cdot c\cdot\beta')\cdot A$$

(recall that $c = c' = \bar{c}$ on $[\![0, \tau]\!]$ and $\Gamma'' \subset [\![0, \tau]\!]$). Since $\tilde{\beta} = \beta - \beta'$ on Γ'' by 3.19, we deduce 3.25.

Next, 3.23 gives

$$\begin{aligned}\varphi_\alpha\left(1 + \frac{x}{z_-}, 1 + \frac{y}{z'_-}\right) * \nu^{(z,z')} &= \varphi_\alpha(Y, Y') * \bar{\nu} + \sum_{s\le\cdot} (1 - \bar{a}_s)\varphi_\alpha\left(\frac{1 - a_s}{1 - \bar{a}_s}, \frac{1 - a'_s}{1 - \bar{a}_s}\right)\\ &= \varphi_\alpha(Y, Y') * \bar{\nu} + \sum_{s\le\cdot} \varphi_\alpha(1 - a_s, 1 - a'_s)1_{\{\bar{a}_s<1\}}\end{aligned}$$

because $b\varphi_\alpha(u/b, v/b) = \varphi_\alpha(u, v)$. Then 3.18 and the property $\varphi_\alpha(0, 0) = 0$ allow to deduce 3.26. Finally, 3.23 and 3.18 again, plus the fact that $\psi(1) = 0$ by 1.40, immediately give 3.27. □

The property 3.21 is particularly difficult to check, unless some (rather strong) additional assumptions are made on Q, or on P and P'. Below, we present two examples.

2. The first example concerns the case where all Q-local martingales have the representation property, relative to X, in the sense of III.4.22.

3.28 **Theorem.** *In addition to* $P \overset{\text{loc}}{\ll} Q$ *and* $P' \overset{\text{loc}}{\ll} Q$ *and to the fact that* X *is a* Q*-semimartingale, assume that all local martingales on the basis* $(\Omega, \mathscr{F}, \mathbf{F}, Q)$ *have the representation property relative to* X. *Then*

a) *if* $\alpha \in [0,1)$ *the process* $h^0(\alpha)$ *defined by 3.9 or 3.10 is a version of the Hellinger process* $h(\alpha; P, P')$;

b) *if* ψ *meets 1.40, the process* $\imath^0(\psi)$ *defined by 3.11 is a version of the process* $\imath(\psi; P, P')$ *of 1.46.*

Proof. Theorem III.5.19 implies that z and z' meet 3.21, so the claim follows from 3.20. □

Again, the representation property above is difficult to check, because a-priori we have no simple way of choosing a "nice" measure Q that dominates P and P'. There are, however, some cases where this is possible, and we expound such an example (the case of point processes) in Section 4.

3. The case where $P' \overset{\text{loc}}{\ll} P$. In addition to the assumptions of § 3a, we suppose from now on until the end of this subsection that $P' \overset{\text{loc}}{\ll} P$. Then

3.29
$$\begin{cases} B'^i = B^i + \left(\sum_{j \le d} c^{ij}\beta^j\right) \cdot A + h^i(x)(Y-1) * \nu \\ C' = C \\ \nu' = Y \cdot \nu \end{cases}$$

holds up to a P'-evanescent set, where $\beta = (\beta^i)_{i \le d}$ is a predictable process and Y is a nonnegative $\tilde{\mathscr{P}}$-measurable function on $\tilde{\Omega}$.

We denote by Z the *density process* of P', relative to P, and

3.30
$$\tau_n = \inf(t: Z_t < 1/n).$$

We also recall that in III.5.10 we have introduced a process N, defined on a random set Δ which contains $\bigcup_n [\![0, \tau_n]\!]$ (see III.5.9) such that

3.31
$$N^{\tau_n} = (\beta 1_{[\![0,\tau_n]\!]}) \cdot X^c + \left(Y - 1 + \frac{a' - a}{1 - a} 1_{\{a<1\}}\right) 1_{[\![0,\tau_n]\!]} * (\mu - \nu);$$

the stochastic integrals above are relative to P, and are meaningful. Furthermore, we have given *two conditions* (Theorems III.5.19 and III.5.32) *under which*

3.32 $$Z_t = \begin{cases} Z_0 \exp\left(N_t - \frac{1}{2}(\beta\cdot c\cdot\beta)\cdot A_t\right) \prod_{s\le t} (1+\Delta N_s) e^{-\Delta N_s} & \text{if } t \le \tau_n \text{ for some } n \in \mathbb{N}^* \\ 0 & \text{otherwise.} \end{cases}$$

3.33 **Theorem.** *Assume that $P' \overset{\text{loc}}{\ll} P$ and that 3.32 holds.*
a) *If $\alpha \in (0,1)$, a version of $h(\alpha; P, P')$ is*

3.34 $$h(\alpha) = \frac{\alpha(1-\alpha)}{2}(\beta\cdot c\cdot\beta)\cdot A + \varphi_\alpha(1, Y) * \nu + \sum_{s\le\cdot} \varphi_\alpha(1-a_s, 1-a'_s).$$

In particular,

3.35 $$h\left(\frac{1}{2}\right) = \frac{1}{8}(\beta\cdot c\cdot\beta)\cdot A + \frac{1}{2}(1-\sqrt{Y})^2 * \nu + \frac{1}{2}\sum_{s\le\cdot} (\sqrt{1-a_s} - \sqrt{1-a'_s})^2$$

b) *If ψ satisfies 1.40, a version of $\imath(\psi; P, P')$ is*

3.36 $$\imath(\psi) = Y\psi\left(\frac{1}{Y}\right) * \nu + \sum_{s\le\cdot} (1-a'_s)\psi\left(\frac{1-a_s}{1-a'_s}\right)$$

(recall that in the present situation $h(0; P, P') = 0$ by 1.55c).

Proof. We start with $Q = P$, hence $z = 1$ and $z' = Z$. Then z obviously satisfies 3.21 with 0 and 1 instead of β and Y. Moreover, 3.31 and 3.32 and the fact that $\{a = 1\} \cap (\bigcup_n [\![0, \tau_n]\!]) \subset \{a' = 1\}$ up to a P-evanescent set show that (with $0/0 = 0$):

$$z'^{\tau_n} = z'_0 + (z'_-\beta 1_{[\![0,\tau_n]\!]})\cdot X^c + z'_-\left(Y - 1 + \frac{a'-a}{1-a}\right) 1_{[\![0,\tau_n]\!]} * (\mu - \nu).$$

Since $z'_- = 0$ outside $\bigcup_n [\![0, \tau_n]\!]$, we deduce that z' satisfies 3.21 with β and Y instead of β' and Y': therefore, Theorem 3.20 applies.

Now, in 3.4 we choose $\lambda = \nu$, and we use 3.29: 3.5 holds with $U = 1$ and $U' = Y$, and 3.7 holds with $\tilde\beta = -\beta$, while we know that $\bigcup_n [\![0, \tau_n]\!]$ (which here equals Γ'') is contained in Σ. Therefore the right-hand sides of 3.34 and 3.9 (resp. 3.36 and 3.11) coincide on Γ'', and this is enough for the claimed properties. □

3.37 **Remark.** In III.5.7 we have introduced a process H which is "comparable" to $h(1/2)$. Indeed we have seen in III.5.17 that $\bigcup_n [\![0, \tau_n]\!] \subset [\![0, \sigma[\![$ where $\sigma = \inf(t\colon a_t = 1 \text{ and } a'_t < 1)$, so $h'(1/2) = 1_{[\![0,\sigma[\![}\cdot h(1/2)$ is another version of $h(\frac{1}{2}; P, P')$, and under 3.32 we have:

$$2h'(\tfrac{1}{2}) \le H \le 8h'(\tfrac{1}{2}).$$ □

This remark shows that Theorem III.5.34 is closely related to Corollary 2.8. Indeed we have:

3.38 **Theorem.** *Assume that 3.29 holds identically, that III.5.29 holds, that 3.1 holds, and that local uniqueness holds for the martingale problem $\mathfrak{s}(\mathscr{H}, X|P'_H; B', C', \nu')$, with P' as its unique solution. Assume further that $P' \overset{\text{loc}}{\ll} P$, and define H by III.5.17. Then $P' \ll P$ if and only if $P'(H_\infty < \infty) = 1$.*

Proof. The sufficient part comes from III.5.34. We also have 3.32 (apply Theorem III.5.32), hence $2h'(1/2) \leq H \leq 8h'(1/2)$: therefore both the necessary condition and the sufficient conditions follow from 2.8. □

4. In general, the process $h^0(\alpha)$ of 3.9 is *not* a version of $h(\alpha; P, P')$. However, with no assumption at all (but, of course, within the setting of § 3a), we have the following:

3.39 **Theorem.** a) *If* $\alpha \in [0,1)$, *there is a version $h(\alpha)$ of $h(\alpha; P, P')$ such that $h(\alpha) - h^0(\alpha)$ is an increasing nonnegative process.*

b) *If ψ meets 1.40 and is convex, there is a version $\imath(\psi)$ of $\imath(\psi; P, P')$ such that $\imath(\psi) - \imath^0(\psi)$ is an increasing nonnegative process.*

Proof. (i) Let Q be any measure such that $P \overset{\text{loc}}{\ll} Q$ and $P' \overset{\text{loc}}{\ll} Q$ and X is a Q-semimartingale (e.g.: $Q = (P + P')/2$). We use all the notation of the beginning of the present § 3b. We set $J = \{\bar{a} > 0\}$ and we call (T_n) a sequence of predictable times that exhausts Q-a.s. the thin predictable set J. We choose λ as in the proof of 3.20, i.e. $\lambda = \bar{\nu}$ and $U = Y$ and $U' = Y'$ on Γ'', and we also choose $\tilde{\beta} = \beta - \beta'$ on Γ''.

(ii) We will prove (a) and (b) at the same time. We call K the process $h^0(\alpha)$ (resp. $\imath^0(\psi)$), and we call H the process 1.34 (resp. 1.48): it clearly suffices to prove that $1_{\Gamma''} \cdot H - 1_{\Gamma''} \cdot K$ is increasing nonnegative.

We set $\theta(u,v) = \varphi_\alpha(u,v)$ in case (a), $\theta(u,v) = v\psi(\frac{u}{v})$ in case (b). The function θ is convex on $\mathbb{R}^2_+$. Due to 3.9 and 3.11, we have

$$1_{\Gamma''} \cdot K = \frac{\alpha(1-\alpha)}{2} K^1 + K^2 + \sum_n k_n 1_{[\![T_n, \infty[\![}, \quad \text{where}$$

$$K^1 = (\tilde{\beta} \cdot c \cdot \tilde{\beta}) 1_{\Gamma''} \cdot A \quad (\text{resp.} = 0) \text{ in case (a) (resp. (b))},$$

$$K^2 = \theta(Y, Y') 1_{\Gamma'' \cap J^c} * \bar{\nu}$$

$$k_n = 1_{\Gamma''}(T_n) \left[\int \bar{\nu}(\{T_n\} \times dx) \theta(Y, Y')(T_n, x) + \theta(1 - a_{T_n}, 1 - a'_{T_n}) \right].$$

Similarly, 1.34 and 1.48 give

$$1_{\Gamma''}\cdot H = \frac{\alpha(1-\alpha)}{2} H^1 + H^2 + \sum_n h_n 1_{[\![T_n,\infty[\![}, \quad \text{where}$$

$$H^1 = 1_{\Gamma''}\cdot\left\{\frac{1}{z_-^2}\cdot\langle z^c, z^c\rangle - \frac{2}{z_- z'_-}\cdot\langle z^c, z'^c\rangle + \frac{1}{z'^2_-}\cdot\langle z'^c, z'^c\rangle\right\}$$

(resp. $= 0$) in case (a) (resp. (b)),

$$H^2 = \theta\left(1+\frac{x}{z_-}, 1+\frac{y}{z'_-}\right) 1_{\Gamma''\cap J^c} * \nu^{(z,z')}$$

$$h_n = 1_{\Gamma''}(T_n)\int \nu^{(z,z')}(\{T_n\}, dx, dy)\theta\left(1+\frac{x}{z_{T_n-}}, 1+\frac{y}{z'_{T_n-}}\right).$$

It is clearly enough to prove that $H^i - K^i$ is increasing nonnegative for $i = 1, 2$, and that $k_n \le h_n$ Q-a.s. on $\{T_n < \infty\}$.

(iii) In general, we do not have 3.21, but we deduce from III.5.17a that (with $0/0 = 0$; recall 3.18):

3.40
$$\begin{cases} z = z_0 + (z_-\beta)\cdot X^c + z_-\left(Y - 1 + \dfrac{a-\bar a}{1-\bar a}\right) * (\mu - \bar\nu) + \tilde z \\[2ex] z' = z'_0 + (z'_-\beta')\cdot X^c + z'_-\left(Y' - 1 + \dfrac{a'-\bar a}{1-\bar a}\right) * (\mu - \bar\nu) + \tilde z', \end{cases}$$

where $\tilde z$ and $\tilde z'$ are two Q-local martingales with

3.41
$$\langle \tilde z^c, (X^i)^c\rangle = \langle \tilde z'^c, (X^i)^c\rangle = 0, \quad M_\mu^Q(\Delta\tilde z|\tilde{\mathscr P}) = M_\mu^Q(\Delta\tilde z'|\tilde{\mathscr P}) = 0.$$

(iv) The same computation that gave 3.25 now yields, because of 3.41:

$$\begin{aligned} H^1 &= K^1 + 1_{\Gamma''}\cdot\left\{\frac{1}{z_-^2}\cdot\langle \tilde z^c, \tilde z^c\rangle - \frac{2}{z_- z'_-}\cdot\langle \tilde z^c, \tilde z'^c\rangle + \frac{1}{z'^2_-}\cdot\langle \tilde z'^c, \tilde z'^c\rangle\right\} \\ &= K^1 + [M, M], \end{aligned}$$

(we have used $\tilde\beta = \beta - \beta'$), where $M = \dfrac{1}{z_-} 1_{\Gamma''}\cdot\tilde z^c - \dfrac{1}{z'_-} 1_{\Gamma''}\cdot\tilde z'^c$ is a "Q-local martingale on Γ''". It follows that $H^1 - K^1$ is increasing nonnegative.

(v) We shall now prove that for every nonnegative predictable process L,

3.42
$$E_Q(L\cdot K^2_\infty) \le E_Q(L\cdot H^2_\infty).$$

Since H^2 and K^2 are predictable and increasing, this will obviously imply that $H^2 - K^2$ is increasing nonnegative.

Exactly like in the proof of 3.22, we deduce from 3.40 that

3.43
$$\Delta z_t = z_{t-}\left[(Y(t,\Delta X_t) - 1)1_{\{\Delta X_t \neq 0\}} + \frac{\bar a_t - a_t}{1-\bar a_t} 1_{\{\Delta X_t = 0\}}\right] + \Delta\tilde z_t,$$

and a similar formula for $\Delta z'_t$. Then we deduce from the definitions of $\nu^{(z,z')}$ and of H^2 that

$$\begin{aligned}E_Q(L\cdot H^2_\infty) &= E_Q\left(\sum_{(s)} 1_{\Gamma''\cap J^c}(s)L_s\theta(1+\Delta z_s/z_{s-}, 1+\Delta z'_s/z'_{s-})\right)\\ &\geq E_Q\left(\sum_{(s)} 1_{\Gamma''\cap J^c}(s)L_s 1_{\{\Delta X_s\neq 0\}}\theta(1+\Delta z_s/z_{s-}, 1+\Delta z'_s/z'_{s-})\right)\\ &= M^Q_\mu(1_{\Gamma''\cap J^c}L\theta(Y+\Delta\tilde{z}/z_-, Y'+\Delta\tilde{z}'/z'_-)).\end{aligned}$$

The function θ being convex, Jensen's inequality yields on Γ'':

$$\begin{aligned}&M^Q_\mu(\theta(Y+\Delta\tilde{z}/z_-, Y'+\Delta\tilde{z}'/z'_-)|\tilde{\mathscr{P}})\\ &\quad\geq \theta\left(Y+\frac{1}{z_-}M^Q_\mu(\Delta\tilde{z}|\tilde{\mathscr{P}}), Y'+\frac{1}{z'_-}M^Q_\mu(\Delta\tilde{z}'|\tilde{\mathscr{P}})\right),\end{aligned}$$

which equals $\theta(Y, Y')$ by 3.41 (recall that J, L, Y, Y', Γ'' are predictable). Hence

$$\begin{aligned}E_Q(L\cdot H^2_\infty) &\geq M^Q_\mu(1_{\Gamma''\cap J^c}L\theta(Y, Y'))\\ &= E_Q\left(\sum_{(s)} 1_{\Gamma''\cap J^c}(s)L_s\theta(Y, Y')(s, \Delta X_s)1_{\{\Delta X_s\neq 0\}}\right)\\ &= E_Q(L1_{\Gamma''\cap J^c}\theta(Y, Y'))*\bar{\nu}_\infty] = E_Q(L\cdot K^2_\infty)\end{aligned}$$

by definition of $\bar{\nu}$ and of K^2. Thus we have proved 3.42.

(vi) It remains to prove that $h_n\geq k_n$. For simplicity, set $T=T_n$ and

3.44
$$\begin{cases}\delta = Y(T,\Delta X_T)1_{\{T<\infty,\Delta X_T\neq 0\}} + \dfrac{1-a_T}{1-\bar{a}_T}1_{\{T<\infty,\Delta X_T=0\}}\\[2ex] \delta' = Y'(T,\Delta X_T)1_{\{T<\infty,\Delta X_T\neq 0\}} + \dfrac{1-a'_T}{1-\bar{a}_T}1_{\{T<\infty,\Delta X_T=0\}},\end{cases}$$

and also $\mathscr{G}=\mathscr{F}_{T-}\vee\sigma(\Delta X_T)$. From II.1.18 and the definition of h_n, from 3.43 and its analogue for $\Delta z'$, and from the $\mathscr{G}$-measurability of δ and δ', we deduce that on the $\mathscr{F}_{T-}$-measurable set $\{T\in\Gamma''\}$,

$$\begin{aligned}h_n &= E_Q[\theta(1+\Delta z_T/z_{T-}, 1+\Delta z'_T/z'_{T-})|\mathscr{F}_{T-}]\\ &= E_Q\{E_Q[\theta(1+\Delta z_T/z_{T-}, 1+\Delta z'_T/z'_{T-})|\mathscr{G}]|\mathscr{F}_{T-}\}\end{aligned}$$

3.45
$$\geq E_Q\left[\theta\left(\delta+\frac{1}{z_{T-}}E_Q(\Delta\tilde{z}_T|\mathscr{G}), \delta'+\frac{1}{z'_{T-}}E_Q(\Delta\tilde{z}'_T|\mathscr{G})\right)\middle|\mathscr{F}_{T-}\right],$$

again because of the convexity of θ.

Now, any $\mathscr{G}$-measurable random variable L is of the form $L(\omega)=\bar{L}(\omega,\Delta X_T(\omega))$, with $\bar{L}$ being $\mathscr{F}_{T-}\otimes\mathscr{R}^d$-measurable on $\Omega\times\mathbb{R}^d$. Then $W(\omega,t,x)=\bar{L}(\omega,x)1_{\{t=T(\omega)\}}$ and $W'(\omega,t,x)=\bar{L}(\omega,0)1_{\{t=T(\omega)\}}$ are $\tilde{\mathscr{P}}$-measurable, and

$$\begin{aligned}
E_Q(\Delta\tilde{z}_T L 1_{\{T<\infty\}}) &= E_Q[\Delta\tilde{z}_T 1_{\{T<\infty, \Delta X_T \neq 0\}}(\bar{L}(\cdot, \Delta X_T) - \bar{L}(\cdot, 0))] \\
&\quad + E_Q(\Delta\tilde{z}_T 1_{\{T<\infty\}}\bar{L}(\cdot, 0)) \\
&= M^Q_\mu(\Delta\tilde{z}(W - W')) + E_Q(\Delta\tilde{z}_T 1_{\{T<\infty\}}\bar{L}(\cdot, 0)) = 0
\end{aligned}$$

because of 3.41 and of $E_Q(\Delta\tilde{z}_T|\mathscr{F}_{T-}) = 0$. Hence $E_Q(\Delta\tilde{z}_T|\mathscr{G}) = 0$, and similarly $E_Q(\Delta\tilde{z}'_T|\mathscr{G}) = 0$, on $\{T < \infty\}$. Thus 3.45 yields

$$h_n \geq E_Q(\theta(\delta, \delta')|\mathscr{F}_{T-}) \quad \text{on } \{T \in \Gamma''\}.$$

Next, we use 3.44 and the property $E_Q(V(T, \Delta X_T) 1_{\{\Delta X_T \neq 0\}}|\mathscr{F}_{T-}) = \int \bar{\nu}(\{T\} \times dx) V(T, x)$ on $\{T < \infty\}$ for all $\tilde{\mathscr{P}}$-measurable V, and we obtain on $\{T \in \Gamma''\}$:

$$\begin{aligned}
h_n &\geq E_Q[\theta(Y, Y')(T, \Delta X_T) 1_{\{\Delta X_T \neq 0\}}|\mathscr{F}_{T-}] \\
&\quad + \theta\left(\frac{1 - a_T}{1 - \bar{a}_T}, \frac{1 - a'_T}{1 - \bar{a}_T}\right) Q(\Delta X_T = 0|\mathscr{F}_{T-}) \\
&= \int \bar{\nu}(\{T\} \times dx)\theta(Y, Y')(T, x) + (1 - \bar{a}_T)\theta\left(\frac{1 - a_T}{1 - \bar{a}_T}, \frac{1 - a'_T}{1 - \bar{a}_T}\right),
\end{aligned}$$

which equals k_n because $b\theta(u/b, v/b) = \theta(u, v)$ for $b > 0$ in both cases (a) and (b), and 3.18 holds: so we are finished. □

§ 3c. The Case Where Local Uniqueness Holds

As seen before, the key point in the previous subsection is the representation property 3.21 for z and z'. This is insured by the representation of Q-local martingales, as in Theorem 3.28. However, Theorem III.5.32 also in principle implies a representation of type 3.21, provided we have *local uniqueness* for the martingale problems $\mathscr{s}(\mathscr{H}, X|P_H; B, C, \nu)$ and $\mathscr{s}(\mathscr{H}, X|P'_H; B', C', \nu')$: these are much more reasonable assumptions than the representation property for Q-local martingales.

However, we pointed out in III.5.33 that in order to apply III.5.32 unmistakenly, one ought to have 3.17 everywhere (and not only on Γ or Γ'). So we cannot just take an arbitrary measure Q that dominates P and P': the example in III.5.33 provides an instance where $h(\alpha; P, P')$ is not given by $h^0(\alpha)$ in 3.9, because the latter gives $h^0(\alpha) = 0$!

So in this subsection, we start the other way around, with the following a-priori new set of characteristics:

3.46 $$\bar{B} = \frac{B + B'}{2}, \quad \bar{C} = \frac{C + C'}{2}, \quad \bar{\nu} = \frac{\nu + \nu'}{2}.$$

These terms obviously satisfy 3.2, and $\bar{C} = \bar{c} \cdot A$ with the same process A than in 3.3 and $\bar{c} = (c + c')/2$, and $\bar{a}_t := \bar{\nu}(\{t\} \times \mathbb{R}^d) = (a_t + a'_t)/2$.

Remember that 3.1 holds, and we suppose in addition that

3.47 $$\sigma(X_0) \subset \mathscr{H} \,(\Leftrightarrow \mathscr{F}_0^0 = \mathscr{H}).$$

3.48 **Hypothesis.** Let $Q_H = (P_H + P'_H)/2$ (recall that P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$). Then the martingale problem $\mathscr{s}(\mathscr{H}, X|Q_H; \bar{B}, \bar{C}, \bar{\nu})$ has at least one solution. □

This hypothesis again might seem difficult to check; we shall see in fact that in many cases it is indeed fulfilled. We also need a technical assumption:

3.49 **Hypothesis.** If Σ, τ, $h^0(1/2)$ are defined by 3.6, 3.8 and 3.12, the predictable random interval $\Sigma' = \Sigma \cap \{h^0(\frac{1}{2}) < \infty\} \cap [\![0, \tau]\!]$ has the form $\Sigma' = \bigcup_n [\![0, \tau_n]\!]$ up to a $(P + P')$-evanescent set, for a sequence (τ_n) of stopping times that are *strict* (see III.2.35) and *predictable.* □

Recall that predictable and positive implies strict, for a stopping time. The reader might be surprised to see $\Sigma \cap [\![0, \tau]\!]$ in the definition of Σ', as $\tau \leq \inf(t\colon t \notin \Sigma)$ by 3.8: we always have $[\![0, \tau[\![\subset \Sigma$, but it may happen that $[\![0, \tau]\!] \not\subset \Sigma$.

Observe that 3.49 is met when Σ' is deterministic, and also under the following:

3.50 The "generalized increasing processes" $h^0(1/2)$ and $K := |h(U - U')| * \lambda$ do not jump to infinity (see III.5.8), and the stopping time

$$\hat{\tau} = \begin{cases} \tau & \text{if } \tau < \infty,\, K_\tau < \infty,\, h^0(1/2)_\tau < \infty \\ +\infty & \text{otherwise} \end{cases}$$

is strict and predictable (then 3.49 is met with $\tau_n = \hat{\tau} \wedge \inf(t\colon K_t + h^0(1/2)_t \geq n)$). □

These assumptions allow to compute $h(\alpha; P, P')$. These processes play a central rôle in the results of Section 2. But in them, we also have to evaluate the probability $P'(G_0)$, where G_0 is defined in 2.2. This may seem as a simple matter! however, in most cases the σ-field $\mathscr{H}$ is *strictly* included into $\mathscr{F}_0$ (typically $\mathscr{H} = \mathscr{F}_0^0 = \sigma(X_0)$), and the "natural" data concerning time 0 are P_H and P'_H. And of course, even if $P'_H \ll P_H$ it is not clear whether $P'_0 \ll P_0$ $(\Leftrightarrow P'(G_0) = 1)$.

We will settle this question together with the computation of $h(\alpha; P, P')$. For that, and likewise in 2.2, we introduce a set:

3.51 $$G_H \in \mathscr{H} \text{ such that } \begin{cases} P_H \sim P'_H & \text{in restriction to } G_H \\ P_H \perp P'_H & \text{in restriction to } (G_H)^c, \end{cases}$$

so that $P'_H \ll P_H \Leftrightarrow P'_H(G_H) = 1$, and $P'_H \perp P_H \Leftrightarrow P'_H(G_H) = 0$.

3.52 **Theorem.** *Set* $\tau' = \inf(t\colon t \notin \Sigma')$, *where* Σ' *is defined in 3.49* (we also have $\tau' = \tau \wedge \inf(t\colon h^0(1/2)_t = \infty)$).

a) *We have* $G_0 \subset G_H \cap \{\tau' > 0\} \quad (P + P')$*-a.s.*

b) *Assume 3.47, 3.48 and 3.49, and also that local uniqueness holds for the two martingale problems* $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ *and* $\mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$. *Then*

(i) $G_0 = G_H \cap \{\tau' > 0\} \quad (P + P')$*-a.s.*

(ii) *Let* $\alpha \in [0, 1)$ *and let* ψ *meet 1.40 and* $\psi(x) < \psi(0)$ *for all* $x > 0$. *Then if* $h^0(\alpha)$ *and* $\imath^0(\psi)$ *are defined by 3.9, 3.10 and 3.11, the following are versions of* $h(\alpha; P, P')$ *and* $\imath(\psi; P, P')$:

3.53
$$\begin{cases} h(\alpha) = 1_{[\![0,\tau'[\![} \cdot h^0(\alpha) + 1_{\{\tau'>0\}} 1_{[\![\tau',\infty[\![} \\ \imath(\psi) = 1_{[\![0,\tau'[\![} \cdot \imath^0(\psi) + \psi(0) 1_{\{\tau'>0\}} 1_{[\![\tau',\infty[\![}. \end{cases}$$

3.54 **Remark.** This result does not contradict 3.20. As a matter of fact, if the assumptions of both theorems 3.20 and 3.52b are met, we have $\Delta h^0(\alpha)_{\tau'} = 1$ and $\Delta \imath^0(\psi)_{\tau'} = \psi(0)$ on the set $\{0 < \tau' \in \Gamma''\}$. □

3.55 **Remark.** The stopping time τ' plays a crucial rôle here. We shall see in the proof that the representation property 3.21 (for a suitable measure Q) holds on Σ', except perhaps at time τ'. Furthermore, the measures P and P' might be equivalent on $\mathscr{F}_{\tau'-}$, but they are always singular on $\mathscr{F}_{\tau'} \cap \{\tau' < \infty\}$.

In fact, one could prove the following statements, which generalize (i) above. For any $t \in \mathbb{R}_+$ define the sets (see 3.1 for $\mathscr{F}_t^0$):

$$G_t^0 \in \mathscr{F}_t^0 \text{ with } \begin{cases} P \sim P' & \text{on } G_t^0 \cap \mathscr{F}_t^0 \\ P \perp P' & \text{on } (G_t^0)^c \cap \mathscr{F}_t^0 \end{cases}, \quad G_t \in \mathscr{F}_t \text{ with } \begin{cases} P \sim P' & \text{on } G_t \cap \mathscr{F}_t \\ P \perp P' & \text{on } (G_t)^c \cap \mathscr{F}_t \end{cases}.$$

Then, under the assumptions of 3.52b, we have $G_t = G_t^0 \cap \{\tau' > t\}$ $(P + P')$-a.s. □

We begin with a lemma.

3.56 **Lemma.** a) *If* $Q = (P + P')/2$ *and if* Γ'' *is associated to* P, P', Q *by 1.15, we have* $\Gamma'' \subset \Sigma'$ *up to a* Q*-evanescent set.*

b) $G_0 \subset G_H \cap \{\tau' > 0\}$ $\quad Q$-a.s. (i.e., 3.52a holds).

Proof. a) We already know by Lemma 3.19 that $\Gamma'' \subset \Sigma \cap [\![0, \tau]\!]$ Q-a.s. With the notation of 1.15 and 1.18, the Q-martingale $M(1/2)$ is uniformly integrable, and $Y(1/2)_- \geq 1/n$ on $[\![0, S_n]\!]$, hence

$$E_Q(h(\tfrac{1}{2}; P, P')_{S_n}) \leq n E_Q(M(\tfrac{1}{2})_{S_n}) < \infty.$$

Thus $h(\frac{1}{2}; P, P')$ is Q-a.s. finite on Γ''. Since by Theorem 3.39 we have a version of $h(\frac{1}{2}; P, P')$ that majorizes $h^0(\frac{1}{2})$, it follows that $\Gamma'' \subset \{h^0(\frac{1}{2}) < \infty\}$ up to a Q-evanescent set, so $\Gamma'' \subset \Sigma'$.

b) That $(G_H)^c \subset (G_0)^c$ Q-a.s., and thus $G_0 \subset G_H$ Q-a.s. as well, are obvious. Now, if $\omega \in G_0$ we have $S(\omega) > 0$ (notation 1.15), hence $[0, \varepsilon] \subset \Gamma''$ for some $\varepsilon(\omega) > 0$. Then (a) yields $G_0 \subset \{\tau' > 0\}$ Q-a.s. □

Proof of Theorem 3.52. a) Due to the previous lemma, it remains to prove 3.52b only, so we assume 3.48 and 3.49 and local uniqueness for the martingale problems associated to P and P'. Let $\bar{Q} \in \mathscr{s}(\mathscr{H}, X | Q_H; \bar{B}, \bar{C}, \bar{\nu})$ as in 3.48. In 3.4 we choose $\lambda = \bar{\nu}$. We can always assume that 3.5 and 3.7 holds also $\bar{Q}$-a.s. Recalling that $c = c' = \bar{c}$ on $[\![0, \tau]\!]$, 3.7 gives:

$$B - B' = \sum_{j \le d} (c^{\cdot j} \tilde{\beta}^j) \cdot A + h(x)(U - U') * \bar{\nu} \quad (P + P' + \bar{Q})\text{-a.s. on } \Sigma \cap [\![0, \tau]\!].$$

Moreover $B + B' = 2\bar{B}$ and $U + U' = 2$ (because $\nu + \nu' = 2\bar{\nu}$). Hence if $\beta = \tilde{\beta}/2$,

3.57 $$\begin{cases} B = \bar{B} + \left(\sum_{j \le d} c^{\cdot j} \beta^j \right) \cdot A + h(x)(U - 1) * \bar{\nu} \\ C = \bar{C}, \qquad \nu = U \cdot \bar{\nu} \end{cases} \quad \text{on } \Sigma \cap [\![0, \tau]\!],$$

$(P + \bar{Q})$-a.s. We can always modify $\bar{B}, \bar{C}, \bar{\nu}, U$ on a $(P + \bar{Q})$-null set, so that 3.57 holds identically on $\Sigma \cap [\![0, \tau]\!]$.

We shall now apply the analysis of Section III.5 to the pair $(\bar{Q}, P)$ (instead of (P, P')). Recall that $a \le 1$ identically, and set:

3.58 $$\begin{cases} \sigma = \inf(t \colon \bar{a}_t = 1 \text{ and } a_t < 1) \\ H = (\beta \cdot c \cdot \beta) 1_{[\![0,\sigma[\![} \cdot A + (1 - \sqrt{U})^2 1_{[\![0,\sigma[\![} * \bar{\nu} \\ \qquad + \sum_{s \le \cdot} (\sqrt{1 - \bar{a}_s} - \sqrt{1 - a_s})^2 1_{\{s < \sigma\}} \end{cases}$$

Since $\bar{a} = (a + a')/2$ and $0 \le a, a' \le 1$, we immediately get $\sigma \equiv \infty$. By 3.12 and the properties $U + U' = 2$ and $\tilde{\beta} = 2\beta$, we also have

3.59 $$\begin{aligned} h^0\left(\frac{1}{2}\right) &= \frac{1}{2}(\beta \cdot c \cdot \beta) 1_\Sigma \cdot A + \frac{1}{2}(\sqrt{2 - U} - \sqrt{U})^2 * \bar{\nu} \\ &\quad + \frac{1}{2} \sum_{s \le \cdot} (\sqrt{1 - a_s} - \sqrt{1 - a'_s})^2. \end{aligned}$$

Now, it is easy to find a constant θ such that

$$\begin{aligned} 0 \le x \le 2 \ &\Rightarrow\ (1 - \sqrt{x})^2 \le \theta(\sqrt{2 - x} - \sqrt{x})^2 \\ 0 \le x,\ y \le 1 \ &\Rightarrow\ (\sqrt{1 - (x + y)/2} - \sqrt{1 - x})^2 \le \theta(\sqrt{1 - x} - \sqrt{1 - y})^2. \end{aligned}$$

Hence, comparing 3.58 and 3.59, we obtain $H \le 2\theta h^0(\frac{1}{2})$ on Σ. Therefore, if $\Delta = [\![0, \sigma[\![\cap \{H < \infty\}$ is the set defined in III5.9 we have $\Sigma' \subset \Delta$.

b) Due to III.5.10 there is a process N, unique up to a $\bar{Q}$-evanescent set on Δ, such that for all stopping times S with $[\![0, S]\!] \subset \Delta$, N^S is the following $\bar{Q}$-local martingale:

$$N^S = (\beta 1_{[\![0,S]\!]}) \cdot X^c + \left(U - 1 + \frac{\bar{a} - a}{1 - \bar{a}} \right) 1_{[\![0,S]\!]} * (\mu - \bar{\nu})$$

(recall that $0/0 = 0$; X^c is the $\bar{Q}$-continuous martingale part of X).

We define $\tilde{Z}$ by III.5.21 with $\tilde{Z}_0 = z_H := dP_H/dQ_H$, so that for all stopping times S with $[\![0, S]\!] \subset \Delta$:

3.60 $$\tilde{Z}^S = z_H + (\tilde{Z}^S_- \beta 1_{[\![0,S]\!]}) \cdot X^c + \tilde{Z}^S_- \left(U - 1 + \frac{\bar{a} - a}{1 - \bar{a}} \right) 1_{[\![0,S]\!]} * (\mu - \bar{\nu}).$$

Now if we examine the proof of III.5.30 (see part (b) of it) we see that if the σ_n's in III.5.29 are predictable, then so are the S_n's in III.5.30. In our present situation, $\Sigma' = \bigcup_n [\![0, \tau_n]\!] \subset \Delta$, all τ_n being predictable and strict stopping times; then III.5.30, applied to Σ' instead of Δ, yields a sequence (ρ_n) of predictable strict stopping times, with $\Sigma' = \bigcup_n [\![0, \rho_n]\!]$ $\bar{Q}$-a. s., and

3.61 $\tilde{Z}^{\rho_n}$ is a $\bar{Q}$-uniformly integrable nonnegative martingale, $E_{\bar{Q}}(\tilde{Z}_{\rho_n}) = 1$.

c) Similarly, (B', C', ν') satisfies 3.57 with $\beta' = -\beta$ and U'. So the same proof than above yields a sequence (ρ'_n) of predictable strict stopping times with $\Sigma' = \bigcup_n [\![0, \rho'_n]\!]$ $\bar{Q}$-a. s., and 3.61 for $\tilde{Z}'^{\rho'_n}$, where $\tilde{Z}'$ is the solution of 3.60 with $z'_H = dP'_H/dQ_H$, β', U' instead of z_H, β, U.

Then we set $\theta_n = \rho_n \wedge \rho'_n$, which again is a predictable and strict stopping time, and

3.62 $$\Sigma' = \bigcup_n [\![0, \theta_n]\!] \quad \bar{Q}\text{-a. s.}$$

Moreover, $\tilde{Z}^{\theta_n}$ and $\tilde{Z}'^{\theta_n}$ are $\bar{Q}$-uniformly integrable. Then 3.57 and 3.60 and $[\![0, \theta_n]\!] \subset \Delta$ allow to apply III.5.27: the probability measures $P^n(d\omega) = \bar{Q}(d\omega)\tilde{Z}_{\theta_n}(\omega)$ and $P'^n(d\omega) = \bar{Q}(d\omega)\tilde{Z}'_{\theta_n}(\omega)$ are solutions of the stopped problems $\mathscr{s}(\mathscr{H}, X^{\theta_n} | P_H; B^{\theta_n}, C^{\theta_n}, \nu^{\theta_n})$ and $\mathscr{s}(\mathscr{H}, X^{\theta_n} | P'_H; B'^{\theta_n}, C'^{\theta_n}, \nu'^{\theta_n})$. Then the local uniqueness yields

3.63 $$P^n = P \quad \text{and} \quad P'^n = P' \quad \text{on the } \sigma\text{-field } \mathscr{F}^0_{\theta_n}.$$

In particular, $P \ll \bar{Q}$ and $P' \ll \bar{Q}$ in restriction to $\mathscr{F}^0_{\theta_n}$. Since θ_n is $\mathscr{F}^0_{\theta_n}$-measurable, 3.62 gives

3.64 $$\Sigma' = \bigcup_n [\![0, \theta_n]\!] \quad (P + P')\text{-a. s.}$$

d) For each $n \in \mathbb{N}$, we define a new (càd) filtration $\mathbf{G}^n$ by

$$\mathscr{G}^n_t \cap \{\theta_n > 0\} = (\mathscr{F}_t \cap \mathscr{F}_{(\theta_n)-}) \cap \{\theta_n > 0\}$$
$$\mathscr{G}^n_t \cap \{\theta_n = 0\} = \mathscr{F}^0_0 \cap \{\theta_n = 0\},$$

and 3.63 implies $P^n = P \ll \bar{Q}$ and $P'^n = P' \ll \bar{Q}$ in restriction to $\mathscr{G}^n_{\infty-}$. We call $\bar{z}(n)$ and $\bar{z}'(n)$ the density processes of $P^n = P$ and $P'^n = P'$ relative to $\bar{Q}$, and for the filtration $\mathbf{G}^n$. Since θ_n is predictable, and by definition of P^n, $\bar{z}(n)_\infty = \bar{z}(n)_{\theta_n -} = \tilde{Z}_{\theta_n -}$ on $\{\theta_n > 0\}$, while $\bar{z}(n)_\infty = \bar{z}_H$ on $\{\theta_n = 0\}$. Hence, due to the definition of $\mathbf{G}^n$, it is obvious that $\bar{z}(n) = \tilde{Z}$ on $[\![0, \theta_n[\![$, and in particular 3.60 yields

3.65 $$\bar{z}(n) = z_H + \bar{z}(n)_- \beta 1_{[\![0,\theta_n[\![} \cdot X^c + \bar{z}(n)_- \left(U - 1 + \frac{\bar{a} - a}{1 - \bar{a}} \right) 1_{[\![0,\theta_n[\![} * (\mu - \bar{\nu}).$$

(the change of filtration does not affect the stochastic integrals, because the integrands are 0 on $[\![\theta_n, \infty[\![$, neither does it affect the $\bar{Q}$-continuous martingale part X^c on $[\![0, \theta_n[\![$). $\bar{z}'(n)$ also satisfies the same equation 3.65, with z'_H, β', U', a'.

In other words, P and P' satisfy the assumptions of Theorem 3.20, including 3.21, relatively to $\bar{Q}$ and to the filtration $\mathbf{G}^n$. Observing that $\tilde{\beta} = \beta - \beta'$ on $[\![0, \theta_n]\!]$ and that $[\![0, \theta_n]\!] \subset \Sigma$, Theorem 3.20 shows that versions of $h(\alpha; P_{[\theta_n]}, P'_{[\theta_n]})$ and $\imath(\psi; P_{[\theta_n]}, P'_{[\theta_n]})$, where $P_{[\theta_n]}$ and $P'_{[\theta_n]}$ denote the restrictions of P and P' to $\mathscr{G}^n_{\infty-}$, are

3.66 $$h(\alpha; P_{[\theta_n]}, P'_{[\theta_n]}) = 1_{[\![0,\theta_n[\![} \cdot h^0(\alpha), \quad \imath(\psi; P_{[\theta_n]}, P'_{[\theta_n]}) = 1_{[\![0,\theta_n[\![} \cdot \imath^0(\psi).$$

e) Set $Q = (P + P')/2$, call z and z' the density processes of P and P' relative to Q (and to the filtration $\mathbf{F}$), and use the notation Γ'', S_n, S of 1.15. Remember that we have $z_H = dP_H/dQ_H$ and $z'_H = dP'_H/dQ_H$. Therefore if $z(n)$ and $z'(n)$ are the density processes of P and P' with respect to Q, but relative to the filtration $\mathbf{G}^n$ this time, it immediately follows from 1.14 that

$$z(n)_t = \begin{cases} z_H & \text{if } \theta_n = 0 \le t \\ z_t & \text{if } 0 \le t < \theta_n, \\ z_{\theta_n -} & \text{if } 0 < \theta_n \le t \end{cases} \quad z'(n)_t = \begin{cases} z'_H & \text{if } \theta_n = 0 \le t \\ z'_t & \text{if } 0 \le t < \theta_n \\ z'_{\theta_n -} & \text{if } 0 < \theta_n \le t. \end{cases}$$

Hence we deduce from the characterizations 1.20 and 1.42 (or equivalently from the explicit forms 1.34 and 1.48) that $1_{[\![0,\theta_n[\![} \cdot h(\alpha; P, P')$ and $1_{[\![0,\theta_n[\![} \cdot \imath(\psi; P, P')$ are versions of $h(\alpha; P_{[\theta_n]}, P'_{[\theta_n]})$ and $\imath(\psi; P_{[\theta_n]}, P'_{[\theta_n]})$. In other words, 3.66 yields

3.67 $$\begin{cases} h^0(\alpha) & \text{is a version of } h(\alpha; P, P') \quad \text{on } \bigcup_n [\![0, \theta_n[\![\\ \imath^0(\psi) & \text{is a version of } \imath(\psi; P, P') \quad \text{on } \bigcup_n [\![0, \theta_n[\![. \end{cases}$$

So, if we define $h(\alpha)$ and $\imath(\psi)$ by 3.53, we have $h(\alpha) = h(\alpha; P, P')$ and $\imath(\psi) = \imath(\psi; P, P')$ on $\bigcup_n [\![0, \theta_n[\![= [\![0, \tau'[\![$, and in particular on $\Gamma'' \backslash [\![S]\!]$ by 3.56. In order to obtain 3.52b(ii), it remains to prove that $\Delta h(\alpha)_S = \Delta h(\alpha; P, P')_S$ and $\Delta \imath(\psi)_S = \Delta \imath(\psi; P, P')_S$ if $S \in \Gamma''$ and $S \ge \theta_n$ for all n; in this case, $S = \tau'$ and furthermore $S = S_p = \theta_n$ for some n, p large enough (depending upon ω).

In other words, we have to prove that $\Delta h(\alpha)_S = \Delta h(\alpha; P, P')_S$ and $\Delta \imath(\psi)_S = \Delta \imath(\psi; P, P')_S$ on $F = \bigcup_{n,p} \{S = \tau = \theta_n = S_p < \infty\}$. Recall that for $\alpha = 0$, $h(0; P, P') = \imath(\psi; P, P')$ with $\psi(x) = 1_{\{x=0\}}$, so this function meets 1.40 and $\psi(x) < \psi(0) = 1$ for $x > 0$. By 3.53 we have $\Delta h(\alpha)_S = 1$ and $\Delta \imath(\psi)_S = \psi(0)$ on F. On the other hand, $\tau' = \sup_n \tau_n$ is predictable by I.2.9, so I.2.10 and I.2.11 imply that

$$\theta'_n = \begin{cases} \theta_n & \text{if } \theta_n = \tau' \text{ (equivalently, if } \theta_n \ge \tau') \\ +\infty & \text{otherwise} \end{cases}$$

is a predictable time, and $\theta'_n \ge S$. Hence $\Delta h(\alpha; P, P')_S = 1$ and $\Delta \imath(\psi; P, P')_S = \psi(0)$ on $\bigcup_p \{S_p = \tau'_n < \infty\} = \bigcup_p \{S = \tau = \theta_n = S_p < \infty\}$ by 1.30 and 1.59: thus we are finished with the proof of 3.52b(ii).

f) It remains to prove 3.52b(i). In view of 3.56, one needs only to prove that $G_H \cap \{\tau' > 0\} \subset G_0$ $(P + P')$-a. s., and even that $G_H \cap \{\theta_n > 0\} \subset G_0$ $(P + P')$-a. s. for all n.

But 3.65 yields $\bar{z}(n)_0 = z_H$, and similarly $\bar{z}'(n)_0 = z'_H$. Remember also that a version of G_H is $G_H = \{z_H > 0, z'_H > 0\}$. Then $G_H \cap \{\theta_n > 0\} \subset \{\bar{z}(n)_0 > 0, \bar{z}'(n)_0 > 0\}$, and $\mathscr{F}_0 \cap \{\theta_n > 0\} = \mathscr{G}_0^n \cap \{\theta_n > 0\}$, so $\bar{z}(n)_0 = dP_0/dQ_0$ and $\bar{z}'(n)_0 = dP'_0/dQ_0$ on the set $\{\theta_n > 0\}$: it obviously follows that $P_0 \sim P'_0$ on the set $G_H \cap \{\theta_n > 0\}$, and so $G_H \cap \{\theta_n > 0\} \subset G_0$ $(P + P')$-a. s. □

3.68 **Corollary.** *Assume 3.47, 3.48, and local uniqueness for both problems* $\mathscr{s}(\mathscr{H}, X | P_H; B, C, \nu)$ *and* $\mathscr{s}(\mathscr{H}, X | P'_H; B', C', \nu')$. *Assume that the process* $h^0(1/2)$ *does not jump to infinity* (see III.5.8) *and that* $\tau \equiv +\infty$. *Then if* $\alpha \in (0, 1)$, *the process* $h^0(\alpha)$ *is a version of* $h(\alpha; P, P')$.

Proof. Since $K_t < \infty$ for all $t \in \mathbb{R}_+$, 3.50 implies that 3.49 is met. Moreover, $h^0(1/2)_{\tau'-} = \infty$ on $\{\tau' < \infty\}$ by hypothesis, so 3.9 and Lemma 2.19 imply that $h^0(\alpha)_{\tau'-} = \infty$ as well on $\{\tau' < \infty\}$ for all $\alpha \in (0, 1)$, so the process $h(\alpha)$ defined by 3.53 equals $h^0(\alpha)$. □

4. Examples

After the rather tedious Section 3, it is time to show that we can indeed compute Hellinger processes and derive absolute continuity or singularity results for, at least, some of the most usual processes. In fact, we consider only:

1) point processes and multivariate point processes, in § 4a.

2) the case where P (resp. P') is the law of the standard Wiener process (resp. of a Wiener process plus a (random) drift): obviously, much more general results would be available!

3) the case where P and P' are the law of two processes with independent increments: then we compute the Hellinger processes, and also the Hellinger integrals (being able to do so is quite an untypical situation); then we give necessary and sufficient conditions for absolute continuity (this is expected from Theorem 2.6), and also for singularity (this is less expected).

For simplicity, we place ourselves (except for multivariate point processes) in the *canonical setting* III.2.13: Ω is the canonical space of all $\mathbb{R}^d$-valued càdlàg processes, with X the canonical process and **F** the canonical filtration (filtration generated by X), and $\mathscr{F} = \mathscr{F}_{\infty-}$, and $\mathscr{H} = \sigma(X_0) = \mathscr{F}_0^0$: so in particular 3.1 and 3.47 hold.

§ 4a. Point Processes and Multivariate Point Processes

1. For our first example, $d = 1$ and the process X is a. s. a *point process* (see § III.1c) for the two measures P and P'. We call A and A' the compensators of X under P and P' (they are predictable càdlàg increasing processes with $A_0 = A'_0 = 0$).

Let $\bar{A}$ be any increasing predictable càdlàg process such that $dA \ll d\bar{A}$ and $dA' \ll d\bar{A}$ (e.g. $\bar{A} = A + A'$). There are two nonnegative predictable processes g and g' such that

4.1 $$A = g \cdot \bar{A}, \quad A' = g' \cdot \bar{A} \quad (P + P')\text{-a.s.}$$

4.2 **Theorem.** a) *If* $\alpha \in [0, 1)$, *a version of* $h(\alpha; P, P')$ *is*

4.3 $$h(\alpha) = \varphi_\alpha(g, g') \cdot \bar{A} + \sum_{s \leq \cdot} \varphi_\alpha(1 - \Delta A_s, 1 - \Delta A'_s),$$

and in particular

4.4 $$h(0) = g' 1_{\{g=0\}} \cdot \bar{A} + \sum_{s \leq \cdot} (1 - \Delta A'_s) 1_{\{\Delta A_s = 1\}}.$$

b) *If* ψ *satisfies 1.40, a version of* $\imath(\psi; P, P')$ *is*

4.5 $$\imath(\psi) = g' \psi\left(\frac{g}{g'}\right) \cdot \bar{A} + \sum_{s \leq \cdot} (1 - \Delta A'_s) \psi\left(\frac{1 - \Delta A_s}{1 - \Delta A'_s}\right).$$

Proof. We choose a truncation function $h \in \mathscr{C}_t^1$ such that $h(1) = 0$, so that the characteristics of X under P and P' are

$$\begin{cases} B = 0, \quad C = 0, \quad \nu(dt, dx) = dA_t \, \varepsilon_1(dx) \\ B' = 0, \quad C' = 0, \quad \nu'(dt, dx) = dA'_t \, \varepsilon_1(dx). \end{cases}$$

Then in 3.4 we take $\lambda(dt, dx) = d\bar{A}_t \, \varepsilon_1(dx)$ and 3.5 holds with $U(\omega, t, x) = g_t(\omega)$ and $U'(\omega, t, x) = g'_t(\omega)$. Moreover $\tilde{B} = 0$ and $\tilde{\beta} = 0$ in 3.6 and 3.7, and $a = \Delta A$ and $a' = \Delta A'$: therefore 4.3, 4.4, 4.5 are nothing else than 3.9, 3.10, 3.11.

It remains to check that we can apply Theorem 3.28. But, take $Q = (P + P')/2$. Then for Q (as for any measure under which X is a.s. a point process), the martingale representation property holds (see III.4.37), and therefore 3.28 applies. □

4.6 **Theorem.** a) *In order that* $P' \ll P$, *it is necessary and sufficient that there is a version of* A' *with the following three conditions*:

4.7 $$A' = k \cdot A \quad \text{for some nonnegative predictable process } k,$$

4.8 $$\Delta A_t = 1 \Rightarrow \Delta A'_t = k_t \Delta A_t = 1,$$

4.9 $$P'(H_\infty < \infty) = 1, \quad \text{where}$$
$$H = (1 - \sqrt{k})^2 \cdot A + \sum_{s \leq \cdot} (\sqrt{1 - \Delta A_s} - \sqrt{1 - \Delta A'_s})^2.$$

b) *Under 4.7, if* $P'(H_\infty = \infty) = 1$ *we have* $P' \perp P$.

Proof. This theorem will be a consequence of 2.6a,c. Firstly, we observe that $X_0 = 0$ P-a.s. and P'-a.s.; moreover III.1.29 implies that $\mathscr{F}_0$ is Q-trivial, if $Q = (P + P')/2$. Hence $P_0 = P'_0$ and, with the notation 2.2, we have $P'(G_0) = 1$.

Next, we assume that 4.7 holds. Then in 4.1 we can take $\bar{A} = A, g = 1, g' = k$. Replacing in 4.3 yields $h(1/2) = H/2$. Then (b) immediately follows from 2.6c.

Suppose further that 4.8 holds, in addition to 4.7. Then 4.4 gives $h(0) = 0$. Hence with the notation 2.5, $G_\infty = G_0 \cap \{H_\infty < \infty\}$, and the sufficient condition in (a) follows from 2.6a.

Finally, we suppose that $P' \ll P$. Then III.3.17 implies that there is a version of A' that satisfies 4.7 and 4.8, then $h(1/2) = H/2$ and $h(0) = 0$ from above, and 4.9 follows from 2.6a. □

We leave to the reader the "localization" (conditions for $P' \overset{\text{loc}}{\ll} P$) of the previous result.

2. Our second example concerns *multivariate point processes*, as in § III.1c. We start with an E-valued multivariate point process μ (see III.1.23) on some space Ω; we call $\mathbf{F}$ the smallest filtration for which μ is optional, and $\mathscr{F} = \mathscr{F}_{\infty-}$, so that III.1.25 holds with $\mathscr{H} = \{\emptyset, \Omega\}$.

We consider two probability measures P and P' on $(\Omega, \mathscr{F})$, and we call ν and ν' "nice" versions of the compensators of μ under P and P', such that $a \leq 1$ and $a' \leq 1$ identically, where

4.10 $$a_t = \nu(\{t\} \times E), \quad a'_t = \nu'(\{t\} \times E).$$

P (resp. P') is a solution of the martingale problem associated to μ and ν (resp. μ and ν'), in the sense of III.1.9 (here, the initial condition is trivial, because $\mathscr{H} = \{\emptyset, \Omega\}$). So the analysis of Section 3 does not formally apply. However, it is clear that the same sort of analysis does indeed apply, provided we consider only μ, ν, ν'. Moreover, the key points on which 4.2 and 4.6 are hinging, namely III.4.37 and III.1.29 and III.3.17, are valid for multivariate point processes as well as for simple point processes. Thus the following extensions of Theorems 4.2 and 4.6 are valid (we state them without a formal proof):

Firstly, we consider an arbitrary predictable measure λ such that $\nu \ll \lambda$ and $\nu' \ll \lambda$ (e.g. $\lambda = \nu + \nu'$). There are two nonnegative predictable functions U and U' on $\tilde{\Omega} = \Omega \times \mathbb{R}_+ \times E$ such that

4.11 $$\nu = U \cdot \lambda, \quad \nu' = U' \cdot \lambda \quad (P + P')\text{-a.s.}$$

Then we have:

4.12 **Theorem.** a) *If* $\alpha \in [0, 1)$, *a version of* $h(\alpha; P, P')$ *is*

4.13 $$h(\alpha) = \varphi_\alpha(U, U') * \lambda + \sum_{s \leq \cdot} \varphi_\alpha(1 - a_s, 1 - a'_s)$$

and in particular

4.14 $$h(0) = U'1_{\{U=0\}} * \lambda + \sum_{s \leq \cdot} (1 - a'_s)1_{\{a_s=1\}}.$$

b) *If ψ meets 1.40, a version of $\imath(\psi; P, P')$ is*

4.15 $$\imath(\psi) = U'\psi\left(\frac{U}{U'}\right) * \lambda + \sum_{s \leq \cdot} (1 - a'_s)\psi\left(\frac{1 - a_s}{1 - a'_s}\right).$$

4.16 **Theorem.** a) *In order that $P' \ll P$ it is necessary and sufficient that there exists a version of ν' with the following three properties*:

4.17 $\nu' = Y \cdot \nu$ for some nonnegative predictable function Y on $\tilde{\Omega}$

4.18 $$a_t = 1 \Rightarrow a'_t = 1$$

4.19 $$P'(H_\infty < \infty) = 1, \quad \text{where}$$
$$H = (1 - \sqrt{Y})^2 * \lambda + \sum_{s \leq \cdot} (\sqrt{1 - a_s} - \sqrt{1 - a'_s})^2.$$

b) *Under 4.17, if $P'(H_\infty = \infty) = 1$ we have $P' \perp P$.*

§ 4b. Generalized Diffusion Processes

Despite our promising title, we only consider here a very elementary case, which can be readily generalized (the details of the generalization are left to the reader). Again $d = 1$. Under P as well as under P', we want X to be a standard Wiener process, plus a drift that is absolutely continuous with respect to Lebesgue measure, and to avoid trivial complications we assume that X starts from the same point $x \in \mathbb{R}$ under P and P'. In other words, we have

4.20 $$\begin{cases} B_t = \int_0^t \beta_s\,ds, \quad C_t = t, \quad \nu = 0, \quad P(X_0 = x) = 1 \\ B'_t = \int_0^t \beta'_s\,ds, \quad C'_t = t, \quad \nu' = 0, \quad P'(X_0 = x) = 1, \end{cases}$$

and *we suppose that X has the characteristics given by 4.20, under P and P'.* Another way of writing 4.20 is:

4.21 $$\begin{cases} X_t = x + \int_0^t \beta_s\,ds + W_t, \quad W \text{ is a } P\text{-standard Wiener process} \\ X_t = x + \int_0^t \beta'_s\,ds + W'_t, \quad W' \text{ is a } P'\text{-standard Wiener process.} \end{cases}$$

We define three "generalized increasing predictable processes" (see III.5.8):

4.22 $$K_t = \int_0^t (\beta_s)^2\,ds, \quad K'_t = \int_0^t (\beta'_s)^2\,ds, \quad \tilde{K}_t = \int_0^t (\beta_s - \beta'_s)^2\,ds.$$

4.23 Theorem. *We assume that the processes K and K'* (and so $\tilde{K}$ as well) *do not jump to infinity.*

a) $P' \overset{\text{loc}}{\ll} P \Rightarrow P'(\tilde{K}_t < \infty) = 1$ *for all* $t \in \mathbb{R}_+$.

b) $P'(K'_t < \infty) = P'(\tilde{K}_t < \infty) = 1$ *for all* $t \in \mathbb{R}_+ \Rightarrow P' \overset{\text{loc}}{\ll} P$.

c) *If* $P(K_t < \infty) = P'(K'_t < \infty) = 1$ *for all* $t \in \mathbb{R}_+$, *the process* $\frac{\alpha(1-\alpha)}{2}\tilde{K}$ *is a version of the Hellinger process* $h(\alpha; P, P')$. *Moreover for all stopping times T, we have*:

(i) $P'_T \ll P_T \Leftrightarrow P'(\tilde{K}_T < \infty) = 1$

(ii) $P'_T \perp P_T \Leftrightarrow P'(\tilde{K}_T < \infty) = 0$.

Proof. a) Assume $P' \overset{\text{loc}}{\ll} P$. We are exactly in the situation of Section III.5: we have III.5.5 with $\beta' - \beta$ instead of β, and σ in III.5.6 is infinite, and III.5.7 gives $H = \tilde{K}$, so the set Δ of III.5.9 is $\Delta = \{\tilde{K} < \infty\}$ (because $\tilde{K}$ does not jump to infinity). Then Lemma III.5.17b gives $\{Z_- > 0\} = \{\tilde{K} < \infty\}$ up to a P-evanescent set, where Z is the density process of P' with respect to P. But $P' \overset{\text{loc}}{\ll} P$, so $\{Z_- > 0\} = \{\tilde{K} < \infty\}$ up to a P'-evanescent set as well, while $\inf_t Z_t > 0$ P'-a.s. by III.3.5. Therefore, the claim follows.

b) Set $T_n = \inf(t\colon K_t + K'_t \geq n)$, which is a strict stopping time, due to our assumption on K and K'. For each integer n, consider the following P-local martingales:

$$N(n) = -\beta 1_{[\![0,T_n]\!]} \cdot (X - B), \quad Z(n) = \mathscr{E}(N(n)) = \exp(N(n) - \tfrac{1}{2}K^{T_n}).$$

Then $Z(n)^2 = \exp(2N(n) - 2K^{T_n})\exp(K^{T_n}) = \mathscr{E}(2N(n))\exp(K^{T_n})$, and $K^{T_n}_\infty \leq n$, while $\mathscr{E}(2N(n))$ is again a nonnegative P-local martingale, and hence a P-supermartingale. Thus

$$\sup_t E_P[Z(n)_t^2] \leq \sup_t E_P[\mathscr{E}(2N(n))_t]e^n = e^n,$$

and we deduce that $Z(n)$ is a P-uniformly integrable martingale. Similarly, if we define (relatively to P'):

$$N'(n) = -\beta' 1_{[\![0,T_n]\!]} \cdot (X - B'), \quad Z'(n) = \exp[N'(n) - \tfrac{1}{2}K'^{T_n}],$$

then $Z'(n)$ is a P'-uniformly integrable martingale.

Next, we define two new probability measures $Q^n = Z(n)_\infty \cdot P$ and $Q'^n = Z'(n)_\infty \cdot P'$. Lemma III.5.27 implies that both Q^n and Q'^n are solutions of the stopped martingale problem $\mathscr{s}(\mathscr{H}, X^{T_n}|P_x; 0, C^{T_n}, 0)$ (P_x = measure on $\mathscr{H}$ such that $P_x(X_0 = x) = 1$). Local uniqueness for $\mathscr{s}(\mathscr{H}, X|P_x; 0, C, 0)$ yields $Q^n = Q'^n$ on $\mathscr{F}^0_{T_n}$. Moreover $Z(n) > 0$ and $Z'(n) > 0$, so $Q^n \sim P$ and $Q'^n \sim P$ on each $\mathscr{F}_t$ ($t < \infty$). Hence $P \sim P'$ on $\mathscr{F}^0_{n \wedge T_n}$. Since $T_n \uparrow \infty$ P'-a.s. as $n \uparrow \infty$ by hypothesis, we immediately deduce that $P' \overset{\text{loc}}{\ll} P$.

c) Call Q the unique measure under which $X - x$ is a standard Wiener process. Applying (b) to the pair (Q, P) instead of (P, P') (so K, K', $\tilde{K}$ become 0, K, K), we get $P \overset{\text{loc}}{\ll} Q$, and similarly $P' \overset{\text{loc}}{\ll} Q$. Since all Q-martingales have the representation property with respect to X (see III.4.33), Theorem 3.28 applies: a version of $h(\alpha; P, P')$ is $h^0(\alpha)$. But for $\alpha \in (0,1)$ $h^0(\alpha) = \dfrac{\alpha(1-\alpha)}{2}\tilde{K}$ is obvious, and $h^0(0) = 0$. Finally, apply III.4.33b: we have $z_0 = E_Q(z_0) = 1$ and $z'_0 = E_Q(z'_0) = 1$ in the representation 3.21, so $G_T = \{\tilde{K}_T < \infty\}$ (notation 2.5), and (i) and (ii) follow from 2.8. □

§ 4c. Processes with Independent Increments

In this subsection, we assume throughout that $X - X_0$ is a PII under P and under P', so its characteristics (B, C, ν) and (B', C', ν') may be chosen *deterministic.*

1. We introduce the same notation than in § 3a: λ, U, U', Σ, $\tilde{B}$, $\tilde{\beta}$, τ; but here, all those terms are deterministic (hence, in particular, the equalities in 3.5 and 3.7 are true *everywhere*).

4.24 **Theorem.** *Let* $\alpha \in [0,1)$ *and let* ψ *be a function meeting 1.40 and* $\psi(x) < \psi(0)$ *for all* $x > 0$. *Then the processes* $h(\alpha; P, P')$ *and* $i(\psi; P, P')$ *have deterministic versions, that are given by 3.53.*

Proof. The set $\Sigma' = \Sigma \cap \{h^0(\frac{1}{2}) < \infty\} \cap [\![0,\tau]\!]$ of 3.49 is deterministic, so hypothesis 3.49 is fulfilled. By III.2.42 local uniqueness holds for both problems $\mathscr{s}(\mathscr{H}, X \mid P_H; B, C, \nu)$ and $\mathscr{s}(\mathscr{H}, X \mid P'_H; B', C', \nu')$ (recall that $\mathscr{H} = \mathscr{F}_0^0$ and that P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$). Moreover, the characteristics $(\bar{B}, \bar{C}, \bar{\nu})$ of 3.46 are also deterministic, so II.5.2b implies Hypothesis 3.48. Therefore, the result follows from Theorem 3.52. □

Next, we intend to explicitely compute the *Hellinger integrals* $H(\alpha; P_t, P'_t)$ and $H(\alpha; P_{t-}, P'_{t-})$ (see 1.7), for the restrictions P_t, P'_t and P_{t-}, P'_{t-} of P, P' to $\mathscr{F}_t$ and $\mathscr{F}_{t-}$. This is possible here *because* $X - X_0$ *is a PII under P and P'*, although this is based upon a multiplicative decomposition of the process $Y(\alpha)$ of 1.18 which is also valid for more general processes, as we shall see in the next chapter (V.4.16).

Here, $\alpha \in (0,1)$ and $h(\alpha)$ is *any deterministic version of* $h(\alpha; P, P')$, *which is càdlàg and meets* $\Delta h(\alpha) \le 1$ *everywhere* (as is, for instance, the version $h(\alpha)$ constructed in 3.53).

We introduce the following function, defined for $t \in [0, \infty]$ (of course, $h(\alpha)_\infty = \lim_{t\uparrow\infty} h(\alpha)_t$):

4.25 $$\mathscr{E}[-h(\alpha)]_t = \begin{cases} e^{-h(\alpha)_t} \prod_{s\le t} (1 - \Delta h(\alpha)_s) e^{-\Delta h(\alpha)_s} & \text{if } h(\alpha)_t < \infty \\ 0 & \text{if } h(\alpha)_t = \infty. \end{cases}$$

If we compare to I.4.63, we observe that $\mathscr{E}[-h(\alpha)]$ is the Doléans-Dade exponential of $-h(\alpha)$, except that this function may take the value $+\infty$ at a finite time. In fact, if we apply I.4.61 to the stopped function $-h(\alpha)^t$ we see that

4.26 $$\mathscr{E}[-h(\alpha)]_t = 1 - \int_0^t \mathscr{E}[-h(\alpha)]_s \, dh(\alpha)_s \quad \text{if } h(\alpha)_t < \infty.$$

4.27 **Lemma.** a) $\mathscr{E}[-h(\alpha)]$ *is a nonnegative nonincreasing càdlàg function, starting from* 0 *at time* 0.

b) *Let* $T = \inf(t\colon h(\alpha)_t = \infty)$ *and* $T' = \inf(t\colon \Delta h(\alpha)_t = 1)$. *Then*

(i) $\mathscr{E}[-h(\alpha)] > 0$ *and* $\mathscr{E}[-h(\alpha)]_- > 0$ *on* $[0, T \wedge T')$;

(ii) $\mathscr{E}[-h(\alpha)] = 0$ *on* $[T \wedge T', \infty)$;

(iii) $\mathscr{E}[-h(\alpha)]_{T-} = 0$ *if and only if* $T' < T$ *or* $h(\alpha)_{T-} = \infty$.

Proof. For simplicity, set $h = h(\alpha)$. (a) is obvious, as well as the first two claims of (b): use I.4.61 and 4.26 for (i), and 4.25 for (ii).

$\mathscr{E}(-h)_{T-} = 0$ if $T' < T$ follows from (ii), so we assume further that $T' \geq T$. Then if

$$h_t^d = \sum_{s \leq t} \Delta h_s, \quad h^c = h - h^d$$

we have $\mathscr{E}(-h)_t = \exp(-h_t^c) \exp \sum_{s \leq t} \mathrm{Log}(1 - \Delta h_s)$ for $t < T$ (recall that $0 \leq \Delta h_s < 1$ for $s < T$). Then $\mathscr{E}(-h)_{T-} = 0$ if and only if at least one of the following two conditions holds: (1) $h_{T-}^c = \infty$, or (2) $-\sum_{s<T} \mathrm{Log}(1 - \Delta h_s) = \infty$.

Since $-\sum_{s<T} \mathrm{Log}(1 - \Delta h_s) \geq \sum_{s<T} \Delta h_s = h_{T-}^d$, it is clear that if $h_{T-} = \infty$, then at least one of (1) or (2) holds.

Conversely, if $h_{T-} < \infty$ then (1) obviously does not hold. Moreover h has only finitely many jumps of size bigger than 1/2 on $[0, T)$, and the biggest size is $b < 1$. Then $-\sum_{s<T} \mathrm{Log}(1 - \Delta h_s) \leq \dfrac{\mathrm{Log}(1-b)}{b} \sum_{s<T} \Delta h_s$, which equals $\dfrac{\mathrm{Log}(1-b)}{b} h_{T-}^d$, which is finite: hence (2) does not hold either. This finishes the proof. $\square$

We also need to take into account the Hellinger integrals at time 0. Recall that $\mathscr{F}_0^0 = \mathscr{H}$ and that P_H and P'_H are the restrictions of P and P' to $\mathscr{H}$ (so they represent the laws of the variable X_0 under P and P'). The following is an extension of Remark 1.25.

4.28 **Theorem.** *With the above notation, and if* $\tau' = \tau \wedge \inf(t\colon h^0(1/2)_t = \infty)$ (this is again a deterministic time), *the Hellinger integrals of order* $\alpha \in (0, 1)$ *are*:

4.29 $$t \in [0, \infty] \Rightarrow H(\alpha; P_t, P'_t) = \begin{cases} H(\alpha; P_H, P'_H) \mathscr{E}[-h(\alpha)]_t & \text{if } \tau' > 0 \\ 0 & \text{if } \tau' = 0. \end{cases}$$

4.30 $$t \in (0,\infty) \Rightarrow H(\alpha; P_{t-}, P'_{t-}) = \begin{cases} H(\alpha; P_H, P'_H)\mathscr{E}(-h(\alpha))_{t-} & \text{if } \tau' > 0 \\ 0 & \text{if } \tau' = 0. \end{cases}$$

Of course, if $t \geq \tau'$ and $\tau' < \infty$ (resp. $t > \tau'$), 4.29 (resp. 4.30) equals 0 as well. The "discrete time version" of this result has already been proved in 1.73a.

Note that for "standard" PII, for which $X_0 = 0$ P-a.s. and P'-a.s., we have $P_H = P'_H$, and so $H(\alpha; P_H, P'_H) = 1$.

Proof. Let $Q = (P + P')/2$; we use the notation of 1.18, and for simplicity we set $h(\alpha) = h$ and $Y(\alpha) = Y$. Set also $H_t = H(\alpha; P_t, P'_t)$. Then we deduce from 1.7 that $H_t = E_Q(Y_t)$; hence H is càdlàg, and $H_{t-} = E_Q(Y_{t-}) = H(\alpha; P_{t-}, P'_{t-})$.

Now, apply 1.28:

4.31 $$\begin{aligned} H_t = H_0 - E_Q\left(\int_0^t Y_{s-}\, dh_s\right) &= H_0 - \int_0^t E_Q(Y_{s-})\, dh_s \quad \text{(Fubini's Theorem)} \\ &= H_0 - \int_0^t H_{s-}\, dh_s. \end{aligned}$$

Let also $T = \inf(t\colon h_t = \infty)$ and $T' = \inf(t\colon \Delta h_t = 1)$. If we compare 4.31 and 4.26, and because of the uniqueness of the solution of the linear equation 4.31 in H, we obtain $H_t = H_0\mathscr{E}(-h)_t$ for $t < T$. Since the Hellinger integral H_t decreases as t increases, and is nonnegative, we deduce from 4.27b that $H_t = H_0\mathscr{E}(-h)_t$ for all $t \in [0,\infty]$ (recall that $H_\infty := \lim_{t\uparrow\infty} H_t$ is also $H(\alpha; P, P') = H(\alpha; P_\infty, P'_\infty)$).

It remains to compute H_0. Set $z_H = dP_H/dQ_H$ and $z'_H = dP'_H/dQ_H$, and also $G_0 = \{z_0 > 0, z'_0 > 0\}$, $G_H = \{z_H > 0, z'_H > 0\}$. If $\tau' = 0$, $Q(G_0) = 0$ by 3.52a and so $H_0 = E_Q(z_0^\alpha z_0'^{1-\alpha}) = 0$. Suppose now that $\tau' > 0$; then 3.52b(i) gives $G_0 = G_H$ Q-a.s. Moreover, we have the 0.1 law for PII's (see III.4.34b): hence $z_0 = Z$ P-a.s. for some $\mathscr{H}$-measurable random variable Z. Since $P'_0 \sim P_0$ on G_0, $z_0 = Z$ Q-a.s. on the set $G_0 = G_H$. Since $E_Q(z_0|\mathscr{H}) = z_H$, it follows that $z_0 = Z = z_H$ Q-a.s. on G_H. Similarly, $z'_0 = z'_H$ Q-a.s. on G_H, and we get

$$\begin{aligned} H_0 = E_Q(z_0^\alpha z_0'^{1-\alpha}) &= E_Q(z_0^\alpha z_0'^{1-\alpha} 1_{G_H}) = E_Q(z_H^\alpha z_H'^{1-\alpha} 1_{G_H}) \\ &= E_Q(z_H^\alpha z_H'^{1-\alpha}) = H(\alpha; P_H, P'_H). \end{aligned}$$ □

2. *Absolute continuity and singularity.* Now we apply the previous results in order to obtain criteria for absolute continuity and singularity.

4.32 **Theorem.** *In order that $P' \ll P$ it is necessary and sufficient that all the following hold: there are two Borel functions $Y\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ and $\beta\colon \mathbb{R}_+ \to \mathbb{R}^d$, such that*

(i) $P'_H \ll P_H$;
(ii) $\nu' = Y \cdot \nu$
(iii) $a_t = 1 \Rightarrow a'_t = 1$ (recall that $a_t = \nu(\{t\} \times \mathbb{R}^d)$, $a'_t = \nu'(\{t\} \times \mathbb{R}^d)$);

(iv) $|h(x)(Y-1)| * \nu_t < \infty$ for all $t < \infty$;

(v) $B' = B + h(x)(Y-1) * \nu + (\sum_{j \le d} c^{\cdot j}\beta^j) \cdot A$ (recall that c and A satisfy 3.3 and are deterministic here);

(vi) $C' = C$;

(vii) $H_\infty < \infty$, *where* $H = (\beta \cdot c \cdot \beta) \cdot A + (\sqrt{Y} - 1)^2 * \nu + \sum_{s \le \cdot} (\sqrt{1 - a_s} - \sqrt{1 - a'_s})^2$.

Proof. a) Assume first $P' \ll P$. (i) is obvious. III.3.17 implies that one can find Y and β such that (ii)–(vi) hold (all the characteristics being deterministic, Y and β also are deterministic, and the "a.s." disappears) Then in § 3a we may take $\lambda = \nu$, $U = 1$, $U' = Y$, then $\Sigma = \Omega \times \mathbb{R}_+$, $\tau = \infty$ and $\tilde{\beta} = \beta$: thus 3.53 reduces to $h(\frac{1}{2}) = h^0(\frac{1}{2})$, while (as already seen in 3.37) we deduce from 3.12 that $2h^0(1/2) \le H \le 8h^0(1/2)$. Then (vii) follows from 2.6a.

b) Conversely, assume (i)–(vii). Here again, (ii) allows to take $\lambda = \nu$, $U = 1$, $U' = Y$. Then (iv), (v), (vi) yield $\Sigma = \Omega \times \mathbb{R}_+$ and $\tau = \infty$ and $\tilde{\beta} = \beta$, so 3.12 yields $h^0(1/2) \le H/2$, hence $h^0(1/2)_\infty < \infty$ by (vii). In particular, $\Sigma' = \Omega \times \mathbb{R}_+$ and $\tau' = \infty$ (notation of 3.52), so 4.24 implies that $h^0(\alpha) = h(\alpha)$ is a version of $h(\alpha; P, P')$. Furthermore, $h^0(0) = 0$ because $U = 1$ and because of (iii), and with the notation 2.5, G_∞ reduces to $G_0 \cap \{h^0(1/2)_\infty < \infty\} = G_0$. Finally, as seen in the proof of 4.28, (i) implies $P'_0 \ll P_0$, that is $P'(G_0) = 1$. Hence $P'(G_\infty) = 1$, and 2.6a gives $P' \ll P$. □

4.33 **Theorem.** *In order that $P' \perp P$ it is necessary and sufficient that at least one of the following conditions hold*:

(i) $P'_H \perp P_H$;

(ii) $\tau < \infty$ (τ is defined by 3.8);

(iii) $h^0(\frac{1}{2})_\infty = \infty$;

(iv) *there is* $t \in \mathbb{R}_+$ *such that the two measures* $\nu(\{t\} \times \cdot)$ *and* $\nu'(\{t\} \times \cdot)$ *on* $\mathbb{R}^d \backslash \{0\}$ *are mutually singular, and at least one of these has mass* 1 (i.e. $a_t = 1$ or $a'_t = 1$).

Proof. a) Call ν_t, ν'_t, λ_t the measures $\nu(\{t\} \times \cdot)$, $\nu'(\{t\} \times \cdot)$, $\lambda(\{t\} \times \cdot)$, and $U_t(x) = U(t,x)$, $U'_t(x) = U'(t,x)$. Set $\tau' = \tau \wedge \inf(t\colon h^0(1/2)_t = \infty)$, and assume that $t < \tau'$. Then 3.12 yields

$$\Delta h^0(\tfrac{1}{2})_t = \tfrac{1}{2}\lambda_t[(\sqrt{U_t} - \sqrt{U'_t})^2] + \tfrac{1}{2}(\sqrt{1 - a_t} - \sqrt{1 - a'_t})^2,$$

while $\nu_t = U_t \cdot \lambda_t$, $\nu'_t = U'_t \cdot \lambda_t$, $a_t = \lambda_t(U_t)$, $a'_t = \lambda_t(U'_t)$. Then

$$\Delta h^0(\tfrac{1}{2})_t = 1 - \lambda_t(\sqrt{U_t U'_t}) - \sqrt{1 - a_t}\sqrt{1 - a'_t}.$$

Therefore $\Delta h^0(\frac{1}{2})_t = 1$ if and only if $\lambda_t(\sqrt{U_t U'_t}) = 0$ and $\sqrt{1 - a_t}\sqrt{1 - a'_t} = 0$. Clearly, $\lambda_t(\sqrt{U_t U'_t}) = 0 \Leftrightarrow U_t U'_t = 0$ λ_t-a.s. $\Leftrightarrow \nu_t \perp \nu'_t$, then

4.34 $$\text{if } t < \tau', \quad \Delta h^0(\tfrac{1}{2})_t = 1 \Leftrightarrow \text{(iv) holds at time } t.$$

b) We call $h(\alpha)$ the (deterministic) version of $h(\alpha; P, P')$ given by 3.53. Then we clearly have

4.35 $$h(\tfrac{1}{2})_\infty = \infty \Leftrightarrow h^0(\tfrac{1}{2})_{\tau'-} = \infty,$$

4.36 $$\Delta h(\tfrac{1}{2})_t = 1 \text{ for some } t \Leftrightarrow \begin{cases} \text{either } \Delta h^0(\tfrac{1}{2})_t = 1 & \text{for some } t < \tau' \\ \text{or } 0 < \tau' < \infty & \text{and } h^0(\tfrac{1}{2})_{\tau'-} < \infty. \end{cases}$$

On the other hand, $P' \perp P \Leftrightarrow H(\frac{1}{2}; P, P') = 0$ by 1.11b, and using 4.29 and 4.27b we deduce that $P' \perp P$ if and only if at least one of the following holds:

(1) $H(1/2; P_H, P'_H) = 0$
(2) $\tau' = 0$
(3) $\Delta h(1/2)_t = 1$ for some $t \in \mathbb{R}_+$
(4) $h(1/2)_\infty = \infty$.

(1)$\Leftrightarrow$(i) by 1.11b. The following are obvious: (2)$\Rightarrow$(ii) or (iii), and (3)$\Rightarrow$(ii) or (iii) or (iv) (use 4.34, 4.36 and the fact that $h^0(1/2)_{\tau'-} < \infty$ implies either $h^0(1/2)_\infty = \infty$ or $\tau = \tau'$), and (4)$\Rightarrow$(iii) (by 4.35). Finally, if none of (2), (3), (4) is true, we have $\tau' > 0$, and 4.35 gives $h^0(1/2)_{\tau'-} < \infty$, while 4.36 gives that $0 < \tau' < \infty$ and $h^0(1/2)_{\tau'-} < \infty$ together are impossible: therefore $\tau' = \infty$ (hence (ii) is wrong), and 4.36 again and 4.34 imply that (iv) is wrong: therefore, we have: [(2) or (3) or (4)] $\Leftrightarrow$ [(ii) or (iii) or (iv)], and this finishes the proof. □

Contrarily to what was obtained in general (Theorem 2.6 and Corollary 2.7), we do have here a "predictable" (actually, a deterministic) criterion for singularity. We have already seen in § 2c an example of these theorems: 2.37 is a particular case of 4.32 and 4.33 (the reader will check by himself that the conditions in 2.37a and 4.32, and in 2.37b and 4.33, are indeed the same in the discrete case).

4.37 **Remark.** Similar results are obviously available for the restrictions P_t and P'_t, and also for P_{t-} and P'_{t-}: they are the same conditions just "stopped" at time t (resp. at "$t-$", i.e. consider the restrictions of all processes to $[0, t)$). We leave the details to the reader (note however that we do not obtain "deterministic" criteria if we wish to compare P_T and P'_T for a stopping time T that is not deterministic). □

3. *The case of PIIS.* In this paragraph we assume that $X_0 = 0$ P-a. s. and P'-a. s., and that X is a PIIS under P and P'. Then the characteristics take the following form (see II.4.19):

4.38 $$\begin{cases} B_t = bt, & C_t = ct, \quad \nu(dt, dx) = dtK(dx) \\ B'_t = b't, & C'_t = c't, \quad \nu'(dt, dx) = dtK'(dx). \end{cases}$$

4.39 **Theorem.** *Suppose that X is a PIIS under P and under P', with the characteristics 4.38.*

a) *Either $P' = P$ or $P' \perp P$.*

b) *Either for all $t \in \mathbb{R}_+$, P_t and P'_t are not singular, or for all $t \in \mathbb{R}_+$, $P_t \perp P'_t$.*

c) *We have $P' \overset{\text{loc}}{\ll} P$ if and only if all the following conditions hold:*

(i) $K' = k \cdot K$ *for some Borel function* $k: \mathbb{R}^d \to \mathbb{R}_+$
(ii) $\int |h(x)(1 - k(x))| K(dx) < \infty$;
(iii) $b' = b + \int h(x)(k(x) - 1) K(dx) + \sum_{j \le d} c^{\cdot j} \beta^j$, *for some* $\beta \in \mathbb{R}^d$;
(iv) $c' = c$;
(v) $\int (1 - \sqrt{k(x)})^2 K(dx) < \infty$.

Proof. a) We easily deduce from 4.38 that in § 3a everything is "homogeneous in time": i.e. Σ is either $[\![0]\!]$ or $\Omega \times \mathbb{R}_+$; in the latter case $\tilde{B}_t = \tilde{B}t$ for a vector $\tilde{B} \in \mathbb{R}^d$, so in 3.7, $\tilde{\beta}$ and $\tilde{b}$ are constant and $\tilde{B}'_t = \tilde{B}'t$, so τ is either 0 or $+\infty$. Moreover, $h^0(1/2)_t = \gamma t$ for some constant $\gamma \in [0, \infty]$.

If $\tau = \infty$ and $\gamma = 0$, we have $c = c'$ and $\beta = 0$ and $\varphi_\alpha(U, U') * \lambda = 0$, which implies $U = U' = 1$ λ-a.s., hence $K = K'$ and so $b = b'$ as well: thus $P' = P$. If $\tau < \infty$ or if $\gamma > 0$, 4.33 yields $P \perp P'$.

b) We have $P'_H = P_H$ and $a_t = a'_t = 0$ for all t. If we stop all the characteristics and processes at time t, 4.33 implies that $P'_t \perp P_t$ if and only if $\tau \le t$ or $h^0(1/2)_t = \infty$. Then it follows from the "homogeneity" properties in (a) that these conditions are independent from $t \in \mathbb{R}_+$.

c) This follows easily from 4.32 applied to processes and characteristics stopped at an abitrary time t, and from using again the homogeneity in time. □

4.40 **Remark.** Let $\bar{K}$ be a positive measure with $K \ll \bar{K}$ and $K' \ll \bar{K}$, and let $k = dK/d\bar{K}$, $k' = dK'/d\bar{K}$.

We leave as an exercise for the reader to check that the first alternative in 4.39(b) (P'_t and P_t are not singular) holds if and only if we have all the following:
(i) $\int |h(x)(k(x) - k'(x))| \bar{K}(dx) < \infty$
(ii) $b' = b + \int h(x)(k'(x) - k(x)) \bar{K}(dx) + \sum_{j \le d} c^{\cdot j} \tilde{\beta}^j$ for some $\tilde{\beta} \in \mathbb{R}^d$
(iii) $\frac{1}{2} \int (\sqrt{k(x)} - \sqrt{k'(x)})^2 \bar{K}(dx) = \rho^2(K, K') < \infty$. □

4. Gaussian PII.

4.41 **Theorem.** *We suppose that* X *is a Gaussian process and that* $X - X_0$ *is a PII, under* P *and under* P'. *Then either* $P' \sim P$ *or* $P' \perp P$.

Proof. We suppose that P and P' are not mutually singular.

P_H and P'_H are (isomorphic to) Gaussian measures on $\mathbb{R}^d$, and 4.33 implies that they are not mutually singular: then $P_H \sim P'_H$.

Next, set $x(t) = E_P(X_t - X_0)$, which is well-defined, and $V_t = X_t - X_0 - x(t)$ is obviously a Gaussian martingale under P. Then, we easily deduce from Theorem II.4.36 that $\nu(dt, dx) = \sum_{s>0} 1_{\{a_s>0\}} \varepsilon_s(dt) K_s(dx)$, where each K_s is a centered Gaussian measure on $\mathbb{R}^d$. We have a similar representation for ν', with K'_s. It follows that a_t and a'_t are either 0 or 1. Since our assumption implies that 4.33(iv) does not hold, we must have:

$$a_t = 1 \Leftrightarrow a'_t = 1, \text{ in which case } K_t \sim K'_t.$$

This implies $\nu' \sim \nu$. Then we take $\lambda = \nu$, $U = 1$, and we have $\nu' = Y \cdot \nu$ with $Y = U'$.

So far, we have proved that 4.32(i, ii, iii) hold. Moreover 4.33 implies $\tau = \infty$, hence 4.32(iv, v, vi) hold. If H is like in 4.32(vii), we have $H \leq 8h^0(1/2)$, and 4.33 again yields $h^0(1/2)_\infty < \infty$, so $H_\infty < \infty$. Then 4.32 yields $P' \ll P$. Exchanging the rôles of P and P' gives $P \ll P'$, hence the result. □

5. We end this subsection with a result of some interest.

4.42 **Proposition.** *Assume that $X - X_0$ is a PII under P and under P'. There exists a probability measure Q such that $P \ll Q$ and $P' \ll Q$ (resp. $P \overset{\text{loc}}{\ll} Q$ and $P' \overset{\text{loc}}{\ll} Q$) and that $X - X_0$ is a PII under Q, if and only if*

(i) $\tau = \infty$,

(ii) $h^0(\frac{1}{2})_\infty < \infty$ *(resp. $h^0(\frac{1}{2})_t < \infty$ for all $t \in \mathbb{R}_+$).*

If this is the case, one may take for Q the unique measure such that $\bar{Q}_H = (P_H + P'_H)/2$ and that the Q-characteristics of X are $\bar{B} = (B + B')/2$, $\bar{C} = (C + C')/2$ and $\bar{\nu} = (\nu + \nu')/2$.

Proof. a) Let $\bar{Q}$ be the measure described above, and suppose that $\tau = \infty$ and that $h^0(1/2)_\infty < \infty$. Then we apply III.5.34 to the pair $(\bar{Q}, P)$ (instead of (P, P')): the process H of III.5.7 has $H_\infty < \infty$ (this goes as the beginning of the proof of 3.52: we have $H \leq 2\theta h^0(1/2)$ on Σ, and here $\Sigma = \Omega \times \mathbb{R}_+$ because $\tau = \infty$). Then III.5.34 yields $P \ll \bar{Q}$, and similarly $P' \ll \bar{Q}$.

b) Assume conversely that $P \ll Q$, $P' \ll Q$ and $X - X_0$ is a Q-PII with characteristics $(\hat{B}, \hat{C}, \hat{\nu})$. We apply 4.32 to the two pairs (Q, P) and (Q, P'), calling $\hat{\tau}$, $\hat{H}$ and $\hat{\tau}'$, $\hat{H}'$ the analogues of τ, H in 4.32. Then 4.32(ii) implies that one may take $\lambda = \hat{\nu}$, and then 4.32(iv, v, vi) yields that $\tau \geq \hat{\tau} \wedge \hat{\tau}' = \infty$. Moreover 3.12 implies $h^0(1/2) \leq 2(\hat{H} + \hat{H}')$, so 4.32(vii) yields $h^0(1/2)_\infty < \infty$.

Finally, the "local" conditions are obtained by stopping all processes and characteristics at arbitrarily large t. □

Chapter V. Contiguity, Entire Separation, Convergence in Variation

We examine here two apparently disconnected sorts of problems. The relation between them essentially comes from the fact that, in order to solve both of them, we use the same tool, namely the Hellinger processes introduced in the previous chapter.

The notion of contiguity for two sequences (P^n) and (P'^n) of measures has been introduced by LeCam, in relation to asymptotic problems in statistics. Loosely speaking, the sequence (P'^n) is contiguous to the sequence (P^n) if "at the limit" P'^n is absolutely continuous with respect to P^n. The opposite notion that "at the limit" P^n and P'^n are mutually singular is termed "entire separation".

Here we are mainly interested in finding criteria for contiguity when P^n and P'^n are defined on filtered spaces. For the same reason as for absolute continuity, we seek for "predictable criteria" which can indeed be expressed in terms of the Hellinger processes (and some related processes introduced in § IV.1d). In Section 1 we introduce the notions of contiguity and entire separation, and we prove a number of criteria based upon Hellinger integrals and density processes. Section 2 is devoted to our predictable criteria, and Section 3 to various examples (similar to those of the previous chapter).

The last section concerns convergence in variation. As is well known, the Hellinger-Kakutani distance defines the same topology as the variation metric. So the Hellinger processes naturally allow to study the distance in variation $\|P - P'\|$ between two measures defined on a filtered space: as a matter of fact, we obtain various estimates of $\|P - P'\|$, both from below and from above, in terms of the Hellinger processes. These estimates in turn give criteria for convergence in variation. Then we give some examples of convergence in variation for (multivariate) point processes and diffusion processes.

1. Contiguity and Entire Separation

§ 1a. General Facts

1. Firstly, we define the concepts of contiguity and entire separation in the context of measurable spaces, without filtration. For each $n \in \mathbb{N}^*$ we consider a measurable space $(\Omega^n, \mathscr{F}^n)$ endowed with two probability measures P^n and P'^n. The corresponding expectations are E_{P^n} and $E_{P'^n}$.

1.1 **Definitions.** a) We say that *the sequence* (P'^n) *is contiguous to the sequence* (P^n), and we write $(P'^n) \lhd (P^n)$, if for all sequences $A^n \in \mathscr{F}^n$ of sets such that $P^n(A^n) \to 0$ as $n \uparrow \infty$, we have $P'^n(A^n) \to 0$.

b) We say that *the two sequences* (P^n) *and* (P'^n) *are entirely separated*, and we write $(P^n) \triangle (P'^n)$, if there is a sequence $n_k \uparrow \infty$ as $k \uparrow \infty$ and for each $k \in \mathbb{N}^*$ a set $A^{n_k} \in \mathscr{F}^{n_k}$ such that $P^{n_k}(A^{n_k}) \to 1$ and $P'^{n_k}(A^{n_k}) \to 0$ as $k \uparrow \infty$. □

Observe that entire separation is a symmetrical concept: if $(P^n) \triangle (P'^n)$, then $(P'^n) \triangle (P^n)$. These concepts are due to LeCam, and they extend absolute continuity and singularity: indeed if we have a "stationary sequence", which means that $(\Omega^n, \mathscr{F}^n) = (\Omega, \mathscr{F})$ and $P^n = P$ and $P'^n = P'$ for all n, then obviously

$$(P'^n) \lhd (P^n) \Leftrightarrow P' \ll P$$

$$(P'^n) \triangle (P^n) \Leftrightarrow P' \perp P.$$

The concepts of tight sequences and uniformly integrable sequences of random variables are well known. Here, however, we have two sequences of measures, so in order that no confusion arises we use the following terminology:

1.2 **Definitions.** For each $n \in \mathbb{N}^*$ let ξ^n be an $\overline{\mathbb{R}}$-valued random variable on $(\Omega^n, \mathscr{F}^n)$, and let Q^n be a probability measure on $(\Omega^n, \mathscr{F}^n)$.

a) The sequence $(\xi^n | Q^n)$ is $\mathbb{R}$-*tight* if $\lim_{N \uparrow \infty} \limsup_n Q^n(|\xi^n| > N) = 0$.

b) The sequence $(\xi^n | Q^n)$ is *uniformly integrable* if $\lim_{N \uparrow \infty} \sup_n E_{Q^n}(|\xi^n| 1_{\{|\xi^n| > N\}}) = 0$ (or equivalently: if ξ^n is Q^n-integrable for all n and $\lim_{N \uparrow \infty} \limsup_n E_{Q^n}(|\xi^n| 1_{\{|\xi^n| > N\}}) = 0$). □

We also denote by $\mathscr{L}(\xi^n | Q^n)$ the law of ξ^n under Q^n, i.e. the image on $\overline{\mathbb{R}}$ of the measure Q^n under the mapping ξ^n. If we may anticipate on the next chapter, we observe that $(\xi^n | Q^n)$ is $\mathbb{R}$-tight if and only if the sequence $\{\mathscr{L}(\xi^n | Q^n)\}$ is weakly relatively compact and all its limit points are probability measures which charge $\mathbb{R}$ only.

2. In order to derive criteria for contiguity and entire separation, we introduce the following:

1.3 $$Q^n = \frac{P^n + P'^n}{2}, \quad \zeta^n = \frac{dP^n}{dQ^n}, \quad \zeta'^n = \frac{dP'^n}{dQ^n}.$$

We also introduce the "density of P'^n with respect to P^n" as being the $[0, \infty]$-valued random variable defined by (with the convention $2/0 = \infty$):

1.4 $$Z^n = \frac{\zeta'^n}{\zeta^n}.$$

Z^n is the usual density dP'^n/dP^n when $P'^n \ll P^n$, and we have already encountered this "extended" density in § IV.2a.

Firstly, we observe that $P^n(\zeta^n \leq 1/N) = E_{Q^n}(\zeta^n 1_{\{\zeta^n \leq 1/N\}}) \leq \frac{1}{N}$ and that $Z^n \leq 2/\zeta^n$ (recall that $\zeta^n + \zeta'^n = 2$), so the following is always true:

1.5 The sequences $\left(\frac{1}{\zeta^n}\middle| P^n\right)$ and $(Z^n|P^n)$ are $\mathbb{R}$-tight.

Secondly, we give several criteria for contiguity. Recall that the *Hellinger integral* of order $\alpha \in (0,1)$ is defined in IV.1.7 by $H(\alpha; P^n, P'^n) = E_{Q^n}[(\zeta^n)^\alpha(\zeta'^n)^{1-\alpha}]$. The equivalences (i) $\Leftrightarrow$ (ii) $\Leftrightarrow$ (iii) below extend Lemma IV.1.11a.

1.6 Lemma. *The following statements are equivalent:*

(i) $(P'^n) \lhd (P^n)$.

(ii) *The sequence* $\left(\frac{1}{\zeta^n}\middle| P'^n\right)$ *is* $\mathbb{R}$*-tight.*

(ii′) *The sequence* $(Z^n|P'^n)$ *is* $\mathbb{R}$*-tight.*

(iii) $\lim_{\alpha\downarrow\downarrow 0} \liminf_n H(\alpha; P^n, P'^n) = 1$.

Proof. (i) $\Rightarrow$ (ii): If (ii) fails, there exist $\varepsilon > 0$ and a sequence $n_k \uparrow \infty$ such that $P'^{n_k}(\zeta^{n_k} < 1/n_k) \geq \varepsilon$; since $P^{n_k}(\zeta^{n_k} < 1/n_k) \leq 1/n_k \to 0$ as $k \uparrow \infty$ (see before 1.5), this contradicts $(P'^n) \lhd (P^n)$.

(ii) $\Leftrightarrow$ (ii′): This is trivial, once noticed that $Z^n = \frac{2}{\zeta^n} - 1$.

(ii) $\Rightarrow$ (i): Let $A^n \in \mathscr{F}^n$ with $P^n(A^n) \to 0$ as $n \uparrow \infty$. We have $\zeta'^n \leq 2$, hence

$$\begin{aligned} P'^n(A^n) &\leq P'^n(\zeta^n \leq \varepsilon) + E_{Q^n}(\zeta'^n 1_{A^n \cap \{\zeta^n > \varepsilon\}}) \\ &\leq P'^n(\zeta^n \leq \varepsilon) + \frac{2}{\varepsilon} E_{Q^n}(\zeta^n 1_{A^n}) = P'^n(\zeta^n \leq \varepsilon) + \frac{2}{\varepsilon} P^n(A^n), \end{aligned}$$

hence $\limsup_n P'^n(A^n) \leq \limsup_n P'^n(\zeta^n \geq \varepsilon)$ for all $\varepsilon > 0$. But (ii) is equivalent to saying that $\lim_{\varepsilon\downarrow\downarrow 0} \limsup_n P'^n(\zeta^n \geq \varepsilon) = 0$, so we deduce that it implies $P'^n(A^n) \to 0$, and (i) holds.

(ii) $\Rightarrow$ (iii): Let $\varepsilon > 0$. We have

$$H(\alpha; P^n, P'^n) = E_{Q^n}[(\zeta^n)^\alpha(\zeta'^n)^{1-\alpha}] \geq E_{Q^n}\left[\left(\frac{\zeta^n}{\zeta'^n}\right)^\alpha 1_{\{\zeta^n \geq \varepsilon\}} 1_{\{\zeta'^n > 0\}} \zeta'^n\right]$$

1.7 $$= E_{P'^n}\left[\left(\frac{\zeta^n}{\zeta'^n}\right)^\alpha 1_{\{\zeta^n \geq \varepsilon\}}\right] \geq \left(\frac{\varepsilon}{2}\right)^\alpha P'^n(\zeta^n \geq \varepsilon)$$

because $\zeta^n + \zeta'^n = 2$. Thus

$$\begin{aligned} \liminf_{\alpha\downarrow\downarrow 0} \liminf_n H(\alpha; P^n, P'^n) &\geq \liminf_{\alpha\downarrow\downarrow 0} \left(\frac{\varepsilon}{2}\right)^\alpha \liminf_n P'^n(\zeta^n \geq \varepsilon) \\ &= \liminf_n P'^n(\zeta^n \geq \varepsilon) \end{aligned}$$

for all $\varepsilon > 0$. But (ii) yields $\lim_{\varepsilon\downarrow\downarrow 0} \liminf_n P'^n(\zeta^n \geq \varepsilon) = 1$, hence since $H(\alpha; P^n, P'^n) \leq 1$ we obtain (iii).

(iii) $\Rightarrow$ (ii): Let $\delta \in (0, 1)$. We have

$$\begin{aligned} H(\alpha; P^n, P'^n) &= E_{Q^n}[(\zeta^n)^\alpha(\zeta'^n)^{1-\alpha}1_{\{\zeta^n<\varepsilon\}}] + E_{Q^n}[(\zeta^n)^\alpha(\zeta'^n)^{1-\alpha}1_{\{\zeta^n\geq\varepsilon,\zeta'^n\leq\delta\}}] \\ &\quad + E_{Q^n}[(\zeta^n)^\alpha(\zeta'^n)^{1-\alpha}1_{\{\zeta^n\geq\varepsilon,\zeta'^n>\delta\}}] \\ &\leq 2\varepsilon^\alpha + 2\delta^{1-\alpha} + E_{Q^n}\left[\zeta'^n\left(\frac{\zeta^n}{\zeta'^n}\right)^\alpha 1_{\{\zeta^n\geq\varepsilon,\zeta'^n>\delta\}}\right] \end{aligned}$$

1.8 $$\leq 2\varepsilon^\alpha + 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha P'^n(\zeta^n \geq \varepsilon).$$

Thus

$$\liminf_{\varepsilon\downarrow\downarrow 0} \liminf_n P'^n(\zeta^n \geq \varepsilon) \geq \left(\frac{\delta}{2}\right)^\alpha \liminf_n H(\alpha; P^n, P'^n) - \frac{2}{2^\alpha}\delta$$

for all $\alpha \in (0, 1)$, $\delta \in (0, 1)$. Letting $\alpha \downarrow 0$, using (iii), then letting $\delta \downarrow 0$, give $\liminf_{\varepsilon\downarrow\downarrow 0} \liminf_n P'^n(\zeta^n \geq \varepsilon) \geq 1$, so $\liminf_n P'^n(\zeta^n \geq \varepsilon)$, which increases when ε decreases, goes to 1 and we have (ii). $\square$

Next, our criteria for entire separation extend IV.1.11b.

1.9 Lemma. *The following statements are equivalent:*

(i) $(P'^n) \,\Delta\, (P^n)$.

(ii) $\liminf_n P'^n(\zeta^n \geq \varepsilon) = 0$ *for all* $\varepsilon > 0$.

(ii′) $\limsup_n P'^n(Z^n \leq N) = 0$ *for all* $N \in \mathbb{R}_+$.

(iii) $\lim_{\alpha\downarrow\downarrow 0} \liminf_n H(\alpha; P^n, P'^n) = 0$.

(iv) $\liminf_n H(\alpha; P^n, P'^n) = 0$ *for all* $\alpha \in (0, 1)$.

(v) $\liminf_n H(\alpha; P^n, P'^n) = 0$ *for some* $\alpha \in (0, 1)$.

Proof. (i) $\Rightarrow$ (ii): Assume that $(P'^n) \,\Delta\, (P^n)$, and let $n_k \uparrow \infty$ and $A^{n_k} \in \mathscr{F}^{n_k}$ such that $P^{n_k}(A^{n_k}) \to 1$ and $P'^{n_k}(A^{n_k}) \to 0$. Then

$$\begin{aligned} P'^{n_k}(\zeta^{n_k} \geq \varepsilon) &\leq P'^{n_k}(A^{n_k}) + E_{Q^{n_k}}\left(\zeta^{n_k}\frac{\zeta'^{n_k}}{\zeta^{n_k}} 1_{(A^{n_k})^c}1_{\{\zeta^{n_k}\geq\varepsilon\}}\right) \\ &= P'^{n_k}(A^{n_k}) + E_{P^{n_k}}\left(\frac{\zeta'^{n_k}}{\zeta^{n_k}} 1_{(A^{n_k})^c}1_{\{\zeta^{n_k}\geq\varepsilon\}}\right) \leq P'^{n_k}(A^{n_k}) + \frac{2}{\varepsilon}P^{n_k}((A^{n_k})^c) \end{aligned}$$

because $\zeta^n + \zeta'^n = 2$. Hence $P'^{n_k}(\zeta^{n_k} \geq \varepsilon) \to 0$ and we deduce (ii).

(ii) $\Rightarrow$ (i): Under (ii) there exists a sequence $n_k \uparrow \infty$ such that $P'^{n_k}(\zeta^{n_k} \geq 1/k) \leq \frac{1}{k} \to 0$ as $k \uparrow \infty$. Since $P^{n_k}(\zeta^{n_k} \geq 1/k) \geq 1 - \frac{1}{k}$ (see before 1.5), we deduce (i).

(ii) $\Leftrightarrow$ (ii′): This is trivial, once noticed that $Z^n = 2/\zeta^n - 1$.

(ii) $\Rightarrow$ (iv): It follows from (ii) and 1.8 that

$$\liminf_n H(\alpha; P^n, P'^n) \le 2\varepsilon^\alpha + 2\delta^{1-\alpha}$$

for arbitrary ε, δ in $(0, 1)$: then (iv) obtains.

(iv) $\Rightarrow$ (iii) and (iv) $\Rightarrow$ (v) are obvious. Finally 1.7 yields

$$\liminf_n P'^n(\zeta^n \ge \varepsilon) \le \left(\frac{2}{\varepsilon}\right)^\alpha \liminf_n H(\alpha; P^n, P'^n).$$

Then (ii) immediately follows from (v), and also from (iii) because $(2/\varepsilon)^\alpha \to 1$ as $\alpha \downarrow 0$ for $\varepsilon \in (0, 1)$. □

3. We end this subsection with some other criteria for contiguity which will not be used in the remainder of the present chapter (some of them are used in Chapter X, though).

1.10 **Lemma.** *The following statements are equivalent:*

(i) $(P'^n) \lhd (P^n)$.

(ii) *The sequence* $(Z^n|P^n)$ *is uniformly integrable, and* $P'^n(Z^n = \infty) \to 0$.

(ii′) *The sequence* $(Z^n|P^n)$ *is uniformly integrable, and* $E_{P^n}(Z^n) \to 1$.

(iii) *For any subsequence* (n_k) *such that* $(Z^{n_k}|P^{n_k})$ *converges in law, the sequence* $(Z^{n_k}|P^{n_k})$ *is uniformly integrable and* $P'^{n_k}(Z^{n_k} = \infty) \to 0$.

(iii′) *For any subsequence* (n_k) *such that* $(Z^{n_k}|P^{n_k})$ *converges in law, the limiting random variable has expectation* 1.

Proof. (ii) $\Rightarrow$ (iii) and (ii′) $\Rightarrow$ (iii′) are evident.

(i) $\Rightarrow$ (ii): 1.3 and 1.4 yield

$$\begin{aligned} E_{P^n}(Z^n 1_{\{Z^n \ge N\}}) &= E_{Q^n}(Z^n \zeta^n 1_{\{\zeta^n > 0\}} 1_{\{Z^n \ge N\}}) \\ &= E_{Q^n}(\zeta'^n 1_{\{N \le Z^n < \infty\}}) = P'^n(N \le Z^n < \infty). \end{aligned} \tag{1.11}$$

If (i) holds, 1.6 implies that the sequence $(Z^n|P'^n)$ is $\mathbb{R}$-tight, so $P'^n(Z^n = \infty) \to 0$ and $\lim_{N\uparrow\infty} \limsup_n P'^n(N \le Z^n < \infty) = 0$; since by 1.11 $E_{P^n}(Z^n) \le 1$ for all n, we deduce from 1.11 that the sequence $(Z^n|P^n)$ is uniformly integrable, and so (ii) holds.

(iii) $\Rightarrow$ (i): Suppose that (i) fails. Then there exists $\varepsilon > 0$ and an infinite subsequence (n') and sets $A^{n'} \in \mathscr{F}^{n'}$, such that $P'^{n'}(A^{n'}) \ge \varepsilon$ for all n' and $P^{n'}(A^{n'}) \to 0$ as $n' \uparrow \infty$. By 1.5 the sequence $(Z^{n'}|P^{n'})$ is $\mathbb{R}$-tight, so it has a further subsequence denoted by (n_k), such that $\mathscr{L}(Z^{n_k}|P^{n_k})$ weakly converges to a probability on $\mathbb{R}$, and we deduce from 1.11 that

$$P'^{n_k}(Z^{n_k} \ge N) = E_{P^{n_k}}(Z^{n_k} 1_{\{Z^{n_k} \ge N\}}) + P'^{n_k}(Z^{n_k} = \infty).$$

Thus (iii) implies that $\lim_{N\uparrow\infty} \limsup_k P'^{n_k}(Z^{n_k} \ge N) = 0$. Hence $(Z^{n_k}|P'^{n_k})$ is $\mathbb{R}$-tight, and 1.6 yields $(P'^{n_k}) \lhd (P^{n_k})$. Since $P^{n_k}(A^{n_k}) \to 0$ we must have $P'^{n_k}(A^{n_k}) \to 0$, which contradicts the property $P'^{n_k}(A^{n_k}) \ge \varepsilon$.

(ii) $\Rightarrow$ (ii'): 1.11 with $N = 0$ yields $E_{P^n}(Z^n) = P'^n(Z^n < \infty)$, hence the claim.

(iii') $\Rightarrow$ (iii): It is enough to prove that if $(Z^n|P^n)$ converges in law to a variable Z defined on some probability space $(\Omega, \mathscr{F}, P)$ and if $E(Z) = 1$, then (ii) holds. For each $N > 0$ such that $P(Z = N) = 0$ we have $E_{P^n}(Z^n 1_{\{Z^n<N\}}) \to E(Z 1_{\{Z<N\}})$ (again, we anticipate on the next chapter; see e.g. [12]). Let $\varepsilon > 0$; there exists $N > 0$ with $P(Z = N) = 0$ and $E(Z1_{\{Z<N\}}) \geq 1 - \varepsilon$, because $E(Z) = 1$. Then there exists $n_0 \in \mathbb{N}^*$ such that $E_{P^n}(Z^n 1_{\{Z^n<N\}}) \geq 1 - 2\varepsilon$ for all $n \geq n_0$; since $\varepsilon > 0$ is arbitrary and $E_{P^n}(Z^n) \leq 1$ we deduce that $E_{P^n}(Z^n) \to 1$ and that $\lim_{N\uparrow\infty} \sup_n E_{P^n}(Z^n 1_{\{Z^n \geq N\}}) = 0$: so we have (ii). □

1.12 **Corollary.** *Assume that $(Z^n|P^n)$ converges in law to a random variable Z. Then $(P'^n) \lhd (P^n)$ if and only if $E(Z) = 1$.*

Proof. The claim follows from the equivalence (i) $\Leftrightarrow$ (iii') in 1.10. □

According to Hájek and Šidak [82], this corollary is known as "LeCam's first lemma", and "LeCam's third lemma" is the following:

1.13 **Lemma.** a) *There is equivalence between:*

(i) *$\mathscr{L}(Z^n|P^n)$ weakly converges to a probability measure η on $\mathbb{R}_+$, and $(P'^n) \lhd (P^n)$.*

(ii) *$\mathscr{L}(Z^n|P'^n)$ weakly converges to a probability measure η' on $\mathbb{R}_+$.*

b) *Assume* (i) *or* (ii) *above. Then $\eta'(dx) = x\eta(dx)$. Moreover if for every n there is a random variable X^n on $(\Omega^n, \mathscr{F}^n)$ with values in a metric space E and if $\mathscr{L}((Z^n, X^n)|P^n)$ weakly converges to a probability measure $\overline{\eta}$ on $\mathbb{R}_+ \times E$, then $\mathscr{L}((Z^n, X^n)|P'^n)$ weakly converges to $\overline{\eta}'$, where $\overline{\eta}'(dx, dy) = x\overline{\eta}(dx, dy)$.*

Proof. Let f be a continuous bounded function on $[0, \infty] \times E$. Then

$$\begin{aligned} E_{P'^n}[f(Z^n, X^n)] &= E_{P'^n}[f(\infty, X^n)1_{\{Z^n=\infty\}}] + E_{Q^n}[\zeta'^n f(Z^n, X^n)1_{\{\zeta^n>0\}}] \\ &= E_{P'^n}[f(\infty, X^n)1_{\{Z^n=\infty\}}] + E_{Q^n}[\zeta^n Z^n f(Z^n, X^n)] \\ &= E_{P'^n}[f(\infty, X^n)1_{\{Z^n=\infty\}}] + E_{P^n}[Z^n f(Z^n, X^n)]. \end{aligned} \tag{1.14}$$

Assume that $(P'^n) \lhd (P^n)$ and that $\mathscr{L}((Z^n, X^n)|P^n) \to \overline{\eta}$ weakly, where $\overline{\eta}$ is a probability measure on $\mathbb{R}_+ \times E$. 1.10 gives $P'^n(Z^n = \infty) \to 0$ and the uniform integrability of the sequence $(Z^n f(Z^n, X^n)|P^n)$. Then, passing to the limit in 1.14 yields

$$E_{P'^n}[f(Z^n, X^n)] \to \int \overline{\eta}(dx, dy)xf(x, y).$$

This proves the implication (i) $\Rightarrow$ (ii) in (a) (take $X^n \equiv 1$ and $E = \{1\}$), and it also proves (b).

Assume now (ii). Then the sequence $(Z^n|P'^n)$ is $\mathbb{R}$-tight, and $(P'^n) \lhd (P^n)$ follows from 1.6. Moreover 1.10 implies that from any subsequence one may extract a further subsequence $\mathscr{L}(Z^{n_k}|P^{n_k})$ that converges weakly to a probability measure on $\mathbb{R}_+$, say η. Applying the first part of the proof, we get that $\eta'(dx) = x\eta(dx)$, which completely determines the measure η: in other words the tight sequence $\{\mathscr{L}(Z^n|P^n)\}$ has only one limit point η, and so it converges. □

1.15 **Corollary.** *Assume that $\mathscr{L}(Z^n|P^n)$ weakly converges to a probability measure η on $\mathbb{R}_+$.*

a) *If $\eta(\{0\}) = 0$, then $(P^n) \lhd (P'^n)$.*

b) *If $\eta(\{0\}) = 0$ and if $\mathscr{L}(Z^n|P'^n)$ also weakly converges to a probability measure on $\mathbb{R}_+$, then $(P^n) \lhd (P'^n) \lhd (P^n)$.*

Proof. a) The assumption implies that $\mathscr{L}(1/Z^n|P^n)$ also weakly converges to a probability measure on $\mathbb{R}_+$. Since $1/Z^n$ is the process associated with P^n and P'^n, in the reverse order, by 1.4, we deduce $(P^n) \lhd (P'^n)$ from the implication (ii$'$) $\Rightarrow$ (i) of 1.6.

b) We already know that $(P^n) \lhd (P'^n)$. That $(P'^n) \lhd (P^n)$ follows from 1.13. □

§ 1b. Contiguity and Filtrations

Now we suppose that for each $n \in \mathbb{N}^*$ the space $(\Omega^n, \mathscr{F}^n)$ is equipped with two probability measures P^n and P'^n, and with a *filtration* $\mathbf{F}^n$. In order to avoid uninteresting complications, *we assume further that* $\mathscr{F}^n = \mathscr{F}^n_{\infty-} = \bigvee_t \mathscr{F}^n_t$. Recall that if T^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ we denote by $P^n_{T^n}$ and $P'^n_{T^n}$ the restrictions of P^n and P'^n to the σ-field $\mathscr{F}^n_{T^n}$.

Let $Q^n = \dfrac{P^n + P'^n}{2}$. We denote by z^n and z'^n the *density processes* of P^n and P'^n with respect to Q^n (see § III.3.1). We also consider the $[0, \infty]$-valued process

1.16 $$Z^n = \frac{z'^n}{z^n} \quad \text{(with } 2/0 = +\infty\text{)}.$$

If T^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, all the results of § 1a are of course valid for $P^n_{T^n}, P'^n_{T^n}$ (with $\zeta^n = z^n_{T^n}$ and $\zeta'^n = z'^n_{T^n}$). We also have the following specific results:

1.17 **Lemma.** *The sequences* $\left(\sup_s \dfrac{1}{z^n_s} \middle| P^n\right)$ *and* $(\sup_s Z^n_s | P^n)$ *are $\mathbb{R}$-tight.*

Proof. Since $Z^n \leq 2/z^n$, only the first claim needs proving. Let $N > 0$ and $T^n = \inf(t: z^n_t < 1/N)$. Then T^n is a stopping time, and $z^n_{T^n}$ is the density $dP'^n_{T^n}/dP^n_{T^n}$ (see IV.1.44), hence

1.18 $$P^n\left(\sup_s \frac{1}{z^n_s} > N\right) = P^n(T^n < \infty) = E_{Q^n}(z^n_{T^n} 1_{\{T^n < \infty\}}) \leq \frac{1}{N}.$$

Therefore $\lim_{N\uparrow\infty} \sup_n P^n(\sup_s 1/z^n_s > N) = 0$. □

1.19 **Lemma.** *For each* $n \in \mathbb{N}^*$ *let* T^n *be a stopping time on* $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. *There is equivalence between*:

(i) $(P'^n_{T^n}) \lhd (P^n_{T^n})$.

(ii) *The sequence* $\left(\sup_{s \le T^n} \frac{1}{z^n_s} \,\middle|\, P'^n\right)$ *is* $\mathbb{R}$-*tight.*

(iii) *The sequence* $(\sup_{s \le T^n} Z^n_s | P'^n)$ *is* $\mathbb{R}$-*tight.*

Proof. The implications (ii) ⇒ (i) and (iii) ⇒ (i) are trivial consequences of 1.6. The equivalence (ii) ⇔ (iii) follows from the equality $Z^n = 2/z^n - 1$.

If (ii) fails, there exist $\varepsilon > 0$ and a sequence $n_k \uparrow \infty$ such that $P'^{n_k}(\sup_{s \le T^{n_k}} 1/z^{n_k}_s \ge 1/n_k) \ge \varepsilon$. But 1.18 implies that

$$P^{n_k}\left(\sup_{s \le T^{n_k}} 1/z^{n_k}_s \ge \frac{1}{n_k}\right) \le \frac{1}{n_k} \to 0,$$

and thus (i) cannot hold: hence (i) ⇒ (ii). □

1.20 **Lemma.** *For each* $n \in \mathbb{N}^*$ *let* T^n *be a stopping time on* $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. *There is equivalence between*:

(i) $(P'^n_{T^n}) \,\triangle\, (P^n_{T^n})$.

(ii) *For all* $\varepsilon > 0$, $\liminf_n P'^n(\inf_{s \le T^n} z^n_s \ge \varepsilon) = 0$.

(ii′) *For all* $N \in \mathbb{R}_+$, $\limsup_n P'^n(\sup_{s \le T^n} Z^n_s \le N) = 1$.

(iii) $\lim_{\varepsilon \downarrow\downarrow 0} \liminf_n P'^n(\inf_{s \le T^n} z^n_s \ge \varepsilon) = 0$.

Proof. (ii) ⇔ (ii′) follows from $Z^n = 2/z^n - 1$, and (i) ⇒ (ii) immediately follows from 1.9, while (ii) ⇒ (iii) is trivial.

Finally, assume (iii). There are two sequences $n_k \uparrow \infty$ and $\varepsilon_k \downarrow 0$ such that $P'^{n_k}\left(\inf_{s \le T^{n_k}} z^{n_k}_s \ge \varepsilon_k\right) \le \frac{1}{k} \to 0$, while 1.18 yields

$$P^{n_k}\left(\inf_{s \le T^{n_k}} z^{n_k}_s \ge \varepsilon_k\right) \ge 1 - \varepsilon_k,$$

so (i) follows. □

2. Predictable Criteria for Contiguity and Entire Separation

§ 2a. Statements of the Results

We start with a sequence of filtered spaces, say $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ for every $n \in \mathbb{N}^*$, each one being endowed with two probability measures P^n and P'^n. For simplicity, we assume that $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$.

In a way similar to what has been done in Section IV.2, we wish to obtain criteria for $(P'^n_{T^n}) \lhd (P^n_{T^n})$ (contiguity) and $(P'^n_{T^n}) \mathbin{\Delta} (P^n_{T^n})$ (entire separation) in terms of the behaviour of suitable increasing predictable processes, where for each n, T^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$.

One of these predictable processes is (a version of) the Hellinger process $h(\alpha; P^n, P'^n)$ of order $\alpha \in (0,1)$, defined in IV.1.24: in all this section, *we denote by $h^n(\alpha)$ an arbitrary version of $h(\alpha; P^n, P'^n)$.*

We also need to consider some of the processes $\imath(\psi; P^n, P'^n)$ introduced in IV.1.46. More precisely, for every $\beta \in [0,1)$ we consider the following function, which obviously meets IV.1.40:

2.1 $$\psi_\beta(x) = 1_{[0,\beta]}(x),$$

and *we denote by $\imath^n(\beta)$ an arbitrary version of $\imath(\psi_\beta; P^n, P'^n)$.*

It is perhaps worthwhile recalling an explicit form for these processes: let $Q^n = (P^n + P'^n)/2$, with z^n and z'^n the density processes of P^n and P'^n with respect to Q^n (so $z^n + z'^n = 2$), and ν^{z^n} be the third Q^n-characteristic of z^n, and $z^{n,c}$ be the continuous martingale part of z^n (relative to Q^n). Then if φ_α is given by IV.1.32,

2.2 $$\begin{cases} h'^n(\alpha) = \dfrac{\alpha(1-\alpha)}{2}\left(\dfrac{1}{(z^n_-)^2} + \dfrac{1}{(z'^n_-)^2}\right) \cdot \langle z^{n,c}, z^{n,c}\rangle + \varphi_\alpha\left(1 + \dfrac{x}{z^n_-}, 1 - \dfrac{x}{z'^n_-}\right) * \nu^{z^n} \\ \imath'^n(\beta) = \left(1 - \dfrac{x}{z'^n_-}\right) 1_{\{1 + x/z^n_- \le \beta(1 - x/z'^n_-)\}} * \nu^{z^n} \end{cases}$$

are the "strict" versions of $h^n(\alpha)$ and $\imath^n(\beta)$, and all other versions meet $h^n(\alpha) \ge h'^n(\alpha)$ and $\imath^n(\beta) \ge \imath'^n(\beta)$. Moreover, $h'^n(\alpha)$ and $\imath'^n(\beta)$ do not depend on the "dominating" measure Q^n.

2.3 **Theorem.** *For each $n \in \mathbb{N}^*$ let T^n be a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, and assume that $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$. There is equivalence between properties* (i), (ii) *and* (iii) *below*:

(i) $(P'^n_{T^n}) \lhd (P^n_{T^n})$.

(ii) (1) $(P'^n_0) \lhd (P^n_0)$;

(2) $\lim_{N \uparrow \infty} \limsup_n P'^n(h^n(\frac{1}{2})_{T^n} > N) = 0$ (equivalently, the sequence $(h^n(1/2)_{T^n} | P'^n)$ is $\mathbb{R}$-tight);

(3) $\lim_{\beta \downarrow\downarrow 0} \limsup_n P'^n(\imath^n(\beta)_{T^n} > \eta) = 0$ *for all* $\eta > 0$.

(iii) (1) $(P'^n_0) \lhd (P^n_0)$;

(2) $\lim_{\alpha \downarrow\downarrow 0} \limsup_n P'^n(h^n(\alpha)_{T^n} > \eta) = 0$ *for all* $\eta > 0$.

If $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n) = (\Omega, \mathscr{F}, \mathbf{F})$ and $P^n = P$ and $P'^n = P'$ for all n (stationary case), these three statements (i, ii, iii) exactly are statements (i, ii, iii) of IV.2.1: this is obvious for (i) and (iii), and for (ii) it comes from the fact that $\imath'^n(\beta)$ decreases to $\imath'^n(0) = h'^n(0)$ as $\beta \downarrow 0$.

Similarly, the following theorem reduces to IV.2.6b,c in the stationary case.

2.4 **Theorem.** *For each* $n \in \mathbb{N}^*$ *let* T^n *be a stopping time on* $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, *and assume that* $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$.

a) *If* $(P'^n_{T^n}) \Delta (P^n_{T^n})$ *we have the following properties*:

(i) $\lim_{\eta\downarrow\downarrow 0} \limsup_n P'^n(z^n_0 < \varepsilon \text{ or } h^n(\frac{1}{2})_{T^n} > N \text{ or } \imath^n(\beta)_{T^n} > \eta) = 1$ *for all* $N > 0$, $\varepsilon > 0$, $\beta \in (0,1)$;

(ii) $\lim_{\eta\downarrow\downarrow 0} \liminf_{\alpha\downarrow\downarrow 0} \limsup_n P'^n(z^n_0 < \varepsilon \text{ or } h^n(\alpha)_{T^n} > \eta) = 1$ *for all* $\varepsilon > 0$;

(iii) $\lim_{\eta\downarrow\downarrow 0} \limsup_n P'^n(z^n_0 < \varepsilon \text{ or } h^n(\alpha)_{T^n} > \eta) = 1$ *for all* $\varepsilon > 0$, $\alpha \in (0,1)$.

b) *If* $(P'^n_0) \Delta (P^n_0)$ *or if* $\limsup_n P'^n(h^n(\frac{1}{2})_{T^n} > N) = 1$ *for all* $N > 0$, *then* $(P'^n_{T^n}) \Delta (P^n_{T^n})$.

Unlike in Section IV.2 (see IV.2.8), the additional assumption that $P'^n \overset{\text{loc}}{\ll} P^n$ for all n does not entail any simplification in the previous statements. We do not know about a possible extension of Theorem IV.2.13 either. But we do have a "non-predictable" criterion for contiguity that is similar to Theorem IV.2.15. For this, we consider the density processes z^n and z'^n as before 2.2, and we set

$$Z^n_t = \frac{z'^n_t}{z^n_t}, \quad \alpha^n_t = \begin{cases} Z^n_t/Z^n_{t-} & \text{if } 0 < Z^n_{t-} < \infty \\ 0 & \text{if } Z^n_{t-} = 0 \\ +\infty & \text{if } Z^n_{t-} = \infty. \end{cases}$$

2.5 **Theorem.** *For each* $n \in \mathbb{N}^*$ *let* T^n *be a stopping time on* $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, *and assume that* $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$. *There is equivalence between* (i) *and* (ii) *below*:

(i) $(P'^n_{T^n}) \lhd (P^n_{T^n})$;

(ii) (1) $(P'^n_0) \lhd (P^n_0)$,

(2) $\lim_{N\uparrow\infty} \limsup_n P'^n(h^n(1/2)_{T^n} > N) = 0$,

(3) $\lim_{N\uparrow\infty} \limsup_n P'^n(\sup_{t\le T^n} \alpha^n_t > N) = 0$.

(2) and (3) above are exactly $\mathbb{R}$-tightness for the two sequences $(h^n(1/2)_{T^n} | P'^n)$ and $(\sup_{t\le T^n} \alpha^n_t | P'^n)$.

2.6 **Remark.** In the previous statements, one could replace everywhere $h^n(1/2)$ by $h^n(\beta)$ for a *fixed* $\beta \in (0,1)$. □

2.7 **Remark.** Instead of $\imath^n(\beta)$ we could use the following processes

$$\bar{\imath}'^n(\beta) = \varphi_{1/2}\left(1 + \frac{x}{z^n_-}, 1 - \frac{x}{z'^n_-}\right) 1_{\{1+x/z^n_- \le \beta(1-x/z'^n_-)\}} * \nu^{z^n},$$

which would be more natural in a sense. In fact, we will see later (2.8) that $\frac{1}{2}(1-\sqrt{\beta})^2 v \le \varphi_{1/2}(u,v) \le \frac{v}{2}$ if $u \le \beta v$ and $\beta \le 1$, so

$$\tfrac{1}{2}(1-\sqrt{\beta})^2 \imath'^n(\beta) \le \bar{\imath}'^n(\beta) \le \tfrac{1}{2}\imath'^n(\beta). \qquad \square$$

§ 2b. The Proofs

Before proceeding to the proofs themselves, we wish to emphazise that, if written for the stationary case $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n, P'^n) = (\Omega, \mathscr{F}, \mathbf{F}, P, P')$, they would be (slightly) simpler than the proofs in § IV.2b (but they do not give Lemma IV.2.12, nor Theorem IV.2.13).

We set $Q^n = (P^n + P'^n)/2$ and we use the notation explained before 2.2. Also, for each $n \in \mathbb{N}^*$, T^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. We set $Y^n(\alpha) = (z^n)^\alpha (z'^n)^{1-\alpha}$ and $H^n(\alpha)_{T^n} = H(\alpha; P^n_{T^n}, P'^n_{T^n})$, and we recall that $H^n(\alpha)_{T^n} = E_{Q^n}[Y^n(\alpha)_{T^n}]$. We begin with some auxiliary lemmas.

2.8 **Lemma.** a) $\varphi_{1/2}(u, v) \le \dfrac{v}{2}$ *for all* $0 \le u \le v$.

b) $\varphi_\alpha(u, v) \ge (\alpha\beta + 1 - \alpha - \beta^\alpha)v$ *for all* $\alpha \in (0,1)$, $\beta \in (0,1)$, $0 \le u \le \beta v$.

Proof. (a) immediately follows from $\varphi_{1/2}(u,v) = \frac{1}{2}(\sqrt{u} - \sqrt{v})^2$. For (b), we observe that the function $u \rightsquigarrow \varphi_\alpha(u,v)$ decreases when u increases from 0 to v, and $\varphi_\alpha(\beta v, v) = (\alpha\beta + 1 - \alpha - \beta^\alpha)v$. □

2.9 **Lemma.** *If R^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ such that $h^n(\alpha)_{R^n} \le \eta$ identically, then*

$$0 \le H^n(\alpha)_0 - H^n(\alpha)_{R^n} \le 2\eta. \tag{2.10}$$

Proof. Since $0 \le Y^n(\alpha) \le 2$, the claim is immediate from IV.1.28. □

2.11 **Lemma.** *If R^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ such that $R^n \le T^n$, then for all $\rho \ge 0$, $\delta > 0$ we have*:

$$\begin{aligned} 0 &\le H^n(\alpha)_{R^n} - H^n(\alpha)_{T^n} \\ &\le E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n \le \rho\}}) + 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha P'^n(R^n < T^n, z_0^n > \rho) \end{aligned} \tag{2.12}$$

Proof. The first inequality follows from IV.1.28. The proof of the second inequality is somewhat similar to 1.8. Firstly, since $Y^n(\alpha) \ge 0$,

$$H^n(\alpha)_{R^n} - H^n(\alpha)_{T^n} = E_{Q^n}(Y^n(\alpha)_{R^n} - Y^n(\alpha)_{T^n}) \le E_{Q^n}(Y^n(\alpha)_{R^n} 1_{\{R^n < T^n\}}).$$

Let $A = \{R^n < T^n, z_0^n > \rho\}$. Then

$$\begin{aligned} H^n(\alpha)_{R^n} - H^n(\alpha)_{T^n} &\le E_{Q^n}(Y^n(\alpha)_{R^n} 1_{\{z_0^n \le \rho\}}) + E_{Q^n}(Y^n(\alpha)_{R^n} 1_A) \\ &\le E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n \le \rho\}}) + E_{Q^n}(Y^n(\alpha)_{R^n} 1_A) \end{aligned} \tag{2.13}$$

because $Y^n(\alpha)$ is a Q^n-supermartingale (IV.1.17). Moreover

$$E_{Q^n}(Y^n(\alpha)_{R^n}1_A) = E_{Q^n}[(z^n_{R^n})^\alpha(z'^n_{R^n})^{1-\alpha}1_A 1_{\{z'^n_{R^n}\le\delta\}}] + E_{Q^n}[(z^n_{R^n}/z'^n_{R^n})^\alpha z'^n_{R^n}1_A 1_{\{z'^n_{R^n}>\delta\}}]$$
$$\le 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha E_{Q^n}(z'^n_{R^n}1_A) = 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha P'^n(A)$$

because $z^n \le 2$ and $A \in \mathscr{F}^n_{R^n}$. This, plus 2.13, give the result. □

Proof of 2.4b. If $(P'^n_0) \Delta (P^n_0)$ the claim is trivial.

Assume now that $\limsup_n P'^n(h^n(1/2)_{T^n} > N) = 1$ for all $N > 0$. Set $S^n_k = \inf\left(t\colon z^n_t < \frac{1}{k} \text{ or } z'^n_t < \frac{1}{k}\right)$. We have $Y^n\left(\frac{1}{2}\right)_- \ge \frac{1}{k}$ on $[\![0, S^n_k]\!]$, so for all $N > 0$:

$$P'^n\left(h^n\left(\frac{1}{2}\right)_{T^n} > N\right) \le P'^n(S^n_k < T^n) + P'^n\left(Y^n\left(\frac{1}{2}\right)_- \cdot h^n\left(\frac{1}{2}\right)_{T^n} > \frac{N}{k}\right)$$
$$\le P'^n(S^n_k < T^n) + \frac{2k}{N} E_{Q^n}\left(Y^n\left(\frac{1}{2}\right)_- \cdot h^n\left(\frac{1}{2}\right)_{T^n}\right)$$
$$\le P'^n(S^n_k < T^n) + \frac{2k}{N}$$

(the second inequality comes from Tchebycheff's inequality and from $P'^n \le 2Q^n$; the third inequality comes from IV.1.28 and the fact that $H(\frac{1}{2}; P^n_0, P'^n_0) \le 1$ and $H(\frac{1}{2}; P^n_{T^n}, P'^n_{T^n}) \ge 0$). Then our assumption yields that $\limsup_n P'^n(S^n_k < T^n) \ge 1 - 2k/N$. This being true for all $N > 0$ we deduce

2.14 $$\limsup_n P'^n(S^n_k < T^n) = 1.$$

But

$$P'^n(S^n_k < T^n) \le P'^n\left(\inf_{t\le T^n} z'^n_t \le \frac{1}{k}\right) + P'^n\left(\inf_{t\le T^n} z^n_t \le \frac{1}{k}\right)$$

and 1.17 yields $\lim_{k\uparrow\infty}\limsup_n P'^n\left(\inf_t z'^n_t \le \frac{1}{k}\right) = 0$. Therefore 2.14 yields

$$\lim_{k\uparrow\infty}\limsup_n P'^n\left(\inf_{t\le T^n} z^n_t \le \frac{1}{k}\right) = 0,$$

which, in view of 1.20, yields $(P'^n_{T^n}) \Delta (P^n_{T^n})$. □

Next, we show that in order to prove 2.3, 2.4a and 2.5, it suffices to consider the "strict" processes $h'^n(\alpha)$ and $\imath'^n(\beta)$ given by 2.2, instead of arbitrary versions $h^n(\alpha)$ and $\imath^n(\beta)$. To see this, call (i′), (ii′), (iii′) the conditions in these theorems, with $h'^n(\alpha)$ and $\imath'^n(\beta)$ instead of $h^n(\alpha)$ and $\imath^n(\beta)$.

Since $h'^n(\alpha) \le h^n(\alpha)$ and $\imath'^n(\beta) \le \imath^n(\beta)$, we clearly have (i′) ⇒ (i), (ii′) ⇒ (ii) and (iii′) ⇒ (iii) in 2.4, so it is enough to prove that $(P'^n_{T^n}) \Delta (P^n_{T^n})$ implies (i′), (ii′), (iii′) in 2.4.

Similarly, in 2.3 and 2.5 we have (ii) $\Rightarrow$ (ii') and (iii) $\Rightarrow$ (iii'). Conversely, assume that we have proved the implications (i) $\Rightarrow$ (ii'), (iii') in 2.3 and 2.5. Define S_k^n as in the previous proof, and assume that $(P'^n_{T^n}) \lhd (P^n_{T^n})$. By 1.17 and 1.19 both sequences $(\sup_{t \le T^n} 1/z_t'^n | P'^n)$ and $(\sup_{t \le T^n} 1/z_t^n | P'^n)$ are $\mathbb{R}$-tight, thus

2.15 $$\lim_{k \uparrow \infty} \limsup_n P'^n(S_k^n < T^n) = 0.$$

Moreover, $h^n(\alpha) = h'^n(\alpha)$ and $\imath^n(\beta) = \imath'^n(\beta)$ on $\bigcup_k [\![0, S_k^n]\!]$. Then

$$P'^n(h^n(\alpha)_{T^n} > \rho) \le P'^n(h'^n(\alpha)_{T^n} > \rho) + P'^n(S_k^n < T^n),$$

and similarly for $\imath^n(\beta)$. So in view of 2.15, we clearly have (ii') $\Rightarrow$ (ii) and (iii') $\Rightarrow$ (iii) in 2.3 and 2.5. (under (i), of course!)

Therefore, due to what precedes, we can and *will assume that* $h^n(\alpha) = h'^n(\alpha)$ *and* $\imath^n(\beta) = \imath'^n(\beta)$ until the end of this subsection.

2.16 **Lemma.** *We have the implications*: a) 2.3(ii) $\Rightarrow$ 2.3(iii),
b) 2.4(ii) $\Rightarrow$ 2.4(i).

Proof. We have seen in IV.2.24 that for all $N \ge e$, $\alpha \in (0,1)$,

$$h^n(\alpha) \le 8\alpha(\operatorname{Log} N) h^n\left(\frac{1}{2}\right) + 8\varphi_{1/2}\left(1 + \frac{x}{z_-^n}, 1 - \frac{x}{z_-'^n}\right) 1_{\{N(1+x/z_-^n) \le 1 - x/z_-'^n\}} * \nu^{z^n}$$

and so 2.8a and 2.1 yield

2.17 $$h^n(\alpha) \le 8\alpha(\operatorname{Log} N) h^n\left(\frac{1}{2}\right) + 4\imath^n\left(\frac{1}{N}\right).$$

a) Assume 2.3(ii), and let $\varepsilon > 0$, $\eta > 0$. There is $N \in \mathbb{R}_+$ such that

$$\limsup_n P'^n(h^n(1/2)_{T^n} > N) \le \frac{\varepsilon}{2}$$

$$\limsup_n P'^n(\imath^n(1/N)_{T^n} > \eta/8) \le \frac{\varepsilon}{2},$$

so 2.17 yields

$$\limsup_n P'^n(h^n(\alpha)_{T^n} > \eta) \le \limsup_n P'^n\left(h^n\left(\frac{1}{2}\right)_{T^n} > \frac{\eta}{16\alpha \operatorname{Log} N}\right)$$
$$+ \limsup_n P'^n\left(\imath^n\left(\frac{1}{N}\right)_{T^n} > \frac{\eta}{8}\right),$$

which is smaller than ε whenever $\dfrac{\eta}{16\alpha \operatorname{Log} N} \ge N$, i.e. for $\alpha \le \dfrac{\eta}{16 N \log N}$. So 2.3(iii) immediately follows.

b) Assume 2.4(ii). Since $\beta \rightsquigarrow \imath^n(\beta)$ is non-decreasing, it is clearly enough to prove 2.4(i) for $\beta = 1/N$. Set

$$u_n(N,\eta,\varepsilon) = P'^n\left(z_0^n < \varepsilon, \text{ or } h^n\left(\frac{1}{2}\right)_{T^n} > N, \text{ or } \imath^n\left(\frac{1}{N}\right)_{T^n} > \eta\right).$$

Then 2.17 yields for all $\alpha \in (0,1)$:

2.18 $$u_n(N,\eta,\varepsilon) \geq P'^n(z_0^n < \varepsilon, \text{ or } h^n(\alpha)_{T^n} > 4\eta + 8\alpha N \operatorname{Log} N).$$

Let $\rho > 0$. By 2.4(ii) there exist $\theta > 0$ and $\alpha_0 > 0$ such that

$$\alpha \in (0,\alpha_0] \Rightarrow \limsup_n P'^n(z_0^n < \varepsilon, \text{ or } h^n(\alpha)_{T^n} > \theta) \geq 1 - \rho.$$

Then if $\eta_0 = \theta/8$ and $\alpha = \alpha_0 \wedge \dfrac{\theta}{16N \operatorname{Log} N}$, 2.18 yields

$$\eta \leq \eta_0 \Rightarrow \limsup_n u_n(N,\eta,\varepsilon) \geq 1 - \rho.$$

Since $\rho > 0$ is arbitrary, 2.4(i) follows. □

2.19 **Lemma.** *We have the implications*: a) 2.3(iii) $\Rightarrow$ 2.3(ii), b) 2.4(i) $\Rightarrow$ 2.4(iii).

Proof. Using 2.2 and 2.8b and IV.2.21b, we obtain for all $\alpha, \beta \in (0,1)$:

2.20 $$\imath^n(\beta) \leq \frac{1}{\alpha\beta + 1 - \alpha - \beta^\alpha} h^n(\alpha), \quad h^n\left(\frac{1}{2}\right) \leq \tilde{\gamma}_\alpha h^n(\alpha).$$

a) Assume 2.3(iii). Let $\varepsilon > 0$, $\eta > 0$. There exists α_0 such that

2.21 $$\alpha \leq \alpha_0 \Rightarrow \limsup_n P'^n(h^n(\alpha)_{T^n} > \eta) \leq \varepsilon.$$

Then 2.20 yields that $\limsup_n P'^n(h^n(1/2)_{T^n} > N) \leq \varepsilon$ for all $N \geq \eta\tilde{\gamma}_\alpha$, so 2.3(ii.2) holds. 2.20 and 2.21 also yields that $\limsup_n P'^n(\imath^n(\beta)_{T^n} > 2\eta) \leq \varepsilon$ for all β such that $\alpha\beta + 1 - \alpha - \beta^\alpha \geq \frac{1}{2}$ for some $\alpha \leq \alpha_0$; but $\lim_{\beta \downarrow\downarrow 0}(\alpha\beta + 1 - \alpha - \beta^\alpha) = 1 - \alpha$, so $\limsup_n P'^n(\imath^n(\beta)_{T^n} > 2\eta) \leq \varepsilon$ for all β small enough, and we deduce 2.3(ii.3).

b) Assume 2.4(i). In view of 2.20 (with $\beta = 1/2$),

2.22 $$\begin{aligned} &P'^n(z_0^n < \varepsilon, \text{ or } h^n(\alpha)_{T^n} > \eta) \\ &\geq P'^n\left(z_0^n < \varepsilon, \text{ or } h^n\left(\frac{1}{2}\right)_{T^n} > \eta\tilde{\gamma}_\alpha, \text{ or } \imath^n\left(\frac{1}{2}\right)_{T^n} > \eta\left(1 - \frac{\alpha}{2} - 2^{-\alpha}\right)^{-1}\right). \end{aligned}$$

Let $\alpha \in (0,1)$ and $\rho > 0$. There exists $\theta > 0$ such that

$$\limsup_n P'^n\left(z_0^n < \varepsilon, \text{ or } h^n\left(\frac{1}{2}\right)_{T^n} > 1, \text{ or } \imath^n\left(\frac{1}{2}\right)_{T^n} > \theta\right) \geq 1 - \rho.$$

Then 2.22 implies

$$\limsup_n P'^n(z_0^n < \varepsilon, \text{ or } h^n(\alpha)_{T^n} > \eta) \geq 1 - \rho$$

whenever $\eta\tilde{\gamma}_\alpha \leq 1$ and $\eta \leq \left(1 - \frac{\alpha}{2} + 2^{-\alpha}\right)\theta$: these inequalities are true for all η small enough, hence since $\rho > 0$ is arbitrary we deduce 2.4(iii). □

2.23 **Lemma.** *We have* 2.3(i) $\Rightarrow$ 2.3(iii).

Proof. Set $S_k^n = \inf\left(t\colon z_t^n < \frac{1}{k} \text{ or } z_t'^n < \frac{1}{k}\right)$. Using Lemma 1.6, we deduce from our assumption $(P_{T^n}'^n) \lhd (P_{T^n}^n)$ that for all $\rho > 0$, there exists $\alpha_\rho \in (0, 1)$ such that

$$\alpha \leq \alpha_\rho \Rightarrow \limsup_n \,[1 - H^n(\alpha)_{T^n}] \leq \rho.$$

(Recall that $H^n(\alpha)_{T^n} = H(\alpha; P_{T^n}^n, P_{T^n}'^n)$). Then IV.1.28 yields

$$\begin{aligned} E_{Q^n}(Y^n(\alpha)_- \cdot h^n(\alpha)_{T^n \wedge S_k^n}) &\leq E_{Q^n}(Y^n(\alpha)_- \cdot h^n(\alpha)_{T^n}) \\ &= H^n(\alpha)_0 - H^n(\alpha)_{T^n} \leq 1 - H^n(\alpha)_{T^n}, \end{aligned}$$

while $Y^n(\alpha)_- \geq 1/k$ on $[\![0, S_k^n]\!]$; therefore

$$\alpha \leq \alpha_\rho \Rightarrow \limsup_n E_{Q^n}(h^n(\alpha)_{T^n \wedge S_k^n}) \leq k\rho.$$

Using Tchebycheff's inequality and $P'^n \leq 2Q^n$, we deduce for all $\eta > 0$:

2.24 $$\alpha \leq \alpha_\rho \Rightarrow \limsup_n P'^n(h^n(\alpha)_{T^n \wedge S_k^n} > \eta) \leq 2\frac{k\rho}{\eta}.$$

Moreover, we have already seen that the contiguity assumption yields 2.15; thus if $\varepsilon > 0$ there exists $k \in \mathbb{N}^*$ such that $\limsup_n P'^n(S_k^n < T^n) \leq \varepsilon$. This inequality and 2.24 imply

$$\alpha \leq \alpha_\rho \Rightarrow \limsup_n P'^n(h^n(\alpha)_{T^n} > \eta) \leq 2\frac{k\rho}{\eta} + \varepsilon = 2\varepsilon$$

if $\rho = \varepsilon\eta/2k$. Hence we obtain 2.3(iii.2). As for 2.3(iii.1), it is obvious. □

Proof of Theorems 2.3 and 2.4. In view of the results already proven, it remains us to show that: 2.3(iii) $\Rightarrow (P_{T^n}'^n) \lhd (P_{T^n}^n)$ and that $(P_{T^n}'^n) \,\Delta\, (P_{T^n}^n) \Rightarrow$ 2.4(ii).

Let $\eta > 0$, $\alpha \in (0, 1)$. Since the Hellinger process in the strict sense $h^n(\alpha)$ does not "jump to infinity" (see III.5.8), each stopping time $R^n(\alpha, \eta) = \inf(t\colon h^n(\alpha)_t \geq \eta)$ is predictable, and so is Q^n-a. s. announced by a sequence $(R_p^n(\alpha, \eta))_{p \geq 1}$ of stopping times. We have

$$\begin{aligned} 1 - H^n(\alpha)_{T^n} = {} & [1 - H^n(\alpha)_0] + [H^n(\alpha)_0 - H^n(\alpha)_{R_p^n(\alpha,\eta) \wedge T^n \wedge t}] \\ & + [H^n(\alpha)_{R_p^n(\alpha,\eta) \wedge T^n \wedge t} - H^n(\alpha)_{T^n \wedge t}] + [H^n(\alpha)_{T^n \wedge t} - H^n(\alpha)_{T^n}]. \end{aligned}$$

Moreover $h^n(\alpha)_{R_p^n(\alpha,\eta)\wedge T^n\wedge t} \leq \eta$ by construction, so 2.10 and 2.11 yield for all $\varepsilon \geq 0$, $\delta > 0$:

$$\begin{aligned}1 - H^n(\alpha)_{T^n} \leq 1 - H^n(\alpha)_0 &+ 2\eta + E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n \leq \varepsilon\}}) + 2\delta^{1-\alpha} \\ &+ \left(\frac{2}{\delta}\right)^\alpha P'^n[(R_p^n(\alpha,\eta) \wedge t) < (T^n \wedge t), z_0^n > \varepsilon] \\ &+ H^n(\alpha)_{T^n\wedge t} - H^n(\alpha)_{T^n}.\end{aligned}$$

We observe that $\lim_{p\uparrow\infty} \downarrow \{(R_p^n(\alpha,\eta) \wedge t) < (T^n \wedge t)\} = \{R^n(\alpha,\eta) \leq t \wedge T^n\} \subset \{h^n(\alpha)_{T^n} \geq \eta\}$, and IV.1.28 immediately yields $H^n(\alpha)_{T^n\wedge t} \to H^n(\alpha)_{T^n}$ as $t \uparrow \infty$. Then if we let $p \uparrow \infty$ first, then $t \uparrow \infty$, in the above inequality, we obtain:

2.25
$$\begin{aligned}1 - H^n(\alpha)_{T^n} \leq 1 - H^n(\alpha)_0 &+ 2\eta + E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n \leq \varepsilon\}}) + 2\delta^{1-\alpha} \\ &+ \left(\frac{2}{\delta}\right)^\alpha P'^n(h^n(\alpha)_{T^n} \geq \eta, z_0^n > \varepsilon).\end{aligned}$$

a) Assume 2.3(iii). Then 1.6 yields $\lim_{\alpha\downarrow\downarrow 0} \liminf_n H^n(\alpha)_0 = 1$. Hence applying 2.25 with $\varepsilon = 0$ (so $Y^n(\alpha)_0 1_{\{z_0^n \leq \varepsilon\}} = 0$) gives

$$\begin{aligned}&\limsup_{\alpha\downarrow\downarrow 0} \limsup_n [1 - H^n(\alpha)_{T^n}] \\ &\leq 2\eta + \limsup_{\alpha\downarrow\downarrow 0} \left[2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha \limsup_n P'^n(h^n(\alpha)_{T^n} \geq \eta)\right] \\ &= 2\eta + 2\delta + \lim_{\alpha\downarrow\downarrow 0} \limsup_n P'^n(h^n(\alpha)_{T^n} \geq \eta) \\ &= 2\eta + 2\delta,\end{aligned}$$

the last equality coming from 2.3(iii.2). Since η and δ are arbitrary, we deduce that $\lim_{\alpha\downarrow\downarrow 0} \liminf_n H^n(\alpha)_{T^n} = 1$, hence $(P'^n_{T^n}) \lhd (P^n_{T^n})$ follows from 1.6.

b) Assume that $(P'^n_{T^n}) \Delta (P^n_{T^n})$. If $z_0^n > \varepsilon$ we have $z_0^n/z_0'^n \geq \frac{\varepsilon}{2}$, and

$$z_0'^n - Y^n(\alpha)_0 = z_0'^n[1 - (z_0^n/z_0'^n)^\alpha] \leq z_0'^n\left[1 - \left(\frac{\varepsilon}{2}\right)^\alpha\right].$$

Thus

$$\begin{aligned}&1 - H^n(\alpha)_0 + E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n \leq \varepsilon\}}) \\ &= 1 - E_{Q^n}(Y^n(\alpha)_0 1_{\{z_0^n > \varepsilon\}}) \\ &= P'^n(z_0^n \leq \varepsilon) + E_{Q^n}[(z_0'^n - Y^n(\alpha)_0 1_{\{z_0^n > \varepsilon\}}] \\ &\leq P'^n(z_0^n \leq \varepsilon) + \left[1 - \left(\frac{\varepsilon}{2}\right)^\alpha\right] E_{Q^n}(z_0'^n) \\ &\leq \left(\frac{2}{\delta}\right)^\alpha P'^n(z_0^n \leq \varepsilon) + 1 - \left(\frac{\varepsilon}{2}\right)^\alpha\end{aligned}$$

whenever $\delta \in (0,2)$. Then 2.25 yields

$$\begin{aligned}
&1 - H^n(\alpha)_{T^n} \\
&\quad \leq 2\eta + 1 - \left(\frac{\varepsilon}{2}\right)^\alpha + 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha [P'^n(z_0^n \leq \varepsilon) + P'^n(h^n(\alpha)_{T^n} \geq \eta, z_0^n > \varepsilon)] \\
&\quad = 2\eta + 1 - \left(\frac{\varepsilon}{2}\right)^\alpha + 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha P'^n(z_0^n \leq \varepsilon, \text{ or } h^n(\alpha)_{T^n} \geq \eta).
\end{aligned}$$

Now we apply 1.9(iii): $\lim_{\alpha \downarrow\downarrow 0} \liminf_n H^n(\alpha)_{T^n} = 0$. We deduce

$$\begin{aligned}
1 &\leq 2\eta + \liminf_{\alpha \downarrow\downarrow 0} \limsup_n \left[1 - \left(\frac{\varepsilon}{2}\right)^\alpha + 2\delta^{1-\alpha} + \left(\frac{2}{\delta}\right)^\alpha P'^n(z_0^n \leq \varepsilon, \text{ or } h^n(\alpha)_{T^n} \geq \eta)\right] \\
&= 2\eta + 2\delta + \liminf_{\alpha \downarrow\downarrow 0} \limsup_n P'^n(z_0^n \leq \varepsilon, \text{ or } h^n(\alpha)_{T^n} \geq \eta).
\end{aligned}$$

Since δ is arbitrary in $(0,2)$, we obtain $\liminf_{\alpha \downarrow\downarrow 0} \limsup_n P'^n(z_0^n \leq \varepsilon, \text{or } h^n(\alpha)_{T^n} \geq \eta) \geq 1 - 2\eta$ for all $\eta > 0$, which clearly implies 2.4(ii). □

Proof of Theorem 2.5. (i) ⇒ (ii): That (ii.1,2) hold comes from 2.3. We have seen in 1.19 that the sequence $(U^n | P^n)$ is $\mathbb{R}$-tight, where $U^n = \sup_{t \leq T^n} Z_t^n$. Moreover $V^n := \sup_{t \leq T^n}(1/Z_t^n) \leq 2 \sup_{t \leq T^n}(1/z_t'^n)$, so $(V^n | P'^n)$ is $\mathbb{R}$-tight by 1.17. Since $\sup_{t \leq T^n} \alpha_t^n \leq U^n V^n$, we immediately get (ii.3).

(ii) ⇒ (i): In view of Theorem 2.3, it is enough to prove that (ii) ⇒ 2.3(ii.3). If $N \in \mathbb{R}_+$ and if $\psi_{1/N}$ is defined by 2.1, we define an increasing process by:

$$A^{n,N} = \sum_{s \leq \cdot} \psi_{1/N}\left(\frac{z_s^n / z_{s-}^n}{z_s'^n / z_{s-}'^n}\right).$$

From the definition of α_t^n, it is easy to deduce

2.26 $$\sup_{t \leq T^n} \alpha_t^n \geq N \Leftrightarrow A_{T^n}^{n,N} > 0$$

Now, with notation IV.1.41, we have $\mathscr{j}^n(\psi_{1/N}) = (z'^n / z_-'^n) \cdot A^{n,N}$, and $\imath^n(1/N)$ is the Q^n-compensator of $\mathscr{j}^n(\psi_{1/N})$. Thus, using I.3.12, we obtain for all stopping times U^n on $(\Omega^n, \mathscr{F}^n. \mathbf{F}^n)$:

$$\begin{aligned}
E_{P'^n}\left(\imath^n\left(\frac{1}{N}\right)_{U^n}\right) &= E_{Q^n}\left(z_\infty'^n \imath^n\left(\frac{1}{N}\right)_{U^n}\right) = E_{Q^n}\left(z_-'^n \cdot \imath^n\left(\frac{1}{N}\right)_{U^n}\right) \\
&= E_{Q^n}(z_-'^n \cdot \mathscr{j}^n(\psi_{1/N})_{U^n}) = E_{Q^n}(z'^n \cdot A_{U^n}^{n,N}) \\
&= E_{Q^n}(z_\infty'^n A_{U^n}^{n,N}) = E_{P'^n}(A_{U^n}^{n,N}).
\end{aligned}$$

Hence $\imath^n(1/N)$ is L-dominated (see § I.3c) by $A^{n,N}$, for the measure P'^n. So Lenglart's inequality I.3.32 yields for all $\varepsilon > 0$, $\eta > 0$ (since $\Delta A^{n,N} \leq 1$):

$$P'^n\left(\imath^n\left(\frac{1}{N}\right)_{T^n} > \eta\right) \leq \frac{1}{\eta}\left[\varepsilon + E_{P'^n}\left(\sup_{s\leq T^n} \Delta A_s^{n,N}\right)\right] + P'^n(A_{T^n}^{n,N} > \varepsilon)$$

$$\leq \frac{1}{\eta}[\varepsilon + P'^n(A_{T^n}^{n,N} > 0)] + P'^n(A_{T^n}^{n,N} > \varepsilon).$$

Letting $\varepsilon \downarrow\downarrow 0$ above and using 2.26 give

$$P'^n\left(\imath^n\left(\frac{1}{N}\right)_{T^n} > \eta\right) \leq \left(\frac{1}{\eta} + 1\right) P'^n\left(\sup_{t\leq T^n} \alpha_t^n \geq N\right)$$

and (ii) $\Rightarrow$ 2.3(ii.3) immediately follows. □

§ 2c. The Discrete Case

1. Now we translate the results of § 2a in the discrete-time setting. For every $n \in \mathbb{N}^*$ we have a discrete-time basis $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n = (\mathscr{F}_p^n)_{p\in\mathbb{N}})$ endowed with two probability measures P^n and P'^n, and we suppose that $\mathscr{F}^n = \mathscr{F}_{\infty-}^n$. Let $Q^n = (P^n + P'^n)/2$, and call $z^n = (z_p^n)_{p\in\mathbb{N}}$ and $z'^n = (z_p'^n)_{p\in\mathbb{N}}$ the density processes of P^n and P'^n relatively to Q^n. For all $p \geq 1$ we set

$$\beta_p^n = z_p^n/z_{p-1}^n, \quad \beta_p'^n = z_p'^n/z_{p-1}'^n$$

(with $0/0 = 0$). In view of IV.1.63 and IV.1.66, Theorems 2.3 and 2.4 become

2.27 Theorem. *For each $n \in \mathbb{N}^*$ let T^n be a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. The following statements are equivalent:*

(i) $(P_{T^n}'^n) \lhd (P_{T^n}^n)$;

(ii) (1) $(P_0'^n) \lhd (P_0^n)$,

(2) $\lim_{N\uparrow\infty} \limsup_n P'^n(\sum_{1\leq p\leq T^n} E_{Q^n}((\sqrt{\beta_p^n} - \sqrt{\beta_p'^n})^2 | \mathscr{F}_{p-1}^n) > N) = 0$,

(3) *for all $\eta > 0$ we have*

$$\lim_{\gamma\downarrow\downarrow 0} \limsup_n P'^n\left(\sum_{1\leq p\leq T^n} E_{Q^n}(\beta_p'^n 1_{\{\beta_p^n \leq \gamma\beta_p'^n\}} | \mathscr{F}_{p-1}^n) > \eta\right) = 0.$$

(iii) (1) $(P_0'^n) \lhd (P_0^n)$,

(2) *for all $\eta > 0$ we have*

$$\lim_{\alpha\downarrow\downarrow 0} \limsup_n P'^n\left(\sum_{1\leq p\leq T^n} E_{Q^n}(1 - (\beta_p^n)^\alpha (\beta_p'^n)^{1-\alpha} | \mathscr{F}_{p-1}^n) > \eta\right) = 0.$$

2.28 Remark. As in Remark 2.7, we could replace (ii.3) above by

$$\lim_{\gamma\downarrow\downarrow 0} \limsup_n P'^n\left(\sum_{1\leq p\leq T^n} E_{Q^n}[(\sqrt{\beta_p^n} - \sqrt{\beta_p'^n})^2 1_{\{\beta_p^n \leq \gamma\beta_p'^n\}} | \mathscr{F}_{p-1}^n) > \eta\right) = 0$$

for all $\eta > 0$. □

2.29 Theorem. *For each $n \in \mathbb{N}^*$ let T^n be a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$.*
a) *If $(P'^n_{T^n}) \triangle (P^n_{T^n})$ we have*

$$\text{(i)}\quad \lim_{\eta \downarrow\downarrow 0} \limsup_n P'^n\Bigg(z^n_0 < \varepsilon \text{ or } \sum_{1 \le p \le T^n} E_{Q^n}((\sqrt{\beta^n_p} - \sqrt{\beta'^n_p})^2 | \mathscr{F}^n_{p-1}) > N$$
$$\text{or } \sum_{1 \le p \le T^n} E_{Q^n}(\beta'^n_p 1_{\{\beta^n_p \le \gamma \beta'^n_p\}} | \mathscr{F}^n_{p-1}) > \eta\Bigg) = 1$$

for all $N \in \mathbb{R}_+$, $\gamma \in (0,1)$, $\varepsilon > 0$;

$$\text{(ii)}\quad \lim_{\eta \downarrow\downarrow 0} \liminf_{\alpha \downarrow\downarrow 0} \limsup_n P'^n\Bigg(z^n_0 < \varepsilon \text{ or }$$
$$\sum_{1 \le p \le T^n} E_{Q^n}(1 - (\beta^n_p)^\alpha (\beta'^n_p)^{1-\alpha} | \mathscr{F}^n_{p-1}) > \eta\Bigg) = 1$$

for all $\varepsilon > 0$;

$$\text{(iii)}\quad \lim_{\eta \downarrow\downarrow 0} \limsup_n P'^n\Bigg(z^n_0 < \varepsilon \text{ or } \sum_{1 \le p \le T^n} E_{Q^n}(1 - (\beta^n_p)^\alpha (\beta'^n_p)^{1-\alpha} | \mathscr{F}^n_{p-1}) > \eta\Bigg) = 1$$

for all $\varepsilon > 0$, $\alpha \in (0,1)$.
b) *If $\limsup_n P'^n(\sum_{1 \le p \le T^n} E_{Q^n}((\sqrt{\beta^n_p} - \sqrt{\beta'^n_p})^2 | \mathscr{F}^n_{p-1}) > N) = 1$ for all $N \in \mathbb{R}_+$, then $(P'^n_{T^n}) \triangle (P^n_{T^n})$.*

Finally, the version of Theorem 2.5 goes as follows. Set

$$Z^n_p = z'^n_p / z^n_p, \quad \alpha^n_p = \begin{cases} Z^n_p / Z^n_{p-1} & \text{if } 0 < Z^n_{p-1} < \infty \\ 0 & \text{if } Z^n_{p-1} = 0 \\ +\infty & \text{if } Z^n_{p-1} = +\infty. \end{cases}$$

2.30 Theorem. *For each $n \in \mathbb{N}^*$ let T^n be a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. We have $(P'^n_{T^n}) \lhd (P^n_{T^n})$ if and only if $(P'^n_0) \lhd (P^n_0)$ and 2.27(ii.2) holds and the sequence $(\sup_{1 \le p \le T^n} \alpha^n_p | P'^n)$ is $\mathbb{R}$-tight.*

2. We now consider the case of "independent random variables", for which (like in § IV.2.3) we can get necessary and sufficient conditions for entire separation as well as for contiguity. The setting is as in § IV.1e.*2*: more precisely, we assume that

$$\text{2.31}\quad \begin{cases} \Omega = (\mathbb{R}^d)^{\mathbb{N}^*}, & \xi_p(\omega) = p^{\text{th}}\text{-coordinate of } \omega \in \Omega \text{ (so } \xi_p(\omega) \in \mathbb{R}^d), \\ \mathscr{F} = (\mathscr{R}^d)^{\mathbb{N}^* \otimes}, & \mathscr{F}_0 = \text{trivial } \sigma\text{-field}, \ \mathscr{F}_p = \sigma(\xi_1, \dots, \xi_p) \text{ for } p \ge 1, \\ P^n = \bigotimes_{p \in \mathbb{N}^*} \rho^n_p, & P'^n = \bigotimes_{p \in \mathbb{N}^*} \rho'^n_p \end{cases}$$

where ρ^n_p, ρ'^n_p are probability measures on $\mathbb{R}^d$: then under P^n (resp. P'^n) the random variables $(\xi_p)_{p \ge 1}$ are independent, with distributions ρ^n_p (resp. ρ'^n_p).

2.32 **Theorem.** *Assume 2.31, and for each $n \in \mathbb{N}^*$ let $k_n \in \bar{\mathbb{N}} = \{1, 2, \dots, +\infty\}$.*

a) *We have $(P'^n_{k_n}) \lhd (P^n_{k_n})$ if and only if*

2.33
$$\lim_{\alpha \downarrow\downarrow 0} \limsup_n \sum_{1 \le p \le k_n} [1 - H(\alpha; \rho^n_p, \rho'^n_p)] = 0.$$

b) *We have $(P'^n_{k_n}) \,\Delta\, (P^n_{k_n})$ if and only if the following holds for at least one $\alpha \in (0,1)$:*

2.34
$$\begin{cases} \text{either (i)} & \limsup_n \sum_{1 \le p \le k_n} [1 - H(\alpha; \rho^n_p, \rho'^n_p)] = \infty \\ \text{or (ii)} & \liminf_n \inf_{1 \le p \le k_n} H(\alpha; \rho^n_p, \rho'^n_p) = 0, \end{cases}$$

and in this case 2.34 holds for all $\alpha \in (0,1)$.

Proof. a) Set $a^n_p(\alpha) = 1 - H(\alpha; \rho^n_p, \rho'^n_p)$. Recall from IV.1.73 that a version of $h(\alpha; P^n, P'^n)$ is

$$h^n(\alpha)_p = \sum_{1 \le q \le p} a^n_q(\alpha),$$

which is deterministic. Since $P'^n_0 = P^n_0$ is the trivial probability on $\mathscr{F}_0$ the claim follows from the equivalence (i)$\Leftrightarrow$(iii) in Theorem 2.3.

Note also that (see IV.1.74)

2.35
$$H(\alpha; P^n_{k_n}, P'^n_{k_n}) = \prod_{1 \le p \le k_n} (1 - a^n_p(\alpha)),$$

so the claim could also be proved (more directly) by using 1.6.

b) 2.34(ii) and 2.35 obviously yield $\liminf_n H(\alpha; P^n_{k_n}, P'^n_{k_n}) = 0$. Since $\mathrm{Log}(1-x) \le -x$ we have

$$\liminf_n \mathrm{Log}\, H(\alpha; P^n_{k_n}, P'^n_{k_n}) \le -\limsup_n \sum_{1 \le p \le k_n} a^n_p(\alpha),$$

so 2.34(i) also yields $\liminf_n H(\alpha; P^n_{k_n}, P'^n_{k_n}) = 0$, and $(P'^n_{k_n}) \,\Delta\, (P^n_{k_n})$ follows from 1.9.

Conversely, assume that $(P'^n_{k_n}) \,\Delta\, (P^n_{k_n})$, so 1.9 again implies that $\liminf_n H(\alpha; P^n_{k_n}, P'^n_{k_n}) = 0$ for all $\alpha \in (0,1)$, and assume also that 2.34(ii) does not hold, for some α. Therefore there is $\eta < 1$ such that $0 \le a^n_p(\alpha) \le \eta$ for all $p \le k_n$, for all n large enough. But there is a constant C_η such that $\mathrm{Log}(1-x) \ge -C_\eta x$ for $x \in [0, \eta]$, and so by 2.35

$$\limsup_n \sum_{1 \le p \le k_n} a^n_p(\alpha) \ge -\frac{1}{C_\eta} \liminf_n \mathrm{Log}\, H(\alpha; P^n_{k_n}, P'^n_{k_n}) = +\infty;$$

therefore 2.34(i) holds. $\square$

Finally we end this section with two very closely related examples.

2.36 **Example.** We have 2.31 with $d = 1$ and

$$\rho^n_p(dx) = 1_{[0,1]}(x)\,dx, \quad \rho'^n_p(dx) = \frac{1}{1 - a_n} 1_{[a_n, 1]}(x)\,dx$$

where $a_n \in [0, 1)$. Then (using $1_{[0,1]}(x)\,dx$ as the dominating measure) a very simple computation shows that

$$H(\alpha; \rho_p^n, \rho_p'^n) = (1 - a_n)^\alpha$$

for $\alpha \in (0,1)$. Then if $k_n = n$, 2.32 easily gives (recalling that $H(\alpha; P, P') = H(1 - \alpha; P', P)$):

a) $(P_n'^n) \lhd (P_n^n) \Leftrightarrow \limsup_n na_n < \infty \left(\text{i.e. } a_n = 0\left(\frac{1}{n}\right)\right)$

b) $(P_n^n) \lhd (P_n'^n) \Leftrightarrow \limsup_n na_n = 0 \left(\text{i.e. } a_n = o\left(\frac{1}{n}\right)\right)$

c) $(P_n'^n) \,\Delta\, (P_n^n) \Leftrightarrow \limsup_n na_n = \infty$. □

2.37 **Example.** The same as 2.36, except that $\rho_p'^n(dx) = 1_{[a_n, 1+a_n]}(x)\,dx$. Then $H(\alpha; \rho_p^n, \rho_p'^n) = 1 - a_n$, and

a) $(P_n'^n) \lhd (P_n^n) \Leftrightarrow (P_n^n) \lhd (P_n'^n) \Leftrightarrow \limsup_n na_n = 0$

b) $(P_n'^n) \,\Delta\, (P_n^n) \Leftrightarrow \limsup_n na_n = \infty$. □

3. Examples

This section is the analogue of Section IV.4, where we studied absolute continuity for point processes, diffusion processes and PII's, but here we examine contiguity and entire separation.

We place ourselves again within the *canonical setting* III.2.13: Ω is the canonical space of all $\mathbb{R}^d$-valued càdlàg functions on $\mathbb{R}_+$, with X the canonical process and $\mathbf{F}$ the canonical filtration (generated by X) and $\mathscr{F} = \mathscr{F}_{\infty-}$, and $\mathscr{H} = \mathscr{F}_0^0 = \sigma(X_0)$. We also consider two sequences (P^n) and (P'^n) of probability measures on $(\Omega, \mathscr{F})$.

§ 3a. Point Processes

Here $d = 1$ and X is a. s. a *point process* under each one of the measures P^n and P'^n. We call A^n and A'^n the compensators of X under P^n and P'^n respectively. Let $\bar{A}^n$ be any increasing predictable process such that $dA_t^n \ll d\bar{A}_t^n$ and $dA_t'^n \ll d\bar{A}_t^n$, and let g^n, g'^n be two predictable nonnegative processes such that

3.1 $$A^n = g^n \cdot \bar{A}^n, \quad A'^n = g'^n \cdot \bar{A}^n \quad (P^n + P'^n)\text{-a. s.}$$

Then Theorems 2.3 and IV.4.2 give (recall that $P_0^n = P_0'^n$ here):

3.2 **Theorem.** *For each* $n \in \mathbb{N}^*$ *let* T^n *be a stopping time. Then* $(P'^n_{T^n}) \lhd (P^n_{T^n})$ *if and only if*

3.3 $$\lim_{N\uparrow\infty} \limsup_n P'^n\Big((\sqrt{g^n} - \sqrt{g'^n})^2 \cdot \bar{A}^n_{T^n} + \sum_{s \le T^n} (\sqrt{1 - \Delta A^n_s} - \sqrt{1 - \Delta A'^n_s})^2 > N\Big) = 0$$

3.4 $$\lim_{\beta\downarrow\downarrow 0} \limsup_n P'^n\Big(g'^n 1_{\{g^n \le \beta g'^n\}} \cdot \bar{A}^n_{T^n} + \sum_{s \le T^n} (1 - \Delta A'^n_s) 1_{\{1-\Delta A^n_s \le \beta(1-\Delta A'^n_s)\}} > \eta\Big) = 0$$

for all $\eta > 0$.

§ 3b. Generalized Diffusion Processes

Here again $d = 1$. We suppose that under P^n and P'^n respectively, X has the form

3.5 $$\begin{cases} X_t = x^n + \int_0^t \beta^n_s\, ds + W^n_t, & W^n \text{ is a } P^n\text{-standard Wiener process} \\ X_t = x^n + \int_0^t \beta'^n_s\, ds + W'^n_t, & W'^n \text{ is a } P'^n\text{-standard Wiener process,} \end{cases}$$

and we define the following generalized increasing predictable processes:

3.6 $$K^n_t = \int_0^t (\beta^n_s)^2\, ds, \quad K'^n_t = \int_0^t (\beta'^n_s)^2\, ds, \quad \tilde{K}^n_t = \int_0^t (\beta^n_s - \beta'^n_s)^2\, ds.$$

3.7 **Theorem.** *Assume that for each* $n \in \mathbb{N}^*$ *the processes* K^n *and* K'^n (and so $\tilde{K}^n$ as well) *do not jump to infinity* (see III.5.8), *and that* $P^n(K^n_t < \infty) = P'^n(K'^n_t < \infty) = 1$ *for all* $t \in \mathbb{R}_+$, *and let* T^n *be a stopping time.*

a) $(P'^n_{T^n}) \lhd (P^n_{T^n})$ *if and only if the sequence* $(\tilde{K}^n_{T^n} | P'^n)$ *is* $\mathbb{R}$*-tight.*

b) $(P'^n_{T^n}) \,\Delta\, (P^n_{T^n})$ *if and only if* $\limsup_n P'^n(\tilde{K}^n_{T^n} > N) = 1$ *for all* $N \in \mathbb{R}_+$.

Proof. In view of IV.4.23c, our assumptions imply that $\tilde{K}^n/8$ is a version of $h(1/2; P^n, P'^n)$. Moreover, the proof of IV.4.23 yields that Theorem IV.3.28 applies, so a version of $\imath(\psi; P^n, P'^n)$ is given by IV.3.11, i.e. $\imath(\psi; P^n, P'^n) = 0$ (because here the measures ν, ν', λ appearing in IV.3.11 are equal to 0). Finally $P^n_0 = P'^n_0$ because $X_0 = x^n$ P^n-a.s. and P'^n-a.s., thus with the notation of § 2a we have $z^n_0 = 1$.

Then (a) follows from the equivalence (i) $\Leftrightarrow$ (ii) of Theorem 2.3, and (b) follows from 2.4(a.i, b). □

§ 3c. Processes with Independent Increments

Now we assume that $X - X_0$ is a PII under P^n and P'^n (for all $n \in \mathbb{N}^*$), with deterministic characteristics (B^n, C^n, ν^n) and (B'^n, C'^n, ν'^n) respectively, and relative to the same *truncation function h*.

Let λ^n be any nonnegative measure on $\mathbb{R}_+ \times \mathbb{R}^d$ such that $\nu^n \ll \lambda^n$ and $\nu'^n \ll \lambda^n$, and let $U^n = d\nu^n/d\lambda^n$ and $U'^n = d\nu'^n/d\lambda^n$ be the Radon-Nikodym derivatives. We also set

3.8 $$a_t^n = \nu^n(\{t\} \times \mathbb{R}^d), \quad a_t'^n = \nu'^n(\{t\} \times \mathbb{R}^d)$$

and we consider a continuous increasing function A^n and a Borel function c^n on $\mathbb{R}_+$, taking values in the set of all nonnegative symmetric $d \times d$ matrices, such that

3.9 $$C^{n,ij} = c^{n,ij} \cdot A^n.$$

Finally, recall that P_H^n and $P_H'^n$ are the restrictions of P^n and P'^n to the initial σ-field $\mathscr{H} = \sigma(X_0)$.

3.10 **Theorem.** $(P'^n) \lhd (P^n)$ *if and only if all the following conditions hold*:

(i) $(P_H'^n) \lhd (P_H^n)$;

(ii) $\limsup_n [(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_\infty^n + \sum_{(s)} (\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2] < \infty$;

(iii) $\lim_{\beta \downarrow\downarrow 0} \limsup_n [U'^n 1_{\{U^n \leq \beta U'^n\}} * \lambda_\infty^n + \sum_{(s)} (1 - a_s'^n) 1_{\{1 - a_s^n \leq \beta(1 - a_s'^n)\}} = 0$;

(iv) *There is* $n_0 \in \mathbb{N}^*$ *such that for all* $n \geq n_0$,

(1) $C_t'^n = C_t^n$ *for all* $t \in \mathbb{R}_+$,

(2) $|h(x)(U^n - U'^n)| * \lambda_t^n < \infty$ *for all* $t \in \mathbb{R}_+$,

(3) *there are Borel functions* $\tilde{\beta}^n$: $\mathbb{R}_+ \to \mathbb{R}^d$ *such that for all* $t \in \mathbb{R}_+$

3.11 $$B_t^n - B_t'^n = h(x)(U^n - U'^n) * \lambda_t^n + \left(\sum_{j \leq d} c^{n,\cdot j} \tilde{\beta}^{n,j} \right) \cdot A_t^n$$

(4) $\limsup_n (\tilde{\beta}^n \cdot c^n \cdot \tilde{\beta}^n) \cdot A_\infty^n < \infty$.

Proof. Let us define the following (deterministic) times:

3.12 $$\tau^n = \inf(t\colon C_t'^n \neq C_t^n, \text{ or } |h(x)(U^n - U'^n)| * \lambda_t^n = \infty, \text{ or } B_t^n - B_t'^n \text{ fails to be of the form 3.11 for a suitable function } \tilde{\beta}^n)$$

and the increasing functions:

3.13 $$\begin{cases} h^{n,0}\left(\frac{1}{2}\right)_t = \frac{1}{8}(\tilde{\beta} \cdot c^n \cdot \tilde{\beta}^n) 1_{[0,\tau^n)} \cdot A_t^n + \frac{1}{2}(\sqrt{U^n} - \sqrt{U'^n})^2 1_{[0,\tau^n)} * \lambda_t^n \\ \qquad + \sum_{s \leq t, s < \tau^n} \frac{1}{2}(\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2 \\ \imath^{n,0}(\beta)_t = U'^n 1_{\{U^n \leq \beta U'^n\}} 1_{[0,\tau^n)} * \lambda_t^n + \sum_{s \leq t, s < \tau^n} (1 - a_s'^n) 1_{\{1 - a_s^n \leq \beta(1 - a_s'^n)\}}. \end{cases}$$

Finally, set

3.14 $$\tau'^n = \tau^n \wedge \inf(t\colon h^{n,0}(\tfrac{1}{2})_t = \infty).$$

Then, in virtue of IV.4.24, the following functions are versions of $h(1/2; P^n, P'^n)$ and $\imath(\psi_\beta; P^n, P'^n)$ (ψ_β is defined in 2.1, with $\beta \in (0, 1)$):

3.15 $$\begin{cases} h^n(\tfrac{1}{2})_t = 1_{[0,\tau'^n)} \cdot h^{n,0}(\tfrac{1}{2})_t + 1_{\{\tau'^n \le t\}} \\ \imath^n(\beta)_t = 1_{[0,\tau'^n)} \cdot \imath^{n,0}(\beta)_t + 1_{\{\tau'^n \le t\}}. \end{cases}$$

Moreover, IV.4.28 yields

3.16 $$H(\alpha; P_0^n, P_0'^n) = \begin{cases} H(\alpha; P_H^n, P_H'^n) & \text{if } \tau'^n > 0 \\ 0 & \text{if } \tau'^n = 0. \end{cases}$$

Now, since $h^n(1/2)$ and $\imath^n(\beta)$ are deterministic, the equivalence (i)$\Leftrightarrow$(ii) of Theorem 2.3 reads as follows (apply 2.3 for $T^n \equiv \infty$): we have $(P'^n) \lhd (P^n)$ if and only if

3.17 $$\begin{cases} \text{(a) } (P_0'^n) \lhd (P_0^n) \\ \text{(b) } \limsup_n h^n(1/2)_\infty < \infty \\ \text{(c) } \lim_{\beta \downarrow\downarrow 0} \limsup_n \imath^n(\beta)_\infty = 0. \end{cases}$$

After these preliminaries, let us assume 3.17. Then (c) implies $\tau'^n = \infty$ for all $n \ge n_0$, where n_0 is some integer (so (iv.1, 2, 3) holds) and also implies (iii) (use 3.13 and 3.15). Since $\tau'^n = \infty$ for $n \ge n_0$, (a) and 1.6 and 3.16 imply (i). Moreover if $\tau'^n = \infty$ we also have

3.18 $$\begin{aligned} h^n\left(\frac{1}{2}\right)_t &= \frac{1}{8}(\tilde\beta^n \cdot c^n \cdot \tilde\beta^n) \cdot A_t^n + \frac{1}{2}(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_t^n \\ &\quad + \frac{1}{2} \sum_{s \le t} (\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2 \end{aligned}$$

and therefore (b) yields (ii) and (iv.4).

Conversely assume that (i)–(iv) hold. Then (iv.1, 2, 3) imply $\tau^n = \infty$ for $n \ge n_0$, and 3.13 and (ii) and (iv.4) imply that $\limsup_n h^{n,0}(1/2)_\infty < \infty$, hence $\tau'^n = \infty$ for all n large enough. Thus 3.16 and (i) and 1.6 yields 3.17(a). Furthermore, we have 3.18 and $\imath^n(\beta) = \imath^{n,0}(\beta)$ for all n large enough, so 3.17(b) holds, and 3.17(c) follows from (iii): hence 3.17 is met, and thus $(P'^n) \lhd (P^n)$. □

3.19 **Remark.** If $P^n = P$ and $P'^n = P'$ for all n, one easily checks that 3.10(i–iv) reduce to conditions (i–vii) of IV.4.32. □

3.20 **Remark.** Let us consider the following measures on $\mathbb{R}_+ \times \mathbb{R}^d$:

$$\chi^n(dt, dx) = \nu^n(dt, dx) + \sum_{s: a_s^n > 0} (1 - a_s^n)\varepsilon_{(s,0)}(dt, dx)$$

$$\chi'^n(dt, dx) = \nu'^n(dt, dx) + \sum_{s: a_s'^n > 0} (1 - a_s'^n)\varepsilon_{(s,0)}(dt, dx)$$

$$\bar\chi^n(dt, dx) = \lambda^n(dt, dx) 1_{\{x \ne 0\}} + \sum_{s: a_s^n > 0 \text{ or } a_s'^n > 0} \varepsilon_{(s,0)}(dt, dx).$$

Then $\chi^n \ll \bar{\chi}^n$, $\chi'^n \ll \bar{\chi}^n$, and the derivatives $\zeta^n = d\chi^n/d\bar{\chi}^n$, $\zeta'^n = d\chi'^n/d\bar{\chi}^n$ are

$$\zeta^n(t,x) = U^n(t,x)1_{\{x\neq 0\}} + 1_{\{a_t^n \neq 0\}}(1 - a_t^n)1_{\{x=0\}}$$
$$\zeta'^n(t,x) = U'^n(t,x)1_{\{x\neq 0\}} + 1_{\{a_t'^n \neq 0\}}(1 - a_t'^n)1_{\{x=0\}}.$$

So if we define the Hellinger distance $\rho(\chi^n, \chi'^n)$ by IV.1.4, we obtain

$$\rho^2(\chi^n, \chi'^n) = \frac{1}{2}(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_\infty^n + \frac{1}{2}\sum_{(s)}(\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2.$$

Hence we obtain:

$$3.10(\text{ii}) \Leftrightarrow \limsup_n \rho^2(\chi^n, \chi'^n) < \infty$$

$$3.10(\text{iii}) \Leftrightarrow \lim_{\beta\downarrow\downarrow 0} \limsup_n \chi'^n(\zeta^n \leq \beta\zeta'^n) = 0.$$

The last condition also reads: the sequence $(\zeta'^n/\zeta^n|\chi'^n)$ is $\mathbb{R}$-tight, so in view of 1.6ii′ it could be called "$(\chi'^n) \lhd (\chi^n)$", or "(χ'^n) is contiguous to (χ^n)" (in fact, this definition of contiguity is not equivalent to the definition 1.1 here, because the measures χ^n and χ'^n may be infinite). □

Obviously, there is a version of this theorem which gives necessary and sufficient conditions for having $(P_t'^n) \lhd (P_t^n)$, for any $t \in \mathbb{R}_+$: we leave it to the reader. We turn now to entire separation: like in 2.32, and unlike in 2.5, we have a necessary and sufficient condition here again.

3.21 **Theorem.** *In order that $(P'^n) \,\Delta\, (P^n)$ it is necessary and sufficient that at least one of the following conditions hold*:

(i) $(P_H'^n) \,\Delta\, (P_H^n)$;

(ii) *there is a sequence $n_k \uparrow \infty$ for which at least one of the conditions 3.10(*iv*) (1, 2, 3) fails,*

(iii) $\limsup_n [(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_\infty^n + \sum_{(s)}(\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2] = \infty$,

(iv) $\limsup_n \inf_t [\int \lambda^n(\{t\} \times dx)\sqrt{U^n(t,x)U'^n(t,x)} + \sqrt{1 - a_t^n}\sqrt{1 - a_t'^n}] = 0$,

(v) *if 3.10(*iv.1, 2, 3*) holds, then* $\limsup_n(\tilde{\beta}^n \cdot c^n \cdot \tilde{\beta}^n) \cdot A_\infty^n = \infty$.

Note that neither $\int \lambda^n(\{t\} \times dx)\sqrt{U^n(t,x)U'^n(t,x)}$ nor (iii) depend on the measure λ^n that dominates ν^n and ν'^n. When $P^n = P$ and $P'^n = P'$ for all n, this theorem reduces to IV.4.33.

Proof. Define τ^n and τ'^n and $h^n(\frac{1}{2})$ by 3.12, 3.14, 3.15. Then IV.4.28 gives

$$H(\tfrac{1}{2}; P^n, P'^n) = \begin{cases} H(\tfrac{1}{2}; P_H^n, P_H'^n)\mathscr{E}(-h^n(\tfrac{1}{2}))_\infty & \text{if } \tau'^n > 0 \\ 0 & \text{if } \tau'^n = 0, \end{cases}$$

where $\mathscr{E}(-h(1/2))_\infty$ is given by IV.4.25, i.e.:

$$\mathscr{E}\left(-h^n\left(\frac{1}{2}\right)\right)_\infty = \begin{cases} e^{-h^n(1/2)_\infty} \prod_{(s)} \left[\left(1 - \Delta h^n\left(\frac{1}{2}\right)_s\right) e^{-\Delta h^n(1/2)_s}\right] & \text{if } h^n\left(\frac{1}{2}\right)_\infty < \infty \\ 0 & \text{if } h^n\left(\frac{1}{2}\right)_\infty = \infty \end{cases}$$

Then we easily deduce that $\liminf_n H(\frac{1}{2}; P^n, P'^n) = 0$, which by 1.9 is equivalent to $(P'^n) \mathbin{\Delta} (P^n)$, is satisfied if and only if at least one of the following conditions holds:

(1) $\liminf_n H(\frac{1}{2}; P^n_H, P'^n_H) = 0$,

(2) $\tau^{n_k} < \infty$ for a sequence $n_k \uparrow \infty$,

(3) $\tau^n = \infty$ for all n large enough, and $\limsup_n h^n(\frac{1}{2})_\infty = \infty$,

(4) $\tau^n = \infty$ for all n large enough, and $\liminf_n \sup_t \Delta h^n(\frac{1}{2})_t = 1$.

Now, (1) = (i) by 1.9, and (2) = (ii) is obvious. If (ii) fails, for all n large enough we have 3.18; therefore if (ii) fails we have (3) if and only if either (iii) or (v) holds. Finally, if 3.18 holds we have for all $t \in \mathbb{R}_+$:

$$\Delta h^n\left(\frac{1}{2}\right)_t = 1 - \int \lambda^n(\{t\} \times dx)\sqrt{U^n(t,x)U'^n(t,x)} - \sqrt{1 - a^n_t}\sqrt{1 - a'^n_t}.$$

Hence if (ii) fails, we have (4) if and only if (iv) holds. Thus (1)–(4) is equivalent to (i)–(v), and the claim is proved. □

3.22 **Corollary.** *Assume that for all P^n and P'^n the process X is a PIIS* (so $X_0 = 0$ a. s.). *Then either $P'^n = P^n$ for all n large enough, or $(P'^n) \mathbin{\Delta} (P^n)$.*

Proof. We have $P'^n_H = P^n_H$. If 3.21(ii) holds we know that $(P'^n) \mathbin{\Delta} (P^n)$. Assume that 3.21(ii) fails, and that there is a sequence $n_k \uparrow \infty$ such that $P'^{n_k} \neq P^{n_k}$. Because of the homogeneity of X under P^n and P'^n, we have

$$(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda^n_t = \alpha^n t, \quad (\tilde{\beta}^n \cdot c^n \cdot \tilde{\beta}^n) \cdot A^n_t = \gamma^n t$$

for some nonnegative constants α^n, γ^n. Moreover we have $C'^n = C^n$ for all large enough (because 3.21(ii) fails), and $\alpha^n = 0$ if and only if $\nu'^n = \nu^n$, and $\gamma^n = 0$ if and only if $B'^n = B^n$. Therefore $\alpha^{n_k} > 0$ or $\gamma^{n_k} > 0$ for all k, which clearly implies that either 3.21(iii) or 3.21(v) holds: so again $(P'^n) \mathbin{\Delta} (P^n)$. □

4. Variation Metric

Here we examine the relations between the *total variation* $\|P' - P\|$ of the difference of two probability measures on a filtered space, and the Hellinger processes $h(\alpha; P, P')$. This allows us to derive our first *limit theorems*: namely, conditions for having $\|P^n - P\| \to 0$. Of course the conditions (some of them being necessary and sufficient) are very stringent, and this kind of results is not widely applicable, except perhaps for point processes, and for "generalized diffusion processes" in the sense of § 3b: these two examples are studied in § 4c,d, while the first two subsections are concerned with generalities.

§ 4a. Variation Metric and Hellinger Integrals

1. $(\Omega, \mathscr{F})$ denotes a measurable space. If μ is a finite signed measure on it, we define its *total variation* $\|\mu\|$ as

4.1 $$\|\mu\| = \sup\{|\mu(\varphi)|: \varphi\ \mathscr{F}\text{-measurable function on } \Omega \text{ with } |\varphi| \leq 1\}.$$

This obviously defines a norm on the space of all finite signed measures on $(\Omega, \mathscr{F})$, and the following is well known (see any book on measure theory for the proof!)

4.2 **Lemma.** a) *If μ is a positive measure,* $\|\mu\| = \mu(\Omega)$.

b) *If $\mu = \mu_+ - \mu_-$ is the Jordan-Hahn decomposition of μ, then* $\|\mu\| = \|\mu_+\| + \|\mu_-\| = \mu_+(\Omega) + \mu_-(\Omega)$.

c) $\|\mu\| = \sup\{\sum_{i=1}^n |\mu(A_i)|: (A_i)_{1\leq i\leq n}\ \mathscr{F}\text{-measurable partition}\}$.

Now, P and P' are two probability measures on $(\Omega, \mathscr{F})$. $\|P - P'\|$ is called the *distance in variation between P and P'*. Next, let $Q = \dfrac{P + P'}{2}$ and call $z = dP/dQ$ and $z' = dP'/dQ$ the Radon-Nikodym derivatives of P and P', with respect to Q

4.3 **Lemma.** a) $\|P' - P\| = 2\sup_{A\in\mathscr{F}} |P(A) - P'(A)|$.

b) $\|P' - P\| = E_Q(|z - z'|) = 2E_Q(|1 - z|)$.

Proof. a) $P(A) - P'(A) = P'(A^c) - P(A^c)$ for all $A \in \mathscr{F}$; hence

$$2|P(A) - P'(A)| = |P(A) - P'(A)| + |P(A^c) - P'(A^c)| \leq \|P - P'\|$$

by 4.2c. The Jordan-Hahn decomposition $P - P' = \mu_+ - \mu_-$ has the form $\mu_+ = 1_A \cdot (P - P')$, $\mu_- = 1_{A^c} \cdot (P' - P)$ for some $A \in \mathscr{F}$. Then $\mu_+(\Omega) = P(A) - P'(A)$ and $\mu_-(\Omega) = P'(A^c) - P(A^c)$, and we deduce $\|P - P'\| = 2|P(A) - P'(A)|$.

b) For all $\mathscr{F}$-measurable functions φ with $|\varphi| \leq 1$, we have by definition of z and z':

$$|E_P(\varphi) - E_{P'}(\varphi)| = |E_Q(\varphi(z - z'))| \leq E_Q(|\varphi||z - z'|) \leq E_Q(|z - z'|),$$

and if $\varphi = \operatorname{sign}(z' - z)$ ($=1$ if $z' \geq z$, $= -1$ if $z' < z$), we get $|E_P(\varphi) - E_{P'}(\varphi)| = E_Q(|z - z'|)$: hence $\|P - P'\| = E_Q(|z - z'|)$. The second claimed inequality follows from $z + z' = 2$. $\square$

Next, recall that the Kakutani-Hellinger distance $\rho(P, P')$ is defined in IV.1.4, and the Hellinger integrals $H(\alpha; P, P')$ in IV.1.7 for $\alpha \in (0,1)$, and also IV.1.6 for $\alpha = 1/2$. We also write $H(P, P') = H(1/2; P, P')$.

4.4 **Proposition.** a) *We have*

4.5 $$2[1 - H(P,P')] \le \|P - P'\| \le \sqrt{8(1 - H(P,P'))}$$

4.6 $$\|P - P'\| \le 2\sqrt{1 - H(P,P')^2}.$$

b) *If* $\alpha \in (0,1)$ *there is a constant* $c_\alpha \in \mathbb{R}_+$ *such that*

4.7 $$2[1 - H(\alpha; P,P')] \le \|P - P'\| \le \sqrt{c_\alpha(1 - H(\alpha; P,P'))}.$$

In particular, in view of IV.1.5, we can write 4.5 and 4.6 as

4.8 $$2\rho^2(P,P') \le \|P - P'\| \le 2\sqrt{2}\rho(P,P'),$$

and therefore the Kakutani-Hellinger distance defines the same topology than the variation metric on the space of all probability measures on $(\Omega, \mathscr{F})$.

In the sequel we use mostly 4.5. However, 4.6 is sharper than the second inequality in 4.5, since $H(P, P') \le 1$ and $1 - x^2 \le 2(1 - x)$ for $0 \le x \le 1$.

Proof. (i) We consider the function $f_\alpha(x) = 1 - x^\alpha(2 - x)^{1-\alpha}$ for $x \in [0,2]$. Since $z + z' = 2$, the definition IV.1.7 of $H(\alpha;\, P,\, P')$ gives

4.9 $$1 - H(\alpha;\, P,\, P') = E_Q[f_\alpha(z)].$$

The inequality $f_\alpha(x) \le |x-1|$ for $x \in [0,2]$ being trivial, we deduce the left-hand inequality in 4.5 and 4.7 from 4.3b and 4.9.

(ii) We have seen in Lemma IV.2.19 that $\varphi_{1/2} \le \gamma_\alpha \varphi_\alpha$ for some constant $\gamma_\alpha > 0$, so the second inequality in 4.7 follows from the second inequality in 4.5 and from 4.9.

(iii) Since $1 - H(P,P')^2 \le 2[1 - H(P,P')]$, only 4.6 remains to be proved. Using 4.3b, we get

$$\begin{aligned}\tfrac{1}{2}\|P - P'\| &= E_Q(|1 - z|) \le \sqrt{E_Q(|1 - z|^2)} = \sqrt{1 - E_Q(z(2 - z))}\\ &= \sqrt{1 - E_Q(zz')} \le \sqrt{1 - (E_Q\sqrt{zz'})^2} = \sqrt{1 - H(P,P')^2}. \quad \square\end{aligned}$$

It is obvious that $P' \perp P \Leftrightarrow \|P - P'\| = 2$. Hence 4.5 and 4.7, plus the fact that $\|P - P'\| \le 2$ in all cases, yield:

$$P \perp P' \Leftrightarrow H(P,P') = 0 \Leftrightarrow \exists \alpha \in (0,1) \text{ with } H(\alpha; P,P') = 0,$$

and so we recover parts of Lemma IV.1.11b.

The following corollary also is obvious:

4.10 **Corollary.** a) *If* $(P^n)_{n \ge 1}$ *and* P *are probability measures on* $(\Omega, \mathscr{F})$, *we have as* $n \uparrow \infty$:

(i) $\|P^n - P\| \to 0 \Leftrightarrow H(P^n, P) \to 1$,
(ii) $\|P^n - P\| \to 2 \Leftrightarrow H(P^n, P) \to 0$.

b) *If for each* $n \in \mathbb{N}^*$, P^n *and* P'^n *are two probability measures on a space* $(\Omega^n, \mathscr{F}^n)$, *we have as* $n \uparrow \infty$:
(i) $\|P^n - P'^n\| \to 0 \Leftrightarrow H(P^n, P'^n) \to 1$,
(ii) $\|P^n - P'^n\| \to 2 \Leftrightarrow H(P^n, P'^n) \to 0$.

2. The variation metric has an interesting (and well known) statistical interpretation. P and P' can be viewed as a statistical hypothesis (H) (the null hypothesis) and its alternative (H'). A test is a measurable function $\varphi: \Omega \to [0,1]$ ($\varphi(\omega)$ is the probability with which we reject (H) if the observation is ω). Then the errors are:

$$\alpha(\varphi) = \text{error of the first kind} = \Pr(\text{reject } H \text{ if true}) = E_P(\varphi)$$

$$\beta(\varphi) = \text{error of the second kind} = \Pr(\text{accept } H \text{ if wrong}) = E_{P'}(1-\varphi),$$

and in case (H) and (H') play a symmetrical rôle we are interested in finding a test φ that minimizes the sum $\alpha(\varphi) + \beta(\varphi)$, and we set

4.11
$$\mathscr{E}_r(P, P') = \inf_{\text{all tests } \varphi} [\alpha(\varphi) + \beta(\varphi)].$$

Then, with the same notation as above, we have:

$$\begin{aligned}\mathscr{E}_r(P, P') &= \inf_{\varphi:\mathscr{F}\text{-measurable}, 0 \le \varphi \le 1} E_Q(z\varphi + z'(1-\varphi)) \\ &= 1 + \inf_{\varphi:\mathscr{F}\text{-measurable}, 0 \le \varphi \le 1} E_Q((z - z')\varphi).\end{aligned}$$

The infimum is achieved for $\varphi = 1_{\{z' < z\}}$, and since $E_Q(z - z') = 0$ we get

4.12
$$\mathscr{E}_r(P, P') = 1 - \tfrac{1}{2} E_Q(|z - z'|) = 1 - \tfrac{1}{2}\|P - P'\|.$$

Therefore, 4.4 yields:

4.13
$$\tfrac{1}{2} H(P, P')^2 \le 1 - \sqrt{1 - H(P, P')^2} \le \mathscr{E}_r(P, P') \le H(P, P').$$

4.14 **Example.** Suppose that $P^n = Q^{n\otimes}$ and $P'^n = Q'^{n\otimes}$, where Q and Q' are two probability measures on $(\Omega, \mathscr{F})$: we observe n i.i.d. random variables, with values in $(\Omega, \mathscr{F})$, with the distribution Q (resp. Q') under hypothesis (H) (resp. (H')). Then IV.1.74 implies $H(P^n, P'^n) = H(Q, Q')^n = e^{-\lambda n}$, where $\lambda = -\operatorname{Log} H(Q, Q') \ge \rho^2(Q, Q')$. Then 4.13 gives

$$\tfrac{1}{2} e^{-2\lambda n} \le \mathscr{E}_r(P^n, P'^n) \le e^{-\lambda n} \le e^{-n\rho^2(Q, Q')}. \qquad \square$$

§ 4b. Variation Metric and Hellinger Processes

1. We suppose that $(\Omega, \mathscr{F}, \mathbf{F})$ is a filtered space, endowed with two probability measures P and P'. We set $Q = (P + P')/2$, and we call z and z' the density processes of P and P' with respect to Q (see IV.1.13).

Let $h(\alpha)$ be any version of $h(\alpha; P, P')$. According to IV.4.25 we set

4.15 $$\mathscr{E}[-h(\alpha)]_t = \begin{cases} e^{-h(\alpha)_t} \prod_{s \le t} (1 - \Delta h(\alpha)_s) e^{\Delta h(\alpha)_s} & \text{if } h(\alpha)_t < \infty \\ 0 & \text{if } h(\alpha)_t = \infty. \end{cases}$$

This process is predictable if $h(\alpha)$ is so; it has the properties of IV.4.27 if $h(\alpha)$ is càdlàg with $\Delta h(\alpha) \le 1$ everywhere (since Lemma IV.4.27 describes "pathwise" properties only). The next result is a sort of generalization of IV.4.28.

4.16 **Proposition.** *Let $\alpha \in (0,1)$. There is a nonnegative process $N(\alpha)$ with the following properties:*

a) $z^\alpha z'^{1-\alpha} = N(\alpha)\mathscr{E}[-h(\alpha)]$, *whatever version $h(\alpha)$ of $h(\alpha; P, P')$ is chosen.*

b) $N(\alpha)^T$ *is a Q-local martingale if T is a stopping time such that $[\![0, T]\!] \subset \{\mathscr{E}(-h(\alpha)) > 0\}$, whatever version $h(\alpha)$ of $h(\alpha; P, P')$ is chosen.*

c) $N(\alpha)$ *is a Q-supermartingale.*

Proof. a) We use the notation Γ'', S of IV.1.15. All versions h of $h(\alpha; P, P')$ coincide on Γ'' and satisfy $h_t < \infty$ for $t \in \Gamma''$ and $\Delta h_t < 1$ for $t < S$ (IV.1.30), while $Y = z^\alpha z'^{1-\alpha}$ has $Y = 0$ on $[\![S, \infty[\![$. Hence if we set $W = \mathscr{E}(-h)$ and

$$N_t = \begin{cases} Y_t / W_t & \text{if } t < S \\ 0 & \text{if } t \ge S, \end{cases}$$

then N does not depend upon the version h of $h(\alpha; P, P')$, and (a) is met with $N(\alpha) = N$.

b) Let h be the Hellinger process $h(\alpha; P, P')$ in the strist sense, and $W = \mathscr{E}(-h)$, and let T be a stopping time such that $[\![0, T]\!] \subset \{W > 0\}$. h is predictable, càdlàg, with $\Delta h \le 1$, so IV.4.27 implies $h_t < \infty$ if $t \le T$ and $t < \infty$, and $\Delta h_t < 1$ for $t \le T$, and so IV.4.26 yields

$$W^T = 1 - W_- \cdot h^T.$$

If $V = 1/W^T$, Ito's formula for processes with finite variation yields

$$V = 1 - \frac{1}{W_-^2} \cdot W^T + \sum_{s \le \cdot \wedge T} \left(\frac{1}{W_s} - \frac{1}{W_{s-}} + \frac{\Delta W_s}{W_{s-}^2} \right)$$

4.17 $$= 1 - \frac{1}{W W_-} \cdot W^T = 1 + \frac{1}{W_-(1 - \Delta h)} \cdot h^T$$

(recall that $\Delta h < 1$ on $[\![0, T]\!]$), and then

4.18 $$V = V_- + \Delta V = V_- + \frac{\Delta h^T}{W_-(1 - \Delta h)} = \frac{V_-}{1 - \Delta h^T}.$$

On the other hand IV.1.18 gives $Y = M - Y_- \cdot h$ for some Q-martingale M. Hence Ito's formula for $N^T = Y^T V$, plus the predictability of V (because h, hence

W, are predictable) give:

$$\begin{aligned}
N^T &= N_0 + V\cdot Y^T + Y_-\cdot V = N_0 + V\cdot M^T - VY_-\cdot h^T + Y_-\cdot V \\
&= N_0 + V\cdot M^T - VY_-\cdot h^T + \frac{Y_-}{W_-(1-\Delta h)}\cdot h^T \qquad \text{(by 4.17)} \\
&= N_0 + V\cdot M^T \qquad \text{(by 4.18 and } V_- = 1/W_- \text{ on } [\![0,T]\!]).
\end{aligned}$$

Hence N^T is a Q-local martingale.

Let now h' be another version of $h(\alpha;P,P')$, and $W' = \mathscr{E}(-h')$. We shall show that if $W'_s > 0$ for all $s \le t$, then $W_t > 0$ (and so $W_s > 0$ for all $s \le t$ as well), and this will be enough to finish the proof of (b). To this end, assume $W'_s > 0$ for $s \le t$ and $W_t = 0$. Then 4.15 yields $h'_t < \infty$, and $h \le h'$ implies $h_t < \infty$; then we necessarily have $\Delta h_s = 1$ for some $s \le t$ (see IV.4.27): thus IV.1.30 yields $s \ge S$, while $h = 1_{\Gamma''}\cdot h$, so $s = S \in \Gamma''$. However, $h' = h$ on Γ'', so $\Delta h_s = 1$ as well and 4.15 clearly implies $W'_s = 0$, which is a contradiction.

c) Let again h be the Hellinger process $h(\alpha;P,P')$ in the strict sense, and $\tau = \inf(t\colon W_t = 0)$, and (τ_k) be a sequence of stopping times that Q-a.s. announces the predictable time τ (recall that $W = \mathscr{E}(-h)$ is predictable). We have $\{W > 0\} = \bigcup_n [\![0,\tau_n]\!]$, so each N^{τ_n} is a Q-local martingale, which is localized by a sequence $(T(n,k))_{k\ge 1}$ of stopping times. Since $N \ge 0$, and $N = 0$ on $[\![\tau,\infty[\![$, we get for $s \le t$:

$$\begin{aligned}
E_Q(N_t|\mathscr{F}_s) &= E_Q(N_t 1_{\{t<\tau\}}|\mathscr{F}_s) = E_Q\Big(\lim_{n,k\uparrow\infty} N_t^{\tau_n\wedge T(n,k)} 1_{\{t<\tau\}}|\mathscr{F}_s\Big) \\
&\le \liminf_{n,k\uparrow\infty} E_Q(N_t^{\tau_n\wedge T(n,k)} 1_{\{t<\tau\}}|\mathscr{F}_s) \qquad \text{(Fatou's lemma)} \\
&\le \liminf_{n,k\uparrow\infty} E_Q(N_t^{\tau_n\wedge T(n,k)} 1_{\{s<\tau\}}|\mathscr{F}_s) \\
&= \liminf_{n,k\uparrow\infty} N_s^{\tau_n\wedge T(n,k)} 1_{\{s<\tau\}} = N_s.
\end{aligned}$$

□

4.19 **Corollary.** *Let $0 < \alpha < \beta < 1$; let $h(\alpha)$ be any version of the Hellinger process $h(\alpha;P,P')$ such that $\Delta h(\alpha) \le 1$. Then for all stopping times T*

4.20 $$H(\beta;P_T,P'_T) \le H(\alpha;P_0,P'_0)^{(1-\beta)/(1-\alpha)}\{E_P(\mathscr{E}(-h(\alpha))_T^{(1-\beta)/(\beta-\alpha)})\}^{(\beta-\alpha)/(1-\alpha)}.$$

(as usual, P_T and P'_T denote the restrictions of P and P' to $\mathscr{F}_T$). The assumption $\Delta h(\alpha) \le 1$ (which is not a real restriction) is here in order that $\mathscr{E}(-h(\alpha)) \ge 0$; if it is not met, one should put $|\mathscr{E}(-h(\alpha))|$ in 4.20.

Proof. Let $p = \frac{1-\alpha}{1-\beta}$ and $q = \frac{1-\alpha}{\beta-\alpha}$, so $\frac{1}{p} + \frac{1}{q} = 1$. That $z^\beta z'^{1-\beta} = (z^\alpha z'^{1-\alpha})^{1/p} z^{1/q}$ is obvious, and 4.16 yields

$$z_T^\beta z_T'^{1-\beta} = N(\alpha)_T^{1/p}[\mathscr{E}(-h(\alpha))_T^{q/p} z_T]^{1/q}.$$

Therefore Hölder's inequality gives

$$\begin{aligned} E_Q(z_T^\beta z_T'^{1-\beta}) &\le [E_Q(N(\alpha)_T)]^{1/p}[E_Q(\mathscr{E}(-h(\alpha))_T^{q/p} z_T)]^{1/q} \\ &\le [E_Q(N(\alpha)_0)]^{1/p}[E_P(\mathscr{E}(-h(\alpha))_T^{q/p})]^{1/q} \end{aligned}$$

(for the last inequality, we use 4.16c and IV.1.14). Since $N(\alpha)_0 = z_0^\alpha z_0'^{1-\alpha}$, the claim follows from Definition IV.1.7. □

2. Now we are in a position to state and prove our main estimates for the variation $\|P - P'\|$.

4.21 **Theorem.** *Let h be a version of $h(\frac{1}{2}; P, P')$ such that $\Delta h \le 1$ identically. Then for all stopping times T and all $\varepsilon > 0$:*

4.22 $$2[1 - \sqrt{H(P_0, P_0')E_P(\exp -h_T)}] \le \|P_T - P_T'\| \le \|P_0 - P_0'\| + 4\sqrt{E_P(h_T)}$$

4.23 $$\|P_T - P_T'\| \le \tfrac{3}{2}\|P_0 - P_0'\| + 3\sqrt{2\varepsilon} + 2P(h_T \ge \varepsilon).$$

4.24 **Remark.** There are many other possible estimates. For instance, Valkeila and Vostrikova [239] have proved:

4.25 $$\begin{cases} 2\rho^2(P_T, P_T') \le 4\varepsilon + 2P(|1 - \sqrt{z_0}| \ge \varepsilon) + 2\sqrt{P(h_T \ge \varepsilon/2)} \\ 2P(|1 - \sqrt{z_0}| \ge \varepsilon) + P(h_T \ge \varepsilon/2) \le \varepsilon^2 + (1 + 2/\varepsilon^2)\sqrt{2}\rho(P_T, P_T'), \end{cases}$$

which, in virtue of 4.8, gives other estimates for $\|P_T - P_T'\|$. There also exist other (apparently less useful) estimates in terms of the measure $Q = (P + P')/2$. □

We begin with two lemmas:

4.26 **Lemma.** *Let $\langle z, z\rangle$ denote the predictable variation of the Q-bounded martingale z, relatively to the measure Q. Then*

a) $\langle z, z\rangle \le 2h$ (h is like in 4.21),

b) $\langle z, z\rangle \le 4z_- \cdot h$.

Proof. We can of course assume that h is the Hellinger process $h(\frac{1}{2}; P, P')$ in the strict sense. Then IV.1.36 gives

$$\begin{aligned} h &= \frac{1}{8}\left(\frac{1}{z_-} + \frac{1}{z'_-}\right)^2 \cdot \langle z^c, z^c\rangle + \frac{1}{2}(\sqrt{1 + x/z_-} - \sqrt{1 - x/z'_-})^2 * \nu^z \\ &= \frac{1}{2}\left(\frac{1}{z_- z'_-}\right)^2 \cdot \langle z^c, z^c\rangle + \frac{1}{z_- z'_-}\frac{2x^2}{(\sqrt{z'_-(z_- + x)} + \sqrt{z_-(z'_- - x)})^2} * \nu^z \end{aligned}$$

by an elementary computation, using $z + z' = 2$. This also implies that $z_- z'_- \le 1$, and it is easy to check that $\sqrt{z'_-(z_- + x)} + \sqrt{z_-(z'_- - x)} \le 2$ for $-z_- \le x \le z'_-$. Hence we deduce that if

$$k = \langle z^c, z^c \rangle + x^2 * \nu^z,$$

then $k \leq 2h$, and also $k \leq 4z_- \cdot h$ (because $z'_- \leq 2$). On the other hand, we can use II.2.29 for $X = z$ to compute $\langle z, z \rangle$: then $\langle z, z \rangle$ is given by II.2.31 with ν^z, and in particular $\langle z, z \rangle \leq k$. □

4.27 **Lemma.** *With the notation of 4.21 and 4.26, we have*

$$\|P_T - P'_T\| \leq \|P_0 - P'_0\| + 2\sqrt{E_Q(\langle z, z \rangle_T)}.$$

Proof. The following is straightforward (recall $z + z' = 2$):

$$|z_T - z'_T| \leq |z_0 - z'_0| + |(z_T - z_0) - (z'_T - z'_0)| \leq |z_0 - z'_0| + 2|z_T - z_0|.$$

Then, by 4.3,

$$\begin{aligned}\|P_T - P'_T\| &= E_Q(|z_T - z'_T|) \leq \|P_0 - P'_0\| + 2E_Q(|z_T - z_0|)\\ &\leq \|P_0 - P'_0\| + 2\sqrt{E_Q((z_T - z_0)^2)}\\ &\leq \|P_0 - P'_0\| + 2\sqrt{E_Q(\langle z, z \rangle_T)}.\end{aligned}$$

□

Proof of Theorem 4.21. a) Firstly, consider the second inequality in 4.22. We can always assume that h is the Hellinger process in the strict sense (this diminishes the right-hand side). Hence, applying the predictability of h and I.3.12 (as already done several times in the same context), we get $E_Q(z_- \cdot h_T) = E_Q(z_T h_T) = E_P(h_T)$. Thus the claimed inequality readily follows from 4.26b and 4.27.

b) In order to obtain the first inequality in 4.22, we apply 4.20 to $\alpha = 1/2$ and $\beta = 3/4$, thus getting

$$H(\tfrac{3}{4}; P_T, P'_T) \leq H(P_0, P'_0)^{1/2} \{E_P(\mathscr{E}(-h)_T)\}^{1/2}.$$

Now, the first inequality in 4.7 and the trivial majoration $\mathscr{E}(-h) \leq e^{-h}$ (because $\Delta h \leq 1$) readily give the result.

c) In order to obtain 4.23, we first prove an auxiliary result. Let S be another stopping time with $S \leq T$. Then

$$\begin{aligned}\|P_T - P'_T\| &= E_Q(|z_T - z'_T|) = E_Q(|z_T - z'_T| 1_{\{T=S\}}) + E_Q(|z_T - z'_T| 1_{\{S<T\}})\\ &\leq E_Q(|z_S - z'_S|) + 2Q(S < T) = \|P_S - P'_S\| + 2Q(S < T),\end{aligned}$$

because $|z - z'| \leq 2$. Moreover, $\{S < T\} \in \mathscr{F}_S$, hence $|P(S < T) - P'(S < T)| \leq \frac{1}{2}\|P_S - P'_S\|$, and

$$\begin{aligned}2Q(S < T) = P(S < T) + P'(S < T) &\leq |P(S < T) - P'(S < T)| + 2P(S < T)\\ &\leq \tfrac{1}{2}\|P_S - P'_S\| + 2P(S < T).\end{aligned}$$

Thus

4.28 $$\|P_T - P'_T\| \leq \tfrac{3}{2}\|P_S - P'_S\| + 2P(S < T).$$

Now we consider the predictable stopping time $S = \inf(t: \langle z, z\rangle_t \geq 2\varepsilon)$, which is Q-a. s. announced by a sequence (S_n) of stopping times. Let $t \in \mathbb{R}_+$. Applying 4.28 to $S_n \wedge t \wedge T$ and $t \wedge T$, and 4.27, and $\langle z, z\rangle_{S_n} \leq 2\varepsilon$, we obtain

4.29
$$\begin{aligned}\|P_{T\wedge t} - P'_{T\wedge t}\| &\leq \tfrac{3}{2}\|P_{T\wedge S_n\wedge t} - P'_{T\wedge S_n\wedge t}\| + 2P(S_n \wedge t < T \wedge t)\\ &\leq \tfrac{3}{2}\|P_0 - P'_0\| + 3\sqrt{2\varepsilon} + 2P(S_n \wedge t < T \wedge t).\end{aligned}$$

Moreover $P(S_n \wedge t < T \wedge t) \downarrow P(S \leq T \wedge t)$ as $n \uparrow \infty$, and if $S \leq T \wedge t$ we have $\langle z, z\rangle_T \geq 2\varepsilon$, so $h_T \geq \varepsilon$ by 4.26a. Thus 4.29 yields

$$\|P_{T\wedge t} - P'_{T\wedge t}\| \leq \tfrac{3}{2}\|P_0 - P'_0\| + 3\sqrt{2\varepsilon} + 2P(h_T \geq \varepsilon).$$

Finally, $\|P_{T\wedge t} - P'_{T\wedge t}\| = 2E_Q(|1 - z_{T\wedge t}|)$ by 4.3, and this goes to $2E_Q(|1 - z_T|) = \|P_T - P'_T\|$ as $t \uparrow \infty$. Therefore 4.23 follows. □

3. In the sequel, we use the following notation: if for each $n \in \mathbb{N}^*$, P^n is a measure on $(\Omega^n, \mathscr{F}^n)$, which is endowed with a random variable V^n taking its values in the metric space (E, d), and if $v \in E$, we write:

4.30 $$V^n \xrightarrow{P^n} v, \quad \text{if } P^n(d(V^n, v) > \varepsilon) \to 0 \quad \text{for all } \varepsilon > 0.$$

4.31 **Theorem.** *Let $(P^n)_{n\geq 1}$ and P be probability measures on $(\Omega, \mathscr{F})$, and for each $n \in \mathbb{N}^*$ let h^n be a version of $h(\frac{1}{2}; P, P^n)$ that satisfies $\Delta h^n \leq 1$ identically. Then for all stopping times T we have, as $n \uparrow \infty$:*

(i) $\|P^n_T - P_T\| \to 0 \Leftrightarrow \|P^n_0 - P_0\| \to 0$ *and* $h^n_T \xrightarrow{P} 0$;
(ii) $\|P^n_T - P_T\| \to 0 \Leftrightarrow \|P^n_0 - P_0\| \to 0$ *and* $h^n_T \xrightarrow{P^n} 0$;
(iii) $h^n_T \xrightarrow{P} +\infty$ (i.e. $P(h^n_T < N) \to 0$ for all $N < \infty$) $\Rightarrow \|P^n_T - P_T\| \to 2$;
(iv) $h^n_T \xrightarrow{P^n} +\infty \Rightarrow \|P^n_T - P_T\| \to 2$.

Proof. (i) Assume that $\|P^n_T - P'_T\| \to 0$. Since $H(P^n_0, P_0) \leq 1$ and $E_P(\exp - h^n_T) \leq 1$, the first inequality in 4.22 implies that $H(P^n_0, P_0) \to 1$, hence $\|P^n_0 - P_0\| \to 0$ by 4.10a, and also that $E_P(\exp - h^n_T) \to 1$, which in turn implies $\exp - h^n_T \xrightarrow{P} 1$ and so $h^n_T \xrightarrow{P} 0$.

Conversely, if $\|P^n_0 - P_0\| \to 0$ and $h^n_T \xrightarrow{P} 0$, it is immediate to deduce $\|P^n_T - P_T\| \to 0$ from 4.23.

(iii) Assume that $h^n_T \xrightarrow{P} +\infty$, hence $E_P(\exp - h^n_T) \to 0$. Since $\|P^n_T - P_T\| \leq 2$, it follows from the first inequality in 4.22 that $\|P^n_T - P_T\| \to 2$.

Finally, (ii) and (iv) are proved similarly to (i) and (iii). □

In particular, if $P^n = P'$ for all n, (iv) reads as: $P(h_T = \infty) = 1 \Rightarrow \|P_T - P'_T\| = 2 \Leftrightarrow P_T \perp P'_T$. Hence we obtain again IV.2.6c (with a more complicated proof, though!)

More interesting, perhaps, is the case where we compare two sequences of measures (P^n) and (P'^n). So for each $n \in \mathbb{N}^*$ we consider two probability measures P^n and P'^n on some filtered space $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. We call h^n a version of $h(1/2; P^n, P'^n)$ which satisfies $\Delta h^n \leq 1$ identically. Then the same proof than above yields:

4.32 Theorem. *In the setting just described, and if for each n, T^n is a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, one has as $n \uparrow \infty$:*

(i) $\|P^n_{T^n} - P'^n_{T^n}\| \to 0 \Leftrightarrow \|P^n_0 - P'^n_0\| \to 0$ *and* $h^n_{T^n} \xrightarrow{P^n} 0$;

(ii) $h^n_{T^n} \xrightarrow{P^n} +\infty \Rightarrow \|P^n_{T^n} - P'^n_{T^n}\| \to 2$.

§ 4c. Examples: Point Processes and Multivariate Point Processes

1. We suppose now that the space Ω is endowed with a *point process X* and that $\mathbf{F}$ is the smallest filtration to which X is adapted and that $\mathscr{F} = \mathscr{F}_{\infty-}$.

We consider two measures P and P' on $(\Omega, \mathscr{F})$ and we call A and A' the compensators of X under P and P'. Recall that $\operatorname{Var}(A - A')$ is the variation process of $A - A'$ (see § I.3a), i.e. $\operatorname{Var}(A - A')_t = \|d(A^t)_s - d(A'^t)_s\|$, with notation 4.1.

4.33 Theorem. *For all $\varepsilon > 0$ and all stopping times T we have*

4.34 $$2\left[1 - \left\{E_P\left(\exp - \frac{\operatorname{Var}(A - A')_T^2}{8A_T + 4\operatorname{Var}(A - A')_T}\right)\right\}^{1/2}\right] \leq \|P_T - P'_T\| \leq 4\sqrt{E_P(\operatorname{Var}(A - A')_T)}$$

4.35 $$\|P_T - P'_T\| \leq 3\sqrt{2\varepsilon} + 2P(\operatorname{Var}(A - A')_T > \varepsilon).$$

Proof. Let $\bar{A} = A + A'$, and g, g' be two predictable nonnegative processes such that $A = g \cdot \bar{A}$ and $A' = g' \cdot \bar{A}$ $(P + P')$-a. s.: see IV.4.1. Then IV.4.3 gives a version h of $h(1/2; P, P')$ which meets $\Delta h \leq 1$, namely

4.36 $$h = \frac{1}{2}(\sqrt{g} - \sqrt{g'})^2 \cdot \bar{A} + \frac{1}{2} \sum_{s \leq \cdot} (\sqrt{1 - \Delta A_s} - \sqrt{1 - \Delta A'_s})^2.$$

Now, $(\sqrt{x} - \sqrt{y})^2 \leq |x - y|$ for $x, y \geq 0$. Hence

$$h \leq \frac{1}{2}|g - g'| \cdot \bar{A} + \frac{1}{2} \sum_{s \leq \cdot} |\Delta A_s - \Delta A'_s|.$$

It is obvious that $|g - g'| \cdot \bar{A} = \operatorname{Var}(A - A')$ (same proof than 4.3b), and that $\sum_{s \leq t} |\Delta A_s - \Delta A'_s| \leq \operatorname{Var}(A - A')_t$, so

4.37 $$h \leq \operatorname{Var}(A - A').$$

On the other hand, we have

$$\begin{aligned}\operatorname{Var}(A - A') &= |g - g'| \cdot \bar{A} = |\sqrt{g} - \sqrt{g'}|(\sqrt{g} + \sqrt{g'}) \cdot \bar{A} \\ &\leq \sqrt{\tfrac{1}{2}(\sqrt{g} - \sqrt{g'})^2 \cdot \bar{A}}\sqrt{2(\sqrt{g} + \sqrt{g'})^2 \cdot \bar{A}} \\ &\leq \sqrt{\tfrac{1}{2}(\sqrt{g} - \sqrt{g'})^2 \cdot \bar{A}}\sqrt{4(g + g') \cdot \bar{A}}.\end{aligned}$$

Moreover $(g + g')\cdot \bar{A} \leq (2g + |g' - g|)\cdot \bar{A} = 2A + \mathrm{Var}(A - A')$, hence 4.36 implies

4.38 $$\mathrm{Var}(A - A') \leq \sqrt{h}\sqrt{8A + 4\,\mathrm{Var}(A - A')}.$$

Finally, we have seen (in the proof of IV.4.6 for instance) that $P_0 = P'_0$, so $\|P_0 - P'_0\| = 0$ and $H(P_0, P'_0) = 1$. Hence if we plug 4.37 and 4.38 into 4.22 and 4.23, we obtain 4.34 and 4.35. □

4.39 **Corollary.** *Let P, $(P^n)_{n\geq 1}$ be probability measures on $(\Omega, \mathscr{F})$, and call A, A^n the corresponding compensators of X. Then for all stopping times T one has as $n \uparrow \infty$:*

a) $\mathrm{Var}(A^n - A)_T \xrightarrow{P} 0 \Rightarrow \|P^n_T - P_T\| \to 0$.

b) $\|P^n_T - P_T\| \to 0$ *and* $P(A_T < \infty) = 1 \Rightarrow \mathrm{Var}(A^n - A)_T \xrightarrow{P} 0$.

Proof. (a) immediately follows from 4.35. Conversely, if $\|P^n_T - P_T\| \to 0$, 4.34 gives that

$$E_P(\exp - Y^n) \to 1, \quad \text{where } Y^n = \mathrm{Var}(A^n - A)^2_T/[8A_T + 4\,\mathrm{Var}(A^n - A)_T],$$

which in turn implies $Y^n \xrightarrow{P} 0$. If $A_T < \infty$ P-a.s., the latter obviously yields $\mathrm{Var}(A^n - A)_T \xrightarrow{P} 0$. □

Now we compare two sequences of measures. For each $n \in \mathbb{N}^*$, we consider a point process X^n defined on a space Ω^n, and $\mathbf{F}^n$ the smallest filtration to which X^n is adapted, and $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$, and P^n and P'^n two probability measures on $(\Omega^n, \mathscr{F}^n)$, and A^n and A'^n the corresponding compensators of X^n.

4.40 **Corollary.** *For each $n \in \mathbb{N}^*$ let T^n be a stopping time on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. Then, as $n \uparrow \infty$:*

a) $\mathrm{Var}(A^n - A'^n)_{T^n} \xrightarrow{P^n} 0 \Rightarrow \|P^n_{T^n} - P'^n_{T^n}\| \to 0$;

b) *if* $\|P^n_{T^n} - P'^n_{T^n}\| \to 0$ *and if the sequence* $(A^n_{T^n}|P^n)$ *is* $\mathbb{R}$*-tight* (see 1.2), *then* $\mathrm{Var}(A^n - A'^n)_{T^n} \xrightarrow{P^n} 0$.

Proof. It is absolutely similar to the proof of 4.39 (replace the pair A, A^n by the pair A^n, A'^n). □

2. *Application to empirical processes.* An interesting and very simple consequence of the previous results concerns the convergence of the suitably normalized empirical processes to a Poisson point process, when the size of the experiment increases. We consider the setting of § II.3c.3: let $(Z_n)_{n\geq 1}$ be a sequence of i.i.d. random variables taking values in $(0, \infty)$, with a distribution G *admitting a density* g. The empirical process of size n is

$$X^n_t = \frac{1}{n} \sum_{1\leq i\leq n} 1_{\{Z_i \leq t\}}.$$

We call Ω the canonical space of all point processes, with the canonical process N and the canonical filtration $\mathbf{F}$ and $\mathscr{F} = \mathscr{F}_{\infty-}$.

4.41 **Theorem.** *Let P^n be the distribution of the point process $\bar{Y}_t^n = nX_{t/n}^n$ and let P be the distribution of the Poisson process with intensity function $g(0)t$. If P_t^n and P_t denote the restrictions of P^n and P to the σ-field $\mathscr{F}_t$, then*

$$4.42 \quad \|P_t^n - P_t\| \le 4\left\{\int_0^t \left[\frac{G((0,s/n])g(s/n)}{G([s/n,\infty))} + \left|\frac{g(s/n)}{G([s/n,\infty))} - g(0)\right|\right] ds\right\}^{1/2},$$

and in particular if g is continuous from the right at 0, we have $\|P_t^n - P_t\| \to 0$ for all $t \in \mathbb{R}_+$ as $n \uparrow \infty$.

Proof. In virtue of II.3.34, the compensator of N under P^n is

$$A_t^n = \int_0^t \left(1 - \frac{N_s}{n}\right) \frac{g(s/n)}{G([s/n,\infty))} ds,$$

and $A_t = g(0)t$ is the compensator of N under P. Moreover, $E_{P^n}(N_s/n) = E(X_{s/n}^n) = G((0,s/n])$. Therefore

$$\begin{aligned} E_{P^n}(\mathrm{Var}(A^n - A)_t) &= E_{P^n}\left(\int_0^t \left|\left(1 - \frac{N_s}{n}\right)\frac{g(s/n)}{G([s/n,\infty))} - g(0)\right| ds\right) \\ &\le E_{P^n}\left(\int_0^t \left[\frac{N_s}{n}\frac{g(s/n)}{G([s/n,\infty))} + \left|\frac{g(s/n)}{G([s/n,\infty))} - g(0)\right|\right] ds\right) \\ &= \int_0^t \left[E_{P^n}(N_s/n)\frac{g(s/n)}{G([s/n,\infty))} + \left|\frac{g(s/n)}{G([s/n,\infty))} - g(0)\right|\right] ds \\ &= \int_0^t \left(\frac{G((0,s/n])g(s/n)}{G([s/n,\infty))} + \left|\frac{g(s/n)}{G([s/n,\infty))} - g(0)\right|\right) ds \end{aligned}$$

and 4.42 follows from 4.34.

Finally if g is continuous at 0, the Lebesgue dominated convergence theorem immediately yields that the right-hand side of 4.42 tends to 0 as $n \uparrow \infty$, hence the claim. □

3. Multivariate point processes. Here we suppose that the space Ω is endowed with an *E-valued multivariate point process* μ (III.1.23), and $\mathbf{F}$ is the smallest filtration for which μ is optional, and $\mathscr{F} = \mathscr{F}_{\infty-}$.

We consider two measures P and P' on $(\Omega, \mathscr{F})$ and we call ν and ν' the compensators of μ under P and P'. We denote by $\mathrm{Var}(\nu - \nu')_t$ the variation distance between the two measures $\nu_{|[0,t]\times E}$ and $\nu'_{|[0,t]\times E}$, i.e.:

$$\mathrm{Var}(v - v')_t(\omega) = \sup\left(\left|\int \varphi(s,x)(v - v')(\omega; ds \times dx)\right| : \varphi \text{ an } \mathscr{R} \otimes \mathscr{E}\text{-measurable function on } \mathbb{R}_+ \times E \text{ such that } |\varphi| \le 1 \text{ and } \varphi(s,x) = 0 \text{ for } s > t\right).$$

4.43 **Theorem.** *For all* $\varepsilon > 0$ *and all stopping times* T *we have*

4.44
$$2\left[1 - \left\{E_P\left(\exp - \frac{\mathrm{Var}(v - v')_T^2}{8v([0,T] \times E) + 4\,\mathrm{Var}(v - v')_T}\right)\right\}^{1/2}\right] \le \|P_T - P'_T\| \le 4\sqrt{E_P(\mathrm{Var}(v - v')_T)}$$

4.45
$$\|P_T - P'_T\| \le 3\sqrt{2\varepsilon} + 2P(\mathrm{Var}(v - v')_T > \varepsilon).$$

Proof. Let $\lambda = \dfrac{v + v'}{2}$ and let U, U' be two nonnegative predictable functions such that $v = U \cdot \lambda$ and $v' = U' \cdot \lambda$ $(P + P')$-a.s. Set also $a_t = v(\{t\} \times E)$ and $a'_t = v'(\{t\} \times E)$. Then IV.4.12 gives a version h of $h(1/2; P, P')$ which meets $\Delta h \le 1$, namely

$$h = \frac{1}{2}(\sqrt{U} - \sqrt{U'})^2 * \lambda + \frac{1}{2} \sum_{s \le \cdot} (\sqrt{1 - a_s} - \sqrt{1 - a'_s})^2.$$

As in 4.33, we obtain

$$h \le \frac{1}{2}|U - U'| * \lambda + \frac{1}{2} \sum_{s \le \cdot} |a_s - a'_s|$$

and $|U - U'| * \lambda_t = \mathrm{Var}(v - v')_t$ (as in 4.3b) and $\sum_{s \le t} |a_s - a'_s| \le \mathrm{Var}(v - v')_t$. Then $h \le \mathrm{Var}(v - v')_t$, and from this point the rest of the proof goes exactly as in the proof of 4.33. □

As for 4.39 and 4.40, we can write the following corollaries:

4.46 **Corollary.** *Let* P, $(P^n)_{n \ge 1}$ *be probability measures on* $(\Omega, \mathscr{F})$, *and call* v, v^n *the corresponding compensators of* μ. *Then for all stopping times* T *one has* as $n \uparrow \infty$:

a) $\mathrm{Var}(v^n - v)_T \xrightarrow{P} 0 \Rightarrow \|P^n_T - P_T\| \to 0$;

b) $\|P^n_T - P_T\| \to 0$ *and* $P(v([0,T] \times E) < \infty) = 1 \Rightarrow \mathrm{Var}(v^n - v)_T \xrightarrow{P} 0$.

4.47 **Corollary.** *For each* $n \in \mathbb{N}^*$ *let* Ω^n *be a space equipped with an* E^n*-valued multivariate point process* μ^n, *and* $\mathbf{F}^n$ *be the filtration generated by* μ^n, *and* $\mathscr{F}^n = \mathscr{F}^n_{\infty-}$. *Let* P^n, P'^n *be two probability measures on* $(\Omega^n, \mathscr{F}^n)$, *and call* v^n *and* v'^n *the corresponding compensators of* μ^n. *Finally, let* T^n *be a stopping time on* $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. *Then, as* $n \uparrow \infty$:

a) $\mathrm{Var}(v^n - v'^n)_{T^n} \xrightarrow{P^n} 0 \Rightarrow \|P^n_{T^n} - P'^n_{T^n}\| \to 0$;

b) *if* $\|P^n_{T^n} - P'^n_{T^n}\| \to 0$ *and if the sequence* $(v^n([0, T^n] \times E^n)|P^n)$ *is* $\mathbb{R}$*-tight, then* $\mathrm{Var}(v^n - v'^n)_{T^n} \xrightarrow{P^n} 0$.

§ 4d. Example: Generalized Diffusion Processes

Here we consider the canonical setting $(\Omega, \mathscr{F}, \mathbf{F})$ with the canonical 1-dimensional process X (as in § 3b). We suppose that P^n and P'^n are measures on $(\Omega, \mathscr{F})$ such that 3.5 holds:

4.48
$$\begin{cases} X_t = x^n + \int_0^t \beta^n_s\, ds + W^n_t, & W^n \text{ is a } P^n\text{-standard Wiener process} \\ X_t = x^n + \int_0^t \beta'^n_s\, ds + W'^n_t, & W'^n \text{ is a } P'^n\text{-standard Wiener process,} \end{cases}$$

and we consider the same processes as in 3.6:

$$K^n_t = \int_0^t (\beta^n_s)^2\, ds, \quad K'^n_t = \int_0^t (\beta'^n_s)^2\, ds, \quad \tilde{K}^n_t = \int_0^t (\beta^n_s - \beta'^n_s)^2\, ds.$$

4.49 **Theorem.** *Assume that for each* $n \in \mathbb{N}^*$ *the processes* K^n *and* K'^n (and so $\tilde{K}^n$ as well) *do not jump to infinity, and that* $P^n(K^n_t < \infty) = P'^n(K'^n_t < \infty) = 1$ *for all* $t \in \mathbb{R}_+$, *and let* T^n *be a stopping time.*

a) $\|P^n_{T^n} - P'^n_{T^n}\| \to 0 \Leftrightarrow \tilde{K}^n_{T^n} \xrightarrow{P^n} 0$.

b) $\tilde{K}^n_{T^n} \xrightarrow{P^n} \infty \Rightarrow \|P^n_{T^n} - P'^n_{T^n}\| \to 2$.

Proof. In the course of the proof of IV.4.23c we have shown that $P^n_0 = P'^n_0$, so $\|P^n_0 - P'^n_0\| = 0$. Then the claims readily follow from 4.32 and from the fact (see IV.4.23c) that $\tilde{K}^n/8$ is a version of $h(1/2; P^n, P'^n)$. □

As an example, we consider the case of "strict" diffusion processes, in the sense of § III.2c, and even the homogeneous case!

More precisely, let b^n, b be measurable and locally bounded functions on $\mathbb{R}$. Suppose that

4.50
$$\beta^n_s = b^n(X_s), \quad \beta'^n_s = b(X_s), \quad x^n = x \in \mathbb{R}$$

and all P'^n are equal to the same measure P (so, instead of comparing two sequences of measures, we give the "convergence" version of the result).

4.51 **Corollary.** *Under the previous assumptions, and if moreover the sequence* (b^n) *uniformly converges to* b *over all compact subsets of* $\mathbb{R}$, *then* $\|P^n_t - P_t\| \to 0$ *for all* $t \in \mathbb{R}_+$.

Proof. The processes K^n and $K'^n = K$ and $\tilde{K}^n$ are finite-valued, so all the assumptions of 4.49 are in force. The local uniform convergence of $b^n \to b$ readily gives

$$\int_0^t (b^n(X_s) - b(X_s))^2 \, ds \xrightarrow{P} 0,$$

and then 4.49a gives the result. □

We thus see that, despite the strength of the convergence in variation, one happens to obtain such a convergence (at least, locally in time) for diffusion processes, under "reasonably" weak assumptions. We shall obtain weak convergence under similar assumptions in Chapter IX, with a major difference, though: here, the "diffusion coefficients" are the same for all processes, while in Chapter IX they are allowed to vary as is the drift coefficient here.

Chapter VI. Skorokhod Topology and Convergence of Processes

In this chapter, we lay down the last cornerstone that is needed to derive functional limit theorems for processes. Namely, we consider the space $\mathbb{D}(\mathbb{R}^d)$ of all càdlàg functions: $\mathbb{R}_+ \to \mathbb{R}^d$; we need to provide this space with a topology, such that: (1) the space is Polish (so we can apply classical limit theorems on Polish spaces); (2) the Borel σ-field is exactly the σ-field generated by all evaluation maps (because the "law" of a process is precisely a measure on this σ-field).

Skorokhod introduced a topology (which he called J_1-topology) for this purpose. We will recall the definition and the main properties, essentially following the book [12], with a difference, though: Skorokhod and Billingsley only speak about spaces of functions defined on a finite interval, while for us it is more natural to consider functions defined on $\mathbb{R}_+$; for this (simple) extension, we follow Stone [230] and Lindvall [155].

As a matter of fact, a commonly shared feeling is that constructing Skorokhod topology and deriving tightness criteria are rather tedious. We agree... However, it seems that no other, simpler topology on $\mathbb{D}(\mathbb{R}^d)$ can serve the same purposes, despite several attempts in this direction: the (locally) uniform topology is too strong (see however the book [198] by Pollard); the weaker topology of convergence in measure (relative to Lebesgue measure, for example) has recently been proposed by Meyer and Zheng [185], but it lacks simple criteria for characterizing compact sets.

So, if we felt obliged to write the present chapter, we would not like to inflict the whole story on our reader. We advise her or him to read only the following:

—§ 1a for some notation and § 1b for the main properties of the Skorokhod topology (but not § 1c, which contains the proofs); section 2 should be considered as a reference section;

—Section 3 for some more notation, and reminder about weak convergence of probability measures;

—Theorem 4.18, and perhaps § 5a for a reasonably general tightness criterion;

—Statements 6.1, 6.6 and 6.7.

1. The Skorokhod Topology

§ 1a. Introduction and Notation

The first two sections of this chapter are purely non-probabilistic.

1.1 **Definitions.** a) We denote by $\mathbb{D}(\mathbb{R}^d)$ the space of all *càdlàg functions*: $\mathbb{R}_+ \to \mathbb{R}^d$ (it is called the *Skorokhod space*).

b) If $\alpha \in \mathbb{D}(\mathbb{R}^d)$ we denote by $\alpha(t)$ the value of α at time t, and by $\alpha(t-)$ its left-hand limit at time t (with $\alpha(0-) = \alpha(0)$ by convention), and $\Delta\alpha(t) = \alpha(t) - \alpha(t-)$.

c) $\mathscr{D}_t^0(\mathbb{R}^d)$ denotes the σ-field generated by all maps: $\alpha \rightsquigarrow \alpha(s)$ for $s \leq t$, and $\mathscr{D}(\mathbb{R}^d) = \bigvee_{t \geq 0} \mathscr{D}_t^0(\mathbb{R}^d)$, and $\mathscr{D}_t(\mathbb{R}^d) = \bigcap_{s>t} \mathscr{D}_s^0(\mathbb{R}^d)$: hence $\mathbf{D}(\mathbb{R}^d) = (\mathscr{D}_t(\mathbb{R}^d))_{t \geq 0}$ is a filtration. □

We wish to endow $\mathbb{D}(\mathbb{R}^d)$ with a topology for which it is a Polish space (a complete separable metric space), such that $\mathscr{D}(\mathbb{R}^d)$ be its Borel σ-field (the σ-field generated by all open subsets). At first glance, a candidate would be the *local uniform topology* associated with the metric

1.2
$$\begin{cases} \delta_{lu}(\alpha, \beta) = \sum_{N \in \mathbb{N}^*} 2^{-N}(1 \wedge \|\alpha - \beta\|_N), \\ \text{where } \|\alpha\|_\theta = \sup_{s \leq \theta} |\alpha(s)| \end{cases}$$

(the subscript "lu" stands for "local uniform").

The space $\mathbb{D}(\mathbb{R}^d)$ is obviously a complete space under δ_{lu}. However, it fails to be *separable*: for example, the functions α_s defined by

$$\alpha_s(t) = 1_{[s,\infty)}(t)$$

for all $s \in [0, 1)$ are uncountable many, while $\delta_{lu}(\alpha_s, \alpha_{s'}) = \frac{1}{2}$ for $s \neq s'$.

However, a particularly important subspace of $\mathbb{D}(\mathbb{R}^d)$ is nicely topologized by the metric δ_{lu}, namely:

1.3 $\mathbb{C}(\mathbb{R}^d) =$ the space of all *continuous* functions: $\mathbb{R}_+ \to \mathbb{R}^d$.

It is immediate that $\mathbb{C}(\mathbb{R}^d)$ is Polish for the local uniform topology and that the corresponding σ-field is the trace of $\mathscr{D}(\mathbb{R}^d)$ on $\mathbb{C}(\mathbb{R}^d)$. Furthermore one knows how to characterize all compact subsets; to this end, with each function α: $\mathbb{R}_+ \to \mathbb{R}^d$ we associate:

1.4
$$\begin{cases} w(\alpha; I) = \sup_{s,t \in I} |\alpha(s) - \alpha(t)|, \quad I \text{ an interval of } \mathbb{R}_+ \\ w_N(\alpha, \theta) = \sup\{w(\alpha; [t, t+\theta]) : 0 \leq t \leq t+\theta \leq N\}, \quad \theta > 0, N \in \mathbb{N}^*. \end{cases}$$

Then, the Ascoli-Arzela's Theorem reads:

1.5 **Theorem.** *A subset A of* $\mathbb{C}(\mathbb{R}^d)$ *is relatively compact* (i.e., it has a compact closure) *for the local uniform topology if and only if*

(i) $\sup_{\alpha\in A} |\alpha(0)| < \infty$;

(ii) *for all* $N\in\mathbb{N}^*$, $\lim_{\theta\downarrow 0} \sup_{\alpha\in A} w_N(\alpha,\theta) = 0$.

Moreover, it is worth noticing that a function $\alpha\colon \mathbb{R}_+ \to \mathbb{R}^d$ belongs to $\mathbb{C}(\mathbb{R}^d)$ if and only if

1.6 $$\lim_{\theta\downarrow 0} w_N(\alpha,\theta) = 0 \quad \text{for all } N\in\mathbb{N}^*.$$

As for $\mathbb{D}(\mathbb{R}^d)$, it will be a Polish space under the Skorokhod topology, to be introduced in § 1b below. Before we introduce some notation. Firstly we set for all $\alpha\in\mathbb{D}(\mathbb{R}^d)$:

1.7 $$\begin{cases} J(\alpha) = \{t>0\colon \Delta\alpha(t)\neq 0\} & \text{(set of discontinuity points of } \alpha)\\ U(\alpha) = \{u>0\colon |\Delta\alpha(t)| = u & \text{for some } t>0\}, \end{cases}$$

which both are at most countable.

Secondly, as 1.6 shows, the moduli of continuity w_N are well adapted to continuous functions. For càdlàg functions, the good moduli of "continuity" are the following ones, where $N\in\mathbb{N}^*$ and $\theta>0$ and α is any function from $\mathbb{R}_+$ into $\mathbb{R}^d$:

1.8 $$w'_N(\alpha,\theta) = \inf\left\{\max_{i\le r} w(\alpha;[t_{i-1},t_i))\colon 0=t_0<\cdots<t_r=N,\ \inf_{i<r}(t_i-t_{i-1})\ge\theta\right\}.$$

For all $\alpha\in\mathbb{D}(\mathbb{R}^d)$ the following is obvious:

1.9 $$\begin{cases} \text{(i) } \theta \rightsquigarrow w'_N(\alpha,\theta) \quad \text{and} \quad N\rightsquigarrow w'_N(\alpha,\theta) \text{ are non-decreasing,}\\ \text{(ii) } w'_N(\alpha,\theta)\le w_N(\alpha,2\theta). \end{cases}$$ □

1.10 **Remarks.** 1) In the book [12] for example, another modulus is introduced under the notation $w'(\alpha,\theta)$, which looks very much alike $w'_1(\alpha,\theta)$ in 1.8, and more precisely is given by

$$w'(\alpha,\theta) = \inf\left\{\max_{i\le r} w(\alpha;[t_{i-1},t_i))\colon 0=t_0<\cdots<t_r=1,\ \inf_{i\le r}(t_i-t_{i-1})\ge\theta\right\}$$

The difference is that in 1.8 we do not require $t_i - t_{i-1}\ge\theta$ for $i=r$. The reason for this difference is that in the Skorokhod topology on the space of càdlàg functions: $[0,1]\to\mathbb{R}^d$ the end-point 1 plays a specific rôle, while here the points $N\in\mathbb{N}^*$ should not play any specific rôle.

2) Instead of considering $\mathbb{D}(\mathbb{R}^d)$ one could consider the space $\mathbb{D}(E)$ of all càdlàg functions: $\mathbb{R}_+\to E$, where E is a Polish space. All the results would remain valid (provided the usual length on $\mathbb{R}^d$ is replaced by the metric of E). □

The following easy lemma should be compared to 1.6:

1.11 **Lemma.** *A function* α: $\mathbb{R}_+ \to \mathbb{R}^d$ *belongs to* $\mathbb{D}(\mathbb{R}^d)$ *if and only if for all* $N \in \mathbb{N}^*$ *we have*:

(i) $\sup_{s \le N} |\alpha(s)| < \infty$;
(ii) $\lim_{\theta \downarrow 0} w'_N(\alpha, \theta) = 0$.

Proof. a) Let $\alpha \in \mathbb{D}(\mathbb{R}^d)$ and $N \in \mathbb{N}^*$ and $\varepsilon > 0$. That (i) holds is obvious. Let $s_0 = 0$, $\dots, s_{n+1} = \inf(t > s_n\colon |\alpha(t) - \alpha(s_n)| > \varepsilon/2)$; then $s_n \uparrow \infty$ because α is càdlàg, so there exists $p \in \mathbb{N}$ such that $s_p \le N < s_{p+1}$. Moreover, $w(\alpha; [s_i, s_{i+1})) \le \varepsilon$ by construction, hence $w'_N(\alpha, \theta) \le \varepsilon$ if $\theta \le \inf_{i \le p}(s_i - s_{i-1})$, and thus (ii) holds.

b) Conversely, assume that (i) and (ii) hold for all $N \in \mathbb{N}^*$. If α does not belong to $\mathbb{D}(\mathbb{R}^d)$ there exists $t \in \mathbb{R}_+$ and an integer $i \le d$ such that the i^{th} coordinate α^i either (1) has no left-hand limit in $\mathbb{R}$ at time t, or (2) is not right-continuous at time t.

In case (1), either $\limsup_{s \uparrow\uparrow t} |\alpha^i(s)| = \infty$, which contradicts (i), or $a := \liminf_{s \uparrow\uparrow t} \alpha^i(s)$ is smaller than $b := \limsup_{s \uparrow\uparrow t} \alpha^i(s)$, in which case it is easy to deduce that $w'_N(\alpha, \theta) \ge b - a > 0$ for all $N \ge t$ and all $\theta > 0$, and so (ii) is contradicted.

In case (2), either $a := \limsup_{s \downarrow t} \alpha^i(s) > b := \alpha^i(t)$, or $c := \liminf_{s \downarrow t} \alpha^i(s) < b$; then $w(\alpha; [u, v)) \ge a - b$ (resp. $b - c$) for all u, v such that $u \le t < v$, thus $w'_N(\alpha, \theta) \ge a - b$ (resp. $b - c$) for all $N > t$, $\theta > 0$, which again contradicts (ii). □

Finally, the following lemma gives another useful expression for $w'_N(\alpha, \theta)$:

1.12 **Lemma.** *If* α *is a function*: $\mathbb{R}_+ \to \mathbb{R}^d$ *we have*

$$w'_N(\alpha, \theta) = \inf\left\{\max_{i \le r} w(\alpha; [t_{i-1}, t_i))\colon 0 = t_0 < \cdots < t_r = N, \right.$$
$$\left. \theta \le t_i - t_{i-1} \le 2\theta \text{ if } i \le r - 1, \text{ and } t_r - t_{r-1} \le 2\theta \right\}.$$

Proof. Let $0 = t_0 < \cdots < t_r = N$ with $t_i - t_{i-1} \ge \theta$ for $i \le r - 1$. If $t_i - t_{i-1} > 2\theta$ for some $i \le r$ we can further subdivize $[t_{i-1}, t_i]$ into $t_{i-1} = s_i^0 < \cdots < s_i^p = t_i$ in such a way that $\theta \le s_i^k - s_i^{k-1} \le 2\theta$, except when $i = r$ in which case we may have $s_p^k - s_p^{k-1} < \theta$; of course, $w(\alpha; [s_i^{k-1}, s_i^k)) \le w(\alpha; [t_{i-1}, t_i))$, so if we compare to 1.8 we immediately obtain the result. □

§ 1b. The Skorokhod Topology: Definition and Main Results

We summarize all the needed results in a single theorem, after introducing the following notation:

1.13 Λ is the set of all continuous functions λ: $\mathbb{R}_+ \to \mathbb{R}_+$ that are strictly increasing, with $\lambda(0) = 0$ and $\lambda(t) \uparrow \infty$ as $t \uparrow \infty$ (we say that such a function λ is a *change of time*). □

1.14 **Theorem.** a) *There is a metrizable topology on* $\mathbb{D}(\mathbb{R}^d)$, *called the Skorokhod topology, for which this space is Polish, and which is characterized as follows: a sequence* (α_n) *converges to* α *if and only if there is a sequence* $\{\lambda_n\} \subset \Lambda$ *such that*

1.15
$$\begin{cases} \text{(i)} \ \sup_s |\lambda_n(s) - s| \to 0 \\ \text{(ii)} \ \sup_{s \le N} |\alpha_n \circ \lambda_n(s) - \alpha(s)| \to 0 \quad \text{for all } N \in \mathbb{N}^*. \end{cases}$$

b) *A subset A of* $\mathbb{D}(\mathbb{R}^d)$ *is relatively compact for the Skorokhod topology if and only if*

1.16
$$\begin{cases} \text{(i)} \ \sup_{\alpha \in A} \sup_{s \le N} |\alpha(s)| < \infty \quad \text{for all } N \in \mathbb{N}^*, \\ \text{(ii)} \ \lim_{\theta \downarrow 0} \sup_{\alpha \in A} w'_N(\alpha, \theta) = 0 \quad \text{for all } N \in \mathbb{N}^*. \end{cases}$$

c) *The Borel* σ*-field equals the* σ*-field* $\mathscr{D}(\mathbb{R}^d)$ *of 1.1; moreover, for all* $t > 0$ *the* σ*-field* $\mathscr{D}_{t-}(\mathbb{R}^d) = \bigvee_{s<t} \mathscr{D}_s(\mathbb{R}^d)$ *is generated by all the functions on* $\mathbb{D}(\mathbb{R}^d)$ *that are* $\mathscr{D}_{t-}(\mathbb{R}^d)$*-measurable and Skorokhod-continuous* (i.e., continuous for the Skorokhod topology).

The proof of this theorem is provided below in § 1c, where a distance defining the topology and under which the space is complete is explicitely given. It should be emphazised that this theorem is essentially contained in the book [12] (except that here we consider "$\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$" instead of "$\mathbb{D}([0,1], \mathbb{R}^d)$"), and we strongly advise the rader to skip the proof, which sheds very little light on the significance of this topology.

In fact, the meaning of the Skorokhod topology is fully embodied in the characterization 1.15, and perhaps the following simple results and examples may provide some insight.

1.17 **Proposition.** a) *The Skorokhod topology is weaker than the local uniform topology.*

b) *If* α *is a continuous function, a sequence* (α_n) *converges to* α *for the Skorokhod topology if and only if it converges to* α *locally uniformly.*

Proof. a) If $\delta_{lu}(\alpha_n, \alpha) \to 0$, then 1.15 is met with $\lambda_n(t) = t$ for all n.

b) Assume that $\alpha_n \to \alpha$ for the Skorokhod topology, and let $\{\lambda_n\} \subset \Lambda$ meet 1.15. Then

1.18 $$|\alpha_n(t) - \alpha(t)| \le |\alpha_n \circ \lambda_n \circ \lambda_n^{-1}(t) - \alpha \circ \lambda_n^{-1}(t)| + |\alpha \circ \lambda_n^{-1}(t) - \alpha(t)|.$$

Since α is uniformly continuous over all finite intervals, 1.15(i) yields $\|\alpha \circ \lambda_n^{-1} - \alpha\|_N \to 0$ (recall notation 1.2); moreover $\|\alpha_n \circ \lambda_n \circ \lambda_n^{-1} - \alpha \circ \lambda_n^{-1}\|_N \le \|\alpha_n \circ \lambda_n - \alpha\|_{N+1}$ whenever $\sup_s |\lambda_n(s) - s| \le 1$, and $\|\alpha_n \circ \lambda_n - \alpha\|_N \to 0$ by 1.15(ii). Hence we deduce that $\|\alpha_n - \alpha\|_N \to 0$ from 1.18, and so $\delta_{lu}(\alpha_n, \alpha) \to 0$. □

Here are two important examples, which readily follow from 1.15:

1.19 *Example.* Let $\alpha_n(s) = x_n 1_{\{t_n \le s\}}$. Then (α_n) tends to a limit α in $\mathbb{D}(\mathbb{R}^d)$ if and only if

i) either $t_n \to \infty$; then $\alpha = 0$;

ii) or $t_n \to t < \infty$ and $x_n \to x$; then $\alpha(s) = x 1_{\{t \le s\}}$. □

1.20 *Example.* Let $\alpha_n(s) = x_n 1_{\{t_n \le s\}} + y_n 1_{\{r_n \le s\}}$, with $t_n < r_n$. Then (α_n) tends to a limit α if and only if

i) either $t_n \to \infty$ (hence $r_n \to \infty$); then $\alpha = 0$;

ii) or $t_n \to t < \infty$, $r_n \to \infty$, $x_n \to x$; then $\alpha(s) = x 1_{\{t \le s\}}$;

iii) or $t_n \to t < \infty$, $r_n \to r < \infty$, $x_n \to x$, $y_n \to y$, *and* $t < r$ *if* $x \ne 0 \ne y$; then $\alpha(s) = x 1_{\{t \le s\}} + y 1_{\{r \le s\}}$. □

The last example shows two important facts:

1.21 As "abstract" spaces, one may of course identify $\mathbb{D}(\mathbb{R}^d)$ with the cartesian product $\mathbb{D}(\mathbb{R})^d$. But as topological spaces, for the Skorokhod topology, *the topology of* $\mathbb{D}(\mathbb{R}^d)$ *is strictly finer than the product topology of* $\mathbb{D}(\mathbb{R})^d$ (however, the σ-fields $\mathscr{D}(\mathbb{R}^d)$ and $\mathscr{D}(\mathbb{R})^{\otimes d}$ coincide). □

1.22 We may have $\alpha_n \to \alpha$, $\beta_n \to \beta$ in $\mathbb{D}(\mathbb{R}^d)$ and yet $\alpha_n + \beta_n$ not converging to $\alpha + \beta$: $\mathbb{D}(\mathbb{R}^d)$ *is not a topological vector space.* □

However, we have the following (to be improved later, in 2.2):

1.23 **Proposition.** *If* $\alpha_n \to \alpha$ *and* $\beta_n \to \beta$ *for the Skorokhod topology and if* β *is continuous, then* $\alpha_n + \beta_n \to \alpha + \beta$ *for the Skorokhod topology.*

Proof. Let $\{\lambda_n\} \subset \Lambda$ meet 1.15 for $\{\alpha_n\}$. It suffices to prove that it also satisfies 1.15(ii) with $\{\beta_n\}$. But $|\beta_n \circ \lambda_n - \beta| \le |\beta_n \circ \lambda_n - \beta \circ \lambda_n| + |\beta \circ \lambda_n - \beta|$, so the uniform continuity of β on finite intervals, plus 1.17b and 1.15i, imply the claim. □

§ 1c. Proof of Theorem 1.14

The proof is broken into many steps, and we try to follow the shortest possible route. Our first task is to define a distance compatible with the convergence 1.15.

To this end, we define for each $N \in \mathbb{N}^*$ the following function k_N:

1.24 $$k_N(t) = \begin{cases} 1 & \text{if } t \le N \\ N + 1 - t & \text{if } N < t < N + 1 \\ 0 & \text{if } t \ge N + 1. \end{cases}$$

Then, for all $\lambda \in \Lambda$ we set

1.25 $$|||\lambda||| = \sup_{s<t} \left| \operatorname{Log} \frac{\lambda(t) - \lambda(s)}{t - s} \right|.$$

Finally, for $\alpha, \beta \in \mathbb{D}(\mathbb{R}^d)$ we set

1.26 $$\begin{cases} \delta_N(\alpha, \beta) = \inf_{\lambda \in \Lambda} (|||\lambda||| + \|(k_N\alpha) \circ \lambda - k_N\beta\|_\infty) \\ \delta(\alpha, \beta) = \sum_{N \in \mathbb{N}^*} 2^{-N}(1 \wedge \delta_N(\alpha, \beta)) \end{cases}$$

($\|\cdot\|_\infty$ is defined in 1.2, and $k_N\alpha$ is the product of the real-valued function k_N with the $\mathbb{R}^d$-valued function α).

1.27 **Remark.** We will see below that δ is a distance under which $\mathbb{D}(\mathbb{R}^d)$ is complete. This distance has been introduced by Prokhorov. The "original" distance of Skorokhod, once transposed for functions defined on $\mathbb{R}_+$ instead of $[0, 1]$, is

$$\delta'(\alpha, \beta) = \sum_{N \in \mathbb{N}^*} 2^{-N}(1 \wedge \delta'_N(\alpha, \beta)), \quad \text{where}$$

$$\delta'_N(\alpha, \beta) = \inf_{\lambda \in \Lambda} (\|\lambda - I\|_\infty + \|(k_N\alpha) \circ \lambda - k_N\beta\|_\infty)$$

and where $I \in \Lambda$ is the identity mapping: $I(t) = t$. Although δ and δ' define the same topology, $\mathbb{D}(\mathbb{R}^d)$ is not complete under δ'. □

The following properties are obvious (there, $\lambda, \mu \in \Lambda$ and $I \in \Lambda$ is the identity: $I(t) = t$):

1.28 $$\begin{cases} |||\lambda||| = |||\lambda^{-1}|||, \quad |||\lambda \circ \mu||| \le |||\lambda||| + |||\mu||| \\ \|\lambda - I\|_t \le t(e^{|||\lambda|||} - 1), \quad \|\mu \circ (\lambda - I)\|_t \le \|\lambda - I\|_t e^{|||\mu|||}. \end{cases}$$

1.29 $$\|(k_N\alpha) \circ \lambda \circ \mu - (k_N\beta) \circ \mu\|_\infty = \|(k_N\alpha) \circ \lambda - k_N\beta\|_\infty.$$

1.30 **Lemma.** *δ is nonnegative, symmetric, and satisfies the triangular inequality.*

Proof. It suffices to prove the claims for each δ_N. The first claim is obvious, and the second claim comes from 1.28 and 1.29. Let $a = \delta_N(\alpha, \beta)$, $b = \delta_N(\beta, \gamma)$, and $\varepsilon > 0$. There exists $\lambda, \mu \in \Lambda$ such that

$$|||\lambda||| + \|(k_N\alpha) \circ \lambda - k_N \circ \beta\|_\infty \le a + \varepsilon, \quad |||\mu||| + \|(k_N\beta) \circ \mu - k_N\gamma\|_\infty \le b + \varepsilon.$$

Then $\lambda \circ \mu \in \Lambda$, and 1.28 and 1.29 yield

$$\begin{aligned}&|||\lambda \circ \mu||| + \|(k_N\alpha) \circ \lambda \circ \mu - k_N\gamma\|_\infty \\ &\quad \leq |||\lambda||| + |||\mu||| + \|(k_N\alpha) \circ \lambda \circ \mu - (k_N\beta) \circ \mu\|_\infty + \|(k_N\beta) \circ \mu - k_N\gamma\|_\infty \\ &\quad \leq a + b + 2\varepsilon\end{aligned}$$

and since $\varepsilon > 0$ is arbitrary we deduce $\delta_N(\alpha, \gamma) \leq \delta_N(\alpha, \beta) + \delta_N(\beta, \gamma)$. □

1.31 **Lemma.** *If $\delta(\alpha_n, \alpha) \to 0$ there exists a sequence $\{\lambda_n\} \subset \Lambda$ such that 1.15 holds.*

Proof. We have $\delta_N(\alpha_n, \alpha) \to 0$ for all $N \in \mathbb{N}^*$, so there are sequences $\{\lambda_n^N\}_{n \geq 1} \subset \Lambda$ such that

$$a_n^N := |||\lambda_n^N||| + \|(k_N\alpha_n) \circ \lambda_n^N - k_N\alpha\|_\infty \to 0 \quad \text{as } n \uparrow \infty.$$

Hence there is an increasing sequence $(m_N)_{N \geq 1}$ such that $m_N \geq N$ and $a_n^N \leq \frac{1}{N}$ for all $n \geq m_N$. Put $r_n = \sup(N\colon m_N \leq n)$. Then $r_n < \infty$ and $r_n \uparrow \infty$ as $n \uparrow \infty$ and $a_n^{r_n} \leq 1/r_n$. Set

$$\lambda_n(t) = \begin{cases} \lambda_n^{r_n}(t) & \text{if } t \leq \sqrt{r_n} \\ t + \lambda_n^{r_n}(\sqrt{r_n}) - \sqrt{r_n} & \text{if } t > \sqrt{r_n}. \end{cases}$$

Obviously we have $\lambda_n \in \Lambda$ and, using 1.28,

$$\begin{aligned}\|\lambda_n - I\|_\infty = \|\lambda_n - I\|_{\sqrt{r_n}} &\leq \sqrt{r_n}(\exp|||\lambda_n^{r_n}||| - 1) \\ &\leq \sqrt{r_n}(e^{1/r_n} - 1) \to 0 \quad \text{as } n \uparrow \infty,\end{aligned}$$

so $\{\lambda_n\}$ meets 1.15(i). Moreover, let $N \in \mathbb{N}^*$ be fixed; for all n large enough we have $\lambda_n(t) = \lambda_n^{r_n}(t)$ and $k_N \circ \lambda_n(t) = k_N(t) = 1$ for all $t \leq N$, hence

$$\|\alpha_n \circ \lambda_n - \alpha\|_N \leq \|(k_N\alpha_n) \circ \lambda_n^{r_n} - k_N\alpha\|_\infty \leq a_n^{r_n} \leq \frac{1}{r_n} \to 0 \quad \text{as } n \uparrow \infty,$$

and thus 1.15(ii) is met. □

1.32 **Corollary.** *δ is a distance on $\mathbb{D}(\mathbb{R}^d)$.*

Proof. In view of 1.30, it remains to prove that $\delta(\alpha, \beta) = 0$ implies $\alpha = \beta$. Now, if $\delta(\alpha, \beta) = 0$, 1.31 implies the existence of a sequence $\{\lambda_n\} \subset \Lambda$ meeting 1.15(i) and $\|\alpha \circ \lambda_n - \beta\|_N \to 0$ for all $N \in \mathbb{N}^*$. Thus $\alpha \circ \lambda_n(t) \to \beta(t)$ for all t, while 1.15(i) yields that $\alpha \circ \lambda_n(t) \to \alpha(t)$ for all $t \notin J(\alpha)$ (notation 1.7), and we deduce that $\alpha = \beta$. □

1.33 **Lemma.** *$\mathbb{D}(\mathbb{R}^d)$ is complete under δ.*

Proof. We need to prove that a sequence (α_n) satisfying $\delta(\alpha_n, \alpha_{n+1}) \leq 2^{-n}$ converges for the distance δ to a limit α in $\mathbb{D}(\mathbb{R}^d)$.

For each $N \in \mathbb{N}^*$ we have $\delta_N(\alpha_n, \alpha_{n+1}) \leq 2^{N-n}$, so there is a change of time λ_n^N such that

1.34 $$|||\lambda_n^N||| + \|(k_N\alpha_n)\circ\lambda_n^N - k_N\alpha_{n+1}\|_\infty \leq 2^{N+1-n}.$$

Set $\rho_{n,p}^N = \lambda_n^N \circ \lambda_{n+1}^N \circ \ldots \circ \lambda_p^N$ for $p \geq n$. Then 1.34 and 1.28 yield

1.35 $$|||\rho_{n,p}^N||| \leq 2^{N+2-n}.$$

If $n \leq p < q$, $\rho_{n,q}^N - \rho_{n,p}^N = \rho_{n,p}^N \circ (\rho_{p+1,q}^N - I)$, so 1.35 and 1.28 give

$$\|\rho_{n,q}^N - \rho_{n,p}^N\|_{N+2} \leq (N+2)e^{2^{N+2-n}}(e^{2^{N+1-p}} - 1),$$

which goes to 0 as $p, q \uparrow \infty$. Hence if $\tilde{\rho}_{n,p}^N$ denotes the restriction of $\rho_{n,p}^N$ to $[0, N+2]$, the sequence $(\tilde{\rho}_{n,p}^N)_{p\geq n}$ is Cauchy for the uniform topology, and so it converges uniformly to a non-decreasing continuous function $\tilde{\mu}_n^N$ on $[0, N+2]$. We extend $\tilde{\mu}_n^N$ to $\mathbb{R}_+$ by setting

$$\mu_n^N(t) = \begin{cases} \tilde{\mu}_n^N(t) & \text{if } t \leq N+2 \\ t - (N+2) + \tilde{\mu}_n^N(N+2) & \text{if } t > N+2, \end{cases}$$

so μ_n^N is non-decreasing and continuous with $\mu_n^N(0) = 0$ and $\|\mu_n^N - \rho_{n,p}^N\|_{N+2} \to 0$ as $p \uparrow \infty$. Moreover the slope of μ_n^N equals 1 on $[N+2, \infty)$, so $|||\mu_n^N||| = \sup_{s<t\leq N+2} |\mathrm{Log}[(\mu_n^N(t) - \mu_n^N(s))/(t-s)]|$; therefore $|||\mu_n^N||| \leq \limsup_p |||\rho_{n,p}^N|||$, and 1.35 implies

1.36 $$|||\mu_n^N||| \leq 2^{N+2-n}.$$

In particular, μ_n^N is strictly increasing, and thus belongs to Λ.

$\lambda_n^N \circ \rho_{n+1,p}^N = \rho_{n,p}^N$ by construction, so letting $p \uparrow \infty$ we obtain that $\lambda_n^N \circ \mu_{n+1}^N(t) = \mu_n^N(t)$ for $t \leq N+2$. Moreover 1.34 and 1.36 and 1.28 imply that $\|\lambda_n^N - I\|_{N+2} \to 0$ and $\|\mu_n^N - I\|_{N+2} \to 0$ as $n \uparrow \infty$, so for all n large enough we have $\lambda_n^N \circ \mu_{n+1}^N(t) > N+1$ and $\mu_n^N(t) > N+1$ if $t > N+2$; therefore, recalling 1.24 we obtain that for all n large enough, $(k_N\alpha_n)\circ\lambda_n^N\circ\mu_{n+1}^N(t) = (k_N\alpha_n)\circ\mu_n^N(t)$ for all $t \in \mathbb{R}_+$. Hence 1.34 and 1.29 yield for n large enough:

$$\|(k_N\alpha_n)\circ\mu_n^N - (k_N\alpha_{n+1})\circ\mu_{n+1}^N\|_\infty \leq 2^{N+1-n}.$$

In other words, the sequence $\{(k_N\alpha_n)\circ\mu_n^N\}_{n\geq 1}$ is Cauchy for the uniform topology, and thus it converges to a limit $\gamma_N \in \mathbb{D}(\mathbb{R}^d)$. In particular, for n large enough,

1.37 $$\|(k_N\alpha_n)\circ\mu_n^N - \gamma_N\|_\infty \leq 2^{N+2-n}.$$

Now, $\|(k_N\alpha_n)\circ\mu_n^N - \gamma_N\|_\infty = \|(k_N\alpha_n) - \gamma_N\circ(\mu_n^N)^{-1}\|_\infty$, and $(\mu_n^N)^{-1} \to I$ locally uniformly (use 1.36 and 1.28). Therefore $\gamma_N\circ(\mu_n^N)^{-1}(t) \to \gamma_N(t)$ for all $t \notin J(\gamma_N)$ (notation 1.7), and so

1.38 $$k_N(t)\alpha_n(t) \to \gamma_N(t) \quad \text{for all } t \notin J(\gamma_N), \quad \text{as } n \uparrow \infty.$$

Since the set $\mathbb{R}_+ \setminus (\bigcup_{N\in\mathbb{N}^*} J(\gamma_N))$ is dense in $\mathbb{R}_+$, we immediately deduce from 1.38 the existence of a function $\alpha \in \mathbb{D}(\mathbb{R}^d)$ such that each γ_N has the form $\gamma_N = k_N\alpha$. Thus 1.36 and 1.37 yield for n large enough:

$$\delta_N(\alpha_n, \alpha) \le |||\mu_n^N||| + \|(k_N\alpha_n)\circ\mu_n^N - \gamma_N\|_\infty \le 2^{N+3-n}.$$

So $\delta_N(\alpha_n, \alpha) \to 0$ as $n \uparrow \infty$ for all $N \in \mathbb{N}^*$, thus $\delta(\alpha_n, \alpha) \to 0$ and we are finished. □

1.39 **Lemma.** *All relatively compact subsets of $\mathbb{D}(\mathbb{R}^d)$ for the topology induced by δ satisfy 1.16.*

Proof. Let A be a relatively compact subset. If 1.16 fails, there is a sequence $\{\alpha_n\} \in A$, which can be assumed to converge to a limit α (not necessarily in A), and an integer N, such that

(1) either $\|\alpha_n\|_N \to \infty$ as $n \uparrow \infty$,

(2) or $w'_N(\alpha_n, \theta) \ge \varepsilon$ for all $\theta > 0$, and some $\varepsilon > 0$.

By 1.31 there is a sequence $\{\lambda_n\} \subset \Lambda$ satisfying 1.15. Then for all n large enough, $\|\lambda_n - I\|_\infty \le 1$, and so $\|\alpha_n\|_N \le \|\alpha\|_{N+1} + \|\alpha_n \circ \lambda_n - \alpha\|_{N+1}$, which goes to $\|\alpha\|_{N+1}$ as $n \uparrow \infty$, hence (1) cannot be true.

Moreover, by 1.11 there exists $\theta \in (0, 1/4)$ such that $w'_{N+1}(\alpha, 2\theta) \le \varepsilon/6$, so there is a sequence $0 = s_0 < \cdots < s_p = N + 1$ with $s_i - s_{i-1} \ge 2\theta$ for $i < p$ and $s_{p-1} \ge N - 1/2$ and $w(\alpha; [s_i, s_{i+1})) \le \varepsilon/3$. For all n large enough we have $\|\lambda_n - I\|_\infty \le \theta/2$ and $\|\alpha_n \circ \lambda_n - \alpha\|_{N+\theta} \le \varepsilon/6$, thus if $t_i^n = \lambda_n(s_i)$ we have $t_i^n - t_{i-1}^n \ge \theta$ for $i < p$ and $t_{p-1}^n \ge N$ and $w(\alpha_n; [t_i^n, t_{i+1}^n)) \le 2\varepsilon/3$: it follows that $w'_N(\alpha_n, \theta) \le 2\varepsilon/3$, which contradicts (2). □

For the next lemma, we need some more notation. Let $N \in \mathbb{N}^*$, $k \in \mathbb{N}^*$ and $\theta > 0$. We denote by $C_{\theta,k}$ a *finite* subset of $\mathbb{R}^d$ such that all points of the ball $\{x \in \mathbb{R}^d\colon |x| \le \theta\}$ are at most at a distance $1/k$ of $C_{\theta,k}$. We denote by $\mathscr{A}(N, \theta, k)$ the *finite* subset of $\mathbb{D}(\mathbb{R}^d)$ consisting in all càdlàg functions taking their values in $C_{\theta,k}$ and which are piecewise constant and which jump at times $\frac{i}{k}$ only, for $1 \le i \le kN$.

1.40 **Lemma.** *Let $N \in \mathbb{N}^*$, $k \in \mathbb{N}^*$ with $k \ge 4$, $\theta > 0$. If $\alpha \in \mathbb{D}(\mathbb{R}^d)$ satisfies $\|\alpha\|_{N+3} \le \theta$ and $w'_{N+3}\left(\alpha, \frac{1}{k}\right) \le \frac{1}{k}$, there exists a function $\beta \in \mathscr{A}(N + 3, \theta, k^2)$ such that $\delta_{N'}(\alpha, \beta) \le 6/k$ for all $N' \le N$.*

Proof. Since $w'_{N+3}(\alpha, 1/k) \le 1/k$, there exists a subdivision $0 = t_0 < \cdots < t_p = N + 3$ with $t_i - t_{i-1} \ge 1/k$ for $i < p$, and $w(\alpha; [t_i, t_{i+1})) \le 1/k$; if $t_{p-1} < N + 2$ it is always possible to add an additional point (namely $N + 2 + 1/k$, recall that $k \ge 4$) so that we may assume $t_{p-1} \ge N + 2$.

Let $s_0 = 0$, and for $1 \le i < p$ let s_i be of the form j/k^2 for some $j = 1, 2, \ldots, k^2(N + 3)$ and such that $|s_i - t_i| \le 1/k^2$; thus $N + 1 \le s_{p-1} \le N + 3$. We consider the change of time λ defined as such: $\lambda(s_i) = t_i$ for $i < p$, and λ is affine between s_i and s_{i+1} for $i < p - 1$, and λ is affine with slope 1 on (s_{p-1}, ∞). Since

$|s_i - t_i| \leq 1/k^2$ and $t_i - t_{i-1} \geq 1/k$ for $i < p$, we clearly have (since $k \geq 4$):

1.41 $$|||\lambda||| \leq -\mathrm{Log}\left(1 - \frac{2}{k}\right) \leq \frac{4}{k}.$$

Now, there exists $\beta \in \mathscr{A}(N + 3, \theta, k^2)$ with the following properties: β is constant on each interval $[s_i, s_{i+1})$ for $i < p - 1$ and on $[s_{p-1}, \infty)$, and $|\beta(s_i) - \alpha(t_i)| \leq 1/k$ for $i < p$. Since $w(\alpha; [t_i, t_{i+1})) \leq 1/k$ we deduce that $|\beta(s) - \alpha \circ \lambda(s)| \leq 2/k$ for $s \in [s_i, s_{i+1})$, $i < p - 1$; since $\lambda(t_{p-1}) = s_{p-1} \geq N + 1$ it follows that for all $N' \leq N$,

1.42 $$\|k_{N'}\beta - (k_{N'}\alpha) \circ \lambda\|_\infty \leq \frac{2}{k}.$$

Putting together 1.41 and 1.42 gives the result. □

1.43 **Corollary.** a) *The space $\mathbb{D}(\mathbb{R}^d)$ is separable for the topology induced by δ.*
b) *A subset A meeting 1.16 is relatively compact for this topology.*

Proof. a) Let $\alpha \in \mathbb{D}(\mathbb{R}^d)$ and $N \in \mathbb{N}^*$ with $N \geq 2$. By 1.11 there exist $p \in \mathbb{N}^*$ with $p \geq \|\alpha\|_{N+3}$ and $k \in \mathbb{N}^*$ with $k \geq N^2$ and $w'_{N+3}\left(\alpha, \frac{1}{k}\right) \leq \frac{1}{k}$. Then 1.40 implies the existence of $\beta \in \mathscr{A}(N + 3, p, k^2)$ such that $\delta_{N'}(\alpha, \beta) \leq \frac{6}{k}$ for all $N' \leq N$; thus

$$\delta(\alpha, \beta) \leq \sum_{1 \leq N' \leq N} 2^{-N'}\frac{6}{k} + \sum_{N' \geq N+1} 2^{-N'} \leq \frac{6N}{k} + 2^{-N} \leq \frac{6}{N} + 2^{-N}.$$

Since N is arbitrary large, we deduce that the countable set

$$\mathscr{A} = \bigcup\nolimits_{N \in \mathbb{N}^*, k \in \mathbb{N}^*, p \in \mathbb{N}^*} \mathscr{A}(N + 3, p, k^2)$$

is dense in $\mathbb{D}(\mathbb{R}^d)$.

b) Assume that A meets 1.16. Since $\mathbb{D}(\mathbb{R}^d)$ is complete by 1.33, in order to prove the claim it suffices to show that for each $\varepsilon > 0$, there is a finite covering of A with balls of radius ε.

Let $N \in \mathbb{N}^*$ with $N \geq 2$. By 1.16 there exist $p, k \in \mathbb{N}^*$ with $p \geq \|\alpha\|_{N+3}$ and $k \geq N^2$ and $w'_{N+3}\left(\alpha, \frac{1}{k}\right) \leq \frac{1}{k}$ for all $\alpha \in A$. Then, exactly as in (a), for each $\alpha \in A$ there is a $\beta \in \mathscr{A}(N + 3, p, k^2)$ such that $\delta(\alpha, \beta) \leq \frac{6}{N} + 2^{-N}$: in other words, A is covered with the balls centered at all points of $\mathscr{A}(N + 3, p, k^2)$ and with radius $\frac{6}{N} + 2^{-N}$. Since N is arbitrary and since $\mathscr{A}(N + 3, p, k^2)$ is finite, we deduce the claim. □

1.44 **Lemma.** *Let $\alpha_n, \alpha \in \mathbb{D}(\mathbb{R}^d)$. If there is a sequence $\{\lambda_n\} \subset \Lambda$ such that 1.15 holds, we have $\delta(\alpha_n, \alpha) \to 0$.*

Proof. It is easy to deduce from 1.15 that $\|\lambda_n^{-1} - I\|_\infty \to 0$ and $\|\alpha_n - \alpha \circ \lambda_n^{-1}\|_N \to 0$ for all $N \in \mathbb{N}^*$. Then $\alpha \circ \lambda_n^{-1}(t) \to \alpha(t)$ and thus $\alpha_n(t) \to \alpha(t)$ for all $t \notin J(\alpha)$. Hence α is the only limit point of the sequence (α_n), for the distance δ.

Therefore it remains to prove that the sequence (α_n) is relatively compact. In view of 1.43, it is enough to prove that the set $A = \{\alpha_n : n \in \mathbb{N}^*\}$ meets 1.16. Let $\varepsilon > 0$ and $N \in \mathbb{N}^*$. By 1.11 we have $\|\alpha\|_{N+1} < \infty$ and $w'_{N+1}(\alpha, \theta) \leq \varepsilon$ for some $\theta \in (0,1)$, while 1.15 yields an integer n_0 such that $\|\lambda_n - I\|_\infty \leq \theta/4$ and $\|\alpha_n \circ \lambda_n - \alpha\|_{N+1} \leq \varepsilon$ for all $n \geq n_0$. Then

$$\|\alpha_n\|_N \leq \|\alpha\|_{N+1} + \|\alpha_n \circ \lambda_n - \alpha\|_{N+1} \leq \varepsilon + \|\alpha\|_{N+1}$$

for $n \geq n_0$, while $\sup_{n < n_0} \|\alpha_n\|_N < \infty$ by 1.11 again, so A meets 1.16(i)

Since $w'_{N+1}(\alpha, \theta) \leq \varepsilon$, there is a subdivision $0 = t_0 < \cdots < t_p = N + 1$ with $t_i - t_{i-1} \geq \theta$ for $i < p$ and $w(\alpha; [t_i, t_{i+1})) \leq 2\varepsilon$ for $i < p$. Then if $s_i^n = \lambda_n(t_i)$ we have $s_i^n - s_{i-1}^n \geq \theta/2$ for $i < p$, and $s_p^n \geq N$, and $w(\alpha_n; [s_i^n, s_{i+1}^n)) \leq 4\varepsilon$ for $i < p$ (because $\|\alpha_n \circ \lambda_n - \alpha\|_{N+1} \leq \varepsilon$): it is easy to deduce that $w'_N(\alpha_n, \theta/2) \leq 4\varepsilon$ for $n \geq n_0$. Moreover $\lim_{\rho \downarrow 0} \sup_{n < n_0} w'_N(\alpha_n, \rho) = 0$ by 1.11, so there exists $\theta' \in (0, \theta/2]$ such that $w'_N(\alpha_n, \theta') \leq 4\varepsilon$ for all $n \geq 1$, hence A meets 1.16(ii). □

At this stage, we have proved parts (a) and (b) of Theorem 1.14. The proof of 1.14(c) is also divided into two steps.

1.45 **Lemma.** *The σ-field $\mathscr{D}_{t-}(\mathbb{R}^d)$ is generated by all functions on $\mathbb{D}(\mathbb{R}^d)$ that are $\mathscr{D}_{t-}(\mathbb{R}^d)$-measurable and continuous for the topology induced by δ; in particular, $\mathscr{D}(\mathbb{R}^d)$ is contained into the Borel σ-field.*

Proof. Note that $\mathscr{D}_{t-}(\mathbb{R}^d)$ is generated by the mappings $\alpha \rightsquigarrow f(\alpha(s))$, for $s < t$ and f continuous and bounded on $\mathbb{R}^d$: so for the first claim it is enough to prove that such a mapping is the pointwise limit of a sequence of continuous functions on $\mathbb{D}(\mathbb{R}^d)$ that are $\mathscr{D}_{t-}(\mathbb{R}^d)$-measurable. The sequence

$$\rho_n(\alpha) = n \int_s^{t \wedge (s+1/n)} f(\alpha(r))\, dr$$

is clearly $\mathscr{D}_{t-}(\mathbb{R}^d)$-measurable, and it converges to $f(\alpha(s))$ for all α. Moreover, if $\alpha_p \to \alpha$ we have seen in the proof of 1.44 that $\alpha_p(s) \to \alpha(s)$ for all $s \notin J(\alpha)$. Since $J(\alpha)$ is at most countable, we deduce from Lebesgue's convergence theorem that $\rho_n(\alpha_p) \to \rho_n(\alpha)$ as $p \uparrow \infty$, so ρ_n is a continuous function on $\mathbb{D}(\mathbb{R}^d)$.

Finally, the last claim follows by taking $t = \infty$. □

1.46 **Lemma.** *The Borel σ-field of* $\mathbb{D}(\mathbb{R}^d)$ (for the topology induced by δ) *is contained into* $\mathscr{D}(\mathbb{R}^d)$.

Proof. Since the space $\mathbb{D}(\mathbb{R}^d)$ is Polish, it is enough to prove that all balls are $\mathscr{D}(\mathbb{R}^d)$-measurable, or equivalently that $\alpha \rightsquigarrow \delta(\alpha, \beta)$ is $\mathscr{D}(\mathbb{R}^d)$-measurable for all β, or even that $\alpha \rightsquigarrow \delta_N(\alpha, \beta)$ is $\mathscr{D}(\mathbb{R}^d)$-measurable for all $N \in \mathbb{N}^*$, $\beta \in \mathbb{D}(\mathbb{R}^d)$.

Let us fix $N \in \mathbb{N}^*$ and $\beta \in \mathbb{D}(\mathbb{R}^d)$. Let $t_1, t_2, \ldots, t_n, \ldots$ be an enumeration of $\mathbb{Q}_+$, and set for $\theta > 0$, $\varepsilon > 0$:

$$A_n(\theta, \varepsilon) = \{\alpha: \text{there exists } \lambda \in \Lambda \text{ with } |||\lambda||| \leq \theta \text{ and } \sup_{i \leq n} |(k_N \beta) \circ \lambda(t_i) - (k_N \alpha)(t_i)| < \varepsilon\}.$$

We will first prove that $A_n(\theta, \varepsilon) \in \mathscr{D}(\mathbb{R}^d)$. Indeed the subset $B(n, \varepsilon, \lambda)$ of $(\mathbb{R}^d)^n$ consisting in all points $(x_1, \ldots, x_n)$ (where $x_i \in \mathbb{R}^d$) such that $\sup_{i \leq n} |x_i - (k_N \beta) \circ \lambda(t_i)| < \varepsilon$ is open, and so is $B'(n, \theta, \varepsilon) = \bigcup_{\lambda \in \Lambda: |||\lambda||| \leq \theta} B(n, \varepsilon, \lambda)$. Since

$$A_n(\theta, \varepsilon) = \{\alpha: (k_N(t_1)\alpha(t_1), \ldots, k_N(t_n)\alpha(t_n)) \in B'(n, \theta, \varepsilon)\}$$

it follows that $A_n(\theta, \varepsilon) \in \mathscr{D}(\mathbb{R}^d)$ from the definition of this σ-field.

Secondly, let $a > 0$ and $C = \{\alpha: \delta_N(\alpha, \beta) < a\}$. It remains to prove that $C \in \mathscr{D}(\mathbb{R}^d)$, and in view of what precedes it is enough to prove that

1.47 $$C = \bigcup_{b,c \in \mathbb{Q}: 0 < b < c < a} \; \bigcap_{n \in \mathbb{N}^*} A_n(b, c - b).$$

That C is included into the right-hand side of 1.47 is immediate from 1.26. Conversely, let $0 < b < c < a$ and let $\alpha \in \bigcap_n A_n(b, c - b)$. For all $n \in \mathbb{N}^*$ there exists $\lambda_n \in \Lambda$ such that $|||\lambda_n||| \leq b$ and

1.48 $$|(k_N \beta) \circ \lambda_n(t_i) - (k_N \alpha)(t_i)| < c - b \quad \text{for } i \leq n.$$

It is easy to deduce from $|||\lambda_n||| \leq b$ and from 1.25 that $w(\lambda_n; [s, t]) \leq (t - s)e^b$. Therefore Theorem 1.6 yields that the seqence (λ_n) is relatively compact for the local uniform topology in $\mathbb{C}(\mathbb{R})$, so there is a subsequence (λ_{n_k}) which converges locally uniformly. In particular the restrictions λ'_{n_k} of λ_{n_k} to $[0, N + 2]$ converge uniformly to a limit λ', which is a continuous non-decreasing function on $[0, N + 2]$ with $\lambda'(0) = 0$. Put

$$\lambda(t) = \begin{cases} \lambda'(t) & \text{if } t \leq N + 2 \\ t - (N + 2) + \lambda'(N + 2) & \text{if } t > N + 2. \end{cases}$$

Then, exactly as in the proof of 1.33, we deduce from $|||\lambda_{n_k}||| \leq b$ that $|||\lambda||| \leq b$. In particular, $\lambda \in \Lambda$.

Moreover, $\lambda_{n_k}(t_i) \to \lambda(t_i)$ for all $t_i \leq N + 2$, and $k_N(t) = 0$ if $t \geq N + 1$. Hence the sequence $\{(k_N \beta) \circ \lambda_{n_k}(t_i)\}_{k \geq 1}$ has either one or two limit points, belonging to the set $\{(k_N \beta) \circ \lambda(t_i), (k_N \beta)(\lambda(t_i)-)\}$. Thus we deduce from 1.48 that for every $i \geq 1$ we have at least one of the following two inequalities:

$$|(k_N \beta)(\lambda(t_i)) - (k_N \alpha)(t_i)| \leq c - b, \quad \text{or} \quad |(k_N \beta)(\lambda(t_i)-) - (k_N \alpha)(t_i)| \leq c - b.$$

Since $\bigcup_{i \geq 1} \{t_i\} = \mathbb{Q}_+$, it is easy to obtain that $|(k_N \beta) \circ \lambda(t) - (k_N \alpha)(t)| \leq c - b$ for all $t \in \mathbb{R}_+$. Hence we have proved so far that if $\alpha \in \bigcap_n A_n(b, c - b)$ there exists $\lambda \in \Lambda$ such that

$$|||\lambda||| \leq b, \quad \|(k_N \beta) \circ \lambda - k_N \alpha\|_\infty \leq c - b,$$

hence $\delta_N(\alpha, \beta) \leq c < a$ by 1.26, and so $\alpha \in C$. This finishes to prove 1.47. □

2. Continuity for the Skorokhod Topology

It is important to decide whether a function defined on $\mathbb{D}(\mathbb{R}^d)$ is continuous for the Skorokhod topology. As a matter of fact, not very many functions are so, hence the question becomes: at which points of $\mathbb{D}(\mathbb{R}^d)$ is a given function continuous?

We study many "useful" functions in § 2a, and we devote a special subsection (§ 2b) to the properties of the restriction of the Skorokhod topology to an important subset of $\mathbb{D}(\mathbb{R})$, namely the subset of all increasing functions on $\mathbb{R}_+$.

In all this section, writing $\alpha_n \to \alpha$ when $\alpha_n, \alpha \in \mathbb{D}(\mathbb{R}^d)$ implicitely means *convergence for the Skorokhod topology* (unless otherwise stated).

§ 2a. Continuity Properties of some Functions

1. We begin with a simple but very important result. The most useful part of it is the statement (b.5), which has already been proved in the beginning of the proof of Lemma 1.44.

2.1 **Proposition.** *Let $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ and $t \geq 0$.*

a) *There is a sequence $t_n \to t$ such that $\alpha_n(t_n) \to \alpha(t)$ and $\alpha_n(t_n -) \to \alpha(t-)$, and so $\Delta\alpha_n(t_n) \to \Delta\alpha(t)$.*

b) *Let (t_n) be any sequence with $t_n \to t$ and $\Delta\alpha_n(t_n) \to \Delta\alpha(t)$. If $\Delta\alpha(t) \neq 0$, any other sequence (t'_n) with the same properties coincides with t_n for all n large enough. Moreover,*

(b.1) *$t''_n \to t$ and $t''_n < t_n$ imply $\alpha_n(t''_n) \to \alpha(t-)$.*

(b.2) *$t''_n \to t$ and $t''_n \leq t_n$ imply $\alpha_n(t''_n -) \to \alpha(t-)$.*

(b.3) *$t''_n \to t$ and $t''_n \geq t_n$ imply $\alpha_n(t''_n) \to \alpha(t)$.*

(b.4) *$t''_n \to t$ and $t''_n > t_n$ imply $\alpha_n(t''_n -) \to \alpha(t)$.*

(b.5) *$t''_n \to t$ and $\Delta\alpha(t) = 0$ imply $\alpha_n(t''_n) \to \alpha(t)$ and $\alpha_n(t''_n -) \to \alpha(t)$*

(b.6) *$\alpha'_n \to \alpha'$ in $\mathbb{D}(\mathbb{R}^d)$, where $\alpha'_n(s) = \alpha_n(s) - \Delta\alpha_n(t_n)1_{\{t_n \leq s\}}$ and $\alpha'(s) = \alpha(s) - \Delta\alpha(t)1_{\{t \leq s\}}$.*

(b.7) $\lim_{\eta \downarrow 0} \limsup_n w(\alpha'_n, [t - \eta, t + \eta]) = 0$, *where α'_n is as above.*

Proof. There is a sequence $\{\lambda_n\} \subset \Lambda$ satisfying 1.15. Then $t_n = \lambda_n(t)$ obviously satisfies the claims in (a).

Assume that $t''_n \to t$ and $t''_n < t_n = \lambda_n(t)$ for all n. There is a sequence $\varepsilon_n > 0$ with $\varepsilon_n \downarrow 0$ and $\alpha_n(t''_n - \varepsilon_n) - \alpha_n(t''_n -) \to 0$, and another sequence $\varepsilon'_n > 0$ with $\varepsilon'_n \downarrow 0$ and $t''_n - \varepsilon_n = \lambda_n(t - \varepsilon'_n)$. Then

$$|\alpha_n(t''_n -) - \alpha(t-)| \leq |\alpha_n(t''_n - \varepsilon_n) - \alpha_n(t''_n -)| + |\alpha_n \circ \lambda_n(t - \varepsilon'_n) - \alpha(t - \varepsilon'_n)| + |\alpha(t - \varepsilon'_n) - \alpha(t-)|,$$

which goes to 0 by 1.15(i). Therefore $t_n = \lambda_n(t)$ satisfies (b.1), and one shows similarly that it meets (b.2, b.3, b.4).

Let $(\hat{t}_n)$ be another sequence converging to t, with $\Delta\alpha_n(\hat{t}_n) \to \Delta\alpha(t)$. If $\hat{t}_n$ is not equal to $t_n = \lambda_n(t)$ for all n large enough, there is an infinite sequence such that either $\hat{t}_{n_k} < t_{n_k}$ for all k, or $\hat{t}_{n_k} > t_{n_k}$ for all k. In the former case, (b.1) and (b.2) imply that $\Delta\alpha_{n_k}(\hat{t}_{n_k}) = \alpha_{n_k}(\hat{t}_{n_k}) - \alpha_{n_k}(\hat{t}_{n_k}-) \to 0$, in the latter case (b.3) and (b.4) similarly imply $\Delta\alpha_{n_k}(\hat{t}_{n_k}) \to 0$: hence we must have $\Delta\alpha(t) = 0$, and so the first claim of (b) is proved.

Obviously (b.5) follows from (b.i) for $i = 1, 2, 3, 4$. It remains to prove (b.6) and (b.7). Assume first that $\Delta\alpha(t) = 0$. Then $\alpha' = \alpha$, while the functions: $s \rightsquigarrow \Delta\alpha_n(t_n)1_{\{t_n \le s\}}$ converge uniformly to 0, hence 1.23 implies $\alpha'_n \to \alpha'$. If $\Delta\alpha(t) \neq 0$ we have $t_n = \lambda_n(t)$ for all n large enough. Then

$$|\alpha'_n \circ \lambda_n(s) - \alpha'(s)| \begin{cases} = |\alpha_n \circ \lambda_n(s) - \alpha(s)| & \text{if } s < t \\ \le |\alpha_n \circ \lambda_n(s) - \alpha(s)| + |\Delta\alpha_n(t_n) - \Delta\alpha(t)| & \text{if } s \ge t, \end{cases}$$

implying that (α'_n, α') satisfies condition 1.15(ii), and so $\alpha'_n \to \alpha'$. Finally $w(\alpha'_n; [t-\eta, t+\eta]) \le w(\alpha'_n \circ \lambda_n; [t-2\eta, t+2\eta])$ if $\sup_s |\lambda_n(s) - s| \le \eta$, and $w(\alpha'_n \circ \lambda_n; [t-2\eta, t+2\eta]) \le w(\alpha'; [t-2\eta, t+2\eta]) + 2\sup_{s \le t+3\eta} |\alpha'_n \circ \lambda_n(s) - \alpha'(s)|$. Since $w(\alpha'; [t-2\eta, t+2\eta]) \downarrow 0$ as $\eta \downarrow 0$ because α' is continuous at point t, we immediately deduce (b.7). □

The next result, which is a partial corollary to 2.1, improves upon 1.23.

2.2 **Proposition.** a) *Let $\alpha_n \to \alpha$ and $\beta_n \to \beta$ in $\mathbb{D}(\mathbb{R}^d)$. Assume that for each $t > 0$ there is a sequence $t_n \to t$, with $\Delta\alpha_n(t_n) \to \Delta\alpha(t)$ and $\Delta\beta_n(t_n) \to \Delta\beta(t)$. Then $\alpha_n + \beta_n \to \alpha + \beta$ in $\mathbb{D}(\mathbb{R}^d)$.*

b) *Let $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ and $\beta_n \to \beta$ in $\mathbb{D}(\mathbb{R}^{d'})$. Assume that for each $t > 0$ there is a sequence $t_n \to t$, with $\Delta\alpha_n(t_n) \to \Delta\alpha(t)$ and $\Delta\beta_n(t_n) \to \Delta\beta(t)$. Then $(\alpha_n, \beta_n) \to (\alpha, \beta)$ in $\mathbb{D}(\mathbb{R}^{d+d'})$.*

Proof. (b) is a simple corollary of (a): indeed, we associate $\bar{\alpha}_n \in \mathbb{D}(\mathbb{R}^{d+d'})$ to α_n by $\bar{\alpha}_n = (\alpha_n, 0)$, and similarly $\bar{\beta}_n = (0, \beta_n)$. It is obvious from 1.15 that $\bar{\alpha}_n \to \bar{\alpha} = (\alpha, 0)$ and $\bar{\beta}_n \to \bar{\beta} = (0, \beta)$ in $\mathbb{D}(\mathbb{R}^{d+d'})$, and $\bar{\alpha}_n + \bar{\beta}_n = (\alpha_n, \beta_n) \to (\alpha, \beta) = \bar{\alpha} + \bar{\beta}$ in $\mathbb{D}(\mathbb{R}^{d+d'})$ by (a).

It remains to prove (a). Since $\alpha_n(t) + \beta_n(t) \to \alpha(t) + \beta(t)$ for all $t \notin J(\alpha) \cup J(\beta)$, the only possible limit for the sequence $\gamma_n = \alpha_n + \beta_n$ is $\gamma = \alpha + \beta$. So it suffices to check that the sequence $\{\gamma_n\}$ is relatively compact.

This sequence obviously fulfills 1.16(i). Now, if 1.16(ii) were wrong, there would exist $N > 0$, $\varepsilon > 0$ and a sequence (n_k) such that $w'_N(\gamma_{n_k}, 1/k) > \varepsilon$. Due to 1.12, this means that we would be in one of the following two cases:

(a) there are numbers $s_k \le \frac{2}{k}$ with $|\gamma_{n_k}(s_k) - \gamma_{n_k}(0)| \ge \varepsilon$; but this is excluded, because 2.1 implies $\alpha_{n_k}(s_k) \to \alpha(0)$ and $\alpha_{n_k}(0) \to \alpha(0)$, and similarly for β;

(b) or, there are numbers with $0 < s_k^1 < s_k^2 < s_k^3 \le N$ and $\liminf_k s_k^1 > 0$ and $s_k^3 - s_k^1 \le 4/k$ and $|\gamma_{n_k}(s_k^{i+1}) - \gamma_{n_k}(s_k^i)| \ge \varepsilon$ for $i = 1, 2$. By taking a subsequence, we may assume that s_k^i converges to some limit $t \in (0, N]$, to which the assumption associates a sequence (t_n). Taking a further subsequence, we may assume that for all k, t_{n_k} is in the same interval $(0, s_k^1)$, or $[s_k^1, s_k^2)$, or $[s_k^2, s_k^3)$, or $[s_k^3, \infty)$. But if $s_k^i < t_{n_k}$ (resp. $s_k^i \geq t_{n_k}$) for all k, then 2.1 and the property of the sequence (t_n) imply that $\gamma_{n_k}(s_k^i) \to \gamma(t-)$ (resp. $\gamma(t)$). Hence among the three limits $a^i = \lim_k \gamma_{n_k}(s_k^i)$, one necessarily have $a^1 = a^2$, or $a^2 = a^3$. However, we have $|a^{i+1} - a^i| \geq \varepsilon$ by construction: hence there is a contradiction. □

2. Now, we start studying the continuity of various functions on $\mathbb{D}(\mathbb{R}^d)$. For example, we can write 2.1(b.5) as such:

2.3 The functions $\alpha \rightsquigarrow \alpha(t)$ and $\alpha \rightsquigarrow \alpha(t-)$ are continuous on $\mathbb{D}(\mathbb{R}^d)$ at each point α such that $t \notin J(\alpha)$ (i.e., such that $\Delta\alpha(t) = 0$).

2.4 **Proposition.** *The functions $\alpha \rightsquigarrow \sup_{t \leq a} |\alpha(t)|$ and $\alpha \rightsquigarrow \sup_{t \leq a} |\Delta\alpha(t)|$ are continuous at each point α such that $a \notin J(\alpha)$.*

Proof. Let $M^\alpha(t) = \sup_{s \leq t} |\alpha(s)|$ and $N^\alpha(t) = \sup_{s \leq t} |\Delta\alpha(s)|$. Let $\alpha_n \to \alpha$, with $a \notin J(\alpha)$. Let $\{\lambda_n\} \subset \Lambda$ be as in 1.15. Since λ_n is continuous and strictly increasing, $M^{\alpha_n}(a) = M^{\alpha_n \circ \lambda_n}(\lambda_n^{-1}(a))$ and $N^{\alpha_n}(a) = N^{\alpha_n \circ \lambda_n}(\lambda_n^{-1}(a))$. Then 1.15(ii) yields

$$|M^{\alpha_n}(a) - M^\alpha(\lambda_n^{-1}(a))| \to 0, \quad |N^{\alpha_n}(a) - N^\alpha(\lambda_n^{-1}(a))| \to 0.$$

Moreover $\lambda_n^{-1}(a) \to a$, and since $\Delta\alpha(a) = 0$ it follows clearly that $M^\alpha(\lambda_n^{-1}(a)) \to M^\alpha(a)$ and $N^\alpha(\lambda_n^{-1}(a)) \to N^\alpha(a)$ as $n \uparrow \infty$. Therefore $M^{\alpha_n}(a) \to M^\alpha(a)$ and $N^{\alpha_n}(a) \to N^\alpha(a)$, and we are done. □

2.5 **Lemma.** *If $\alpha_n \to \alpha$ we have for all $a < b$:*

$$\limsup_n \sup_{a \leq t \leq b} |\Delta\alpha_n(t)| \leq \sup_{a \leq t \leq b} |\Delta\alpha(t)|.$$

Proof. Let u be the right-hand side above. If the inequality were false there would exist $\varepsilon > 0$ and a sequence (n_k) in $\mathbb{N}$ and a sequence $\{t_k\} \subset [a, b]$ tending to a limit t, such that $|\Delta\alpha_{n_k}(t_k)| \geq u + \varepsilon$ for all k. Then 2.1 would imply $|\Delta\alpha(t)| \geq u + \varepsilon$, thus bringing a contradiction. □

Recall that $U(\alpha)$ is defined in 1.7, and set for $u > 0$:

2.6
$$\begin{cases} t^0(\alpha, u) = 0, \quad t^{p+1}(\alpha, u) = \inf(t > t^p(\alpha, u): |\Delta\alpha(t)| > u) \\ \alpha^u(s) = \alpha(s) - \sum_{p \geq 1} \Delta\alpha[t^p(\alpha, u)] 1_{\{t^p(\alpha, u) \leq s\}}. \end{cases}$$

Since $\alpha \in \mathbb{D}(\mathbb{R}^d)$ we have $\lim_{(p)} \uparrow t^p(\alpha, u) = \infty$, hence $\alpha^u \in \mathbb{D}(\mathbb{R}^d)$.

2.7 **Proposition.** *The functions $\alpha \rightsquigarrow t^p(\alpha, u)$ from $\mathbb{D}(\mathbb{R}^d)$ into $\overline{\mathbb{R}}_+$, $\alpha \rightsquigarrow \alpha^u$ from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^d)$, and $\alpha \rightsquigarrow \alpha(t^p(\alpha, u))$, $\alpha \rightsquigarrow \alpha(t^p(\alpha, u)-)$, $\alpha \rightsquigarrow \Delta\alpha(t^p(\alpha, u))$*

from $\mathbb{D}(\mathbb{R}^d)$ *into* $\mathbb{R}^d$, *are continuous at each point* α *such that* $u \notin U(\alpha)$, *and that* $t^p(\alpha, u) < \infty$ *for the three last ones.*

Proof. Let $\alpha_n \to \alpha$, with $u \notin U(\alpha)$. We need to prove that $t_n^p = t^p(\alpha_n, u)$ converges to $t^p = t^p(\alpha, u)$, and $\alpha_n(t_n^p) \to \alpha(t^p)$, $\alpha_n(t_n^p-) \to \alpha(t^p-)$ if $t^p < \infty$, and $\alpha_n^u \to \alpha^u$.

Suppose that $t_n^p \to t^p < \infty$. Then $\liminf_n t_n^{p+1} \geq t^p$. If $\liminf_n t_n^{p+1} = t^p$ there would exist a subsequence $t_{n_k}^{p+1} \to t^p$ and since $t_{n_k}^p \to t^p$ that would contradict 2.1b. Thus $\liminf_n t_n^{p+1} > t^p$.

Next, $\sup_{t \in I} |\Delta\alpha(t)| < u$ for each closed interval $I \subset (t^p, t^{p+1})$, because $u \notin U(\alpha)$. Then 2.5 implies that: $\limsup_n \sup_{t \in I} |\Delta\alpha_n(t)| < u$, which in turn implies that: $\liminf_n t_n^{p+1} \geq t^{p+1}$. Finally if we apply 2.1 it is easy to deduce from what precedes that $t_n^{p+1} \to t^{p+1}$, and that $\alpha_n(t_n^{p+1}) \to \alpha(t^{p+1})$ and $\alpha_n(t_n^{p+1}-) \to \alpha(t^{p+1}-)$ if $t^{p+1} < \infty$. Therefore we obtain by induction on p that, for all $p \in \mathbb{N}^*$, $t_n^p \to t^p$, and $\alpha_n(t_n^p) \to \alpha(t^p)$ and $\alpha_n(t_n^p-) \to \alpha(t^p-)$ if $t^p < \infty$.

Let us call $\alpha^{u,q}$ the function defined as in 2.6, but summing p from 1 to q only. Then what precedes, plus 2.2(b.6) and an induction on q, imply that $\alpha_n^{u,q} \to \alpha^{u,q}$ for all $q \geq 1$. Now, for each $N > 0$ there are q and n_0 large enough so that $t^q > N$, and $t_n^q > N$ for $n \geq n_0$; since $\alpha^{u,q} = \alpha^u$ on $[0, t^q)$ and $\alpha_n^{u,q} = \alpha_n^u$ on $[0, t_n^q)$, we easily deduce that $\alpha_n^u \to \alpha^u$ in $\mathbb{D}(\mathbb{R}^d)$. □

2.8 **Corollary.** *Let* $g\colon \mathbb{R}^d \to \mathbb{R}^{d'}$ *be continuous and vanishing on a neighbourhood of* 0, *and for* $\alpha \in \mathbb{D}(\mathbb{R}^d)$ *set* $\check{\alpha}^g(t) = \sum_{s \leq t} g(\Delta\alpha(s))$. *Then the map* $\alpha \rightsquigarrow (\alpha, \check{\alpha}^g)$ *is continuous from* $\mathbb{D}(\mathbb{R}^d)$ *into* $\mathbb{D}(\mathbb{R}^{d+d'})$.

Proof. Let $\alpha_n \rightsquigarrow \alpha$. Let $u \notin U(\alpha)$ be such that $g(x) = 0$ for $|x| \leq u$. Set

$$\check{\alpha}_n^{g,q}(t) = \sum_{1 \leq p \leq q} g(\Delta\alpha_n(t^p(\alpha_n, u))) 1_{\{t^p(\alpha_n, u) \leq t\}}$$

for $q \in \mathbb{N}^*$, and similarly for $\check{\alpha}^{g,q}$. Then 2.7 and Example 1.19 and an induction on q show that $\check{\alpha}_n^{g,q} \to \check{\alpha}^{g,q}$ in $\mathbb{D}(\mathbb{R}^{d'})$. Moreover $\check{\alpha}^g = \check{\alpha}^{g,q}$ on $[0, t^q(\alpha, u)]$, and similarly for $\check{\alpha}_n^{g,q}$, hence the same argument than in the end of the proof of 2.7 shows that $\check{\alpha}_n^g \to \check{\alpha}^g$ in $\mathbb{D}(\mathbb{R}^{d'})$. Finally since all jump times of $\check{\alpha}_n^g$ (resp. $\check{\alpha}^g$) are also jump times of α_n (resp. α), we deduce $(\alpha_n, \check{\alpha}_n^g) \to (\alpha, \check{\alpha}^g)$ in $\mathbb{D}(\mathbb{R}^{d+d'})$ from 2.2b. □

For the remainder of this subsection, we are interested in "stopping" the function α. This will be useful in Chapter IX. For $a \geq 0$, set

2.9 $$S_a(\alpha) = \inf(t\colon |\alpha(t)| \geq a \text{ or } |\alpha(t-)| \geq a).$$

2.10 **Lemma.** a) *For all* $a \geq 0$, *the map* $\alpha \rightsquigarrow S_a(\alpha)$ *is a strict stopping time for the filtration* $\mathbf{D}(\mathbb{R}^d)$ *defined in 1.1* (recall that it means that $\{S_a \leq t\} \in \mathscr{D}_t^0(\mathbb{R}^d)$ for all t: see III.2.35).

b) *For all* α, *the function* $a \rightsquigarrow S_a(\alpha)$ *is nondecreasing and càg.*

c) *The set* $V(\alpha) = \{a > 0: S_a(\alpha) < S_{a+}(\alpha)\}$ *is at most countable.*

d) *The set* $V'(\alpha) = \{a > 0: S_a(\alpha) \in J(\alpha)$ and $|\alpha(S_a(\alpha)-)| = a\}$ *is at most countable.*

Proof. a) If $M^\alpha(t) = \sup_{s \le t} |\alpha(s)|$, then $\alpha \rightsquigarrow M^\alpha(t)$ is obviously $\mathscr{D}_t^0(\mathbb{R}^d)$-measurable, and $t \rightsquigarrow M^\alpha(t)$ is nondecreasing and càd. Moreover, $S_a(\alpha) > t \Leftrightarrow M^\alpha(t) < a$, hence the claim.

b) That $a \rightsquigarrow S_a(\alpha)$ is non-decreasing is obvious; its left-hand continuity comes from $S_a(\alpha) = \inf(t: M^\alpha(t) \ge a)$ and from the right-continuity of $M^\alpha(\cdot)$, which implies $M^\alpha(S_a(\alpha)) \ge a$.

c) This is obvious, since $a \rightsquigarrow S_a(\alpha)$ has at most countably many discontinuities.

d) Let (t_n) be the sequence of all jump times of α. If $a \in V'(\alpha)$ then $S_a(\alpha) = t_n$ for some n, and $|\alpha(t_n -)| = a$: there are obviously at most countably many such numbers a. □

2.11 **Proposition.** *The function* $\alpha \rightsquigarrow S_a(\alpha)$ *is continuous at each point* α *such that* $a \notin V(\alpha)$ (notation 2.10c).

Proof. Let $\alpha_n \to \alpha$, with $a \notin V(\alpha)$. We use the notation $M^\alpha(t)$ of the previous proof. Then 2.4 yields $M^{\alpha_n}(t) \to M^\alpha(t)$ for all $t \notin J(\alpha)$.

If $t < S_a(\alpha)$, then $M^\alpha(t) < a$; if moreover $t \notin J(\alpha)$, $M^{\alpha_n}(t) < a$ as well for all n large enough, so $t < S_a(\alpha_n)$ as well. Since $\mathbb{R}_+ \setminus J(\alpha)$ is dense in $\mathbb{R}_+$, we deduce that $\liminf_n S_a(\alpha_n) \ge S_a(\alpha)$.

If $t > S_a(\alpha)$, then $t > S_{a+}(\alpha)$ because $a \notin V(\alpha)$, so $t \ge S_b(\alpha)$ for some $b > a$, and $M^\alpha(t) \ge b > a$; if moreover $t \notin J(\alpha)$, $M^{\alpha_n}(t) > a$ as well for all n large enough, so $S_a(\alpha_n) \le t$. We deduce that $\limsup_n S_a(\alpha_n) \le S_a(\alpha)$, and thus $S_a(\alpha_n) \to S_a(\alpha)$. □

2.12 **Proposition.** *To each* $\alpha \in \mathbb{D}(\mathbb{R}^d)$ *we associate the stopped function* α^{S_a} *defined by* $\alpha^{S_a}(t) = \alpha(t \wedge S_a(\alpha))$. *Then* $\alpha \rightsquigarrow (\alpha, \alpha^{S_a})$ *is continuous from* $\mathbb{D}(\mathbb{R}^d)$ *into* $\mathbb{D}(\mathbb{R}^{2d})$ *at each point* α *such that* $a \notin V(\alpha) \cup V'(\alpha)$.

Proof. Let $\alpha_n \to \alpha$, with $a \notin V(\alpha) \cup V'(\alpha)$.

a) The inequalities $\sup_{s \le N} |(\alpha_n, \alpha_n^{S_a})(s)| \le 2 \sup_{s \le N} |\alpha_n(s)|$ and $w((\alpha_n, \alpha_n^{S_a}); I) \le 2w(\alpha_n; I)$ for all intervals I are obvious, and thus $w'_N((\alpha_n, \alpha_n^{S_a}), \theta) \le 2w'_N(\alpha_n, \theta)$. Since the sequence $\{\alpha_n\}$ is relatively compact, a double application of 1.14b yields that the sequence $\{(\alpha_n, \alpha_n^{S_a})\}$ is also relatively compact, and so it remains to prove that its only limit point is (α, α^{S_a}). Hence, up to extracting a subsequence, we may assume that $(\alpha_n, \alpha_n^{S_a})$ converges to a limit, which obviously has the form (α, β), and we need to prove that $\beta = \alpha^{S_a}$.

b) By 2.11, $S_n := S_a(\alpha_n)$ converges to $S := S_a(\alpha)$. Then for all $t < s$ with $t \notin J(\alpha) \cup J(\beta)$, $\alpha_n^{S_a}(t)$ equals $\alpha_n(t)$ for all n large enough and thus converges to $\alpha(t) = \alpha^{S_a}(t)$, and also to $\beta(t)$ (apply 2.3): hence $\beta = \alpha$ on $[0, S)$. For all $t > S$

with $t \notin J(\beta)$, $\alpha_n^{S_a}(t)$ equals $\alpha_n(S_a)$ for all n large enough, and also converges to $\beta(t)$: hence $\beta(t) = \beta(S) = \lim_n \alpha_n(S_n)$ for all $t \geq S$ (because β is right-continuous).

c) So it remains to prove that $\alpha_n(S_n) \to \alpha(S)$ when $S < \infty$. If $S \notin J(\alpha)$, this immediately follows from 2.1(b.5) (recall that $S_n \to S$).

Finally, we assume that $S \in J(\alpha)$. By 2.1 there is a sequence $t_n \to S$ such that $\alpha_n(t_n) \to \alpha(S)$ and $\alpha_n(t_n-) \to \alpha(S-)$. Moreover, $a \notin V'(\alpha)$, so from the definition of $S = S_a(\alpha)$ we must have $|\alpha(S-)| < a \leq |\alpha(S)|$. Then if $t_{n_k} > S_{n_k}$ for an infinite subsequence, 2.1 yields that $|\alpha_{n_k}(S_{n_k})|$ and $|\alpha_{n_k}(S_{n_k}-)|$ both converge to $|\alpha(S-)| < a$; since by definition of S_n we have $|\alpha_n(S_n-)| \vee |\alpha_n(S_n)| \geq a$, this brings a contradiction. Hence we deduce that $t_n \leq S_n$ for all n large enough, in which case 2.1 again yields that $\alpha_n(S_n) \to \alpha(S)$, and we are finished. □

§ 2b. Increasing Functions and the Skorokhod Topology

In this subsection we study an important subset of $\mathbb{D}(\mathbb{R})$, namely:

2.13 $\mathscr{V}^+ =$ the set of all nonnegative càd non-decreasing functions on $\mathbb{R}_+$ null at 0.

Let us also consider the space $\mathscr{V}^{+,1}$ of all *counting functions*, that is functions having the form

2.14 $$\alpha(s) = \sum_{n \geq 1} 1_{\{t_n \leq s\}}$$

where (t_n) increases to $+\infty$, $t_1 > 0$, and $t_n < t_{n+1}$ if $t_n < \infty$. This is the space of sample paths of point processes, as introduced in §I.3b (see I.3.25). We have $\mathscr{V}^{+,1} \subset \mathscr{V}^+$.

Obviously, $\mathscr{V}^+$ and $\mathscr{V}^{+,1}$ are closed in $\mathbb{D}(\mathbb{R})$ for the Skorokhod topology. Our main result is the following.

2.15 **Theorem.** *Let $\alpha_n, \alpha \in \mathscr{V}^+$.*

a) *We have $\alpha_n \to \alpha$ for the Skorokhod topology if and only if there is a dense subset D of $\mathbb{R}_+$, such that*

2.16 $$t \in D \Rightarrow \alpha_n(t) \to \alpha(t)$$

2.17 $$t \in D \Rightarrow \sum_{0<s\leq t} |\Delta\alpha_n(s)|^2 \to \sum_{0<s\leq t} |\Delta\alpha(s)|^2.$$

Then, these conditions are fulfilled with $D = \mathbb{R}_+ \setminus J(\alpha)$.

b) *Let f be any function: $\mathbb{R}_+ \to \mathbb{R}_+$ which is strictly convex and meets $f(0) = f'(0) = 0$. Then $\alpha_n \to \alpha$ if and only if there is a dense subset D of $\mathbb{R}_+$, such that 2.16 and*

2.18 $$t \in D \Rightarrow \sum_{0<s\leq t} f(\Delta\alpha_n(s)) \to \sum_{0<s\leq t} f(\Delta\alpha(s))$$

hold. Then, these conditions are fulfilled with $D = \mathbb{R}_+ \setminus J(\alpha)$.

c) *For $\alpha_n \to \alpha$ it is sufficient that 2.16 holds for a dense subset D of $\mathbb{R}_+$ in the following two cases:*

(i) *α is continuous;*

(ii) *$\alpha_n, \alpha \in \mathscr{V}^{+,1}$.*

d) *Suppose that the sequence $\{\beta_n\} \subset \mathscr{V}^+$ is relatively compact for the Skorokhod topology, and that $\beta_n - \alpha_n \in \mathscr{V}^+$ for all n. Then*

(i) *the sequence (α_n) is relatively compact,*

(ii) *2.16 is sufficient to insure that $\alpha_n \to \alpha$.*

This theorem will be proved through a long string of lemmas. We introduce two other conditions first, in which $D \subset \mathbb{R}_+$.

2.19 $$r \in D, r' > r \Rightarrow \limsup_{(n)} \sup_{r<t\leq r'} \Delta\alpha_n(t) \leq \sup_{r<t\leq r'} \Delta\alpha(t).$$

2.20 If $t \geq 0$ there is a sequence (t_n) with:

(i) $t_n \to t$, and $t_n \leq t$ if $t \in D$,

(ii) $\Delta\alpha_n(t_n) \to \Delta\alpha(t)$.

2.21 **Lemma.** *If D is dense in $\mathbb{R}_+$, we have*: 2.16 $\Rightarrow$ 2.19.

Proof. Let $r \in D$, $r' > r$, and set $a := \limsup_{(n)} \sup_{r<t\leq r'} \Delta\alpha_n(t)$ and $b := \sup_{r<t\leq r'} \Delta\alpha(t)$. There are a sequence (n_k) and a sequence $\{s_k\} \subset (r, r']$ which converges to a limit $s \in [r, r']$ and such that $\Delta\alpha_{n_k}(s_k) \to a$.

We assume 2.16. $\varepsilon > 0$ being given, we may find $t, t' \in D$ with $r \leq t \leq s < t'$, and $t < s$ whenever $r < s$, such that $\alpha(t') - \alpha(t) \leq \varepsilon + \Delta\alpha(s)1_{\{r<s\}}$. We have $s_k \in (t, t']$ for k large enough, hence 2.16 gives

$$a = \lim_{(k)} \Delta\alpha_{n_k}(s_k) \leq \lim_{(k)} [\alpha_{n_k}(t') - \alpha_{n_k}(t)] = \alpha(t') - \alpha(t)$$

$$\leq \varepsilon + \Delta\alpha(s)1_{\{r<s\}} \leq \varepsilon + b$$

and since $\varepsilon > 0$ is arbitrary, we have $a \leq b$. □

2.22 **Lemma.** *Let f be a function like in 2.15b, and D be a dense subset of $\mathbb{R}_+$. If 2.16 holds, then 2.18 and 2.20 are equivalent.*

Proof. a) We suppose first 2.18. In order to check 2.20 it clearly suffices to examine the case $t > 0$ with $a := \Delta\alpha(t) > 0$. For each $m \in \mathbb{N}^*$ there exist $s_m, s'_m \in D$ with $s_m < t \leq s'_m$, and $s'_m = t$ if $t \in D$, and $s'_m - s_m \leq \dfrac{1}{m}$, and $\alpha(s'_m) - \alpha(s_m) \leq a + 1/m$; moreover, we can assume that (s_m) (resp. (s'_m)) is non-decreasing (resp. non-increasing).

Let $F_m = (s_m, s'_m]$. We choose a time $r_n(m) \in F_m$ which realizes the maximum of $\Delta\alpha_n(s)$ for $s \in F_m$. Set

$$\underline{a}_m = \liminf_n \Delta\alpha_n(r_n(m)), \quad \bar{a}_m = \limsup_n \Delta\alpha_n(r_n(m))$$

$$\underline{a} = \lim_m \downarrow \underline{a}_m, \qquad \bar{a} = \lim_m \downarrow \bar{a}_m$$

(the two sequences $\underline{a}_m$ and $\bar{a}_m$ are non-increasing, because F_m itself is so). Then $\bar{a} \le \bar{a}_m \le a + 1/m$, hence $\bar{a} \le a$. Let also $K > 0$ such that $f(x) \le Kx^2$ for $0 \le x \le a + 1$. Then 2.18 gives

$$f(a) \le \sum_{r \in F_m} f(\Delta\alpha(r)) = \lim_n \sum_{r \in F_m} f(\Delta\alpha_n(r)),$$

hence

$$f(a) \le K\underline{a}_m \limsup_n \sum_{r \in F_m} \Delta\alpha_n(r) \le K\underline{a}_m\left(a + \frac{1}{m}\right)$$

$$\begin{aligned} f(a) &\le \liminf_n f(\Delta\alpha_n(r_n(m))) + \limsup_n f\left(\sum_{r \in F_m, r \neq r_n(m)} \Delta\alpha_n(r)\right) \\ &\le f(\underline{a}_m) + \limsup_n f(\alpha_n(s'_m) - \alpha_n(s_m) - \Delta\alpha_n(r_n(m))) \\ &\le f(\underline{a}_m) + f\left(a + \frac{1}{m} - \underline{a}_m\right) \end{aligned}$$

(we apply first that f is convex, then that it is increasing and continuous). Letting $n \uparrow \infty$ we obtain

$$f(a) \le K\underline{a}a, \quad f(a) \le f(\underline{a}) + f(a - \underline{a}).$$

But f is strictly convex, so $f(a) \le f(\underline{a}) + f(a - \underline{a})$ implies that either $\underline{a} = 0$, or $\underline{a} = a$; since $\underline{a} = 0$ contradicts $f(a) \le K\underline{a}a$, we have $\underline{a} = a$, hence $\bar{a} = \underline{a} = a$. Thus for each $q \in \mathbb{N}^*$ there is $k(q) \ge q$ such that

$$a - 1/q \le \underline{a}_{k(q)} \le \bar{a}_{k(q)} \le a + 1/q$$

and there exists $l(q) \ge q$ such that $|a - \Delta\alpha_n(r_n(k(q)))| \le 2/q$ for $n \ge l(q)$. It remains to set $m(n) = \sup(q\colon n \ge l(q))$ and $t_n = r_n(k(m(n)))$ in order to obtain $t_n \to t$ and $\Delta\alpha_n(t_n) \to a = \Delta\alpha(t)$.

b) Conversely, we suppose 2.20. Let $\varepsilon > 0$, $t \in D$. Set $s^0 = 0, \ldots, s^{p+1} = \inf(t > s^p\colon \Delta\alpha(t) > \varepsilon), \ldots$, and $q = \sup(p\colon s^p \le t)$. To each s^i we associate a sequence (s^i_n) satisfying 2.20. Set

$$\alpha'(s) = \alpha(s) - \sum_{i \le q} \Delta\alpha(s^i) 1_{\{s^i \le s\}}$$

$$\alpha'_n(s) = \alpha_n(s) - \sum_{i \le q} \Delta\alpha_n(s^i_n) 1_{\{s^i_n \le s\}}.$$

It is obvious that $\alpha'_n(s) \to \alpha'(s)$ for $s \in D$, then 2.21 and the definition of α' imply that $\limsup_n \sup_{s \le t} \Delta\alpha'_n(s) \le \varepsilon$. Therefore

$$
\begin{aligned}
\limsup_n \left| \sum_{0<s\le t} f(\Delta\alpha_n(s)) - f(\Delta\alpha(s)) \right| &\le \lim_n \sum_{i\le q} |f(\Delta\alpha_n(s_n^i)) - f(\Delta\alpha(s^i))| \\
&\quad + \limsup_n \sum_{0<s\le t} [f(\Delta\alpha'_n(s)) + f(\Delta\alpha'(s))] \\
&\le 0 + \limsup_n g\left(\sup_{0<s\le t} \Delta\alpha'_n(s)\right)\alpha'_n(t) \\
&\quad + g\left(\sup_{0<s\le t} \Delta\alpha'(s)\right)\alpha'(t) \\
&\le 2g(\varepsilon)\alpha'(t)
\end{aligned}
$$

where $g(x) = \sup_{y\le x} \dfrac{f(y)}{y}$ is continuous and $g(x) \downarrow 0$ as $x \downarrow 0$ (because $f(0) = f'(0) = 0$). Since $\varepsilon > 0$ is arbitrary, 2.18 follows. □

2.23 **Lemma.** *If D is dense in $\mathbb{R}_+$, 2.16 implies that*

$$\limsup_{(n)} \sup_{s\le t} |\alpha_n(s) - \alpha(s)| \le 2 \sup_{s\le t} \Delta\alpha(s).$$

Proof. Let $\varepsilon > 0$. We can assume that $0 \in D$. We may find $t_i \in D$ such that $0 = t_0 < \cdots < t_{r-1} \le t < t_r$ and $\alpha(t_i) - \alpha(t_{i-1}) \le u + \varepsilon$, where $u = \sup_{s\le t} \Delta\alpha(s)$. Since α_n and α are increasing, we have for $s \in [t_{i-1}, t_i]$:

$$
\begin{aligned}
|\alpha_n(s) - \alpha(s)| &\le |\alpha_n(t_i) - \alpha(t_{i-1})| + |\alpha_n(t_{i-1}) - \alpha(t_i)| \\
&\le |\alpha_n(t_i) - \alpha(t_i)| + |\alpha_n(t_{i-1}) - \alpha(t_{i-1})| + 2|\alpha(t_i) - \alpha(t_{i-1})|.
\end{aligned}
$$

Since $\alpha_n(t_i) \to \alpha(t_i)$ for all i, and $\varepsilon > 0$ is arbitrary, we easily deduce the result. □

2.24 **Corollary.** *Let $\alpha_n, \alpha \in \mathscr{V}^+$ with α continuous. Let D be a dense subset of $\mathbb{R}_+$. Then*: 2.16 $\Rightarrow$ 2.17.

Proof. If 2.16 holds, the previous lemma implies that α_n tends to α locally uniformly. Hence $\sup_{s\le t} \Delta\alpha_n(s) \to 0$ for all t, and $\sum_{s\le t} |\Delta\alpha_n(s)|^2 \le \alpha_n(t) \sup_{s\le t} \Delta\alpha_n(s)$ also goes to 0: thus 2.17 holds. □

2.25 **Lemma.** *If 2.16 and 2.20 hold for a dense subset D of $\mathbb{R}_+$, the sequence (α_n) is relatively compact for the Skorokhod topology.*

Proof. We will prove that $A = \{\alpha_n\}_{n\ge1}$ satisfies 1.16, in which we can of course assume that $N \in D$ (instead of $N \in \mathbb{N}^*$). That 1.16(i) holds follows from 2.16 and from the fact that each α_n is increasing. We can again assume that $0 \in D$.

Now let $\varepsilon > 0$ and consider the notation (s^i, s_n^i), α', α'_n introduced in part (b) of the proof of Lemma 2.22 with $q = \sup(i: s^i \le N)$. Since $\Delta\alpha' \le \varepsilon$ one may find a subdivision $0 = t_0 < \cdots < t_r = N$ with $t_i \in D$ and

2.26 $$\alpha'(t_i) - \alpha'(t_{i-1}) \le \varepsilon/2.$$

We consider the subdivision $0 = v_0 < \cdots < v_p = N < v_{p+1} = s^{q+1}$ consisting of all points $(t_i)_{0 \le i \le r}$ and $(s^i)_{1 \le i \le q+1}$. We set $\theta = \inf_{i \le p+1}(v_i - v_{i-1})$.

We have seen in 2.22 that $s_n^i \to s^i$ and that (α'_n, α') satisfy 2.16, so there exists $n_0 \in \mathbb{N}$ such that:

2.27 $$n \ge n_0 \Rightarrow |\alpha'_n(t_i) - \alpha'(t_i)| \le \varepsilon \quad \text{for } i \le r, |s_n^i - s^i| \le \frac{\theta}{4} \text{ for } i \le q+1.$$

Now we fix $n \ge n_0$ and we associate the subdivision $0 = v_0^n < \cdots < v_p^n = N$ by $v_i^n = s_n^j$ if $v_i = s^j$, and $v_i^n = v_i$ otherwise. From 2.27 it follows that $v_i^n - v_{i-1}^n \ge \theta/2$ for all $i \le p$. From 2.27 again, from 2.26 and from the fact that α'_n is increasing, we have $w(\alpha'_n; [v_{i-1}^n, v_i^n)) \le 4\varepsilon$ and it follows from the definition of α'_n that we also have (for $n \ge n_0$): $w(\alpha_n; [v_{i-1}^n, v_i^n)) \le 4\varepsilon$. Thus

$$n \ge n_0 \Rightarrow w'_N\left(\alpha_n, \frac{\theta}{2}\right) \le 4\varepsilon.$$

Now, the finite family $(\alpha_n)_{n \le n_0}$ is relatively compact, so by 1.16 there is $\theta' > 0$ such that $w'_N(\alpha_n, \theta') \le 4\varepsilon$ for all $n \le n_0$. Then if $\theta'' = \theta' \wedge \frac{\theta}{2}$ we get

$$\sup_{(n)} w'_N(\alpha_n, \theta'') \le 4\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, A satisfies condition 1.16ii, and the proof is finished. □

Proof of Theorem 2.15. (a) and (b): of course, (a) is a particular case of (b) (with $f(x) = x^2$). Assume that 2.16 and 2.18 hold for some D dense in $\mathbb{R}_+$ and containing 0. Lemmas 2.22 and 2.25 imply that the sequence (α_n) is relatively compact in $\mathbb{D}(\mathbb{R})$; moreover, 2.16 clearly implies that α is the only possible limit point of this sequence, hence $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R})$.

Conversely, suppose that $\alpha_n \to \alpha$ in $\mathbb{D}(\mathbb{R})$, and put $D = \mathbb{R}_+ \setminus J(\alpha)$. Then 2.1 gives conditions 2.16 and 2.20, and 2.18 follows from 2.22.

(c) When α is continuous, we have seen in 2.24 that 2.16 implies 2.17. When $\alpha_n, \alpha \in \mathscr{V}^{+,1}$ we have

$$\alpha(t) = \sum_{0 < s \le t} \Delta\alpha(s)^2$$

and similarly for α_n, hence 2.16 and 2.17 are equivalent.

(d) The assumption $\beta_n - \alpha_n \in \mathscr{V}^+$ obviously yields that $|\alpha_n(t)| \le |\beta_n(t)|$ and that $w'_N(\alpha_n, \theta) \le w'_N(\beta_n, \theta)$. Thus the statement (i) immediately follows from 1.14b, and (ii) is proved exactly like in (a) above. □

3. Weak Convergence

In this section we remind the reader with some basic facts about weak convergence of measures and convergence in law for variables and processes. All unproved statements may be found in [12] or [196] for example, or in many other text-books.

§ 3a. Weak Convergence of Probability Measures

1. Let us be given a Polish space E, with its Borel σ-field $\mathscr{E}$. We denote by $\mathscr{P}(E)$ the space of all *probability measures* on $(E, \mathscr{E})$.

We endow $\mathscr{P}(E)$ with the *weak topology*: this is the coarsest topology for which the mappings: $\mu \rightsquigarrow \mu(f)$ are continuous for all bounded continuous functions f on E. Then $\mathscr{P}(E)$ itself is a *Polish space* for this topology.

If $\mu_n \to \mu$ weakly (or: in $\mathscr{P}(E)$), not only do we have $\mu_n(f) \to \mu(f)$ for every bounded continuous f, but:

3.1 If $\mu_n \to \mu$ in $\mathscr{P}(E)$ and if F is a closed set in E, then

$$\limsup_{(n)} \mu_n(F) \le \mu(F). \qquad \square$$

3.2 If $\mu_n \to \mu$ in $\mathscr{P}(E)$ and if f is a bounded function on E that is μ-a. s. continuous, then $\mu_n(f) \to \mu(f)$. $\square$

Let E' be another Polish space, and $h: E \to E'$. We denote by $\mu \circ h^{-1}$ the image of $\mu \in \mathscr{P}(E)$ by h; we have $\mu \circ h^{-1} \in \mathscr{P}(E')$. And:

3.3 The mapping: $\mu \rightsquigarrow \mu \circ h^{-1}$ is continuous from $\mathscr{P}(E)$ into $\mathscr{P}(E')$ at each point μ such that h is μ-a. s. continuous. $\square$

A *convergence-determining class* is a set $\mathscr{H}$ of continuous bounded functions on E with the following property: if μ_n, $\mu \in \mathscr{P}(E)$ and if $\mu_n(h) \to \mu(h)$ for all $h \in \mathscr{H}$, then $\mu_n \to \mu$ in $\mathscr{P}(E)$.

3.4 There exists a *countable* convergence-determining class on E (for example if $E = \mathbb{R}^d$, $\mathscr{H} = \{e^{iu \cdot x}\colon u \in \mathbb{Q}^d\}$ is convergence-determining). $\square$

Finally, recall that a subset A of $\mathscr{P}(E)$ is called *tight* if for every $\varepsilon > 0$ there exists a compact subset K in E such that $\mu(E \setminus K) \le \varepsilon$ for all $\mu \in A$. Then the famous and essential Prokhorov's Theorem reads as:

3.5 *A subset A of $\mathscr{P}(E)$ is relatively compact* (for the weak topology) *if and only if it is tight.*

3.6 **Remark.** We shall occasionally need to consider the space $\mathscr{M}^+(E)$ of all *positive finite measures* on $(E, \mathscr{E})$. This space is also Polish for the weak topology, and 3.5 remains valid, provided we add in the definition of a tight subset A that $\sup_{\mu \in A} \mu(E) < \infty$. □

2. Now we consider random variables. Let X be an E-valued random variable on some triple $(\Omega, \mathscr{F}, P)$. Then $P \circ X^{-1}$, the image of P under X, belongs to $\mathscr{P}(E)$; it is called the *law*, or the *distribution*, of X and it is also denoted by $\mathscr{L}(X)$, or by $\mathscr{L}(X|P)$ if there may be some ambiguity as to the measure P (as we have already done in § V.1a).

Consider now a sequence (X^n) of E-valued random variables: they may be defined on different probability spaces, say X^n on $(\Omega^n, \mathscr{F}^n, P^n)$. The law $\mathscr{L}(X^n)$ is of course $\mathscr{L}(X^n|P^n) = (P^n) \circ (X^n)^{-1}$. We say that (X^n) *converges in law* (or in distribution) *to* X, and we write

3.7
$$X^n \xrightarrow{\mathscr{L}} X,$$

if $\mathscr{L}(X^n) \to \mathscr{L}(X)$ weakly in $\mathscr{P}(E)$. This is equivalent to saying that $E_{P^n}(f(X^n)) \to E_P(f(X))$ for all bounded continuous functions f on E, where E_{P^n} denotes the expectation with respect to P^n. We can obviously transpose all previous results in terms of random variables. For instance, 3.2 and 3.3 have the following translation:

3.8 Assume that $X^n \xrightarrow{\mathscr{L}} X$ and that $P(X \in C) = 1$, where C is the continuity set of the function $h: E \to E'$. Then

(i) if $E' = \mathbb{R}$ and h is bounded, then $E_{P^n}(h(X^n)) \to E_P(h(X))$;

(ii) if E' is Polish, then $h(X^n) \xrightarrow{\mathscr{L}} h(X)$.

(Note that C is not necessarily Borel in E, but the (possibly nonmeasurable) set $\{\omega: X(\omega) \notin C\}$ has to be P-negligible). □

Finally, in this book we say that *the sequence* (X^n) (or $(X^n|P^n)$ if there is ambiguity: see V.1.2) *is tight* if the sequence of distributions $\mathscr{L}(X^n)$ is tight (i.e., if for every $\varepsilon > 0$ there is a compact subset K of E such that $P^n(X^n \notin K) \leq \varepsilon$ for all n). Then 3.5 implies:

3.9 *The sequence* $\{\mathscr{L}(X^n)\}$ *is relatively compact in* $\mathscr{P}(E)$ *if and only if the sequence* (X^n) *is tight.*

§ 3b. Application to Càdlàg Processes

1. In this subsection we consider only $\mathbb{R}^d$-valued càdlàg processes. Let X be such a process, defined on a triple $(\Omega, \mathscr{F}, P)$. Then it may be considered as a random variable taking its values in the Polish space $\mathbb{D}(\mathbb{R}^d)$, supposedly equipped with Skorokhod topology. Consequently its law $\mathscr{L}(X)$ is an element of $\mathscr{P}(\mathbb{D}(\mathbb{R}^d))$.

By analogy with 1.7 and 2.6 we set

3.10 $$\begin{cases} J(X) = \{t \geq 0 \colon P(\Delta X_t \neq 0) > 0\} \\ U(X) = \{u > 0 \colon P(|\Delta X_t| = u \text{ for some } t > 0) > 0\}. \end{cases}$$

3.11 $$T_0(X,u) = 0, \ldots, T_{p+1}(X,u) = \inf(t > T_p(X,u) \colon |\Delta X_t| > u).$$

3.12 **Lemma.** *The sets $J(X)$ and $U(X)$ are at most countable* (this generalizes II.4.2).

Proof. For any $\mathbb{R}$-valued random variable Z, the set $\{t \colon P(Z = t) > 0\}$ is at most countable. Then the result follows from the identities:

$$J(X) = \bigcup_{n,p \geq 1} \{t \colon P(T_p(X, 1/n) = t) > 0\}$$

$$U(X) = \bigcup_{n,p \geq 1} \{u \colon P(|\Delta X_{T_p(X,1/n)}| = u,\ T_p(X, 1/n) < \infty) > 0\}. \qquad \square$$

Now we consider a sequence (X^n) of $\mathbb{R}^d$-valued càdlàg processes, each X^n being defined on some space $(\Omega^n, \mathscr{F}^n, P^n)$. Accordingly to 3.7, write

$$X^n \xrightarrow{\mathscr{L}} X$$

if $\mathscr{L}(X^n) \to \mathscr{L}(X)$ in $\mathscr{P}(\mathbb{D}(\mathbb{R}^d))$. If D is a subset of $\mathbb{R}_+$, we also write:

3.13 $$X^n \xrightarrow{\mathscr{L}(D)} X \quad \text{if } (X^n_{t_1}, \ldots, X^n_{t_k}) \xrightarrow{\mathscr{L}} (X_{t_1}, \ldots, X_{t_k}), \quad \forall t_i \in D,\ k \in \mathbb{N}^*$$

for the *finite-dimensional convergence along D*.

Of course, most of Section 2 has a counterpart in terms of convergence in law, via 3.8. As an illustration, let us state four among the most useful results:

3.14 **Proposition.** *If $X^n \xrightarrow{\mathscr{L}} X$, then $X^n \xrightarrow{\mathscr{L}(D)} X$ for $D = \mathbb{R}_+ \setminus J(X)$* (but not for $D = \mathbb{R}_+$, in general; apply 2.3).

3.15 **Proposition.** *If $X^n \xrightarrow{\mathscr{L}} X$, then for all $u \notin U(X)$, $k \geq 1$, we have*

$$(g(T_i(X^n, u), X^n_{T_i(X^n,u)}, \Delta X^n_{T_i(X^n,u)}))_{1 \leq i \leq k} \xrightarrow{\mathscr{L}} (g(T_i(X,u), X_{T_i(X,u)}, \Delta X_{T_i(X,u)}))_{1 \leq i \leq k}$$

where g is any function on $[0,\infty] \times \mathbb{R}^d \times \mathbb{R}^d$ that is continuous and satisfies $g(\infty, x, y) = 0$. (Apply 2.7).

3.16 **Proposition.** *If $X^n \xrightarrow{\mathscr{L}} X$ and if g is a continuous function on $\mathbb{R}^d$, vanishing in a neighbourhood of 0, then the processes $(X^n, \sum_{s \leq \cdot} g(\Delta X^n_s))$ converge in law to the process $(X, \sum_{s \leq \cdot} g(\Delta X_s))$* (Apply 2.8).

3.17 **Proposition.** *If* $X^n \xrightarrow{\mathscr{L}} X$ *and if* β *is a continuous* $\mathbb{R}^d$*-valued function on* $\mathbb{R}_+$, *then* $X^n + \beta \xrightarrow{\mathscr{L}} X + \beta$ (apply 1.23).

This may be false if β is discontinuous.

Now, consider the question of proving that $X^n \xrightarrow{\mathscr{L}} X$. The most common method, initiated by Prokhorov, goes through the following procedure:

3.18 (i) prove (X^n) is tight (or: $\{\mathscr{L}(X^n)\}$ is relatively compact in $\mathscr{P}(\mathbb{D}(\mathbb{R}^d))$),

(ii) prove that $\mathscr{L}(X)$ is the only possible limit point for the sequence $\{\mathscr{L}(X^n)\}$. □

(3.18 is actually necessary and sufficient for: $X^n \xrightarrow{\mathscr{L}} X$). For proving (ii) there are several different methods, one of these being based upon:

3.19 **Lemma.** *Let* D *be a dense subset of* $\mathbb{R}_+$, *and* X, Y *be two càdlàg processes satisfying* $\mathscr{L}(X_{t_1},\dots,X_{t_k}) = \mathscr{L}(Y_{t_1},\dots,Y_{t_k})$ *for all* $t_i \in D$, $k \in \mathbb{N}^*$. *Then* $\mathscr{L}(X) = \mathscr{L}(Y)$.

Proof. Since D is dense, $\mathscr{D}(\mathbb{R}^d)$ is generated by the mappings: $\alpha \rightsquigarrow \alpha(t)$ for $t \in D$. Then a monotone class argument yields the result. □

So in place of 3.18 one may go through the following:

3.20 (i) prove (X^n) is tight,

(ii) prove $X^n \xrightarrow{\mathscr{L}(D)} X$ for some dense subset D of $\mathbb{R}_+$, □

which again is necessary and sufficient for $X^n \xrightarrow{\mathscr{L}} X$ to hold. Of course (ii) above is not the only way of identifying the limit: as a matter of fact (ii) is achievable only when one has some grasp on the finite-dimensional distributions, which is not often! In Chapter IX (see also VII.1) we shall see another way, through "martingale problems".

Now we turn to tightness. Here we derive some general criteria, while in Sections 4 and 5 we will exhibit more specific results.

3.21 **Theorem.** *The sequence* (X^n) *is tight if and only if*

(i) *for all* $N \in \mathbb{N}^*$, $\varepsilon > 0$, *there are* $n_0 \in \mathbb{N}^*$ *and* $K \in \mathbb{R}_+$ *with*

3.22 $$n \geq n_0 \Rightarrow P^n\Big(\sup_{t \leq N} |X^n_t| > K\Big) \leq \varepsilon;$$

(ii) *for all* $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$, *there are* $n_0 \in \mathbb{N}^*$ *and* $\theta > 0$ *with*

3.23 $$n \geq n_0 \Rightarrow P^n(w'_N(X^n, \theta) \geq \eta) \leq \varepsilon.$$

Then these conditions also hold with $n_0 = 1$.

(Note that these conditions, though expressed in terms of the P^n's, in fact only depend on the laws $\mathscr{L}(X^n)$; recall that w'_N is defined in 1.8).

Proof. Necessary condition. Let $\varepsilon > 0$. We deduce from Prokhorov Theorem 3.9 the existence of a compact subset $\tilde{K}$ in $\mathbb{D}(\mathbb{R}^d)$ such that $P^n(X^n \notin \tilde{K}) \leq \varepsilon$ for all n. Now we apply 1.14b: let $N \in \mathbb{N}^*$, $\eta > 0$. Then $K := \sup_{t \leq N, \alpha \in \tilde{K}} |\alpha(t)|$ is finite, and there exists $\theta > 0$ with $\sup_{\alpha \in \tilde{K}} w'_N(\alpha, \theta) \leq \eta$, and thus

$$3.24 \qquad P^n\left(\sup_{t \leq N} |X_t^n| > K\right) \leq \varepsilon, \quad P^n(w'_N(X^n, \theta) > \eta) \leq \varepsilon$$

for all n, thus implying (i) and (ii) with $n_0 = 1$.

Sufficient condition. We suppose that (i) and (ii) hold. The finite family $(X^n)_{1 \leq n \leq n_0}$ being tight, it follows from above that it satisfies 3.24 for some $K' < \infty$ and $\theta' > 0$. Hence, replacing K by $K \vee K'$ and θ by $\theta \wedge \theta'$, we obtain that (i) and (ii) hold with $n_0 = 1$.

Fix $\varepsilon > 0$, $N \in \mathbb{N}^*$. Let $K_{N\varepsilon} < \infty$ and $\theta_{N\varepsilon k} > 0$ satisfy

$$\sup_{(n)} P^n\left(\sup_{t \leq N} |X_t^n| > K_{N\varepsilon}\right) \leq \frac{\varepsilon}{2} 2^{-N}$$

$$\sup_{(n)} P^n(w'_N(X^n, \theta_{N\varepsilon k}) > 1/k) \leq \frac{\varepsilon}{2} 2^{-N-k}.$$

Then $A_{N\varepsilon} = \{\alpha \in \mathbb{D}(\mathbb{R}^d) \colon \sup_{t \leq N} |\alpha(t)| \leq K_{N\varepsilon}, w'_N(\alpha, \theta_{N\varepsilon k}) \leq 1/k \text{ for all } k \in \mathbb{N}^*\}$ satisfies for all N:

$$P^n(X^n \notin A_{N\varepsilon}) \leq P^n\left(\sup_{t \leq N} |X_t^n| > K_{N\varepsilon}\right) + \sum_{k \geq 1} P^n\left(w'_N(X^n, \theta_{N\varepsilon k}) > \frac{1}{k}\right) \leq \varepsilon 2^{-N}.$$

Therefore $A_\varepsilon = \bigcap_{N \geq 1} A_{N\varepsilon}$ satisfies: $\sup_{(n)} P^n(X^n \notin A_\varepsilon) \leq \varepsilon$. On the other hand, A_ε satisfies 1.16 by construction, so it is relatively compact in $\mathbb{D}(\mathbb{R}^d)$. This being true for all $\varepsilon > 0$, Prokhorov Theorem 3.9 implies tightness of (X^n). □

The following property arises often enough to deserve a name of its own:

3.25 **Definition.** A sequence (X^n) of processes is called *C-tight* if it is tight, and if all limit points of the sequence $\{\mathscr{L}(X^n)\}$ are laws of continuous processes (i.e.: if a subsequence $\{\mathscr{L}(X^{n_k})\}$ converges to a limit P in $\mathscr{P}(\mathbb{D}(\mathbb{R}^d))$, then P charges only the set $\mathbb{C}(\mathbb{R}^d)$). □

3.26 **Proposition.** *There is equivalence between*

(i) *the sequence* (X^n) *is C-tight.*

(ii) *Condition 3.21.i holds, and for all* $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$ *there are* $n_0 \in \mathbb{N}^*$ *and* $\theta > 0$ *with* (recall 1.4 for the definition of w_N):

$$3.27 \qquad n \geq n_0 \Rightarrow P^n(w_N(X^n, \theta) > \eta) \leq \varepsilon.$$

(iii) *The sequence* (X^n) *is tight, and for all* $N \in \mathbb{N}^*$, $\varepsilon > 0$ *we have*

$$3.28 \qquad \lim_n P^n\left(\sup_{t \leq N} |\Delta X_t^n| > \varepsilon\right) = 0.$$

Proof. (i) $\Rightarrow$ (iii): Under (i), the sequence (X^n) is tight, so it suffices to prove 3.28 for any convergent subsequence. So we may assume that $X^n \xrightarrow{\mathcal{L}} X$, with X some continuous process. Then 3.8 and 2.4 imply that $\sup_{t\le T}|\Delta X_t^n| \xrightarrow{\mathcal{L}} \sup_{t\le T}|\Delta X_t|$ for all $T \notin J(X)$. Since X is continuous, $J(X) = \varnothing$ and $\sup_{t\le T}|\Delta X_t| = 0$, so we have 3.28.

(iii) $\Rightarrow$ (ii): This implication follows from 3.21 and from the following, easy, inequality:

3.29 $$w_N(\alpha, \theta) \le 2w'_N(\alpha, \theta) + \sup_{t\le N} |\Delta\alpha(t)|.$$

(ii) $\Rightarrow$ (i): 1.9 and 3.21 yield that the sequence (X^n) is tight. It remains to prove that if a subsequence, still denoted by (X^n), converges in law to some X, then X is a. s. continuous. But $\sup_{t\le N}|\Delta\alpha(t)| \le w_N(\alpha, \theta)$ trivially, so 3.27 implies: $\sup_{t\le N}|\Delta X_t^n| \xrightarrow{\mathcal{L}} 0$. We have also been above that: $\sup_{t\le s}|\Delta X_t^n| \xrightarrow{\mathcal{L}} \sup_{t\le s}|\Delta X_t|$ for all $s \notin J(X)$. Therefore $\sup_{t\le N}|\Delta X_t| = 0$ a. s. for all $N \in \mathbb{N}^*$, and this obviously implies that X is a. s. continuous. □

2. We shall now derive a criterion for tightness: it looks very technical and it is rather superficial, but it is also very useful. The situation is as above: for each n we have a probability space $(\Omega^n, \mathcal{F}^n, P^n)$; all processes defined on this space show a superscript "n" and are all $\mathbb{R}^d$-valued.

Our first remark is quite superficial. Suppose that

3.30 $$\forall N > 0, \quad \forall \varepsilon > 0, \quad \lim_{(n)} P^n\left(\sup_{t\le N} |Z_s^n| > \varepsilon\right) = 0.$$

Then certainly the sequence (Z^n) converges in law to the process $Z = 0$. More generally we have:

3.31 **Lemma.** *If the sequence (Z^n) satisfies 3.30 and if the sequence (Y^n) is tight (resp. converges in law to Y), then the sequence $(Y^n + Z^n)$ is tight (resp. converges in law to Y).*

Proof. Suppose that the statement concerning tightness has been proved, and that $Y^n \xrightarrow{\mathcal{L}} Y$. Then 3.14 yields $Y^n \xrightarrow{\mathcal{L}(D)} Y$, where $D = \mathbb{R}_+ \setminus J(Y)$, and 3.30 implies that $Y^n + Z^n \xrightarrow{\mathcal{L}(D)} Y$ as well. So using the procedure 3.20, we see that $Y^n + Z^n \xrightarrow{\mathcal{L}} Y$.

That $(Y^n + Z^n)$ is tight is easily proved directly; it is also a consequence of the next lemma, with $U^{nq} = Y^n$, $V^{nq} = 0$, $W^{nq} = Z^n$. □

3.32 **Lemma.** *Suppose that for all $n, q \in \mathbb{N}^*$ we have a decomposition*

$$X^n = U^{nq} + V^{nq} + W^{nq}$$

with (i) *the sequences $(U^{nq})_{n\ge1}$ are tight;*

(ii) *the sequences $(V^{nq})_{n\ge1}$ are tight and there is a sequence (a_q) of real numbers with*: $\lim_q a_q = 0$, $\lim_n P^n(\sup_{t\le N}|\Delta V_t^{nq}| > a_q) = 0 \ \forall N \in \mathbb{N}^*$.

(iii) *for all* $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\lim_q \limsup_n P^n(\sup_{t \leq N} |W_t^{nq}| > \varepsilon) = 0$.
Then, the sequence (X^n) *is tight.*

Proof. That (X^n) satisfies condition 3.21i is trivial, as well as the two following inequalities:

$$w_N(\alpha, \theta) \leq 2 \sup_{t \leq N} |\alpha(t)|$$

$$w'_N(\alpha + \beta, \theta) \leq w'_N(\alpha, \theta) + w_N(\beta, 2\theta).$$

Then, these inequalities and 3.29 give:

$$\begin{aligned} w'_N(X^n, \theta) &\leq w'_N(U^{nq} + V^{nq}, \theta) + w_N(W^{nq}, 2\theta) \\ &\leq w'_N(U^{nq}, \theta) + 2w'_N(V^{nq}, 2\theta) + \sup_{t \leq N} |\Delta V_t^{nq}| + 2 \sup_{t \leq N} |W_t^{nq}|. \end{aligned}$$

Let $\varepsilon > 0$, $\eta > 0$. We choose q so that: $\limsup_n P^n(\sup_{t \leq N} |W_t^{nq}| > \eta) \leq \varepsilon$ and $a_q \leq \eta$. Then, applying 3.21ii, we may choose $n_0 \in \mathbb{N}^*$ and $\theta > 0$ such that

$$n \geq n_0 \Rightarrow \begin{cases} P^n(w'_N(U^{nq}, \theta) > \eta) \leq \varepsilon, & P^n(w'_N(V^{nq}, 2\theta) > \eta) \leq \varepsilon, \\ P^n\left(\sup_{t \leq N} |W_t^{nq}| > \eta\right) \leq 2\varepsilon, & P^n\left(\sup_{t \leq N} |\Delta V_t^{nq}| > 2\eta\right) \leq \varepsilon, \end{cases}$$

Then $P^n(w'_N(X^n, \theta) > 7\eta) \leq 5\varepsilon$ for $n \geq n_0$, and we deduce that (X^n) satisfies 3.21ii. □

3.33 **Corollary.** *Let* (Y^n) *be a C-tight sequence of d-dimensional processes; let* (Z^n) *be a tight (resp. C-tight) sequence of d'-dimensional processes.*

a) *If* $d = d'$, *then* $(Y^n + Z^n)$ *is tight (resp. C-tight).*

b) *the sequence* $\{(Y^n, Z^n)\}$ *of* $(d + d')$*-dimensional processes is tight (resp. C-tight).*

Proof. a) It suffices to apply the preceding lemma with $U^{nq} = Z^n$, $V^{nq} = Y^n$, $a_q = 1/q$, and $W^{nq} = 0$, and to use 3.26.

b) With obvious notation, $w_N((0, Z^n), \theta) = w_N(Z^n, \theta)$ and $w'_N((0, Z^n), \theta) = w'_N(Z^n, \theta)$, so 3.21 and 3.26 imply that if (Z^n) is tight (resp. C-tight), then so is the sequence of $(d + d')$-dimensional processes $\{(0, Z^n)\}$. The same holds for $\{(Y^n, 0)\}$, and the claim follows from (a) applied to $(Y^n, Z^n) = (Y^n, 0) + (0, Z^n)$. □

3. *Increasing processes.* When we restrict our attention to increasing processes, we may obtain some nice and simple results. Recall that "increasing process" means: nonnegative, non-decreasing, càd, and null at 0.

3.34 **Definition.** Let X and Y be two increasing processes defined on the same stochastic basis. We say that X *strongly majorizes* Y, and we write $Y \prec X$, if the process $X - Y$ is itself increasing. This implies $dY \ll dX$ (absolute continuity of the measure dY_t with respect to dX_t), and in fact is much stronger. □

Just as before, X^n and Y^n are processes defined on the space $(\Omega^n, \mathscr{F}^n, P^n)$. Compare the next result with Theorem 2.15d:

3.35 **Proposition.** *Suppose that for every $n \in \mathbb{N}^*$, X^n is an increasing process that strongly majorizes the increasing process Y^n. If the sequence (X^n) is tight (resp. C-tight), then so is the sequence (Y^n).*

Proof. This follows immediately from 3.21 and 3.26, once noticed that $|Y_t^n| \leq |X_t^n|$, that $w'_N(Y^n, \theta) \leq w'_N(X^n, \theta)$, and that $w_N(Y^n, \theta) \leq w_N(X^n, \theta)$. □

3.36 **Proposition.** a) *Let (X^n) be a sequence of d-dimensional processes with finite variation and $X_0^n = 0$. If the sequence $(\sum_{i \leq d} \operatorname{Var}(X^{n,i}))$ is tight (resp. each sequence $(\operatorname{Var}(X^{n,i}))$ is C-tight), then the sequence (X^n) is tight (resp. C-tight).*

b) *Let (Y^n) be a sequence of $(d \times d)$-dimensional processes, such that $Y_t^n - Y_s^n$ is a symmetric nonnegative matrix for all $s \leq t$. If the sequence $(\sum_{i \leq d} Y^{n,ii})$ is tight (resp. each sequence $(Y^{n,ii})_{n \geq 1}$ is C-tight), then the sequence (Y^n) is tight (resp. C-tight).*

Proof. (a) is proved as 3.35, using 3.33 to obtain first that $(\sum_{i \leq d} \operatorname{Var}(X^{n,i}))$ is C-tight when each $(\operatorname{Var}(X^{n,i}))_{n \geq 1}$ is so.

b) The property of Y^n yields that $\operatorname{Var}(Y^{n,ij}) \prec 2Y^{n,ii} + 2Y^{n,jj}$. Hence $\sum_{i,j \leq d} \operatorname{Var}(Y^{n,ij}) \prec (2d - 1) \sum_{i \leq d} Y^{n,ii}$, so the result follows from (a). □

Let us end with a theorem that shows that in some (very rare!) cases, tightness is indeed an easy problem to solve.

3.37 **Theorem.** *Let X^n, X be increasing processes, such that*

(a) *either X is continuous,*

(b) *or all X^n and X are point processes (i.e. their paths lie in the set $\mathscr{V}^{+,1}$*: see §I3.b or 2.14).

Then if $X^n \xrightarrow{\mathscr{L}(D)} X$ for some dense subset D of $\mathbb{R}_+$, we also have $X^n \xrightarrow{\mathscr{L}} X$.

Proof. From 3.20 tightness of (X^n) is the only thing to prove. Since X^n is increasing and $X_t^n \xrightarrow{\mathscr{L}} X_t$ for all $t \in D$, that (X^n) satisfies condition 3.21i is trivial. To check condition 3.21ii we separate the two cases. We can of course assume $0 \in D$.

(a) Let $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$ be fixed. Consider a subdivision $0 = t_0 < \cdots < t_{r-1} < N \leq t_r$ with $t_i \in D$ and $t_i - t_{i-1} \geq \theta$ for $1 \leq i \leq r$. Set $A = \{x = (x_i)_{0 \leq i \leq r} \in \mathbb{R}^{r+1}: x_{i+1} - x_{i-1} < \eta \text{ for } 1 \leq i \leq r - 1\}$ and associate to it the set $\tilde{A} = \{\alpha \in \mathbb{D}(\mathbb{R}): (\alpha(t_i))_{0 \leq i \leq r} \in A\}$. It is immediate to check that

3.38 $$\alpha \in \mathscr{V}^+ \cap \tilde{A} \Rightarrow w_N(\alpha, \theta) < \eta.$$

Next, X being continuous, one easily finds a $\theta > 0$ and a subdivision as above, such that $P(X \notin \tilde{A}) \leq \varepsilon$. Then, $\{X \notin \tilde{A}\} = \{(X_{t_i})_{0 \leq i \leq r} \notin A\}$ and similarly for X^n; since A is open in $\mathbb{R}^{r+1}$, 3.38 and 3.1 imply:

$$\limsup_{(n)} P^n(w_N(X^n, \theta) \geq \eta) \leq \limsup_{(n)} P^n(X^n \notin \tilde{A}) \leq P(X \notin \tilde{A}) \leq \varepsilon.$$

Hence (X^n) satisfies 3.26ii and a fortiori 3.21ii.

(b) The proof is the same, except that for A we take

$$A = \{(x_i)_{0 \leq i \leq r} \in \mathbb{R}^{r+1}: x_{i+1} - x_{i-1} < \tfrac{3}{2} \text{ for } 1 \leq i \leq r-1\}.$$

Then, due to the particular structure of $\mathscr{V}^{+,1}$, 3.38 is replaced by

$$\alpha \in \mathscr{V}^{+,1} \cap \tilde{A} \Rightarrow w'_N(\alpha, \theta) < \eta, \quad \forall \eta > 0.$$

Finally one easily finds a $\theta > 0$ and a subdivision (t_i) as above, such that $P(X \notin \tilde{A}) \leq \varepsilon$: indeed, if $T_1, \ldots, T_p, \ldots$ are the successive jump times of X, choose p such that $P(T_p \leq N + 1) \leq \varepsilon/2$, then $\theta > 0$ such that $P(T_{i+1} - T_i > 4\theta$ for $i \leq p - 1) \geq 1 - \varepsilon/2$, then finally choose any t_i in D with $\theta \leq t_i - t_{i-1} \leq 2\theta$. The rest of the proof goes on like for (a), with w_N replaced by w'_N. □

Of course, this theorem should be compared to 2.14c.

4. Criteria for Tightness: The Quasi-Left Continuous Case

In this section, we are given a sequence (X^n) of $\mathbb{R}^d$-valued càdlàg processes, each X^n being defined on the space $(\Omega^n, \mathscr{F}^n, P^n)$. We wish to derive criteria for tightness of the sequence (X^n), that are more easy to use than the general theorem 3.21.

Here is an example of such a criterion, which is adapted from Billingsley [12, p. 128] for the case of processes indexed by $\mathbb{R}_+$:

4.1 **Theorem.** *Assume that*

(i) *the sequence* (X_0^n) *is tight* (in $\mathbb{R}^d$);

(ii) $\lim_{\delta \downarrow 0} \limsup_n P^n(|X_\delta^n - X_0^n| > \varepsilon) = 0$ *for all* $\varepsilon > 0$;

(iii) *there is an increasing continuous function F on $\mathbb{R}_+$ and two constants $\gamma \geq 0$, $\alpha > 1$ such that*

4.2
$$\begin{gathered}\forall \lambda > 0, \quad \forall s < r < t, \quad \forall n \in \mathbb{N}^*, \\ P^n(|X_r^n - X_s^n| \geq \lambda, |X_t^n - X_r^n| \geq \lambda) \leq \lambda^{-\delta}[F(t) - F(s)]^\alpha.\end{gathered}$$

Then, the sequence (X^n) is tight.

This criterion is reasonnably general for applications: for instance it works when the X^n's are diffusion processes (possibly with jumps), when all parameters (or coefficients) are bounded uniformly in n. However, it suffers from two important limitations:

1) the majoration 4.2 is *uniform in n*;

2) the majoration 4.2 is a sort of *deterministic* control of the increments of X^n.

Limitation (1) above could be weakened by allowing the function F to depend on n (provided we impose a sort of convergence of the functions F_n's that replace F). But limitation (2) is intrinsic to criteria of the type of 4.1; however, in many cases (like when the X^n's are diffusions with unbounded coefficients) it is impossible to obtain a deterministic control of the increments.

This is why we presently proceed to prove more general criteria, based on ideas taken from Aldous.

§ 4a. Aldous' Criterion for Tightness

This criterion supposes an additional structure on the probability spaces:

4.3 For each n, $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$ is a stochastic basis (see I.1.1), on which is defined an $\mathbb{R}^d$-valued càdlàg *adapted* process X^n. For $N \in \mathbb{N}^*$, we denote by $\mathscr{T}_N^n$ the set of all $\mathbf{F}^n$-stopping times that are bounded by N. □

4.4 *Condition*: For all $N \in \mathbb{N}^*$, $\varepsilon > 0$ we have:

$$\lim_{\theta \downarrow 0} \limsup_n \sup_{S,T \in \mathscr{T}_N^n: S \leq T \leq S+\theta} P^n(|X_T^n - X_S^n| \geq \varepsilon) = 0. \qquad \square$$

4.5 **Theorem.** *If conditions 3.21i and 4.4 are satisfied, the sequence (X^n) of adapted processes is tight.*

4.6 **Remark.** Tightness has *a*-priori nothing to do with filtrations, so this theorem may appear a little strange at first glance. But:

1) Bigger are the filtrations $\mathbf{F}^n$, and more stringent is condition 4.4. So it is judicious to take $\mathbf{F}^n$ as small as possible, namely the smallest filtration such that $\mathscr{F}_t^n \supset \sigma(X_s^n: s \leq t)$ (recall that a filtration is right-continuous, so we cannot simply take $\mathscr{F}_t^n = \sigma(X_s^n: s \leq t)$). Hence the "weakest possible" condition 4.4 is a condition based upon the processes X^n themselves, and nothing else.

2) Now, it may happen that considering bigger filtrations facilitates the computations, as we shall see in some applications of this theorem.

3) If, instead of considering stopping times, we had used all "measurable times" in 4.4 (or, equivalently, if $\mathbf{F}^n$ is the biggest possible filtration, i.e. $\mathscr{F}_t^n = \mathscr{F}^n$ for all t), then this condition would have implied C-tightness. □

4.7 **Remark.** Although the X^n's are not supposed to be quasi-left-continuous, the title of this section comes from the fact that if 4.4 holds, the X^n's are "asymptotically quasi-left-continuous", in a sense made precise by Aldous [2], but which roughly speaking means that all limit points of the sequence $\{\mathscr{L}(X^n)\}$ are laws of processes that are quasi-left-continuous for their natural filtrations.

In order to be simple, let us consider the "stationary" case where $\mathscr{B}^n = \mathscr{B}$ and $X^n = X$ for all n. Then the "sequence" (X^n) obviously satisfies 3.21i, and it satisfies 4.4 *if and only if the process X is quasi-left-continuous.* Suppose indeed that X is not quasi-left-continuous; there exists a predictable time $T \in \mathscr{T}_N$ for some $N \in \mathbb{N}^*$ and $\eta > 0$, $\varepsilon > 0$ with $P(|\Delta X_T| > 2\eta) \geq 3\varepsilon$; there also exists $\delta > 0$ with $P(\sup_{T-\delta \leq s \leq T} |X_s - X_{T-}| > \eta) \leq \varepsilon$, and a sequence (S_n) of stopping times increasing to T and with $S_n < T$; then there is n with $P(S_n < T - \delta) \leq \varepsilon$ and

$$P(|X_{(S_n+\delta)\wedge T} - X_{S_n}| > \eta) \geq P\left(|\Delta X_T| \geq 2\eta, S_n \geq T - \delta, \sup_{T-\delta \leq s \leq T} |X_s - X_{T-}| \leq \eta\right)$$
$$\geq \varepsilon$$

which contradicts 4.4. The converse implication (that quasi-left continuity implies 4.4) is proved similarly. □

Proof of Theorem 4.5. The only thing to prove is that 3.21ii holds under 4.4. Fix $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$. Then 4.4 implies that for all $\rho > 0$ there are $\delta(\rho) > 0$ and $n(\rho) \in \mathbb{N}^*$ such that

4.8 $\quad n \geq n(\rho), \quad S, T \in \mathscr{T}_N^n, \quad S \leq T \leq S + \delta(\rho) \Rightarrow P^n(|X_T^n - X_S^n| \geq \eta) \leq \rho.$

Define the following stopping times: $S_0^n = 0$, $S_{k+1}^n = \inf(t > S_k^n: |X_t^n - X_{S_k^n}^n| \geq \eta)$. Applying 4.8 to $\rho = \varepsilon$ and $S = S_k^n \wedge N$ and $T = S_{k+1}^n \wedge (S_k^n + \delta(\rho)) \wedge N$, and noticing that $|X_{S_k^n}^n - X_{S_{k+1}^n}^n| \geq \eta$ if $S_{k+1}^n < \infty$ show that:

4.9 $$n \geq n_1, k \geq 1 \Rightarrow P^n(S_{k+1}^n \leq N, S_{k+1}^n \leq S_k^n + \delta) \leq \varepsilon,$$

where $n_1 = n(\varepsilon)$ and $\delta = \delta(\varepsilon)$.

Then we choose $q \in \mathbb{N}^*$ such that $q\delta > 2N$. The same argument as above shows that if $\theta = \delta(\varepsilon/q)$ and $n_2 = n_1 \vee n(\varepsilon/q)$,

4.10 $$n \geq n_2, k \geq 1 \Rightarrow P^n(S_{k+1}^n \leq N, S_{k+1}^n \leq S_k^n + \theta) \leq \frac{\varepsilon}{q}.$$

Since $S_q^n = \sum_{1 \leq k \leq q} (S_k^n - S_{k-1}^n)$, we have for $n \geq n_1$:

$$\begin{aligned} NP^n(S_q^n < N) &\geq E_{P^n}\left(\sum_{1 \leq k \leq q} (S_k^n - S_{k-1}^n) 1_{\{S_q^n \leq N\}}\right) \\ &\geq \sum_{1 \leq k \leq q} E_{P^n}[(S_k^n - S_{k-1}^n) 1_{\{S_q^n \leq N, S_k^n - S_{k-1}^n > \delta\}}] \\ &\geq \sum_{1 \leq k \leq q} \delta[P^n(S_q^n \leq N) - P^n(S_q^n \leq N, S_k^n - S_{k-1}^n \leq \delta)] \\ &\geq \delta q P^n(S_q^n \leq N) - \delta q \varepsilon, \end{aligned}$$

the last inequality coming from 4.9. Since $q\delta > 2N$, we deduce that

4.11 $$n \geq n_1 \Rightarrow P^n(S_q^n < N) \leq 2\varepsilon.$$

Next, set $A^n = \{S_q^n \geq N\} \cap [\bigcap_{1 \leq k \leq q} \{S_{k+1}^n > \inf(N, S_k^n + \theta)\}]$. By 4.10 and 4.11 we obtain:

4.12 $$n \geq n_2 \Rightarrow P^n(A^n) \geq 1 - 3\varepsilon.$$

Now if we pick $\omega \in A^n$ and consider the subdivision $0 = t_0 < \cdots < t_r = N$ with $t_i = S_i^n(\omega)$ if $i \leq r - 1$ and $r = \inf(i: S_i^n(\omega) \geq N)$, we have $w(X^n(\omega); [t_{i-1}, t_i)) \leq 2\eta$ by construction of the S_j^n's, and also $t_i - t_{i-1} \geq \theta$ for $i \leq r - 1$. Consequently $w'_N(X^n(\omega), \theta) \leq 2\eta$. Thus 4.12 yields

$$n \geq n_2 \Rightarrow P^n(w'_N(X^n, \theta) > 2\eta) \leq 3\varepsilon,$$

and 3.21ii is proved. □

§ 4b. Application to Martingales and Semimartingales

1. The set-up is again 4.3, and we assume in addition that the process $X^n - X_0^n$ is a locally square integrable martingale on $\mathscr{B}^n$ (that is, the components $X^{n,j} - X_0^{n,j}$ belong to the set $\mathscr{H}_{\text{loc}}^2(\mathscr{B}^n)$ defined in I.1.39). The predictable quadratic covariation $\langle X^{n,j}, X^{n,j} \rangle$ is defined in I.4.2.

4.13 **Theorem.** *We suppose that $X^n - X_0^n$ is a locally square-integrable martingale on $\mathscr{B}^n$ for each n, and we set $G^n = \sum_{j \leq d} \langle X^{n,j}, X^{n,j} \rangle$. Then for the sequence (X^n) to be tight, it is sufficient that*:

(i) *the sequence (X_0^n) is tight* (in $\mathbb{R}^d$);

(ii) *the sequence (G^n) is C-tight* (in $\mathbb{D}(\mathbb{R})$).

Proof. Let $M^n = X^n - X_0^n$. Then $(M^{n,j})^2$ is L-dominated by the predictable increasing process G^n (see I.3.29), and I.3.30 implies that for all $a > 0$, $b > 0$, $N > 0$ we have:

$$P^n\left(\sup_{t \leq N} |M_t^i| \geq \frac{a}{d}\right) \leq \frac{bd^2}{a^2} + P^n(G_N^n \geq b).$$

Thus

4.14 $$P^n\left(\sup_{t \leq N} |X_t^n| \geq 2a\right) \leq P^n(|X_0^n| \geq a) + \sum_{i \leq d} P^n\left(\sup_{s \leq N} |M_s^i| \geq \frac{a}{d}\right)$$
$$\leq P^n(|X_0^n| \geq a) + (bd^3)/a^2 + dP^n(G_N^n \geq b).$$

Therefore conditions (i) and (ii) imply, via 3.21i for (G^n), that (X^n) satisfies 3.21i (choose first b, then a, so that 4.14 is as small as one wishes to).

Similarly, let $S, T \in \mathscr{T}_N^n$ with $S \le T$. If $N_t^n = X_t^n - X_{t\wedge S}^n$ and $\tilde{G}_t^n = G_t^n - G_{t\wedge S}^n$, then each process $(N^{n,i})^2$ is L-dominated by $\tilde{G}^n$, and I.3.30 implies that for all $\varepsilon > 0$, $\eta > 0$,

4.15 $$P^n(|X_T^n - X_S^n| \ge \varepsilon) \le \sum_{i\le d} P^n\left(\sup_{s\le T} |N_s^{n,i}| \ge \frac{\varepsilon}{d}\right)$$
$$\le \sum_{i\le d}\left[\frac{d^2\eta}{\varepsilon^2} + P^n(\tilde{G}_T^n \ge \eta)\right] \le \frac{d^3\eta}{\varepsilon^2} + dP^n(G_T^n - G_S^n \ge \eta).$$

Now, condition (ii) and 3.31 imply the existence of $n_1 \in \mathbb{N}^*$, $\theta > 0$ such that

4.16 $$n \ge n_1 \Rightarrow P^n(w_N(G^n, \theta) \ge \eta) \le \eta.$$

If S, T are as above and $T \le S + \theta$ we have $G_T^n - G_S^n \le \eta$ if $w_N(G^n, \theta) \le \eta$. Therefore it follows from 4.15 and 4.16 that

$$n \ge n_1 \Rightarrow \sup_{S,T\in\mathscr{T}_N^n, S\le T\le S+\theta} P^n(|X_T^n - X_S^n| \ge \varepsilon) \le \eta(d^3/\varepsilon^2 + d).$$

Since $\eta > 0$ is arbitrary. we deduce that (X^n) satisfies condition 4.4. Thus the result follows from Theorem 4.5. □

2. Finally, we consider the case where each X^n is a d-dimensional semimartingale, and we shall heavily use the characteristics of X^n, as defined in section II.2: we pick a troncation function h in the class $\mathscr{C}_t^d$ defined in II.2.3, and we consider the characteristics $(B^n = B^n(h), C^n, \nu^n)$ of X^n on the basis $\mathscr{B}^n$, associated to this function. We also introduce the modified second characteristic of X^n (see II.2.16 and II.2.18), namely

4.17 $$\tilde{C}^{n,ij} = \tilde{C}^{n,ij}(h) = C^{n,ij} + (h^i h^j) * \nu^n - \sum_{s\le\cdot} \Delta B_s^{n,i} \Delta B_s^{n,j}.$$

4.18 **Theorem.** *With the above notation, for the sequence (X^n) to be tight it suffices that*

(i) *the sequence* (X_0^n) *is tight* (in $\mathbb{R}^d$);

(ii) *for all* $N > 0$, $\varepsilon > 0$,

4.19 $$\lim_{a\uparrow\infty} \limsup_{(n)} P^n[\nu^n([0,N] \times \{x: |x| > a\}) > \varepsilon] = 0;$$

(iii) *each one of the following sequences of processes is C-tight*:

1—(B^n),

2—$(\tilde{C}^n)$,

3—$(g_p * \nu^n)_{n\ge1}$, *with* $g_p(x) = (p|x| - 1)^+ \wedge 1$ *and* $p \in \mathbb{N}^*$.

Moreover, (i) *and* (ii) *are also necessary for tightness of* (X^n).

4.20 **Remarks.** 1) Due to 3.36b, we could replace (iii.2) above by: each sequence $(\tilde{C}^{n,ii})$ is C-tight.

2) Consider the following increasing predictable process:

4.21 $$F^n = \sum_{i\leq d} [\mathrm{Var}(B^{n,i}) + C^{n,ii}] + (|x|^2 \wedge 1) * \nu^n.$$

Then *C-tightness of* (F^n) *implies 4.18iii*: notice that $\tilde{C}^{n,ii} \prec aF^n$ for some $a > 0$ and $g_p * \nu^n \prec p^2 F^n$, then apply 3.35 and 3.36. □

We begin with two lemmas.

4.22 **Lemma.** a) *For all* $N > 0$, $a > 0$, *there is equivalence between*:
(i) $\lim_{(n)} P^n(\sup_{s\leq N} |\Delta X_s^n| > a) = 0$,
(ii) $\lim_{(n)} P^n(\nu^n([0,N] \times \{x: |x| > a\}) > \varepsilon) = 0$ *for all* $\varepsilon > 0$.
b) *For all* $N > 0$ *there is equivalence between*
(i) $\lim_{a\uparrow\infty} \limsup_{(n)} P^n(\sup_{s\leq N} |\Delta X_s^n| > a) = 0$,
(ii) $\lim_{a\uparrow\infty} \limsup_{(n)} P^n(\nu^n([0,N] \times \{x: |x| > a\}) > \varepsilon) = 0$ *for all* $\varepsilon > 0$.

Proof. This is a simple consequence of Lenglart domination property. Set

$$A_t^n = \sum_{0<s\leq t} 1_{\{|\Delta X_s^n| > a\}}$$

$$\tilde{A}_t^n = \nu^n([0,t] \times \{x: |x| > a\}).$$

Then $\tilde{A}^n$ is the compensator of A^n on $\mathscr{B}^n$, so A^n is L-dominated by $\tilde{A}^n$, and $\tilde{A}^n$ is L-dominated by A^n. Thus I.3.30a gives

$$P^n(A_N^n \geq 1) \leq \varepsilon + P^n(\tilde{A}_N^n \geq \varepsilon)$$

and since $\{A_N^n \geq 1\} = \{\sup_{s\leq N} |\Delta X_s^n| > a\}$ we easily deduce the implications (ii) ⇒ (i) in (a) and in (b).

Applying I.3.30b with $\eta = \rho\varepsilon$ gives for all $\rho > 0$, $\varepsilon > 0$:

$$P^n(\tilde{A}_N^n \geq \varepsilon) \leq \rho + \frac{1}{\varepsilon} E^n\left(\sup_{s\leq N} \Delta A_s^n\right) + P^n(A_N^n \geq \varepsilon\rho).$$

But $\Delta A^n \leq 1$ and $\{\sup_{s\leq N} \Delta A_s^n > 0\} = \{A_N^n \geq \varepsilon\rho \wedge 1\} = \{\sup_{s\leq N} |\Delta X_s^n| > a\}$, hence

$$P^n(\tilde{A}_N^n \geq \varepsilon) \leq \rho + \left(\frac{1}{\varepsilon} + 1\right) P^n\left(\sup_{s\leq N} |\Delta X_s^n| > a\right).$$

Since $\varepsilon > 0$ and $\rho > 0$ are arbitrary, we deduce the implications (i) ⇒ (ii) in (a) and in (b). □

4.23 **Lemma.** *Condition 4.18iii does not depend on the choice of h in* $\mathscr{C}_t^d$.

Proof. Suppose that 4.18.iii is satisfied for some $h \in \mathscr{C}_t^d$, and let h' be another function in $\mathscr{C}_t^d$. There are two constants $a > 0$, $b > 0$ such that $|h| \leq a$, $|h'| \leq a$,

$h(x) = h'(x) = x$ for $|x| \leq b$. Choose $p \in \mathbb{N}^*$ such that $2/p \leq b$. Hence $|h - h'| \leq 2ag_p$ and $(|h|^2 - |h'|^2 \leq 2a^2 g_p$ and it follows from 3.35 and 3.36 and from C-tightness of $(g_p * \nu^n)$ that the sequences $\{(h'-h)*\nu^n\}$ and $\{(|h|^2 - |h'|^2 * \nu^n\}$ are also C-tight.

From II.2.25, $B^n(h') = B^n(h) + (h' - h) * \nu^n$, hence 3.33 implies that the sequence $\{B^n(h')\}$ is C-tight. Similarly, II.2.26 gives

$$\tilde{C}^{n,jj}(h') = \tilde{C}^{n,jj}(h) + [(h^j)^2 - (h'^j)^2] * \nu^n + H^{n,jj}$$

with

$$H_t^{n,jj} = \sum_{s \leq t} [\Delta B_s^{n,j}(h)^2 - \Delta B_s^{n,j}(h')^2],$$

hence $\{\tilde{C}^{n,jj}(h')\}$ will be C-tight, provided $(H^{n,jj})$ itself is C-tight. But $|\Delta B^n(h)| \leq a$ and $|\Delta B^n(h')| \leq a$, thus from II.2.25 again,

$$\begin{aligned}\operatorname{Var}(H^{n,jj}) &\prec \sum_{s \leq \cdot} |\Delta B_s^n(h') - \Delta B_s^n(h)| [|\Delta B_s^n(h')| + |\Delta B_s^n(h)|] \\ &\prec 2a|h - h'| * \nu^n \prec 4a^2(g_p * \nu^n),\end{aligned}$$

and the result follows from 3.36 and 4.20.1. □

Proof of Theorem 4.18. a) We prove first the sufficient condition. Let $h \in \mathscr{C}_t^d$ be fixed, and set $h_q(x) = qh(x/q)$ for all $q \in \mathbb{N}^*$: then $h_q \in \mathscr{C}_t^d$. We will apply Lemma 3.32 to the decompositions $X^n = U^{nq} + V^{nq} + W^{nq}$, with

$$U^{nq} = X_0^n + M^n(h_q), \quad V^{nq} = B^n(h_q), \quad W^{nq} = \check{X}^n(h_q),$$

where we have used the notation II.2.4 and II.2.5.

First, 4.23 implies that the sequence $(V^{nq})_{n \geq 1}$ is C-tight, so it satisfies 3.32ii (with $a_q = 1/q$ for example). Secondly, 4.23 again implies that the sequence $(\sum_{j \leq d} \tilde{C}^{n,j}(h_q))_{n \geq 1}$ is C-tight, while $\tilde{C}^{n,j}(h_q) = \langle M^{n,j}(h_q), M^{n,j}(h_q) \rangle$; so Theorem 4.13 yields that the sequence $(U^{nq})_{n \geq 1}$ is tight. Thirdly, there is a constant $a > 0$ such that $h(x) = x$ for $|x| \leq a$, so $h_q(x) = x$ for $|x| \leq aq$ and by definition of $\check{X}^n(h_q)$ we obtain that

$$P^n\left(\sup_{t \leq N} |W_t^{nq}| > 0\right) \leq P^n\left(\sup_{t \leq N} |\Delta X_t^n| > aq\right).$$

Then condition 4.18ii clearly implies, via Lemma 4.22b, that the family (W^{nq}) satisfies condition 3.32iii. Hence, Lemma 3.32 insures that the sequence (X^n) is tight.

b) Conversely, suppose that (X^n) is tight. Then 3.21 gives

$$\lim_{a \uparrow \infty} \limsup_{(n)} P^n\left(\sup_{t \leq N} |X_t^n| > a\right) = 0$$

for all $N > 0$. Then we have (i), and since $|\Delta X_t^n| \leq 2 \sup_{s \leq t} |X_s^n|$ we also have (ii) (apply Lemma 4.22b). □

5. Criteria for Tightness: The General Case

§ 5a. Criteria for Semimartingales

Notation and assumptions are the same as in § 4c. We will prove a result similar to Theorem 4.18, but under weaker conditions. To this end, we first introduce a series of conditions on a sequence (G^n) of increasing processes, each G^n being defined on the basis $\mathscr{B}^n$.

5.1 *Condition* (C1). (G^n) converges in law to a deterministic process. □

This is equivalent to:

5.2 *Condition* (C′1). There is a càdlàg increasing function g on $\mathbb{R}_+$ and a dense subset D in $\mathbb{R}_+$, such that

5.3
$$t \in D \Rightarrow \begin{cases} G_t^n \xrightarrow{\mathscr{L}} g(t) \\ \sum_{0<s\leq t} (\Delta G_s^n)^2 \xrightarrow{\mathscr{L}} \sum_{0<s\leq t} \Delta g(s)^2. \end{cases}$$ □

(the equivalence of these two conditions follows from 2.15, which immediately gives (C1) $\Rightarrow$ (C′1) with $D = \mathbb{R}_+ \setminus J(g)$. Conversely assume (C′1); we may assume that D is countable; then since convergence in law or in measure are the same when the limit is deterministic, any subsequence contains a further subsequence for which 5.3 holds identically outside a null set; hence 2.15 implies that this sub-sub-sequence converges almost surely to g in $\mathbb{D}(\mathbb{R})$, which in turn implies: $G^n \xrightarrow{\mathscr{L}} g$).

5.4 *Condition* (C2). (G^n) converges in law to a process G, whose all paths are strongly majorized (see 3.34) by the same (deterministic) increasing càdlàg function F. □

5.5 *Condition* (C3). (i) We have $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ for all n (the filtrations $\mathbf{F}^n$ may differ); set $\mathscr{F}_t = \bigcap_n \mathscr{F}_t^n$;

(ii) there is a process G on Ω such that (G^n) tends to G in measure for Skorokhod topology;

(iii) there is an $\mathbf{F}$-predictable process F which strongly majorizes G. □

Again an application of 2.15 shows that (C3) is equivalent to:

5.6 *Condition* (C′3). We have (i) and (iii) of (C3), where G is an increasing process with the following property: there is a dense subset D of $\mathbb{R}_+$, such that

$$t \in D \Rightarrow \begin{cases} G_t^n \xrightarrow{P} G_t \\ \sum_{0 \le s \le t} (\Delta G_s^n)^2 \xrightarrow{P} \sum_{0 \le s \le t} (\Delta G_s)^2 \end{cases}$$

($\xrightarrow{P}$ means: convergence in measure). □

For the last conditions that we want to introduce, we need some additional pieces of notation. Recall that the filtration $\mathbf{D}(\mathbb{R}^d)$ is defined in 1.1. If $P \in \mathscr{P}(\mathbb{D}(\mathbb{R}^d))$, we denote by $\mathbf{D}(\mathbb{R}^d)^P$ the completion of the filtration $\mathbf{D}(\mathbb{R}^d)$ with respect to P, in the sense of I.1.4. We also denote by ξ the canonical process on $\mathbb{D}(\mathbb{R}^d)$, defined by $\xi_t(\alpha) = \alpha(t)$ for all $t \geq 0$, $\alpha \in \mathbb{D}(\mathbb{R}^d)$.

5.7 *Condition* (C4). The sequence $\{\mathscr{L}(G^n)\}$ converges in $\mathscr{P}(\mathbb{D}(\mathbb{R}))$ to a limit P, and the canonical process ξ is predictable with respect to the filtration $\mathbf{D}(\mathbb{R})^P$. □

The previous condition is easy enough to state, but not so easy to verify in practice. So we introduce another condition which trivially generalize (C4) and is met much more often.

5.8 *Condition* (C5). Each basis $\mathscr{B}^n$ supports a d-dimensional process Y^n, and we consider the $(d+1)$-dimensional process (Y^n, G^n). Then the sequence $\{\mathscr{L}(Y^n, G^n)\}$ converges in $\mathscr{P}(\mathbb{D}(\mathbb{R}^{d+1}))$ to a limit P; moreover, the $(d+1)$th component ξ^{d+1} of the canonical process is predictable with respect to the filtration $\mathbf{D}(\mathbb{R}^{d+1})^P$. □

The following implications are trivial:

5.9 $$\text{(C1)} \Rightarrow \text{(C2)}, \quad \text{(C1)} \Rightarrow \text{(C4)} \Rightarrow \text{(C5)},$$

We may now state our generalization of Theorem 4.18. The assumptions are the same: for each $n \in \mathbb{N}^*$, X^n is a *d-dimensional semimartingale* on $\mathscr{B}^n$, with characteristics (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$ given by 4.17 (the truncation function $h \in \mathscr{C}_t^d$ is fixed). We also set $g_p(x) = (p|x| - 1)^+ \wedge 1$ for $p \in \mathbb{N}^*$.

5.10 **Theorem.** *With the above notation, for the sequence* (X^n) *to be tight it suffices that*

(i) *the sequence* (X_0^n) *is tight* (in $\mathbb{R}^d$);

(ii) *for all* $N > 0$, $\varepsilon > 0$, *we have*

$$\lim_{a \uparrow \infty} \limsup_{(n)} P^n(\nu^n([0, N] \times \{x : |x| > a\}) > \varepsilon) = 0;$$

(iii) *the sequence* (B^n) *is tight*;

(iv) *for all* $n \in \mathbb{N}^*$, $p \in \mathbb{N}^*$ *there exists a predictable increasing process* $G^{n,p}$ *on* $\mathscr{B}^n$ *that strongly majorizes* $\sum_{j \le d} \tilde{C}^{n,jj} + g_p * \nu^n$, *with the following: for each* $p \in \mathbb{N}^*$, *for any subsequence extracted from* $(G^{n,p})_{n\geq 1}$ *there is a further subsequence that satisfies* (C1) *or* (C2) *or* (C3) *or* (C4) *or* (C5).

We prove this theorem in the next subsection. We shall also see that these conditions do not depend on the choice of h in $\mathscr{C}_t^d$.

5.11 **Remark.** Theorem 4.18 is a particular case of this theorem: take $G^{n,p} = \tilde{C}^n + g_p * \nu^n$ and note that C-tightness of $(G^{n,p})_{n\geq 1}$ implies that from each subsequence one may extract a further subsequence that satisfies condition (C4) (because: continuous and adapted imply predictable). □

5.12 **Remark.** The decisive improvement brought up here, upon 4.18, is that no more "asymptotic quasi-left-continuity" is implicit in its assumptions.

For instance, take a stationary sequence $X^n = X$. Then it satisfies all the conditions of the theorem (take $G^{n,p} = \sum_{j\leq d} \tilde{C}^{jj} + g_p * \nu$ in (iv), with condition (C3)). □

5.13 **Remark.** Define F^n by 4.21 and suppose that for each n there is an increasing predictable process G^n on $\mathscr{B}^n$ that strongly majorizes F^n, such that from any subsequence on (G^n) on may extract a further subsequence satisfying one of the (Ci)'s: then condition (iii) and (iv) are satisfied; this follows from 3.36 for (iii), and for (iv) it suffices to take $G^{n,p} = pG^n$. □

5.14 **Remark.** The predictability assumptions on the $G^{n,p}$'s could be dispensed with (at the price of some more complications to an already messy proof). But *predictability* of F in (C3) or ξ^{d+1} in (C5) is essential.

Let us give an example. N denotes a standard Poisson process (see § I.3b) on a basis $\mathscr{B}$, and $A(a)_t = N_{(t-a)^+}$. Then $A(a)$ is a predictable process in $\mathscr{V}^+$ for each $a > 0$, while $A(0) = N$. Set $X^n = N + A\left(\frac{1}{n}\right)$ and $\mathscr{B}^n = \mathscr{B}$, and choose $h \in \mathscr{C}^1$ such that $h(1) = 0$. Then a simple computation shows that $B^n = 0$, $\tilde{C}^n = 0$, $g_p * \nu_t^n = t + A\left(\frac{1}{n}\right)$: it is thus natural to take $G_t^{n,p} = t + A\left(\frac{1}{n}\right)_t$. The sequence $(G^{n,p})_{n\geq 1}$ would satisfy (C3) with $G_t = t + N_t$, *except* that the limit G is not strongly majorized by a predictable increasing process. And (X^n) is not tight, because each X^n has only jumps of magnitude 1, while all limit processes would have the same distribution as $2N$, with only jumps of magnitude 2.

This example shows us three more things:

1) The (perhaps surprising) fact that the pointwise limit for Skorokhod topology of a sequence of predictable processes may not be predictable;

2) (C3) does not imply (C4) or (C5): take $G^n = A\left(1 + \frac{1}{n}\right)$, which meets (C3) but not (C5).

3) tightness of each sequence (B^n) and $(\tilde{C}^n + g_p * \nu^n)_{n\geq 1}$, plus conditions (i) and (ii), are not sufficient for obtaining tightness of (X^n) (while C-tightness is sufficient, by 4.18). □

§ 5b. An Auxiliary Result

Exactly as in Section 4 we need an auxiliary result, similar to 4.13. We suppose that each X^n is a locally square-integrable semimartingale (see II.2.27), i.e. that

5.15
$$X^n = X_0^n + M^n + A^n, \quad M^{n,i} \in \mathscr{H}^2_{\mathrm{loc}}(\mathscr{B}^n),$$
$$A^{n,i} \in \mathscr{P} \cap \mathscr{V}(\mathscr{B}^n), \quad M_0^n = A_0^n = 0$$

and we associate with it the following predictable process in $\mathscr{V}^+(\mathscr{B}^n)$:

5.16
$$\hat{G}^n = \sum_{i\leq d} [\mathrm{Var}(A^{n,i}) + \langle M^{n,i}, M^{n,i}\rangle].$$

5.17 **Theorem.** *With the above assumptions, it suffices for* (X^n) *to be tight that*:

(i) *the sequence* (X_0^n) *be tight* (in $\mathbb{R}^d$);

(ii) *for each n there is an increasing predictable process* G^n *on* $\mathscr{B}^n$ *that strongly majorizes* $\hat{G}^n$, *and such that the sequence* (G^n) *satisfies* (C1) *or* (C2) *or* (C3) *or* (C4) *or* (C5).

We defer the proof to the next subsection and we proceed now to deduce Theorem 5.10 from the above. We begin with a lemma.

5.18 **Lemma.** *If conditions* 5.10iii, iv *are satisfied for some* $h \in \mathscr{C}_t^d$, *they are also satisfied for any other function in* $\mathscr{C}_t^d$.

Proof. a) We suppose that conditions 5.10iii, iv are satisfied for $h \in \mathscr{C}_t^d$, and $B^n = B^n(h)$, $\tilde{C}^n = \tilde{C}^n(h)$.

Let $h' \in \mathscr{C}^d$, and $B'^n = B^n(h')$, $\tilde{C}'^n = \tilde{C}^n(h')$. There are two constants $a \geq 1$, $b > 0$ such that $|h| \leq a$, $|h'| \leq a$, $h(x) = h'(x) = x$ for $|x| \leq b$. We choose $p \in \mathbb{N}^*$ such that $\frac{2}{p} \leq b$. Then the same proof as in 4.23 shows that (with the same process H^n)

$$\begin{aligned}\sum_{j\leq d} \tilde{C}'^{n,jj} &= \sum_{j\leq d} \tilde{C}^{n,jj} + (|h'|^2 - |h|^2) * \nu^n + \sum_{j\leq d} H^{n,jj} \\ &\prec \sum_{j\leq d} \tilde{C}^{n,jj} + 2a^2(1+2d)(g_p * \nu^n).\end{aligned}$$

Thus, since $a \geq 1$ and $g_p \leq g_q$ (resp. $\geq g_q$) for $p \leq q$ (resp. $\geq q$) we obtain that

$$\sum_{j\leq d} \tilde{C}'^{n,jj} + g_q * \nu^n \prec \begin{cases} 2a^2(1+2d)G^{n,q} & \text{if } q \geq p \\ 2a^2(1+2d)G^{n,p} & \text{if } q < p \end{cases}$$

and we deduce that (iv) is fulfilled for h'.

b) It remains to prove that (B'^n) is tight. Let $A^n = B'^n - B^n$ and $H^{n,q} = g_q * \nu^n$. Since $H^{n,q} \prec G^{n,q}$, 3.35 implies that each sequence $(H^{n,q})_{n \geq 1}$ is tight: indeed, this follows from the tightness of $(G^{n,q})$, which in turn follows from the following convergence in (iv): from each subsequence of $(G^{n,q})_{n \geq 1}$ one may extract a further subsequence that satisfies one of the (Ci)'s and hence that converges in law.

Moreover, $A^n = (h' - h) * \nu^n$ by II.2.25, hence $\sum_{i \leq d} \mathrm{Var}(A^{n,i}) \prec 2aH^{n,p}$ and 3.36 implies that (A^n) is tight. Since (B^n) is also tight, it is immediate that (B'^n) satisfies 3.21i. So we are left to prove that 3.21ii is fulfilled by the sequence (B'^n).

Let $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$. We pick a $q \in \mathbb{N}^*$ with $2/q \leq b \wedge \eta$. We have $|h| \leq a$ and $h(x) = x$ for $|x| \leq 2/q$, thus $|h| \leq 2/q + ag_q \leq \eta + ag_q$. From II.2.11 and II.2.14 it follows that

$$|\Delta B_s^n| = \left| \int \nu^n(\{s\} \times dx) h(x) \right| \leq \eta + a\Delta(g_q * \nu^n).$$

Since $A^n = (h' - h) * \nu^n$ and $|h' - h| \leq 2ag_q$, we also have $|\Delta A^n| \leq 2a\Delta(g_q * \nu^n)$. Recalling that $H^{n,q} = g_q * \nu^n$, we thus obtain:

5.19 $$|\Delta B^n| \leq \eta + a\Delta(H^{n,q}), \quad |\Delta A^n| \leq 2a\Delta(H^{n,q}).$$

Since (A^n), (B^n), $(H^{n,q})_{n \geq 1}$ are tight, 3.21 implies the existence of $n_0 \in \mathbb{N}^*$ and $\theta > 0$ such that

5.20 $$\begin{cases} n \geq n_0 \Rightarrow P^n(K^n) \geq 1 - \varepsilon, \quad \text{with} \\ K^n = \left\{ w'_N(A^n, \theta) < \eta, w'_N(B^n, \theta) < \eta, w'_N(H^{n,q}, \theta) < \dfrac{\eta}{a} \right\}. \end{cases}$$

Now we consider a point ω in K^n. From 1.12 there exists a subdivision $0 = t_0 < \cdots < t_r = N$ with $\theta \leq t_i - t_{i-1} \leq 2\theta$ for $1 \leq i \leq r - 1$, and $t_r - t_{r-1} \leq 2\theta$, and $w(H^{n,q}; [t_{i-1}, t_i)) \leq \eta/a$. Hence $\Delta(H^{n,q}) \leq \eta/a$ on each interval (t_{i-1}, t_i) and 5.19 gives:

5.21 $$s \in (t_{i-1}, t_i), \quad \omega \in K^n \Rightarrow |\Delta B_s^n| \leq 2\eta, \quad |\Delta A_s^n| \leq 2\eta.$$

For the same point $\omega \in K^n$ we have $w'_N(B^n, \theta) < \eta$: hence there is a subdivision $0 = s_0 < \cdots < s_{r'} = N$ with $s_i - s_{i-1} \geq \theta$ for $1 \leq i \leq r' - 1$ and $w(B^n; [s_{i-1}, s_i)) \leq \eta$. Now, there at most two points s_j in each interval $[t_{i-1}, t_i)$, and so 5.21 gives:

$$w(B^n; [t_{i-1}, t_i)) \leq \eta + 2\eta + \eta + 2\eta + \eta = 7\eta.$$

For the same reasons, we also have $w(A^n; [t_{i-1}, t_i)) \leq 7\eta$. Finally, if we recall that $B'^n = B^n + A^n$, it follows that $w(B'^n; [t_{i-1}, t_i)) \leq 14\eta$. Thus we have proved that

$$K^n \subset \{ w'_N(B'^n, \theta) \leq 14\eta \}.$$

Hence 5.20 yields

$$n \geq n_0 \Rightarrow P^n(w'_N(B'^n, \theta) > 14\eta) \leq \varepsilon.$$

Since $\varepsilon > 0$ and $\eta > 0$ are arbitrary, the sequence (B'^n) satisfies condition 3.21ii, and the proof is finished. □

Proof of Theorem 5.10. Let $h \in \mathscr{C}_t^d$ be fixed, and set $h_b(x) = bh(x/b)$ for all $b > 0$, so that $h_b \in \mathscr{C}_t^d$ as well. We will apply Lemma 3.32 to the decompositions $X^n = U^{nq} + V^{nq} + W^{nq}$, with

$$U^{nq} = X_0^n + M^n(h_q) + B^n(h_q) - B^n(h_{1/q})$$
$$V^{nq} = B^n(h_{1/q}), \quad W^{nq} = \check{X}^n(h_q),$$

where we have used the notation II.2.4 and II.2.5 and where $q \in \mathbb{N}^*$ (note that this decomposition is not exactly the same than for proving 4.18).

That the family (W^{nq}) satisfies 3.32iii follows from 5.10ii exactly like in 4.18. By 5.18, each sequence $(V^{nq})_{n \geq 1}$ is tight, and we have $|\Delta V^{nq}| \leq a_q := a/q$ where a is an upper bound for $|h|$: therefore the family (V^{nq}) satisfies 3.32ii. Finally, U^{nq} is a process of type 5.15 and the increasing process $\hat{G}^{nq}$ associated to it by 5.16 is

$$\hat{G}^{nq} = \sum_{j \leq d} \tilde{C}^{n,jj}(h_q) + |h_q - h_{1/q}| * \nu^n.$$

(use II.2.25 and II.2.16). Since $h_q - h_{1/q}$ is bounded and is 0 on a neighbourhood of 0, it follows from 5.10iv that $\hat{G}^{nq}$ is strongly majorized by a predictable increasing process $\tilde{G}^{nq}$, where the family $(\tilde{G}^{nq})$ satisfies property 5.10iv. Therefore, if we apply Theorem 5.17, we obtain that from any subsequence of $(U^{nq})_{n \geq 1}$ one may extract a further subsequence that is tight. This in turn implies that the sequence $(U^{nq})_{n \geq 1}$ is tight, and we are finished. □

§ 5c. Proof of Theorem 5.17.

1. Throughout all this subsection, we assume that *the hypotheses of Theorem 5.17 are in force.* Under each one of the conditions (Ci), the sequence (G^n) converges in law to an increasing process G^∞:

5.22 $$G^n \xrightarrow{\mathscr{L}} G^\infty$$

($G^\infty = G$ under (C2) or (C3)). Recalling 3.10, we set:

5.23 $$U = U(G^\infty)$$

5.24 $$\begin{cases} T_0^n(u) = 0, \quad T_{p+1}^n(u) = \inf(t > T_p^n(u): \Delta G_t^n > u), & n \in \bar{\mathbb{N}}, u > 0 \\ G_t^n(u) = G_t^n - \sum_{i \geq 1} \Delta G_{T_i^n(u)}^n 1_{\{T_i^n(u) \leq t\}}, & n \in \bar{\mathbb{N}}, u > 0 \\ X_t^n(u) = X_t^n - \sum_{i \geq 1} \Delta X_{T_i^n(u)}^n 1_{\{T_i^n(u) \leq t\}}, & n \in \mathbb{N}, u > 0 \end{cases}$$

(with $G^n(\infty) = G^n$, $X^n(\infty) = X^n$, $T_i^n(\infty) = \infty$ for $i \geq 1$).

5.25 **Lemma.** *Set* $f(\eta) = \dfrac{d^2}{\eta^2} + \dfrac{d}{2\eta}$ *for* $\eta > 0$. *Then for all* $\varepsilon > 0, \eta > 0, u \in (0, \infty]$, $n \in \mathbb{N}^*$, *and all finite stopping times* $S^n \leq T^n$ *on* $\mathscr{B}^n$, *we have*

5.26 $$P^n\left(\sup_{S^n \le s \le T^n} |X^n_s(u) - X^n_{S^n}(u)| \ge \eta\right)$$
$$\le 4d[\varepsilon f(\eta) + P^n(G^n_{T^n}(u) - G^n_{S^n}(u) \ge \varepsilon)].$$

Proof. a) We consider first a locally square-integrable semimartingale X:

$$X = X_0 + M + A: \quad M^i \in \mathscr{H}^2_{\mathrm{loc}}(\mathscr{B}), \quad A^i \in \mathscr{P} \cap \mathscr{V}(\mathscr{B}), \quad M_0 = A_0 = 0$$

on some basis $\mathscr{B}$, and the associated increasing process (see 5.16):

5.27 $$\hat{G} = \sum_{i \le d} [\mathrm{Var}(A^i) + \langle M^i, M^i \rangle].$$

Let $\varepsilon > 0$, $\eta > 0$ and consider two finite stopping times $S \le T$ on $\mathscr{B}$. Since for every real-valued variable Z one has $P(|Z| \ge a) \le \frac{b}{a} + P(|Z| \ge b)$ $(a, b > 0)$, we get

$$P\left(\sup_{S < s \le T} |A^i_s - A^i_S| \ge \frac{\eta}{2d}\right) \le \frac{2d\varepsilon}{\eta} + P(\mathrm{Var}(A^i)_T - \mathrm{Var}(A^i)_S \ge \varepsilon).$$

Next, set

$$N^i_t = M^i_t - M^i_{t \wedge S}, \quad H^i_t = \langle M^i, M^i \rangle_t - \langle M^i, M^i \rangle_{t \wedge S}.$$

Then $H^i = \langle N^i, N^i \rangle$, and $(N^i)^2$ is L-dominated by H^i, so I.3.30 yields:

$$P\left(\sup_{S < s \le T} |M^i_s - M^i_S| \ge \frac{\eta}{2d}\right) = P\left(\sup_{s \le T} (N^i_s)^2 \ge \frac{\eta^2}{4d^2}\right)$$
$$\le \frac{4d^2 \varepsilon}{\eta^2} + P(\langle M^i, M^i \rangle_T - \langle M^i, M^i \rangle_S \ge \varepsilon).$$

Finally, if $|X_s - X_S| \ge \eta$ we have $|A^i_s - A^i_S| \ge \eta/2d$ or $|M^i_s - M^i_S| \ge \eta/2d$ for at least one $i \le d$. Then, summing up the preceding results shows that

5.28 $$P\left(\sup_{S < s \le T} |X_s - X_S| \ge \eta\right) \le 4d[\varepsilon f(\eta) + P(G_T - G_S \ge \varepsilon)].$$

b) We turn back to the situation at hand. Let $u \in (0, \infty]$ and $H^n(u) = 1 - \sum_{i \ge 1} 1_{[\![T^n_i(u)]\!]}$ (and $H^n(\infty) = 1$). Then $H^n(u)$ is a predictable process (because G^n, and hence $T^n_i(u)$, are predictable) on $\mathscr{B}^n$, and by construction

$$G^n(u) = H^n(u) \cdot G^n, \quad X^n(u) = H^n(u) \cdot X^n.$$

Now it is clear that $X^n(u)$ is again a locally square-integrable semimartingale, and the process $\hat{G}^n(u)$ associated to $X^n(u)$ by 5.27 is $\hat{G}^n(u) = H^n(u) \cdot \hat{G}^n$, where $\hat{G}^n$ is given by 5.16. Therefore 5.26 follows from 5.28 and from the property $\hat{G}^n(u) \prec G^n(u)$. □

2. Now we derive two easy consequences of Property 5.22.

5.29 **Lemma.** *For all $N > 0$, $u > 0$, we have*

$$\lim_{\theta \downarrow 0} \limsup_{(n)} P^n\left(\bigcup_{i \geq 1} \{T_i^n(u) \leq N, T_i^n(u) - T_{i-1}^n(u) < \theta\} \right) = 0.$$

Proof. A glance at the definitions of w'_N and $T_i^n(u)$ shows that

$$\bigcup_{i \geq 1} \{T_i^n(u) \leq N, T_i^n(u) - T_{i-1}^n(u) < \theta\} \subset \{w'_{N+1}(G^n, \theta) > u\}$$

for $\theta \leq 1$, hence the result follows from 3.21 and 5.22. □

Recall that $\mathscr{T}_N^n$ denotes the set of all $\mathbf{F}^n$-stopping times that are bounded by N.

5.30 **Lemma.** *For all $N > 0$, $u \notin U$, $\eta > 0$, we have*

$$\lim_{\theta \downarrow 0} \limsup_{(n)} \sup_{S \in \mathscr{T}_N^n} P^n\{G_{S+\theta}^n(u) - G_S^n(u) \geq u + \eta\} = 0.$$

Proof. We can reproduce part (a) of the proof of 3.37: since $\Delta G^\infty(u) \leq u$ it is easy to find a subdivision $0 = t_0 < \cdots < t_{r-1} \leq N + 1 < t_r$ with $t_i \notin J(G^\infty)$, such that $P(G^\infty(u) \notin \tilde{A}) \leq \varepsilon$ for a given $\varepsilon > 0$, with the notation:

$$\tilde{A} = \{\alpha \in \mathbb{D}(\mathbb{R}) \colon (\alpha(t_i))_{0 \leq i \leq r} \in A\}$$

$$A = \{x = (x_i)_{0 \leq i \leq r} \in \mathbb{R}^{r+1} \colon x_{i+1} - x_{i-1} < u + \eta \quad \text{for } 1 \leq i \leq r - 1\}.$$

Then, 5.22 and $u \notin U(G^\infty)$ allow to apply 2.7 to obtain: $G^n(u) \xrightarrow{\mathscr{L}} G^\infty(u)$. Since A is open in $\mathbb{R}^{r+1}$, and since 3.38 holds (with $u + \eta$ instead of η), we get

$$\begin{aligned} \limsup_{(n)} P^n(w_{N+1}(G^n(u), \theta) \geq u + \eta) &\leq \limsup_{(n)} P^n(G^n(u) \notin \tilde{A}) \\ &\leq P(G^\infty(u) \notin \tilde{A}) \leq \varepsilon, \end{aligned}$$

where $\theta = \inf_{1 \leq i \leq r}(t_i - t_{i-1})$. Since $G_{S+\theta}^n(u) - G_S^n(u) \leq u + \eta$ whenever $\theta \leq 1$, $S \leq N$ and $w_{N+1}(G^n(u), \theta) \leq u + \eta$, and since $\varepsilon > 0$ is arbitrary, we have the result. □

3. An auxiliary condition. We consider the following condition:

5.31 *Condition.* For all $N > 0$, $u \notin U$, $i \geq 1$, $\varepsilon > 0$, there is an integer $J \in \mathbb{N}^*$ such that if $\delta > 0$ there exist $n_0 \in \mathbb{N}^*$ and $\sigma \in (0, \delta)$ and, for each $n \geq n_0$, a finite family $(R_j^n)_{1 \leq j \leq J}$ of elements of $\mathscr{T}_N^n$, such that

$$P^n[T_i^n(u) \leq N, T_i^n(u) \notin \textstyle\bigcup_{1 \leq j \leq J}(R_j^n + \sigma, R_j^n + \delta)] \leq \varepsilon. \qquad \square$$

This condition means that, with a large probability, and for all $n \geq n_0$, there is a finite family of stopping times R_j^n such that one of the R_j^n's is "uniformly" (in

n) close to $T_i^n(u)$ on the left, but at the same time "uniformly not too close" from $T_i^n(u)$. Note that when $X^n = X$ does not depend on n, 5.31 is trivially met (with $J = 1$) because $T_i^n(u)$ is predictable.

We will now prove that (Ci) $\Rightarrow$ 5.31 for all $i = 1, 2, 3, 4, 5$.

5.32 **Lemma.** *Condition* (C2) *implies condition 5.31.*

Proof. Take N, u, i, ε, δ like in 5.31. Recall that $G^\infty = G \prec F$, where F is deterministic. So if we call $t_1, t_2, \dots$ the successive times where $\Delta F > u$, then $T_i^\infty(u)$ takes only the values $t_1, t_2, \dots$

If $t_1 > N$, then $T_i^\infty(u) > N$. But $u \notin U$, so we saw in 5.15 that $T_i^n(u) \xrightarrow{\mathscr{L}} T_i^\infty(u)$, hence $P^n(T_i^n(u) \leq N) \to 0$ as $n \uparrow \infty$, and 5.31 is trivial.

Suppose now that $t_1 \leq N$, and let $J \in \mathbb{N}^*$ be such that $t_J \leq N < t_{J+1}$ (so J depends on N and u only). Using again $T_i^n(u) \xrightarrow{\mathscr{L}} T_i^\infty(u)$ we see that there is $n_0 \in \mathbb{N}^*$ such that

$$n \geq n_0 \Rightarrow P^n\left[T_i^n(u) \leq N, T_i^n(u) \notin \bigcup_{1 \leq j \leq J} \left(t_j - \frac{\delta}{3}, t_j + \frac{\delta}{3}\right)\right] \leq \varepsilon$$

(use again that $T_i^\infty(u)$ takes only the values $t_1, \dots, t_J$ when $T_i^\infty(u) \leq N$). Then $\sigma = \delta/3$ and $R_j^n = t_j - 2\delta/3$ clearly meet 5.31. $\square$

5.33 **Lemma.** *Condition* (C3) *implies condition 5.31.*

Proof. Take N, u, i, ε, δ like in 5.31. Recall that all processes are defined on the same space $(\Omega, \mathscr{F}, P)$, and that $G^\infty = G \prec F$ where F is an increasing process that is predictable with respect to the filtration $\mathbf{F} = \bigcap_n \mathbf{F}^n$.

Set $S_0 = 0$, $S_{j+1} = \inf(t > S_j : \Delta F_t \geq u)$. These times S_j are predictable. Hence each S_j admits an announcing sequence, and thus we easily find an $\mathbf{F}$-stopping time R_j such that $R_j \leq N$ and $R_j < S_j$ a.s. and $P(R_j \leq S_j - \delta/2, S_j \leq N) \leq \varepsilon/8J$, where J is the smallest integer with $J \geq 1$ and $P(S_{J+1} \leq N) \leq \varepsilon/4$ (so J *depends on* N, u, ε *only*). Since $R_j < S_j$ a.s. we can also find $\theta \in (0, \delta/2)$ such that $P(R_j \geq S_j - \theta) \leq \varepsilon/8J$ for all $j \leq J$. Hence

$$\text{5.34} \quad P\left(\bigcup_{j \leq J} \left\{S_j \notin \left(R_j + \theta, R_j + \frac{\delta}{2}\right)\right\} \cup \{S_{J+1} \leq N\}\right)$$
$$\leq P(S_{J+1} \leq N) + \sum_{j \leq J} \left[P\left(R_j \leq S_j - \frac{\delta}{2}, S_j \leq N\right) + P(R_j \geq S_j - \theta)\right] \leq \frac{\varepsilon}{2}.$$

Next, $u \notin U$ and so 2.7 yields $T_i^n(u) \xrightarrow{P} T_i^\infty(u)$. Moreover, since $G \prec F$, $[\![T_i^\infty(u)]\!] \subset \bigcup_{j \geq 1} [\![S_j]\!]$. Therefore there exists $n_0 \in \mathbb{N}^*$ such that

$$n \geq n_0 \Rightarrow P\left(T_i^n(u) \leq N, |T_i^n(u) - S_j| \geq \frac{\theta}{2} \text{ for all } j \geq 1\right) \leq \frac{\varepsilon}{2}.$$

Hence 5.34 yields for $n \geq n_0$:

$$P\left[T_i^n(u) \leq N, T_i^n(u) \notin \bigcup_{j \leq t}\left(R_j + \frac{\theta}{2}, R_j + \frac{\theta}{2} + \frac{\delta}{2}\right)\right] \leq \varepsilon.$$

Therefore taking $\sigma = \theta/2$ and $R_j^n = R_j$ for all $n \geq n_0$ yields 5.31. □

5.35 **Lemma.** *Condition* (C5) *implies condition 5.31.*

Proof. Take N, u, i, ε, δ like in 5.31. Recall that ξ is the canonical process on $\mathbb{D}(\mathbb{R}^{d+1})$, and we set $\zeta(\alpha) = t^i(\alpha^{d+1}, u)$ where α^{d+1} is the $(d+1)$th-component of α (see 2.6). Recall that by 2.7, $\alpha \rightsquigarrow \zeta(\alpha)$ is continuous on $\mathbb{D}(\mathbb{R}^{d+1})$ whenever $u \notin U(\alpha^{d+1})$.

For simplicity, we write $\mathbb{D} = \mathbb{D}(\mathbb{R}^{d+1})$, and accordingly we write the σ-fields $\mathscr{D}$, $\mathscr{D}_t$ and $\mathscr{D}_t^P$ (the latter is the P-completion in the sense of I.1.4, where P is the limiting measure showing in (C5)). The proof will go through several steps.

a) Let us define a positive measure μ on $(\mathbb{D} \times \mathbb{R}_+, \mathscr{D} \otimes \mathscr{R}_+)$ by

$$\mu(\phi) = \int P(d\alpha) 1_{\{\zeta(\alpha) < \infty\}} \phi(\alpha, \zeta(\alpha)).$$

We call C_{ad} the set of all positive bounded functions on $\mathbb{D} \times \mathbb{R}_+$ that are continuous for the product topology, and are adapted to $(\mathscr{D}_t)_{t \geq 0}$ when considered as processes on $\mathbb{D}$. We will prove that:

5.36 $$\mu(B) = \begin{cases} \sup\{\mu(\phi = 0) : \phi \in C_{ad}, \{\phi = 0\} \subset B\} \\ \inf\{\mu(\phi > 0) : \phi \in C_{ad}, B \subset \{\phi > 0\}\}. \end{cases}$$

for every set B belonging to the predictable σ-field $\mathscr{P}$ on $\mathbb{D} \times \mathbb{R}_+$ associated to the filtration $\mathbf{D}^P$. Passing to complements, we see that it suffices to prove the second equality in 5.36.

Since the collection $(\{\phi > 0\} : \phi \in C_{ad})$ is stable under finite intersection and countable union, it is enough to prove 5.36 for a semi-algebra generating $\mathscr{P}$. Thus apply I.2.2ii: after noticing that $A \times (s,t] = \bigcap_{(n)} A \times (s + \frac{1}{n}, t]$, we obtain that $\mathscr{P}$ is generated by the semi-algebra of all sets of the form $A \times \{0\}$ $(A \in \mathscr{D}_0^P)$ or $A \times (s,t]$ $(r < s < t, A \in \mathscr{D}_r^P)$. Since $\mu(\mathbb{D}(\mathbb{R}) \times \{0\}) = 0$ by construction (we have $\zeta > 0$ identically), we see that it is enough to prove 5.36 for $B = A \times (s,t]$, with $A \in \mathscr{D}_r^P$ and $r < s < t$.

Next, consider 1.14c: $\mathscr{D}_r^P$ is contained in the completion of the σ-field generated by all sets of the form $\{f > 0\}$, where f belongs to the class C_s of all $\mathscr{D}_s$-measurable bounded positive functions that are continuous for Skorokhod topology. Since the collection $(\{f > 0\} : f \in C_s)$ is stable under finite intersection and countable union, it follows by a monotone class argument that

$$P(A) = \inf(P(f > 0) : f \in C_s, \{f > 0\} \supset A).$$

So let (f_n) be a decreasing sequence in C_s, such that $A \subset \{f_n > 0\}$ and that $P(A) = \lim_{(n)} \downarrow P(f_n > 0)$. Let also (h_n) be a decreasing sequence of positive continuous functions on $\mathbb{R}_+$, such that $\{h_n > 0\} = (s, t + \frac{1}{n})$, and put $\phi_n(\alpha, v) = f_n(\alpha)h_n(v)$. It is clear that $\phi_n \in C_{ad}$ and that $B \subset \{\phi_n > 0\}$, and also that

$$\mu(B) = \lim_{(n)} \downarrow \mu(\phi_n > 0).$$

This finishes the proof of 5.36.

b) Condition (C5) implies that ξ is $\mathbf{D}^P$-predictable and P-a. s. increasing, hence $[\![\zeta]\!] \in \mathscr{P}$, where as usual $[\![\zeta]\!] = \{(\alpha, t): t < \infty, t = \zeta(\alpha)\}$. Thus we may apply 5.36 to $B = [\![\zeta]\!]$, hence obtaining a function ϕ in C_{ad} with

5.37 $$\{\phi = 0\} \subset [\![\zeta]\!]$$

5.38 $$P(\alpha: \zeta(\alpha) < \infty, \phi(\alpha, \zeta(\alpha)) > 0) \leq \frac{\varepsilon}{4}.$$

As $\phi \in C_{ad}$, the function $\psi(\alpha, t) = \inf_{s \leq t} \phi(\alpha, s)$ is also in C_{ad} and 5.38 *a-fortiori* holds for ψ. In view of 5.37, we have $\psi(\alpha, s) > 0$ for $s < \zeta(\alpha)$, whereas $\phi(\alpha, \zeta(\alpha)-) = 0$ if $\phi(\alpha, \zeta(\alpha)) = 0$. The function $\psi(\alpha, \cdot)$ being decreasing, as a consequence of 5.38 one can find $a > 0$ and $\sigma \in (0, \delta)$ such that

5.39 $$P(\{\zeta \leq N\} \cap \{\alpha: \psi(\alpha, \zeta(\alpha) - \delta) \leq a \text{ or } \psi(\alpha, \zeta(\alpha) - \sigma) \geq a\}) \leq \frac{\varepsilon}{2}.$$

c) Set $\rho(\alpha) = \inf(t: \psi(\alpha, t) \leq a) \wedge N$. We define the following functions on Ω^n: $R^n = \rho((Y^n, G^n))$ and $T^n = \zeta((Y^n, G^n))$ (recall that (Y^n, G^n) may be considered as a mapping: $\Omega^n \to \mathbb{D}$). By definition of ζ, we have $T^n = T_i^n(u)$. Since $\psi \in C_{ad}$ and ψ is decreasing in time, we have $\{R^n \leq t\} = \{\omega: \psi((Y^n(\omega), G^n(\omega)), t) \leq a\}$ (resp. $= \Omega^n$) for $t < N$ (resp. $t \geq N$): hence $\{R^n \leq t\} \in \mathscr{F}_t^n$ and so $R^n \in \mathscr{T}_N^n$.

Since $u \notin U$, $\zeta(\cdot)$ is P-a. s. continuous, and thus $\psi(\cdot, \zeta(\cdot) - \delta)$ and $\psi(\cdot, \zeta(\cdot) - \sigma)$ also are P-a. s. continuous. Then the convergence $\mathscr{L}(Y^n, G^n) \to P$ and 5.39 imply

5.40 $$\begin{aligned}\limsup_n P^n[\{T^n \leq N\} \cap \{&\psi((Y^n, G^n), T^n - \delta) \leq a \text{ or} \\ &\psi((Y^n, G^n), T^n - \sigma) \geq a\}] \leq \frac{\varepsilon}{2}.\end{aligned}$$

The definition of R^n and the continuity and decreasingness of $\psi(\alpha, \cdot)$ give

$$\begin{aligned}\{T^n \leq N, R^n \notin (T^n - \delta, T^n - \sigma)\} \subset \{T^n \leq N\} \cap \{&\psi((Y^n, G^n), T^n - \delta) \leq a \text{ or} \\ &\psi((Y^n, G^n), T^n - \sigma) \geq a\}.\end{aligned}$$

This inclusion, together with 5.40, yield 5.31 with $J = 1$ and $R_1^n = R^n$ (recall once more that $T^n = T_i^n(u)$). □

4. In view of 5.9, the three preceding lemmas show that under the assumptions of 5.17, we have 5.31. Before proceeding to the proof of 5.17, we derive an easy consequence of 5.31.

5.41 **Lemma.** *Condition 5.31 implies that for all $N > 0, \eta > 0, u \notin U, S^n \in \mathscr{T}_N^n$, we have*

$$\lim_{\sigma \downarrow 0} \limsup_{(n)} P^n\left(\bigcup_{i\geq 1} \{T_i^n(u) - \sigma < S^n < T_i^n(u) + \sigma, \right.$$

$$\left. S^n \neq T_i^n(u), |\Delta X_{S^n}^n| \geq \eta\}\right) = 0.$$

Proof. Fix $N > 0$, $\eta > 0$, $u \notin U$, $S^n \in \mathscr{T}_N^n$, and let $\varepsilon > 0$. According to 5.29 there exist $q, n_0 \in \mathbb{N}^*$ such that

$$n \geq n_0 \Rightarrow P^n\left(\bigcup_{i\geq 1}\left\{T_i^n(u) \leq N+1, T_i^n(u) - T_{i-1}^n(u) < \frac{N+1}{q}\right\}\right) \leq \varepsilon,$$

hence

5.42 $$n \geq n_0 \Rightarrow P^n(T_q^n(u) \leq N+1) \leq \varepsilon.$$

Next we consider 5.31: to $N+1$, u, i, $\frac{\varepsilon}{q}$ we associate the integer $J_i\left(= J\left(N+1, u, i, \frac{\varepsilon}{q}\right)\right)$ defined in 5.31, and we set $J = \sum_{1\leq i\leq q} J_i$. Before applying 5.31, we choose $u' \notin U$, smaller than $u \wedge [\varepsilon/8Jdf(\eta/2)]$. 5.29 implies the existence of $\theta > 0$, and $n_1 \geq n_0$ with

5.43 $$n \geq n_1 \Rightarrow P^n(A^n) \leq \varepsilon, \text{ where}$$

$$A^n = \bigcup_{i\geq 1} \{T_i^n(u') \leq N+1, T_i^n(u') - T_{i-1}^n(u') < \theta\}.$$

Now, 5.30 implies the existence of $\theta' > 0$ and $n_2 \geq n_1$ with

5.44 $$n \geq n_2,\ V^n \in \mathscr{T}_{N+1}^n \Rightarrow P^n\left(G_{V^n+\theta'}^n(u') - G_{V^n}^n(u') > \frac{\varepsilon}{4dJf(\eta/2)}\right) \leq \frac{\varepsilon}{4dJ}.$$

Then we can apply 5.31, simultaneously for all $i \leq q$, with $\delta = \theta \wedge \frac{\theta'}{2} \wedge 1$. There exists $n_3 \geq n_2$ and $\sigma \in (0, \delta)$ and families $(R_j^n)_{j\leq J}$ of elements of $\mathscr{T}_N^n$ with

5.45 $$n \geq n_3, i \leq q \Rightarrow P^n(B_i^n) \leq \frac{\varepsilon}{q}, \quad \text{where}$$

$$B_i^n = \left\{T_i^n(u) \leq N+1, T_i^n(u) \notin \bigcup_{j\leq J} (R_j^n + \sigma, R_j^n + \delta)\right\}.$$

Then, use Lemma 5.25: applying the domination inequality 5.26 and 5.44 to $V^n = R_j^n$ yields

5.46 $$n \geq n_3, j \leq J \Rightarrow P^n(C_j^n) \leq \frac{2\varepsilon}{J}, \quad \text{where}$$

$$C_j^n = \left\{ \sup_{R_j^n \leq s \leq R_j^n + \theta'} |X_s^n(u') - X_{R_j^n}^n(u')| \geq \frac{\eta}{2} \right\}.$$

Therefore if we set

$$D^n = \{T_q^n(u) < N + 1\} \cup A^n \cup \left(\bigcup_{i \leq q} B_i^n \right) \cup \left(\bigcup_{j \leq J} C_j^n \right),$$

we deduce from 5.42, 5.43, 5.45 and 5.46: $n \geq n_3 \Rightarrow P^n(D^n) \leq 5\varepsilon$.

Now let us consider $\omega \notin D^n$. Suppose that $T_i^n(u) - \sigma < S^n < T_i^n(u) + \sigma$ for some i, and $S^n \neq T_i^n(u)$, and $|\Delta X_{S^n}^n| \geq \eta$. Since $T_q^n(u) \geq N + 1$ and $\sigma \leq 1$, we have $i \leq q$. Since $u' \leq u$, $T_i^n(u) = T_k^n(u')$ for some k. But $\sigma \leq \theta$, and $\omega \notin A^n$, so $|S^n - T_i^n(u)| < \sigma$ implies that S^n is different from all $T_j^n(u')$: hence $\Delta X_{S^n}^n = \Delta X_{S^n}^n(u')$. But then, $\omega \notin C_j^n$ while $|\Delta X_{S^n}^n(u')| \geq \eta$, so S^n does not belong to any of the intervals $(R_j^n, R_j^n + \theta']$. However, $\omega \notin B_i^n$, so $T_i^n(u) \in (R_j^n + \sigma, R_j^n + \delta)$ for some $j \leq J$ and thus $S^n \in (R_j^n, R_j^n + \delta + \sigma)$; since $\delta + \sigma \leq \theta'$, we get $S^n \in (R_j^n, R_j^n + \theta']$, thus bringing a contradiction. Therefore,

$$H^n := \bigcup_{i \geq 1} \{T_i^n(u) - \sigma < S^n < T_i^n(u) + \sigma, S^n \neq T_i^n(u), |\Delta X_{S^n}^n| \geq \eta\} \subset D^n,$$

hence $P^n(H^n) \leq 5\varepsilon$ if $n \geq n_3$. Since $\varepsilon > 0$ is arbitrary, we have the result. □

4. *Proof of Theorem 5.17.* We will prove that (X^n) satisfies the conditions of Theorem 3.21.

Firstly, we deduce from 5.25 (with $u = \infty$) that for $N > 0$,

$$P^n\left(\sup_{s \leq N} |X_s^n| > 2a \right) \leq P^n(|X_0^n| > a) + P^n\left(\sup_{s \leq N} |X_s^n - X_0^n| > a \right)$$

$$\leq P^n(|X_0^n| > a) + 4d\varepsilon f(a) + 4d\, P^n(G_N^n \geq \varepsilon).$$

The sequence (G^n) is tight, so if $\eta > 0$ there exists $\varepsilon > 0$ such that $4dP^n(G_N^n \geq \varepsilon) \leq \eta$ for all $n \in \mathbb{N}^*$. Since $\lim_{a \uparrow \infty} f(a) = 0$ we may choose, by using condition 5.17(i), $a > 0$ such that $4d\varepsilon f(a) \leq \eta$ and that $P^n(|X_0^n| > a) \leq \eta$ for all $n \in \mathbb{N}^*$. Therefore

$$\sup_{(n)} P^n\left(\sup_{s \leq N} |X_s^n| > 2a \right) \leq 3\eta$$

and we deduce that condition 3.21i is fulfilled.

It remains to prove 3.21ii, and this will be achieved through a number of steps.

a) Let $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$ be given. If $u_0 \notin U$, Lemmas 5.25 and 5.30 and the fact that $G^n(u) \prec G^n(u_0)$ for $u \leq u_0$, give:

$$\lim_{\theta\downarrow 0}\limsup_{(n)}\sup_{u\in(0,u_0];S,T\in\mathcal{T}_N^n,S\le T\le S+\theta} P^n(|X_T^n(u)-X_S^n(u)|\ge\eta)$$

$$\le 4d\inf_{\rho>u_0}\left[\rho f(\eta)+\lim_{\theta\downarrow 0}\limsup_{(n)}\sup_{S\in\mathcal{T}_N^n} P^n(G_{S+\theta}^n(u_0)-G_S^n(u_0)\ge\rho)\right]$$

5.47 $$\le 4du_0 f(\eta).$$

Thus we choose $u_0\notin U$ with $u_0<\varepsilon/4df(\eta)$; there exist $\delta>0$, $n_0\in\mathbb{N}^*$ so that

5.48 $$n\ge n_0\Rightarrow \sup_{u\in(0,u_0];S,T\in\mathcal{T}_N^n,S\le T\le S+\delta} P^n(|X_T^n(u)-X_S^n(u)|\ge\eta)\le\varepsilon.$$

Next we choose $q\in\mathbb{N}^*$ such that $q\delta>2N$, and $u\notin U$ such that $u\le u_0$ and $u<\varepsilon/4dqf(\eta)$. We set $S_0^n=0$, $S_{k+1}^n=\inf(t>S_k^n:|X_t^n(u)-X_{S_k^n}^n(u)|\ge\eta)$. Since $u\le u_0$, 5.48 yields

5.49 $$n\ge n_0,\ k\ge 1\Rightarrow P^n(S_{k+1}^n\le N, S_{k+1}^n\le S_k^n+\delta)\le\varepsilon.$$

Similarly, using 5.47 and since $u<\varepsilon/4dqf(\eta)$ we may find $n_1\ge n_0$ and $\theta>0$ such that

5.50 $$n\ge n_1,\ k\ge 1\Rightarrow P^n(S_{k+1}^n\le N, S_{k+1}^n\le S_k^n+\theta)\le\frac{\varepsilon}{q}.$$

We remark that 5.49 and 5.50 are exactly 4.9 and 4.10, hence we deduce 4.12, that is:

$$\begin{cases} n\ge n_1\Rightarrow P^n(A^n)\ge 1-3\varepsilon, \quad \text{where} \\ A^n=\{S_q^n\ge N\}\cap\left(\bigcap_{1\le k\le q}\{S_k^n-S_{k-1}^n>\theta\}\right). \end{cases}$$

Now we use 5.29 to obtain $n_2\ge n_1$ and $\sigma'>0$ such that

$$n\ge n_2\Rightarrow P^n(B^n)\le\varepsilon,\quad \text{where } B^n=\bigcup_{i\ge 1}\{T_i^n(u)\le N, T_i^n(u)-T_{i-1}^n(u)<\sigma'\}.$$

Finally we use 5.41 to obtain $n_3\ge n_2$ and $\sigma\in(0,\frac{\theta}{2}\wedge\sigma']$ such that

$$\begin{cases} n\ge n_3\Rightarrow P^n(C^n)\le\varepsilon, \quad \text{where} \\ C^n=\displaystyle\bigcup_{i\ge 1,1\le k\le q}\{T_i^n(u)-\sigma<S_k^n<T_i^n(u)+\sigma, S_k^n\ne T_i^n(u), S_k^n\le N, |\Delta X_{S_k^n}^n|>\eta\}. \end{cases}$$

Putting the last three inequalities together yields

5.51 $$n\ge n_3\Rightarrow P^n(D^n)\ge 1-5\varepsilon,\quad \text{where } D^n=A^n\cap(B^n)^c\cap(C^n)^c.$$

b) Now we proceed to the construction, for each fixed $\omega\in D^n$, of a subdivision $0=t_0<\cdots<t_r=N$ satisfying $t_i-t_{i-1}\ge\sigma$ for $i\le r-1$, and

5.52 $$w(X^n(\omega);[t_{i-1},t_i))\le 8\eta \quad \text{for } i\le r.$$

This will imply that $D^n \subset \{w'_N(X^n, \sigma) \leq 8\eta\}$, and then 5.51 and the arbitrariness of $\varepsilon > 0$ will give condition 3.21ii, thus finishing the proof.

The subdivision consists in all points $T_i^n(u) < N$, and in all points $S_k^n < N$ which do not belong to any of the intervals $(T_i^n(u) - \sigma, T_i^n(u) + \sigma)$, and the point N, all rearranged in the natural order so as to form the sequence $0 = t_0 < \cdots < t_r = N$. Then each interval $[t_{p-1}, t_p)$ thus obtained belongs to one of the following four classes:

(i) $[S_k^n, S_{k+1}^n \wedge N)$ for some $k \leq q$;

(ii) $[S_k^n, T_i^n(u) \wedge N)$, in which case $S_k^n + \sigma \leq T_i^n(u) < S_{k+2}^n$ (recall that $\theta \geq 2\sigma$ and that $\omega \in A^n$), and $T_i^n(u) - \sigma < S_{k+1}^n$

(iii) $[T_i^n(u), S_k^n \wedge N)$, in which case $S_{k-2}^n < T_i^n(u) \leq S_k^n - \sigma$ and $S_{k-1}^n < T_i^n(u) + \sigma$;

(iv) $[T_i^n(u), T_{i+1}^n(u) \wedge N)$, in which case there are at most two points S_k^n and S_{k+1}^n inside the interval and each one of these lies at a distance less than σ from either $T_i^n(u)$ or $T_{i+1}^n(u)$.

Every interval $[t_{p-1}, t_p)$ is actually included into some interval $[T_{i-1}^n(u), T_i^n(u))$ for some i, thus $w(X^n; [t_{p-1}, t_p)) = w(X^n(u); [t_{p-1}, t_p))$. This, and the definition of S_{k+1}^n, imply that in case (i) above:

$$w(X^n; [t_{p-1}, t_p)) \leq 2\eta. \tag{5.53}$$

In case (ii), either $S_{k+1}^n \geq T_i^n(u)$ and we have 5.53, or $T_i^n(u) - \sigma < S_{k+1}^n < T_i^n(u)$ and $S_{k+2}^n > T_i^n(u)$; in the latter case $|\Delta X_{S_{k+1}^n}^n| \leq \eta$ because $\omega \notin C^n$, hence

$$w(X^n; [t_{p-1}, t_p)) \leq w(X^n; [S_k^n, S_{k+1}^n)) + \eta + w(X^n; [S_{k+1}^n, S_{k+2}^n)) \leq 5\eta. \tag{5.54}$$

In case (iii) we also obtain $w(X^n; [t_{p-1}, t_p)) \leq 5\eta$. Finally consider case (iv). Then either there is no S_k^n in the interval and 5.53 holds; or there is one and the same argument as above shows 5.54; or there are two: $T_i^n(u) < S_k^n < T_i^n(u) + \sigma$, and $T_{i+1}^n(u) - \sigma < S_{k+1}^n < N$; then $|\Delta X_{S_k^n}^n| \leq \eta$ and $|\Delta S_{S_{k+1}^n}^n| \leq \eta$ and the same argument again shows that

$$\begin{aligned} w(X^n; [t_{p-1}, t_p)) &\leq w(X^n; [T_i^n(u), S_k^n)) + \eta + w(X^n; [S_k^n, S_{k+1}^n)) \\ &\quad + \eta + w(X^n; [S_{k+1}^n, T_{i+1}^n(u))) \\ &\leq 8\eta. \end{aligned}$$

Hence, in all cases we have 5.52, and this finishes the proof. □

6. Convergence, Quadratic Variation, Stochastic Integrals

An aim of this section is to prove that when a sequence (X^n) of semimartingales converges in law to a semimartingale X, then the quadratic variations $[X^n, X^n]$ also converge in law to $[X, X]$, under a mild additional assumption.

Another aim is to study the convergence of stochastic integrals: as we will see, these two questions are closely related one with the other, and the solution of both relies on a technical condition which we introduce in § 6a below; see also Chapter IX, Section 5, for another method to deal with the same question.

§ 6a. The P-UT Condition

For each integer n let $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$ be a stochastic basis. We denote by $\mathscr{H}^n$ the set of all predictable processes H^n on $\mathscr{B}^n$ having the form

$$H_t^n = Y_0 1_{\{0\}} + \sum_{i=1}^{k} Y_i 1_{(s_i, s_{i+1}]}(t)$$

with $k \in \mathbb{N}$, $0 = s_0 < \ldots < s_{k+1}$, and Y_i is $\mathscr{F}_{s_i}^n$–measurable with $|Y_i| \leq 1$. If X^n is any 1–dimensional process on $\mathscr{B}^n$ and if $H^n \in \mathscr{H}^n$ is as above, we define the "elementary" stochastic integral process $H^n \cdot X^n$ (see I-4.30) by

$$H^n \cdot X_t^n = \sum_{i=1}^{k} Y_i (X_{t \wedge s_{i+1}}^n - X_{t \wedge s_i}^n).$$

6.1 **Definition.** A sequence $(X^n = (X^{n,i})_{1 \leq i \leq d})_{n \in \mathbb{N}}$ of *adapted càdlàg d–dimensional processes, each X^n on the basis $\mathscr{B}^n$, is said to be P-UT* (for "Predictably Uniformly Tight") if for every $t > 0$ the family of random variables $(\sum_{1 \leq i \leq d} H^{n,i} \cdot X_t^{n,i} : n \in \mathbb{N}, H^{n,i} \in \mathscr{H}^n)$ is tight in $\mathbb{R}$, that is

$$\lim_{a \uparrow \infty} \sup_{H^{n,i} \in \mathscr{H}^n, n \in \mathbb{N}} P^n(|\sum_{i=1}^{d} H^{n,i} \cdot X_t^{n,i}| > a) = 0.$$

6.2 **Remark.** Suppose that $\mathscr{B}^n = \mathscr{B}$ and $X^n = X$ for all n. The result of Bichteler, Dellacherie and Mokobodzki referred to in Remark I-4.32 states that X is a semimartingale if and only if the "stationary" sequence X has P-UT. So in 6.1, each X^n is necessarily a semimartingale on $\mathscr{B}^n$. In fact we do not want to use this difficult result here, so in each statement we will add the (unnecessary) condition that the X^n's are semimartingales. □

From the very definition we deduce some simple properties:

6.3 (X^n) is P-UT if and only if each sequence $(X^{n,i})$ is P-UT.

6.4 If (X^n) and (Y^n) are P-UT, then so is $(X^n + Y^n)$.

6.5 (X^n) is P-UT if and only if $(X^n - X_0^n)$ is P-UT.

6.6 If each X^n is locally of bounded variation with variation process denoted by $\mathrm{Var}(X^n)$, then (X^n) is P-UT as soon as for each $t > 0$ the sequence $(\mathrm{Var}(X^n)_t)_{n \in \mathbb{N}}$ is tight.

Next, we give deeper results.

6.7 **Proposition.** *If (X^n) is a sequence of 1–dimensional semimartingales having P-UT, for each $t > 0$ the two sequences $(\sup_{s\leq t}|X^n_s - X^n_0|)_{n\in\mathbb{N}}$ and $([X^n, X^n]_t)_{n\in\mathbb{N}}$ are tight.*

Proof. Due to 6.4 we may and will assume $X^n_0 = 0$. Fix $t > 0$ and $\varepsilon > 0$ and set $A^n_a = \{\sup_{s\leq t}|X^n_s| > a\}$. 6.1 yields the existence of $a > 0$ such that

6.8 $$\forall n \in \mathbb{N},\ \forall H^n \in \mathscr{H}^n, \qquad P^n(|H^n \cdot X^n_t| > a) \leq \varepsilon.$$

a) Let $T^n = \inf(s : |X^n_s| > a) \wedge t$, and $S^n_p = \frac{i+1}{2^p}t$ if $\frac{i}{2^p}t < T^n \leq \frac{i+1}{2^p}t$, for some $i \geq 0$. Each S^n_p is a stopping time on $\mathscr{B}^n$, taking finitely many values, so $1_{]\!]0,S^n_p]\!]}$ belongs to $\mathscr{H}^n$ and $1_{]\!]0,S^n_p]\!]} \cdot X^n_t = X^n_{S^n_p}$, and 6.8 yields $P^n(|X^n_{S^n_p}| > a) \leq \varepsilon$. But S^n_p decreases to T^n as $p \uparrow \infty$, hence $X^n_{S^n_p} \to X_{T^n}$ and by 3.1 we get

6.9 $$P^n(A^n_a) = P^n(|X^n_{T^n}| > a) \leq \liminf_{p\to\infty} P^n(|X^n_{S^n_p}| > a) \leq \varepsilon,$$

hence the first claim.

b) The proof of the second claim is based upon the following construction of the quadratic variation, given in I-4.47a: recall that $X^n_0 = 0$, and set

6.10 $$S_p(X^n) = \sum_{1\leq i\leq p} (X^n_{\frac{i}{p}t} - X^n_{\frac{i-1}{p}t})^2 = (X^n_t)^2 + H^{n,p} \cdot X^n_t,$$

where $H^{n,p} = -2\sum_{1\leq i\leq p} X^n_{\frac{i-1}{p}t} 1_{(\frac{i-1}{p}t,\frac{i}{p}t]}$. Then $S_p(X^n)$ converges to $[X^n, X^n]_t$ in P^n–measure as $p \to \infty$, so for all $c > 0$:

6.11 $$P^n([X^n, X^n]_t > c) \leq \liminf_{p\to\infty} P^n(S_p(X^n) > c)$$

Set $H^{n,p,b} = H^{n,p} \vee (b) \wedge (-b)$. Then $S_p(X^n) = (X^n_t)^2 + H^{n,p,b} \cdot X^n_t$ on the complement $(A^n_b)^c$, and $H^{n,p,b}/2b$ belongs to $\mathscr{H}_n$. Then

$$P^n(S_p(X^n) > c) \leq P^n(A^n_b) + P^n\left(\left|\frac{H^{n,p,b}}{2b} \cdot X^n_t\right| > \frac{c - b^2}{2b}\right)$$

which is smaller than 2ε by 6.8 and 6.9 whenever $c - b^2 > 2ba$ and $b > a$. Using 6.11 we get $P^n([X^n, X^n]_t > c) \leq 2\varepsilon$ if $c > 3a^2$, hence the second claim. □

6.12 **Proposition.** *Suppose that each process X^n is predictable on $\mathscr{B}^n$, with locally finite variation. The sequence (X^n) is P-UT if and only if for each $t > 0$ the sequence $(\mathrm{Var}(X^n)_t)_{n\in\mathbb{N}}$ is tight.*

Proof. Here again we may and will assume $X^n_0 = 0$. The sufficient part comes from 6.6. For the necessary part we set $Y^n = \mathrm{Var}(X^n)$. There is a predictable set A_n on $\mathscr{B}^n$ such that, if $H^n = 1_{A_n} - 1_{(A_n)^c}$, then $X^n = H^n \cdot Y^n$ and $Y^n = H^n \cdot X^n$.

Let $\varepsilon > 0$ and $t > 0$. Since (X^n) is P-UT there exists $a > 0$ with 6.8. By virtue of Lemma I-3.10, there is a stopping time T^n on $\mathscr{B}^n$ such that $E^n(Y^n_{T^n}) < \infty$ and $P^n(T^n < t) \le \varepsilon$. Now, the class of all finite unions of sets of the form $A \times \{0\}$ for $A \in \mathscr{F}^n_0$, or $A \times (s, r]$ for $0 \le s < r$ and $A \in \mathscr{F}^n_s$, is an algebra which generates the predictable σ–field on $\mathscr{B}^n$. Since $E^n(Y^n_{T^n}) < \infty$ we deduce the existence of a set B_n in this algebra, such that

$$E^n(|1_{A_n} - 1_{B_n}| \cdot Y^n_{T^n}) \le \frac{a\varepsilon}{2}.$$

From this, from $P^n(T^n < t) \le \varepsilon$, from 6.8 and the fact that $H'^n = 1_{B_n} - 1_{(B_n)^c}$ belongs to $\mathscr{H}^n$, and using the identity $H^n \cdot X^n = H'^n \cdot X^n + (1 - H^n H'^n) \cdot Y^n = H'^n \cdot X^n + 2|1_{B_n} - 1_{A_n}| \cdot Y^n$, we deduce that

$$\begin{aligned} P^n(Y^n_t > 2a) &\le P^n(T^n < t) + P^n(T^n \ge t, H^n \cdot X^n_t > 2a) \\ &\le \varepsilon + P^n(T^n \ge t, H'^n \cdot X^n_t + 2|1_{B_n} - 1_{A_n}| \cdot Y^n_t > 2a) \\ &\le \varepsilon + P^n(H'^n \cdot X^n_t > a) + P^n(|1_{B_n} - 1_{A_n}| \cdot Y^n_{T^n} > \frac{a}{2}) \\ &\le \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{aligned}$$

and the result follows. □

6.13 **Proposition.** *Suppose that each process X^n is a local martingale on $\mathscr{B}^n$, with $|\Delta X^n| \le A$ identically for some constant A. The following four properties are equivalent:*

(i) The sequence (X^n) is P-UT.

(ii) For each $t > 0$ the sequence $([X^n, X^n]_t)$ is tight.

(iii) For each $t > 0$ the sequence $(\langle X^n, X^n\rangle_t)$ is tight.

(iv) For each $t > 0$ the sequence $(\sup_{s \le t} |X^n_s - X^n_0|)$ is tight.

Proof. Let $U^n_t = \sup_{s \le t} |X^n_s - X^n_0|^2$. (iv) is equivalent to the tightness of the sequence $(U^n_t)_{n \in \mathbb{N}}$ for all $t > 0$. By Doob's inequality I-1.43, and I-4.4 and I-4.50 applied to $X^n - X^n_0$, we get for all stopping times T on $\mathscr{B}^n$:

6.14 $$E^n(U^n_T) \le 4E(\langle X^n, X^n\rangle_T) = 4E^n([X^n, X^n]_T) \le 4E^n(U^n_T).$$

Since $\Delta[X^n, X^n] \le A^2$, a straightforward application of Lenglart's inequalities I-3.31 and I-3.32 yields (ii)⇔(iii)⇒(iv). Similarly if $H^n \in \mathscr{H}^n$ we have for any stopping time T:

$$E^n((H^n \cdot X^n_T)^2) = E^n((H^n)^2 \cdot [X^n, X^n]_T) \le E(\langle X^n, X^n\rangle_T),$$

so I-3.31 again yields (iii)⇒(i). We have seen (i)⇒(ii) in 6.7.

Assume finally (iv), and let $T^n_p = \inf(t : U^n_t > p)$. Then

$$\lim_p \sup_n P^n(T^n_p < t) = 0.$$

On the other hand $\Delta U^n_{s \wedge T^n_p} \leq A(A + 2\sqrt{p})$ identically, so 6.14 and I-3.32 again give

$$\lim_{a\uparrow\infty} \sup_n P^n([X^n, X^n]_t > a) = 0$$

and, combining with the previous inequality, we obtain (ii). □

We end this paragraph with a description of P-UT in terms of the characteristics: for each n we have a d–dimensional semimartingale X^n on $\mathscr{B}^n$, with characteristics (B^n, C^n, ν^n) and second modified characteristic $\widetilde{C}^n$ (w.r.t. some given truncation $h \in \mathscr{C}_t^d$), and with associated jump measure μ^n. Recall the following processes:

$$\check{X}^n_t = \sum_{s\leq t}(\Delta X^n_s - h(\Delta X^n_s)) = (x - h(x)) \star \mu^n_t.$$

6.15 **Theorem.** *With the previous notation, the sequence (X^n) is P-UT if and only if the following three properties hold:*

(i) For each $t > 0$ the sequence $(\mathrm{Var}(\check{X}^n)_t)$ is tight.

(ii) For each $t > 0$, $i = 1, \dots, d$, the sequence $(\widetilde{C}^{n,ii}_t)$ is tight.

(iii) For all $t > 0$, $i = 1, \dots, d$, the sequence $(\mathrm{Var}(B^{n,i})_t)$ is tight.

Observe that, due to the result itself, these conditions do not depend on the choice of the truncation function h. This result is not quite in terms of the characteristics yet, but we also have:

6.16 **Theorem.** *With the previous notation, the sequence (X^n) is P-UT if and only if we have (ii) and (iii) of 6.13, and also the following two properties:*

6.17 *For all $t > 0$, $a > 0$, the sequence $(\nu^n([0,t] \times \{x : |x| > a\}))$ is tight,*

6.18 *For all $t > 0$, $\varepsilon > 0$, we have $\lim_{a\uparrow\infty} \sup_n P^n(\nu^n([0,t] \times \{x : |x| > a\}) > \varepsilon) = 0$.*

Proof of 6.15 and 6.16. a) Let us first state some auxiliary equivalences, which are proved exactly as Lemma 4.22:

$$6.19 \quad \begin{cases} 6.17 \iff \forall t > 0, \forall a > 0, \text{ the sequence } \left(\sum_{s\leq t} 1_{\{|\Delta X^n_s| > a\}}\right) \text{ is tight,} \\ 6.18 \iff \lim_{a\uparrow\infty} \sup_n P^n(\sup_{s\leq t} |\Delta X^n_s| > a) = 0 \end{cases}$$

b) Let $A > 0$ be such that $h(x) = x$ when $|x| \leq 1/A$ and $|h(x)| \leq A$. Recall that

$$X^n = X^n_0 + B^n + M^n + \check{X}^n,$$

where M^n is a local martingale on $\mathscr{B}^n$ with $|\Delta M^n| \leq A$ and $\langle M^{n,i}, M^{n,j}\rangle = \widetilde{C}^{n,ij}$. The sufficient condition in 6.15 then follows from 6.4, 6.5, 6.6 and 6.13.

b) Assume P-UT for the sequence (X^n). First, $|x - h(x)| \leq (A + A^3)|x|^2$, hence $\operatorname{Var}(\check{X}^{n,i}) \leq (A+A^3)[X^n, X^n]$; so 6.7 yields (i) and 6.6 implies that $(\check{X}^n)$ is P-UT. Then the sequence $X'^n = B^n + M^n$ is P-UT as well by 6.4 and 6.5, while $|\Delta X'^n| \leq 2A$ by construction.

Now, let μ'^n be the jump measure of X'^n, and (B'^n, C'^n, ν'^n) and $\widetilde{C}'^n$ be its characteristics and second modified characteristic, w.r.t. some truncation function h' which satisfies $h'(x) = x$ when $|x| \leq 2A$: then clearly $B'^n = B^n$, and $C'^n = C^n$, and $\widetilde{C}'^n = \widetilde{C}^n$. Then $\widetilde{C}^{n,ii}$ is smaller than $C^{n,ii} + |x^i|^2 \star \nu'^n$, which is the compensator (on $\mathscr{B}^n$) of the process $F^n = [X'^{n,i}, X'^{n,i}] = C^{n,ii} + |x^i|^2 \star \mu'^n$: applying 6.7 and the P-UT property of (X'^n), we obtain that each sequence $(F^n_t)_{n\in\mathbb{N}}$ is tight; then I-3.32 and the fact that $\Delta F^n \leq 4A^2$ give (ii). Then, 6.13 and 6.4 allow to deduce that (B^n) is P-UT, and (iii) follows from 6.12.

Next, let $h_a(x) = x 1_{\{|x|\leq a\}}$, and $\check{X}^n(h_a)$ be the process $\check{X}^n$ when $h = h_a$. Since (i) holds for all truncation functions and since $\sum_{s\leq t} 1_{\{|\Delta X^n_s|>a\}} \leq \frac{1}{a}\sum_{1\leq i\leq d} \operatorname{Var}(\check{X}^{n,i}(h_a))$, we readily deduce 6.17 from (i) and 6.19. Finally (i) implies the right side of the second equivalence in 6.19, hence we have 6.18 as well.

c) We have $\operatorname{Var}(\check{X}^{n,i})_t \leq (A + a)\sum_{s\leq t} 1_{\{|\Delta X^n_s|\geq 1/A\}}$ whenever $\sup_{s\leq t}|\Delta X^n_s| \leq a$, so under 6.17 and 6.18, Property (i) is readily deduced from 6.19. □

6.20 **Corollary.** *Let (X^n) be a sequence of 1–dimensional semimartingales being P-UT. Then*

a) *For all $t > 0$, $A > 0$ the family $(|H^n \cdot X^n_t| : n \in \mathbb{N},\ H^n$ predictable on $\mathscr{B}^n$, $|H^n| \leq A)$ is tight.*

b) *If H^n is predictable on $\mathscr{B}^n$ and if for each $t > 0$ the sequence $(\sup_{s\leq t}|H^n_s|)_{n\in\mathbb{N}}$ is tight, the the sequence $(H^n \cdot X^n)$ is P-UT.*

Proof. If $|H^n| \leq A$ we have, with the notation of the previous proof:

$$|H^n \cdot X^n_t| \leq A\big(\operatorname{Var}(\check{X}^n)_t + \operatorname{Var}(B^n)_t\big) + |H^n \cdot M^n_t|\big).$$

By 6.15 the sequences $(A(\operatorname{Var}\check{X}^n)_t + \operatorname{Var}B^n)_t))_n$ and $((H^n)^2 \cdot \langle M^n, M^n\rangle_t)_n$ are tight (the latter because it is smaller than $A^2\langle M^n, M^n\rangle_t$), so the sequence $(H^n \cdot M^n_t)_n$ is tight as well (use once more I-3.31), and we obtain (a).

Since $K^n \cdot (H^n \cdot X^n) = (K^n H^n) \cdot X^n$, (a) implies (b) if we have $|H^n| \leq A$ for some constant A. For the general case, set $H^{n,a} = H^n \bigwedge(a) \bigvee(-a)$ and $A^n_a = \{\sup_{s\leq t}|H^n_s| > a\}$. We have seen that each sequence $(H^{n,a} \cdot X^n)_n$ is P-UT, and for all K^n we have $K^n \cdot (H^n \cdot X^n) = K^n \cdot (H^{n,a} \cdot X^n)$ on the complement $(A^n_a)^c$, while $\lim_{a\uparrow\infty} \sup_n P^n(A^n_a) = 0$ by hypothesis: then clearly the sequence $(H^n \cdot X^n)$ is P-UT. □

§ 6b. Tightness and the P-UT Property

For a sequence (X^n) of semimartingales, P-UT is not implied by tightness, and does not imply tightness either: for example if $X^n_t = \frac{1}{\sqrt{n}} \sum_{i \geq 1} (nt - i) 1_{[i/n,(i+1)/n)}(t)$, the sequence (X^n) converges uniformly towards 0 (hence is tight) but is not P-UT. On the opposite if $X^n_t = 1_{[1,1+1/n)}(t)$, the sequence (X^n) is P-UT but not tight.

Here is however a connection between these two notions:

6.21 Theorem. *Let (X^n) be a sequence of d–dimensional semimartingales, with characteristics and second modified characteristics (B^n, C^n, ν^n) and $\widetilde{C}^n$. If the sequence (X^n) is tight, then:*

a) *We have 6.17 and 6.18.*

b) *The sequence (X^n) is P-UT if and only if it satisfies 6.15(iii).*

Proof. a) In view of 3.22 and of Lemma 4.22, the tightness of (X^n) implies the two right sides of 6.19, hence the claim.

b) The necessary part comes from 6.15. Conversely, assume 6.15(iii). From (a) and the proof of 6.16, we have 6.15(i), so with the notation $X^n = X^n_0 + \check{B}^n + M^n + \check{X}^n$, the sequences $(\sup_{s \leq t} |\check{X}^n_s|)_{n \in \mathbb{N}}$ and $(\sup_{s \leq t} |\check{B}^n_s|)_{n \in \mathbb{N}}$ are tight, as well as $(\sup_{s \leq t} |X^n_s - X^n_0|)_{n \in \mathbb{N}}$ by hypothesis: thus the sequence $(\sup_{s \leq t} |M^n_s|)_{n \in \mathbb{N}}$ is tight, and since $|\Delta M^n| \leq A$ identically for some A we deduce from 6.13 that 6.15(ii) holds, and using once more 6.15 we obtain the sufficient part. □

§ 6c. Convergence of Stochastic Integrals and Quadratic Variation

We now proceed to the main results of this section. Here again, for each $n \in \overline{\mathbb{N}}$ we have a stochastic basis $\mathscr{B}^n$, endowed with a d–dimensional semimartingale X^n. For each n, let K^n be a predictable d–dimensional and locally bounded process on $\mathscr{B}^n$, so that the integral process $K^n \cdot X^n = \sum_{1 \leq i \leq d} K^{n,i} \cdot X^{n,i}$ is well defined. We are interested in the convergence of $K^n \cdot X^n$ towards $K^\infty \cdot X^\infty$, and for that we need of course that the pair (X^n, K^n) converges to (X^∞, K^∞) in some sense.

Even in the 1–dimensional deterministic case and with $K^n_t = k(t)$ independent of n, the Skorokhod convergence of X^n to X^∞ is not enough to yield $\int_0^t k(s) dX^n_s \to \int_0^t k(s) dX^\infty_s$ (here all X^n's have locally finite variation): we should either strenghten the convergence of X^n (to the convergence in variation; see also Section IX-5), or add some continuity property for k. The latter is what we do here.

More precisely we will assume that $K^n = H^n_-$, the left-hand limit of a càdlàg process H^n. For futher use we also introduce an auxiliary càdlàg process Z^n which plays no effective rôle but costs nothing. The next theorem is in fact the very reason for introducing the P-UT property.

6.22 **Theorem.** *For each $n \in \overline{\mathbb{N}}$ let X^n be a d–dimensional semimartingales on $\mathscr{B}^n$, and let H^n (resp. Z^n) be a $q \times d$–dimensional (resp. r–dimensional) adapted càdlàg process on $\mathscr{B}^n$, and set $(H^n_- \cdot X^n)^i = \sum_{1\leq j\leq d} H^{n,ij}_- \cdot X^{n,j}$ for $i = 1, \dots, q$. Assume that the sequence $(X^n)_{n\in\mathbb{N}}$ is P-UT.*

(a) If the sequence of $(d + dq + r)$–dimensional processes (X^n, H^n, Z^n) is tight for the Skorokhod topology, then the sequence of $(d+q+r)$–dimensional processes $(X^n, H^n_- \cdot X^n, Z^n)$ is also tight.

(b) If we have $(X^n, H^n, Z^n) \xrightarrow{\mathscr{L}} (X^\infty, H^\infty, Z^\infty)$, then $(X^n, H^n, Z^n, H^n_- \cdot X^n)$ $\xrightarrow{\mathscr{L}} (X^\infty, H^\infty, Z^\infty, H^\infty_- \cdot X^\infty)$.

(c) If further $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ for all n (the filtrations may differ) *and if $(X^n, H^n, Z^n) \xrightarrow{P} (X^\infty, H^\infty, Z^\infty)$* (convergence in measure, for the Skorokhod topology), *then $(X^n, H^n, Z^n, H^n_- \cdot X^n) \xrightarrow{P} (X^\infty, H^\infty, Z^\infty, H^\infty_- \cdot X^\infty)$.*

Proof. Set $Y^n = (X^n, H^n, Z^n)$. Since a sequence of processes is tight if and only if from each subsequence one may extract a sub-subsequence which is tight, we deduce (b)⇒(a). We will prove (b) and (c) at the same time, assuming that $Y^n \xrightarrow{\mathscr{L}} Y^\infty$ (case (b)) or $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ and $Y^n \xrightarrow{\mathrm{P}} Y^\infty$ (case (c)).

Let us recall the sets $V(\alpha)$ and $V'(\alpha)$ of 2.10. Let $p \in \mathbb{N}^*$. We construct stopping times $T^{n,p}_i$ for $i = 0, 1, \dots$ on $\mathscr{B}^n$, inductively on i, as follows: set $T^{n,p}_0 = 0$. Assuming $T^{n,p}_i$ known, set $Y^{n,p,i}_t = Y^n_{t \wedge T^{n,p}_i}$ and

$$V(n, p, i) = \{a > 0 : P^n(a \in V(Y^{n,p,i}) \cup V'(Y^{n,p,i})) = 0\};$$

this set is dense in $\mathbb{R}_+$, so we can choose $a(n, p, i) \in V(n, p, i) \cap (1/2p, 1/p]$ and set

$$T^{n,p}_{i+1} = \inf(t > T^{n,p}_i : |Y^n_t - Y^n_{T^{n,p}_i}| \geq a(n, p, i) \ \text{ or } |Y^n_{t-} - Y^n_{T^{n,p}_i}| \geq a(n, p, i)),$$

that is $T^{n,p}_{i+1} = S_{a(n,p,i)}(Y^n - Y^{n,p,i})$ with the notation 2.9. An induction on i, using Propositions 2.11 and 2.12, give that

6.23
$$\begin{aligned}&(Y^n, Y^{n,p,1}, \dots, Y^{n,p,i}, T^{n,p}_1, \dots, T^{n,p}_i)\\ &\quad\to (Y^\infty, Y^{\infty,p,1}, \dots, Y^{\infty,p,i}, T^{\infty,p}_1, \dots, T^{\infty,p}_i)\end{aligned}$$

in law (case (b)) or in measure (case (c)), for the product topology on $\mathbb{D}(\mathbb{R}^{(i+1)(d+dq)}) \times \mathbb{R}^i_+$.

Now, set

$$H^{n,p} = \sum_{i\geq 0} H^n_{T^{n,p}_i} 1_{]\!]T^{n,p}_i, T^{n,p}_{i+1}]\!]},$$

$$W^{n,p}_t = H^{n,p} \cdot X^n_t = \sum_{i\geq 0} H^n_{T^{n,p}_i} \left(X^n_{t\wedge T^{n,p}_{i+1}} - X^n_{t\wedge T^{n,p}_i}\right),$$

$$W^{n,p,j}_t = \sum_{0\leq i\leq j} H^n_{T^{n,p}_i} \left(X^n_{t\wedge T^{n,p}_{i+1}} - X^n_{t\wedge T^{n,p}_i}\right).$$

Then 6.23 gives, for the Skorokhod topology:

6.24 $$\forall p \in \mathbb{N}^\star,\ \forall j \in \mathbb{N}^\star, \qquad (Y^n, W^{n,p,j}) \to (Y^\infty, W^{\infty,p,j})$$

in law (case (b)) or in measure (case (c)).

Next, observe that if $w'_{t+1}(Y^n, \theta) \leq 1/2p$ for some $\theta \in (0, 1/2)$, then necessarily $T^{n,p}_{i+2} - T^{n,p}_i > \theta$ for all i such that $T^{n,p}_{i+2} \leq t$; hence $T^{n,p}_{2i} \geq t$ as soon as $i \geq t/\theta$. So the tightness of the sequence (Y^n) and Theorem 3.21 yield for all $t > 0$:

$$\forall p \in \mathbb{N}^\star, \qquad \lim_{i\uparrow\infty} \sup_n P^n(T^{n,p}_i < t) = 0.$$

Since $W^{n,p,i}_t = W^{n,p}_t$ for $t \leq T^{n,p}_i$, we deduce:

6.25 $$\forall p \in \mathbb{N}^\star, \qquad \lim_{i\uparrow\infty} \sup_n P^n(\sup_{s\leq t} |W^{n,p,i}_s - W^{n,p}_s| > 0) = 0.$$

By construction $|H^n_- - H^{n,p}| \leq 1/p$ identically, hence the P-UT property of (X^n) and 6.7 and 6.20(a) yield for all $\eta > 0$:

$$\lim_{p\uparrow\infty} \sup_n P^n(\sup_{s\leq t} |(H^n_- - H^{n,p}) \cdot X^n_s| > \eta) = 0.$$

This and 6.25 give

$$\lim_{p\uparrow\infty} \lim_{i\uparrow\infty} \sup_n P^n(\sup_{s\leq t} |(Y^n_s, H^n_- \cdot X^n_s) - (Y^n_s, H^{n,p,i} \cdot X^n_s)| > \eta) = 0.$$

Then, by 6.24 and 1.23, we deduce that $(Y^n, H^n_- \cdot X^n)$ converges to $(Y^\infty, H^\infty_- \cdot X^\infty)$ in law (resp. in measure). □

As a consequence we obtain a result on the convergence of quadratic variations. As before, each d–dimensional semimartingale X^n (resp. X) is defined on its stochastic basis $\mathscr{B}^n$ and admits characteristics (B^n, C^n, ν^n) w.r.t. the same truncation function h. Further, $[X^n, X^n]$ denotes the $\mathbb{R}^d \otimes \mathbb{R}^d$–valued process whose components are the co-quadratic variations $[X^{n,j}, X^{n,k}]$ (cf. § I.4e).

6.26 **Theorem.** *Assume that $X^n \xrightarrow{\mathscr{L}} X^\infty$, and that the sequence (X^n) is P-UT (or equivalently that 6.15(iii) holds). Then $(X^n, [X^n, X^n]) \xrightarrow{\mathscr{L}} (X^\infty, [X^\infty, X^\infty])$ in $\mathbb{D}(\mathbb{R}^d \times (\mathbb{R}^d \otimes \mathbb{R}^d))$* (and in particular, $[X^n, X^n] \xrightarrow{\mathscr{L}} [X^\infty, X^\infty]$).

Proof. That P-UT and 6.15(iii) are equivalent (for every choice of the truncation h) follows from Theorem 6.21, under the assumption $X^n \xrightarrow{\mathscr{L}} X^\infty$.

Then it is enough to apply the previous theorem with $r = d^2$ and Z^n given by its coordinates $Z^{n,ij} = X^{n,i}X^{n,j} - X^{n,i}_0 X^{n,j}_0$, and also $q = d^2$ and H^n given by its coordinates $H^{n,ij,k} = -X^{n,j}$ (resp. $= -X^{n,i}$, resp. $= 0$) if $k = i$ (resp. $k = j$, resp. otherwise); since $[X^n, X^n] = Z^n + H^n \cdot X^n$, with matrix notation, and since $X^n \xrightarrow{\mathscr{L}} X^\infty$ implies that $(X^n, H^n, Z^n) \xrightarrow{\mathscr{L}} (X^\infty, H^\infty, Z^\infty)$, our result holds. □

6.27 *Example.* It is an example where all "processes" are deterministic. Set

$$X_t^n = \sum_{1\le k\le [n^2 t]} (-1)^k \frac{1}{n}.$$

We have $|X^n| \le 1/n$, so $X^n \xrightarrow{\mathscr{L}} X$ where $X = 0$. We also have

$$[X^n, X^n]_t = \sum_{1\le k\le [n^2 t]} \left(\frac{1}{n}\right)^2 = \frac{[n^2 t]}{n^2},$$

which converges to t, while of course $[X, X] = 0$. Condition 6.15(iii) is not satisfied here, because $\mathrm{Var}(B^n)_t = [n^2 t]/n$ for all n big enough. □

6.28 **Remark.** There is a version of this theorem where the X^n's are not necessarily semimartingales (and so with P-UT replaced by another – weaker – condition, but have a quadratic variation (in the sense that the variables $S_p(X^n)$ converge in measure as $p \to \infty$ to a limit, denoted by $[X^n, X^n]_t$, for all t). This comes from the fact that the quadratic variation exists in the above sense has a-priori nothing to do with the filtration.

6.29 **Corollary.** *Suppose that the X^n's are local martingales satisfying $|\Delta X_t^n(\omega)| \le c$ identically for some constant c. Then $X^n \xrightarrow{\mathscr{L}} X^\infty$ implies that the sequence (X^n) is P-UT, and thus $(X^n, [X^n, X^n]) \xrightarrow{\mathscr{L}} (X^\infty, [X^\infty, X^\infty])$.*

Proof. Taking a truncation h satisfying $h(x) = x$ for $|x| \le c$, the first characteristic of X^n becomes $B^n = 0$, so P-UT for (X^n) follows from Theorem 6.21. □

Here is a deeper result:

6.30 **Corollary.** *Suppose that the X^n's are local martingales satisfying*

6.31 $$\sup_n E_{P^n}\left(\sup_{s\le t} |\Delta X_s^n|\right) < \infty \qquad \textit{for all } t \ge 0.$$

Then $X^n \xrightarrow{\mathscr{L}} X^\infty$ implies that the sequence (X^n) is P-UT, and thus $(X^n, [X^n, X^n]) \xrightarrow{\mathscr{L}} (X^\infty, [X^\infty, X^\infty])$.

Proof. Let $h \in \mathscr{C}_t^d$ be continuous and satisfy $|h(x)| \le |x|$. We use the notation $\check{X}^n$ (see before 6.15) and $A^n = \mathrm{Var}(\check{X}^n)$. Since $|h(x) - x| \le 2|x|$, if K_t denotes the left side of 6.31 we have

$$E_{P^n}\left(\sup_{s\le t} |\Delta A_s^n|\right) \le 2K_t.$$

Now, by II.2.30 we have $B^n = [h(x) - x] \star \nu^n$, hence

$$\mathrm{Var}(B^{n,j}) \leq \widetilde{A}^n := |h(x) - x| \star \nu^n$$

and by definition of A^n and $\widetilde{A}^n$ we have $E_{P^n}(A^n_T) = E_{P^n}(\widetilde{A}^n_T)$ for every stopping time T. Therefore Lenglart's inequality I.3.32 yields

6.32 $$P^n(\mathrm{Var}(B^{n,j})_t > b) \leq P^n(\widetilde{A}^n_t > b) \leq \frac{1}{b}(\eta + 2K_t) + P^n(A^n_t > \eta)$$

for every $\eta > 0$. Now, from $X^n \xrightarrow{\mathscr{L}} X^\infty$ and 3.16, we deduce $\check{X}^n \xrightarrow{\mathscr{L}} \check{X}^\infty$ and also $A^n \xrightarrow{\mathscr{L}} A^\infty$. Thus if $\varepsilon > 0$ and $t \geq 0$ are given, 3.21 implies the existence of $\eta > 0$ such that $P^n(A^n_t > \eta) \leq \varepsilon$, and there is a $b > 0$ such that $\frac{1}{b}(\eta + 2K_t) \leq \varepsilon$. After plugging this into 6.32, we see that condition 6.15(iii) is met and we conclude by 6.21 and 6.26 again □

§ 6c. Some Additional Results

We end this chapter with two additional results. The first one is another version of Theorem 3.21. For each n we have a d–dimensional càdlàg process X^n on a basis $\mathscr{B}^n$, and for each $p \in \mathbb{N}$ we consider a strictly increasing sequence of random times $(T^{n,p}_i)_{i\geq 0}$ with $T^{n,p}_0 = 0$ and $\lim_i T^{n,p}_i = \infty$; we also consider strictly positive reals ϱ^N_p. Two properties will be of interest:

6.33 $$\lim_p \liminf_n P^n\Big(\inf_{i: T^{n,p}_{i+1} \leq N} (T^{n,p}_{i+1} - T^{n,p}_i) > \varrho^N_p\Big) = 1 \qquad \forall N > 0,$$

6.34 $$\lim_p \limsup_n P^n\big(\sup_i w(X^n, [T^{n,p}_i, T^{n,p}_{i+1}) \cap [0, N]) \geq \varepsilon\big) = 0 \quad \forall N > 0, \forall \varepsilon > 0.$$

6.35 **Proposition.** *a) If the sequence (X^n) satisfies 3.21(i) and 6.34 for random times having 6.33 relative to some positive reals ϱ^N_p, then it is tight.*

b) If the sequence (X^n) converges in law, one has 3.21(i) with $n_0 = 1$, and there are positive numbers ϱ^N_p and $a(n, p, i)$, such that 6.33 and 6.34 are satisfied by the following times, defined recursively on i by $T^{n,p}_0 = 0$ and

6.36 $$\begin{aligned} T^{n,p}_{i+1} = \inf(t > T^{n,p}_i : |X^n_t - X^n_{T^{n,p}_i}| \\ \geq a(n,p,i) \text{ or } |X^n_{t-} - X^n_{T^{n,p}_i}| \geq a(n,p,i)). \end{aligned}$$

If X^n is adapted to some filtration, the $T^{n,q}_i$'s in 6.36 are stopping times. The reason for which we need the convergence in law of the sequence (X^n), rather than mere tightness, is that we want to have an "explicit" form for these radom times, to be able to precisely check that they are stopping times when X^n is adapted.

Proof. Assume first 6.33 and 6.34, and 3.21(i). Set $\alpha(n,N,p) = \inf_{i:T^{n,p}_{i+1}\leq N} (T^{n,p}_{i+1} - T^{n,p}_i)$. Fix $\varepsilon > 0$ and $\eta > 0$ and $N > 1$. From 6.33 and 6.34 we find p and n_0 such that $P^n(A_n) \geq 1-\varepsilon$ for all $n \geq n_0$, where

$$A_n = \{\alpha(n,N,q) > \varrho^N_q, \sup_i w(X^n, [T^{n,q}_i, T^{n,q}_{i+1}) \cap [0,N]) < \eta\}.$$

Fix n and $\omega \in A_n$: setting $t_i = T^{n,p}_i \bigwedge N$ and $k = \inf(i : t_i = N)$, we get $w(X^n, [t_i, t_{i+1})) \leq \eta$ for all i, and $t_{i+1} - t_i > \varrho^N_q$ for $i \leq k-2$, so $w'_N(X^n, \varrho^N_q) < \eta$. Thus $P^n(w'_N(X^n, \varrho^N_q) \geq \eta) \leq \varepsilon$ if $n \geq n_0$, and the sequence (X^n) is tight by Theorem 3.21.

Conversely assume that $X^n \xrightarrow{\mathscr{L}} X^\infty$. We repeat the proof of Theorem 6.22 with $Y^n = X^n$, so that the times $T^{n,p}_i$ of this proof are exactly those in 6.36. By construction $w(X^n, [T^{n,p}_i, T^{n,p}_{i+1})) \leq a(n,p,i) \leq \frac{1}{p}$ identically, so 6.34 holds. We also have 6.23, so $\alpha(n,N,p) \xrightarrow{\mathscr{L}} \alpha(\infty,N,p)$ as soon as N is not a fixed time of discontinuity for X^∞. Then letting $\varrho^N_p > 0$ be such that $P^\infty(\alpha(\infty,N,p) \leq \varrho^N_p) \leq \frac{1}{p}$, we readily deduce 6.33. □

The last result concerns the preservation of tightness (or convergence) by time-changes which are close enough to the identity. In view of the definition of the Skorokhod topology, this is of course not surprising. The setting is as follows: we have d–dimensional càdlàg processes X^n, each one possibly defined on its own probability space $(\Omega^n, \mathscr{F}^n, P^n)$ (no filtration here!), and for each n we have another real-valued càdlàg increasing process τ^n (hence $\tau^n_0 = 0$) defined on the same space as X^n, and we consider the time-changed processes $Y^n = X^n_{\tau^n_t}$.

6.37 **Proposition.** *Assume that τ^n converges in law to the identity, that is*

6.38 $$\lim_{n\to\infty} P^n(\sup_{s\leq N} |\tau^n_s - s| > \varepsilon) = 0 \qquad \text{for all } \varepsilon > 0, N > 0.$$

a) *If the sequence X^n is tight, then so is the sequence Y^n.*

b) *If the sequence X^n converges in law, then the sequence Y^n converges in law to the same limit.*

c) *Assuming further that $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ does not depend on n, if the sequence X^n converges in probability to some limit X, then so does the sequence Y^n.*

d) *Assuming again that $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ does not depend on n, if the sequence X^n converges a.s. to some limit X and if 6.38 is strenghtened as $\sup_{s\leq N}|\tau^n_s - s| \to 0$ a.s. for all $N > 0$, then the sequence Y^n also converges a.s. to X.*

This applies in particular when one discretizes the processes: if $X^n \xrightarrow{\mathscr{L}} X$ and $Y^n_t = X^n_{[nt]/n}$, then $Y^n \xrightarrow{\mathscr{L}} X$ as well.

Proof. Up to taking subsequences, it is clear that (b)$\Rightarrow$(a) and (d)$\Rightarrow$(c). Now the assumption in (b) and 6.38 imply that the pair (X^n, τ^n) converges in law to (X, I) (I is the identity: $I_t = t$). So for both (b) and (d), it is enough to prove that if two sequences $\alpha_n \in \mathbb{D}(\mathbb{R})$ and $\tau_n \in \mathscr{V}^+$ (see 2.13) converge respectively to α and to the identity I, then the sequence $\beta_n(t) = \alpha_n(\tau_n(t))$ converges in $\mathbb{D}(\mathbb{R}^d)$ to α.

For every $p > 0$, set $t_0^p = 0$ and $t_{i+1}^p = \inf(t > t_i^p : |\alpha(t) - \alpha(t_i^p)| \geq \frac{1}{p})$. Then $\varrho_p^N = \frac{1}{2}\inf(t_i^p - t_{i-1}^p : i \geq 0, t_{i-1}^p \leq N) > 0$. Choose $t_i^{n,p}$ in such a way that $\tau^n(t_i^{n,p}-) \leq t_i^p \leq \tau^n(t_i^{n,p})$. Since $\tau_n \to I$ locally uniformly, it is clear that $\inf(t_i^{n,p} - t_{i-1}^{n,p} : i \geq 0, t_{i-1}^{n,p} \leq N)$ converges to $2\varrho^N$, so the times $T_i^{n,p} = t_i^{n,p}$ satisfy 6.33 (there is no probability here). Furthermore $w(\beta_n, [t_i^{n,p}, t_{i-1}^{n,p}) \leq w(\alpha, [t_i^n, t_{i-1}^n) \leq \frac{1}{p}$ by construction, so 6.34 holds, and the family $A = \{\beta_n : n \geq 1\}$ is relatively compact by Proposition 6.35. Since $\beta_n(t) \to \alpha(t)$ when $\Delta\alpha(t) = 0$ by Proposition 2.1, we deduce that $\beta_n \to \alpha$. $\square$

Chapter VII. Convergence of Processes with Independent Increments

With this chapter, at last, we enter the subject which has given its name to the whole book. Our final aim is to prove convergence theorems for a sequence of semimartingales toward a semimartingale. We present the material through three successive steps, corresponding to Chapters VII, VIII and IX: firstly, the pre-limiting processes, as well of course as the limiting process, have independent increments; secondly, only the limiting process has independent increments; thirdly, the limiting process itself belongs to some rather broad class of semi-martingales. This method of exposition, going from the particular to the general, surely leaves way to many redundancies; it also have some advantages:

1) It allows to put together in a single chapter most of the "old" results that are due to Lévy, Khintchine, Kolmogorov, Gnedenko, etc..., although this chapter also contains some recent or new results.

2) The independent increments case shows a very simple structure, from the probabilistic point of view.

3) At the same time, most of the analytical difficulties are already present in this case. Henceforth, our presentation allows to single out the two kinds of problems, analytical and probabilistic.

4) Finally, the simple structure of PII's allows for various necessary and sufficient conditions for convergence, a fact that is untrue in general.

So we devote the present chapter to the first step as described above. That is, we consider a sequence $(X^n)_{n\geq 1}$ of $\mathbb{R}^d$-valued processes with independent increments ("PII": see Definition II.4.1) and its "potential" limit process X, which is also $\mathbb{R}^d$-valued; simple considerations based upon finite-dimensional convergence imply that X is also a PII, with respect to the filtration that it generates. Of course, each one of the processes X^n or X might be defined on its own stochastic basis; however, by taking tensor products, it is always possible to assume that:

(*) $\quad X^n$ and X are PII's defined on the same probability space $(\Omega, \mathscr{F}, P)$.

The purpose of this hypothesis, which is not a restriction, is only to ease notation.

In Section 1 we give a general presentation of the available methods for proving limit theorems, in the particularly simple case of Poisson processes.

Section 2 is devoted to finite-dimensional convergence $X^n \xrightarrow{\mathscr{L}(D)} X$ along a subset D of $\mathbb{R}_+$ in the case when X has no fixed time of discontinuity and all

X^n's are semimartingales. This property easily reduces to 1-dimensional convergence (in the sense that D has only one point), so this section really concerns convergence of sums of independent random variables; it means that we reproduce the parts of the books of Gnedenko and Kolmogorov [65] and Petrov [197] that are related to our subject.

In Section 3 we give a necessary and sufficient condition for "functional" convergence $X^n \xrightarrow{\mathscr{L}} X$. This section constitutes the heart of the chapter, and it is essentially independent from Section 2 (except for the notation and some "elementary" properties, gathered in § 2a).

In Section 4 another necessary and sufficient condition is given, and expressed in terms of the characteristic functions of the increments of the processes. We also examine finite-dimensional convergence in the general case. This section is highly technical, and should not be read at first reading.

Finally in Section 5 we specialize the previous results to the case when the limit is Gaussian (§ 5a, which is very easy), and we also give some "non-classical" conditions (§§ 5b,c,d, which are much more difficult).

1. Introduction to Functional Limit Theorems

The purpose of the present section is to describe on a very simple example (the limit process, as well as the pre-limiting processes, are Poisson processes) the methods available to prove (functional) limit theorems, and to discuss their relative merits.

1. The setting is as follows: we consider a sequence of Poisson processes X^n, with their intensity functions A^n (they are increasing continuous "deterministic" functions). Similarly the limiting process X also is a Poisson process, with intensity A. As said above, we can assume without loss of generality that all X^n and X are defined on the same probability space $(\Omega, \mathscr{F}, P)$, but of course they may generate different filtrations $\mathbf{F}^n$ and $\mathbf{F}$.

We wish to prove the:

1.1 **Theorem.** *If $A_t^n \to A_t$ for all $t \geq 0$, then $X^n \xrightarrow{\mathscr{L}} X$.*

The proof is based upon procedure VI.3.18: prove tightness for the sequence (X^n), then identify all limit points of $\{\mathscr{L}(X^n)\}$ with the law $\mathscr{L}(X)$ of X.

1.2 *Tightness of* (X^n). This property is a straightforward consequence of Theorem VI.4.18. Indeed, choose a truncation function $h \in \mathscr{C}_t^1$ such that $h(1) = 0$; the characteristics and modified second characteristic of X^n, relative to the filtration $\mathbf{F}^n$, are

$$B^n = 0, \quad C^n = 0, \quad \tilde{C}^n = 0, \quad \nu^n(dt, dx) = dA^n(t) \otimes \varepsilon_1(dx)$$

(recall e.g. II.4.15). Moreover $A_t^n \to A_t$ for all $t \geq 0$, and A is continuous, so VI.2.15c implies that $A^n \to A$ for the local uniform topology in $\mathbb{D}(\mathbb{R})$ (or $\mathbb{C}(\mathbb{R})$, here). Then the conditions of VI.4.18 are trivially fulfilled, and the sequence (X^n) is tight. □

Next, we identify the limit, via the following three methods:

1) *Finite-dimensional method.* We begin with an elementary lemma, valid when *X^n and X are arbitrary PII's.* D is a subset of $\mathbb{R}_+$.

1.3 Lemma. *For finite-dimensional convergence along D: $X^n \xrightarrow{\mathscr{L}(D)} X$, it is sufficient* (and also obviously necessary) *that $X_t^n - X_s^n \xrightarrow{\mathscr{L}} X_t - X_s$ for all $s, t \in D \cup \{0\}$.*

Proof. Let $0 = t_0 < \cdots < t_p$ with $t_j \in D$ for $j \geq 1$. If $u_j \in \mathbb{R}^d$ we have

$$E\left(\exp i \sum_{0 \leq j \leq p} u_j \cdot X_{t_j}^n\right) = \prod_{1 \leq j \leq p} E(\exp i v_j \cdot (X_{t_j}^n - X_{t_{j-1}}^n))$$

provided $v_j = \sum_{j \leq k \leq p} u_k$ for $1 \leq j \leq p$ (recall that $X_0^n = 0$), and the same holds with X in place of X^n. The result follows immediately. □

Coming back to our Poisson processes, we observe that $X_t^n - X_s^n$ (resp. $X_t - X_s$) has a Poisson distribution with mean value $A_t^n - A_s^n$ (resp. $A_t - A_s$). Hence the assumption $A_t^n \to A_t$ for all $t \geq 0$ implies that $X_t^n - X_s^n \xrightarrow{\mathscr{L}} X_t - X_s$ for all $s \leq t$, and therefore $X^n \xrightarrow{\mathscr{L}(\mathbb{R}_+)} X$ by the previous lemma. Hence VI.3.20 and 1.2 yield $X^n \xrightarrow{\mathscr{L}} X$.

(*Remark*: here we deal with point processes; hence by VI.3.37, $X^n \xrightarrow{\mathscr{L}(\mathbb{R}_+)} X$ is enough to insure that $X^n \xrightarrow{\mathscr{L}} X$, so 1.2 is useless with this method; of course, this remark does not extend to processes that are not point processes).

2) *Martingale method.* This method is based upon a very important theorem, which will be proved in Chapter IX, and whose we state (in a loose way) a version relevant to the problem at hand.

1.4 Assume $M^n \xrightarrow{\mathscr{L}} M$; if each M^n is a local martingale, and if $|\Delta M^n|$ is uniformly bounded, then M is a local martingale relative to the filtration that it generates. □

(This is "classical", at least if we replace "local martingale" by "uniformly bounded martingale").

Coming back to our problem, let $\tilde{P}$ be a limit point of the sequence $\{\mathscr{L}(X^n)\}$: there is a subsequence (X^{n_k}) such that $\mathscr{L}(X^{n_k}) \to \tilde{P}$ weakly. $\tilde{P}$ is a measure on the canonical space $\mathbb{D}(\mathbb{R})$, and we denote by ξ the canonical process on it, with $\mathbf{D}(\mathbb{R})$ the filtration generated by ξ (see VI.1.1). $M^{n_k} = X^{n_k} - A^{n_k}$ is a local martingale with $|\Delta M^{n_k}| \leq 1$. Since A is continuous, VI.2.15c yields that $A^{n_k} \to A$ uniformly on finite intervals, so by VI.1.23, $\mathscr{L}(M^{n_k})$ converges weakly to the law of $M = \xi - A$ under $\tilde{P}$. Then 1.4 yields that M is a $\tilde{P}$-local martingale for the filtration generated by M, which precisely is $\mathbf{D}(\mathbb{R})$ because A is deterministic.

Moreover the subset $\mathscr{V}^{+,1}$ of all paths of point processes is closed in $\mathbb{D}(\mathbb{R})$ (see VI.2b), so $\tilde{P}(\mathscr{V}^{+,1}) = 1$ and A is the $\tilde{P}$-compensator of ξ. Then II.4.5 implies that ξ is a Poisson process with intensity A. In other words, the law of ξ under $\tilde{P}$ equals $\mathscr{L}(X)$.

At this stage, we have proved that any limit point of $\{\mathscr{L}(X^n)\}$ equals $\mathscr{L}(X)$. In view of 1.2, this implies $X^n \xrightarrow{\mathscr{L}} X$.

3) *A method based upon a necessary condition for convergence.* The following converse to Theorem 1.1 is interesting by itself:

1.5 Theorem. *Assume that $X^n \xrightarrow{\mathscr{L}} Y$. Then Y is an extended Poisson process* (see I.3.26), *and if A' denotes the intensity function of Y then $A^n_t \to A'_t$ for all $t \in D(Y) := \{t \geq 0\colon P(\Delta Y_t \neq 0) = 0\}$.*

As a matter of fact, we will see a better result in Section 3, and in particular Y is indeed a Poisson process (or equivalently A' is continuous): see Corollary 3.14, but this is irrelevant to the present discussion.

Proof. That Y is a point process with independent increments, and so is an extended Poisson process relative to its own filtration, is trivial.

Call $(S^n_i)_{i \geq 1}$ the successive jump times of X^n, and $(S_i)_{i \geq 1}$ those of Y. We have $P(S^n_i \leq t) = P(X^n_t \geq i) \to P(Y_t \geq i) = P(S_i \leq t)$ for all $t \in D(Y)$, so $S^n_i \xrightarrow{\mathscr{L}} S_i$. We also have $\{X^n_{S^n_i \wedge t} \geq j\} = \varnothing$ if $j > i$, and $\{X^n_{S^n_i \wedge t} \geq j\} = \{X^n_t \geq j\}$ if $j \leq i$, and similarly for $Y_{S_i \wedge t}$, hence

$$E(X^n_{S^n_i \wedge t}) = \sum_{1 \leq j \leq i} P(X^n_t \geq j) \to \sum_{1 \leq j \leq i} P(Y_t \geq j) = E(Y_{S_i \wedge t})$$

for all $t \in D(Y)$.

On the other hand, $E(A^n_T) = E(X^n_T)$ and $E(A'_T) = E(Y_T)$ for all stopping times T (for the corresponding filtrations). Hence we get for all $t \in D(Y)$, as $n \uparrow \infty$:

1.6
$$\begin{aligned} \beta^n_i &:= A^n_t + E[(A^n_{S^n_i} - A^n_t)1_{\{S^n_i < t\}}] = E(A^n_{S^n_i \wedge t}) \\ \to \beta_i &:= A'_t + E[(A'_{S_i} - A'_t)1_{\{S_i < t\}}] = E(A'_{S^n_i \wedge t}). \end{aligned}$$

Let $\varepsilon \in (0, 1/2)$ and $t \in D(Y)$. There exists $i \in \mathbb{N}^*$ such that $P(S_i < t + 1) \leq \varepsilon$. Using $S^n_i \xrightarrow{\mathscr{L}} S_i$ and 1.6, we obtain $n_0 \in \mathbb{N}^*$ such that $|\beta^n_i - \beta_i| \leq \varepsilon$ and $P(S^n_i < t) \leq 2\varepsilon$ for $n \geq n_0$. Moreover 1.6 yields $A^n_t \leq \beta^n_i / P(S^n_i \geq t)$ and $\beta_i \leq A'_t$ and

$$|A_t^n - A_t'| \le |\beta_i^n - \beta_i| + A_t^n P(S_i^n < t) + A_t' P(S_i < t)$$
$$\le \varepsilon + 2\varepsilon \frac{A_t' + \varepsilon}{1 - 2\varepsilon} + \varepsilon A_t'$$

for $n \ge n_0$. Since ε is arbitrarily small, we obtain $A_t^n \to A_t'$ for all $t \in D(Y)$. □

Now we can deduce 1.1 from 1.5. If (X^{n_k}) is a subsequence which converges in law to a process Y, then by 1.5 Y is an extended Poisson process, and $A_t^{n_k} \to A_t'$ for all $t \notin J(Y)$, where A' is the intensity function of Y. Since $A_t^{n_k} \to A_t$ for all $t \ge 0$, we must have $A' = A$, so $\mathscr{L}(Y) = \mathscr{L}(X)$. In view of 1.2, we conclude that $X^n \xrightarrow{\mathscr{L}} X$.

2. As far as Poisson processes are concerned, it is clear that the first method above is the simplest one. Let us discuss the range of applicability and also the facility of application.

1) *Finite-dimensional method*: We need to have a grasp on the finite-dimensional distributions, and at least to have a kind of "explicit" form for the finite-dimensional distributions of the *limiting process*.

Needless to say, that is not very often the case, except *when X is a PII*; so theoretically speaking the scope of this method is quite narrow, even though on a practical point of view the limiting process is indeed a Wiener process in many cases, or at least a PII.

Moreover, this method is very simple for Poisson processes, or when all X^n are PII without fixed time of discontinuity; it gets significantly more difficult when the X^n's have fixed discontinuities while X stays without fixed discontinuities: below we derive those results in Section 2. It gets even more difficult in the general case (all X^n's still being PII), as the reader will notice if he reads Section 4!

Let us add that, unlike the other two methods, it does not use the tightness of (X^n) as a constitutive part. Hence it allows to obtain finite-dimensional convergence results when the sequence (X^n) is not tight (and so functional convergence does not hold).

2) *Martingale method*: We feel that it is the most useful and wide-scoped method. However, it needs some technical apparatus, which is essentially the same whether the processes X^n or X are PII or not. In fact, the key requirement is that the law of the limiting process should be characterized as the unique solution of some martingale problem, in the sense of Section III.2.

We do not use this method in the present chapter (neither do we in Chapter VIII, as a matter of fact); however, all the "functional results" below are also *particular cases* of the results of Chapter IX (at least the sufficient conditions for convergence).

3) *The method based upon necessary conditions*: It turns out that the natural sufficient conditions for $X^n \xrightarrow{\mathscr{L}} X$ are also necessary when all X^n's and X are PII, a property that is true only under these assumptions. Hence this method is usable only within the setting of the present chapter.

However, within this setting, we feel that it is the simplest method when no further assumptions are made on the PII's X^n and X: so we use this method in Section 3 below.

2. Finite-Dimensional Convergence

Despite the generality of the title, we will consider here only finite-dimensional convergence toward a PII *without fixed time of discontinuity* (the general case is treated in Section 4). § 2a serves as a refresher, and also is a good introduction to the specific form that our various conditions for limit theorems take. § 2b is technical and is used only for the necessary conditions for finite-dimensional convergence (except for the "classical" Lemma 2.16). § 2c is concerned with the behaviour of triangular arrays, while the PII's are introduced in § 2d only.

§ 2a. Convergence of Infinitely Divisible Distributions

We will take for granted the structure of infinitely divisible distributions on $\mathbb{R}^d$, and in particular the *Lévy-Khintchine formula*, which has already been encountered in § II.4d (see II.4.21).

More precisely, we consider a *truncation function* $h \in \mathscr{C}_t^d$ (see II.2.3: it is a function $h : \mathbb{R}^d \to \mathbb{R}^d$, bounded, satisfying $h(x) = x$ on a neighbourhood of 0). Moreover, we choose h to be *continuous*: this is not essential, but all the same it is not a restriction since h can be chosen arbitrarily in $\mathscr{C}_t^d$, and it considerably simplifies some statements concerning limit theorems. We denote by A a constant such that

2.1 $$|x| \leq \frac{1}{A} \Rightarrow h(x) = x, \quad |h| \leq A, \quad \text{and } A \geq 1$$

Then, a function $\varphi\colon \mathbb{R}^d \to \mathbb{C}$ is the characteristic function of an infinitely divisible distribution on $\mathbb{R}^d$ if and only if there exists a triple (b, c, F) consisting in

2.2 $$\begin{cases} b \in \mathbb{R}^d \\ c, \text{ a } d \times d \text{ symmetric nonnegative matrix} \\ F, \text{ a positive measure on } \mathbb{R}^d \text{ with } F(\{0\}) = 0 \text{ and } \int F(dx)|x|^2 \wedge 1 < \infty, \end{cases}$$

and such that $\varphi = \exp \psi_{b,c,F}$, where

2.3 $$\psi_{b,c,F}(u) = iu\cdot b - \frac{1}{2}u\cdot c\cdot u + \int (e^{iu\cdot x} - 1 - iu\cdot h(x))F(dx).$$

Moreover, we know by Lemma II.2.44 that the function $\psi_{b,c,F}$ uniquely determines the triple (b, c, F); and, if h' is another truncation function, c and F are not affected while b should be replaced by

2.4 $$b' = b + \int [h'(x) - h(x)]F(dx).$$

Another useful characteristic is the following $d \times d$ symmetric nonnegative matrix (the "modified second characteristic"):

2.5 $$\tilde{c} = (\tilde{c}^{ij}), \quad \tilde{c}^{ij} = c^{ij} + \int h^i(x)h^j(x)F(dx);$$

note that $\tilde{c}$ depends on the choice of the truncation function.

The following facts are also well known, and stated without proof (see Gnedenko and Kolmogorov [65]).

2.6 Let $(\mu_n)_{n\geq 1}$ be a sequence of infinitely divisible distributions, with characteristics (b_n, c_n, F_n), which converges weakly to μ; then μ is infinitely divisible, and if (b, c, F) are its characteristics we have $\psi_{b_n,c_n,F_n} \to \psi_{b,c,F}$ uniformly on compact subsets of $\mathbb{R}^d$. □

Before stating the main theorem, we introduce some classes of functions:

2.7 $C_2(\mathbb{R}^d)$ = the set of all continuous bounded functions: $\mathbb{R}^d \to \mathbb{R}$ which are 0 around 0.

$C_3(\mathbb{R}^d)$ = the set of all continuous bounded functions $f\colon \mathbb{R}^d \to \mathbb{R}$ satisfying $f(x) = o(|x|^2)$ when $|x| \to 0$

$C_4(\mathbb{R}^d)$ = the set of all continuous functions $f\colon \mathbb{R}^d \to \mathbb{R}$ which are 0 around 0 and such that $f(x)/|x|^2$ is bounded.

$C_1(\mathbb{R}^d)$ = is a subclass of $C_2(\mathbb{R}^d)$ having only nonnegative functions, which contains all functions $g_a(x) = (a|x| - 1)^+ \wedge 1$ for all positive rationals a, and with the following property: let η_n, η be positive measures on $\mathbb{R}^d$ which do not charge $\{0\}$ and are finite on the complement of any neighborhood of 0; then $\eta_n(f) \to \eta(f)$ for all $f \in C_1(\mathbb{R}^d)$ implies $\eta_n(f) \to \eta(f)$ for all $f \in C_2(\mathbb{R}^d)$ (so, it is a *convergence-determining class* for the weak convergence induced by $C_2(\mathbb{R}^d)$). □

We have the inclusions

2.8 $$C_1(\mathbb{R}^d) \subset C_2(\mathbb{R}^d) \subset C_3(\mathbb{R}^d), \quad C_2(\mathbb{R}^d) \subset C_4(\mathbb{R}^d),$$

and a very important fact for the next chapter is that there exists a class $C_1(\mathbb{R}^d)$ that is *countable*.

2.9 **Theorem.** *Assume that the truncation function h is continuous. Let $(\mu_n)_{n\geq 1}$ and μ be infinitely divisible distributions on $\mathbb{R}^d$ with characteristics (b_n, c_n, F_n) and (b, c, F), and define $\tilde{c}_n$ and $\tilde{c}$ by 2.5. Consider the conditions*:

$[\beta_1]$ $\quad b_n \to b$

$[\gamma_1]$ $\quad \tilde{c}_n \to \tilde{c}$

$[\delta_{1,i}]$ $\quad F_n(g) \to F(g)\quad$ for all $g \in C_i(\mathbb{R}^d)$.

Then $\mu_n \to \mu$ weakly if and only if $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,1}]$ are met. In this case, we have $[\delta_{1,2}]$ and $[\delta_{1,3}]$ as well.

Note that, since the weak convergence $\mu_n \to \mu$ has nothing to do with the truncation function h, the set of conditions $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,i}]$ is also independent of h, provided h is continuous (actually, they remain valid for all $h \in \mathscr{C}_t^d$ which are F-a.s. continuous).

Proof. a) From Definition 2.7 we deduce that $[\delta_{1,1}] \Leftrightarrow [\delta_{1,2}]$. We will prove first that $[\gamma_1] + [\delta_{1,2}] \Rightarrow [\delta_{1,3}]$. Let $g \in C_3(\mathbb{R}^d)$ and $a = \sup|g|$ and $\varepsilon > 0$ be given. There exist two positive rationals K, η such that $\eta < K$ and $F(|x| > K) \leq \varepsilon$, and $|g(x)| \leq \varepsilon |x|^2$ and $h(x) = x$ for $|x| \leq \eta$. Set $g' = g_{2/\eta} - g_{1/2K}$ and $g'' = g_{1/K}$ (see 2.7). Then

$$\begin{aligned}|F_n(gg') - F_n(g)| &\leq \int F_n(dx)|g(x)|1_{\{|x|<\eta\}} + \int F_n(dx)|g(x)|1_{\{|x|>2K\}}\\ &\leq \varepsilon \int F_n(dx)|h(x)|^2 + aF_n(g'')\\ &\leq \varepsilon \sum_{1\leq j\leq d} \tilde{c}_n^{jj} + aF_n(g''),\end{aligned}$$

and similarly for F. Now, since $g'' \in C_2(\mathbb{R}^d)$ and $gg' \in C_2(\mathbb{R}^d)$, $[\gamma_1]$ and $[\delta_{1,2}]$ yield the existence of $N \in \mathbb{N}^*$ such that

$$n \geq N \Rightarrow \begin{cases} F_n(g'') \leq F(g'') + \varepsilon \leq 2\varepsilon, & \sum_{1\leq j\leq d} \tilde{c}_n^{jj} \leq \sum_{1\leq j\leq d} \tilde{c}^{jj} + \varepsilon\\ |F_n(gg') - F(gg')| \leq \varepsilon \end{cases}$$

and we deduce that

$$\begin{aligned}n \geq N \Rightarrow |F_n(g) - F(g)| &\leq |F_n(g) - F_n(gg')| + |F_n(gg') - F(gg')|\\ &\quad + |F(gg') - F(g)|\\ &\leq \varepsilon\left[2 \sum_{1\leq j\leq d} \tilde{c}^{jj} + 3a + 2\right]\end{aligned}$$

and since $\varepsilon > 0$ is arbitrary, we have $F_n(g) \to F(g)$.

b) Secondly, we prove the sufficient condition. Assume $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,2}]$, so $[\delta_{1,3}]$ holds as well, and fix $u \in \mathbb{R}^d$. If $g_u(x) = e^{iu\cdot x} - 1 - iu\cdot h(x) + \frac{1}{2}|u\cdot h(x)|^2$, we have

$$\psi_{b_n,c_n,F_n}(u) = iu\cdot b_n - \tfrac{1}{2}u\cdot \tilde{c}_n\cdot u + F_n(g_u) \tag{1}$$

and similarly for $\psi_{b,c,F}$. Now, g_u is a complex-valued function whose real part and imaginary part are functions belonging to $C_3(\mathbb{R}^d)$ (because h is continuous and $h(x) = x$ for $|x|$ small). Hence $[\beta_1] + [\gamma_1] + [\delta_{1,3}]$ gives $\psi_{b_n,c_n,F_n}(u) \to \psi_{b,c,F}(u)$, which yields the result.

c) Finally, we prove the necessary condition. For each $w \in \mathbb{R}^d\backslash\{0\}$, set

$$\varphi_{w,n}(u) = \psi_{b_n,c_n,F_n}(u) - \frac{1}{2}\int_{-1}^{1} \psi_{b_n,c_n,F_n}(u + sw)\,ds$$

which, following the proof of Lemma II.2.44, is the characteristic function of the following positive finite measure:

$$G_{w,n}(dx) = \left[\frac{1}{6}w\cdot c_n\cdot w\right]\varepsilon_0(dx) + \left(1 - \frac{\sin w\cdot x}{w\cdot x}\right)\cdot F_n(dx).$$

We construct φ_w and G_w similarly with b, c, F. The assumption $\mu_n \to \mu$ weakly and 2.6 imply that $\varphi_{w,n}(u) \to \varphi_w(u)$ for all u, so $G_{w,n} \to G_w$ weakly.

Let $k(x) = \sum_{1\le j\le d}(1 - (\sin w_j\cdot x)/(w_j\cdot x)) = \sum_{1\le j\le d}(1 - (\sin x^j)/x^j)$, where $w_1, \dots, w_d$ is the usual orthonormal basis of $\mathbb{R}^d$, and set $G_n = \sum_{1\le j\le d} G_{w_j,n}$ and $G = \sum_{1\le j\le d} G_{w_j}$. Then $G_n \to G$ weakly, while k is continuous, bounded on $\mathbb{R}^d$, strictly positive outside 0, and $k(x) = 0(|x|^2)$ when $|x| \to 0$; moreover, by construction of $G_{w,n}$ and G_n, we have $F_n(dx) = \dfrac{1}{k(x)}1_{\{x\neq 0\}}\cdot G_n(dx)$, and similarly $F(dx) = \dfrac{1}{k(x)}1_{\{x\neq 0\}}\cdot G(dx)$. Then let $g \in C_3(\mathbb{R}^d)$: it is obvious that $g'(x) = \dfrac{g(x)}{k(x)}1_{\{x\neq 0\}}$ is continuous and bounded on $\mathbb{R}^d$, so $G_n(g') \to G(g')$, which yields $F_n(g) \to F(g)$. Therefore, we have $[\delta_{1,3}]$.

Consider formula (1): we have $g_u \in C_3(\mathbb{R}^d)$, so $F_n(g_u) \to F(g_u)$ and it follows that

$$iu\cdot b_n - \tfrac{1}{2}u\cdot\tilde{c}_n\cdot u \to iu\cdot b - \tfrac{1}{2}u\cdot\tilde{c}\cdot u \quad \text{for all } u \in \mathbb{R}^d,$$

which clearly implies $[\beta_1]$ and $[\gamma_1]$. □

2.10 **Remark.** Under $[\gamma_1]$, the two equivalent conditions $[\delta_{1,2}]$ and $[\delta_{1,3}]$ are also equivalent to the following one: the measures $|x|^3 \wedge 1\cdot F_n(dx)$ converge weakly to the measure $|x|^3 \wedge 1\cdot F(dx)$.

When $d = 1$, still another condition is equivalent to $[\gamma_1] + [\delta_{1,2}]$, namely that the measures $c_n\varepsilon_0(dx) + |x|^2 \wedge 1\cdot F_n(dx)$ converge weakly to $c\varepsilon_0(dx) + |x|^2 \wedge 1\cdot F(dx)$ (in fact, in this version of the conditions for convergence, it is more usual to replace $x^2 \wedge 1$ by $x^2/(1+x^2)$). But when $d \geq 2$, c_n and c are matrices and this condition makes no sense. □

In some cases it is not necessary to use a truncation function (which amounts to saying that we use the "truncation" function $h(x) = x$). More precisely, let μ be an infinitely divisible distribution with characteristics (b, c, F), and suppose that

2.11 $$\int |x|^2 F(dx) < \infty \quad \left(\text{known to be equivalent to: } \int |x|^2 \mu(dx) < \infty\right).$$

Then we set

2.12 $$\begin{cases} b' = b + \int [x - h(x)] F(dx) \\ \tilde{c}'^{jk} = c^{jk} + \int x^j x^k F(\mathrm{d}x), \end{cases}$$

and of course we have

2.13 $$\psi_{b,c,F}(u) = \psi'_{b',c,F}(u) := iu \cdot b' - \frac{1}{2} u \cdot c \cdot u + \int (e^{iu \cdot x} - 1 - iu \cdot x) F(dx).$$

2.14 **Theorem.** *In the situation of 2.9, suppose that μ_n and μ satisfy 2.11, and define b', b'_n and $\tilde{c}'$, $\tilde{c}'_n$ by 2.12. Suppose that*

2.15 $$\lim_{a \uparrow \infty} \limsup_n \int |x|^2 1_{\{|x| > a\}} F_n(dx) = 0.$$

Then $\mu_n \to \mu$ weakly if and only if we have

$[\beta'_1]$ $b'_n \to b'$

$[\gamma'_1]$ $\tilde{c}'_n \to \tilde{c}'$

and $[\delta_{1,1}]$. In this case, $[\delta_{1,i}]$ holds also for $i = 2, 3, 4$.

Proof. The functions $x^j - h^j(x)$ and $x^j x^k - h^j(x) h^k(x)$ belong to $C_4(\mathbb{R}^d)$, so under $[\delta_{1,4}]$ we have the equivalences $[\beta_1] \Leftrightarrow [\beta'_1]$ and $[\gamma_1] \Leftrightarrow [\gamma'_1]$. We know that $[\delta_{1,1}] \Leftrightarrow [\delta_{1,2}]$, and it is obvious that 2.15 and $[\delta_{1,2}]$ imply $[\delta_{1,4}]$.

Then the claims readily follow from Theorem 2.9. □

§ 2b. Some Lemmas on Characteristic Functions

In this subsection we recall three classical results on characteristic functions. The two first ones are well known; the third one is borrowed from Gnedenko and

Kolmogorov [65] (see also [30] or [197]), except that we replace the median by the truncated mean.

2.16 **Lemma.** *For each $\theta > 0$ there are universal constants $C_1(\theta)$ and $C_2(\theta)$ such that every probability measure μ on $\mathbb{R}^d$, with characteristic function φ, satisfies*

2.17 $$\int |x|^2 \wedge 1 \mu(dx) \le C_1(\theta) \int_{|u| \le \theta} [1 - \operatorname{Re} \varphi(u)]\, du$$

2.18 $$\sup_{|u| \le \theta} |1 - \varphi(u)| \le C_2(\theta) \int |x| \wedge 1 \mu(dx).$$

Proof. a) We have

(1) $$\int_{|u| \le \theta} [1 - \operatorname{Re} \varphi(u)]\, du = \int \mu(dx) \int_{|u| \le \theta} [1 - \cos(u \cdot x)]\, du.$$

For each $x \in \mathbb{R}^d \backslash \{0\}$ we denote by $C(x, \theta)$ an hypercube in $\mathbb{R}^d$ which is contained in the ball $\{u \colon |u| \le \theta\}$, whose an edge is parallel to the vector x, and whose edges have length $b = 2\theta/\sqrt{d}$. Then if $x \ne 0$ we have

$$\begin{aligned}
\int_{|u| \le \theta} [1 - \cos(u \cdot x)]\, du &\ge \int_{C(x,\theta)} [1 - \cos(u \cdot x)]\, du \\
&= b^{d-1} \int_{-b/2}^{b/2} [1 - \cos(s|x|)]\, ds \\
&= b^d \left(1 - \frac{2}{b|x|} \sin \frac{b|x|}{2}\right)
\end{aligned}$$

There is a constant C such that $1 - \dfrac{\sin t}{t} \ge C(t^2 \wedge 1)$ for all $t \in \mathbb{R}$, so

(2) $$\int_{|u| \le \theta} [1 - \cos(u \cdot x)]\, du \ge C(2\theta)^d d^{-d/2} \left[\left(\frac{\theta^2}{d} |x|^2\right) \wedge 1\right].$$

This inequality is trivially true for $x = 0$ also. Finally, we have $(s^2|x|^2) \wedge 1 \ge (s^2 \wedge 1)(|x|^2 \wedge 1)$ for all $s \in \mathbb{R}$, so (1) and (2) yield

$$\int_{|u| \le \theta} [1 - \operatorname{Re} \varphi(u)]\, du \ge \frac{1}{C_1(\theta)} \int |x|^2 \wedge 1 \mu(dx),$$

where $C_1(\theta) = \{C(2\theta/\sqrt{d})^d [(\theta^2/d^2) \wedge 1]\}^{-1}$, which gives 2.17.

b) One easily finds a constant $C_2(\theta)$ such that $|e^{iu \cdot x} - 1| \le C_2(\theta)[|x| \wedge 1]$ for all $|u| \le \theta$, and 2.18 follows. □

2.19 **Lemma.** *For all $\theta > 0$, $A \ge 1$ there is a universal constant $C(\theta, A)$ with the following property: let μ be any probability measure on $\mathbb{R}^d$ such that $\mu(|x| > A) =$*

0, *and let* $\delta = \int x\mu(dx)$ *and* φ *be the characteristic function of* μ; *then*

2.20 $$\int |x-\delta|^2\mu(dx) \le C(\theta, A)\int_{|u|\le\theta} [1-|\varphi(u)|^2]\,du.$$

Proof. Let $\tilde{\mu}$ be the symmetrized measure, defined by

2.21 $$\tilde{\mu}(g) = \int g(x-y)\mu(dx)\mu(dy),$$

and whose characteristic function is $\tilde{\varphi} = |\varphi|^2$. By definition of δ, for all $y \in \mathbb{R}^d$ we have $\int |x-\delta|^2\mu(dx) \le \int |x-y|^2\mu(dx)$, and so

$$\int |x|^2\tilde{\mu}(dx) = \int \mu(dy)\int \mu(dx)|x-y|^2 \ge \int |x-\delta|^2\mu(dx).$$

Moreover $\tilde{\mu}(|x| > 2A) = 0$, thus $\int |x|^2\tilde{\mu}(dx) \le 4A^2\int(|x|^2 \wedge 1)\tilde{\mu}(dx)$. Then 2.20, with $C(\theta, A) = 4A^2C_1(\theta)$ follows from 2.17 applied to $\tilde{\mu}$. □

The next lemma is an improvement of the previous one:

2.22 **Lemma.** *For all* $\theta > 0$, $A \ge 1$, *there is a universal constant* $C'(\theta, A)$ *with the following property: if* h *is a truncation function satisfying 2.1 and if* μ *is a probability measure on* $\mathbb{R}^d$ satisfying

2.23 $$\mu\left(|x| > \frac{1}{4A}\right) \le \frac{1}{4A}$$

and if φ *is the characteristic function of* μ *and* $\delta = \int h(x)\mu(dx)$, *then*

2.24 $$\int |x-\delta|^2 \wedge 1\mu(dx) \le C'(\theta, A)\int_{|u|\le\theta} [1-|\varphi(u)|^2]\,du.$$

Proof. a) The function $f(y) = \int |x-y|^2 \wedge 1\mu(dx)$ is continuous and bounded, and by 2.23 it satisfies

$$\begin{cases} f(0) \le \dfrac{1}{16A^2} + \dfrac{1}{4A} \le \dfrac{5}{16} \\ |y| \ge 3/2 \Rightarrow f(y) \ge \mu\left(|x| \le \dfrac{1}{4A}\right) \ge 1 - \dfrac{1}{4A} \ge \dfrac{3}{4}. \end{cases}$$

Hence f reaches its minimum for some y_0 with $|y_0| \le 3/2$.

For each $x \in \mathbb{R}^d$, we have $|x-\delta|^2 \le |x-y_0|^2 + 2(y_0-\delta)\cdot(x-\delta)$. Since 2.23 implies that $|\delta| \le 1/2$ (recall that $A \ge 1$), we obtain

$$\int |x-\delta|^2 \wedge 1\mu(dx) \le \int_{|x|\le 1/2A} |x-\delta|^2\mu(dx) + \mu(|x| > 1/2A)$$
$$\le \int_{|x|\le 1/2A} |x-y_0|^2\mu(dx)$$

2.25 $$+ 2(y_0-\delta)\cdot\int_{|x|\le 1/2A} (x-\delta)\mu(dx) + \mu\left(|x| > \frac{1}{2A}\right)$$

Now, $h(x) = x$ if $|x| \le 1/2A$, hence

2.26 $$\int_{|x|\le 1/2A} (x-\delta)\mu(dx) + \int_{|x|>1/2A} (h(x)-\delta)\mu(dx) = 0$$

by definition of δ. Since $|h| \le A$, $|\delta| \le 1/2$ and $|y_0| \le 3/2$, it follows from 2.25 and 2.26 that

$$\int |x-\delta|^2 \wedge 1\mu(dx)$$
$$\le f(y_0) + 2|y_0-\delta| \int_{|x|>1/2A} |h(x)-\delta|\mu(dx) + \mu\left(|x| > \frac{1}{2A}\right)$$

2.27 $$\le f(y_0) + [4(A+1)+1]\mu\left(|x| > \frac{1}{2A}\right)$$

b) Introduce the symmetrized measure $\tilde{\mu}$ defined by 2.21. Then 2.23 yields:

$$\tilde{\mu}\left(|x| > \frac{1}{4A}\right) \ge \int \mu(dx)\mu(dy)1_{\{|x|>1/2A,|y|\le 1/4A\}}$$
$$\ge \mu\left(|x| > \frac{1}{2A}\right)\mu\left(|x| \le \frac{1}{4A}\right) \ge \frac{3}{4}\mu\left(|x| > \frac{1}{2A}\right)$$

$\left(\text{because } 1 - \frac{1}{4A} \ge \frac{3}{4}\right)$; hence

2.28 $$\mu\left(|x| > \frac{1}{2A}\right) \le \frac{4}{3}\tilde{\mu}\left(|x| > \frac{1}{4A}\right) \le \frac{64}{3}A^2 \int |x|^2 \wedge 1\tilde{\mu}(dx).$$

Next, by definition of f and of y_0 we have

$$f(y_0) \le \int \mu(dx) \int \mu(dy)|x-y|^2 \wedge 1 = \int |x|^2 \wedge 1\tilde{\mu}(dx),$$

so that 2.27 and 2.28 give

2.29 $$\int |x-\delta|^2 \wedge 1\mu(dx) \le \left\{1 + \frac{64}{3}A^2[4(A+1)+1]\right\} \int |x|^2 \wedge 1\tilde{\mu}(dx).$$

Finally, the characteristic function of $\tilde{\mu}$ is $\tilde{\varphi} = |\varphi|^2$, so 2.24 with $C'(\theta, A) = \{1 + \frac{64}{3}A^2[4(A+1)+1]\}C_1(\theta)$ follows from 2.29 and 2.17. □

§ 2c. Convergence of Rowwise Independent Triangular Arrays

1. We will investigate the limiting behaviour of sums of independent random variables. The setting is such:

2.30 **Definition.** A *rowwise independent d-dimensional triangular array scheme* is a sequence (K^n) of elements of $\bar{\mathbb{N}}^*$ and a sequence $(\Omega^n, \mathscr{F}^n, P^n)$ of probability triples, each one being equipped with an independent sequence $(\chi_k^n)_{1 \le k \le K^n}$ of $\mathbb{R}^d$-valued random variables. □

Of course, by taking the tensor product of all spaces $(\Omega^n, \mathscr{F}^n, P^n)$, it is always possible to assume that all rows of the scheme are defined on the same probability space (Ω, F, P), a property which will be assumed thereafter.

In this subsection, we restrict our attention to rowwise independent schemes which satisfy:

2.31
$$\begin{cases} \sum_{1 \le k \le K^n} |E[h(\chi_k^n)]| < \infty \\ \sum_{1 \le k \le K^n} E(|\chi_k^n|^2 \wedge 1) < \infty \end{cases}$$

for each n, where h is a given truncation function (obviously, 2.31 does not depend on the choice of h in $\mathscr{C}_t^d$). Of course if $K^n < \infty$ we have 2.31. We have proved in Theorem II.3.11 that condition 2.31 is sufficient for

2.32
$$\xi^n = \sum_{1 \le k \le K^n} \chi_k^n$$

to be well-defined (actually, since the χ_k^n's are independent, it follows also from the "three series Theorem" that 2.31 is sufficient for the series 2.32 to converge a. s., independently on the order of summation: see II.3.17).

Moreover, since we are interested only in the sum ξ^n, we can always replace χ_k^n by 0 if $k > K^n$, and then sum up on all $k \in \mathbb{N}^*$: in other words it is not a restriction to suppose that $K^n = \infty$.

Now, the limiting behaviour of ξ^n may be anything: just take $\chi_k^n = 0$ for $k \ge 2$ and χ_1^n be arbitrary random variables. However, if each individual variable χ_k^n becomes small, uniformly in k, when $n \uparrow \infty$ and if the sequence ξ^n converges in law, then the limit is necessarily infinitely divisible. Moreover, one can derive necessary and sufficient conditions for ξ^n to converge toward any given infinitely divisible law. The property of "uniform smallness" of the χ_k^n's is precisely the following:

2.33 **Definition.** The rowwise independent scheme (χ_k^n) is called *infinitesimal* if for all $\varepsilon > 0$ we have

$$\lim_n \sup_{k\geq 1} P(|\chi_k^n| > \varepsilon) = 0. \qquad \square$$

Here are some equivalent properties:

2.34 **Lemma.** *The following properties*: (a), *or* (b), *or* (c_β) (for any $\beta > 0$), *are equivalent to infinitesimality*:

(a) *for all* $\theta > 0$, $\sup_k \sup_{|u|\leq\theta} |\varphi_k^n(u) - 1| \to 0$,

(b) *for all* $\theta > 0$, $\sup_k \sup_{|u|\leq\theta} [1 - \operatorname{Re} \varphi_k^n(u)] \to 0$,

(c_β) $\sup_k E(|\chi_k^n|^\beta \wedge 1) \to 0$;

where φ_k^n *denotes the characteristic function of* χ_k^n.

Proof. The equivalence: $(c_\beta) \Leftrightarrow$ infinitesimality follows from the inequalities: $E(|X|^\beta \wedge 1) \leq \varepsilon^\beta + P(|X| > \varepsilon)$ and $P(|X| > \varepsilon) \leq \varepsilon^{-\beta} E(|X|^\beta \wedge 1)$, for $0 < \varepsilon \leq 1$ and for any variable X. $(c_1) \Rightarrow$ (a) follows from 2.18; (a) $\Rightarrow$ (b) is trivial, and (b) $\Rightarrow (c_2)$ follows from 2.17. $\square$

For the purpose of deriving finite-dimensional convergence for PII's, we need to investigate the limiting behaviour of the sum $\xi^n + \zeta^n$ where ξ^n is given by 2.32, and where ζ^n is another variable, independent from $(\chi_k^n)_{k\geq 1}$, and which is not necessarily small but which is infinitely divisible. This is an innocuous, but useful, generalization of the usual results of Gnedenko and Kolmogorov [65].

2.35 **Theorem.** *We suppose that the rowwise independent scheme* (χ_k^n) *is infinitesimal and satisfies 2.31, and* $\xi^n = \sum_k \chi_k^n$; *let* ζ^n *be an* $\mathbb{R}^d$*-valued infinitely divisible variable with characteristics* (b^n, c^n, F^n) *with respect to a continuous truncation function h, and independent from* $(\chi_k^n)_{k\geq 1}$; *define* $\tilde{c}^n$ *by 2.5* (from c^n and F^n and h).

a) *If* $\mathscr{L}(\xi^n + \zeta^n) \to \mu$ *weakly, then* μ *is infinitely divisible.*

b) *In order that* $\mathscr{L}(\xi^n + \zeta^n) \to \mu$ *weakly, where* μ *is infinitely divisible with characteristics* (b, c, F) (and $\tilde{c}$ defined by 2.5), *it is necessary and sufficient that the following three conditions hold*:

$$[\beta_2] \quad b^n + \sum_k E[h(\chi_k^n)] \to b$$

$$[\gamma_2] \quad \tilde{c}^{n,jl} + \sum_k \{E(h^j h^l(\chi_k^n) - E[h^j(\chi_k^n)]E[h^l(\chi_k^n)]\} \to \tilde{c}^{jl}$$

$$[\delta_{2,i}] \quad F^n(g) + \sum_k E[g(\chi_k^n)] \to F(g) \quad \text{for all } g \in C_i(\mathbb{R}^d).$$

for either $i = 1$ *or* $i = 2$.

Exactly as for Theorem 2.9, we can notice that the set of conditions $[\beta_2]$, $[\gamma_2]$, $[\delta_{2,i}]$ does not depend on the function h, as long as it is continuous (or, even, F-a. s. continuous).

We saw that Theorem 2.9 has a version (Theorem 2.14) for the "square-integrable" case. Here, we also have such a version:

2.36 **Theorem.** *In the situation of 2.35* (in which, in particular, we assume infinitesimality) *we suppose in addition that F^n and F satisfy 2.11 and that*

2.37 $$\lim_{a\uparrow\infty} \limsup_n \left[\int |x|^2 1_{\{|x|>a\}} F^n(dx) + \sum_k E(|\chi_k^n|^2 1_{\{|\chi_k^n|>a\}})\right] = 0.$$

Define b', b'^n and $\tilde{c}'$, $\tilde{c}'^n$ by 2.12. Then in order that $\mathscr{L}(\xi^n + \zeta^n) \to \mu$ weakly, it is necessary and sufficient that

$[\beta_2']$ $$b'^n + \sum_k E(\chi_k^n) \to b'$$

$[\gamma_2']$ $$\tilde{c}'^{n,jl} + \sum_k \{E(\chi_k^{n,j}\chi_k^{n,l}) - E(\chi_k^{n,j})E(\chi_k^{n,l})\} \to \tilde{c}'^{jl}$$

and $[\delta_{2,1}]$ hold. In this case, we also have $[\delta_{2,2}]$ and $[\delta_{2,4}]$.

The reader should now jump (for a while) to § 5a, in order to immediately read Lindeberg's Theorem (Theorem 5.2), which gives a version of the above in which the limit is Gaussian.

2. We turn now to the proofs of these theorems. We begin by deducing 2.36 from 2.35.

Proof of Theorem 2.36. From the definition 2.7 of $C_1(\mathbb{R}^d)$, we have $[\delta_{2,1}] \Leftrightarrow [\delta_{2,2}]$ (because $F'^n(A) = F^n(A) + \sum_k P(\chi_k^n \in A\backslash\{0\})$ defines a positive measure F'^n on $\mathbb{R}^d$ which does not charge 0 and puts a finite mass on the complement of any neighbourhood of 0, by 2.31). It is also obvious that under 2.37, $[\delta_{2,2}] \Leftrightarrow [\delta_{2,4}]$. So in view of 2.35 it suffices to prove that if $\hat{b}^n$ (resp. $\hat{c}^{n,jl}$) denotes the difference between the left-hand side of $[\beta_2']$ (resp. $[\gamma_2']$) and the left-hand side of $[\beta_2]$ (resp. $[\gamma_2]$), then $[\delta_{2,4}]$ implies:

(1) $$\hat{b}^n \to b' - b = \int [x - h(x)] F(dx)$$

(2) $$\hat{c}^{n,jl} \to \hat{c}^{jl} := \int [x^j x^l - h^j h^l(x)] F(dx).$$

The components of $f(x) = x - h(x)$ are in $C_4(\mathbb{R}^d)$ and $\hat{b}^n = F^n(f) + \sum_k E(f(\chi_k^n))$, so (1) immediately follows from $[\delta_{2,4}]$. Similarly, $g^{jl}(x) = x^j x^l - h^j h^l(x)$ belongs to $C_4(\mathbb{R}^d)$ and

$$\hat{c}^{n,jl} = F^n(g^{jl}) + \sum_k E(g^{jl}(\chi_k^n)) + \sum_k [E[h^j(\chi_k^n)]E[h^l(\chi_k^n)] - E(\chi_k^{n,j})E(\chi_k^{n,l})].$$

Hence (2) will follow from $[\delta_{2,4}]$ and from the property $\gamma^{n,jl} \to 0$, where

$$\gamma^{n,jl} := \sum_k |E[h^j(\chi_k^n)]E[h^l(\chi_k^n)] - E(\chi_k^{n,j})E(\chi_k^{n,l})|$$
$$= \sum_k |E[h^j(\chi_k^n) - \chi_k^{n,j}]E[h^l(\chi_k^n)] + E(\chi_k^{n,j})E[h^l(\chi_k^n) - \chi_k^{n,l}]|$$
$$\leq [\sum_k E(|f(\chi_k^n)|)] \sup_k [E(|\chi_k^n|) + E(|h(\chi_k^n)|)]$$

(recall that $f(x) = x - h(x)$). Since $|f| \in C_4(\mathbb{R}^d)$, $[\delta_{2,4}]$ yields $\limsup_n \sum_k E(|f(\chi_k^n)|) \leq F(|f|) < \infty$; moreover the infinitesimality, and 2.37 and 2.1 imply that $\sup_k[E(|\chi_k^n|) + E(|h(\chi_k^n)|)] \to 0$ as $n \uparrow \infty$: so $\gamma^{n,jl} \to 0$ and (2) holds. □

The proof of Theorem 2.35 proceeds through a long string of lemmas. First, since the components of $h - h'$ are in $C_2(\mathbb{R}^d)$ for any two continuous truncation functions h and h', the same proof than for 2.36 readily shows that under $[\delta_{2,2}]$ the conditions $[\beta_2] + [\gamma_2]$ do not depend on the truncation function h as long as it is continuous: so until the end of § 2c we take h to be *uniformly continuous*. Further, A is a constant satifying 2.1. We set

2.38 $$b_k^n = E[h(\chi_k^n)], \quad Y_k^n = \chi_k^n - b_k^n, \quad b_k'^n = E[h(Y_k^n)].$$

Here is a first lemma, which does not use the infinitesimality of (χ_k^n):

2.39 **Lemma.** *If (χ_k^n) satisfies 2.31, then (Y_k^n) also satisfies 2.31.*

Proof. Let A satisfy 2.1. We have $b_k'^n = E[h(\chi_k^n - b_k^n) + b_k^n - h(\chi_k^n)]$, and $|h(\chi_k^n - b_k^n) + b_k^n - h(\chi_k^n)|$ is always smaller than $3A$, and is equal to 0 when $|b_k^n| \leq 1/2A$ and $|\chi_k^n| \leq 1/2A$. Hence

$$\sum_k |b_k'^n| \leq 3A \sum_k [1_{\{|b_k^n| > 1/2A\}} + P(|\chi_k^n| > 1/2A)]$$
$$\leq 3A \sum_k [1_{\{|b_k^n| > 1/2A\}} + 4A^2 E(|\chi_k^n|^2 \wedge 1)]$$

which is finite by 2.31. Similarly, if $|b_k^n| \leq 1/2$ and $|\chi_k^n| \leq 1/2$ we have $|\chi_k^n - b_k^n|^2 \leq 2|\chi_k^n|^2 \wedge 1 + |b_k^n|$, hence

$$\sum_k E(|Y_k^n|^2 \wedge 1) \leq \sum_k [1_{\{|b_k^n| > 1/2\}} + P(|\chi_k^n| > \tfrac{1}{2}) + 2E(|\chi_k^n|^2 \wedge 1) + |b_k^n|]$$
$$\leq \sum_k [3|b_k^n| + 6E(|\chi_k^n|^2 \wedge 1)] < \infty. \qquad \square$$

In the sequel, we assume that (χ_k^n) satisfies 2.31 and is infinitesimal. Then the properties 2.1 of h yield:

2.40 $$M^n \to 0, \quad \text{where } M^n = \sup_k |b_k^n|.$$

This, in turn, implies that the array (Y_k^n) is infinitesimal. Let us introduce two other conditions:

$[\tilde{\gamma}_2]$ $\tilde{c}^{n,jl} + \sum_k E[h^j h^l(Y_k^n)] \to \tilde{c}^{jl}$

$[\tilde{\delta}_2]$ $F^n(g) + \sum_k E[g(Y_k^n)] \to F(g)$ for all $g \in C_2(\mathbb{R}^d)$.

2.41 **Lemma.** *We have the equivalence* $[\delta_{2,2}] \Leftrightarrow [\tilde{\delta}_2]$.

Proof. It suffices to prove that for $g \in C_2(\mathbb{R}^d)$, we have $\sum_k |\delta_k^n(g)| \to 0$, where $\delta_k^n(g) = E[g(\chi_k^n) - g(Y_k^n)]$, under either $[\delta_{2,2}]$ or $[\tilde{\delta}_2]$.

Assume for example $[\delta_{2,2}]$. Let $\theta > 0$ be such that $g(x) = 0$ for $|x| \le \theta$, and $\varepsilon > 0$.

For each $N > 0$ there is $\eta_N > 0$ with $(|x| \le N, |x - y| \le \eta_N) \Rightarrow |g(x) - g(y)| \le \varepsilon$. Set $K = \sup|g|$. If $|b_k^n| \le \frac{\theta}{2}$ and $|\chi_k^n| \le \frac{\theta}{2}$ we have $g(Y_k^n) = g(\chi_k^n) = 0$, and if $|b_k^n| \le \eta_N$ and $|\chi_k^n| \le N$ then $|g(Y_k^n) - g(\chi_k^n)| \le \varepsilon$, so we get

$$M^n \le \frac{\theta}{s} \bigwedge \eta_N \quad \Rightarrow \quad |\delta_k^n(g)| \le \varepsilon P(|\chi_k^n| > \frac{\theta}{2}) + 2K P(|\chi_k^n| > N).$$

With g_a as in 2.7, we then deduce from 2.40 and $[\delta_{2,2}]$ that

$$\begin{aligned}\limsup_n \sum_k |\delta_k^n(g)| &\le \limsup_n \sum_k E[g_{4/\theta}(\chi_k^n)] + \limsup_n \sum_k E[g_{2/N}(\chi_k^n)] \\ &\le \varepsilon F(g_{4/\theta}) + 2K F(g_{2/N})\end{aligned}$$

for all $N > 0$. Since $\varepsilon > 0$ is arbitrary and $\lim_{N\to\infty} F(g_{2/N}) = 0$, we obtain the result. If conversely we suppose $[\delta_2]$, the same argument holds with Y_k^n in place of χ_k^n. □

2.42 **Lemma.** *We have the equivalence* $[\gamma_2] + [\delta_{2,2}] \Leftrightarrow [\tilde{\gamma}_2] + [\tilde{\delta}_2]$.

Proof. Due to the previous lemma, it suffices to prove the equivalence $[\gamma_2] \Leftrightarrow [\tilde{\gamma}_2]$ under $[\delta_{2,2}]$. And for this, proving that $\sum_k |\gamma_k^{n,jl}| \to 0$ is sufficient, where

$$\begin{aligned}\gamma_k^{n,jl} &= E[h^j h^l(\chi_k^n)] - b_k^{n,j} b_k^{n,l} - E[h^j h^l(Y_k^n)] \\ &= E[h^j h^l(\chi_k^n) - h^j h^l(\chi_k^n - b_k^n) + b_k^{n,j} b_k^{n,l} - b_k^{n,j} h^l(\chi_k^n) - b_k^{n,l} h^j(\chi_k^n)].\end{aligned}$$

Let A satisfy 2.1, and $f(x,y) = h^j h^l(x) - h^j h^l(x - y) + y^j y^l - y^j h^l(x) - y^l h^j(x)$. If $|x|, |y| \le 1/2A$ we have $f(x,y) = 0$; and for $\varepsilon > 0$ there is $\eta > 0$ with $|f(x,y)| \le \varepsilon$ for all $|y| \le \eta$, $x \in \mathbb{R}^d$. Hence if $M^n \le \dfrac{1}{2A} \wedge \eta$ we have

$$|\gamma_k^{n,jl}| \le \varepsilon P(|\chi_k^n| > 1/2A).$$

Then the same argument as in the previous lemma gives the result. □

2.43 **Lemma.** *Assume that* $\sup_n \sum_k P(|Y_k^n| > a) < \infty$ *for all* $a > 0$. *Then*

a) $\sum_k |b_k'^n| \to 0$.

b) *for each* $u \in \mathbb{R}^d$ *we have* $\rho^n(u) \to 1$, *where*

2.44 $$\rho^n(u) = \left\{\exp - \sum_k E(e^{iu\cdot Y_k^n} - 1 - iu\cdot h(Y_k^n))\right\} \prod_k E(e^{iu\cdot Y_k^n}).$$

Note that by Lemma 2.39, the sequence (Y_k^n) satisfies 2.31: hence $\sum_k Y_k^n$ converges, and the infinite product in 2.44 is meaningful; moreover $|e^{iu\cdot x} - 1 - iu\cdot h(x)| \le C_u(|x|^2 \wedge 1)$ for some constant C_u, so the series in the "exp" in 2.44 is also convergent.

Proof. Let $K = \sup_n \sum_k P(|Y_k^n| > \frac{1}{2A})$ with A meeting 2.1. Let $\varepsilon > 0$; there exists $\eta > 0$ such that $|x - y| \le \eta \Rightarrow |h(x) - h(y)| \le \varepsilon$. We have $b_k'^n = E[h(Y_k^n) - h(Y_k^n + b_k^n) + b_k^n]$, and if $M^n \le \frac{1}{2A} \wedge \eta$ we have $h(Y_k^n) - h(Y_k^n + b_k^n) + b_k^n = 0$ if $|Y_k^n| \le 1/2A$ and $|h(Y_k^n) - h(Y_k^n + b_k^n)| \le \varepsilon$ everywhere. Hence $|b_k'^n| \le (\varepsilon + M^n)P(|Y_k^n| > 1/2A)$ if $M^n \le (1/2A) \wedge \eta$, and 2.40 yields:

$$\limsup_n \sum_k |b_k'^n| \le \varepsilon \sum_k P\left(|Y_k^n| > \frac{1}{2A}\right) \le \varepsilon K.$$

Since $\varepsilon > 0$ is arbitrary, we obtain (a).

Now if $u \in \mathbb{R}^d$ is fixed and $\delta_k^n = E(e^{iu\cdot Y_k^n} - 1)$ we have

$$\rho^n(u) = \left\{\exp iu\cdot \sum_k b_k'^n\right\} \prod_k [(1 + \delta_k^n)e^{-\delta_k^n}].$$

Hence in order to obtain (b), and by virtue of (a), it remains to prove that $\bar{\delta}^n := \prod_k[(1 + \delta_k^n)\exp -\delta_k^n] \to 1$.

The array (Y_k^n) being infinitesimal, Lemma 2.34 implies that $\sup_k |\delta_k^n| \to 0$. On the other hand, $|e^{iu\cdot x} - 1 - iu\cdot h(x)| \le C_u(|x|^2 \wedge 1)$, so

$$|\delta_k^n| = |E(e^{iu\cdot Y_k^n} - 1 - iu\cdot h(Y_k^n)) + iu\cdot b_k'^n| \le C_u E(|Y_k^n|^2 \wedge 1) + |b_k'^n|\,|u|$$

and thus $\limsup_n \sum_k |\delta_k^n| \le C_u K$ because of (a) again. It follows that

$$\text{(1)} \qquad \limsup_n \sum_k |\delta_k^n|^2 \le \left\{\limsup_n \sup_k |\delta_k^n|\right\}\left\{\limsup_n \sum_k |\delta_k^n|\right\} = 0$$

Finally, if $\operatorname{Log} x$ denotes the principal determination of the logarithm of the complex number x, we have $|-x + \operatorname{Log}(1 + x)| \le C|x|^2$ for $x \in \mathbb{C}$, $|x| \le 1/2$. Then (1) implies that $|\delta_k^n| \le 1/2$ for all k, for n large enough. For these n's we have

$$\bar{\delta}^n = \exp \sum_k [-\delta_k^n + \operatorname{Log}(1 + \delta_k^n)],$$

$$\sum_k |-\delta_k^n + \operatorname{Log}(1 + \delta_k^n)| \le C \sum_k |\delta_k^n|^2,$$

so $\bar{\delta}^n \to 1$ by (1), and this finishes the proof. □

Proof of Theorem 2.35. (i) Define $\rho^n(u)$ by 2.44 and set

$$\tilde{b}^n = b^n + \sum_k b_k^n$$

$$\tilde{F}^n(A) = F^n(A) + \sum_k P(Y_k^n \in A\backslash\{0\}) \quad (A \in \mathscr{R}^d),$$

so that $\tilde{F}^n(|x|^2 \wedge 1) < \infty$ by 2.39. Let $\varphi^n(u) = E(\exp iu\cdot(\xi^n + \zeta^n))$. A simple computation shows that

2.45 $$\varphi^n(u) = \rho^n(u) \exp \psi_{\tilde{b}^n, c^n, \tilde{F}^n}(u).$$

(ii) Assume that $\mathscr{L}(\xi^n + \zeta^n) \to \mu$ weakly, and denote by φ the characteristic function of μ. There is $\theta > 0$ such that $|\varphi(u)| \geq 3/4$ for all $|u| \leq \theta$. Let A satisfy 2.1. Since $\varphi^n \to \varphi$ uniformly on compact sets, there is $n_0 \in \mathbb{N}^*$ such that $|\varphi^n(u)| \geq 1/2$ for $|u| \leq \theta$ and $n \geq n_0$. Since (χ_k^n) is infinitesimal, there is $n_1 \geq n_0$ such that

2.46 $$n \geq n_1 \Rightarrow P\left(|\chi_k^n| > \frac{1}{4A}\right) \leq \frac{1}{4A} \quad \text{for all } k \geq 1.$$

Since $|\varphi^n(u)| \geq 1/2$ for $|u| \leq \theta$ and $n \geq n_1$, we also have

$$n \geq n_1, |u| \leq \theta \Rightarrow \sum_k -\operatorname{Log}|E(e^{iu\cdot\chi_k^n})| \leq \operatorname{Log} 2.$$

Applying Lemma 2.22 to $\mathscr{L}(\chi_k^n)$, using 2.46 and the inequality $1 - |\varphi|^2 \leq -\frac{1}{2}\operatorname{Log}|\varphi|$, and suming up on k, we get

2.47 $$n \geq n_1 \Rightarrow \sum_k E(|Y_k^n|^2 \wedge 1) \leq \frac{1}{2} C'(\theta, A)(\operatorname{Log} 2)\omega_d \theta^d$$

where ω_d is the volume of the unit sphere in $\mathbb{R}^d$.

Then 2.47 and Lemma 2.39 imply that: $\sup_n \sum_k E(|Y_k^n|^2 \wedge 1) < \infty$, so $\rho^n(u) \to 1$ by 2.43. Hence $\varphi = \lim_n \exp \psi_{\tilde{b}^n, c^n, \tilde{F}^n}$: the closure property of the set of all infinitely divisible distributions implies that $\varphi = \exp \psi_{b,c,F}$ for some triple (b, c, F) satisfying 2.2. Moreover $(\tilde{b}^n, c^n, \tilde{F}^n)$ satisfies $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,2}]$ relatively to (b, c, F) by Theorem 2.9. But with this notation, we have $[\beta_1] = [\beta_2]$, and $[\gamma_1] = [\tilde{\gamma}_2]$, and $[\delta_{1,2}] = [\tilde{\delta}_2]$. It remains to apply Lemma 2.42 in order to obtain $[\gamma_2]$ and $[\delta_{2,2}]$ and this finishes the proof of (a) and of the necessary condition of (b).

(iii) It remains to prove the sufficient condition in (b). We assume $[\beta_2]$, $[\gamma_2]$, $[\delta_{2,1}]$. We know that $[\delta_{2,2}]$ holds as well (see the proof of 2.36), and we have $[\tilde{\gamma}_2]$ and $[\tilde{\delta}_2]$ by 2.42. Clearly $[\tilde{\delta}_2]$ yields $\sup_n \sum_k P(|Y_k^n| > a) < \infty$ for all $a > 0$, hence we deduce from 2.43 that $\rho^n(u) \to 1$, while exactly like in part (ii) above, $[\beta_2] + [\tilde{\gamma}_2] + [\tilde{\delta}_2]$ implies that $\psi_{\tilde{b}^n, c^n, \tilde{F}^n} \to \psi_{b,c,F}$. Due to 2.45, this proves that $\varphi^n \to \exp \psi_{b,c,F}$, and the sufficient condition is proved. □

§ 2d. Finite-Dimensional Convergence of PII-Semimartingales to a PII Without Fixed Time of Discontinuity

We will apply the previous results, and in particular Theorem 2.35, to the finite-dimensional convergence $X^n \xrightarrow{\mathscr{L}(D)} X$ along a subset D of $\mathbb{R}_+$, under the following assumption:

2.48 X^n is a d-dimensional PII-semimartingale, X is a d-dimensional PII without fixed time of discontinuity. □

The assumption that the X^n's are semimartingales is for simplicity.

According to the structure theorem II.4.15, the distribution of the process X^n is characterized by a triplet of characteristics (B^n, C^n, ν^n) relative to some fixed truncation function h. We associate with it the second modified characteristic $\tilde{C}^n$, which is càdlàg and increasing in the set of all $d \times d$ symmetric nonnegative matrices for their natural order, by

2.49
$$\begin{aligned}\tilde{C}_t^{n,jk} &= C_t^{n,jk} + (h^j h^k) * \nu_t^n - \sum_{s \le t} \nu^n(\{s\} \times h^j)\nu^n(\{s\} \times h^k)\\ &= C_t^{n,jk} + (h^j h^k) * \nu_t^n - \sum_{s \le t} \Delta B_s^{n,j} \Delta B_s^{n,k}\end{aligned}$$

(note that $(|x|^2 \wedge 1) * \nu_t^n < \infty$ because of the semimartingale assumption).

Similarly, we denote by (B, C, ν) the characteristics of X, for the same truncation function h; 2.48 implies that $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all t, and B is continuous, and $(|x|^2 \wedge 1) * \nu_t < \infty$ (see II.4.15 if X is a semimartingale, and II.5.2 if it is not). So the second modified characteristic $\tilde{C}$ of X is given by

2.50
$$\tilde{C}^{jk} = C^{jk} + (h^j h^k) * \nu$$

(a particular case of 2.49 if X is a semimartingale, or due to II.5.8 in general).

We introduce a family of conditions, in which D denotes a subset of $\mathbb{R}_+$:

$[\beta_3\text{-}D]$ $\quad B_t^n \to B_t$ for all $t \in D$

2.51 $[\gamma_3\text{-}D]$ $\quad \tilde{C}_t^n \to \tilde{C}_t$ for all $t \in D$

$[\delta_{3,i}\text{-}D]$ $\quad g * \nu_t^n \to g * \nu_t$ for all $t \in D$, $g \in C_i(\mathbb{R}^d)$

(see 2.7; recall that $g * \nu_t^n$ denotes $\int_0^t \int_{\mathbb{R}^d} g(x)\nu^n(ds, dx)$).

2.52 **Theorem.** *We suppose that 2.48 holds, that the truncation function h is continuous, and that D is a subset of* $\mathbb{R}_+$.

a) *Under the assumption*

2.53
$$\limsup_n \sum_{s \le t} \nu^n(\{s\} \times \{|x| > \varepsilon\}) = 0 \quad \text{for all } \varepsilon > 0, t \in D,$$

we have $X^n \xrightarrow{\mathcal{L}(D)} X$ *if and only if* $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$, $[\delta_{3,i}\text{-}D]$, *hold, where either* $i = 1$ *or* $i = 2$.

b) *If in addition D is dense in* $\mathbb{R}_+$, *then* $[\delta_{3,1}\text{-}D] \Rightarrow 2.53$ (so the three conditions $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$, $[\delta_{3,1}\text{-}D]$ are sufficient to insure $X^n \xrightarrow{\mathcal{L}(D)} X$).

Proof. a) Let $J^n = \{s > 0 \colon \nu^n(\{s\} \times \mathbb{R}^d) > 0\}$ be the set of fixed times of discontinuity for X^n, and call $\nu^{n,c}$ the measure $\nu^{n,c}(ds, dx) = \nu^n(ds, dx) 1_{(J^n)^c}(s)$, and $B_t^{n,c} = B^n - \sum_{s \le t} \Delta B_s^n$. Let also $g_u(x) = e^{iu \cdot x} - 1 - iu \cdot h(x)$. Then formula II.4.16 writes as

2.54 $$E(e^{iu\cdot(X_t^n-X_s^n)})=\exp\left\{iu\cdot(B_t^{n,c}-B_s^{n,c})-\frac{1}{2}u\cdot(C_t^n-C_s^n)\cdot u+g_u*\nu_t^{n,c}\right.$$
$$\left.-g_u*\nu_s^{n,c}\right\}\times\prod_{s<r\leq t,r\in J^n}[1+\nu^n(\{r\}\times(e^{iu\cdot x}-1))].$$

Moreover, the law $\eta_t^n=\mathscr{L}(\Delta X_t^n)$ is (see II.4.17):

2.55 $$\eta_t^n(dx)=\nu^n(\{t\}\times dx)+[1-\nu^n(\{t\}\times\mathbb{R}^d)]\varepsilon_0(dx).$$

Now, let $s,t\in D\cup\{0\}$ with $s<t$. Let K^n be the number of points in $J^n\cap(s,t]$ and let $(s_k^n)_{1\leq k\leq K^n}$ be an enumeration of these points. Set

$$b^n=B_t^{n,c}-B_s^{n,c},\quad c^n=C_t^n-C_s^n,\quad F^n(dx)=\nu^{n,c}((s,t]\times dx)$$
$$b=B_t-B_s,\qquad c=C_t-C_s,\qquad F(dx)=\nu((s,t]\times dx)$$
$$\chi_k^n=\begin{cases}\Delta X_{s_k^n}^n & \text{if } 1\leq k\leq K^n\\ 0 & \text{if } k>K^n\end{cases}$$

Then (b,c,F) and (b^n,c^n,F^n) satisfy 2.2, and the array (χ_k^n) satisfies 2.31 (use 2.55, the property $|x|^2\wedge 1*\nu_t^n<\infty$, and $E[h(\chi_k^n)]=\Delta B_{s_k^n}^n$ if $k\leq K^n$ and the fact that B^n has finite variation). Moreover, 2.53 and 2.55 yield infinitesimality for the array (χ_k^n). If $\xi^n=\sum_k\chi_k^n$ and $\zeta^n=X_t^n-X_s^n-\xi^n$, then 2.54 and 2.55 show that ζ^n, which clearly is independent of $(\chi_k^n)_{k\geq 1}$, has a law that is infinitely divisible with characteristics (b^n,c^n,F^n) and a simple computation shows that

2.56 $$\begin{cases}b^n+\sum\limits_k E[h(\chi_k^n)]=B_t^n-B_s^n\\ \tilde{c}^{n,jl}+\sum\limits_k\{E[h^jh^l(\chi_k^n)]-E[h^j(\chi_k^n)]E[h^l(\chi_k^n)]\}=\tilde{C}_t^{n,jl}-\tilde{C}_s^{n,jl}\\ F^n(g)+\sum\limits_k E[g(\chi_k^n)]=g*\nu_t^n-g*\nu_s^n.\end{cases}$$

On the other hand, since X has no fixed time of discontinuity, the formula 2.54 written for X has no infinite product, and $\mathscr{L}(X_t-X_s)$ is infinitely divisible with characteristics (b,c,F).

It remains to apply Theorem 2.35: we have $X_t^n-X_s^n\xrightarrow{\mathscr{L}}X_t-X_s$ if and only if

$$B_t^n-B_s^n\to B_t-B_s$$
$$\tilde{C}_t^n-\tilde{C}_s^n\to\tilde{C}_t-\tilde{C}_s$$
$$g*\nu_t^n-g*\nu_s^n\to g*\nu_t-g*\nu_s\quad\text{for all } g\in C_i(\mathbb{R}^d)\ (i=1\text{ or }i=2).$$

Hence the equivalence: $X^n\xrightarrow{\mathscr{L}(D)}X\Leftrightarrow[\beta_3\text{-}D]+[\gamma_3\text{-}D]+[\delta_{3,i}\text{-}D]$ follows from Lemma 1-3.

b) Assume $[\delta_{3,1}\text{-}D]$, with D dense in $\mathbb{R}_+$. Let $\varepsilon>0$ be rational, and set $g=g_{2/\varepsilon}$ (see 2.7). Since $g*\nu^n$ and $g*\nu$ are increasing and $g*\nu$ is continuous, $[\delta_{3,1}\text{-}D]$ and VI.2.15c imply that $g*\nu^n\to g*\nu$ uniformly on finite intervals. Hence

$\sup_{s \leq t} \Delta(g * \nu^n)_s \to 0$ for all t. But $\Delta(g * \nu^n)_s = \nu^n(\{s\} \times g) \geq \nu^n(\{s\} \times \{|x| > \varepsilon\})$, so 2.53 is met. □

2.57 **Remark.** The two theorems 2.35 and 2.52 are actually very close one to the other; we have used 2.35 to prove 2.52. Conversely, 2.35b is a particular case of 2.52, namely the case where $D = \{1\}$. To see that, consider the situation of Theorem 2.35, with the triangular array (χ_k^n) and the variables ζ^n. Let

$$X_t^n = \zeta^n 1_{\{t \geq 1\}} + \sum_k \chi_k^n 1_{\{t \geq 1/k\}}.$$

Then X^n is a PII, as well as $X_t = \zeta 1_{\{t \geq 1\}}$ where ζ is a variable with $\mathscr{L}(\zeta) = \mu$. Then condition 2.53 (resp. $[\beta_3\text{-}\{1\}]$, $[\gamma_3\text{-}\{1\}]$, $[\delta_{3,i}\text{-}\{1\}]$) are exactly infinitesimality (resp. $[\beta_2]$, $[\gamma_2]$, $[\delta_{2,i}]$). □

2.58 **Remark.** It is possible to have $X^n \xrightarrow{\mathscr{L}(D)} X$ with D dense in $\mathbb{R}_+$, without having 2.53 (nor $[\beta_3\text{-}D]$, of course). For example, consider the (deterministic) processes

$$X_t^n = \begin{cases} 1 & \text{if } \dfrac{1}{n} \leq t < \dfrac{2}{n} \\ 0 & \text{otherwise} \end{cases}$$

and $X_t = 0$. Then $X^n \xrightarrow{\mathscr{L}(\mathbb{R}_+)} X$ and 2.53 fails (note however that in this case, $X^n \xrightarrow{\mathscr{L}} X$ fails as well; if it were true, it would contradicts the conditions for functional convergence proved in the next section). □

The following result is of less importance; its interest lies in the fact that there is no condition on the drift, although the drift of $X'^n = X^n - B^n$ is not 0. We suppose that the truncation function h is uniformly continuous, although the result holds also when h is only continuous.

2.59 **Proposition.** *Assume that the truncation function h is uniformly continuous, and that 2.48 and 2.53 hold for some subset D of $\mathbb{R}_+$. Set $X'^n = X^n - B^n$ and $X' = X - B$. Then $X'^n \xrightarrow{\mathscr{L}D} X'$ if and only if $[\gamma_3\text{-}D]$ and $[\delta_{3,i}\text{-}D]$ for $i = 1$ or $i = 2$ hold.*

Proof. a) The characteristics (B', C', ν') of X' and (B'^n, C'^n, ν'^n) of X'^n are given by II.2.33. In particular, because of 2.48, $B' = 0$ and $C' = C$ and $\nu' = \nu$, and thus $\tilde{C}' = \tilde{C}$.

Let $t \in D$, and $J^n = \{s > 0 \colon \nu^n(\{s\} \times \mathbb{R}^d) > 0\}$, and let $(s_k^n)_{1 \leq k \leq K^n}$ (with $K^n \leq \infty$) be an enumeration of the points of $J^n \cap [0, t]$. We also put

$$F^n(dx) = \nu^{n,c}([0,t] \times dx), \quad F(dx) = \nu([0,t] \times dx),$$

$$\chi_k^n = \begin{cases} \Delta X_{S_k^n}^n & \text{if } 1 \leq k \leq K^n \\ 0 & \text{if } k > K^n \end{cases} \qquad Y_k^n = \chi_k^n - E[h(\chi_k^n)].$$

Then 2.56 yields, if $\tilde{c}^{n,jl} = C_t^{n,jl} + F^n(h^j h^l)$:

$$(1)\qquad \begin{cases} \tilde{C}_t^{n,jl} = \tilde{c}^{n,jl} + \sum_k \{E[h^j h^l(\chi_k^n)] - E[h^j(\chi_k^n)]E[h^l(\chi_k^n)]\} \\ g * \nu_t^n = F^n(g) + \sum_k E[g(\chi_k^n)]. \end{cases}$$

Moreover, II.2.33 easily yields $B_t'^n = \sum_k b_k'^n$, where $b_k'^n = E[h(Y_k^n)]$, and

$$(2)\qquad \begin{cases} \tilde{C}_t'^{n,jl} = \tilde{c}^{n,jl} + \sum_k E[h^j h^l(Y_k^n)] - \sum_k b_k'^{n,j} b_k'^{n,l} \\ g * \nu_t'^n = F^n(g) + \sum_k E[g(Y_k^n)]. \end{cases}$$

Furthermore, as in the proof of 2.52, we see that 2.53 implies that the array (χ_k^n) is infinitesimal, hence so is the array (Y_k^n) (see before 2.41), which in turn is equivalent to (recall that $t \in D$):

$$(3)\qquad \sup_{s \le t} \nu'^n(\{s\} \times \{|x| > \varepsilon\}) = \sup_k P(|Y_k^n| > \varepsilon) \to 0 \quad \text{for all } \varepsilon > 0.$$

Let us denote by $[\beta_3 - D]'$, $[\gamma_3 - D]'$, $[\delta_{3,i} - D]'$ the conditions similar to 2.51, but relative to X'^n and X'. Due to (1) and (2) we have (see 2.35 and 2.41):

$$(4)\qquad [\delta_{3,2} - \{t\}]' = [\delta_{2,2}] \quad \Leftrightarrow \quad [\tilde{\delta}_2] = [\delta_{3,2} - t]'.$$

Further, under $[\tilde{\delta}_2]$ we have $\sum_n \sum_k P(|Y_k^n| > a) < \infty$ for all $a > 0$, hence 2.43 implies

$$(5)\qquad [\tilde{\delta}_2] \quad \Rightarrow \quad [\beta_3 - t]',$$

while 2.43, 2.42 and (2) yield

$$(6)\qquad [\tilde{\delta}_2] \quad \Rightarrow \quad ([\gamma_3 - \{t\}] = [\gamma_2] \quad \Leftrightarrow \quad [\tilde{\gamma}_2] \quad \Leftrightarrow \quad [\gamma_3 - \{t\}]').$$

b) If $X_t'^n \xrightarrow{\mathscr{L}} X_t$ (with $t \in D$), (3) and 2.52 imply $[\gamma_3 - \{t\}]'$ and $[\delta_{3,2} - \{t\}]'$, hence $[\gamma_3 - \{t\}]$ and $[\delta_{3,2} - \{t\}]$ hold by (4) and (6).

c) Suppose conversely $[\gamma_3 - \{t\}]$ and $[\delta_{3,1} - \{t\}]$, hence also $[\delta_{3,2} - \{t\}]$, for some $t \in D$. Then (4), (5) and (6) imply $[\beta_3 - \{t\}]'$, $[\gamma_3 - \{t\}]'$ and $[\delta_{3,2} - \{t\}]'$. We thus have $[\beta_3 - D]'$, $[\gamma_3 - D]'$ and $[\delta_{3,2} - D]'$, so (3) and 2.52 imply $X'^n \xrightarrow{\mathscr{L}(D)} X$. □

Finally, we state the "square-integrable" version of Theorem 2.52. We suppose that each X^n is a locally square-integrable semimartingale in the sense of II.2.27, which amounts to saying that

$$2.60\qquad |x|^2 * \nu_t^n < \infty \quad \text{for all } t \in \mathbb{R}_+.$$

Then we may define a version of the first characteristic "without truncation", say B'^n, which is the only predictable (here, even deterministic) process with finite variation, such that $B_0'^n = 0$ and that $X^n - B'^n$ is a local martingale, and by II.2.29 it is related to B^n by

$$2.61\qquad B'^n = B^n + (x - h(x)) * \nu^n.$$

Similarly, instead of $\tilde{C}^n$ it is natural to consider

2.62 $$\tilde{C}_t'^{n,jl} = C_t^{n,jl} + (x^j x^l) * \nu_t^n - \sum_{s\leq t} \Delta B_s'^{n,j} \Delta B_s'^{n,l}.$$

2.63 **Theorem.** *In the situation of 2.52, we suppose that 2.53 holds and that ν^n and ν satisfy 2.60, and that*

2.64 $$\lim_{a\uparrow\infty} \limsup_n |x|^2 1_{\{|x|>a\}} * \nu_t^n = 0 \quad \text{for all } t \in D.$$

Then if B', B'^n are defined by 2.61 and $\tilde{C}'$, $\tilde{C}'^n$ by 2.62, we have $X^n \xrightarrow{\mathscr{L}(D)} X$ if and only if

$[\beta_3'\text{-}D]$ $\quad B_t'^n \to B_t'$ for all $t \in D$

$[\gamma_3'\text{-}D]$ $\quad \tilde{C}_t'^n \to \tilde{C}_t'$ for all $t \in D$

and $[\delta_{3,1}\text{-}D]$ hold. In this case we also have $[\delta_{3,2}\text{-}D]$ and $[\delta_{3,4}\text{-}D]$.

The proof is exactly the same than for Theorem 2.52, except that one uses 2.36 instead of 2.35.

3. Functional Convergence and Characteristics

The setting is as follows: X^n and X are d-dimensional PII (defined on the same probability space $(\Omega, \mathscr{F}, P)$, but possibly relative to distinct filtrations $\mathbf{F}^n$ and $\mathbf{F}$). We denote by (B^n, C^n, ν^n) (resp. (B, C, ν)) their characteristics, relative to the same *continuous* truncation function $h \in \mathscr{C}_t^d$: see II.4.15 if they are semimartingales, and II.5.2 if not. We also consider the second modified characteristic, defined in general by (see II.5.7):

3.1 $$\tilde{C}_t^{n,jk} = C_t^{n,jk} + [h^j(x) - \Delta B^{n,j}][h^k(x) - \Delta B^{n,k}] * \nu_t^n + \sum_{s\leq t} [1 - \nu^n(\{s\} \times \mathbb{R}^d)] \Delta B_s^{n,j} \Delta B_s^{n,k}$$

and similarly for $\tilde{C} = (\tilde{C}^{jk})$. Let us recall that when X^n is a semimartingale, this reduces to 2.49, and to 2.50 when X^n has no fixed time of discontinuity.

We will give necessary and sufficient conditions for $X^n \xrightarrow{\mathscr{L}} X$, in terms of convergence of (B^n, C^n, ν^n) toward (B, C, ν). The results are stated in § 3a. Since the sufficient condition is by far the most important one, and since in most cases the X^n are semimartingales as well as PII and X has no fixed time of discontinuity, we prove the sufficient part under Assumption 2.48 in § 3b, using "method 1" of Section 1.

Then the necessary condition is proved in § 3c, and the sufficient one (general case) in § 3d, using "method 3" of Section 1: for the general case, this third method is much simpler than the method using finite-dimensional convergence, even taking into account the (not so simple) proof of the necessary part.

§ 3a. The Results

We first introduce a series of conditions on the characteristics. Firstly, remember the conditions of 2.51, where D is a subset of $\mathbb{R}_+$:

$[\beta_3\text{-}D]$ $\quad B_t^n \to B_t \quad$ for all $t \in D$

$[\gamma_3\text{-}D]$ $\quad \tilde{C}_t^n \to \tilde{C}_t \quad$ for all $t \in D$

$[\delta_{3,i}\text{-}D]$ $\quad g * \nu_t^n \to g * \nu_t \quad$ for all $t \in D$, $g \in C_i(\mathbb{R}^d)$.

Next, set

3.2 $$\begin{cases} [\text{Sup-}\beta_3] & \sup_{s \le t} |B_s^n - B_s| \to 0 \quad \text{for all } t \ge 0 \\ [\text{Sup-}\gamma_3] & \sup_{s \le t} |\tilde{C}_s^n - \tilde{C}_s| \to 0 \quad \text{for all } t \ge 0 \\ [\text{Sup-}\delta_{3,i}] & \sup_{s \le t} |g * \nu_s^n - g * \nu_s| \to 0 \quad \text{for all } t \ge 0,\ g \in C_i(\mathbb{R}^d). \end{cases}$$

3.3 $$\begin{cases} [\text{Sk-}\beta_3] & B^n \to B \quad \text{for the Skorokhod topology in } \mathbb{D}(\mathbb{R}^d) \\ [\text{Sk-}\delta_{3,i}] & g * \nu^n \to g * \nu \quad \text{for the Skorokhod topology in } \mathbb{D}(\mathbb{R}), \text{ for all } g \in C_i(\mathbb{R}^d), \\ [\text{Sk-}\beta\gamma\delta_3] & (B^n, \tilde{C}^n, g * \nu^n) \to (B, \tilde{C}, g * \nu) \quad \text{for the Skorokhod topology in } \mathbb{D}(\mathbb{R}^{d+d^2+m}), \text{ for all } g\colon \mathbb{R}^d \to \mathbb{R}^m \text{ whose components belong to } C_2(\mathbb{R}^d). \end{cases}$$

1. We start with the results concerning the case when X has no fixed time of discontinuity (then $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all t, and B and $\tilde{C}$ and $g * \nu$ are continuous functions). The main theorem is:

3.4 **Theorem.** *Assume that X has no fixed time of discontinuity* (and recall that the truncation function h is chosen continuous), *and let D be a dense subset of $\mathbb{R}_+$. There is equivalence between*

a) $X^n \xrightarrow{\mathscr{L}} X$;

b) $[\text{sup-}\beta_3]$, $[\gamma_3\text{-}D]$, $[\delta_{3,1}\text{-}D]$ *hold.*

Moreover, in this case we also have $[\text{Sup-}\gamma_3]$ *and* $[\text{Sup-}\delta_{3,i}]$ *for* $i = 1, 2$.

Here are two interesting corollaries.

3.5 **Corollary.** *Assume that X^n and X have no fixed time of discontinuity. Then $X^n \xrightarrow{\mathscr{L}} X$ if and only if $X^n \xrightarrow{\mathscr{L}(\mathbb{R}_+)} X$ and* $[\text{Sup-}\beta_3]$ *holds.*

Proof. The necessity is obvious from 3.4 and VI.3.14. Conversely, assume that $X^n \xrightarrow{\mathscr{L}(\mathbb{R}_+)} X$, so in particular $X_t^n \xrightarrow{\mathscr{L}} X_t$ for all t. Now, the assumption implies that X_t^n has an infinitely divisible distribution, with characteristics (in the sense of 2.2) given by

$$b^n = B_t^n, \quad c^n = C_t^n, \quad F^n = \nu^n((0,t] \times \cdot),$$

and similarly for X_t (see II.5.2). This being true for all $t \in \mathbb{R}_+$, Theorem 2.9 yields $[\beta_3\text{-}\mathbb{R}_+]$, $[\gamma_3\text{-}\mathbb{R}_+]$, $[\delta_{3,i}\text{-}\mathbb{R}_+]$ for $i = 1, 2, 3$. If moreover [Sup-β_3] holds, we conclude by applying Theorem 3.4. □

3.6 **Corollary.** *Assume that X^n and X are PIIS* (see II.4.1) *with characteristics $B_t^n = b^n t$, $C_t^n = c^n t$, $\nu^n(ds, dx) = dt \otimes F^n(dx)$ and $B_t = bt$, $C_t = ct$, $\nu(dt, dx) = dt \otimes F(dx)$. There is equivalence between*:

a) $X^n \xrightarrow{\mathscr{L}} X$.

b) $X_1^n \xrightarrow{\mathscr{L}} X_1$.

c) *Conditions* $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,i}]$ (for $i = 1$, or 2, or 3) *of Theorem 2.9 hold.*

Proof. Due to the form of the characteristics, $[\beta_1] = [\text{Sup-}\beta_3]$, $[\gamma_1] = [\text{Sup-}\gamma_3]$, $[\delta_{1,i}] = [\text{sup-}\delta_{3,i}]$, so (a)$\Leftrightarrow$(c) follows from 3.4, while (b)$\Leftrightarrow$(c) follows from 2.9. □

Finally, we state the "square-integrable" version of 3.4.

3.7 **Theorem.** *Assume that 2.48 holds, and that ν^n and ν satisfy 2.60, and that 2.64 holds. Define B', B'^n by 2.61 and $\tilde{C}'$, $\tilde{C}'^n$ by 2.62. Then $X^n \xrightarrow{\mathscr{L}} X$ if and only if*

$$[\text{Sup-}\beta_3'] \quad \sup_{s \le t} |B_s'^n - B_s'| \to 0 \quad \text{for all } t \ge 0,$$

$$[\gamma_3'\text{-}D] \quad \tilde{C}_t'^n \to \tilde{C}_t' \quad \text{for all } t \in D$$

and $[\delta_{3,1}\text{-}D]$ *hold, where D is some dense subset of* $\mathbb{R}_+$.

Moreover, in this case we also have $[\text{Sup-}\delta_{3,i}]$ *for* $i = 1, 2, 4$, *and*

$$[\text{Sup-}\gamma_3'] \quad \sup_{s \le t} |\tilde{C}_s'^n - \tilde{C}_s'| \to 0 \quad \text{for all } t \ge 0.$$

Before deducing 3.7 as a corollary of 3.4, we prove a lemma which in particular shows the last claim in both 3.4 and 3.7.

3.8 **Lemma.** *If D is a dense subset of* $\mathbb{R}_+$ *and if X has no fixed time of discontinuity, the following equivalences hold*: $[\gamma_3\text{-}D] \Leftrightarrow [\text{Sup-}\gamma_3]$, $[\gamma_3'\text{-}D] \Leftrightarrow [\text{Sup-}\gamma_3']$, $[\delta_{3,i}\text{-}D] \Leftrightarrow [\text{Sup-}\delta_{3,i}]$ *for* $i = 1, 2, 4$.

Proof. a) We know that $[\delta_{3,1}\text{-}D] \Leftrightarrow [\delta_{3,2}\text{-}D]$. Assume $[\delta_{3,i}\text{-}D]$ for $i = 2$ or $i = 4$. In order to prove $[\text{Sup-}\delta_{3,i}]$, it suffices to prove that $g * \nu^n$ converges uniformly to $g * \nu$ on all compact intervals, where g is a non-negative function of $C_i(\mathbb{R}^d)$

(since g^+ and g^- belong to $C_i(\mathbb{R}^d)$ as well in this case). But this is a consequence of VI.1.17b and VI.2.15c, because $g * \nu^n$ and $g * \nu$ are non-decreasing and the latter is continuous.

b) The same argument also shows that under $[\gamma_3\text{-}D]$, $\tilde{C}^{n,ii} \to \tilde{C}^{ii}$ uniformly over compact intervals. Then an application of VI.3.36 (for the "deterministic" processes $\tilde{C}^n$) shows that $\tilde{C}^n \to \tilde{C}$ uniformly on compact intervals as well. Finally, that $[\gamma'_3\text{-}D] \Leftrightarrow [\text{Sup-}\gamma'_3]$ is proved similarly. □

Proof of Theorem 3.7. Let D be a dense subset of $\mathbb{R}_+$. In view of 2.61,

$$\left|\sup_{s\leq t}|B'^n_s - B'_s| - \sup_{s\leq t}|B^n_s - B_s|\right| \leq \sup_{s\leq t}\left||x - h(x)| * \nu^n_s - |x - h(x)| * \nu_s\right|$$

and the function $|x - h(x)|$ belongs to $C_4(\mathbb{R}^d)$, so under $[\delta_{3,4}\text{-}D]$ one has $[\text{Sup-}\beta_3] \Leftrightarrow [\text{Sup-}\beta'_3]$. Similarly, 2.62 yields

3.9 $$\tilde{C}'^{n,jk} = \tilde{C}^{n,jk} + g^{jk} * \nu^n + \sum_{s\leq\cdot}[\Delta B^{n,j}_s \Delta B^{n,k}_s - \Delta B'^{n,j}_s \Delta B'^{n,k}_s],$$

where $g^{jk}(x) = x^j x^k - h^j h^k(x)$. Then, exactly as in the end of the proof of 2.36 (with ΔB^n_s and $\Delta B'^n_s$ instead of $E[h(\chi^n_k)]$ and $E(\chi^n_k)$), we get that $[\gamma_3\text{-}D] \Leftrightarrow [\gamma'_3\text{-}D]$ under $[\delta_{3,4}\text{-}D]$ (which, because of 2.52b, implies 2.53 and thus $\sup_{s\leq t}(|\Delta B^n_s| + |\Delta B'^n_s|) \to 0$). Finally, $[\delta_{3,1}\text{-}D] \Rightarrow [\delta_{3,2}\text{-}D]$ by 2.7, and $[\delta_{3,2}\text{-}D]$ plus 2.64 obviously yield $[\delta_{3,4}\text{-}D]$: so the claims follow from 3.4 and 3.8. □

3.10 **Remark.** Theorem 3.7 still holds if we relax 2.48, provided X has no fixed time of discontinuity. But then 2.62 may be meaningless, and the correct definition of $\tilde{C}'^n$ is formula 3.9 (the proof is obviously the same). □

3.11 **Corollary** (*Donsker's Theorem*). *Let $(\xi_k)_{k\geq 1}$ be an i.i.d. sequence of real-valued random variables, with $E(\xi_k) = 0$ and $E((\xi_k)^2) = 1$. Then the processes $X^n_t = \frac{1}{\sqrt{n}} \sum_{1\leq k\leq [nt]} \xi_k$ converge in law to a standard Wiener process.*

Proof. Notice that each X^n is a PII, and the corresponding functions B'^n, $\tilde{C}'^n$, $g * \nu^n$ are (cf. II.3.14, 2.61 and 2.62):

$$B'^n_t = \sum_{1\leq k\leq [nt]} E(\xi_k/\sqrt{n}) = 0$$

$$\tilde{C}'^n_t = \sum_{1\leq k\leq [nt]} E[(\xi_k/\sqrt{n})^2] = \frac{[nt]}{n}$$

$$g * \nu^n_t = \sum_{1\leq k\leq [nt]} E[g(\xi_k/\sqrt{n})] = [nt]E[g(\xi_1/\sqrt{n})],$$

while the characteristics of the standard Wiener process are $B = B' = 0$, $C_t = \tilde{C}'_t = t$, $\nu = 0$. So $[\text{Sup-}\beta'_3]$ and $[\gamma'_3\text{-}\mathbb{R}_+]$ are trivially met. Furthermore, if $a > 0$ we have

$$|x|^2 1_{\{|x|>a\}} * v_t^n = [nt]\frac{1}{n}E(|\xi_1|^2 1_{\{|\xi_1|>a\sqrt{n}\}}),$$

which clearly goes to 0 as $n \uparrow \infty$: hence 2.64, and also $[\delta_{3,1}\text{-}\mathbb{R}_+]$ (because every $g \in C_1(\mathbb{R}^d)$ has $|g(x)| \leq C|x|^2 1_{\{|x|>a\}}$ for some C, a) hold. □

Before proceeding to the general case, we state a last result, which complements Proposition 2.59.

3.12 **Proposition.** *Assume that the truncation h is uniformly continuous, that 2.48 holds, let D be a dense subset of $\mathbb{R}_+$, and set $X'^n = X^n - B^n$ and $X' = X - B$.*

a) *If $X'^n \xrightarrow{\mathcal{L}} X'$ and if 2.53 holds, we have $[\gamma_3\text{-}\mathbb{R}_+]$ and $[\delta_{3,2}\text{-}\mathbb{R}_+]$.*

b) *If $[\gamma_3\text{-}D]$ and $[\delta_{3,1}\text{-}D]$ hold, then $X'^n \xrightarrow{\mathcal{L}} X'$.*

In this result again, we could relax 2.48, assuming only that X has no fixed time of discontinuity. Also, the uniform continuity of h may be replaced with continuity only.

Proof. a) Since $X'^n \xrightarrow{\mathcal{L}} X'$ implies $X'^n \xrightarrow{\mathcal{L}(\mathbb{R}_+)} X'$ because X' has no fixed time of discontinuity, the claim is a (weaker) version of the necessary part of 2.59.

b) Conversely, assume that $[\gamma_3\text{-}D] + [\delta_{3,1}\text{-}D]$ holds, so 2.53 also holds (recall 2.52b). Then we have seen in the proof of 2.59 that $[\gamma_3\text{-}D]'$, $[\delta_{3,2}\text{-}D]'$ and $[\text{Sup-}\beta_3]'$ hold (use 2.43 as in the proof of 2.59 for the latter), where the "$'$" means that we consider the conditions relative to X'^n and X'. Then the claim follows from 3.4. □

2. Now we state our general theorem.

3.13 **Theorem.** *There is equivalence between*

a) $X^n \xrightarrow{\mathcal{L}} X$.

b) $[\text{Sk-}\beta_3]$, $[\gamma_3\text{-}D]$, $[\text{Sk-}\delta_{3,1}]$ *hold, where D is some dense subset of $\mathbb{R}_+$. Moreover, in this case we have $[\gamma_3\text{-}D]$ with $D = \mathbb{R}_+ \backslash J(X) = \{t\colon P(\Delta X_t \neq 0) = 0\}$, and also $[\text{Sk-}\beta\gamma\delta_3]$.*

Some comments are in order:

1) The set of conditions in (b) above does not depend on the truncation function h, as long as it is continuous (this follows from the theorem itself); one could even choose h to be continuous $\nu(\mathbb{R}_+ \times \cdot)$-a. s. only.

2) Theorem 3.4 is a particular case of the above: this follows from Lemma 3.8 and from the equivalence $[\text{Sk-}\beta_3] \Leftrightarrow [\text{Sup-}\beta_3]$ when B is continuous.

3) By virtue of VI.2.15, $[\text{Sk-}\delta_{3,1}] \Leftrightarrow [\delta_{3,1} - D] + [\hat{\delta}_{3,1} - D]$, where

$$[\hat{\delta}_{3,1}\text{-}D] \quad \sum_{s \leq t} \nu^n(\{s\} \times g)^2 \to \sum_{s \leq t} \nu(\{s\} \times g)^2 \quad \text{for all } t \in D,\ g \in C_1(\mathbb{R}^d).$$

3.14 **Corollary.** *If $X^n \xrightarrow{\mathcal{L}} X$ and if the X^n's have no fixed times of discontinuity, then X also has no fixed time of discontinuity.*

Proof. The assumption implies that $g * \nu^n$ is continuous for all $g \in C_2(\mathbb{R}^d)$, and also (by 3.13) that $g * \nu^n \to g * \nu$ for Skorokhod topology; then $g * \nu$ is continuous for all $g \in C_2(\mathbb{R}^d)$, which yields $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all $t \geq 0$, and so the claim follows. □

§ 3b. Sufficient Condition for Convergence Under 2.48

Here we suppose that each X^n is a semimartingale, and that X has no fixed time of discontinuity. According to VI.3.20, in order to obtain $X^n \xrightarrow{\mathcal{L}} X$ (i.e., (b) $\Rightarrow$ (a) in 3.4), it suffices to prove:

3.15 $\qquad [\text{Sup-}\beta_3] + [\gamma_3\text{-}D] + [\delta_{3,1}\text{-}D] \Rightarrow$ the sequence (X^n) is tight,

3.16 $\qquad [\text{Sup-}\beta_3] + [\gamma_3\text{-}D] + [\delta_{3,1}\text{-}D] \Rightarrow X^n \xrightarrow{\mathcal{L}(D)} X.$

Now, since D is dense, 3.16 follows from Theorem 2.52. As for 3.15, we state it as a theorem:

3.17 **Theorem.** *We suppose that 2.48 holds and that the truncation function h is continuous. If* $[\text{Sup-}\beta_3]$, $[\gamma_3\text{-}D]$, $[\delta_{3,1}\text{-}D]$ *are satisfied for some dense subset* $D \subset \mathbb{R}_+$, *then the sequence* (X^n) *is tight.*

Proof. We check that all conditions of Theorem VI.4.18 are met. That (i) is met is trivial, because $X_0^n = 0$. By 3.8 we obtain $[\text{Sup-}\gamma_3]$ and $[\text{Sup-}\delta_{3,1}]$, so (iii) follows from the hypotheses (recall that $g_a(x) = (a|x| - 1)^+ \wedge 1$ belongs to $C_1(\mathbb{R}^d)$ for all $a \in \mathbb{Q}_+$, $a > 0$). Finally, for all $N > 0$, $\varepsilon > 0$, there is $a \in \mathbb{Q}_+$, $a > 0$ such that $g_a * \nu_N \leq \nu((0,N] \times \{|x| > 1/a\}) \leq \varepsilon/2$. Since $g_a * \nu_N^n \to g_a * \nu_N$, for all n large enough we have $\nu^n((0,N] \times \{|x| > 2/a\}) \leq g_a * \nu_N^n \leq \varepsilon$, and VI.4.19 follows. □

Therefore (b) $\Rightarrow$ (a) in 3.4 is proved, under assumption 2.48. Note that this proof hinges upon Theorem 2.52, hence on 2.35, which itself is rather difficult to prove (even for the sufficient part), as we have seen.

§ 3c. Necessary Condition for Convergence

1. The key point is the following (recall that X^n and X are PII's):

3.18 **Proposition.** *Assume that* $X^n \xrightarrow{\mathcal{L}} X$ *and that for each* $t \geq 0$ *the sequence of random variables* $\{\sup_{s \leq t} |X_s^n|\}_{n \geq 1}$ *is uniformly integrable. Then if* $\alpha_n(t) = E(X_t^n)$ *and* $\alpha(t) = E(X_t)$, *we have* $\alpha_n \to \alpha$ *for the Skorokhod topology in* $\mathbb{D}(\mathbb{R}^d)$.

3.19 **Remark.** Let us emphazise the fact that the PII property for *each* X^n is crucial in this proposition. Here is a counter-example: let Y be uniformly distributed on $[0,1]$ and $X_t = 1_{\{t \geq 1\}}$ and $X_t^n = 1_{\{t \geq 1+Y/n\}}$. Of course $X^n \xrightarrow{\mathscr{L}} X$ (in fact $X^n_\cdot(\omega) \to X_\cdot(\omega)$ in Skorokhod sense for all ω) and $|X_t^n| \leq 1$. However, α_n does not converge to α in Skorokhod sense, because each α_n is continuous, whereas α has a jump. □

Before proving that proposition, we introduce the sets $J^n = \{t > 0: \nu^n(\{t\} \times \mathbb{R}^d) > 0\}$ and $J = \{t > 0: \nu(\{t\} \times \mathbb{R}^d) > 0\}$ of fixed times of discontinuity for X^n and X, and we state a lemma:

3.20 **Lemma.** *If $X^n \xrightarrow{\mathscr{L}} X$, then* $[\text{Sk-}\delta_{3,2}]$ *holds.*

Proof. It suffices to prove that if $g \in C_2(\mathbb{R}^d)$ meets $0 \leq g \leqslant 1/2$, then $\alpha_n = g * \nu^n$ converges to $\alpha = g * \nu$ in $\mathbb{D}(\mathbb{R})$.

a) We first prove that $\alpha_n(t) \to \alpha(t)$ for all $t \notin J$. Set $X_t'^n = \sum_{s \leq t} g(\Delta X_s^n)$ and $X_t' = \sum_{s \leq t} g(\Delta X_s)$. With the notation VI.3.10, let $u \in (0, \infty) \backslash U(X)$ be such that $g(x) = 0$ for $|x| \leq u$, and call $\{S_i^n = T_i(X^n, u)\}_{i \geq 1}$ and $\{S_i = T_i(X, u)\}_{i \geq 1}$ the successive times where $|\Delta X^n| > u$ and $|\Delta X| > u$. Hence we also have $X_t'^n = \sum_i g(\Delta X_{S_i^n}^n) 1_{\{S_i^n \leq t\}}$ and $X_t' = \sum_i g(\Delta X_{S_i}) 1_{\{S_i \leq t\}}$. Now, Proposition VI.3.15 yields

$$\text{3.21} \qquad \begin{cases} S_i^n \xrightarrow{\mathscr{L}} S_i \\ X'^n_{S_i^n \wedge t} = \sum\limits_{1 \leq j \leq i} g(\Delta X^n_{S_j^n}) 1_{\{S_j^n \leq t\}} \xrightarrow{\mathscr{L}} X'_{S_i \wedge t} \quad \text{if } t \notin J. \end{cases}$$

Moreover, $0 \leq X'^n_{S_i^n \wedge t} \leq i/2$, and similarly for $X'_{S_i \wedge t}$. Hence 3.21 implies

$$E(X'^n_{S_i^n \wedge t}) \to E(X'_{S_i \wedge t}) \quad \text{for } t \notin J.$$

At this stage, we can reproduce the proof of Theorem 1.5. We have $E(\alpha_n(T)) = E(X_T'^n)$ and $E(\alpha(T)) = E(X_T')$ for all stopping times T; then 1.6 holds with $A_t^n = \alpha_n(t)$ and $A_t' = \alpha(t)$, because of what precedes, and we deduce that $\alpha_n(t) \to \alpha(t)$ for all $t \notin J$.

b) Secondly, we will apply Theorem VI.2.15b to the sequence (α_n), with for f a strictly convex function having $f(x) = -x - \operatorname{Log}(1-x)$ for $0 \leq x \leq 1/2$. Due to (a), it remains to prove that

$$a_n(t) := \sum_{s \leq t} f(\Delta \alpha_n(s)) = -\sum_{s \leq t} [\nu^n(\{s\} \times g) + \operatorname{Log}(1 - \nu^n(\{s\} \times g))]$$
$$\to a(t) := \sum_{s \leq t} f(\Delta \alpha(s)) = -\sum_{s \leq t} [\nu(\{s\} \times g) + \operatorname{Log}(1 - \nu(\{s\} \times g))]$$

for all $t \notin J$. For this, we observe that $g'' = -\operatorname{Log}(1-g)$ belongs to $C_2(\mathbb{R}^d)$ and is nonnegative, so VI.3.16 yields that $X''^n = \sum_{s \leq \cdot} g''(\Delta X_s^n)$ converges in law to $X'' = \sum_{s \leq \cdot} g''(\Delta X_s)$, and so $X_t''^n \xrightarrow{\mathscr{L}} X_t''$ if $t \notin J$. Moreover, II.5.29 and the property $1 - e^{-g''} = g$ yields

$$E(\exp - X_t''^n) = \exp\left\{-g 1_{(J^n)^c} * \nu_t^n + \sum_{s \leq t} \operatorname{Log}[1 - \nu^n(\{s\} \times g)]\right\}$$
$$= \exp[-\alpha_n(t) - a_n(t)],$$

and similarly for $E(\exp - X_t'')$. Since $\alpha_n(t) \to \alpha(t)$ for $t \notin J$ by (a), and $E(\exp - X_t'''^n) \to E(\exp - X_t'')$ for $t \notin J$ because then $X_t'''^n \xrightarrow{\mathcal{L}} X_t''$, we deduce that $a_n(t) \to a(t)$ for $t \notin J$, and we are finished. □

Proof of Proposition 3.18. Since $X^n \xrightarrow{\mathcal{L}} X$, we have $X_t^n \xrightarrow{\mathcal{L}} X_t$ for all $t \notin J$. Due to the uniform integrability hypothesis, we deduce that $\alpha_n(t) \to \alpha(t)$ for all $t \notin J$. We will prove that the sequence (α_n) is relatively compact for Skorokhod topology in $\mathbb{D}(\mathbb{R}^d)$: since α is the only possible limit for that sequence, we deduce that $\alpha_n \to \alpha$.

We use the modulus w_N' defined in VI.1.8. Since $\sup_{s \leq t, n} |\alpha_n(s)| \leq \sup_n E(\sup_{s \leq t} |X_s^n|) < \infty$, it suffices to prove that

3.22 $$\lim_{\delta \downarrow 0} \limsup_n w_N'(\alpha_n, \delta) = 0 \quad \text{for all } N > 0,$$

by Theorem VI.1.14 (note that $\lim_{\delta \downarrow 0} w_N'(\alpha_n, \delta) = 0$ for each n, so 3.22 implies: $\lim_{\delta \downarrow 0} \sup_n w_N'(\alpha_n, \delta) = 0$).

Let $N > 0$, $\varepsilon > 0$ be fixed, and $Z^n = \sup_{s \leq t} |X_s^n|$. The uniform integrability hypothesis yields the existence of $\theta > 0$ such that

3.23 $$E(Z^n 1_{\{Z^n > \theta\}}) \leq \varepsilon \quad \text{for all } n.$$

The sequence (X^n) is tight, so by VI.3.21 there exist $\delta_0 > 0$, $n_0 \in \mathbb{N}^*$ such that

3.24 $$n \geq n_0 \Rightarrow P(F^n) \leq \frac{\varepsilon}{\theta}, \quad \text{where } F^n = \{w_N'(X^n, \delta_0) > \varepsilon\}.$$

Let $g \in C_2(\mathbb{R}^d)$ with $0 \leq g \leq 1$ and $g(x) = 1$ for $|x| \geq \varepsilon$, and set $\beta_n = g * \nu^n$ and $\beta = g * \nu$. By 3.20, $\beta_n \to \beta$ for the Skorokhod topology. Then there is a $\delta \in (0, \delta_0/2]$ such that $w_N'(\beta_n, \delta) \leq \varepsilon/\theta$ for all n. Hence, because of VI.1.12, for each n there is a subdivision $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = N$ such that $t_{j+1}^n - t_j^n \geq \delta$ for $j \leq p_n - 2$, and $t_{j+1}^n - t_j^n \leq 2\delta \leq \delta_0$ for $j \leq p_n - 1$, and $\beta_n(t_{j+1}^n -) - \beta_n(t_j^n) \leq 2\varepsilon/\theta$ for $j \leq p_n - 1$.

Now, let $G_j^n = \{\sup_{t_j^n < r < t_{j+1}^n} |\Delta X_r^n| > \varepsilon\}$. Using II.5.10, we can write

3.25 $$P(G_j^n) \leq E\left(\sum_{t_j^n < r < t_{j+1}^n} g(\Delta X_r^n)\right) = \beta_n(t_{j+1}^n -) - \beta_n(t_j^n) \leq \frac{2\varepsilon}{\theta}.$$

Next, let $t_j^n \leq s < t < t_{j+1}^n$; if $\omega \notin G_j^n \cap F^n$, and since $t_{j+1}^n - t_j^n \leq \delta_0$, we have $|X_t^n - X_s^n| \leq 3\varepsilon$. Hence

$$|X_t^n - X_s^n| \leq 3\varepsilon + 2Z^n 1_{G_j^n \cup F^n} \leq 3\varepsilon + 2\theta 1_{G_j^n \cup F^n} + 2Z^n 1_{\{Z^n > \theta\}}$$

and 3.23, 3.24, 3.25 yield for $n \geq n_0$ and $t_j^n \leq s < t < t_{j+1}^n$:

$$|\alpha_n(t) - \alpha_n(s)| \leq 3\varepsilon + 2\theta\left(\frac{\varepsilon}{\theta} + \frac{2\varepsilon}{\theta}\right) + 2\varepsilon = 11\varepsilon.$$

Hence $w(\alpha_n, [t_j^n, t_{j+1}^n]) \leq 11\varepsilon$ and it follows that $w_N'(\alpha_n, \delta) \leq 11\varepsilon$, for all $n \geq n_0$: that is, we have 3.22 and the proof is finished. □

2. We can now proceed to the proof of the implication (a) $\Rightarrow$ (b) in Theorem 3.13. For this, we begin with several lemmas, for which we use the notation of Section II.2, and in particular

3.26
$$\begin{cases} \check{X}^n(h)_t = \sum_{s\le t} [\Delta X^n_s - h(\Delta X^n_s)] \\ X^n(h) = X^n - \check{X}^n(h) \\ M^n(h) = X^n(h) - B^n. \end{cases}$$

Then $M^n(h)$ is a local martingale (and even a martingale, here), whose bracket is

3.27
$$\langle M^{n,j}(h), M^{n,k}(h)\rangle = \tilde{C}^{n,jk}$$

(see II.5.10) where $\tilde{C}^n$ is given by 3.1.

3.28 **Lemma.** *If $X^n \xrightarrow{\mathscr{L}} X$ we have*
(i) $X^n(h) \xrightarrow{\mathscr{L}} X(h)$,
(ii) $\sup_n \tilde{C}^{n,jj}_t < \infty$ *for all* $t \ge 0, j \le d$.

Proof. Since h is continuous, (i) follows from VI.3.16.

Let $\varphi(u)_t = E(\exp iu\cdot X_t)$ and $\varphi^n(u)_t = E(\exp iu\cdot X^n_t)$, and $\nu^{nc}(ds,dx) = \nu^n(ds,dx)1_{(J^n)^c}(s)$ and $\nu^c(ds,dx) = \nu(ds,dx)1_{J^c}(s)$. If we write formula II.4.16 for $s = 0$ and take the modulus, we obtain

3.29
$$|\varphi^n(u)_t| = \exp\left[-\frac{1}{2}u\cdot C^n_t\cdot u - (1-\cos u\cdot x) * \nu^{nc}_t\right] \times \prod_{r\le t} |1 + \nu^n(\{r\}\times(e^{iu\cdot x}-1))|.$$

There exists $\theta > 0$ such that $|\varphi(u)_t| \ge 3/4$ for $|u| \le \theta$. Since $X^n_t \xrightarrow{\mathscr{L}} X_t$, we have $\varphi^n(u)_t \to \varphi(u)_t$ uniformly on $\{u\colon |u|\le\theta\}$, so there is $n_0 \in \mathbb{N}^*$ such that $|\varphi^n(u)_t| \ge 1/2$ for $|u| \le \theta$, $n \ge n_0$. In particular, each factor in 3.29 is bigger or equal to 1/2.

We first apply this fact to the exponential in 3.29:

3.30
$$n \ge n_0,\quad |u| \le \theta \Rightarrow \tfrac{1}{2}u\cdot C^n_t\cdot u + (1-\cos u\cdot x) * \nu^{nc}_t \le \operatorname{Log} 2.$$

Let A satisfy 2.1. Let $u \in \mathbb{R}^d$ with components $u^j = (\pi A)\wedge\theta$ and $u^k = 0$ for $k \ne j$. We have $y^2 \le \dfrac{\pi^2}{2}(1-\cos y)$ if $|y| \le \pi$, so if $\lambda = (\pi A)\wedge\theta$ we get

$$1 - \cos u\cdot x \ge (1-\cos u\cdot x)1_{\{|x|\le 1/A\}} \ge \frac{2}{\pi^2}|u\cdot x|^2 1_{\{|x|\le 1/A\}}$$
$$= \frac{2\lambda^2}{\pi^2}|x^j|^2 1_{\{|x|\le 1/A\}} \ge \frac{2\lambda^2}{\pi^2}|h^j(x)|^2 1_{\{|x|\le 1/A\}}$$

By 3.20 we have [Sk-$\delta_{3,2}$], hence clearly $K := \sup_n 1_{\{|x|>1/A\}} * \nu^n_t < \infty$. Then 3.30 and $|h| \le A$ yield

3.31
$$n \ge n_0 \quad\Rightarrow\quad C^{n,jj}_t + |h^j|^2 * \nu^{nc}_t \le \frac{\pi^2}{2\lambda^2}\operatorname{Log} 2 + A^2 K.$$

Next, we examine the infinite product in 3.29. Call η_s^n the probability measure on $\mathbb{R}^d$ given by $\eta_s^n(dx) = \nu^n(\{s\} \times dx) + [1 - \nu^n(\{s\} \times \mathbb{R}^d)]\varepsilon_0(dx)$ (i.e., $\eta_s^n = \mathscr{L}(\Delta X_s^n)$). The infinite product in 3.29 is bigger than 1/2 for $|u| \leq \theta$, $n \geq n_0$, hence

3.32 $$n \geq n_0, \quad |u| \leq \theta \Rightarrow \sum_{r \leq t} \left[1 - \left|\int \eta_s^n(dx) e^{iu\cdot x}\right|^2\right] \leq 2 \operatorname{Log} 2$$

because $1 - y^2 \leq -2 \operatorname{Log} y$ for all $y > 0$. Let also μ_s^n be the image of η_s^n by the mapping h, i.e. $\mu_s^n(g) = \eta_s^n(g \circ h)$ for all g (we have $\mu_s^n = \mathscr{L}(h(\Delta X_s^n))$). Then

$$\left|\int \mu_s^n(dx) e^{iu\cdot x} - \int \eta_s^n(dx) e^{iu\cdot x}\right| = \left|\int \eta_s^n(dx)(e^{iu\cdot h(x)} - e^{iu\cdot x})\right|$$
$$\leq 2\eta_s^n\left(|x| > \frac{1}{A}\right).$$

Then, since $\sum_{r \leq t} \eta_r^n(|x| > \frac{1}{A}) \leq 1_{\{|x| > 1/A\}} * \nu_t^n \leq K$, we get

$$\sum_{r \leq t} \left|\left|\int \mu_r^n(dx) e^{iu\cdot x}\right|^2 - \left|\int \eta_r^n(dx) e^{iu\cdot x}\right|^2\right|$$
$$\leq 2 \sum_{r \leq t} \left|\int \mu_r^n(dx) e^{iu\cdot x} - \int \eta_r^n(dx) e^{iu\cdot x}\right| \leq 4K.$$

Hence 3.32 yields

$$n \geq n_0, \quad |u| \leq \theta \Rightarrow \sum_{r \leq t} \left[1 - \left|\int \mu_r^n(dx) e^{iu\cdot x}\right|^2\right] \leq 4K + 2 \operatorname{Log} 2.$$

We apply Lemma 2.19 to each μ_s^n: we have $\mu_s^n(|x| > A) = 0$ because $|h| \leq A$, and $\int \mu_s^n(dx) x = \int \eta_s^n(dx) h(x) = \Delta B_s^n$, so

3.33 $$n \geq n_0 \Rightarrow \sum_{r \leq t} \int \mu_r^n(dx) |x - \Delta B_r^n|^2 \leq \omega_d \theta^d C(\theta, A)(4K + 2 \operatorname{Log} 2)$$

where ω_d is the volumn of the unit sphere in $\mathbb{R}^d$.

Next, using the definitions of η_s^n and μ_s^n and formula 3.1, we obtain

$$\tilde{C}_t^{n,jj} = C_t^{n,jj} + (h^j)^2 * \nu_t^{nc} + \sum_{s \leq t} \int_{x \neq 0} \eta_s^n(dx)[h^j(x) - \Delta B_s^j]^2 + \sum_{s \leq t} \eta_s^n(\{0\})(\Delta B_s^j)^2$$
$$= C_t^{n,jj} + (h^j)^2 * \nu_t^{nc} + \sum_{s \leq t} \int \eta_s^n(dx)[h^j(x) - \Delta B_s^j]^2$$
$$= C_t^{n,jj} + (h^j)^2 * \nu_t^{nc} + \sum_{s \leq t} \int \mu_s^n(dx)[x^j - \Delta B_s^j]^2$$

and 3.31 and 3.33 give

$$n \geq n_0 \Rightarrow \tilde{C}_t^{n,jj} \leq \frac{\pi^2}{2\lambda^2} \operatorname{Log} 2 + \omega_d \theta^d C(\theta, A)(4K + 2 \operatorname{Log} 2) + A^2 K,$$

hence we obtain the result. □

The next lemma is about locally square integrable martingales. It admits a (relatively) simple proof based upon Davis-Burkhölder-Gundy inequalities (see [36] or [183] for these), but here we give a proof based upon the more elementary Doob's inequality.

3.34 **Lemma.** *There exist two constants K_1 and K_2 such that every real-valued locally square-integrable martingale M with $M_0 = 0$ satisfies*

3.35 $$E\left(\sup_{s \leq t} M_s^4\right) \leq K_1 a^2 E(\langle M, M\rangle_t^2)^{1/2} + K_2 E(\langle M, M\rangle_t^2),$$

where $a = \sup_{t,\omega} |\Delta M_t(\omega)|$

Proof. Only the case $a < \infty$ needs to be examined. Let $T_n = \inf(t\colon |M_t| \geq n$ or $\langle M, M\rangle_t \geq n)$. The two stopped processes M^{T_n} and $\langle M, M\rangle^{T_n}$ are bounded, by $n + a$ and $n + a^2$ respectively. Suppose that 3.35 holds for all M^{T_n}: in this formula, the integrands on both sides are increasing with n, so if we let $n \uparrow \infty$ we obtain 3.35 for M. Therefore we can assume that M and $\langle M, M\rangle$ are bounded.

Fix $t > 0$ and let $y = E(\sup_{s \leq t} M_s^4)$ and $z = E(\langle M, M\rangle_t^2)$. Then $N = [M, M] - \langle M, M\rangle$ is a martingale satisfying $|\Delta N| \leq 2a^2$ and which has finite variation. Hence $[N, N] = \sum_{s \leq \cdot} (\Delta N_s)^2$ is strongly majorized by $2a^2 \operatorname{Var}(N)$, which itself is strongly majorized by $2a^2([M, M] + \langle M, M\rangle)$. The strong majoration passes to compensators, so we have $\langle N, N\rangle \leq 4a^2 \langle M, M\rangle$ and

3.36 $$\begin{aligned} E([M, M]_t^2) &\leq 2E(N_t^2) + 2E(\langle M, M\rangle_t^2) = 2E(\langle N, N\rangle_t) + 2z \\ &\leq 8a^2 E(\langle M, M\rangle_t) + 2z \leq 8a^2 \sqrt{z} + 2z. \end{aligned}$$

On the other hand, $M^2 = 2M_- \cdot M + [M, M]$ (see I.4.45), so

3.37 $$y \leq 8E\left(\sup_{s \leq t} |M_- \cdot M_s|^2\right) + 2E([M, M]_t^2).$$

Finally the local martingale $Z = M_- \cdot M$ satisfies $\langle Z, Z\rangle = M_-^2 \cdot \langle M, M\rangle$, which is bounded. So Z is square-integrable, and by Doob's inequality I.1.37 we have

3.38 $$\begin{aligned} E\left(\sup_{s \leq t} |M_- \cdot M_s|^2\right) &\leq 4E(Z_t^2) = 4E(\langle Z, Z\rangle_t) \\ &\leq 4E\left[\left(\sup_{s \leq t} M_{s-}^2\right)\langle M, M\rangle_t\right] \leq 4\sqrt{yz} \end{aligned}$$

by Cauchy-Schwarz inequality. Putting together 3.36, 3.37 and 3.38 yields

$$y \leq 32\sqrt{yz} + 16a^2\sqrt{z} + 4z.$$

Since $y < \infty$, this is possible only if $y \leq [16\sqrt{z} + 2(65z + 4a^2\sqrt{z})^{1/2}]^2$. Hence we have $y \leq K_1 a^2\sqrt{z} + K_2 z$ with $K_1 = 32$ and $K_2 = 1032$, and this gives 3.35. □

Proof of (a) ⇒ (b) *in Theorem 3.13.* We suppose that $X^n \xrightarrow{\mathscr{L}} X$. By 3.20 we have $[Sk\text{-}\delta_{3,2}]$. By 3.28 we have $\sup_n \tilde{C}_t^{n,jj} < \infty$ for all $t \geq 0$, $j \leq d$. If we apply the previous lemma to each component $M^{n,j}(h)$ (defined by 3.26, and satisfying 3.27 and $|\Delta M^n(h)| \leq 2A$, where A satisfies 2.1), we have:

$$\sup_n E(|M^n(h)|_t^{*4}) < \infty$$

where $|M^n(h)|_t^* = \sup_{s \leq t} |M^n(h)_s|$. Therefore,

3.39 the sequence $\{|M^n(h)|_t^{*p}\}_{n \geq 1}$ is uniformly integrable if $p < 4$.

By II.5.10 (or 3.39), $B_t^n = E(X^n(h)_t)$ and $B_t = E(X(h)_t)$. By 3.28, $X^n(h) \xrightarrow{\mathscr{L}} X(h)$, and $X^n(h)$ and $X(h)$ are PII's. Hence 3.18 will imply [Sk-β_3], provided we prove that $|X^n(h)|_t^* := \sup_{s \leq t} |X^n(h)_s|$ has

3.40 the sequence $(|X^n(h)|_t^*)_{n \geq 1}$ is uniformly integrable.

Let also $|B^n|_t^* = \sup_{s \leq t} |B_s^n|$. Property 3.39 obviously implies:

3.41 $$\lim_{\theta \uparrow \infty} \sup_n P(|M^n(h)|_t^* > \theta) = 0$$

and since $X^n(h) \xrightarrow{\mathscr{L}} X(h)$ we deduce from VI.3.21i that $|X^n(h)|_t^*$ also satisfies 3.41. By difference, $|B^n|_t^*$ also satisfies 3.41; but the "random variables" $|B^n|_t^*$ are deterministic, so 3.41 for them just means that: $\sup_n |B^n|_t^* < \infty$. Since $|X^n(h)|_t^* \leq |B^n|_t^* + |M^n(h)|_t^*$, 3.39 yields 3.40, and we deduce [Sk-β_3].

Finally $J = J(X)$ obviously contains $J(X(h))$, while $X^n(h) \xrightarrow{\mathscr{L}} X(h)$, so $X^n(h)_t \xrightarrow{\mathscr{L}} X(h)_t$ for all $t \notin J$. Moreover $\Delta B_t = 0$ if $t \notin J$, hence [Sk-δ_3] implies that $B_t^n \to B_t$ for $t \notin J$, from which we deduce that $M^n(h)_t \xrightarrow{\mathscr{L}} M(h)_t$ for $t \notin J$. But $\tilde{C}_t^{n,jk} = E(M^{n,j}(h)_t M^{n,k}(h)_t)$ by II.5.10, so 3.39 implies that $\tilde{C}_t^n \to \tilde{C}_t$ for $t \notin J$: that is, we have [γ_3-D] with $D = \mathbb{R}_+ \backslash J$. □

§ 3d. Sufficient Condition for Convergence

It remains to prove the implications (b) ⇒ (a) and (b) ⇒ [Sk-$\beta\gamma\delta_3$] in 3.13. The main point consists in proving that (b) implies tightness of the sequence (X^n), and we proceed again through several lemmas.

3.42 **Lemma.** *Assume* [Sk-$\delta_{3,1}$]. *For each* $t > 0$ *there is a sequence* (t_n) *converging to* t, *such that* $t_n = t$ *if* $t \in D := \mathbb{R}_+ \backslash J(X)$ *and*

(i) $\nu^n(\{t_n\} \times g) \to \nu(\{t\} \times g)$ *for all* $g \in C_2(\mathbb{R}^d)$;
(ii) $\lim_{\eta \downarrow 0} \limsup_n \nu^n(([t-\eta, t+\eta] \backslash \{t_n\}) \times \{|x| > \varepsilon\}) = 0$ *for all* $\varepsilon > 0$.

Proof. a) Let $g \in C_1(\mathbb{R}^d)$, and set $\alpha_n = g * \nu^n$, $\alpha = g * \nu$. From VI.2.1 there is a sequence $\{t_n^g\}$ converging to t, such that

3.43 $$\Delta\alpha_n(t_n^g) = \nu^n(\{t_n^g\} \times g) \to \Delta\alpha(t) = \nu(\{t\} \times g),$$

3.44 $$\lim_{\eta \downarrow 0} \limsup_n w(\alpha_n'^g; [t-\eta, t+\eta]) = 0,$$

where $\alpha_n'^g(s) = \alpha_n(s) - \Delta\alpha_n(t_n^g) 1_{\{t_n^g \le s\}}$, and where w is the modulus defined in VI.1.4. But then $\alpha_n'^g = (g 1_{\{t_n^g\}^c}) * \nu^n$, which is increasing, so $w(\alpha_n'^g; [t-\eta, t+\eta]) = \nu^n(([t-\eta, t+\eta] \backslash \{t_n^g\}) \times g)$, and 3.44 yields

3.45 $$\lim_{\eta \downarrow 0} \limsup_n \nu^n(([t-\eta, t+\eta] \backslash \{t_n^g\}) \times g) = 0.$$

Recall also that $C_1(\mathbb{R}^d)$ contains all functions $g_a(x) = (a|x| - 1)^+ \wedge 1$ for $a \in \mathbb{Q}$, $a > 0$.

b) Suppose first that $t \notin J(X)$, so $\nu(\{t\} \times g) = 0$ for all g. Then by VI.2.1 one can take $t_n^g = t$ for all $g \in C_1(\mathbb{R}^d)$. Therefore if $t_n = t$, (ii) holds (apply 3.45 with $g = g_a$ and $a \ge 1/\varepsilon$), and the "convergence-determining" property of $C_1(\mathbb{R}^d)$ and 3.43 yield (i).

c) Secondly, assume that $t \in J(X)$, so there exists $q \in \mathbb{N}^*$ with $\nu(\{t\} \times g_q) > 0$. Now let $g, g' \in C_1(\mathbb{R}^d)$ with $g \ge Cg'$ (where $C > 0$) and $\nu(\{t\} \times g') > 0$. Then 3.43 and 3.45 yield

$$\liminf_n \nu^n(\{t_n^{g'}\} \times g) \ge C \lim_n \nu^n(\{t_n^{g'}\} \times g') = C\nu(\{t\} \times g') > 0$$

$$\lim_{\eta \downarrow 0} \limsup_n \nu^n(([t-\eta, t+\eta] \backslash \{t_n^g\}) \times g) = 0,$$

which obviously yield that $t_n^{g'} = t_n^g$ for all n large enough. Then we set $t_n = t_n^{g_q}$, where g_q is as above. If $q' \ge q$ then $g_{q'} \ge Cg_q$ for some $C > 0$, so what precedes shows that $t_n^{g_{q'}} = t_n$ for all n large enough, or in other words one may choose $t_n^{g_{q'}} = t_n$ for all n. Similarly if $g \in C_1(\mathbb{R}^d)$ and $\nu(\{t\} \times g) > 0$, there exists $q' \ge q$ with $g_{q'} \ge Cg$ for some $C > 0$ and again $t_n^g = t_n^{g_{q'}} = t_n$ for all n large enough, so we can again choose $t_n^g = t_n$. Finally, if $\nu(\{t\} \times g) = 0$ it follows from VI.2.1 that once more one may take $t_n^g = t_n$.

Therefore $t_n^g = t_n$ for all $g \in C_1(\mathbb{R}^d)$, and one concludes that (i) and (ii) hold exactly as in (b) above. □

3.46 **Corollary.** *Assume* $[\text{Sk-}\delta_{3,1}]$ *and let* $t > 0$ *and* (t_n) *be the sequence associated with* t *in 3.42. Then*

$$\Delta B_{t_n}^n \to \Delta B_t, \quad \Delta \tilde{C}_{t_n}^n \to \Delta \tilde{C}_t.$$

Proof. Let A satisfy 2.1 and $q > A$. If g_q is as in 2.7 or in the previous proof, we have $g_q h^j \in C_2(\mathbb{R}^d)$, so 3.42 yields

3.47 $$\nu^n(\{t_n\} \times g_q h) \to \nu(\{t\} \times g_q h).$$

We have $|g_q h - h| \le \frac{1}{q}$, and $\Delta B^n_s = \nu^n(\{s\} \times h)$; hence $|\Delta B^n_{t_n} - \nu^n(\{t_n\} \times g_q h)| \le \frac{1}{q}$ and $|\Delta B_t - \nu(\{t\} \times g_q h)| \le \frac{1}{q}$. Then 3.47 and the arbitraryness of q yield that $\Delta B^n_{t_n} \to \Delta B_t$. By 3.1 or II.5.9, $\Delta \tilde{C}^{n,jk}_s = \nu^n(\{s\} \times h^j h^k) - \Delta B^{n,j}_s \Delta B^{n,k}_s$, and the same argument yields

$$\nu^n(\{t_n\} \times h^j h^k) \to \nu(\{t\} \times h^j h^k),$$

so $\Delta \tilde{C}^n_{t_n} \to \Delta \tilde{C}_t$. □

3.48 **Corollary.** *Assume* [Sk-β_3], [γ_3-D] *and* [Sk-$\delta_{3,1}$], *where D is dense in $\mathbb{R}_+$. Then* [Sk-$\beta\gamma\delta_3$] *holds.*

Proof. a) We prove first that $\tilde{C}^{n,jj} \to \tilde{C}^{jj}$ in $\mathbb{D}(\mathbb{R})$. Since these functions are increasing and since they meet VI.2.20 by the previous corollary, we deduce the claim from VI.2.22 and VI.2.15 and [γ_3-D].

b) Next we prove that $\tilde{C}^n \to \tilde{C}$ in $\mathbb{D}(\mathbb{R}^{d\times d})$. To begin with, the components $(\tilde{C}^{n,jk})$ satisfy the conditions of VI.2.2, so $\sum_{j\le d} \tilde{C}^{n,jj} \to \sum_{j\le d} \tilde{C}^{jj}$ in $\mathbb{D}(\mathbb{R})$. Therefore VI.3.36 (applied to the "deterministic" processes $\tilde{C}^n$) implies that the sequence $(\tilde{C}^n)$ is tight. Then [γ_3-D] yields the claim.

c) Finally, let $g\colon \mathbb{R}^d \to \mathbb{R}^m$ have its components in $C_2(\mathbb{R}^d)$. Then all components of the sequence of $(d + d^2 + m)$-dimensional functions $(B^n, \tilde{C}^n, g * \nu^n)$ converge in $\mathbb{D}(\mathbb{R})$ to the relevant component of $(B, \tilde{C}, g * \nu)$ (use (b) and [Sk-$\delta_{3,2}$] and [Sk-β_3]). Moreover , by 3.42 and 3.46 they fulfill the conditions of VI.2.2, hence the result. □

3.49 **Theorem.** *Under* [Sk-$\beta\gamma\delta_3$], *the sequence* (X^n) *is tight.*

Proof. We essentially use Theorem VI.5.10. However, this theorem is stated for semimartingales, whereas here the X^n's are not necessarily semimartingales; on the other hand the characteristics are deterministic, which considerably simplifies the matter: hence we prefer to give a direct proof (one could also check step by step that the proof of VI.5.10 remains valid).

Let A satisfy 2.1. For $b > 0$, set

$$h_b(x) = bh\left(\frac{x}{b}\right),$$

which is again a truncation function. Moreover, the components of $h_b - h$ are in $C_2(\mathbb{R}^d)$. For each $q \in \mathbb{N}^*$ we set

$$A^{nq} = (h_q - h_{1/q}) * \nu^n$$
$$M^n(h_q) = X^n(h) - B^n + \sum_{s \leq \cdot} [h_q(\Delta X_s^n) - h(\Delta X_s^n)] - (h_q - h) * \nu^n$$
$$U^{nq} = A^{nq} + M^n(h_q)$$
$$V^{nq} = B^n + (h_{1/q} - h) * \nu^n$$
$$W^{nq} = \sum_{s \leq \cdot} [\Delta X_s^n - h_q(\Delta X_s^n)],$$

so that for each q we have:

3.50 $$X^n = U^{nq} + V^{nq} + W^{nq}.$$

It remains to prove that these decompositions satisfy the conditions of Lemma VI.3.32.

a) *For each $q \in \mathbb{N}^*$ the sequence $(U^{nq})_{n \geq 1}$ is tight.* By definition we have $M^n(h_q) = X^n(h_q) - B^n(h_q)$, where $B^n(h_q) = B^n + (h_q - h) * \nu^n$ is the first characteristic associated with the truncation function h_q: then $\langle M^{n,j}(h_q), M^{n,j}(h_q)\rangle = \tilde{C}^{n,jj}(h_q)$ is given by 3.1 with h_q instead of h. A simple computation shows that

$$\tilde{C}^{n,jj}(h_q) = \tilde{C}^{n,jj} + [(h_q^j)^2 - (h^j)^2] * \nu^n + \sum_{s \leq \cdot} [\nu^n(\{s\} \times h^j)^2 - \nu^n(\{s\} \times h_q^j)^2].$$

Hence U^{nq} is a locally square-integrable semimartingale and if we set (as in VI.5.16):

$$F^{nq} = \sum_{j \leq d} [\operatorname{Var}(A^{nq,j}) + \langle M^{n,j}(h_q), M^{n,j}(h_q)\rangle]$$

we have the following strong majoration ($\alpha \prec \beta$ stands for: $\beta - \alpha$ is increasing):

$$\begin{aligned} F^{nq} \prec \sum_{j \leq d} \Big\{ & [\tilde{C}^{n,jj} + |(h_q^j)^2 - (h^j)^2| * \nu^n] \\ & + \sum_{s \leq \cdot} \nu^n(\{s\} \times |h_q^j + h^j|)\nu^n(\{s\} \times |h_q^j - h^j|) + |h_q^j - h_{1/q}^j| * \nu^n \Big\} \\ \prec G^{nq} &:= \sum_{j \leq d} \tilde{C}^{n,jj} + \hat{h}_q * \nu^n, \end{aligned}$$

where

$$\hat{h}_q = \sum_{j \leq d} [|(h_q^j)^2 - (h^j)^2| + A(1+q)|h_q^j - h^j| + |h_q^j - h_{1/q}^j|].$$

Moreover $\hat{h}_q \in C_2(\mathbb{R}^d)$, so [Sk-$\beta\gamma\delta_3$] implies that $G^{nq} \to G^q := \sum_{j \leq d} \tilde{C}^{jj} + \hat{h}_q * \nu$ in $\mathbb{D}(\mathbb{R})$. Since in addition $U_0^{nq} = 0$, the conditions of Theorem VI.5.17 are fulfilled by the sequence $(U^{nq})_{n \geq 1}$ (with condition (C1)) and we deduce the result.

b) *(V^{nq}) satisfies VI.3.32ii.* By construction $|\Delta V^{nq}| \leq A/q$, and $V^{nq} \to B + (h_{1/q} - h) * \nu$ in $\mathbb{D}(\mathbb{R}^d)$ by [Sk-$\beta\gamma\delta_3$], hence the result.

c) *(W^{nq}) satisfies VI.3.32iii.* Since $h_q(x) = x$ for $|x| \leq q/A$, we obtain:

$$\sup_{s \leq N} |W_s^{nq}| > 0 \Rightarrow \sum_{s \leq N} g_{2A/q}(\Delta X_s^n) \geq 1.$$

So, using II.5.10,

3.51 $$P\Big(\sup_{s\le N}|W^{nq}_s|>0\Big)\le E\Big(\sum_{s\le N} g_{2A/q}(\Delta X^n_s)\Big)=g_{2A/q}*\nu^n_N.$$

By [Sk-$\beta\gamma\delta_3$], $g_{2A/q}*\nu^n_{N'}\to g_{2A/q}*\nu_{N'}$ for all $N'\ge N$, $N'\in D(X)$. Moreover, $\lim_{q\uparrow\infty} g_{2A/q}*\nu_{N'}=0$. Then VI.3.32iii immediately follows from 3.51. □

Now we are ready to prove (b) ⇒ (a) in 3.13. We assume (b), so [Sk-$\beta\gamma\delta_3$] holds by 3.48, and (X^n) is tight by 3.49. Hence it remains to prove that $\mathscr{L}(X)$ is the only limit point of the sequence $\{\mathscr{L}(X^n)\}$.

To this end, we suppose that a subsequence X^{n_k} converges in law to a limit X', which necessarily is a PII, whose characteristics are denoted by (B',C',ν'). Then by the necessary part of the theorem, we have [Sk-β_3], [γ_3-D'] and [Sk-$\delta_{3,2}$] with $D'=\mathbb{R}_+\backslash J(X')$, and where (B,C,ν) is replaced by (B',C',ν') and the sequence (n) is replaced by the subsequence (n_k). If we identify the two limits of B^{n_k} (resp. $\tilde{C}^{n_k}$, resp. $g*\nu^{n_k}$) in these conditions, we obtain that $B'=B$, that $\tilde{C}'_t=\tilde{C}_t$ for $t\in D'\cap D$ (where $D=\mathbb{R}_+\backslash J(X)$), and that $g*\nu'=g*\nu$ for all $g\in C_2(\mathbb{R}^d)$; we deduce that $\tilde{C}'=\tilde{C}$ (because $D\cap D'$ is still dense in $\mathbb{R}_+$) and $\nu'=\nu$, which in turn imply that $C'=C$. Hence X' and X have the same characteristics, a property which yields $\mathscr{L}(X')=\mathscr{L}(X)$ by Theorem II.5.2. Therefore $\mathscr{L}(X)$ is the only possible limit point of the sequence $\{\mathscr{L}(X^n)\}$ and we deduce that $X^n\xrightarrow{\mathscr{L}}X$. □

4. More on the General Case

In this section we go even further into what we might call pure technicalities. We immediately reassure the reader: these results are not used in the sequel, except in the next section.

At first, we turn back to finite-dimensional convergence, but for general PII, trying to obtain a result similar to Theorem 2.52 (and also a result similar to Theorem 2.35). Of course, if X has fixed times of discontinuity it would be unreasonnable to assume 2.53, and the key for stating a "reasonnable" condition to replace 2.53 is provided by Lemma 3.42: namely, instead of 2.53 we assume that the sequence (ν^n) meets the two conditions (i) and (ii) of 3.42.

Secondly we will deduce in § 4c a new necessary and sufficient condition for functional convergence, expressed in terms of characteristic functions.

§ 4a. Convergence of Non-Infinitesimal Rowwise Independent Arrays

Let us consider a d-dimensional triangular array scheme (χ^n_k) which is rowwise independent: exactly like in § 2c, we may and will assume that $K^n=\infty$, i.e. there are infinitely many entries in each row.

For further reference, we need to consider the case where 2.31 is not necessarily satisfied. However, we suppose that:

4.1 if $b_k^n = E[h(\chi_k^n)]$ and $Y_k^n = \chi_k^n - b_k^n$, then each sequence $(Y_k^n)_{k \geq 1}$ satisfies 2.31. □

In other words, the series $\sum_{(k)} \chi_k^n$ converges a. s. "after centering" independently on the order of summation, and for centering constants we may take b_k^n. Of course, by 2.39, we have the implication: 2.31 $\Rightarrow$ 4.1.

Instead of infinitesimality, we suppose that a suitable selection of "large" variables in row n converges in law (when $n \uparrow \infty$) to non-zero variables, while the "small" variables are uniformly small. A (relatively) simple way of expressing this idea is to consider an auxiliary sequence $(\chi_k^\infty)_{k \geq 1}$ of independent variables, which satisfies 4.1, and such that:

4.2 *Condition.* For each $\varepsilon > 0$ there is a finite integer $p(\varepsilon)$, and for each $n \in \bar{\mathbb{N}}^*$ a sequence $(k_j^n(\varepsilon))_{1 \leq j \leq p(\varepsilon)}$ of distinct indices, such that

(i) $\chi_{k_j^n(\varepsilon)}^n \xrightarrow{\mathscr{L}} \chi_{k_j^\infty(\varepsilon)}^\infty$

(ii) $\limsup_n \sup_{k: k \neq k_j^n(\varepsilon) \text{ for all } j} P(|\chi_k^n| > \varepsilon) \leq \varepsilon$

(here, $\limsup_n a_n = \lim_n \downarrow \sup_{n \leq m \leq \infty} a_m$). □

For each $n \in \bar{\mathbb{N}}^*$ we also consider an $\mathbb{R}^d$-valued variable ζ^n whose distribution is infinitely divisible, with characteristics (b^n, c^n, F^n) relative to some truncation function h, and which is independent from the sequence $(\chi_k^n)_{k \geq 1}$. As usual in this chapter, we assume that this truncation function h is *continuous.* We set

4.3 $$V^n = \zeta^n + \sum_{(k)} Y_k^n, \quad V^n(\varepsilon) = V^n - \sum_{1 \leq j \leq p(\varepsilon)} Y_{k_j^n(\varepsilon)}^n.$$

4.4 **Theorem.** *Assume 4.1 and 4.2.*

a) *In order that $V^n \xrightarrow{\mathscr{L}} V^\infty$ and that $V^n(\varepsilon) \xrightarrow{\mathscr{L}} V^\infty(\varepsilon)$ for all $\varepsilon > 0$, it is necessary and sufficient that the following three conditions hold*:

$[\beta_4]$ $\quad b^n \to b^\infty$;

$[\gamma_4]$ $$\tilde{c}^{n,jl} + \sum_k \{E[h^j h^l(\chi_k^n)] - E[h^j(\chi_k^n)]E[h^l(\chi_k^n)]\}$$
$$\to \tilde{c}^{\infty,jl} + \sum_k \{E[h^j h^l(\chi_k^\infty)] - E[h^j(\chi_k^\infty)]E[h^l(\chi_k^\infty)]\};$$

$[\delta_{4,i}]$ $$F^n(g) + \sum_k E[g(\chi_k^n)] \to F^\infty(g) + \sum_k E[g(\chi_k^\infty)] \quad \text{for all } g \in C_i(\mathbb{R}^d)$$

(here, $\tilde{c}^n$ is defined by 2.5), *for either $i = 1$ or $i = 2$.*

b) *If moreover the characteristic function of V^∞ does not vanish, then $[\beta_4]$, $[\gamma_4]$, $[\delta_{4,i}]$ are necessary and sufficient for $V^n \xrightarrow{\mathscr{L}} V^\infty$* (under 4.1 and 4.2, of course).

Note that this theorem gives conditions for convergence of the sums of the Y_k^n's, in terms of the χ_k^n's. If one can add up the χ_k^n's themselves, namely under 2.31, we have another version of this theorem. For this, set first:

4.5 $$\xi^n = \sum_{(k)} \chi_k^n, \quad \xi^n(\varepsilon) = \xi^n - \sum_{1 \le j \le p(\varepsilon)} \chi_{k_j^n(\varepsilon)}^n.$$

4.6 **Corollary.** *We assume 2.31 and 4.2. Then*

a) *In order that $\xi^n + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty + \zeta^\infty$ and that $\xi^n(\varepsilon) + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty(\varepsilon) + \zeta^\infty$ for all $\varepsilon > 0$, it is necessary and sufficient that we have*

$$[\hat{\beta}_4] \quad b^n + \sum_k E[h(\chi_k^n)] \to b^\infty + \sum_k E[h(\chi_k^\infty)]$$

and $[\gamma_4]$ and $[\delta_{4,i}]$ above (for either $i = 1$ or $i = 2$).

b) *If moreover the characteristic function of $\xi^\infty + \zeta^\infty$ does not vanish, then $[\hat{\beta}_4]$, $[\gamma_4]$, $[\delta_{4,i}]$ are necessary and sufficient for $\xi^n + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty + \zeta^\infty$.*

We remark that Theorem 2.35 is a particular case of this corollary, namely the case when $\chi_k^\infty = 0$ for all $k \ge 1$: indeed, if this is true, 4.2 is clearly equivalent to infinitesimality for the array (χ_k^n), and 2.35 reduces to part (b) of the corollary.

Needless to say, when the characteristic function $E(e^{iu\cdot V^\infty})$ happens to vanish (which amounts to saying that one of the characteristic functions $E(e^{iu\cdot \chi_k^\infty})$ vanishes), this theorem and its corollary are *not* very satisfactory! The following examplifies what can happen in that case:

4.7 **Example.** Let $\chi_k^n = 0$ for $k \ge 2$; let $\chi_1^n = Y$ be a 1-dimensional variable whose characteristic function is

$$\varphi(u) = [1 - |u|]^+.$$

Let ζ^n be as above (with $d = 1$), and suppose that $\psi_{b^n,c^n,F^n}(u) \to \psi_{b^\infty,c^\infty,F^\infty}(u)$ for $|u| \le 1$, but not for all u in $\mathbb{R}$. Then $\xi^n + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty + \zeta^\infty$ but ζ^n does not converge in law to ζ^∞, and $[\hat{\beta}_4]$, $[\gamma_4]$, $[\delta_{4,i}]$ are not all satisfied. □

Proof of Corollary 4.6. Let $\zeta'^n = \zeta^n + \sum_{(k)} b_k^n$, which admits the characteristics (b'^n, c^n, F^n) with $b'^n = b^n + \sum_{(k)} b_k^n$: hence $[\hat{\beta}_4]$ for ζ^n and $[\beta_4]$ for ζ'^n are the same, and conditions $[\gamma_4]$ (resp. $[\delta_{4,i}]$) are the same for ζ^n and ζ'^n. Now, if V'^n and $V'^n(\varepsilon)$ are associated to ζ'^n by 4.3, we have

$$V'^n = \xi^n + \zeta^n, \quad V'^n(\varepsilon) = \xi^n(\varepsilon) + \zeta^n + \sum_{1 \le j \le p(\varepsilon)} b_{k_j^n(\varepsilon)}^n.$$

Now, 4.2 implies that $b_{k_j^n(\varepsilon)}^n \to b_{k_j^\infty(\varepsilon)}^\infty$ for all $j \le p(\varepsilon)$. Hence we have the equivalences:

$$\xi^n + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty + \zeta^\infty \quad \Leftrightarrow \quad V'^n \xrightarrow{\mathcal{L}} V'^\infty$$
$$\xi^n(\varepsilon) + \zeta^n \xrightarrow{\mathcal{L}} \xi^\infty(\varepsilon) + \zeta^\infty \Leftrightarrow V'^n(\varepsilon) \xrightarrow{\mathcal{L}} V'^\infty(\varepsilon).$$

Therefore, we immediately deduce 4.6 from 4.4. □

Now we proceed to the proof of 4.4. It is essentially the same than for 2.35, with additional complications. Exactly as for 2.35, it is enough to prove the result when h is uniformly continuous, an assumption made from now on until the end of § 4a. Set

4.8
$$\begin{cases} K^n(\varepsilon) = \{k_j^n(\varepsilon)\colon 1 \le j \le p(\varepsilon)\} \\ b_k'^n = E[h(Y_k^n)] \\ Z_k^n = Y_k^n - b_k'^n \\ M^n(\varepsilon) = \sup\limits_{k \notin K^n(\varepsilon)} |b_k^n| \\ M'^n(\varepsilon) = \sup\limits_{k \notin K^n(\varepsilon)} |b_k'^n| \end{cases}$$

(recall that (Y_k^n) satisfies 2.31, so (Z_k^n) also by Lemma 2.39).

$[\tilde{\gamma}_4]$ $\quad \tilde{c}^{n,jl} + \sum_k E[h^j h^l(Y_k^n)] \to \tilde{c}^{\infty,jl} + \sum_k E[h^j h^l(Y_k^\infty)]$

$[\tilde{\delta}_4]$ $\quad F^n(g) + \sum_k E[g(Y_k^n)] \to F^\infty(g) + \sum_k E[g(Y_k^\infty)]$ for all $g \in C_2(\mathbb{R}^d)$.

4.9 Lemma. a) $Y_{k_j^n(\varepsilon)}^n \xrightarrow{\mathscr{L}} Y_{k_j^\infty(\varepsilon)}^\infty$ (hence $b_{k_j^n(\varepsilon)}'^n \to b_{k_j^\infty(\varepsilon)}'^\infty$) *and* $Z_{k_j^n(\varepsilon)}^n \xrightarrow{\mathscr{L}} Z_{k_j^\infty(\varepsilon)}^\infty$.

b) *If* $\varepsilon \le 1/2A(1 + A)$ (where A satisfies 2.1), *then* $\limsup_n M^n(\varepsilon) \le \varepsilon(1 + A)$ *and* $\limsup_n M'^n(\varepsilon) \le 3\varepsilon(1 + A)$.

Proof. (a) immediately follows from 4.2i.

b) $|b_k^n| \le \varepsilon + AP(|\chi_k^n| > \varepsilon)$ if $\varepsilon \le 1/A$, because of 2.1: hence the first claim of (b) follows from 4.2ii. Moreover $P(|Y_k^n| > 2\varepsilon(1 + A)) \le P(|\chi_k^n| > \varepsilon)$ if $|b_k^n| \le \varepsilon(1 + 2A)$. Then 4.2ii and the result about $M^n(\varepsilon)$ yield

4.10
$$\limsup_n \sup_{k \notin K_n(\varepsilon)} P(|Y_k^n| > 2\varepsilon(1 + A)) \le \varepsilon,$$

while $|b_k'^n| \le 2\varepsilon(1 + A) + AP(|Y_k^n| > 2\varepsilon(1 + A))$ when $2\varepsilon(1 + A) \le 1/A$. Therefore the second statement in (b) follows from 4.10. □

4.11 Lemma. *We have the equivalence*: $[\delta_{4,i}] \Leftrightarrow [\tilde{\delta}_4]$ *for* $i = 1$ *and* $i = 2$.

Proof. That $[\delta_{4,1}] \Leftrightarrow [\delta_{4,2}]$ follows (as usual) from the properties of $C_1(\mathbb{R}^d)$ in 2.7. Similarly $[\tilde{\delta}_4]$ holds for all $g \in C_2(\mathbb{R}^d)$ as soon as it holds for all $g \in C_1(\mathbb{R}^d)$. So if $\delta_k^n(g) = E[g(\chi_k^n) - g(Y_k^n)]$ it is enough to prove that $\sum_k \delta_k^n(g) \to \sum_k \delta_k^\infty(g)$ for all $g \in C_1(\mathbb{R}^d)$, under either $[\delta_{4,1}]$ or $[\tilde{\delta}_4]$.

Assume for instance $[\delta_{4,1}]$. Let $\varepsilon > 0$, and let $\theta > 0$ such that $g(x) = 0$ for $|x| \le \theta$. There is $\eta > 0$ such that $|x - y| \le \eta \Rightarrow |g(x) - g(y)| \le \varepsilon$, and set $\delta = \frac{\theta}{2} \wedge \eta$. Recall that $Y_k^n = \chi_k^n - b_k^n$, so exactly like in 2.41,

$$|b_k^n| \le \delta \Rightarrow |\delta_k^n(g)| \le \varepsilon P\left(|\chi_k^n| > \frac{\theta}{2}\right).$$

We have $P(|\chi_k^n| > \theta/2) \le E[g_{4/\theta}(\chi_k^n)]$. Let $\varepsilon' = \delta/2(1 + A)$. From 4.9b we deduce that for all n large enough, $|b_k^n| \le M^n(\varepsilon') \le \delta$ for all $k \notin K^n(\varepsilon')$, while $\delta^n_{k_j^n(\varepsilon')}(g) \to \delta^\infty_{k_j^\infty(\varepsilon')}(g)$ by 4.9a. Hence

$$\begin{aligned}\limsup_n \left|\sum_k (\delta_k^n(g) - \delta_k^\infty(g))\right| &\le \limsup_n \left\{\sum_{k\notin K^n(\varepsilon')} |\delta_k^n(g)| + \sum_{k\notin K^\infty(\varepsilon')} |\delta_k^\infty(g)|\right\}\\ &\le \varepsilon \limsup_n \left\{\sum_k E[g_{4/\theta}(\chi_k^n)] + \sum_k E[g_{4/\theta}(\chi_k^\infty)]\right\}\\ &\le \varepsilon\left\{F^\infty(g_{4/\theta}) + 2\sum_k E[g_{4/\theta}(\chi_k^\infty)]\right\},\end{aligned}$$

where the last inequality follows from $[\delta_{4,i}]$. Since $\varepsilon > 0$ is arbitrary, we obtain the result. Conversely, if $[\tilde{\delta}_4]$ holds, the same argument shows the result, after interchanging χ_k^n and Y_k^n. □

4.12 Lemma. *We have the equivalence* $[\gamma_4] + [\delta_{4,i}] \Leftrightarrow [\tilde{\gamma}_4] + [\tilde{\delta}_4]$ *for* $i = 1, 2$.

Proof. Due to 4.11, it suffices to prove that $[\gamma_4] \Leftrightarrow [\tilde{\gamma}_4]$ under $[\delta_{4,2}]$. For this, it is sufficient to prove that $\sum_k \gamma_k^{n,jl} \to \sum_k \gamma_k^{\infty,jl}$, where $\gamma_k^{n,jl}$ is defined like in the proof of 2.42. With the notation of this proof, we also have the majoration

$$|b_k^n| \le \delta := \frac{1}{2A} \wedge \eta \Rightarrow |\gamma_k^{n,jl}| \le \varepsilon P\left(|\chi_k^n| > \frac{1}{2A}\right),$$

while if $\varepsilon' = \delta/2(1 + A)$ we have $|b_k^n| \le \delta$ for $k \notin K^n(\varepsilon')$, for all n large enough, and $\gamma^{n,jl}_{k_j^n(\varepsilon')} \to \gamma^{\infty,jl}_{k_j^\infty(\varepsilon')}$ by Lemma 4.9, and the same argument than in the previous proof gives the result. □

Each sequence $(Y_k^n)_{k\ge 1}$ satisfies 2.31, hence $\sum_k |b_k'^n| < \infty$, and the following functions are well-defined (cf. after Lemma 2.43):

4.13
$$\begin{cases} \rho^n(u) = \prod_k \{E(e^{iu\cdot Y_k^n}) \exp - E(e^{iu\cdot Y_k^n} - 1 - iu\cdot h(Y_k^n))\} \\ \rho_\varepsilon^n(u) = \prod_{k\notin K^n(\varepsilon)} \{E(e^{iu\cdot Y_k^n}) \exp - E(e^{iu\cdot Y_k^n} - 1 - iu\cdot h(Y_k^n))\} \end{cases}$$

4.14 Lemma. *Assume that* $\sup_n \sum_k E(|Z_k^n|^2 \wedge 1) < \infty$. *Then*

a) $\sum_k b_k'^n \to \sum_k b_k'^\infty$;

b) $\sup_n \sum_k E(|Y_k^n|^2 \wedge 1) < \infty$;

c) $\rho^n(u) \to \rho^\infty(u)$ *and* $\rho_\varepsilon^n(u) \to \rho_\varepsilon^\infty(u)$ *for all* $u \in \mathbb{R}^d$, $\varepsilon > 0$.

Proof. We begin with some preliminaries. Set $K = \sup_n \sum_k E(|Z_k^n|^2 \wedge 1)$. We have $b_k^n = E[h(Y_k^n + b_k^n)]$, hence

$$b_k'^n = E[h(Y_k^n) - h(Y_k^n + b_k^n) + b_k^n].$$

Let $\varepsilon > 0$, so there exists $\eta > 0$ such that $|x - y| \leq \eta \Rightarrow |h(x) - h(y)| \leq \varepsilon$, and set $\delta = \eta \wedge \dfrac{1}{2A} \wedge \varepsilon$. Then $|h(Y_k^n) - h(Y_k^n + b_k^n) - b_k^n|$ is smaller than 2ε if $|b_k^n| \leq \delta$, and is equal to 0 if $|b_k^n| \leq \delta$ and $|Y_k^n| \leq 1/2A$; moreover we have $|Y_k^n| \leq 1/2A$ whenever $|Z_k^n| \leq 1/4A$ and $|b_k'^n| \leq 1/4A$. Therefore:

4.15 $$|b_k^n| \leq \delta, \quad |b_k'^n| \leq \frac{1}{4A} \Rightarrow |b_k'^n| \leq 2\varepsilon P\left(|Z_k^n| > \frac{1}{4A}\right) \leq 32A^2\varepsilon E(|Z_k^n|^2 \wedge 1).$$

Now we can proceed to the proof itself.

a) Set $\varepsilon' = \delta/8(1 + A)$. From 4.9b, for all n large enough we have $|b_k^n| \leq \delta$ and $|b_k'^n| \leq \delta/2 \leq 1/4A$ if $k \notin K^n(\varepsilon')$. Since $b'^n_{k_j^n(\varepsilon')} \to b'^\infty_{k_j^\infty(\varepsilon')}$ we have (like in 4.11), using 4.15:

$$\limsup_n \left|\sum_k b_k'^n - \sum_k b_k'^\infty\right| \leq \limsup_n \left[\sum_{k \notin K^n(\varepsilon')} |b_k'^b| + \sum_{k \notin K^n(\varepsilon')} |b_k'^\infty|\right]$$
$$\leq 64A^2K\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we obtain the result.

b) Exactly like in (a), we prove that $\sum_k |b_k'^n| \to \sum_k |b_k'^\infty|$. Thus $K' = \sup_n \sum_k |b_k'^n|$ is finite. Moreover, $|b_k'^n| \leq A$, and $|Y_k^n|^2 \wedge 1 \leq 2(|Z_k^n|^2 \wedge 1) + 2|b_k'^n|^2$, so we have:

$$\sup_n \sum_k E(|Y_k^n|^2 \wedge 1) \leq K'' := 2K + 2AK'.$$

c) Let $u \in \mathbb{R}^d$ be fixed, and $\delta_k^n = E(e^{iu \cdot Y_k^n}) - 1$, so that

$$\rho^n(u) = \bar{\delta}^n \exp iu \cdot \sum_k b_k'^n$$

where $\bar{\delta}^n = \prod_k [(1 + \delta_k^n) \exp -\delta_k^n]$. In virtue of (a) it remains to prove that $\bar{\delta}^n \to \bar{\delta}^\infty$ in order to obtain that $\rho^n(u) \to \rho^\infty(u)$.

There is a constant C_u such that $|e^{iu \cdot x} - 1 - iu \cdot h(x)| \leq C_u(|x|^2 \wedge 1)$, so

4.16 $$\begin{aligned} |\delta_k^n| &= |E(e^{iu \cdot Y_k^n} - 1 - iu \cdot h(Y_k^n)) + iu \cdot b_k'^n| \\ &\leq C_u E(|Y_k^n|^2 \wedge 1) + |b_k'^n| \cdot |u| \end{aligned}$$

and with the notation of (b) it follows that

4.17 $$\sum_k |\delta_k^n| \leq C_1 := C_u K'' + |u| K'.$$

Moreover, let $\varepsilon > 0$, and ε' and δ be as in (a). For all n large enough, we have for all $k \notin K^n(\varepsilon')$: $|b_k^n| \leq \delta \leq \varepsilon$ and $P(|\chi_k^n| > \varepsilon) \leq \varepsilon$ (because $\varepsilon' \leq \varepsilon$), hence $P(|Y_k^n| > 2\varepsilon) \leq \varepsilon$. Therefore we deduce from 2.18 that for all n large enough,

4.18 $$\sup_{k \notin K^n(\varepsilon')} |\delta_k^n| \leq 2\varepsilon + \varepsilon C_2(|u|).$$

Therefore, if $\gamma := C_1(2 + C_2(|u|))$, 4.17 and 4.18 yield

4.19 $$\sum_{k \notin K^n(\varepsilon')} |\delta_k^n|^2 \leq \gamma\varepsilon \quad \text{for all } n \text{ large enough.}$$

Let $\operatorname{Log} x$ be the principal determination of the logarithm of $x \in \mathbb{C}$; we have $|-x + \operatorname{Log}(1 + x)| \leq C|x|^2$ for a constant C, when $|x| \leq 1/2$. There is another constant C' such that $|e^x - 1| \leq C'|x|$ for all $x \in \mathbb{C}$ with $|x| \leq \gamma$. Since $|\delta_k^n| \leq 1/2$ for all $k \notin K^n(\varepsilon')$ by 4.18, provided ε is small enough, we deduce from 4.19:

$$\left|\prod_{k \notin K^n(\varepsilon')} [(1 + \delta_k^n)e^{-\delta_k^n}] - 1\right| = \left|\exp \sum_{k \notin K^n(\varepsilon')} \{-\delta_k^n + \operatorname{Log}(1 + \delta_k^n)\} - 1\right|$$

4.20
$$\leq CC' \sum_{k \notin K^n(\varepsilon')} |\delta_k^n|^2 \leq CC'\gamma\varepsilon.$$

On the other hand, 4.9a implies that $\delta^n_{k_j^n(\varepsilon')} \to \delta^\infty_{k_j^\infty(\varepsilon')}$. Then $\bar{\delta}^n = \bar{\delta}_1^n \bar{\delta}_2^n$ where $\bar{\delta}_1^n = \prod_{k \in K^n(\varepsilon')} [(1 + \delta_k^n)e^{-\delta_k^n}]$ and $\bar{\delta}_2^n = \prod_{k \notin K^n(\varepsilon')} [(1 + \delta_k^n)e^{-\delta_k^n}]$. 4.9a implies $\bar{\delta}_1^n \to \bar{\delta}_1^\infty$, and 4.20 gives $|\bar{\delta}_2^n - 1| \leq CC'\gamma\varepsilon$ for n large enough. Then

$$|\bar{\delta}^n - \bar{\delta}^\infty| \leq |\bar{\delta}_1^n - \bar{\delta}_1^\infty| |\bar{\delta}_2^n| + |\bar{\delta}_1^\infty| (|\bar{\delta}_2^n - 1| + |\bar{\delta}_2^\infty - 1|)$$

$$\limsup_n |\bar{\delta}^n - \bar{\delta}^\infty| \leq \left(\lim_n |\bar{\delta}_1^n - \bar{\delta}_1^\infty|\right)(1 + CC'\gamma\varepsilon) + |\bar{\delta}_1^\infty| 2CC'\gamma\varepsilon$$

4.21
$$\leq 2CC'\gamma\varepsilon|\bar{\delta}_1^\infty|$$

because of 4.20. Since $\varepsilon > 0$ is arbitrarily small, 4.21 yields that $\bar{\delta}^n \to \bar{\delta}^\infty$, and we deduce that $\rho^n(u) \to \rho^\infty(u)$.

In order to finish the proof, it remains to prove that $\rho_\varepsilon^n(u) \to \rho_\varepsilon^\infty(u)$, where $\varepsilon > 0$ is arbitrary. Set

$$\chi_k^n(\varepsilon) = \begin{cases} \chi_k^n & \text{if } k \notin K^n(\varepsilon) \\ 0 & \text{if } k \in K^n(\varepsilon) \end{cases}$$

and associate $Y_k^n(\varepsilon)$ to $\chi_k^n(\varepsilon)$ by 4.8. The array $(\chi_k^n(\varepsilon))$ obviously satisfies 4.2, and the function associated by the first formula in 4.13 is exactly $\varphi_\varepsilon^n(u)$. Moreover, the hypothesis implies that

$$\sup_n \sum_k E(|Z_k^n(\varepsilon)|^2 \wedge 1) = \sup_n \sum_{k \notin K^n(\varepsilon)} E(|Z_k^n|^2 \wedge 1) < \infty.$$

Therefore, an application of what precedes gives: $\rho_\varepsilon^n(u) \to \rho_\varepsilon^\infty(u)$. □

Proof of Theorem 4.4. We denote by φ^n and φ_ε^n the characteristic functions of V^n and $V^n(\varepsilon)$. We set

$$\tilde{F}^n(g) = F^n(g) + \sum_k E[g(Y_k^n)], \quad \tilde{F}_\varepsilon^n(g) = F^n(g) + \sum_{k \notin K^n(\varepsilon)} E[g(Y_k^n)]$$

(since (Y_k^n) satisfies 2.31, $\tilde{F}^n$ and $\tilde{F}_\varepsilon^n$ satisfy 2.2). A simple computation shows that

4.22
$$\varphi^n(u) = \rho^n(u) \exp \psi_{b^n, c^n, \tilde{F}^n}(u), \quad \varphi_\varepsilon^n(u) = \rho_\varepsilon^n(u) \exp \psi_{b^n, c^n, \tilde{F}_\varepsilon^n}(u),$$

and we also have

4.23 $$\psi_{b^n,c^n,\tilde{F}^n}(u) = \psi_{b^n,c^n,\tilde{F}^n_\varepsilon}(u) + \sum_{k\in K^n(\varepsilon)} E(e^{iu\cdot Y^n_k} - 1 - iu\cdot h(Y^n_k))$$

(i) *Necessary condition.* We assume that $V^n \xrightarrow{\mathscr{L}} V^\infty$ and $V^n(\varepsilon) \xrightarrow{\mathscr{L}} V^\infty(\varepsilon)$ for all $\varepsilon > 0$. We have $\varphi^n \xrightarrow{\mathscr{L}} \varphi^\infty$ uniformly on compact sets. There is $\theta > 0$ such that $|\varphi^\infty(u)| \geq 3/4$ for $|u| \leq \theta$, and there is n_0 such that $|\varphi^n(u)| \geq 1/2$ for $|u| \leq \theta$, $n \geq n_0$. Let also $\varepsilon = 1/10A(1 + A)$. Then by 4.2 and 4.9 there is $n_1 \geq n_0$ such that

$$n \geq n_1 \Rightarrow \sup_{k\notin K^n(\varepsilon)} P(|\chi^n_k| > \varepsilon) \leq \varepsilon, \quad M^n(\varepsilon) \leq 2\varepsilon(1 + A)$$

and since $Y^n_k = \chi^n_k - b^n_k$ we obtain

4.24 $$n \geq n_1 \Rightarrow \sup_{k\notin K^n(\varepsilon)} P\left(|Y^n_k| > \frac{1}{4A}\right) \leq \frac{1}{4A}.$$

Since

$$\varphi^n(u) = E(e^{iu\cdot\zeta^n})\left(\prod_{k\in K^n(\varepsilon)} E(e^{iu\cdot Y^n_k})\right)\left(\prod_{k\notin K^n(\varepsilon)} E(e^{iu\cdot Y^n_k})\right)$$

and since $|\varphi^n(u)| \geq 1/2$ for $n \geq n_1$, $|u| \leq \theta$, each term of this product has also a modulus bigger than 1/2, and

$$n \geq n_1, \quad |u| \leq \theta \Rightarrow \sum_{k\notin K^n(\varepsilon)} -\operatorname{Log}|E(e^{iu\cdot Y^n_k})| \leq \operatorname{Log} 2.$$

Applying Lemma 2.22 to $\mathscr{L}(Y^n_k)$ for $k \notin K^n(\varepsilon)$, and $1 - |x|^2 \leq -\frac{1}{2}\operatorname{Log}|x|$ for $|x| \leq 1$ and 4.24 and summing up, we get

4.25 $$n \geq n_1 \Rightarrow \sum_{k\notin K^n(\varepsilon)} E(|Z^n_k|^2 \wedge 1) \leq \frac{1}{2}C'(\theta, A)\omega_d\theta^d \operatorname{Log} 2$$

where ω_d is the volumn of the unit sphere in $\mathbb{R}^d$. Moreover $K^n(\varepsilon)$ has exactly $p(\varepsilon)$ elements, so 4.25 implies

$$n \geq n_1 \Rightarrow \sum_k E(|Z^n_k|^2 \wedge 1) \leq p(\varepsilon) + \frac{1}{2}C'(\theta, A)\omega_d\theta^d \operatorname{Log} 2.$$

Since each sequence $(Z^n_k)_{k\geq 1}$ satisfies 2.31 (by Lemma 2.39), we deduce that

4.26 $$\sup_n \sum_k E(|Z^n_k|^2 \wedge 1) < \infty.$$

It follows from Lemma 4.13 that $\rho^n_\varepsilon(u) \to \rho^\infty_\varepsilon(u)$, while by hypothesis we also have $\varphi^n_\varepsilon(u) \to \varphi^\infty_\varepsilon(u)$, hence we deduce from 4.22 that

4.27 $$\psi_{b^n,c^n,\tilde{F}^n_\varepsilon}(u) \to \psi_{b^\infty,c^\infty,\tilde{F}^\infty_\varepsilon}(u) \quad \text{if } \rho^\infty_\varepsilon(u) \neq 0.$$

Finally, if we denote by $\alpha^n_\varepsilon(u)$ the last sum in the right-hand side of 4.23, then 4.9a implies that $\alpha^n_\varepsilon \to \alpha^\infty_\varepsilon(u)$. Then 4.23 and 4.27 give

4.28 $\psi_{b^n,c^n,\tilde{F}^n}(u) \to \psi_{b^\infty,c^\infty,\tilde{F}^\infty}(u)$ if $\rho_\varepsilon^\infty(u) \neq 0$ for some $\varepsilon \in \left(0, \frac{1}{10A(1+A)}\right)$.

However, $\rho_\varepsilon^\infty(u) = 0$ if and only if $E(e^{iu\cdot Y_k^\infty}) = 0$ for some $k \notin K^\infty(\varepsilon)$, or equivalently if and only if $E(\exp iu\cdot\chi_k^\infty) = 0$ for some $k \notin K^\infty(\varepsilon)$. But in virtue of 2.16b,

$$|E(e^{iu\cdot\chi_k^\infty} - 1)| \leq C_2(|u|)E(|\chi_k^\infty| \wedge 1) \leq C_2(|u|)2\varepsilon.$$

Then for each $u \in \mathbb{R}^d$ there is $\varepsilon > 0$ such that $C_2(|u|)2\varepsilon \leq 1/2$, implying that $E(\exp iu\cdot\chi_k^\infty) \neq 0$ for all $k \notin K^\infty(\varepsilon)$, and thus $\rho_\varepsilon^\infty(u) \neq 0$. Thus we deduce from 4.28 that

4.29 $$\psi_{b^n,c^n,\tilde{F}^n}(u) \to \psi_{b^\infty,c^\infty,\tilde{F}^\infty}(u) \quad \text{for all } u \in \mathbb{R}^d.$$

Now we apply Theorem 2.9: the sequence $(b^n, c^n, \tilde{F}^n)$ satisfies $[\beta_1]$, $[\gamma_1]$, $[\delta_{1,2}]$ relatively to $(b^\infty, c^\infty, \tilde{F}^\infty)$. But with this notation $[\beta_1] = [\beta_4]$ and $[\gamma_1] = [\tilde{\gamma}_4]$ and $[\delta_{1,2}] = [\tilde{\delta}_4]$, and by Lemma 4.12 we also have $[\gamma_4]$ and $[\delta_{1,2}]$.

(ii) *Sufficient condition.* We assume $[\beta_4]$, $[\gamma_4]$, $[\delta_{4,i}]$ (for $i = 1$ or $i = 2$). We have $[\tilde{\gamma}_4]$ and $[\tilde{\delta}_4]$ by 4.12, and therefore we deduce from Theorem 2.9 that 4.29 holds. On the other hand, one easily sees that $[\tilde{\gamma}_4]$ and $[\tilde{\delta}_4]$ imply that $\sup_n \sum_k E(|Y_k^n|^2 \wedge 1) < \infty$. Now, a thorough examination of the proof of Lemma 4.14 reveals that this property is sufficient for its conclusion to hold (as a matter of fact, only the "preliminaries" need to be modified, and they indeed become slightly easier). Thus $\rho^n(u) \to \rho^\infty(u)$ and $\rho_\varepsilon^n(u) \to \rho_\varepsilon^\infty(u)$. These facts, plus 4.29 and 4.22, give that $\varphi^n(u) \to \varphi^\infty(u)$ and $\varphi_\varepsilon^n(u) \to \varphi_\varepsilon^\infty(u)$, which are the wanted results.

(iii) *Part* (b) *of 4.4.* Let $\hat{\varphi}_\varepsilon^n(u) = \prod_{k\in K^n(\varepsilon)} E(e^{iu\cdot Y_k^n})$. Then $\varphi^n = \varphi_\varepsilon^n \hat{\varphi}_\varepsilon^n$, and 4.9a implies that $\hat{\varphi}_\varepsilon^n(u) \to \hat{\varphi}_\varepsilon^\infty(u)$ for all u, ε. Hence if $\varphi^\infty(u) \neq 0$, we have $\varphi_\varepsilon^\infty(u) \neq 0$ for all $\varepsilon > 0$, and we deduce that $\varphi^n(u) \to \varphi^\infty(u)$ if and only if $\varphi_\varepsilon^n(u) \to \varphi_\varepsilon^\infty(u)$ (for any fixed $\varepsilon > 0$) that is, $V^n \xrightarrow{\mathcal{L}} V^\infty$ is equivalent to $V^n(\varepsilon) \xrightarrow{\mathcal{L}} V^\infty(\varepsilon)$. □

§ 4b. Finite-Dimensional Convergence for General PII

In this subsection we state another version of Theorem 4.4, that is suitable for PII's: it bears the same relations to 4.4 than 2.52 did to 2.35.

For each $n \in \bar{\mathbb{N}}^*$ we consider a PII X^n with characteristics (B^n, C^n, ν^n) relative to some truncation function h. We define $\tilde{C}^n$ by 3.1, and D is a subset of $\mathbb{R}_+$. We also write X and (B, C, ν) and $\tilde{C}$ instead of X^∞ and $(B^\infty, C^\infty, \nu^\infty)$ and $\tilde{C}^\infty$.

4.30 *Condition.* For each $\varepsilon > 0$ there is an increasing sequence of (possibly infinite) times $(t_j^n(\varepsilon))_{j\geq 1}$ with $\lim_j \uparrow t_j^n(\varepsilon) = \infty$ and $t_j^n(\varepsilon) < t_{j+1}^n(\varepsilon)$ if $t_j^n(\varepsilon) < \infty$, such that:

(i) if $s, t \in D \cup \{0\}$ with $s < t$ and if $s < t_j^\infty(\varepsilon) \leq t$, we have $s < t_j^n(\varepsilon) \leq t$ for all n large enough, and $\nu^n(\{t_j^n(\varepsilon)\} \times g) \to \nu^\infty(\{t_j^\infty(\varepsilon)\} \times g)$ for all $g \in C_1(\mathbb{R}^d)$;

(ii) $\limsup_n \sup_{s\leq t, s\neq t_j^n(\varepsilon) \text{ for all } j} \nu^n(\{s\} \times \{|x| > \varepsilon\}) \leq \varepsilon$ for all $t \in D$.

We also define the following PII:

4.31 $$X^n(\varepsilon)_t = X^n_t - \sum_{k:\, t^n_k(\varepsilon) \le t}^{\infty} \Delta X^n_{t^n_k(\varepsilon)}.$$

4.32 **Theorem.** *We suppose that 4.30 holds for some subset $D \subset \mathbb{R}_+$, and that h is continuous. Then*

a) $X^n \xrightarrow{\mathscr{L}(D)} X$ *and* $X^n(\varepsilon) \xrightarrow{\mathscr{L}(D)} X(\varepsilon)$ *for all* $\varepsilon > 0$, *if and only if* $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$, $[\delta_{3,i}\text{-}D]$ *hold (either for $i = 1$ or for $i = 2$).*

b) *If moreover, for each $t \in D$ the characteristic function $E(e^{iu\cdot X_t})$ never vanishes, then $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$, $[\delta_{3,1}\text{-}D]$ are necessary and sufficient for having* $X^n \xrightarrow{\mathscr{L}(D)} X$ (under 4.30, of course).

c) *If D is dense in $\mathbb{R}_+$, then $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$, $[\delta_{3,1}\text{-}D]$ are necessary and sufficient for having* $X^n \xrightarrow{\mathscr{L}(D)} X$ (under 4.30).

Of course, this theorem is just as unsatisfactory as Theorem 4.4, except in the case when D is dense, or when $E(\exp iu \cdot X_t)$ does not vanish. If X has no fixed time of discontinuity, 4.30 is equivalent to 2.53, so Theorem 2.52a is a particular case of the above.

Proof. Let $J^n = \{t\colon \nu^n(\{t\} \times \mathbb{R}^d) > 0\}$ be the set of fixed times of discontinuity of X^n, and call ν^{nc} the measure $\nu^{nc}(ds, dx) = \nu^n(ds, dx) 1_{(J^n)^c}(s)$.

Let $s, t \in D \cup \{0\}$ with $s < t$. Let $(s^n_k)_{1 \le k \le K^n}$ be an enumeration of the points of $J^n \cap (s, t]$, and set

$$\chi^n_k = \begin{cases} \Delta X^n_{s^n_k} & \text{if } k \le K^n \\ 0 & \text{otherwise.} \end{cases}$$

Now, II.5.2 yields that the law $\mu^n_k = \mathscr{L}(\chi^n_k)$ is given by II.4.17, i.e.:

4.33 $$\mu^n_k(dx) = \nu^n(\{s^n_k\} \times dx) + [1 - \nu^n(\{s^n_k\} \times \mathbb{R}^d)]\varepsilon_0(dx) \quad \text{for } k \le K^n;$$

hence if $b^n_k = E[h(\chi^n_k)]$ we have $b^n_k = \Delta B^n_{s^n_k}$ for $k \le K^n$ by II.5.5v. Secondly, the process $X'^n = X^n - B^n$ is a PII-semimartingale by II.5.28, and so is the process $1_{J^n} \cdot X'^n = \sum_{s \le \cdot} 1_{J^n}(s) \Delta X'^n_s$. Since $\Delta X'^n_{s^n_k} = \chi^n_k - b^n_k$, we deduce from II.3.11 that the sequence $Y^n_k = \chi^n_k - b^n_k$ satisfies 2.31. Thirdly, if we set $\zeta^n = X^n_t - X^n_s - \sum_{k \le K^n} Y^n_k$, the characteristic function of ζ^n is the exponential part in the formula II.4.16 written for X^n: therefore $\mathscr{L}(\zeta^n)$ is infinitely divisible, with the following characteristics:

4.34 $$b^n = B^n_t - B^n_s, \quad c^n = C^n_t - C^n_s, \quad F^n(dx) = \nu^{nc}((s, t] \times dx)$$

and of course ζ^n is independent of $(\chi^n_k)_{k \ge 1}$.

Finally for each $\varepsilon > 0$ there are indices such that $t^\infty_{m(\varepsilon)}(\varepsilon) \le s < t^\infty_{m(\varepsilon)+1}(\varepsilon)$ and $t^\infty_{m(\varepsilon)+p(\varepsilon)}(\varepsilon) \le t < t^\infty_{m(\varepsilon)+p(\varepsilon)+1}(\varepsilon)$ (with the convention $t^\infty_0(\varepsilon) = 0$). For $1 \le j \le p(\varepsilon)$ we set

$$k^n_j(\varepsilon) = k \quad \text{if } t^n_{m(\varepsilon)+j}(\varepsilon) = s^n_k$$

and we define $k_j^n(\varepsilon)$ arbitrarily (but different from the other $k_i^n(\varepsilon)$'s) if $t_{m(\varepsilon)+j}^n(\varepsilon) \notin J^n \cap (s,t]$. Then 4.30i implies that for all n large enough, we have $t_{m(\varepsilon)+j}^n(\varepsilon) = s_{k_j^n(\varepsilon)}^n$ for all $j \leq p(\varepsilon)$ such that $t_{m(\varepsilon)+j}^\infty(\varepsilon) \in J^\infty$. Then 4.33 and 4.30 imply that the array (χ_k^n) satisfies 4.2, with the $k_j^n(\varepsilon)$ defined above.

Define V^n and $V^n(\varepsilon)$ by 4.3. Then

$$V^n = X_t^n - X_s^n, \quad V^n(\varepsilon) = X^n(\varepsilon)_t - X^n(\varepsilon)_s + \sum_{1 \leq j \leq p(\varepsilon)} b_{k_j^n(\varepsilon)}^n.$$

Since 4.2 holds, we have $\sum_{1 \leq j \leq p(\varepsilon)} b_{k_j^n(\varepsilon)}^n \to \sum_{1 \leq j \leq p(\varepsilon)} b_{k_j^\infty(\varepsilon)}^\infty$, and we deduce the equivalence:

4.35
$$\begin{cases} V^n \xrightarrow{\mathscr{L}} V^\infty \quad \Leftrightarrow \quad X_t^n - X_s^n \xrightarrow{\mathscr{L}} X_t^\infty - X_s^\infty \\ V^n(\varepsilon) \xrightarrow{\mathscr{L}} V^\infty(\varepsilon) \Leftrightarrow X^n(\varepsilon)_t - X^n(\varepsilon)_s \xrightarrow{\mathscr{L}} X^\infty(\varepsilon)_t - X^\infty(\varepsilon)_s. \end{cases}$$

Finally, an immediate computation, based upon 3.1, 4.33, 4.34, gives

$$\tilde{c}^{n,jl} + \sum_k \{E[h^j h^l(\chi_k^n)] - E[h^j(\chi_k^n)]E[h^l(\chi_k^n)]\} = \tilde{C}_t^{n,jl} - \tilde{C}_s^{n,jl}$$

$$F^n(g) + \sum_k E[g(\chi_k^n)] = g * v_t^n - g * v_s^n.$$

Henceforth, conditions $[\beta_4]$, $[\gamma_4]$, $[\delta_{4,i}]$ of 4.4 are exactly the following:

$[\beta_4]$ $\quad B_t^n - B_s^n \to B_t^\infty - B_s^\infty$,

$[\gamma_4]$ $\quad \tilde{C}_t^n - \tilde{C}_s^n \to \tilde{C}_t^\infty - \tilde{C}_s^\infty$,

$[\delta_{4,i}]$ $\quad g * v_t^n - g * v_s^n \to g * v_t^\infty - g * v_s^\infty$ for all $g \in C_i(\mathbb{R}^d)$.

Then, we deduce parts (a) and (b) of the theorem from 4.35, and Theorem 4.4, and Lemma 1.3.

It remains to prove part (c). We suppose that D is dense in $\mathbb{R}_+$. We will prove that if $s, t \in D \cup \{0\}$, $s < t$, we have the implication:

4.36
$$X^n \xrightarrow{\mathscr{L}(D)} X^\infty \Rightarrow X^n(\varepsilon)_t - X^n(\varepsilon)_s \xrightarrow{\mathscr{L}} X^\infty(\varepsilon)_t - X^\infty(\varepsilon)_s.$$

Indeed, let $\varphi^{n,s,t}(u) = E[\exp iu \cdot (X_t^n - X_s^n)]$ and $\varphi_\varepsilon^{n,s,t}(u) = E[\exp iu \cdot (X^n(\varepsilon)_t - X^n(\varepsilon)_s)]$. Then $\varphi^{n,s,t}(u) = \varphi_\varepsilon^{n,s,t}(u)\hat{\varphi}_\varepsilon^{n,s,t}(u)$, where

$$\hat{\varphi}_\varepsilon^{n,s,t}(u) = \prod_{k: s < t_k^n(\varepsilon) \leq t} E(\exp iu \cdot \Delta X_{t_k^n(\varepsilon)}^n)$$

and 4.30 implies that $\hat{\varphi}_\varepsilon^{n,s,t}(u) \to \hat{\varphi}_\varepsilon^{\infty,s,t}(u)$. Hence if $\varphi^{n,s,t}(u) \to \varphi^{\infty,s,t}(u) \neq 0$ we deduce that $\varphi_\varepsilon^{n,s,t}(u) \to \varphi_\varepsilon^{\infty,s,t}(u)$.

Now, suppose that $X^n \xrightarrow{\mathscr{L}(D)} X^\infty$ and let $s, t \in D \cup \{0\}$ with $s < t$. Let $u \in \mathbb{R}^d$. There is a finite subdivision $s = t_0 < t_1 < \cdots < t_q = t$ with $t_j \in D$ and $\varphi^{n,t_j,t_{j+1}}(u) \neq 0$, and the hypothesis implies that $\varphi^{n,t_j,t_{j+1}}(u) \to \varphi^{\infty,t_j,t_{j+1}}(u)$. From what precedes we deduce that $\varphi_\varepsilon^{n,s,t}(u) = \prod_{0 \leq j \leq q-1} \varphi_\varepsilon^{n,t_j,t_{j+1}}(u) \to \varphi_\varepsilon^{\infty,s,t}(u)$. Therefore we have proved 4.36.

But Lemma 1.3 and 4.36 imply that if $X^n \xrightarrow{\mathscr{L}(D)} X$, we also have $X^n(\varepsilon) \xrightarrow{\mathscr{L}(D)} X(\varepsilon)$: then (c) follows from (a). □

In the introduction to this section, we asserted that condition 4.30 which we impose on the X^n's was motivated by the conditions of Lemma 3.42. The following makes this assertion more precise.

4.37 **Lemma.** *Let D be a dense subset of $\mathbb{R}_+$, contained in $\mathbb{R}_+ \backslash J(X) = \{t: P(\Delta X_t \neq 0) = 0\}$. Then under $[\delta_{3,2}\text{-}D]$ the two conditions $[\text{Sk-}\delta_{3,2}]$ (cf. 3.3) and 4.30 are equivalent.*

Proof. a) Suppose first $[\text{Sk-}\delta_{3,2}]$. Let $\varepsilon > 0$, and call $t_1^\infty < \cdots < t_j^\infty < \cdots$ the successive times where $\nu^\infty(\{t_j^\infty\} \times \{|x| > \varepsilon/2\}) \geq \varepsilon$. For each j, call (t_j^n) the sequence associated to t_j^∞ in Lemma 3.42. Since $t_j^n \to t_j^\infty$ we can always assume that $t_j^n < t_{j+1}^n$ if $t_j^n < \infty$, with $\lim_j \uparrow t_j^n = \infty$ (by modifying if necessary the values of t_j^n for a finite number of n's). Then 4.30i comes from 3.42i and from the convergence $t_j^n \to t_j^\infty$ and the fact that $t_j^\infty \notin D$.

Let $g \in C_2(\mathbb{R}^d)$ with $0 \leq g \leq 1$, and $g(x) = 0$ for $|x| \leq \varepsilon/2$, and $g(x) = 1$ for $|x| \geq \varepsilon$. Put $\alpha_n = g * \nu^n$, and $\alpha'_n(t) = \alpha_n(t) - \sum_{t_j^n \leq t} \Delta\alpha_n(t_j^n)$. Since $\alpha_n \to \alpha_\infty$ in $\mathbb{D}(\mathbb{R})$ by $[\text{Sk-}\delta_{3,2}]$ we also have $\alpha'_n \to \alpha'_\infty$ in $\mathbb{D}(\mathbb{R})$ from Proposition VI.2.1, and Lemma VI.2.5 yields that $\limsup_n \sup_{s \leq t} \Delta\alpha'_n(s) \leq \sup_{s \leq t} \Delta\alpha'_\infty(s)$, which is smaller than ε by construction, for all $t \in D$. Then 4.30ii follows from

$$\sup_{s \leq t, s \neq t_j^n} \nu^n(\{s\} \times \{|x| > \varepsilon\}) \leq \sup_{s \leq t, s \neq t_j^n} \nu^n(\{s\} \times g)$$
$$\leq \sup_{s \leq t} \Delta\alpha'_n(s).$$

b) Conversely, suppose that we have 4.30 and $[\delta_{3,2}\text{-}D]$. Let $g \in C_2(\mathbb{R}^d)$ with $g \geq 0$, and set $\alpha_n = g * \nu^n$. If $t \in D(X)$ it is easy to find a sequence $t_n \to t$ with $t_n \leq t$ and $\Delta\alpha_n(t_n) = 0$. If $t \in J(X)$ there is $\varepsilon > 0$, $j \geq 1$ with $t = t_j^\infty(\varepsilon)$. Then $t_n = t_j^n(\varepsilon)$ satisfies $t_n \to t$ because D is dense (apply 4.30i), and $\Delta\alpha_n(t_n) \to \Delta\alpha_\infty(t)$ by 4.30ii. Hence the sequence (α_n) of increasing functions satisfies VI.2.16 by $[\delta_{3,2}\text{-}D]$ and VI.2.20 because $D \subset \mathbb{R}_+ \backslash J(X)$: so by VI.2.15 and VI.2.22, $\alpha_n \to \alpha_\infty$ in $\mathbb{D}(\mathbb{R})$: thus $[\text{Sk-}\delta_{3,2}]$ holds. □

4.38 **Corollary.** *Suppose that $[\text{Sk-}\delta_{3,1}]$ holds, and let $D = \mathbb{R}_+ \backslash J(X)$. Then $X^n \xrightarrow{\mathscr{L}(D)} X$ if and only if $[\beta_3\text{-}D]$ and $[\gamma_3\text{-}D]$ hold.*

4.39 **Remark.** This corollary allows for another "direct" proof of the sufficient condition in Theorem 3.12: just apply Theorem 3.49 and identify the limit points of the sequence $(\mathscr{L}(X^n))$ by this corollary. □

§ 4c. Another Necessary and Sufficient Condition for Functional Convergence

In this subsection we turn back to the problem examined in Section 3 and give another condition for $X^n \xrightarrow{\mathscr{L}} X$. This condition will be expressed in terms of the characteristic functions, and the proof is based upon Theorem 4.32.

Here again, for each $n \in \bar{\mathbb{N}}^*$, X^n is a PII and $X = X^\infty$. We set

4.40 $$g^n(u)_t = E(\exp iu \cdot X_t^n)$$

4.41 $$g_r^n(u)_t = \begin{cases} E(\exp iu \cdot (X_t^n - X_r^n)) & \text{if } r < t \\ 1 & \text{if } r \geq t \end{cases}$$

and in the next theorem, for each q-uple $\mathbf{u} = (u_1, \ldots, u_q) \in (\mathbb{R}^d)^q$ we denote by $g_r^n(\mathbf{u})$ the $\mathbb{C}^q$-valued function whose j-th (complex) component is the function $t \rightsquigarrow g_r^n(u_j)_t$. We also write $g_r(u)_t$ for $g_r^\infty(u)_t$.

4.42 Theorem. a) *Let $D = \mathbb{R}_+ \backslash J(X) = \{t : P(\Delta X_t \neq 0) = 0\}$. In order that $X^n \xrightarrow{\mathscr{L}} X$ it is necessary and sufficient that $g_r^n(\mathbf{u}) \to g_r(\mathbf{u})$ for the Skorokhod topology in $\mathbb{D}(\mathbb{C}^q)$ for all $\mathbf{u} = (u_1, \ldots, u_q) \in (\mathbb{R}^d)^q$ and all $r \in D$.*

b) *If $g(u)_t \neq 0$ for all $t \geq 0$, $u \in \mathbb{R}^d$, then in order that $X^n \xrightarrow{\mathscr{L}} X$ it is necessary and sufficient that $g^n(\mathbf{u}) \to g(\mathbf{u})$ for the Skorokhod topology in $\mathbb{D}(\mathbb{C}^q)$ for all $\mathbf{u} = (u_1, \ldots, u_q) \in (\mathbb{R}^d)^q$.*

The second statement is not surprising: the functions $g^n(u)$ characterize $\mathscr{L}(X^n)$ if they never vanish; when they do vanish we need the functions $g_r^n(u)$ to characterize $\mathscr{L}(X^n)$. Note also that (b) is a trivial consequence of (a): suppose indeed that $g(u)_t \neq 0$ for all $t \geq 0$, $u \in \mathbb{R}^d$, and that $g^n(\mathbf{u}) \to g(\mathbf{u})$; then if $r \in D$, $\Delta g(\mathbf{u})_r = 0$ and hence $g^n(\mathbf{u})_r \to g(\mathbf{u})_r$; now, we have

$$g_r^n(u_j)_t = g^n(u_j)_{t \vee r} / g^n(u_j)_r \quad \text{if } g^n(u_j)_r \neq 0$$

and it is immediate to deduce that $g_r^n(\mathbf{u}) \to g_r(\mathbf{u})$ for the Skorokhod topology.

4.43 Corollary. *Suppose that X has no fixed time of discontinuity. Then $X^n \xrightarrow{\mathscr{L}} X$ if and only if for each $u \in \mathbb{R}^d$ we have: $g^n(u) \to g(u)$ uniformly over finite intervals. If it is the case, we also have:*

4.44 $$\sup_{s \leq t, |u| \leq \theta} |g^n(u)_s - g(u)_s| \to 0 \quad \text{for all } t \geq 0, \theta > 0.$$

Proof of the corollary. Each function $g(u)$ is continuous and non-vanishing. Hence the first part comes from 4.42b, because in virtue of Proposition VI.1.17, $g^n(\mathbf{u}) \to g(\mathbf{u})$ for the Skorokhod topology if and only if $g^n(u_j) \to g(u_j)$ uniformly over finite intervals for $1 \leq j \leq q$.

It remains to prove that $X^n \xrightarrow{\mathscr{L}} X$ implies 4.44. If it were not the case, there would exist a subsequence (n_k), and a sequence $s_k \to s$, and a sequence $u_k \to u$, and $\varepsilon > 0$, such that

4.45 $$|g^{n_k}(u_k)_{s_k} - g(u_k)_{s_k}| \geq \varepsilon \quad \text{for all } k.$$

Now, if $\alpha_k \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ and if $\Delta\alpha(s) = 0$, then $\alpha_k(s_k) \to \alpha(s)$; we easily deduce from this fact and from the hypotheses that $X_{s_k}^{n_k} \xrightarrow{\mathscr{L}} X_s$ so $g^{n_k}(\cdot)_{s_k} \to g(\cdot)_s$ uniformly over compact sets. Thus $|g^{n_k}(u_k)_{s_k} - g(u_k)_s| \to 0$. We also have $g(u_k)_s \to g(u)_s$ and therefore we obtain a contradiction with 4.45. $\square$

Proof of Theorem 4.42. (i) *Necessary condition.* We can notice first that the necessary condition has already been proved just above when $D = \mathbb{R}_+$.

Let us examine the general case. Let $\mathbf{u} = (u_1, \ldots, u_q)$, and Y^n be the $\mathbb{C}^q$-valued process whose components are $Y^{n,j} = \exp iu_j \cdot X^n$, and $\alpha_n(t) = E(Y_t^n) = g^n(\mathbf{u})_t$. We have $\alpha_n(t) \to \alpha_\infty(t)$ for all $t \in D$ and we will prove that the sequence (α_n) is relatively compact in $\mathbb{D}(\mathbb{C}^q)$, thus implying that $g^n(\mathbf{u}) \to g(\mathbf{u})$ in $\mathbb{D}(\mathbb{C}^q)$.

To this end, we reproduce the proof of Proposition 3.18. Since $|\alpha_n(t)| \le q$, it suffices to prove 3.22. Let $N > 0$, $\varepsilon > 0$ be fixed, and $\theta = 1$. We have 3.24 and 3.25 (both with $\theta = 1$) and if $t_j^n \le s < t < t_{j+1}^n$ and $\omega \notin G_j^n \cup F^n$ (with the notation of 3.18) we have $|X_t^n - X_s^n| \le 3\varepsilon$; moreover there is a constant C such that $|e^{iu_j \cdot x} - 1| \le C|x|$ for all $j \le q$. Then for $t_j^n \le s < t < t_{j+1}^n$ we obtain

$$|Y_t^n - Y_s^n| \le 3C\varepsilon + 2q 1_{G_j^n \cup F^n}$$

$$|\alpha_n(t) - \alpha_n(s)| \le 3C\varepsilon + 2q(2\varepsilon + \varepsilon) = \varepsilon(3C + 6q).$$

Hence $w_N'(\alpha_n, \delta) \le \varepsilon(3C + 6q)$ for $n \ge n_0$, and we get 3.22.

Therefore $g^n(\mathbf{u}) \to g(\mathbf{u})$ in $\mathbb{D}(\mathbb{C}^q)$. That $g_r^n(\mathbf{u}) \to g_r(\mathbf{u})$ is proved similarly, when $r \in D$: for example, we can use the fact that $X_t'^n = X_{t \vee r}^n - X_r^n$ is a PII, and that $X'^n \xrightarrow{\mathcal{L}} X'$ because $r \in D$.

(ii) *Sufficient condition.* We suppose now that $g_r^n(\mathbf{u}) \to g_r(\mathbf{u})$ in $\mathbb{D}(\mathbb{C}^q)$ for all $r \in D$, $\mathbf{u} = (u_1, \ldots, u_q) \in (\mathbb{R}^d)^q$. The proof requires several steps.

Step 1: we have 4.30 with $D = D(X)$. Firstly, let $t > 0$ be fixed. We call η_t^n the distribution of ΔX_t^n, and $\hat{\eta}_t^n$ its characteristic function, so

4.46 $$g_r^n(u)_t = g_r^n(u)_{t-} \hat{\eta}_t^n(u) \quad \text{if } r < t.$$

Let $\mathbf{u} = (u_1, \ldots, u_q)$ be fixed. There exists $r \in D$ with $r < t$ and $g_r(u_j)_{t-} \neq 0$ for all $j \le q$ (take r close enough to t). Then by VI.2.1 we deduce from the convergence $g_r^n(\mathbf{u}) \to g_r(\mathbf{u})$:

4.47 if $\hat{\eta}_t(u_j) = 1$ for all $j \le q$, then $\hat{\eta}_{s_n}^n(u_j) \to 1$ for every sequence (s_n) converging to t (because $g_r(\mathbf{u})$ is continuous at t);

4.48 if $\hat{\eta}_t(u_j) \neq 1$ for at least one $j \le q$, there exists a sequence $t_n(\mathbf{u}) \to t$ with $\hat{\eta}_{t_n(\mathbf{u})}^n(u_j) \to \hat{\eta}_t(u_j)$ for all $j \le q$; moreover, if $s_n \to t$ and $s_n > t_n(\mathbf{u})$ (resp. $s_n < t_n(\mathbf{u})$) for all n, then $\hat{\eta}_{s_n}^n(u_j) \to 1$ for all $j \le q$;

(Note that the time r has disappeared in 4.47 and 4.48).

We deduce from 4.47:

4.49 $$t \in D \Rightarrow \eta_{s_n}^n \to \varepsilon_0 \text{ weakly if } s_n \to t.$$

Suppose now that $t \notin D$, and let $u \in \mathbb{R}^d$ with $\hat{\eta}_t^\infty(u) \neq 1$. Call $t_n = t_n(\mathbf{u})$ the sequence associated to $\mathbf{u} = \{u\}$ in 4.48. Let $\mathbf{v} = (v_1, \ldots, v_q)$ and $\mathbf{w} = \mathbf{v} \cup \{u\}$, and suppose that there is at least one $j \le q$ such that $\hat{\eta}_t(v_j) \neq 1$. Then by 4.48 we obtain that $\hat{\eta}_{t_n(\mathbf{w})}^n(v_j)$ and $\hat{\eta}_{t_n(\mathbf{v})}^n(v_j)$ both converge to $\hat{\eta}_t(v_j)$, which implies that $t_n(\mathbf{w}) = t_n(\mathbf{v})$

for all n large enough. Similarly $t_n(\mathbf{w}) = t_n$ $(= t_n(\mathbf{u}))$ for all n large enough, so $t_n(\mathbf{v}) = t_n$ as well for all n large enough. In other words, we have proved:

4.50 if $t \notin D$ there is a sequence (t_n) converging to t, such that $\eta^n_{t_n} \to \eta_t$ weakly, and such that if $s_n \to t$ and $s_n > t_n$ (resp. $s_n < t_n$) for all n, then $\eta^n_{s_n} \to \varepsilon_0$ weakly.

Presently, we shall deduce 4.30 from 4.49 and 4.50, after noticing that

4.51 $$\eta^n_t(dx) = v^n(\{t\} \times dx) + [1 - v^n(\{t\} \times \mathbb{R}^d)]\varepsilon_0(dx).$$

Let $t^\infty_1(\varepsilon) < \cdots < t^\infty_j(\varepsilon) < \cdots$ be the successive times where $\eta^n_t(|x| > \varepsilon) \geq \varepsilon$. By 4.50 we associate to each $t^\infty_j(\varepsilon)$ a sequence $t^n_j(\varepsilon) \to t^\infty_j(\varepsilon)$: up to modifying $t^n_j(\varepsilon)$ for a finite number of n's, we can suppose that the sequence $(t^n_j(\varepsilon))_{j \geq 1}$ is strictly increasing and converges to $+\infty$, without altering 4.50. Then 4.30i is deduced from the property $t^\infty_j(\varepsilon) \notin D$, from $t^n_j(\varepsilon) \to t^\infty_j(\varepsilon)$, from 4.50 and from 4.51.

Suppose that 4.30ii is wrong. Then there is a sequence $n_k \uparrow \infty$, a sequence $s_k \to t$ and $\eta > 0$ with $\eta^{n_k}_{s_k}(|x| > \varepsilon) \geq \varepsilon + \eta$ for all k, and $s_k \neq t^{n_k}_j(\varepsilon)$ for all $j \geq 1$. If $t \in D$ this clearly contradicts 4.49. If $t \notin D$, let (t_n) be the sequence associated to t in 4.50; but then, $\eta^{n_k}_{s_k}(|x| > \varepsilon) \geq \varepsilon + \eta$ and 4.50 clearly imply that $s_k = t_{n_k}$ for all k large enough, in which case $\eta^{n_k}_{s_k} \to \eta_t$ weakly; if $\eta_t(|x| > \varepsilon) < \varepsilon$ we arrive to a trivial contradiction, while if $\eta_t(|x| > \varepsilon) \geq \varepsilon$, we have $t = t^\infty_j(\varepsilon)$ for some j, and so $s_k = t^{n_k}_j(\varepsilon)$ for k large, a property that was excluded. Therefore we always have a contradiction, which proves that 4.30ii is satisfied.

Step 2: We have $g^n_r(u)_t \to g_r(u)_t$ for all $r, t \in D$. Hence by Lemma 1.3,

4.52 $$X^n \xrightarrow{\mathscr{L}(D)} X.$$

Therefore, taking a uniformly continuous truncation function h, Step 1 and Theorem 4.32c yield the conditions $[\beta_3\text{-}D]$, $[\gamma_3\text{-}D]$ and $[\delta_{3,2}\text{-}D]$. Moreover, by Lemma 4.37, we have $[\text{Sk-}\delta_{3,2}]$.

Step 3: Set $X'^n = X^n - B^n$. Since we have $[\beta_3\text{-}D]$, we deduce from 4.52 that

4.53 $$X'^n \xrightarrow{\mathscr{L}(D)} X'.$$

Then II.5.28 gives the characteristics (B'^n, C'^n, v'^n) of the PII X'^n, which is also a semimartingale, and in particular

4.54 $$B'^n_t = \sum_{s \leq t} \left\{ \int v^n(\{s\} \times dx) h(x - \Delta B^n_s) + [1 - v^n(\{s\} \times \mathbb{R}^d)] h(-\Delta B^n_s) \right\}.$$

We will prove that the sequence (X'^n) satisfies 4.30. Let A satisfy 2.1, and $\varepsilon > 0$. Let $\varepsilon' = \varepsilon/2(1 + A)$, and set, with the notation of Step 1 and of 4.30:

$$\begin{cases} J^n_{\varepsilon'} = \{t^n_j(\varepsilon') : j \geq 1\} \\ M^n_t(\varepsilon') = \sup_{s \geq t, s \notin J^n_{\varepsilon'}} |\Delta B^n_s| \\ \eta'^n_t = \mathscr{L}(\Delta X'^n_t). \end{cases}$$

If $\eta_t^n(|x| > \varepsilon') \leq \varepsilon'$ we have $|\Delta B_t^n| \leq \varepsilon'(1 + A)$ because of 2.1. Then we deduce from 4.30 (for X^n) that for all $t < \infty$,

4.55 $$\limsup_n M_t^n(\varepsilon') \leq \varepsilon'(1 + A).$$

Moreover if $|\Delta B_t^n| \leq \varepsilon'(1 + 2A)$, and since $\Delta X_t'^n = \Delta X_t^n - \Delta B_t^n$, and $\varepsilon = \varepsilon' + \varepsilon'(1 + 2A)$, we deduce that $\eta_t'^n(|x| > \varepsilon) \leq \eta_t^n(|x| > \varepsilon')$. So 4.55 and 4.30ii (for X^n) yield

4.56 $$\limsup_n \sup_{s \leq t, s \notin J_{\varepsilon'}^n} \eta_s'^n(|x| > \varepsilon) \leq \varepsilon' \leq \varepsilon.$$

We also have $\eta'^n_{t_j^n(\varepsilon')} \to \eta'_{t_j^\infty(\varepsilon')}$ weakly, by 4.30i. Therefore, using the relationship 4.51 between ν'^n and $\eta_t'^n$, we obtain that the sequence (X'^n) satisfies 4.30 with the times $t_j'^n(\varepsilon) = t_j^n(\varepsilon')$.

Step 4: We have $X'^n \xrightarrow{\mathscr{L}} X'$.

Exactly like in Step 2, we deduce from 4.30 for (X'^n) and from 4.53 that the conditions [β_3'-D], [γ_3'-D], [$\delta_{3,2}'$-D] are met, where the "$'$" means that we use the characteristics (B'^n, C'^n, ν'^n), and Lemma 4.37 yields [Sk-$\delta_{3,2}'$].

Consider 4.54: B'^n is purely discontinuous, and

4.57 $$\Delta B_t'^n = E[h(\Delta X_t^n - \Delta B_t^n)] = E[h(\Delta X_t^n - \Delta B_t^n) - h(\Delta X_t^n) + \Delta B_t^n].$$

Let $\varepsilon > 0$, and $\eta > 0$ such that $|x - y| \leq \eta \Rightarrow |h(x) - h(y)| \leq \varepsilon$. Hence if $\delta = \eta \wedge \dfrac{1}{2A} \wedge \varepsilon$ and if $|\Delta B_t^n| \leq \delta$, we have $|h(\Delta X_t^n - \Delta B_t^n) - h(\Delta X_t^n) + \Delta B_t^n| \leq 2\varepsilon$, and this expression is equal to 0 if in addition $|\Delta X_t^n| \leq 1/2A$. Hence 4.57 implies:

$$|\Delta B_t^n| \leq \delta \Rightarrow |\Delta B_t'^n| \leq 2\varepsilon P\left(|\Delta X_t^n| > \frac{1}{2A}\right).$$

Let $\varepsilon' = \delta/2(1+A)$. From 4.55 we deduce that $M_t^n(\varepsilon') \leq \delta$ for all n large enough, while 4.30 (or 4.50) imply that $\Delta B'^n_{t_j^n(\varepsilon')} \to \Delta B'^\infty_{t_j^\infty(\varepsilon')}$. Thus if $t \in D$,

4.58 $$\begin{aligned} \limsup_n \sup_{s \leq t} |B_s'^n - B_s'^\infty| &\leq \lim_n \sum_{j: t_j^\infty(\varepsilon') \leq t} |\Delta B'^n_{t_j^n(\varepsilon)} - \Delta B'^\infty_{t_j^\infty(\varepsilon')}| \\ &\quad + \limsup_n \left\{ \sum_{s \leq t, s \notin J_{\varepsilon'}^n} |\Delta B_s'^n| + \sum_{s \leq t, s \notin J_{\varepsilon'}^\infty} |\Delta B_s'^\infty| \right\} \\ &\leq 2\varepsilon \limsup_n \left\{ \sum_{s \leq t} P\left(|\Delta X_t^n| > \frac{1}{2A}\right) \right. \\ &\quad \left. + \sum_{s \leq t} P\left(|\Delta X_s^\infty| > \frac{1}{2A}\right) \right\}. \end{aligned}$$

Let $g \in C_2(\mathbb{R}^d)$ with $0 \leq g \leq 1$ and $g(x) = 1$ for $|x| \geq 1/2A$. Then [Sk-$\delta_{3,2}$] and 4.58 yield

$$\limsup_n \sup_{s \leq t} |B_s'^n - B_s'^\infty| \leq 4\varepsilon g * \nu_t^\infty.$$

Since $\varepsilon > 0$ is arbitrary, we deduce that $B'^n \to B'^\infty$ uniformly over finite intervals. So we have [Sk-β'_3] (and even [Sup-β'_3]). Then the claim follows from Theorem 3.13.

Step 5: Define $g'^n(u)_t = E[\exp iu \cdot X'^n_t]$. If $t > 0$, call (t_n) the sequence defined in 4.50 (or $t_n = t$ if $t \in D$); then by construction we have $\Delta g^n(u)_{t_n} \to \Delta g(u)_t$; moreover $\eta^n_{t_n} \to \eta_t$ weakly and $\Delta B^n_{t_n} \to \Delta B^\infty_t$, so $\eta'^n_{t_n} \to \eta'_t$ weakly as well. But by Step 4 and by the necessary condition, $g'^n(u) \to g'(u)$ in $\mathbb{D}(\mathbb{C})$, so 4.50 allows to associate with t another sequence $t'_n \to t$ with $\Delta g'^n(u)_{t'_n} \to \Delta g'(u)_t$ and $\eta'^n_{t'_n} \to \eta'_t$ weakly, and $t'_n = t$ if $t \in D$. The uniqueness in 4.50 implies that $t'_n = t_n$ for all n large enough. So finally

4.59 $\qquad$ there is $t_n \to t$ with $\Delta g^n(u)_{t_n} \to \Delta g(u)_t$ and $\Delta g'^n(u)_{t_n} \to \Delta g'(u)_t$.

Apply Proposition VI.2.2 to α_n and β_n, which are the functions in $\mathbb{D}(\mathbb{C}^2)$ defined by: $\alpha^1_n = g^n(u)$, $\alpha^2_n = 0$, and $\beta^1_n = 0$, $\beta^2_n = g'^n(u)$. It follows from 4.59 that $(g^n(u), g'^n(u)) \to (g^\infty(u), g'^\infty(u))$ in $\mathbb{D}(\mathbb{C}^2)$. Hence if f_u is a function: $\mathbb{C}^2 \to \mathbb{R}^d$ that is continuous and bounded and satisfies $\exp iu \cdot f_u(x, y) = \dfrac{x}{y}$ for $|y| \geq 1/2$ and $\left|\dfrac{x}{y}\right| = 1$, we also have:

4.60 $$\gamma^n_u := f_u(g^n(u), g'^n(u)) \to \gamma^\infty_u := f_u(g^\infty(u), g'^\infty(u)) \quad \text{in } \mathbb{D}(\mathbb{R}^d).$$

But by definition of X'^n we have $g^n(u)_t = g'^n(u)_t \exp iu \cdot B^n_t$. Therefore $B^n_t = \gamma^n_u(t)$ for all t such that $|g'^n(u)_t| \geq 1/2$. For each $N \in D$ there exists $u \in \mathbb{R}^d$ such that $|g'(u)_N| \geq 3/4$ and so $|g'^n(u)_N| \geq 1/2$ for all n large enough because $g'^n(u)_N \to g'(u)_N$. Since the function $t \rightsquigarrow |g'^n(u)_t|$ decreases, it follows that $B^n = \gamma^n_u$ on $[0, N]$ for all n large enough. Then we deduce from 4.60 that $B^n \to B^\infty$ in $\mathbb{D}(\mathbb{R}^d)$.

Hence we have [Sk-β_3], and we have seen in Step 2 that [Sk-$\delta_{3,2}$] and [γ_3-D] were fulfilled. Then Theorem 3.13 implies that $X^n \xrightarrow{\mathscr{L}} X$, and the proof is finished. $\square$

5. The Central Limit Theorem

The purpose of this section is two-fold: firstly we specialize the previous results to the case where the limit is Gaussian: this gives the very easy § 5a below. Secondly we expound a "non-classical" theorem concerning sums of non-infinitesimal variables: it is a version of Theorem 4.4, with condition 4.2 weakened, but when the limit is Gaussian. Finally we give a "functional" version of this result.

§ 5a. The Lindeberg-Feller Theorem

5.1 **Definition.** A rowwise independent array (χ_k^n) satisfies *Lindeberg condition* if for all $\varepsilon > 0$ we have

$$\lim_n \sum_k E(|\chi_k^n|^2 1_{\{|\chi_k^n| > \varepsilon\}}) = 0. \qquad \square$$

Of course, this implies $\sum_k E(|\chi_k^n|^2) < \infty$, provided we have 2.31.

5.2 **Theorem.** *We suppose that the d-dimensional rowwise independent array satisfies 2.31 and Lindeberg condition, and let $\xi^n = \sum_k \chi_k^n$. Then*

a) *If $\mathscr{L}(\xi^n) \to \mu$ weakly, then μ is a Gaussian measure on $\mathbb{R}^d$;*

b) *in order that $\mathscr{L}(\xi^n) \to \mathscr{N}(b, c)$, the Gaussian measure with mean b and covariance matrix c, it is necessary and sufficient that the following two conditions hold:*

$$[\beta_2''] \quad \sum_k E(\chi_k^n) \to b$$

$$[\gamma_2''] \quad \sum_k \left[E(\chi_k^{n,j} \chi_n^{n,l}) - E(\chi_k^{n,j}) E(\chi_n^{n,l})\right] \to c^{jl}.$$

Proof. Lindeberg condition obviously implies infinitesimality for the array (χ_k^n), and also condition $[\delta_{2,2}]$ with $F^n = 0$ and $F = 0$. Then Theorem 2.35a applied with $\zeta^n = 0$ yields that if $\mathscr{L}(\xi^n) \to \mu$, then μ is infinitely divisible with characteristics $(b, c, 0)$, that is $\mu = \mathscr{N}(b, c)$, and we have (a).

Moreover, if $\mu = \mathscr{N}(b, c)$ we have $b' = b$ and $\tilde{c}' = c$ with the notation of 2.36, so $[\beta_2''] = [\beta_2']$ and $[\gamma_2''] = [\gamma_2']$, and 2.37 is obviously implied by Lindeberg condition. Therefore (b) is deduced from Theorem 2.36. $\square$

5.3 **Remark.** Consider the *usual central limit theorem*: $(Y_n)_{n \geq 1}$ is an i.i.d. sequence of random variables with $E(Y_n) = 0$ and $E(Y_n^j Y_n^l) = c^{jl}$. Set $\xi^n = \frac{1}{\sqrt{N}} \sum_{1 \leq p \leq n} Y_p$. Then $\mathscr{L}(\xi^n) \to \mathscr{N}(0, c)$ and this is a particular case of the previous theorem: take $\chi_k^n = Y_k/\sqrt{n}$ (resp. $= 0$) if $k \leq n$ (resp. $k > n$), so that:

$$\sum_k E(\chi_k^n) = 0$$

$$\sum_k \left[E(\chi_k^{n,j} \chi_n^{n,l}) - E(\chi_k^{n,j}) E(\chi_n^{n,l})\right] = nE(Y_1^j Y_1^l / n) = c^{jl}$$

$$\sum_k E(|\chi_k^n|^2 1_{\{|\chi_k^n| > \varepsilon\}}) = nE(|Y_1/\sqrt{n}|^2 1_{\{|Y_1/\sqrt{n}| > \varepsilon\}})$$

$$= E(|Y_1|^2 1_{\{|Y_1| > \varepsilon\sqrt{n}\}}) \to 0. \qquad \square$$

Now we turn to the functional version of this theorem. We suppose that X is a continuous d-dimensional PII, whose characteristics (B, C, ν) necessarily satisfy $\nu = 0$. We also consider a sequence X^n of PII-semimartingales with characteristics (B^n, C^n, ν^n) relative to some truncation function h. We suppose that each ν^n satisfies 2.60, and we define B'^n by $\tilde{C}'^n$ by 2.61 and 2.62 (note that ν also meets 2.60! and $B' = B$ and $\tilde{C}' = \tilde{C} = C$).

5.4 **Theorem.** *Together with the above assumptions and notation, we suppose that*

5.5 $$(|x|^2 1_{\{|x|>\varepsilon\}}) * \nu_t^n \to 0 \quad \text{for all } \varepsilon > 0, t \in D$$

where $D \subset \mathbb{R}_+$. Then

a) $X^n \xrightarrow{\mathscr{L}(D)} X$ *if and only if the following two conditions hold:*

$[\beta_3'\text{-}D]$ $\quad B_t'^n \to B_t \quad$ for all $t \in D$;

$[\gamma_3'\text{-}D]$ $\quad \tilde{C}_t'^n \to C_t \quad$ for all $t \in D$.

b) $X^n \xrightarrow{\mathscr{L}} X$ *if and only if*

$[\text{Sup-}\beta_3']$ $\quad \sup_{s \leq t} |B_s'^n - B_s| \to 0 \quad$ for all $t \geq 0$,

and $[\gamma_3'\text{-}D]$ hold for some dense subset $D \subset \mathbb{R}_+$.

Proof. Note that 5.5 implies 2.64 and $[\delta_{3,2}]$ with $\nu = 0$. Then (a) and (b) follow respectively from Theorems 2.63 and 3.7. □

§ 5b. Zolotarev's Type Theorems

We consider again the problem, already studied in § 4a, of the convergence of sums of triangular arrays which are not infinitesimal. We will replace Condition 4.2 by a condition that is in some sense weaker; on the other hand, the limit will always by Gaussian.

To be simple, we start with a 1-*dimensional* rowwise independent array (χ_k^n) *which satisfies 2.31.*

5.6 *Notation.* We denote by ϕ_Δ the Gaussian measure $\mathscr{N}(0, \Delta)$ (variance $= \Delta$) on $\mathbb{R}$, and by η_k^n the law $\mathscr{L}(\chi_k^n)$; we denote by $\hat{\phi}_\Delta$ and $\hat{\eta}_k^n$ their repartition functions: $\hat{\phi}_\Delta(x) = \phi_\Delta((-\infty, x])$ and $\hat{\eta}_k^n(x) = \eta_k^n((-\infty, x])$. □

We also consider two functions:

5.7
- a truncation function $h \in \mathscr{C}_t^1$ of class C^1, with compact support and $h(x) = -h(-x)$;
- a bounded even function f of class $C^1 : \mathbb{R} \to \mathbb{R}_+$ with bounded first derivative, increasing on $\mathbb{R}_+$ and $f(x) = x^2$ on a neighborhood of 0

We call Δ_k^n the (unique) nonnegative number characterized by

5.8 $$E[f(\chi_k^n)] = \int f(x) \phi_{\Delta_k^n}(dx) \qquad (= \phi_{\Delta_k^n}(f)).$$

5.9 **Theorem.** *Let (χ_k^n) be a 1-dimensional rowwise independent array satisfying 2.31, and (f,h) with 5.7, and Δ_k^n defined by 5.8. Let ζ^n be a variable that is independent from $(\chi_k^n)_{k\geq 1}$ and infinitely divisible with characteristics (b^n, c^n, F^n), and define $\tilde{c}^n$ by 2.5. Suppose that*

$[A_1]$ $F^n(|x| > \varepsilon) \to 0$ for all $\varepsilon > 0$;

$[B_1]$ $|b^n| + \sum_k |E[h(\chi_k^n)]| \to 0$;

$[C_1]$ $\tilde{c}^n + \sum_k \Delta_k^n \to c$;

$[D_1]$ $\sum_k \int_{|x|\geq\varepsilon} |\hat{\eta}_k^n(x) - \hat{\phi}_{\Delta_k^n}(x)|\, dx \to 0$ for all $\varepsilon > 0$.

Then if $\xi^n = \sum_k \chi_k^n$, we have $\mathscr{L}(\xi^n + \zeta^n) \to \mathcal{N}(0,c)$ weakly.

This theorem should be compared to Corollary 4.6: $[B_1]$ is stronger than $[\tilde{\beta}_4]$, and $[C_1]$ is "similar" to $[\gamma_4]$; but $[A_1] + [D_1]$ is considerably weaker than $[\delta_{4,2}]$ plus 4.2.

Proof. Let $\varphi^n(u)$ be the characteristic function of $\xi^n + \zeta^n$: we wish to prove that $\varphi^n(u) \to \exp(-u^2c/2)$. Fix $u \in \mathbb{R}$ and set $\delta_k^n = \eta_k^n(e^{iux})$ and $\tilde{\delta}_k^n = \phi_{\Delta_k^n}(e^{iux}) = \exp(-u^2\Delta_k^n/2)$. We have

$$\varphi^n(u) = \exp \psi_{b^n,c^n,F^n}(u) \left\{\prod_k \delta_k^n\right\}.$$

Set $\alpha^n = \tilde{c}^n + \sum_k \Delta_k^n$ and $\beta^n = F^n(e^{iux} - 1 - iuh(x) + \frac{u^2}{2}h(x)^2)$. Then

5.10 $$I^n := \exp \psi_{b^n,c^n,F^n}(u) \left\{\prod_k \tilde{\delta}_k^n\right\} = \exp\left(iub^n - \frac{u^2}{2}\alpha^n + \beta^n\right).$$

There is a constant C such that

5.11 $$\left|e^{iux} - 1 - iuh(x) + \frac{u^2}{2}h(x)^2\right| \leq C(|x|^3 \wedge 1)$$

so if A satisfies 2.1 and if $\varepsilon \leq 1/A$,

$$\begin{aligned} |\beta^n| \leq CF^n(|x|^3 \wedge 1) &\leq C[F^n(|x| > \varepsilon) + \varepsilon F^n(|x|^2 1_{\{|x|\leq\varepsilon\}})] \\ &\leq C[F^n(|x| > \varepsilon) + \varepsilon F^n(h^2)] \end{aligned}$$

5.12 $$\leq C[F^n(|x| > \varepsilon) + \varepsilon\alpha^n].$$

From $[A_1]$ and $[C_1]$, $\limsup_n |\beta^n| \leq C\varepsilon c$ for all $\varepsilon > 0$, so $\beta^n \to 0$. By $[B_1]$, $b^n \to 0$. Thus by 5.10, $I^n \to \exp(-u^2c/2)$. So it remains to prove that $J^n := \varphi^n(u) - I^n \to 0$.

We have $J^n = \exp \psi_{b^n,c^n,F^n}(u) \{\prod_k \delta_k^n - \prod_k \tilde{\delta}_k^n\}$, and the modulus of $\psi_{b^n,c^n,F^n}(u)$ is smaller than 1, so it remains to prove that $\tilde{J}^n := \prod_k \delta_k^n - \prod_k \tilde{\delta}_k^n \to 0$. Since $|\delta_k^n| \leq 1$ and $|\tilde{\delta}_k^n| \leq 1$, we obviously have

$$\tilde{J}^n \leq \sum_k |\delta_k^n - \tilde{\delta}_k^n|$$

(proved by induction on the number of terms in the two products). Hence if $\mu_k^n = \eta_k^n - \phi_{\Delta_k^n}$ and $\hat{\mu}_k^n(x) = \mu_k^n((-\infty, x])$ we have

$$|\tilde{J}^n| \leq \sum_k |\mu_k^n(e^{iux})| = \sum_k \left| \mu_k^n\left(e^{iux} - 1 - iuh(x) + \frac{u^2}{2} f(x)\right) + iu\mu_k^n(h) \right|$$

because $\mu_k^n(1) = 0$, and $\mu_k^n(f) = 0$ by 5.8). Let $\gamma_k^n = \mu_k^n\left(e^{iux} - 1 - iuh(x) + \frac{u^2}{2} f(x)\right)$. Since h is odd, $\phi_{\Delta_k^n}(h) = 0$, and

$$|\tilde{J}^n| \leq \sum_k |\gamma_k^n| + \sum_k |u|\,|E[h(\chi_k^n)]|.$$

Hence by $[B_1]$ it remains to prove that $\sum_k |\gamma_k^n| \to 0$.

Since the function $\hat{\mu}_k^n(x)$ has limit 0 when $x \to \pm\infty$, by integration by parts we obtain for every bounded function g with bounded derivative g'

5.13
$$\int g(x) \mu_k^n(dx) = -\int \hat{\mu}_k^n(x) g'(x)\, dx.$$

Applying this to $g(x) = e^{iux} - 1 - iuh(x) + \frac{u^2}{2} f(x)$ yields

$$\gamma_k^n = -iu \int \left[e^{iux} - h'(x) - \frac{iu}{2} f'(x) \right] \hat{\mu}_k^n(x)\, dx = -iu(\gamma_k'^n(\delta) + \gamma_k''^n(\delta)),$$

with

$$\begin{cases} \gamma_k'^n(\delta) = \displaystyle\int_{|x| \leq \delta} \left[e^{iux} - h'(x) - \frac{iu}{2} f'(x) \right] \hat{\mu}_k^n(x)\, dx \\ \gamma_k''^n(\delta) = \displaystyle\int_{|x| > \delta} \left[e^{iux} - h'(x) - \frac{iu}{2} f'(x) \right] \hat{\mu}_k^n(x)\, dx \end{cases}$$

and where $\delta > 0$ is such that $h(x) = x$ and $f(x) = x^2$ for $|x| \leq \delta$. We have

5.14
$$|\gamma_k''^n(\delta)| \leq C' \int_{|x| > \delta} |\hat{\mu}_k^n(x)|\, dx$$

for some constant C', and $\hat{\mu}_k^n(x) = \hat{\eta}_k^n(x) - \hat{\phi}_{\Delta_k^n}(x)$, so $[D_1]$ implies

5.15
$$\sum_k |\gamma_k''^n(\delta)| \to 0 \quad \text{for all } \delta > 0.$$

We also have for $|x| \leq \delta$:

$$\left| e^{iux} - h'(x) - \frac{iu}{2} f'(x) \right| = |e^{iux} - 1 - iux| \leq \theta |x|^2$$

for some constant θ, not depending on δ; hence

5.16 $$|\gamma_k'^n(\delta)| \le \delta\theta \int_{|x|\le\delta} |x|\,|\hat{\mu}_k^n(x)|\,dx$$

A new integration by parts yields, because $f'(x) = 2x$ for $|x| \le \delta$:

$$2\int_{-\delta}^{0} |x|\,|\hat{\mu}_k^n(x)|\,dx \le -\int_{-\infty}^{0} f'(x)[\hat{\eta}_k^n(x) + \hat{\phi}_{\Delta_k^n}(x)]\,dx$$
$$= \int_{-\infty}^{0} f(x)[\eta_k^n(dx) + \phi_{\Delta_k^n}(dx)]$$
$$2\int_{0}^{\delta} |x|\,|\hat{\mu}_k^n(x)|\,dx \le \int_{0}^{\infty} f'(x)[(1 - \hat{\eta}_k^n(x)) + (1 - \hat{\phi}_{\Delta_k^n}(x))]\,dx$$
$$= \int_{0}^{\infty} f(x)[\eta_k^n(dx) + \phi_{\Delta_k^n}(dx)].$$

Hence we deduce from 5.16 and 5.8 that

$$|\gamma_k'^n(\delta)| \le \frac{\delta\theta}{2}\int_{-\infty}^{\infty} f(x)[\eta_k^n(dx) + \phi_{\Delta_k^n}(dx)] = \delta\theta\phi_{\Delta_k^n}(f) \le \delta\theta\theta'\Delta_k^n$$

where θ' is a constant such that $f(x) \le \theta' x^2$. Therefore

5.17 $$\sum_k |\gamma_k'^n(\delta)| \le \delta\theta\theta'\alpha^n$$

Since $\delta > 0$ is arbitrarily small, we deduce from 5.15 and 5.17 and $[C_1]$ that $\sum_k |\gamma_k^n| \to 0$. □

Now we state the version of Theorem 5.9 that is suited to the square-integrable case.

5.18 **Theorem.** *In the situation of 5.9, we suppose in addition that $F^n(x^2) < \infty$ and $E[(\chi_k^n)^2] < \infty$ for all n, k. We define b'^n and $\tilde{c}'^n$ by 2.12 and we set*

5.19 $$\Delta_k'^n = E[(\chi_k^n)^2].$$

Then if

$[A_1']$ $F^n(x^2 1_{\{|x|>\varepsilon\}}) \to 0$ for all $\varepsilon > 0$

$[B_1']$ $|b'^n| + \sum_k |E(\chi_k^n)| \to 0$

$[C_1']$ $\tilde{c}'^n + \sum_k \Delta_k'^n \to c$

$[D_1']$ $\sum_k \int_{|x|>\varepsilon} |x|\,|\hat{\eta}_k^n(x) - \hat{\phi}_{\Delta_k'^n}(x)|\,dx \to 0$ for all $\varepsilon > 0$,

we have $\mathscr{L}(\xi^n + \zeta^n) \to \mathscr{N}(0, c)$ weakly.

Proof. Note that $\Delta_k'^n$ satisfies

5.20
$$E[(\chi_k^n)^2] = \int x^2 \phi_{\Delta_k'^n}(dx).$$

Then we can reproduce the proof of 5.9, with the functions $f(x) = x^2$ and $h(x) = x$: thus 5.20 is the same than 5.8, and b'^n and $\tilde{c}'^n$ are the characteristics of $\mathscr{L}(\zeta^n)$ associated to the "truncation" function $h(x) = x$. The only changes that are required are the following ones:

a) use $\Delta_k'^n$, b'^n, $\tilde{c}'^n$ and $\mu_k'^n = \eta_k^n - \phi_{\Delta_k'^n}$ in place of Δ_k^n, b^n, $\tilde{c}^n$, μ_k^n.

b) 5.11 should be replaced by $\left|e^{iux} - 1 - iux + \dfrac{u^2x^2}{2}\right| \le C(|x|^3 \wedge x^2)$, so 5.12 becomes $|\beta^n| \le C[F^n(x^2 1_{\{|x|>\varepsilon\}}) + \varepsilon\alpha'^n]$, where $\alpha'^n = \tilde{c}'^n + \sum_k \Delta_k'^n$; then by $[A_1']$ and $[C_1']$ we conclude that $\beta^n \to 0$.

c) Since x^2 is integrable with respect to η_k^n and $\phi_{\Delta_k'^n}$, the formula 5.13 remains valid when g is not necessarily bounded, but such that $g(x)/(x^2+1)$ is bounded, which is the case of $g(x) = e^{iux} - 1 - iux + \dfrac{u^2x^2}{2}$.

d) 5.14 should be replaced by $|\gamma_k''^n(\delta)| \le C' \int_{|x|>\delta} |x|\,|\hat{\mu}_k'^n(x)|\,dx$, and $[D_1']$ still implies 5.15. □

§ 5c. Finite-Dimensional Convergence of PII's to a Gaussian Martingale

Here again we suppose that the pair (f, h) satisfies 5.7. For each $n \in \mathbb{N}^*$, let X^n be a 1-*dimensional PII-semimartingale*, with characteristics (B^n, C^n, ν^n) relative to h. Let $J^n = \{s: \nu^n(\{s\} \times \mathbb{R}) > 0\}$ be the set of fixed times of discontinuity of X^n. Define the numbers $\Delta_s^n \ge 0$ by

5.21
$$\nu^n(\{s\} \times f) = \int f(x)\phi_{\Delta_s^n}(dx) \quad (= \phi_{\Delta_s^n}(f))$$

(note that $\Delta_s^n = 0$ if $s \notin J^n$, where we use the convention $\phi_0 = \varepsilon_0$). Let ν^{nc} be the measure

$$\nu^{nc}(dt, dx) = \nu^n(dt, dx) 1_{(J^n)^c}(t)$$

and set

$$B_t^{nc} = B_t^n - \sum_{s \le t} \Delta B_s^n.$$

On the other hand, let X be a 1-dimensional Gaussian martingale with $X_0 = 0$, and set

5.22
$$\tilde{C}_t' = E(X_t^2).$$

We have seen in § II.4d that this is equivalent to saying that X is a PII with characteristics (B, C, ν) (relative to h) given by

5.23
$$\begin{cases} B_t = \sum_{s\le t} \phi_{\Delta\tilde{C}'_s}(h) = 0 \quad \text{(because } h \text{ is odd)}, \\ C_t = \tilde{C}'_t - \sum_{s\le t} \Delta\tilde{C}'_s \\ \nu(ds,dx) = \sum_{s>0,\Delta\tilde{C}'_s>0} \varepsilon_s(dt)\otimes\phi_{\Delta\tilde{C}'_s}(dx) \end{cases}$$

(with the notation of § II.4d we have $\tilde{C}'_t = \tilde{c}(t)$ and $\phi_{\Delta\tilde{C}'_s} = K_s$ if $\Delta\tilde{C}'_s > 0$; we use the notation $\tilde{C}'$, because it is the function associated to (B, C, ν) by 2.62).

5.24 **Theorem.** *With the above assumptions and notation, and if $D \subset \mathbb{R}_+$, the following conditions are sufficient for having $X^n \xrightarrow{\mathscr{L}(D)} X$:*

$[A_2\text{-}D]$ $\nu^{nc}([0,t]\times\{|x|>\varepsilon\}) \to 0$ for all $\varepsilon > 0$, $t \in D$;

$[B_2\text{-}D]$ $|B_t^{nc}| + \sum_{s\le t} |\Delta B_s^n| \to 0$ for all $t \in D$;

$[C_2\text{-}D]$ $C_t^n + h^2 * \nu_t^{nc} + \sum_{s\le t} \Delta_s^n \to \tilde{C}'_t$ for all $t \in D$;

$[D_2\text{-}D]$ $\sum_{s\le t} \int_{|x|>\varepsilon} |\hat{\eta}_s^n(x) - \hat{\phi}_{\Delta_s^n}(x)|\,dx \to 0$ for all $\varepsilon > 0$, $t \in D$,

where $\hat{\eta}_s^n(x) = P(\Delta X_s^n \le x)$.

Note that in $[D_2\text{-}D]$ all summands corresponding to $s \notin J^n$ are 0.

Proof. Let $s, t \in D \cup \{0\}$ with $s < t$. Let K^n be the number of points in $J^n \cap [s,t]$, and $(s_k^n)_{1\le k\le K^n}$ be an enumeration of these points. Set

$$b^n = B_t^{nc} - B_s^{nc}, \quad c^n = C_t^n - C_s^n, \quad F^n(dx) = \nu^{nc}([0,t]\times dx)$$

$$\chi_k^n = \begin{cases} \Delta X_{s_k^n}^n & \text{if } 1 \le k \le K^n \\ 0 & \text{otherwise} \end{cases} \qquad \tilde{\Delta}_k^n = \begin{cases} \Delta_{s_k^n}^n & \text{if } 1 \le k \le K^n \\ 0 & \text{otherwise} \end{cases}$$

$$\xi^n = \sum_k \chi_k^n, \quad \zeta^n = X_t^n - X_s^n - \xi^n.$$

We have seen in the proof of 2.52 that $\mathscr{L}(\zeta^n)$ is infinitely divisible with characteristics (b^n, c^n, F^n), and ξ^n is independent from $(\chi_k^n)_{k\ge1}$; moreover we have 2.55 if $\eta_s^n = \mathscr{L}(\Delta X_s^n)$, so 5.8 holds with $\tilde{\Delta}_k^n$ by 5.21, while $\Delta B_{s_k^n}^n = E[h(\chi_k^n)]$ for $k \le K^n$. Finally, since X^n is a semimartingale, the array (χ_k^n) satisfies 2.31.

Therefore $[A_2\text{-}D] \Rightarrow [A_1]$, and $[B_2\text{-}D] \Rightarrow [B_1]$, and $[C_2\text{-}D] \Rightarrow \{[C_1]$ with $\tilde{\Delta}_k^n$ and $c = \tilde{C}'_t - \tilde{C}'_s\}$, and $[D_2\text{-}D] \Rightarrow \{[D_1]$ with $\tilde{\Delta}_k^n$ and $\hat{\eta}_k^n(x) = P(\chi_k^n \le x)\}$. Hence Theorem 5.9 yields

$$X_t^n - X_s^n \xrightarrow{\mathscr{L}} X_t - X_s$$

because $\mathscr{L}(X_t - X_s) = \phi_{\tilde{C}'_t - \tilde{C}'_s}$. We conclude by Lemma 1.3. □

If we use 5.18 instead of 5.9, we obtain the square-integrable version:

5.25 **Theorem.** *In the situation of 5.24, we suppose in addition that 2.60 holds for all $n \in \mathbb{N}^*$, and we define B'^n by 2.61, and $B_t'^{nc} = B_t'^n - \sum_{s \leq t} \Delta B_s'^n$ and $\Delta_s'^n = \nu^n(\{s\} \times x^2)$. Then for having $X^n \xrightarrow{\mathscr{L}(D)} X$ it suffices to have*:

$$[A_2'\text{-}D] \quad (x^2 1_{\{|x|>\varepsilon\}}) * \nu_t^{nc} \to 0 \quad \text{for all } \varepsilon > 0, t \in D;$$

$$[B_2'\text{-}D] \quad |B_t'^{nc}| + \sum_{s \leq t} |\Delta B_s'^n| \to 0 \quad \text{for all } t \in D;$$

$$[C_2'\text{-}D] \quad C_t^n + x^2 * \nu_t^{nc} + \sum_{s \leq t} \Delta_s'^n \to \tilde{C}_t' \quad \text{for all } t \in D;$$

$$[D_2'\text{-}D] \quad \sum_{s \leq t} \int_{|x|>\varepsilon} |x|\,|\hat{\eta}_s^n(x) - \hat{\phi}_{\Delta_s'^n}(x)|\,dx \to 0 \quad \text{for all } \varepsilon > 0, t \in D.$$

§ 5d. Functional Convergence of PII's to a Gaussian Martingale

The situation is the same than in the previous subsection: the pair (f, h) satisfies 5.7; X^n is a 1-dimensional PII-semimartingale with characteristics (B^n, C^n, ν^n) and we define Δ_s^n, ν^{nc}, B^{nc} as in § 5c.

X is a Gaussian martingale with $X_0 = 0$ and $\tilde{C}'$ is defined by 5.22.

5.26 **Theorem.** *Together with the above assumptions and notation, we suppose that D is a dense subset of $\mathbb{R}_+$ and that we have conditions $[A_2\text{-}D]$, $[B_2\text{-}D]$, $[C_2\text{-}D]$, $[D_2\text{-}D]$ and*

$$[\tilde{C}_2\text{-}D] \quad \sum_{s \leq t} (\Delta_s^n)^2 \to \sum_{s \leq t} (\Delta \tilde{C}_s')^2 \quad \text{for all } t \in D.$$

Then $X^n \xrightarrow{\mathscr{L}} X$.

In view of Theorem 5.24, it would be enough to show that the above conditions imply tightness for the sequence (X^n). Instead of doing so, we shall use Theorem 3.13 and give a proof that does not use the previous parts of this section.

5.27 **Remark.** *If* $\alpha^n = C^n + h^2 * \nu^{nc} + \sum_{s \leq \cdot} \Delta_s^n$, we have $\Delta\alpha^n(s) = \Delta_s^n$; hence VI.2.15 implies that

$$[C_2\text{-}D] + [\tilde{C}_2\text{-}D] \Leftrightarrow [\text{Sk-}C_2]\text{: } C^n + h^2 * \nu^{nc} + \sum_{s \leq \cdot} \Delta_s^n \to \tilde{C}' \text{ in } \mathbb{D}(R). \qquad \square$$

Proof. a) $[B_2\text{-}D] \Rightarrow [\text{Sk-}\beta_3]$ is obvious (we even have $[\text{Sup-}\beta_3]$).

b) We will now prove $[\text{Sk-}\delta_{3,1}]$. In 2.7 we can choose the set $C_1(\mathbb{R}^d)$ so that it contains a subset $C_1'(\mathbb{R}^d)$ consisting in functions that are boundedly differentiable, and which itself is convergence-determining (in the sense of 2.7), and also that for every $g \in C_1(\mathbb{R}^d)$ there is $\tilde{g} \in C_1'(\mathbb{R}^d)$ such that $0 \leq g \leq \tilde{g}$. Let $g \in C_1'(\mathbb{R}^d)$ and $\eta_s^n = \mathscr{L}(\Delta X_s^n)$. Then by 2.55,

$$g * \nu_t^n = g * \nu_t^{nc} + \sum_{s \le t} \eta_s^n(g)$$

5.28 $$= g * \nu_t^{nc} + \sum_{s \le t} \int g(x)[\eta_s^n(dx) - \phi_{\Delta_s^n}(dx)] + \sum_{s \le t} \phi_{\Delta_s^n}(g).$$

Let $\theta > 0$ be such that $g(x) = 0$ for $|x| \le \theta$; and $C = \sup |g'|$. Integrating by parts (like in 5.13) yields

$$\left| \int g(x)[\eta_s^n(dx) - \phi_{\Delta_s^n}(dx)] \right| = \left| \int_{-\infty}^{\infty} g'(x)[\hat{\eta}_s^n(x) - \hat{\phi}_{\Delta_s^n}(x)] \right| dx$$

$$\le C \int_{|x| \ge \theta} |\hat{\eta}_s^n(x) - \hat{\phi}_{\Delta_s^n}(x)|\, dx.$$

Hence $[A_2\text{-}D]$ and $[D_2\text{-}D]$ imply that for $t \in D$:

5.29 $$\sup_{s \le t} \left| g * \nu_s^{nc} + \sum_{r \le s} \int g(x)[\eta_r^n(dx) - \phi_{\Delta_r^n}(dx)] \right| \to 0.$$

Let $\alpha^n = C^n + h^2 * \nu^{nc} + \sum_{s \le \cdot} \Delta_s^n$ and $\beta^n = \sum_{s \le \cdot} \phi_{\Delta_s^n}(g)$, and $\beta = \sum_{s \le \cdot} \phi_{\Delta \tilde{C}_s'}(g) = g * \nu$ (the last equality coming from 5.23). For each $s \ge 0$, there exists a sequence $t_n(s) \to s$, with $t_n(s) \le s$ if $s \in D$, and with $\Delta_{t_n(s)}^n \to \Delta \tilde{C}_s'$ (apply VI.2.22 to the sequence (α^n), which converges to $\tilde{C}'$ in $\mathbb{D}(\mathbb{R})$ by $[\text{Sk-}C_2]$: see 5.27). Then $\phi_{\Delta_{t_n(s)}^n}(g) \to \phi_{\Delta \tilde{C}_s'}(g)$, that is:

5.30 the sequence (β^n) satisfies VI.2.20 with the limit β.

Let $\varepsilon > 0$ with $\Delta \tilde{C}_s' \neq \varepsilon$ for all $s > 0$, and call $s_1 < \cdots < s_j < \cdots$ the successive times when $\Delta \tilde{C}_s' > \varepsilon$. Let $K = \{s_j : j \ge 1\}$ and $K^n = \{t_n(s_j) : j \ge 1\}$. We have $\phi_{\Delta_{t_n(j)}^n}(g) \to \phi_{\Delta \tilde{C}_{s_j}'}(g)$, so if $t \in D$:

5.31 $$\limsup_n |\beta^n(t) - \beta(t)| \le \limsup_n \left\{ \sum_{s \notin K^n, s \le t} \phi_{\Delta_s^n}(g) + \sum_{s \notin K, s \le t} \phi_{\Delta \tilde{C}_s'}(g) \right\}.$$

But if $C = \sup |g|$ and $\delta = \int |x|^3 \phi_1(dx)$,

5.32 $$\phi_\Delta(g) \le C \int_{|x| \ge \theta} \phi_\Delta(dx) \le C\theta^{-3} \Delta^{3/2} \delta,$$

while VI.2.7 implies that for $t \in D$,

$$\limsup_n \sup_{s \notin K^n, s \le t} \Delta_s^n = \limsup_n \sup_{s \notin K^n, s \le t} \Delta \alpha^n(s)$$

$$\le \sup_{s \notin K, s \le t} \Delta \tilde{C}_s' \le \varepsilon.$$

Hence 5.31 and 5.32 yield

$$\limsup_n |\beta^n(t) - \beta(t)| \le C\theta^{-3}\delta \left[\sqrt{\varepsilon} \limsup_n \sum_{s \notin K^n, s \le t} \Delta_s^n + \sqrt{\varepsilon} \sum_{s \notin K, s \le t} \Delta \tilde{C}_s' \right]$$

5.33 $$\le C\theta^{-3}\delta\sqrt{\varepsilon} \left[\limsup_n \alpha^n(t) + \tilde{C}_t' \right] \le 2C\theta^{-3}\delta\tilde{C}_t'\sqrt{\varepsilon}.$$

Now, $\varepsilon > 0$ is arbitrary, so 5.33 implies that $\beta^n(t) \to \beta(t)$. This, together with 5.30, implies (in virtue of VI.2.22 and VI.2.15) that $\beta^n \to \beta$ in $\mathbb{D}(\mathbb{R})$. Thus we deduce from 5.23, 5.28 and 5.29 that

$$g * v^n \to g * v \quad \text{in } \mathbb{D}(\mathbb{R}).$$

In particular, $g * v_t^n \to g * v_t$ for all $t \in D$, $g \in C_1'(\mathbb{R}^d)$. Then $g * v_t^n \to g * v_t$ also for all $g \in C_1(\mathbb{R}^d)$, because of the convergence-determining property of $C_1'(\mathbb{R}^d)$. Moreover, if $g \in C_1(\mathbb{R}^d)$ and $\tilde{g} \in C_1'(\mathbb{R}^d)$ with $g \leq \tilde{g}$, then $g * v^n \prec \tilde{g} * v^n$ (strong majoration). Then, from VI.3.35, the sequence $\{g * v^n\}$ is relatively compact in $\mathbb{D}(\mathbb{R})$. Hence $g * v^n \to g * v$ in $\mathbb{D}(\mathbb{R})$, and [Sk-$\delta_{3,1}$] holds.

c) Finally, in view of 3.13, it remains to prove [γ_3-D]. For that, we could use Lemma 4.37, which in view of (b) implies 4.30. Moreover, Theorem 5.24 implies $X^n \xrightarrow{\mathcal{L}(D)} X$, so the necessary part of Theorem 4.32 gives [γ_3-D].

However, we shall give a direct proof, which does not use the (rather difficult) results of Section 4. Firstly, note that

$$\tilde{C}_t^n = C_t^n + h^2 * v_t^{nc} + \sum_{s \leq t} [\eta_s^n(h^2) - (\Delta B_s^n)^2].$$

By [B_2-D] we have $\sum_{s \leq t} (\Delta B_s^n)^2 \to 0$. So, due to [$C_2$-$D$], the only thing to prove is that

$$t \in D \Rightarrow \sum_{s \leq t} [\eta_s^n(h^2) - \Delta_s^n] \to \tilde{C}_t - \tilde{C}_t' = \sum_{s \leq t} [\phi_{\Delta \tilde{C}_s'}(h^2) - \Delta \tilde{C}_s'].$$

We have $h^2 - f \in C_2(\mathbb{R}^d)$, so [$\delta_{3,2}$-$D$] (implied by [$\delta_{3,1}$-$D$]) gives

$$t \in D \Rightarrow \sum_{s \leq t} [\eta_s^n(h^2 - f)] \to (h^2 - f) * v_t = \sum_{s \leq t} \phi_{\Delta \tilde{C}_s'}(h^2 - f),$$

so it remains to prove that

$$t \in D \Rightarrow \sum_{s \leq t} [\eta_s^n(f) - \Delta_s^n] \to \sum_{s \leq t} [\phi_{\Delta \tilde{C}_s'}(f) - \Delta \tilde{C}_s'].$$

By 5.21 we have $\eta_s^n(f) = \phi_{\Delta_s^n}(f)$, so we are left to prove that

5.34 $$t \in D \Rightarrow \gamma^n(t) := \sum_{s \leq t} [\phi_{\Delta_s^n}(f) - \Delta_s^n] \to \gamma(t) := \sum_{s \leq t} [\phi_{\Delta \tilde{C}_s'}(f) - \Delta \tilde{C}_s'].$$

With the same notation than in (b), we have $\Delta_{t_n(s_j)}^n \to \Delta \tilde{C}_{s_j}'$. Hence

$$\limsup_n |\gamma^n(t) - \gamma(t)| \leq \limsup_n \sum_{s \notin K^n, s \leq t} |\phi_{\Delta_s^n}(f) - \Delta_s^n| + \sum_{s \notin K, s \leq t} |\phi_{\Delta \tilde{C}_s'}(f) - \Delta \tilde{C}_s'|.$$

Let $\theta > 0$ such that $f(x) = x^2$ for $|x| \leq \theta$, and $C = \sup |f|$, and $\delta = \int |x|^3 \phi_1(dx)$ and $\delta' = \int |x|^5 \phi_1(dx)$. We have

$$|\phi_\Delta(f) - \Delta| = \left| \int_{|x| \geq 0} (f(x) - x^2) \phi_\Delta(dx) \right| \leq C\theta^{-3} \Delta^{3/2} \delta + \theta^{-5} \Delta^{5/2} \delta'$$

Then the same argument than for obtaining 5.33 yields

$$\limsup_n |\gamma^n(t) - \gamma(t)| \leq 2(C\theta^{-3}\delta + \theta^{-5}\delta')\tilde{C}'_t\sqrt{\varepsilon}.$$

Since $\varepsilon > 0$ is arbitrary, we obtain 5.34. Hence we have [γ_3-D]. It remains to apply Theorem 3.13 to obtain the result. □

5.35 **Remark.** One could prove that for D dense in $\mathbb{R}_+$,

$$[B_2\text{-}D] \Rightarrow \{[A_2\text{-}D] + [C_2\text{-}D] + [D_2\text{-}D] + [\tilde{C}_2\text{-}D] \Leftrightarrow [\gamma_3\text{-}D] + [\text{Sk-}\delta_{3,1}]\}.$$

We have proved the implication $\Rightarrow$. For the reverse implication, one could use Lemma 4.37. □

Finally, the square-integrable version of 5.26 is the following, which we give without proof (see [163]).

5.36 **Theorem.** *In the situation of 5.26, we suppose in addition that 2.60 holds. We define B'^n by 2.61, and $B_t'^{nc} = B_t'^n - \sum_{s\leq t}\Delta B_s'^n$, and $\Delta_s'^n = \nu^n(\{s\} \times x^2)$. Then if D is a dense subset of $\mathbb{R}_+$ and if [A'_2-D], [B'_2-D], [C'_2-D], [D'_2-D] and*

$$[\tilde{C}'_2\text{-}D] \quad \sum_{s\leq t} (\Delta_s'^n)^2 \to \sum_{s\leq t} (\Delta\tilde{C}'_s)^2 \quad \text{for all } t \in D,$$

hold, we have $X^n \xrightarrow{\mathscr{L}} X$.

Chapter VIII. Convergence to a Process with Independent Increments

This chapter constitutes the second step on our way to general limit theorems. We consider a sequence (X^n) of semimartingales, with characteristics (B^n, C^n, ν^n), and a limiting process X which is a PII with characteristics (B, C, ν). Our main objective is to prove that the various conditions of Chapter VII still insure the (functional or finite-dimensional) convergence of (X^n) to X, although the X^n's are no longer PII.

Section 1 contains the key theorem, which allows to reduce all theorems of this chapter to the corresponding results of the previous chapter.

Sections 2 and 4, and parts of Section 3, are reformulations of the main results of Chapter VII in our more general setting. The method, explained in the proof of Theorem 2.4, is always the same; hence these sections are undoubtedly tedious, but relatively easy to read (so we think!). However, we decided to write many variants of the same basic theorem, because of their usefulness for applications: they certainly are more useful than when all X^n's are PII.

In Section 3 we also give some applications: firstly we notice that the conditions (based upon the characteristics) which insure $X^n \xrightarrow{\mathscr{L}} X$ badly fail to be necessary in general, when the X^n's are not PII. However, they are (almost) necessary when X is a continuous PII (i.e., $\nu = 0$) and we examine several variants of the "necessary" conditions in this case. Secondly, we give an application to the convergence of normed sums of i.i.d. semimartingales. Thirdly, we give some examples of central limit theorems for additive functionals of Markov processes and finally central limit theorems for stationary processes (under mixing, or slightly more general, conditions).

Section 5 is devoted to two closely related kinds of results: one is the extension of the previous limit theorems to the case where X is a conditional PII (or, "mixture" of PII's). The second one is the "stable convergence": we recall the definition of this stronger sort of convergence and give some applications.

1. Finite-Dimensional Convergence, a General Theorem

§ 1a. Description of the Setting for This Chapter

For each integer n we consider a *d-dimensional semimartingale* $X^n = (X^{n,i})_{i \le d}$ defined on a stochastic basis $\mathscr{B}^n = ((\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$; for simplicity, we assume

1.1 $$X_0^n = 0$$

(this is by no way essential!). We fix a *continuous* truncation function $h \in \mathscr{C}_t^d$: recall that there is a constant $A \geq 1$ such that

1.2 $$|x| \leq \frac{1}{A} \Rightarrow h(x) = x, \quad |h| \leq A.$$

We denote by (B^n, C^n, ν^n) the *characteristics* of X^n, relative to h, and by $\tilde{C}^n$ its modified second characteristic, defined by

1.3 $$\tilde{C}_t^{n,ij} = C_t^{n,ij} + (h^i h^j) * \nu_t^n - \sum_{s \leq t} \Delta B_s^{n,i} \Delta B_s^{n,j}.$$

The *limiting process* X will always be a *d-dimensional PII* $X = (X^i)_{i \leq d}$ defined on a stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$. We denote by (B, C, ν) its characteristics relative to h (they are deterministic): see II.4.15 if X is a semimartingale, or II.5.2 if it is not. According to II.5.7, its modified second characteristic is

1.4 $$\tilde{C}_t^{ij} = C_t^{ij} + (h^i - \Delta B^i)(h^j - \Delta B^j) * \nu_t + \sum_{s \leq t} [1 - \nu(\{s\} \times \mathbb{R}^d)] \Delta B_s^i \Delta B_s^j$$

(which reduces to 1.3 when X is a semimartingale).

The expectation with respect to P^n is denoted by $E^n(\cdot)$; as a rule, all random elements defined on Ω^n will present the superscript "n"; for example, if we write "the sequence (Y^n)" it is understood that each Y^n is defined on Ω^n. If the Y^n's take their values in a metric space (E, δ) and if Y is a (non-random) element of E we write indifferently $Y^n \xrightarrow{P} Y$ or $Y^n \xrightarrow{\mathscr{L}} Y$ for

1.5 $$P^n(\delta(Y^n, Y) > \varepsilon) \underset{(n)}{\rightarrow} 0 \quad \text{for all } \varepsilon > 0.$$

1.6 **Remark.** Exactly like in the previous chapter, we could assume without loss of generality (by taking the tensor product of all stochastic bases) that all X^n and X are defined on the same basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$. We shall refer to this situation as to "*hypothesis 1.6*". But in the statement of the results we usually prefer to keep the mention of the index n. □

§ 1b. The Basic Theorem

In order to state the next theorem, upon which all the present chapter is based, we need some more notation. Firstly, we set

1.7 $$g(u)_t = E(\exp iu \cdot X_t), \quad t \geq 0, u \in \mathbb{R}^d.$$

Secondly, for each $n \in \mathbb{N}^*$ we also define two predictable complex-valued processes with finite variation (see II.2.40 and II.2.47):

1.8
$$\begin{cases} A^n(u) = iu\cdot B^n - \frac{1}{2}u\cdot C^n\cdot u + (e^{iu\cdot x} - 1 - iu\cdot h(x)) * \nu^n \\ G^n(u) = \mathscr{E}[A^n(u)]_t = \{\exp A(u)_t\} \prod_{s\le t} [(1 + \Delta A^n(u)_s)e^{-\Delta A^n(n)}s]. \end{cases}$$

1.9 **Theorem.** *Assume that X has no fixed time of discontinuity* (so $g(u)_t \neq 0$ for all t, u). *If*

1.10
$$G^n(u)_t \xrightarrow{P} g(u)_t \quad \text{for all } u\in\mathbb{R}^d,$$

for all t in a subset D of $\mathbb{R}_+$, we have $X^n \xrightarrow{\mathscr{L}(D)} X$.

This theorem is trivial when each X^n is a PII, in addition to being a semimartingale: indeed in this case, $G^n(u)_t = E^n(\exp iu\cdot X_t^n)$ by II.4.16.

We shall see in the next section that this theorem is the only ingredient needed to reduce the convergence problems of this chapter to the results of the previous chapter.

Proof. We can of course assume that $0\in D$. Let $0 = t_0 < \cdots < t_p$ with $t_j\in D$. We shall prove that $(X_{t_0}^n,\dots,X_t^n) \xrightarrow{\mathscr{L}} (X_{t_0},\dots,X_{t_p})$ by induction on p. The claim is trivial for $p = 0$. We assume that it is true for $p - 1$, and we will prove that

1.11
$$\begin{aligned} &E^n\left[\exp i\left\{\sum_{0\le j\le p-1} u_j\cdot X_{t_j}^n + u\cdot(X_{t_p}^n - X_{t_{p-1}}^n)\right\}\right] \\ &\quad\to E\left[\exp i\left\{\sum_{0\le j\le p-1} u_j\cdot X_{t_j} + u\cdot(X_{t_p} - X_{t_{p-1}})\right\}\right] \end{aligned}$$

for all u_j, $u\in\mathbb{R}^d$, and this will give us the result.

Set $\zeta^n = \exp i\sum_{0\le j\le p-1} u_j\cdot X_{t_j}^n$ and $\zeta = \exp i\sum_{0\le j\le p-1} u_j\cdot X_{t_j}$. The induction hypothesis implies

1.12
$$E^n(\zeta^n) \to E(\zeta).$$

Set $v^n = E^n[\zeta^n \exp iu\cdot(X_{t_p}^n - X_{t_{p-1}}^n)]$ and

$$v = E[\zeta \exp iu\cdot(X_{t_p} - X_{t_{p-1}})] = E(\zeta)g(u)_{t_p}/g(u)_{t_{p-1}}$$

(the PII property of X and the fact that $g(u)_{t_{p-1}} \neq 0$ have been used for the last equality). Then 1.11 reduces to: $v^n \to v$.

Let $a = |g(u)_{t_p}|$. We have $a > 0$, and $R^n = \inf(t\colon |G^n(u)_t| \le \frac{a}{2})$ is a predictable time on $\mathscr{B}^n$ (apply I.2.13 and the fact that $|G^n(u)|$ is predictable and decreasing). Since $|G^n(u)_t| \xrightarrow{P} a$ by 1.10, and since $|G^n(u)|$ is decreasing, we deduce that

1.13
$$P^n(R^n \le t_p) \to 0.$$

R^n being predictable, there is a stopping time S^n on $\mathscr{B}^n$ such that $S^n < R^n$ and $P^n(S^n \le t_p < R^n) \le 1/n$. Combining this with 1.13, we obtain

1.14
$$P^n(S^n \le t_p) \to 0.$$

We have $|G^n(u)| \geq a/2$ on the interval $[\![0, S^n]\!]$; then we deduce from Theorem II.2.47 that the process $M_t^n = (\exp iu \cdot X_{t \wedge S^n}^n)/G^n(u)_{t \wedge S^n}$ is a local martingale on $\mathscr{B}^n$. Since $|M^n| \leq 2/a$ by construction, the process M^n is even a martingale. Therefore

1.15 $$E^n(\beta^n | \mathscr{F}_{t_{p-1}}^n) = 1 \quad \text{if } \beta^n = M_{t_p}^n / M_{t_{p-1}}^n.$$

Finally, set $\gamma^n = G^n(u)_{t_p}/G^n(u)_{t_{p-1}}$, with $0/0 = 0$, and $\gamma = g(u)_{t_p}/g(u)_{t_{p-1}}$. Since ζ^n is $\mathscr{F}_{t_{p-1}}^n$-measurable, 1.15 yields

$$\begin{aligned} v^n &= E^n[\zeta^n 1_{\{S^n \leq t_p\}} \exp iu \cdot (X_{t_p}^n - X_{t_{p-1}}^n) + \zeta^n 1_{\{S^n > t_p\}} \exp iu \cdot (X_{t_p \wedge S^n}^n - X_{t_{p-1} \wedge S^n}^n)] \\ &= E^n[\zeta^n 1_{\{S^n \leq t_p\}} \exp iu \cdot (X_{t_p}^n - X_{t_{p-1}}^n)] + E^n(\zeta^n 1_{\{S^n > t_p\}} \beta^n \gamma^n) \\ &= E^n[\zeta^n 1_{\{S^n \leq t_p\}} \exp iu \cdot (X_{t_p}^n - X_{t_{p-1}}^n)] + E^n[\zeta^n \beta^n (\gamma^n 1_{\{S^n > t_p\}} - \gamma)] + \gamma E^n(\zeta^n). \end{aligned}$$

Since $|\zeta^n| = 1$ we have

1.16 $$|v^n - v| \leq P^n(S^n \leq t_p) + E^n(|\beta^n| |\gamma^n 1_{\{S^n > t_p\}} - \gamma|) + |\gamma| |E^n(\zeta^n) - E(\zeta)|.$$

1.10 and 1.14 imply that $\gamma^n 1_{\{S^n > t_p\}} \xrightarrow{P} \gamma$. Moreover $1 \leq |M^n| \leq \frac{2}{a}$, hence $|\beta^n| \leq 2/a$, and it follows that $\delta^n := \beta^n(\gamma^n 1_{\{S^n > t_p\}} - \gamma) \xrightarrow{P} 0$. If $S^n > t_p$ we have $|\beta^n \gamma^n| = 1$, therefore $|\delta^n| \leq 1 + 2|\gamma|/a$ and we deduce that $E^n(|\delta^n|) \to 0$. Using this, and 1.14 and 1.12, we deduce from 1.16 that $v^n \to v$, and the result is proved. □

§ 1c. Remarks and Comments

Upon carefully examining the previous proof, several comments are in order:

1) The PII property of X is absolutely crucial (indeed, a conditional PII property would also give the result: see Section 5). If X were only a semimartingale, to which we would associate $G(u)$ accordingly to 1.8, the convergence $G^n(u)_t \xrightarrow{P} G(u)_t$ for all $u \in \mathbb{R}^d$, $t \in D$ would *not* (in general) imply $X^n \xrightarrow{\mathscr{L}(D)} X$. As a matter of fact, a fair portion of the next chapter is devoted to studying which kind of stronger convergence of $G^n(u)$ to $G(u)$ would do the job.

2) Contrarywise, the property that X has no fixed time of discontinuity is not fully used, but only through the fact that $g(u)_t \neq 0$ for all u, t. Indeed, in order to obtain that $(X_{t_0}^n, \ldots, X_{t_p}^n) \xrightarrow{\mathscr{L}} (X_{t_0}, \ldots, X_{t_p})$ it is sufficient that 1.11 holds for all u_j, u in $\mathbb{R}^d$, except on a set of zero Lebesgue measure. Hence the previous proof works under the following:

1.17 the set $U_t = \{u \in \mathbb{R}^d : g(u)_t = 0\}$ has Lebesgue measure zero (since $t \rightsquigarrow |g(u)_t|$ decreases, $t \rightsquigarrow U_t$ increases).

Then we have:

1.18 **Theorem.** *Assume that 1.10 and 1.17 hold for all $t \in D$. Then we have $X^n \xrightarrow{\mathscr{L}(D)} X$.*

There is also a version of this theorem, in case 1.17 fails, but it is much more complicated (see [105]).

3) Now, the fact that X^n is a semimartingale has not been fully used either. The useful property of X^n is the following one:

1.19 *Hypothesis*: $G^n(u)$ is a $\mathbb{C}$-valued process such that

(i) $|G^n(u)|$ is decreasing, predictable, and $G^n(u)_0 = 1$;

(ii) if $T^n(u) = \inf(t\colon G^n(u)_t = 0)$, the process $(e^{iu\cdot X^n}/G^n(u))1_{[\![0,T^n(u)[\![}$ is a local martingale on $[\![0, T^n(u)[\![$ (see II.2.46; compare to II.2.47).

1.20 **Theorem.** *Assume that each X^n is an adapted càdlàg process to which a process $G^n(u)$ satisfying 1.19 is associated for each $u \in \mathbb{R}^d$. Assume also that 1.10 and 1.17 hold for all $t \in D$. Then $X^n \xrightarrow{\mathscr{L}(D)} X$.*

If X^n is any process, $G^n(u) = \exp iu\cdot X^n$ satisfies 1.19! but for this particular choice 1.10 cannot be satisfied, unless $C = 0$, $\nu = 0$ and $X^n_t \xrightarrow{\mathscr{L}} B_t$ for all $t \in D$: this would give an absolutely uninteresting theorem.

4) Here is a case where 1.20 is a true improvement upon 1.9 or 1.18. Assume that

$$X^n = X'^n + A^n \tag{1.21}$$

where X'^n is a semimartingale to which we associate $G'^n(u)$ by 1.8. Then 1.19 holds with $G^n(u) = G'^n(u)\exp iu\cdot A^n$, and 1.10 reads as follows:

$$G'^n(u)_t e^{iu\cdot A^n_t} \xrightarrow{P} g(u)_t \quad \text{for all } u \in \mathbb{R}^d. \tag{1.22}$$

5) Another case of interest is when there exist processes $G^n(u)$ satisfying 1.19 and *predictable*; $|G^n(u)|$ has finite variation by 1.19(i), but this is not necessarily the case for $G^n(u)$ itself. For example, suppose that X^n is a PII, but not a semimartingale; then $g^n(u)_t = E(\exp iu\cdot X^n_t)$ satisfies 1.19, although it has not finite variation (at least for some $u \in \mathbb{R}^d$).

2. Convergence to a PII Without Fixed Time of Discontinuity

All notation and assumptions of § 1a are in force, in particular each X^n is a semimartingale on $\mathscr{B}^n$ with $X^n_0 = 0$ and X is a PII. We first introduce a series of conditions on the characteristics, to be used throughout the whole chapter. Recall that $C_i(\mathbb{R}^d)$ (for $i = 1,2,3,4$) denotes a class of functions on $\mathbb{R}^d$ and is defined in VII.2.7. D is a subset of $\mathbb{R}_+$ and we set

$$\begin{cases} [\beta_5\text{-}D] & B^n_t \xrightarrow{P} B_t \quad \text{for all } t \in D \\ [\gamma_5\text{-}D] & \tilde{C}^n_t \xrightarrow{P} \tilde{C}_t \quad \text{for all } t \in D \\ [\delta_{5,i}\text{-}D] & g * \nu^n_t \xrightarrow{P} g * \nu_t \quad \text{for all } t \in D,\ g \in C_i(\mathbb{R}^d) \end{cases} \tag{2.1}$$

2.2 $$\begin{cases} [\text{Sup-}\beta_5] & \sup_{s\le t} |B^n_s - B_s| \xrightarrow{P} 0 \quad \text{for all } t \in \mathbb{R}_+ \\ [\text{Sup-}\gamma_5] & \sup_{s\le t} |\tilde{C}^n_s - \tilde{C}_s| \xrightarrow{P} 0 \quad \text{for all } t \in \mathbb{R}_+ \\ [\text{Sup-}\delta_{5,i}] & \sup_{s\le t} |g * \nu^n_s - g * \nu_s| \xrightarrow{P} 0 \quad \text{for all } t \in \mathbb{R}_+,\, g \in C_i(\mathbb{R}^d). \end{cases}$$

Observe that when all X^n's are PII, which amounts to saying that B^n, C^n, ν^n and $\tilde{C}^n$ are deterministic, these conditions are the conditions labelled $[\beta_3\text{-}D]$, etc..., defined in VII.2.51 and VII.3.2. Due to VII.2.8, we have

2.3 $$\begin{cases} [\delta_{5,3}\text{-}D] \Rightarrow [\delta_{5,2}\text{-}D] \Rightarrow [\delta_{5,1}\text{-}D] \\ [\text{Sup-}\delta_{5,3}] \Rightarrow [\text{Sup-}\delta_{5,2}] \Rightarrow [\text{Sup-}\delta_{5,1}]. \end{cases}$$

§ 2a. Finite-Dimensional Convergence

Here is the main theorem:

2.4 **Theorem.** *Assume that X has no fixed time of discontinuity* (equivalently, $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all t, so B, $\tilde{C}$ are continuous). *Let $D \subset \mathbb{R}_+$.*

a) *Assume that*

2.5 $$\sup_{s\le t} \nu^n(\{s\} \times \{|x| > \varepsilon\}) \xrightarrow{P} 0 \quad \text{for all } t \in D,\, \varepsilon > 0$$

and $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ *hold. Then* $X^n \xrightarrow{\mathscr{L}(D)} X$ (finite-dimensional convergence in law, along D) *and we also have* $[\delta_{5,2}\text{-}D]$.

b) *If D is dense in $\mathbb{R}_+$, then $[\delta_{5,1}\text{-}D]$ implies 2.5* (so $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ implies $X^n \xrightarrow{\mathscr{L}(D)} X$).

When each X^n is a PII, the conditions above are exactly the same than the conditions of VII.2.52 (with 2.5 = VII.2.53).

Proof. As said before, this theorem is the first of a series of results which are proved by reduction to the results of Chapter VII, modulo a typical argument that for once we explain in full details.

We begin by proving (b), which is simpler than (a). We can of course suppose that D is countable. Let $\varepsilon > 0$, $t \in D$ be fixed. In order to prove the convergence in 2.5, it suffices to show that from any infinite subsequence (n') one can extract a further subsequence (n'') such that

2.6 $$\sup_{s\le t} \nu^{n''}(\{s\} \times \{|x| > \varepsilon\}) \xrightarrow{P} 0.$$

We may assume that all processes are defined on the same probability space (see 1.6). By definition of $C_1(\mathbb{R}^d)$ there is a function g in $C_1(\mathbb{R}^d)$ such that $0 \le g \le 1$ and $g(x) = 1$ for $|x| \ge \varepsilon$. Let (n') be an infinite subsequence. By a diagonal argument we can extract a further subsequence (n'') and find a set A with $P(A) = 1$ such that $g * \nu_s^{n''}(\omega) \to g * \nu_s$ for all $s \in D$, $\omega \in A$. Since $s \rightsquigarrow g * \nu_s$ is continuous, the convergence is uniform over compact intervals, while $\Delta(g * \nu^{n''})_s = \nu^{n''}(\{s\} \times g)$, so in particular: $\sup_{s \le t} \nu^{n''}(\omega; \{s\} \times g) \to 0$ if $\omega \in A$. But $\nu^{n''}(\{s\} \times \{|x| > \varepsilon\}) \le \nu^{n''}(\{s\} \times g)$, and 2.6 follows.

Next, we prove (a). We choose (as we can) a family $C_1(\mathbb{R}^d)$ that is countable. Let $t \in D$ be fixed and let (n') be an infinite subsequence. By a diagonal argument, there is a further subsequence (n'') and a set A with $P(A) = 1$, such that

2.7 $$\omega \in A \Rightarrow \begin{cases} \sup\limits_{s \le t} \nu^{n''}(\omega; \{s\} \times \{|x| > \varepsilon\}) \to 0 & \text{for all } \varepsilon > 0 \\ B_t^{n''}(\omega) \to B_t \\ \tilde{C}_t^{n''}(\omega) \to \tilde{C}_t \\ g * \nu_t^{n''}(\omega) \to g * \nu_t & \text{for all } g \in C_1(\mathbb{R}^d). \end{cases}$$

Therefore if $\omega \in A$ is fixed, the sequence $(B^{n''}(\omega), C^{n''}(\omega), \nu^{n''}(\omega))$ satisfies VII.2.53 and $[\beta_3\text{-}\{t\}]$, $[\gamma_3\text{-}\{t\}]$, $[\delta_{3,1}\text{-}\{t\}]$.

Applying VII.2.52, we deduce first that the above sequence meets $[\delta_{3,2}\text{-}\{t\}]$: in other words, for all $g \in C_2(\mathbb{R}^d)$, and all infinite sequence (n') there is a subsequence such that $g * \nu_t^{n''} \to g * \nu_t$ a. s.; thus $g * \nu_t^n \xrightarrow{P} g * \nu_t$ and $[\delta_{5,2}\text{-}D]$ holds.

Fix again $\omega \in A$, and come back to 2.7. In virtue of II.5.2, there is a PII, say $Z^{n'',\omega}$ (defined on some auxiliary stochastic basis) admitting the characteristics $(B^{n''}(\omega), C^{n''}(\omega), \nu^{n''}(\omega))$, and the expectation of $\exp iu \cdot Z_t^{n'',\omega}$ is then $G^{n''}(u)_t(\omega)$, as defined in 1.8. Then Theorem VII.2.52 yields that $\mathscr{L}(Z_t^{n'',\omega}) \to \mathscr{L}(X_t)$, which implies $G^{n''}(u)_t(\omega) \to g(u)_t$. Therefore for all $t \in D$, $u \in \mathbb{R}^d$ and all infinite subsequences (n') there is a further subsequence (depending on t) with $G^{n''}(u)_t \xrightarrow{P} g(u)_t$ a. s., and we deduce that 1.10 holds. Hence the convergence $X^n \xrightarrow{\mathscr{L}(D)} X$ follows from Theorem 1.9. □

2.8 **Remark.** Contrarily to what happens in 1.18, this theorem absolutely requires that X has no fixed time of discontinuity. □

2.9 **Remark.** Let A^n be another adapted càdlàg d-dimensional process on $\mathscr{B}^n$, with $A_0^n = 0$. Then under 2.5, $[\gamma_5\text{-}D]$, $[\delta_{5,1}\text{-}D]$ and

$$B_t^n + A_t^n \xrightarrow{P} B_t \quad \text{for all } t \in D,$$

we have $X^n + A^n \xrightarrow{\mathscr{L}(D)} X$; the proof is similar, but uses 1.20 (and § 1c.4) instead of 1.9 (see [105]). □

2.10 **Remark.** Contrarily to what happens in VII.2.52 the conditions $[\beta_5\text{-}D]$, $[\gamma_5\text{-}D]$ and $[\delta_{5,1}\text{-}D]$ are not necessary for $X^n \xrightarrow{\mathscr{L}(D)} X$, even under 2.5 (see 3.39, or after Theorem 3.65). We can hope for necessary conditions only for the functional convergence $X^n \xrightarrow{\mathscr{L}} X$, and under very specific conditions: some of them will be examined in Section 3. □

Now we state the "square-integrable" version of Theorem 2.4. We suppose that each X^n is a locally square-integrable semimartingale in the sense of II.2.27, which amounts to saying that

2.11 $$|x|^2 * \nu_t^n < \infty \quad \text{for all } t \geq 0.$$

We can put

2.12 $$B'^n = B^n + (x - h(x)) * \nu^n$$

2.13 $$\tilde{C}_t'^{n,jk} = C_t^{n,jk} + (x^j x^k) * \nu_t^n - \sum_{s \leq t} \Delta B_s'^{n,j} \Delta B_s'^{n,k}.$$

Then $M'^n = X^n - B'^n$ is a locally square-integrable martingale, with $\tilde{C}'^{n,ij} = \langle M'^{n,i}, M'^{n,j} \rangle$.

Similarly, assume that X is a PII-semimartingale that is locally square-integrable, so ν meets 2.11 and we define B' and $\tilde{C}'$ by 2.12 and 2.13 (starting with (B, C, ν)).

2.14 **Theorem.** *On top of the above assumptions, assume that X has no fixed time of discontinuity. Let $D \subset \mathbb{R}_+$, and assume that*

2.15 $\lim_{a \uparrow \infty} \limsup_n P^n(|x|^2 1_{\{|x|>a\}} * \nu_t^n > \eta) = 0$ for all $\eta > 0$, $t \in D$

$[\beta_5'\text{-}D]$ $\quad B_t'^n \xrightarrow{P} B_t'$ $\quad$ for all $t \in D$

$[\gamma_5'\text{-}D]$ $\quad \tilde{C}_t'^n \xrightarrow{P} \tilde{C}_t'$ $\quad$ for all $t \in D$

and 2.5 and $[\delta_{5,1}\text{-}D]$ hold. Then $X^n \xrightarrow{\mathcal{L}(D)} X$.

Proof. The proof goes exactly as for 2.4, except that we use VII.2.63 instead of VII.2.52. The only modification is the following: let $t \in D$ and (n') be an infinite subsequence. Then we need to find a further subsequence (n'') and a set A with $P(A) = 1$, such that for all $\omega \in A$ we have 2.7 and

2.16 $$\lim_{a \uparrow \infty} \limsup_{n''} |x|^2 1_{\{|x|>a\}} * \nu_t^{n''}(\omega) = 0$$

(then VII.2.64 will be met by $(B^{n''}(\omega), C^{n''}(\omega), \nu^{n''}(\omega))$ for time t). Indeed, we first find a subsequence $(\tilde{n})$ of (n') and A_0 such that $P(A_0) = 1$ and 2.7 holds for $\omega \in A_0$ and for $(\tilde{n})$. Then if $\rho(n, a) = |x|^2 1_{\{|x|>a\}} * \nu_t^n$, 2.15 yields that for each $k \in \mathbb{N}^*$ there are $a_k > 0$, $\tilde{n}_k \geq k$ such that

$$P\left(\rho(\tilde{n}_k, a_k) > \frac{1}{k}\right) \leq 2^{-k}.$$

Then $A_1 = \liminf_k \{\rho(\tilde{n}_k, a_k) \leq 1/k\}$ has $P(A_1) = 1$ by Borel-Cantelli. Hence $A = A_0 \cap A_1$ satisfies $P(A) = 1$, and the sequence $(n'') = (\tilde{n}_k)$ satisfies 2.7 for all $\omega \in A$, and it remains to prove that it meets 2.16 as well.

If it were not the case for some $\omega \in A$, there would exist $\varepsilon > 0$ and a sequence (k_m) going to $+\infty$, such that $\rho(\tilde{n}_{k_m}, a)(\omega) \geq \varepsilon$ for all $a > 0$ (recall that $a \rightsquigarrow \rho(n, a)$

is decreasing); since $\rho(\tilde{n}_{k_m}, a_{k_m})(\omega) \leq 1/k_m$ for all k_m large enough (by definition of A), we obtain a contradiction. □

§ 2b. Functional Convergence

2.17 **Theorem.** *Assume that X has no fixed time of discontinuity, and that D is a dense subset of $\mathbb{R}_+$. Then* [Sup-β_5] + [γ_5-D] + [$\delta_{5,1}$-D] *imply* $X^n \xrightarrow{\mathscr{L}} X$.

Moreover, in this case we also have [Sup-γ_5] *and* [Sup-$\delta_{5,2}$].

This generalizes the implication (b) ⇒ (a) of Theorem VII.3.4. The *converse is not true* in general when the X^n's are not PII.

Proof. The proof of the last claim goes according to the same scheme as in 2.4: we restrict D so that it becomes countable, while staying dense in $\mathbb{R}_+$; if (n') is an infinite subsequence, there is a further subsequence (n'') such that $\tilde{C}_t^{n''}(\omega) \to \tilde{C}_t$ for all $t \in D$, $\omega \in A$, where $P(A) = 1$. Then Lemma VII.3.8 yields that $\sup_{s \leq t} |\tilde{C}_s^{n''}(\omega) - \tilde{C}_s| \to 0$ for all $t \in \mathbb{R}_+$, $\omega \in A$, and we deduce that [Sup-γ_5] holds. By 2.4 we know that [$\delta_{5,2}$-D] holds, and the same argument (using again VII.3.8) shows that indeed [Sup-$\delta_{5,2}$] holds.

For the first claim, in view of 2.4 it remains to prove that the sequence (X^n) is tight. For this we use Theorem VI.4.18. Since $X_0^n = 0$ the first condition of this theorem is met, and the third condition is met as well because [Sup-β_5], [Sup-γ_5], [Sup-$\delta_{5,2}$] respectively imply that the sequences (B^n), $(\tilde{C}^n)$, $(g * \nu^n)$ for $g \in C_2(\mathbb{R}^d)$ are C-tight. Let $g_q(x) = (q|x| - 1)^+ \wedge 1$ and observe that $\nu^n([0,N] \times \{|x| > a\}) \leq g_{2/a} * \nu_N^n$. Let $\varepsilon > 0$ and $\eta > 0$ and pick $a > 0$ such that $g_{2/a} * \nu_N \leq \varepsilon/2$; then apply [Sup-$\delta_{5,2}$] to find n_0 such that

$$n > n_0 \Rightarrow P^n(g_{2/a} * \nu_N^n > \varepsilon) \leq \eta.$$

It follows that Condition (ii) of VI.4.18 is met, and the theorem is proved. □

The square-integrable version is proved similarly, using 2.14 and VI.4.13 instead of 2.4 and VI.4.18:

2.18 **Theorem.** *Assume that X is a PII-semimartingale without fixed time of discontinuity, and that ν^n and ν meet 2.11. Let D be a dense subset of $\mathbb{R}_+$. Define B'^n, B' by 2.12 and $\tilde{C}'^n$, C' by 2.13. If*

$$[\text{Sup-}\beta'_5] \quad \sup_{s \leq t} |B_s'^n - B_s'| \xrightarrow{P} 0 \quad \text{for all } t \geq 0$$

and [γ'_5-D] + [$\delta_{5,1}$-D] *and 2.15 hold, then* $X^n \xrightarrow{\mathscr{L}} X$.

Moreover, in this case we also have [Sup-$\delta_{5,2}$] *and*

$$[\text{Sup-}\gamma'_5] \quad \sup_{s \leq t} |\tilde{C}_s'^n - \tilde{C}_s'| \xrightarrow{P} 0 \quad \text{for all } t \geq 0.$$

For further reference, we state the following corollary of Propositions VII.2.59 and VII.3.12.

2.19 **Proposition.** *Assume that X is a PII-semimartingale without fixed time of discontinuity, let D be a subset of $\mathbb{R}_+$, and put $X'^n = X^n - B^n$ and $X' = X - B$, with a uniformly continuous truncation function h.*

a) *If $[\gamma_5\text{-}D]$, $[\delta_{5,1}\text{-}D]$ and 2.5 hold, we have $X'^n \xrightarrow{\mathscr{L}(D)} X'$.*

b) *If D is dense in $\mathbb{R}_+$, and if $[\gamma_5\text{-}D]$ and $[\delta_{5,1}\text{-}D]$ hold, we have $X'^n \xrightarrow{\mathscr{L}} X'$.*

Proof. We can assume that D is countable, and we call (B'^n, C'^n, ν'^n) and (B', C', ν') the characteristics of X'^n and X' (so $B' = 0$, $C' = C, \nu' = \nu$), from which we compute the modified second characteristics $\tilde{C}'^n$ and $\tilde{C}'$.

Let (n') be a subsequence. Exactly as in 2.4 we find a set A with $P(A) = 1$ and a further subsequence (n'') such that 2.7 holds for all $t \in D$ (except the statement about $B^{n''}$ and B). Then the same argument than in 2.4, but based on VII.2.59 and VII.3.12, shows that for $\omega \in A$, $t \in D$:

$$\text{2.20} \qquad \begin{cases} \tilde{C}'^{n''}_t(\omega) \to \tilde{C}'_t, \quad g * \nu'^{n''}_t(\omega) \to g * \nu_t \quad \text{for } g \in C_1(\mathbb{R}^d) \\ \sup_{s \leq t} \nu'^{n''}(\omega; \{s\} \times \{|x| > \varepsilon\}) \to 0 \quad \text{for } \varepsilon > 0 \\ \sup_{s \leq t} |B'^{n''}_s(\omega)| \to 0 \end{cases}$$

By the usual argument, 2.20 implies that $[\gamma_5\text{-}D]$, $[\delta_{5,1}\text{-}D]$, 2.5, $[\beta_5\text{-}D]$, and $[\text{Sup-}\beta_5]$ in case (b), are met by (X'^n) and the limiting process X': so the claims follow from 2.4 and 2.17. □

2.21 **Remarks.** We have seen in Chapter VII that under $[\delta_{3,1}\text{-}D]$, the conditions $[\beta_3\text{-}D]$, $[\text{Sup-}\beta_3]$, $[\gamma_3\text{-}D]$ do not depend upon the (continuous) truncation function h. This fact also applies here, namely that under $[\delta_{5,1}\text{-}D]$ the conditions $[\beta_5\text{-}D]$, $[\text{Sup-}\beta_5]$, $[\gamma_5\text{-}D]$ do not depend upon h (the proof still goes along the same scheme than in 2.4. □

§ 2c. Application to Triangular Arrays

Very often one has to consider limits of partial sums of triangular arrays that are not rowwise independent. In order to fit within the previous setting, we will consider only those triangular arrays that give rise to semimartingales, so we set:

2.22 **Definition.** A *d-dimensional semimartingale triangular array scheme* consists in the following: for all $n \in \mathbb{N}^*$ we have a discrete-time stochastic basis $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, (\mathscr{F}^n_p)_{p \in \mathbb{N}}, P^n)$ endowed with an adapted sequence $(U^n_k)_{k \geq 1}$ of d-dimensional random variables, and

(i) either a change of time $\sigma^n = (\sigma^n_t)_{t \geq 0}$ on $\mathscr{B}^n$ (see II.3.5)

(ii) or a single stopping time K^n, also denoted by $K^n = \sigma_1^n$, such that for all $t \in \mathbb{R}_+$ in case (i), and for $t = 1$ in case (ii),

2.23
$$\begin{cases} \sum_{1 \leq k \leq \sigma_t^n} |E[h(U_k^n)|\mathscr{F}_{k-1}^n]| < \infty \quad \text{a.s.} \\ \sum_{1 \leq k \leq \sigma_t^n} E(|U_k^n|^2 \wedge 1|\mathscr{F}_{k-1}^n) < \infty \quad \text{a.s.} \end{cases}$$ □

In case (i), we define the continuous-time filtration $\mathbf{G}^n$ by $\mathscr{G}_t^n = \mathscr{F}_{\sigma_t^n}^n$, and the process

2.24
$$X_t^n = \sum_{1 \leq k \leq \sigma_t^n} U_k^n.$$

The series in 2.24 converges in measure, and X^n is a semimartingale on $(\Omega^n, \mathscr{F}^n, \mathbf{G}^n, P^n)$ by II.3.11. In case (ii), we just have the following random variable:

2.25
$$Z^n = \sum_{1 \leq k \leq K^n} U_k^n.$$

There is a way of reducing case (ii) to case (i), which is as follows: let (v_n) be a strictly increasing sequence of numbers, with $v_0 = 0$ and $\lim_n \uparrow v_n = 1$. Set

2.26
$$\sigma_t^n(\omega) = \begin{cases} m \wedge K^n(\omega) & \text{if } v_m \leq t < v_{m+1} \\ K^n(\omega) & \text{if } t \geq 1. \end{cases}$$

Then $(\sigma_t^n)_{t \geq 0}$ is a change of time of $\mathscr{B}^n$, with $\sigma_1^n = K^n$, and with notation 2.24 and 2.25 we have $Z^n = X_1^n$.

Now we state the versions of 2.4 and 2.17 relative to the present situation. For simplicity, the "finite-dimensional" version is stated with $D = \{1\}$, and this corresponds to the behaviour of Z^n (case (ii) above).

2.27 **Theorem.** *We consider a d-dimensional semimartingale triangular array scheme with case* (ii), *and Z^n defined by 2.25. Let μ be an infinitely divisible distribution on $\mathbb{R}^d$ with characteristics (b, c, F), and $\tilde{c}$ be defined by VI.2.5. If*

2.28
$$\sup_{1 \leq k \leq K^n} P^n(|U_k^n| > \varepsilon|\mathscr{F}_{k-1}^n) \xrightarrow{P} 0 \quad \text{for all } \varepsilon > 0$$

$[\beta_6]$
$$\sum_{1 \leq k \leq K^n} E^n[h(U_k^n)|\mathscr{F}_{k-1}^n)] \xrightarrow{P} b$$

$[\gamma_6]$
$$\sum_{1 \leq k \leq K^n} \{E^n[h^j h^l(U_k^n)|\mathscr{F}_{k-1}^n] - E^n[h^j(U_k^n)|\mathscr{F}_{k-1}^n] E^n[h^l(U_k^n)|\mathscr{F}_{k-1}^n]\} \xrightarrow{P} \tilde{c}^{jl}$$

$[\delta_{6,1}]$
$$\sum_{1 \leq k \leq K^n} E^n[g(U_k^n)|\mathscr{F}_{k-1}^n] \xrightarrow{P} F(g) \quad \text{for all } g \in C_1(\mathbb{R}^d),$$

then $\mathscr{L}(Z^n) \to \mu$ weakly.

Proof. Apply 2.4 to X^n defined by 2.24, with σ_t^n defined by 2.26, and $D = \{1\}$, and to the PIIS X satisfying $\mathscr{L}(X_1) = \mu$, and use II.3.14. □

2.29 **Theorem.** *We consider a d-dimensional semimartingale triangular array scheme with case* (i), *and X^n defined by 2.24. Let X be a PII without fixed time of discontinuity, with characteristics (B, C, ν). If*

$$[\text{Sup-}\beta_6] \quad \sup_{s \le t} \left| \sum_{1 \le k \le \sigma_s^n} E^n[h(U_k^n)|\mathscr{F}_{k-1}^n] - B_s \right| \xrightarrow{P} 0 \quad \text{for all } t \ge 0$$

$$[\gamma_6\text{-}D] \quad \sum_{1 \le k \le \sigma_t^n} \{E^n[h^j h^l(U_k^n)|\mathscr{F}_{k-1}^n] - E^n[h^j(U_k^n)|\mathscr{F}_{k-1}^n] \times E^n[h^l(U_k^n)|\mathscr{F}_{k-1}^n]\} \xrightarrow{P} \tilde{C}_t^{jl} \quad \text{for all } t \in D$$

$$[\delta_{6,1}\text{-}D] \quad \sum_{1 \le k \le \sigma_t^n} E^n[g(U_k^n)|\mathscr{F}_{k-1}^n] \xrightarrow{P} g * \nu_t \quad \text{for all } t \in D, g \in C_1(\mathbb{R}^d),$$

hold for some dense subset D of $\mathbb{R}_+$, then $X^n \xrightarrow{\mathscr{L}} X$.

We leave to the reader the versions of 2.14 and 2.18 for triangular arrays.

§ 2d. Other Conditions for Convergence

Now we come back to the general setting of this section. In Theorem 1.9 we gave a condition on $G^n(u)$ (defined in 1.8) to the effect that $X^n \xrightarrow{\mathscr{L}(D)} X$. The following gives a stronger condition on $G^n(u)$, in order that $X^n \xrightarrow{\mathscr{L}} X$, and it is an easy extension of Corollary VII.4.43.

2.30 **Theorem.** *Assume that X has no fixed time of discontinuity. If*

2.31
$$\sup_{|u| \le \theta} \sup_{s \le t} |G^n(u)_s - g(u)_s| \xrightarrow{P} 0 \quad \text{for all } t > 0, \theta > 0,$$

then $X^n \xrightarrow{\mathscr{L}} X$.

In view of Theorem 2.17, this is a corollary of the following:

2.32 **Lemma.** *If X has no fixed time of discontinuity, there is equivalence between 2.31 and the three conditions* $[\text{Sup-}\beta_5] + [\gamma_5\text{-}\mathbb{R}_+] + [\delta_{5,1}\text{-}\mathbb{R}_+]$.

Proof. Suppose first 2.31. Let (n') be an infinite subsequence. A diagonal argument gives A with $P(A) = 1$ and a further subsequence (n'') with

2.33
$$\omega \in A \Rightarrow \sup_{|u| \le \theta} \sup_{s \le t} |G^{n''}(u)_s(\omega) - g(u)_s| \to 0, \quad \text{all } t > 0, \theta > 0.$$

Then VII.4.43 and the implication (a) $\Rightarrow$ (b) of VII.3.4 yield

2.34
$$\omega \in A \Rightarrow \begin{cases} \sup_{s \le t} |B_s^{n''}(\omega) - B_s| \to 0 & \text{for all } t > 0 \\ \tilde{C}_t^{n''}(\omega) \to \tilde{C}_t & \text{for all } t > 0 \\ g * \nu_t^{n''}(\omega) \to g * \nu_t & \text{for all } t > 0, g \in C_1(\mathbb{R}^d) \end{cases}$$

and we deduce $[\text{Sup-}\beta_5]$, $[\gamma_5\text{-}\mathbb{R}_+]$ and $[\delta_{5,1}\text{-}\mathbb{R}_+]$ as in the proof of 2.4.

Conversely, suppose that [Sup-β_5], [γ_5-$\mathbb{R}_+$], [$\delta_{5,1}$-$\mathbb{R}_+$] hold. Then by 2.17 we also have [Sup-γ_5] and [Sup-$\delta_{5,2}$]. Hence if (n') is an infinite subsequence there is a further subsequence (n'') and a set A with $P(A) = 1$, such that 2.34 holds (recall that we can choose $C_1(\mathbb{R}^d)$ to be countable). Then we deduce that 2.33 holds from the implication (b) $\Rightarrow$ (a) of VII.3.4 and from VII.4.43, and 2.31 follows. □

2.35 **Remark.** Note the inversion of the argument: we use 1.9 for 2.4, and 2.17 for 2.30! actually, the key point for proving 2.4 is the following implication, valid under 2.5:

$$[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D] \Rightarrow G^n(u)_t \xrightarrow{P} g(u)_t \quad \text{for all } t \in D, u \in \mathbb{R}^d.$$

We do not know whether this is an equivalence, but we do have the following equivalence, under 2.5 again:

$$[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D] \Leftrightarrow \sup_{|u| \leq \theta} |G^n(u)_t - g(u)_t| \xrightarrow{P} 0 \quad \text{for all } t \in D, \theta > 0. \quad □$$

Now, the previous theorem is perhaps not so useful for applications, because Condition 2.31, based upon the rather complicated process $G^n(u)$, is likely to be very difficult to check. It would be more useful to have a condition based upon the processes $A^n(u)$ of 1.8. A general result in that direction is not available, and we will restrict to the very particular case where $G^n(u) = \exp A^n(u)$, which amounts to saying that all X^n are quasi-left continuous.

2.36 **Theorem.** *Assume that X is a PII-semimartingale without fixed time of discontinuity, and that all X^n are quasi-left-continuous. Define $A^n(u)$ and $A(u)$ by 1.8 ($A(u)$* is computed from (B, C, ν)), *and let $D \subset \mathbb{R}_+$.*

a) *If*

2.37 $$A^n(u)_t \xrightarrow{P} A(u)_t \quad \text{for all } t \in D, u \in \mathbb{R}^d,$$

then $X^n \xrightarrow{\mathcal{L}(D)} X$.

b) *Assume that D is dense in $\mathbb{R}_+$. If 2.37 and* [Sup-β_5] *hold, then $X^n \xrightarrow{\mathcal{L}} X$.*

c) *If*

2.38 $$\sup_{s \leq t, |u| \leq \theta} |A^n(u)_s - A(u)_s| \xrightarrow{P} 0 \quad \text{for all } t \in \mathbb{R}_+, \theta \in \mathbb{R}_+,$$

then $X^n \xrightarrow{\mathcal{L}} X$.

Proof. By hypothesis, $G^n(u)_t = \exp A^n(u)_t$ and $g(u)_t = \exp A(u)_t$. Then 2.37 and 2.38 respectively imply 1.10 and 2.31, so (a) and (c) follow from Theorems 1.9 and 2.30.

It remains to prove (b). In view of 2.17, it suffices to prove that 2.37 implies [γ_5-D] and [$\delta_{5,1}$-D], which amounts to proving [γ_5-$\{t\}$] and [$\delta_{5,1}$-$\{t\}$] for all $t \in D$. So we fix $t \in D$.

2.37 implies $G^n(u)_t \xrightarrow{P} g(u)_t$ and $|G^n(u)_t| \leq 1$, hence for all $p \in \mathbb{N}^*$

2.39 $$E^n\left(\int_{|u|\leq 1}\left|G^n\left(\frac{u}{p}\right)_t - g\left(\frac{u}{p}\right)_t\right| du\right) = \int_{|u|\leq 1} E^n\left(\left|G^n\left(\frac{u}{p}\right)_t - g\left(\frac{u}{p}\right)_t\right|\right) du \to 0.$$

Let (n') be an infinite subsequence. 2.39 implies the existence of a set A with $P(A) = 1$ and of a further subsequence (n'') such that

2.40 $$\omega \in A \Rightarrow \begin{cases} \displaystyle\int_{|u|\leq 1}\left|G^{n''}\left(\frac{u}{p}\right)_t(\omega) - g\left(\frac{u}{p}\right)_t\right| du \to 0 & \text{for all } p \in \mathbb{N}^* \\ G^{n''}(u)_t(\omega) \to g(u)_t & \text{for all } u \in \mathbb{Q}^d. \end{cases}$$

Now, $u \rightsquigarrow G^{n''}(u)_t(\omega)$ is the characteristic function of an infinitely divisible distribution $\eta^{n''}_{\omega,t}$ whose characteristics (see § VII.2a) are $B^{n''}_t(\omega)$, $C^{n''}_t(\omega)$, $\nu^{n''}(\omega; [0,t] \times \cdot)$. Moreover, the majoration VII.2.17 yields for $p \in \mathbb{N}^*$:

2.41 $$\eta^{n''}_{\omega,t}(|x|>p) \leq \int \eta^{n''}_{\omega,t}(dx)\left(\left|\frac{x}{p}\right|^2 \wedge 1\right) \leq C_1(1)\int_{|u|\leq 1}\left[1 - \operatorname{Re} G^{n''}\left(\frac{u}{p}\right)_t(\omega)\right] du.$$

Fix $\omega \in A$ and $\varepsilon > 0$. There exists $p \in \mathbb{N}^*$ such that $\int_{|u|\leq 1}\left(1 - \operatorname{Re} g\left(\frac{u}{p}\right)_t\right) du \leq \varepsilon$, so 2.40 yields that $\int_{|u|\leq 1}\left[1 - \operatorname{Re} G^{n''}\left(\frac{u}{p}\right)_t(\omega)\right] du \leq 2\varepsilon$ for all n'' large enough, and 2.41 yields for those n'': $\eta^{n''}_{\omega,t}(|x| > p) \leq 2\varepsilon C_1(1)$. Since $\varepsilon > 0$ is arbitrary, we deduce that the sequence $\{\eta^{n''}_{\omega,t}\}_{n''}$ is tight on $\mathbb{R}^d$, and the second property in 2.40 then yields that this sequence actually converges weakly to the infinitely divisible distribution with characteristic function $u \rightsquigarrow g(u)_t$. Therefore VII.2.9 implies:

$$\omega \in A \Rightarrow \begin{cases} \tilde{C}^{n''}_t(\omega) \to \tilde{C}_t \\ g * \nu^{n''}_t(\omega) \to g * \nu_t & \text{for all } g \in C_1(\mathbb{R}^d). \end{cases}$$

Again allowing to our usual method, we deduce that $\tilde{C}^n_t \xrightarrow{P} \tilde{C}_t$ and $g * \nu^n_t \xrightarrow{P} g * \nu_t$ for all $g \in C_1(\mathbb{R}^d)$: hence $[\gamma_5\text{-}\{t\}]$ and $[\delta_{5,1}\text{-}\{t\}]$ hold, and we are finished. □

3. Applications

The aim of this section is twofold:

1) We specialize the previous results to the cases that are most often encountered in practice: firstly when the limiting process X is a continuous PII (in general, a Wiener process), so that we really obtain "functional central limit theorems" (also known as "invariance principles"); moreover it is then possible to get necessary and sufficient conditions for convergence (§§ 3a,b,c). Secondly the case where all X^n's and X are point processes, so in particular X is a Poisson process (§ 3d).

2) We also give three applications: the first one concerns normalized sums of i.i.d. semimartingales: this problem, as natural (and obvious) as it may look, is in fact unsolved so far in general, and we must be content with very partial results (§ 3e).

The second application (§ 3f) concerns the limiting behaviour of some functionals of recurrent Markov processes (other results concerning Markov processes will be obtained in the next chapter). In close relationship to this, our third application concerns the convergence of normalized stationary ergodic processes, under assumptions weaker than the usual mixing conditions (§ 3g). More on this subject will be given in Section 5.

§ 3a. Central Limit Theorem: Necessary and Sufficient Conditions

The setting is as in § 1a, and in addition we suppose that the PII X is *continuous*, with characteristics $(B, C, 0)$. We recall the notation II.2.4 and II.2.5:

3.1
$$\begin{cases} \check{X}^n(h) = \sum_{s \leq \cdot} [\Delta X^n_s - h(\Delta X^n_s)] \\ X^n(h) = X^n - \check{X}^n(h), \quad M^n = X^n(h) - B^n - X^n_0, \end{cases}$$

so $\tilde{C}^{n,ij} = \langle M^{n,i}, M^{n,j} \rangle$. We introduce also the quadratic variation process

3.2
$$\hat{C}^n = (\hat{C}^{n,ij})_{i,j \leq d}, \text{ with } \hat{C}^{n,ij} = [M^{n,i}, M^{n,j}].$$

We use the conditions $[\beta_5\text{-}D]$, $[\gamma_5\text{-}D]$, $[\delta_{5,1}\text{-}D]$ and $[\text{Sup-}\beta_5]$ introduced in 2.1 and 2.2, where $D \subset \mathbb{R}_+$. Since $\nu = 0$ we get $\tilde{C} = C$, so here we have:

3.3
$$[\gamma_5\text{-}D] = \{\tilde{C}^n_t \xrightarrow{P} C_t \quad \text{for all } t \in D\}.$$

We also introduce two other conditions:

3.4
$$\begin{cases} [\hat{\gamma}_5\text{-}D] & \hat{C}^n_t \xrightarrow{P} C_t \quad \text{for all } t \in D; \\ [\hat{\delta}_5\text{-}D] & \nu^n([0,t] \times \{|x| > \varepsilon\}) \xrightarrow{P} 0 \quad \text{for all } t \in D, \varepsilon > 0. \end{cases}$$

Then, using again the property $\nu = 0$, and Lemma VI.4.22, we obtain the equivalence:

3.5
$$[\delta_{5,1}\text{-}D] \Leftrightarrow [\delta_{5,2}\text{-}D] \Leftrightarrow [\hat{\delta}_5\text{-}D] \Leftrightarrow \left\{ \sup_{s \leq t} |\Delta X^n_s| \xrightarrow{P} 0 \text{ for all } t \in D \right\}$$

Our first result concerns finite-dimensional convergence, and it is a paraphrase of Theorem 2.4.

3.6 **Theorem.** *Assume that X is a continuous PII with characteristics $(B, C, 0)$, and let $D \subset \mathbb{R}_+$.*

a) *Under $[\hat{\delta}_5\text{-}D]$ we have: $[\gamma_5\text{-}D] \Leftrightarrow [\hat{\gamma}_5\text{-}D]$.*

b) *Under $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$, or under $[\beta_5\text{-}D] + [\hat{\gamma}_5\text{-}D] + [\hat{\delta}_5\text{-}D]$, we have $X^n \xrightarrow{\mathscr{L}(D)} X$.*

Proof. In view of (a) and 3.5, (b) is just a restatement of 2.4. In order to prove (a), we assume $[\hat{\delta}_5\text{-}D]$. We fix $j, k \le d$ and we set $Y^n := \hat{C}^{n,jk} - \tilde{C}^{n,jk}$; we will prove that $\sup_{s\le t}|Y^n_s| \xrightarrow{P} 0$ for all $t\in D$, under either $[\gamma_5\text{-}D]$ or $[\hat{\gamma}_5\text{-}D]$, and this will imply the claim.

Note that Y^n is a local martingale with finite variation. Let A satisfy 1.2. We have $|\Delta X^n(h)| = |h(\Delta X^n)| \le A$, and $\Delta B^n_t = \nu^n(\{t\}\times h)$, so $|\Delta B^n| \le A$, hence $|\Delta M^n| \le 2A$, hence $|\Delta \hat{C}^{n,jk}| \le 4A^2$ and $|\Delta \tilde{C}^{n,jk}| \le 4A^2$, and finally $|\Delta Y^n| \le 8A^2$. Therefore Y^n is locally square-integrable and for any stopping time T we have $E^n((Y^n_T)^2) \le E^n([Y^n, Y^n]_T)$. Then Lenglart's inequality I.3.32 yields for all $\varepsilon > 0$, $\eta > 0$:

$$P^n\left(\sup_{s\le t}|Y^n_s|^2 \ge \varepsilon\right) \le \frac{1}{\varepsilon}\left(\eta + E^n\left(\sup_{s\le t}\Delta[Y^n, Y^n]_s\right)\right) + P^n([Y^n, Y^n]_t \ge \eta)$$

$$\le \frac{\eta}{\varepsilon} + \left(\frac{64A^4}{\varepsilon} + 1\right)P^n([Y^n, Y^n]_t \ge \eta)$$

because $\Delta[Y^n, Y^n] \le 64A^4$ and $\sup_{s\le t}\Delta[Y^n, Y^n]_s \le [Y^n, Y^n]_t$. It is obvious that the above tends to 0 when $n\uparrow\infty$ for all $\varepsilon > 0$, $t\in D$, provided we have

3.7 $$P^n([Y^n, Y^n]_t \ge \eta) \to 0 \quad \text{for all } t\in D, \eta > 0.$$

So it remains to prove 3.7 under either $[\gamma_5\text{-}D]$ or $[\hat{\gamma}_5\text{-}D]$.

$\hat{C}^{n,jj}$ is L-dominated by $\tilde{C}^{n,jj}$, and vice-versa, and their jumps are smaller than $4A^2$, hence I.3.32 yields

$$P^n(\tilde{C}^{n,jj}_t \ge a) \le \frac{1}{a}(\eta + 4A^2) + P^n(\hat{C}^{n,jj}_t \ge \eta),$$

and the same after interchanging $\tilde{C}^n$ and $\hat{C}^n$. Then we easily deduce that the two following properties are equivalent:

(1) $$\lim_{a\uparrow\infty}\limsup_n P^n(\tilde{C}^{n,jj}_t + \tilde{C}^{n,kk}_t \ge a) = 0$$

(2) $$\lim_{a\uparrow\infty}\limsup_n P^n(\hat{C}^{n,jj}_t + \hat{C}^{n,kk}_t \ge a) = 0.$$

Next, we set

$$\alpha^n = \sup_{s\le t}|\Delta\hat{C}^{n,jk}_s|,\ \beta^n = \sup_{s\le t}|\Delta\tilde{C}^{n,jk}_s|,\ \gamma^n = \sup_{s\le t}|\Delta B^n_s|.$$

From $\Delta B^n_t = \nu^n(\{t\}\times h)$ and 1.2, for all $\varepsilon\in(0, 1/A]$ we have

$$\gamma^n \le \varepsilon + A\nu^n([0,t]\times\{|x| > \varepsilon\}).$$

Hence $[\hat{\delta}_5\text{-}D]$ implies $\gamma^n \xrightarrow{P} 0$. We also deduce from 3.5 that $\sup_{s\le t}|\Delta X^n_s(h)| \xrightarrow{P} 0$ and thus $\sup_{s\le t}|\Delta M^n_s| \xrightarrow{P} 0$, and $\alpha^n \xrightarrow{P} 0$. Finally 1.3 and 1.2 yield

$$\beta^n \le (\gamma^n)^2 + \varepsilon^2 + A^2\nu^n([0,t]\times\{|x| > \varepsilon\})$$

for $\varepsilon\in(0, 1/A]$ and we deduce that $\beta^n \xrightarrow{P} 0$.

We have

$$\begin{aligned}[Y^n, Y^n]_t &= \sum_{s\le t} [\Delta \hat{C}_s^{n,jk} - \Delta \tilde{C}_s^{n,jk}]^2 \\ &\le 2 \sum_{s\le t} [(\Delta \hat{C}_s^{n,jk})^2 + (\Delta \tilde{C}_s^{n,jk})^2] \\ &\le \alpha^n(\hat{C}_t^{n,jj} + \hat{C}_t^{n,kk}) + \beta^n(\tilde{C}_t^{n,jj} + \tilde{C}_t^{n,kk}),\end{aligned}$$

because $\tilde{C}^n$ and $\hat{C}^n$ are symmetric nonnegative matrices. Since $\alpha^n \xrightarrow{P} 0$ and $\beta^n \xrightarrow{P} 0$ and since (1) and (2) hold, under either $[\gamma_5\text{-}D]$ or $[\hat{\gamma}_5\text{-}D]$, we then deduce 3.7 from the above, and the proof is finished. □

Now comes the functional convergence theorem, for which we have a sort of necessary and sufficient condition (except when all X^n are PII, it is probably impossible to obtain similar necessary conditions for finite-dimensional convergence).

3.8 **Theorem.** *Assume that X is a continuous PII with characteristics $(B, C, 0)$, and let D be a dense subset of $\mathbb{R}_+$.*

a) *If $X^n \xrightarrow{\mathcal{L}} X$ then $[\hat{\delta}_5\text{-}D]$ holds.*

b) *Under* $[\text{Sup-}\beta_5]$, *there is equivalence between*

(i) $X^n \xrightarrow{\mathcal{L}} X$;

(ii) $[\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$;

(iii) $[\hat{\gamma}_5\text{-}D] + [\hat{\delta}_5\text{-}D]$.

Moreover, in this case $\tilde{C}^n \xrightarrow{\mathcal{L}} C$ and $\hat{C}^n \xrightarrow{\mathcal{L}} C$ (convergence in law to the deterministic "process" C).

It is important to notice that we can have $X^n \xrightarrow{\mathcal{L}} X$ even when $[\text{Sup-}\beta_5]$ fails: we shall see examples in § 3f. This equivalence also badly fails when X is discontinuous, as we shall see in § 3d.

Proof. a) Since X is continuous, $\sup_{s\le t} |\Delta X_s^n| \xrightarrow{P} 0$ by VI.3.26, thus the claim follows from 3.5.

b) (ii)$\Leftrightarrow$(iii) follows from 3.6a, and (ii)$\Rightarrow$(i) follows from 2.17. Suppose now that $X^n \xrightarrow{\mathcal{L}} X$. By (a) we have $[\hat{\delta}_5\text{-}D]$, and so $\check{X}^n(h) \xrightarrow{\mathcal{L}} 0$. Therefore $X^n(h) \xrightarrow{\mathcal{L}} X$. Since $B^n \xrightarrow{\mathcal{L}} B$ by $[\text{Sup-}\beta_5]$ and since B is continuous, we deduce that $M^n \xrightarrow{\mathcal{L}} M = X - B$ (see VI.3.33). Moreover, we have seen that $|\Delta M^n| \le 2A$, where A satisfies 1.2. Therefore VI.6.29 yields $\hat{C}^n \xrightarrow{\mathcal{L}} C$, which gives $[\hat{\gamma}_5\text{-}\mathbb{R}_+]$. Hence (i) $\Rightarrow$ (iii).

Finally, that $\tilde{C}^n \xrightarrow{\mathcal{L}} C$ under (ii) can be checked, either by using $\hat{C}^n \xrightarrow{\mathcal{L}} C$ and $\hat{C}^n - \tilde{C}^n \xrightarrow{\mathcal{L}} 0$ (proved in 3.6b under $[\hat{\delta}_5\text{-}D] + [\gamma_5\text{-}D]$), or directly from $[\gamma_5\text{-}D]$ (as in Lemma VII.3.8). □

§ 3b. Central Limit Theorem: The Martingale Case

Here we specialize further: we suppose that *each X^n is a local martingale.* It is then natural to assume also that the limiting process X is a continuous Gaussian martingale (see § II.4d), i.e. a PII with characteristics $(0, C, 0)$. Note that by II.2.29a,

3.9
$$\begin{cases} |x|^2 \wedge |x| * v_t^n < \infty & \text{for all } t \leq 0 \\ B^n = [h(x) - x] * v^n. \end{cases}$$

In addition to the conditions introduced in § 3a, it is natural to also introduce the following conditions:

$[\hat{\gamma}_5'\text{-}D]$ $\quad [X^{n,i}, X^{n,j}]_t \xrightarrow{P} C_t^{ij}$ for all $t \in D$

and, when each X^n is locally square-integrable:

$[\gamma_5'\text{-}D]$ $\quad \langle X^{n,i}, X^{n,j}\rangle_t \xrightarrow{P} C_t^{ij}$ for all $t \in D$.

(Note that in the latter case, $\langle X^{n,i}, X^{n,j}\rangle = \tilde{C}'^{n,ij}$, with notation 2.13, so $[\gamma_5'\text{-}D]$ is the same as in 2.14). The same argument as in Lemma VII.3.8 yields, when D is *dense* in $\mathbb{R}_+$ (and with obvious notation for $[X^n, X^n]$ and $\langle X^n, X^n\rangle$):

3.10
$$\begin{cases} [\hat{\gamma}_5'\text{-}D] \Leftrightarrow [X^n, X^n] \xrightarrow{\mathscr{L}} C \\ [\gamma_5'\text{-}D] \Leftrightarrow \langle X^n, X^n\rangle \xrightarrow{\mathscr{L}} C. \end{cases}$$

Our first result is very simple, but is perhaps the most useful.

3.11 Theorem. *Assume that X is a continuous Gaussian martingale with characteristics $(0, C, 0)$, and that each X^n is a local martingale with $|\Delta X^n| \leq K$ identically. If D is a dense subset of $\mathbb{R}_+$, the following are equivalent:*

(i) $X^n \xrightarrow{\mathscr{L}} X$;

(ii) $[\hat{\gamma}_5'\text{-}D]$;

(iii) $[\gamma_5'\text{-}D] + [\hat{\delta}_5\text{-}D]$ (see 3.4).

Proof. We choose a truncation function h such that $h(x) = x$ for $|x| \leq K$. Then $\check{X}^n(h) = 0$, and $B^n = 0$, and $\hat{C}^{n,ij} = [X^{n,i}, X^{n,j}]$, and $\tilde{C}^{n,ij} = \langle X^{n,i}, X^{n,j}\rangle$. Thus $[\text{Sup-}\beta_5]$ obviously holds, and $[\hat{\gamma}_5'\text{-}D] = [\hat{\gamma}_5\text{-}D]$, and $[\gamma_5'\text{-}D] = [\gamma_5\text{-}D]$. Hence the result is a consequence of Theorem 3.8b, provided we prove the implication $[\hat{\gamma}_5'\text{-}D] \Rightarrow [\hat{\delta}_5\text{-}D]$.

We observe that $[\hat{\gamma}_5'\text{-}D]$ implies $[X^{n,i}, X^{n,i}] \xrightarrow{\mathscr{L}} C^{ii}$ (see 3.10). Then *a*-fortiori $\sup_{s \leq t} \Delta[X^{n,i}, X^{n,i}]_s \xrightarrow{P} 0$ (see VI.3.26). Since $|\Delta X^n|^2 = \sum_{j \leq d} \Delta[X^{n,j}, X^{n,j}]$ we deduce $[\hat{\delta}_5\text{-}D]$ from 3.5, and we are finished. □

3.12 Theorem. *Assume that X is a continuous Gaussian martingale with characteristics $(0, C, 0)$, and that each X^n is a local martingale. Let D be a dense subset of $\mathbb{R}_+$, and consider the two conditions:*

3.13 $$\lim_{a\uparrow\infty} \limsup_n P^n(|x|1_{\{|x|>a\}} * \nu_t^n > \eta) = 0 \quad \text{for all } \eta > 0, t > 0$$

3.14 the sequence $\left(\sup_{s\le t} |\Delta X_s^n| \,\Big|\, P^n\right)$ is uniformly integrable for all $t > 0$.

a) *3.14 implies 3.13.*

b) *Under 3.13, we have equivalence between:*

(i) $X^n \xrightarrow{\mathcal{L}} X$;

(ii) $[\hat{\gamma}'_5\text{-}D]$;

(iii) $[\hat{\gamma}_5\text{-}D] + [\hat{\delta}_5\text{-}D]$;

(iv) $[\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$.

We begin with two lemmas.

3.15 **Lemma.** *Under 3.13 and* $[\hat{\delta}_5\text{-}D]$, *we have:*

$[\text{Var-}\beta_5]$ $\quad \text{Var}(B^{n,j})_t \xrightarrow{P} 0 \quad$ for all $t \geq 0, j \leq d$.

Proof. Let A satisfy 1.2. For each $a > A$ we have $|h(x) - x| \leq (A + a)1_{\{|x|>1/A\}} + |x|1_{\{|x|>a\}}$. We deduce from 3.9 that

$$\text{Var}(B^{n,j})_t \leq (A + a)\nu^n\left([0,t] \times \left\{|x| > \frac{1}{A}\right\}\right) + |x|1_{\{|x|>a\}} * \nu_t^n$$

$$P^n(\text{Var}(B^{n,j})_t > \varepsilon) \leq P^n\left(\nu^n\left([0,t] \times \left\{|x| > \frac{1}{A}\right\}\right) > \frac{\varepsilon}{2(A + a)}\right)$$
$$+ P^n\left(|x|1_{\{|x|>a\}} * \nu_t^n > \frac{\varepsilon}{2}\right)$$

and the result easily follows (choose first a, and then n). □

3.16 **Lemma.** *Under* $[\text{Var-}\beta_5]$ *and* $[\hat{\delta}_5\text{-}D]$, *we have* $[\hat{\gamma}_5\text{-}D] \Leftrightarrow [\hat{\gamma}'_5\text{-}D]$.

Proof. We will prove that $Y_t^n = [X^{n,j}, X^{n,k}]_t - \hat{C}_t^{n,jk}$ (for $j, k \leq d$ fixed) satisfies $Y_t^n \xrightarrow{P} 0$ for all t, and this will yields the result. We have

$$Y_t^n = \sum_{s\le t} \Delta X_s^{n,j}\Delta X_s^{n,k} - \sum_{s\le t} [h^j(\Delta X_s^n) - \Delta B_s^{n,j}][h^k(\Delta X_s^n) - \Delta B_s^{n,k}]$$
$$= \sum_{s\le t} [\Delta X_s^{n,j}\Delta X_s^{n,k} - h^j h^k(\Delta X_s^n)]$$
$$+ \sum_{s\le t} [\Delta B_s^{n,j} h^k(\Delta X_s^n) + \Delta B_s^{n,k} h^j(\Delta X_s^n) - \Delta B_s^{n,j}\Delta B_s^{n,k}]$$

Set $Z_t^n = \sup_{s\le t} |\Delta X_s^n|$, and let A satisfy 1.2; then $|\Delta B^n| \leq A$, hence

3.17 $$|Y_t^n| \leq 2A[\text{Var}(B^{n,j})_t + \text{Var}(B^{n,k})_t] \quad \text{on } \{Z_t^n \leq 1/A\}.$$

By 3.5 we have $Z_t^n \xrightarrow{P} 0$ and it follows immediately from 3.17 and $[\text{Var-}\beta_5]$ that $Y_t^n \xrightarrow{P} 0$. □

Proof of Theorem 3.12. a) By definition of the compensator ν^n, for each $a > 0$ the process $U^n(a) = |x| 1_{\{|x|>a\}} * \nu^n$ is L-dominated by the process $V^n(a) = \sum_{s\leq \cdot} |\Delta X_s^n| 1_{\{|\Delta X_s^n|>a\}}$. Set again $Z_t^n = \sup_{s\leq t} |\Delta X_s^n|$. It follows from Lenglart's inequality I.3.32 that for all $\varepsilon > 0$, $\eta > 0$,

$$P^n(U^n(a)_t \geq \eta) \leq \frac{1}{\eta}[\varepsilon + E^n(\sup_{s\leq t} \Delta V^n(a)_s)] + P^n(V^n(a)_t \geq \varepsilon).$$

Observe that $\{V^n(a)_t > 0\} = \{Z_t^n > a\}$, and that $\Delta V^n(a)_s \leq Z_t^n 1_{\{Z_t^n > a\}}$ for all $s \leq t$. Thus, letting $\varepsilon \downarrow 0$ above yields

$$\begin{aligned} P^n(U^n(a)_t \geq \eta) &\leq \frac{1}{\eta} E^n(Z_t^n 1_{\{Z_t^n>a\}}) + P^n(Z_t^n > a) \\ &\leq \left(\frac{1}{\eta} + \frac{1}{a}\right) E^n(Z_t^n 1_{\{Z_t^n>a\}}) \end{aligned}$$

and thus 3.13 easily follows from 3.14.

b) We assume 3.13. By 3.8a, (i) $\Rightarrow$ $[\hat{\delta}_5\text{-}D]$, and we have seen in the proof of 3.11 that $[\hat{\gamma}'_5\text{-}D] \Rightarrow [\hat{\delta}_5\text{-}D]$. So all conditions (i)–(iv) imply $[\hat{\delta}_5\text{-}D]$, and so $[\text{Var-}\beta_5]$ by Lemma 3.15, which in turn implies $[\text{Sup-}\beta_5]$.

Therefore the equivalences (i) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) follow from Theorem 3.8b, whereas the equivalence (ii) $\Leftrightarrow$ (iii) follows from Lemma 3.16. □

The next corollary is of secondary importance. $B_t^{n,c} = B_t^n - \sum_{s\leq t} \Delta B_s^n$ denotes the "continuous part" of B^n.

3.18 Corollary. *Assume that X is a continuous Gaussian martingale with characteristics* $(0, C, 0)$, *and that each X^n is a local martingale. Let D be a dense subset of* $\mathbb{R}_+$. *Under*

$$[\text{Var}'\text{-}\beta_5] \quad \sup_{s\leq t} |B_s^{n,c}| + \sum_{s\leq t} |\Delta B_s^n| \xrightarrow{P} 0 \quad \text{for all } t \geq 0,$$

the four conditions (i)–(iv) *of 3.12b are equivalent.*

Note that if each X^n is quasi-left-continuous, we have $B^n = B^{n,c}$, in which case $[\text{Var}'\text{-}\beta_5] = [\text{Sup-}\beta_5]$.

Proof. We have (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (i) by 3.8, because $[\text{Var}'\text{-}\beta_5] \Rightarrow [\text{Sup-}\beta_5]$. The inequality 3.17 remains valid if one replaces $\text{Var}(B^{n,j})_t$ and $\text{Var}(B^{n,k})_t$ by $\sum_{s\leq t} |\Delta B_s^n|$, hence under $[\hat{\delta}_5\text{-}D]$ and $[\text{Var}'\text{-}\beta_5]$ we have the equivalence $[\hat{\gamma}_5\text{-}D] \Leftrightarrow [\hat{\gamma}'_5\text{-}D]$, whereas $[\hat{\gamma}'_5\text{-}D] \Rightarrow [\hat{\delta}_5\text{-}D]$ (see the proof of 3.11), so (ii) $\Leftrightarrow$ (iii). □

3.19 Remark. Assume the following property, which is weaker than 3.14

3.20 $\quad$ the sequence $\left(\sup_{s\leq t} |\Delta X_s^n| \,\middle|\, P^n\right)$ is bounded in L^1 for each $t \geq 0$.

Then VI.6.30 implies that $X^n \xrightarrow{\mathcal{L}} X \Rightarrow [\hat{\gamma}'_5\text{-}\mathbb{R}_+]$. But the converse is false, as seen by the following *counter-example*. □

3.21 *Counter-example.* We consider a 1-dimensional semimartingale triangular array scheme $(U_k^n, \sigma_t^n, \mathscr{B}^n)$ (see 2.22) which is rowwise independent, with

$$P^n(U_k^n = n) = 1/n^2,\ P^n\left(U_k^n = -\frac{1}{n(1-1/n^2)}\right) = 1 - 1/n^2$$

and $\sigma_t^n = [nt]$ (integer part of nt), and $\mathscr{F}_k^n = \sigma(U_1^n, \dots, U_k^n)$. Then X^n, defined by 2.24, is clearly a martingale. We have $\sup_{s \le t} |\Delta X_s^n| \le \sum_{1 \le k \le [nt]} |U_k^n|$, and $E^n(\sum_{1 \le k \le [nt]} |U_k^n|) = 2[nt]/n$, hence 3.20 is met. We also have

$$[X^n, X^n]_t = \sum_{1 \le k \le [nt]} (U_k^n)^2 \xrightarrow{P} 0$$

(use Laplace transform), so we have $[\hat{\gamma}'_5\text{-}\mathbb{R}_+]$ with $C = 0$. However we do not have $X^n \xrightarrow{\mathscr{L}} 0$ (on easily sees that $X_t^n \xrightarrow{P} -t$).

It should be noted here that $[\gamma_5\text{-}\mathbb{R}_+]$ and $[\hat{\delta}_5\text{-}\mathbb{R}_+]$ hold, but $[\text{Sup-}\beta_5]$ fails because for all n large enough,

$$B_t^n = \sum_{1 \le k \le [nt]} E^n(U_k^n 1_{\{U_k^n = n\}}) = -\frac{[nt]}{n} \to -t. \qquad \square$$

Now we give a "Lindeberg-Feller" type theorem, and some variants.

3.22 **Theorem.** *Assume that X is a continuous Gaussian martingale with characteristics* $(0, C, 0)$, *that each X^n is a locally square-integrable martingale, and that D is a dense subset of* $\mathbb{R}_+$. *Set*

3.23 $$|x|^2 1_{\{|x| > \varepsilon\}} * \nu_t^n \xrightarrow{P} 0 \quad \text{for all } t \ge 0,\ \varepsilon > 0.$$

There is equivalence between:

(i) *3.23 and* $X^n \xrightarrow{\mathscr{L}} X$.
(ii) *3.23 and* $[\gamma'_5\text{-}D]$.
(iii) *3.23 and* $[\hat{\gamma}'_5\text{-}D]$.
(iv) *3.23 and* $[\gamma_5\text{-}D]$.
(v) *3.23 and* $[\hat{\gamma}_5\text{-}D]$.
(vi) $[\gamma'_5\text{-}D]$ *and* $[\hat{\gamma}'_5\text{-}D]$.
(vii) $[\gamma'_5\text{-}D]$ *and* $X^n \xrightarrow{\mathscr{L}} X$.

One can deduce a number of other results from the above. In particular, the following one is of interest:

3.24 **Corollary.** *Assume that X is a continuous Gaussian martingale with characteristics* $(0, C, 0)$, *that each X^n is a locally square-integrable martingale* (recall that $X_0^n = 0$) *and that D is a dense subset of* $\mathbb{R}_+$. *Then if 3.23 holds, there is equivalence between:*

(i) $X^n \xrightarrow{\mathscr{L}} X$.
(ii) $[\hat{\gamma}'_5\text{-}D]$: $[X^n, X^n]_t \xrightarrow{P} C_t$ *for all* $t \in D$.
(iii) $[\gamma'_5\text{-}D]$: $\langle X^n, X^n \rangle_t \xrightarrow{P} C_t$ *for all* $t \in D$.

Proof of Theorem 3.22. a) 3.23 obviously yields $[\hat{\delta}_5\text{-}D]$ and 3.13, so (i)$\Leftrightarrow$(iii)$\Leftrightarrow$(iv)$\Leftrightarrow$(v) follows from Theorem 3.12.

b) We have

$$|x^j| 1_{\{|x|>a\}} * \nu^n \leq \frac{1}{a} |x^j|^2 * \nu^n \leq \frac{1}{a} \tilde{C}'^{n,jj},$$

so we easily deduce that $[\gamma'_5\text{-}D]$ implies 3.13: then (vi)$\Leftrightarrow$(vii) follows from Theorem 3.12 again.

c) Recalling that $B'^n = 0$ (because X^n is a local martingale), we deduce from 2.12 and 2.13 that

3.25 $$\begin{cases} B^n = (h(x) - x) * \nu^n \\ \tilde{C}^{n,jj} - \tilde{C}'^{n,jj} = [h^j(x)^2 - (x^j)^2] * \nu^n - \sum_{s \leq \cdot} (\Delta B_s^{n,j})^2. \end{cases}$$

Therefore 3.23 obviously implies firstly [Var-β_5] (see 3.15), and then $\tilde{C}_t^{n,jj} - \tilde{C}_t'^{n,jj} \xrightarrow{P} 0$ for all $t \in \mathbb{R}_+$: therefore under 3.23, $[\gamma'_5\text{-}D] \Leftrightarrow [\gamma_5\text{-}D]$, and so (ii)$\Leftrightarrow$(iv).

d) Since (ii) + (iii) $\Rightarrow$ (vi) is obvious, it remains to prove that (vi) $\Rightarrow$ 3.23.

So we assume (vi). We have seen in (b) that 3.13 holds, so Theorem 3.12 implies $[\hat{\delta}_5\text{-}D]$ and $[\gamma_5\text{-}D]$, and 3.15 implies [Var-β_5]. Then we deduce from 3.25 and $[\gamma_5\text{-}D]$ + $[\gamma'_5\text{-}D]$ + [Var-β_5] that

3.26 $$[(h^j(x))^2 - (x^j)^2] * \nu_t^n \xrightarrow{P} 0 \quad \text{for all } t \in D.$$

Now the truncation function h is arbitrary, and so we can choose one that satisfies $|h^j(x)| \leq |x^j|$ for all x; if A satisfies 1.2, we then have $|x^j|^2 1_{\{|x|>A\}} \leq (x^j)^2 - (h^j(x))^2$ for all x, and we deduce from 3.26 that

$$|x|^2 1_{\{|x|>A\}} * \nu_t^n \xrightarrow{P} 0 \quad \text{for all } t \in D,$$

and by $[\hat{\delta}_5\text{-}D]$,

$$|x|^2 1_{\{|x|>\varepsilon\}} * \nu_t^n \leq |x|^2 1_{\{|x|>A\}} * \nu_t^n + A^2 \nu^n([0,t] \times \{|x| > \varepsilon\}) \xrightarrow{P} 0$$

for all $t \in D$, $\varepsilon > 0$: thus we have 3.23. □

§ 3c. Central Limit Theorem for Triangular Arrays

In this short subsection, we (partly) reformulate some of the previous result for triangular arrays.

We consider a d-dimensional triangular array scheme in the sense of 2.22: $(U_k^n)_{k \geq 1}$ and $(\sigma_t^n)_{t \geq 0}$ are defined on the discrete-time basis $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, (\mathscr{F}_p^n)_{p \in \mathbb{N}}, P^n)$, and 2.23 holds. We also assume:

3.27 for each n, $(U_k^n)_{k \geq 1}$ is a martingale difference, i.e. $E^n(U_k^n | \mathscr{F}_{k-1}^n) = 0$.

Moreover we introduce the following conditions:

3.28 $$\lim_{a\uparrow\infty} \limsup_n P^n\left(\sum_{1\leq k\leq \sigma_t^n} E^n(|U_k^n| 1_{\{|U_k^n|>a\}} | \mathscr{F}_{k-1}^n) > \eta\right) = 0$$

for all $t \geq 0, \eta > 0$.

3.29 the sequence $\left(\sup_{1\leq k\leq\sigma_t^n} |U_k^n| | P^n\right)$ is uniformly integrable for all $t \geq 0$

3.30 $\sup_{1\leq k\leq\sigma_t^n} |U_k^n| \xrightarrow{P} 0$ for all $t \geq 0$.

3.31 (*Conditional Lindeberg condition*): for all $\varepsilon > 0$, $t \geq 0$ we have

$$\sum_{1\leq k\leq\sigma_t^n} E^n(|U_k^n|^2 1_{\{|U_k^n|>\varepsilon\}} | \mathscr{F}_{k-1}^n) \xrightarrow{P} 0.$$

We have introduced Condition [γ_6-D] in 2.29, and we set:

[$\hat{\gamma}'_6$-D] $\sum_{1\leq k\leq\sigma_t^n} U_k^{n,j} U_k^{n,l} \xrightarrow{P} C_t^{jl}$ for all $t \in D$

[γ'_6-D] $\sum_{1\leq k\leq\sigma_t^n} E^n(U_k^{n,j} U_k^{n,l} | \mathscr{F}_{k-1}^n) \xrightarrow{P} C_t^{jl}$ for all $t \in D$

[$\hat{\delta}_6$-D] $\sum_{1\leq k\leq\sigma_t^n} P^n(|U_k^n| > \varepsilon | \mathscr{F}_{k-1}^n) \xrightarrow{P} 0$ for all $t \in D$, $\varepsilon > 0$.

Then in this setting, Theorem 3.12 reads as

3.32 **Theorem.** *Let X be a continuous Gaussian martingale with characteristics $(0, C, 0)$, and consider a semimartingale triangular array with 3.27, and let D be a dense subset of $\mathbb{R}_+$.*

a) [$\hat{\delta}_6$-D] $\Leftrightarrow$ 3.30.

b) *3.29 implies 3.28.*

c) *Let $X_t^n = \sum_{1\leq k\leq\sigma_t^n} U_k^n$. Then under 3.28 we have $X^n \xrightarrow{\mathscr{L}} X$ if and only if one of the following conditions holds*: [$\hat{\gamma}'_6$-D], *or* [$\hat{\gamma}_6$-D] + [$\hat{\delta}_6$-D], *or* [γ_6-D] + [$\hat{\delta}_6$-D].

As for Corollary 3.24, it becomes:

3.33 **Theorem.** *Under the same hypotheses than in Theorem 3.32, and if Lindeberg's condition 3.31 holds, then $X^n \xrightarrow{\mathscr{L}} X$ if and only if [$\hat{\gamma}'_6$-D] holds, and also if and only if [γ'_6-D] holds.*

3.34 **Remark.** The reader will write by himself the translation of Theorem 3.24. Other results in the same vein can be found in [57] and [211].

§ 3d. Convergence of Point Processes

In this subsection we assume that X, in addition to being a PII without fixed time of discontinuity, is also a point process: that is, X is a *Poisson process* in the sense of I.3.26. We call A its compensator: by I.3.27, A is non-random and continuous, and $A_t = E(X_t)$ (i.e., A is the *intensity* of X). A simple computation shows that the characteristics (B, C, ν) of X are

3.35 $$B_t = h(1)A_t, \quad C_t = 0, \quad \nu(dt, dx) = dA_t \otimes \varepsilon_1(dx)$$

(ε_1 = Dirac measure at point 1).

3.36 **Theorem.** *We suppose that X is a Poisson process with intensity A; we suppose that each X^n is a point process with compensator A^n; let $D \subset \mathbb{R}_+$.*

a) *The following condition implies $X^n \xrightarrow{\mathscr{L}(D)} X$:*

3.37D $$A_t^n \xrightarrow{P} A_t \quad \text{for all } t \in D.$$

b) *If moreover D is dense in $\mathbb{R}_+$, 3.37D implies that $X^n \xrightarrow{\mathscr{L}} X$.*

Proof. a) Like in 3.35, the characteristics of X^n are

3.38 $$B_t^n = h(1)A_t^n, \quad C_t^n = 0, \quad \nu^n(dt, dx) = dA_t^n \otimes \varepsilon_1(dx).$$

Hence [γ_5-D] is satisfied, and [β_5-D] and [$\delta_{5,1}$-D] are both clearly equivalent to 3.37D, so the result follows from Theorem 2.4.

b) Assume that D is dense. Then the result follows from (a) and from Theorem VI.3.37 (we could also apply 2.14, because we have [Sup-$\delta_{5,1}$], which, here, clearly implies [Sup-β_5]). □

3.39 **Remark.** Contrarily to what happens in § 3a, we cannot hope here for a converse to statement (b). Here is a *counter-example.* We call $(T_k)_{k \geq 1}$ the successive jumps of a Poisson process X, defined on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. We define X^n on the same stochastic basis, by

$$X_t^n = \sum_{k \geq 1} 1_{\{T_k + 1/n \leq t\}}.$$

Then obviously $X_{\cdot}^n(\omega) \to X_{\cdot}(\omega)$ for the Skorokhod topology, for all $\omega \in \Omega$, and *a*-fortiori $X^n \xrightarrow{\mathscr{L}} X$. However X^n is predictable with respect to $\mathbf{F}$, hence $A^n = X^n$ and 3.37D fails.

The following should also be noticed: suppose that we choose h such that $h(1) = 0$. Then by 3.35 and 3.38 we have [Sup-β_5] and [γ_5-D]. But we do not have [$\delta_{5,1}$-D], although $X^n \xrightarrow{\mathscr{L}} X$.

It is nevertheless possible to give necessary and sufficient conditions for $X^n \xrightarrow{\mathscr{L}} X$ in terms of the compensators; but these conditions are not of the type 3.37D, but rather in terms of "weak $\sigma(L^1, L^\infty)$ convergence" of A^n to A: see [104]. □

Sometimes one encounters a sequence of processes X^n that are not point processes, but which converge to a point process. Here is a result in this direction (we state only the functional convergence result).

3.40 **Proposition.** *We suppose that X is a Poisson process with intensity A, and that each X^n is a 1-dimensional semimartingale. We also suppose that the truncation function h satisfies $h(1) = 0$. Under the conditions*

(i) $\sup_{s \le t} |B_s^n| \xrightarrow{P} 0$ for all $t \geq 0$;

(ii) $\tilde{C}_t^n \xrightarrow{P} 0$ for all $t \geq 0$;

(iii) $\nu^n([0,t] \times \{x: |x-1| > \varepsilon, |x| > \varepsilon\}) \xrightarrow{P} 0$ for all $t \geq 0, \varepsilon > 0$

(iv) $\nu^n([0,t] \times \{x: |x-1| \le \varepsilon\}) \xrightarrow{P} A_t$ for all $t \geq 0$, $\varepsilon \in (0,1)$,

we have $X^n \xrightarrow{\mathcal{L}} X$.

Proof. Since $h(1) = 0$ we have $B = 0$ and $\tilde{C} = 0$: thus (i) $\Rightarrow$ [Sup-β_5] and (ii) $\Rightarrow$ [γ_5-$\mathbb{R}_+$]. For each $\varepsilon \in (0,1)$, set

$$f_\varepsilon(x) = 1_{\{|x-1| \le \varepsilon\}}, \quad f_\varepsilon'(x) = 1_{\{|x-1|>\varepsilon, |x|>\varepsilon\}}$$

so that (iii) and (iv) imply

3.41 $$f_\varepsilon * \nu_t^n \xrightarrow{P} A_t, \quad f_\varepsilon' * \nu_t^n \xrightarrow{P} 0$$

for all $t \geq 0$. Let $g \in C_1(\mathbb{R}^d)$ and $\eta > 0$. There exists $\varepsilon \in (0,1)$ such that $|g(x) - g(1)| \le \eta$ if $|x-1| \le \varepsilon$, and $g(x) = 0$ if $|x| \le \varepsilon$. Then

$$|g * \nu_t^n - g(1) f_\varepsilon * \nu_t^n| \le \eta f_\varepsilon * \nu_t^n + \|g\| f_\varepsilon' * \nu_t^n.$$

Then 3.41 and the arbitrariness of $\eta > 0$ show that $|g * \nu_t^n - g(1) f_\varepsilon * \nu_t^n| \xrightarrow{P} 0$ and 3.41 again yields that $g * \nu_t^n \xrightarrow{P} g(1) A_t$; since $g(1) A_t = g * \nu_t$, we have [$\delta_{5,1}$-$\mathbb{R}_+$], and the result follows from Theorem 2.14. □

Application: convergence of empirical processes to a Poisson process. Here we consider a sequence $(Z_i)_{i \geq 1}$ of i.i.d. random variables with common distribution G on $(0, \infty]$. According to II.3.34 we set

3.42 $$\bar{Y}_t^n = \sum_{1 \le i \le n} 1_{\{Z_i \le t/n\}}.$$

The following is weaker than Theorem V.4.41, but in a sense its proof is simpler.

3.43 **Theorem.** *Assume that G admits a density g which is continuous from the right at 0. Then $\bar{Y}^n$ converges in law to a Poisson process with intensity $tg(0)$* (cf. I.3.26).

Proof. Recall from II.3.34 that the compensator of $\bar{Y}^n$, relative to the filtration generated by $\bar{Y}^n$ itself, is

$$\bar{A}_t^n = \int_0^t (1 - \bar{Y}_s^n/n) \frac{g(s/n)}{G([s/n, \infty])} ds.$$

Then $\dfrac{g(s/n)}{G([s/n, \infty])} \to g(0)$ uniformly on $[0,t]$ as $n \uparrow \infty$, and $\dfrac{1}{n} E(\bar{Y}_s^n = G([0, s/n]) \to 0$. Thus the Lebesgue convergence theorem yields $\bar{A}_t^n \to g(0)t$ in L^1, and so in measure, and the result follows from Theorem 3.36. □

3.44 **Remark.** There is of course a more elementary way for proving this theorem, which consists in proving the finite-dimensional convergence by hand, and then in applying VI.3.37b. □

§ 3e. Normed Sums of I.I.D. Semimartingales

1. This subsection is devoted to studying the behaviour of normed sums of a sequence of *i.i.d. semimartingales* Y^n, satisfying $Y_0^n = 0$. For simplicity we consider only the 1-dimensional case.

A simple way of constructing i.i.d. semimartingales is as follows.

3.45 *Hypothesis*: We have a sequence $(\mathscr{B}^n, Y^n)_{n\geq 1}$ of identical copies of the pair $(\mathscr{B}, Y)$, where $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$ is a stochastic basis, and Y is a 1-dimensional semimartingale on $\mathscr{B}$ with $Y_0 = 0$. We define the basis $\tilde{\mathscr{B}} = (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbf{F}}, \tilde{P})$ as being the tensor product of all $\mathscr{B}^n$'s, as usual, and any variable or process or σ-field defined on $\mathscr{B}^n$ is naturally extended to $\tilde{\mathscr{B}}$ and the extension is represented by the same symbol (example: $Y^n, \mathscr{F}_t^n, \ldots$). □

Our first theorem below fully deserves the name of "central limit theorem" for processes. It is simple enough, but we shall see that unfortunately it does not extend easily to more general semimartingale situations.

3.46 **Theorem.** *Suppose that we have 3.45, where Y is a local martingale for which the function $C_t = E(Y_t^2)$ is finite-valued and continuous* (it is automatically increasing, with $C_0 = 0$). *Then*

3.47 $$X^n = \frac{1}{\sqrt{n}} \sum_{1\leq p\leq n} Y^p$$

converges in law to a Wiener process X with characteristics $(0, C, 0)$.

The continuity of C is equivalent to the absence of fixed time of discontinuity for Y, although Y is not necessarily a PII.

Proof. a) We begin with some general remarks, which pertain to 3.45. Call (B^Y, C^Y, ν^Y) the characteristics of Y, and $\tilde{C}^Y$, $\tilde{C}'^Y$ (when it exists), $A^Y(u)$ the processes associated by 1.3, 2.13 or 1.8. It is natural to denote by B^{Y^n}, C^{Y^n}, ν^{Y^n}, $\tilde{C}^{Y^n}$, $\tilde{C}'^{Y^n}$, $A^{Y^n}(u)$ the n^{th} copy, and of course $(B^{Y^n}, C^{Y^n}, \nu^{Y^n})$ are the characteristics of Y^n on $\mathscr{B}^n$. That, in particular, implies that any "function" f of the characteristics satisfies:

3.48 $$\{f(B^{Y^n}, C^{Y^n}, \nu^{Y^n})\}_{n\geq 1} \text{ are i.i.d. random variables.}$$

Now, by 3.45 it readily follows that

3.49 $\begin{cases} \cdot \text{ Any (local) martingale on } \mathscr{B}^n \text{ is a (local) martingale on } \tilde{\mathscr{B}}. \\ \cdot \text{ If } n \neq m, \text{ the product of a (local) martingale on } \mathscr{B}^n \text{ and a (local)} \\ \quad \text{martingale on } \mathscr{B}^m \text{ is a (local) martingale on } \tilde{\mathscr{B}}. \end{cases}$

We easily deduce that any semimartingale on $\mathscr{B}^n$ (e.g. Y^n) is a semimartingale on $\tilde{\mathscr{B}}$. Moreover, 3.49 and the martingale characterization of the characteristics (see II.2.21) immediately imply that

3.50 $$(B^{Y^n}, C^{Y^n}, \nu^{Y^n}) \text{ are the characteristics of } Y^n \text{ on } \tilde{\mathscr{B}}.$$

b) Now we proceed to the proof of the theorem itself. Everything in the remainder of the proof is defined on the basis $\tilde{\mathscr{B}}$. By 3.49, each Y^n is a martingale and $\langle Y^n, Y^m \rangle = 0$ if $n \neq m$. Hence X^n is a martingale and

$$\langle X^n, X^n \rangle = \frac{1}{n} \sum_{1 \le p \le n} \langle Y^p, Y^p \rangle = \frac{1}{n} \sum_{1 \le p \le n} \tilde{C}'^{Y^p}.$$

Moreover, $E(\tilde{C}_t'^{Y^p}) = C_t$ for each p, and by 3.48 the variables $(\tilde{C}_t'^{Y^p})_{p \ge 1}$ are i.i.d. Hence the strong law of large numbers yields [γ_5'-$\mathbb{R}_+$].

Call μ^{X^n} and μ^{Y^n} the measures associated with the jumps of X^n and Y^n by II.1.16. Since Y has no fixed time of discontinuity, the Y^n's have P-a.s. no common jump, hence for every measurable function $g \ge 0$ on $\mathbb{R}$ we have

$$g * \mu^{X^n} = \sum_{1 \le p \le n} g\left(\frac{x}{\sqrt{n}}\right) * \mu^{Y^p} \quad \tilde{P}\text{-a.s.}$$

This relationship carries over to the compensators. Thus

3.51 $$g * \nu^n = \sum_{1 \le p \le n} g\left(\frac{x}{\sqrt{n}}\right) * \nu^{Y^p} \quad \tilde{P}\text{-a.s.}$$

where ν^n is the compensator of μ^{X^n}. In particular,

$$|x|^2 1_{\{|x| > \varepsilon\}} * \nu_t^n = \frac{1}{n} \sum_{1 \le p \le n} |x|^2 1_{\{|x| > \varepsilon\sqrt{n}\}} * \nu_t^{Y^p}$$

$\tilde{P}$-a.s., and we deduce that

3.52 $$\tilde{E}(|x|^2 1_{\{|x| > \varepsilon\}} * \nu_t^n) = E(|x|^2 1_{\{|x| > \varepsilon\sqrt{n}\}} * \nu_t^Y).$$

Now, $|x|^2 1_{\{|x| > \varepsilon\sqrt{n}\}} * \nu_t^Y \downarrow 0$ as $n \uparrow \infty$, and these variables remain smaller than $\langle Y, Y \rangle_t$, which is integrable. Hence 3.52 tend to 0 as $n \uparrow \infty$, which implies 3.23, and the result follows from Theorem 3.22. □

2. Now we examine the general case. We assume 3.45, and we consider the normed sums (and the associated processes $A^n(u)$, see 1.8):

3.53 $$X^n = \alpha_n \sum_{1 \le p \le n} Y^p.$$

3.54 **Lemma.** *Suppose that Y has no fixed time of discontinuity. Then*

$$A^n(u) = \sum_{1 \le p \le n} A^{Y^p}(\alpha_n u).$$

Proof. Let $m \neq n$; let Z^n (resp. Z^m) be a semimartingale on $\mathscr{B}^n$ (resp. $\mathscr{B}^m$), with the associated processes $A^{Z^n}(u)$ (resp. $A^{Z^m}(u)$). Assume also that Z^n and Z^m have no common jump.

Set $V^n = \exp iuZ^n$ and $V^m = \exp iuZ^m$. Then Ito's formula yields $V^{n,c} = iuV^n_- \cdot Z^{n,c}$ and $V^{m,c} = iuV^m_- \cdot Z^{m,c}$; since $Z^{n,c}Z^{m,c}$ is a local martingale by 3.49, we have $\langle Z^{n,c}, Z^{m,c}\rangle = 0$, and so $\langle V^{n,c}, V^{m,c}\rangle = 0$. Moreover, V^n and V^m have no common jump, hence we deduce that $[V^n, V^m] = 0$ (see I.4.53), and thus

$$V^n V^m = V^n_- \cdot V^m + V^n_- \cdot V^m.$$

But $V^n - V^n_- \cdot A^{Z^n}(u)$ and $V^m - V^m_- \cdot A^{Z^m}(u)$ are local martingales, hence $V^n V^m - (V^n V^m)_- \cdot (A^{Z^n}(u) + A^{Z^m}(u))$ is a local martingale, and from II.2.42 it follows that $A^{Z^n + Z^m}(u) = A^{Z^n}(u) + A^{Z^m}(u)$.

Now, the hypothesis implies that $Y^1, \dots, Y^n$ have no common jump. Then the previous result and an induction easily yield that if $U^n = Y^1 + \dots + Y^n$, the process $A^{U^n}(u)$ associated to U^n is $A^{U^n}(u) = \sum_{1 \le p \le n} A^{Y^p}(u)$, that is

$$e^{iuU^n} - (e^{iuU^n_-}) \cdot \left(\sum_{1 \le p \le n} A^{Y^p}(u) \right) \quad \text{is a local martingale.}$$

Applying this to $u = \alpha_n u'$ gives

$$e^{iu'X^n} - (e^{iu'X^n_-}) \cdot \left(\sum_{1 \le p \le n} A^{Y^p}(\alpha_n u') \right) \quad \text{is a local martingale,}$$

and we deduce the claimed formula. □

In order to obtain that $X^n \xrightarrow{\mathscr{L}} X$, where X is a suitable PII, it remains to compute the characteristics of X^n by "inverting" the first formula in 1.8, and to check that [Sup-β_5] + [γ_5-D] + [$\delta_{5,1}$-D] holds.... needless to say, this is impossible to do in general!

Nevertheless, we presently give an example for which this programme may actually be fulfilled. It is taken from [63].

3.55 *Hypothesis*: $Y = H \cdot Z$, where H is a *bounded* predictable process on $\mathscr{B}$, and Z is a *PII without fixed time of discontinuity* on $\mathscr{B}$, with characteristics $(\hat{B}, \hat{C}, \hat{\nu})$ and the function $\hat{A}(u)$ associated by 1.8 (then $\exp \hat{A}(u)_t = E(\exp iuZ_t)$). Then, according to II.2.41, we have the factorization

$$\hat{A}(u) = a(u) \cdot A$$

where A is an increasing continuous function and $u \rightsquigarrow a(u)_s$ is (for each $s \geq 0$) of "Lévy-Khintchine" type. In addition, we suppose that

(i) Z is symmetrical (i.e. $\mathscr{L}(Z) = \mathscr{L}(-Z)$), which amounts to saying that one can choose a such that $a(u) = a(-u)$.

(ii) For some $\alpha \in (0, 2]$, we have for all $u \in \mathbb{R}$ and $s \geq 0$:

3.56 $$na_s(un^{-1/\alpha}) \to -|u|^\alpha. \qquad \square$$

Note that (i) and (ii) are both satisfied when Z is a symmetric stable process of order α, in which case $A(u)_t = -|u|^\alpha A_t$.

Of course, on each $\mathscr{B}^n$ is defined a copy (Z^n, H^n) of the pair (Z, H), and $Y^n = H^n \cdot Z^n$.

3.57 **Theorem.** *Under 3.55, the processes*

$$X^n = n^{-1/\alpha} \sum_{1 \leq p \leq n} Y^p$$

converge in law to a PII X which is a (non-homogeneous) symmetric stable process of order α, with

3.58 $$E(e^{iuX_t}) = \exp -|u|^\alpha \int_0^t \delta(s)\, dA_s,$$

where $\delta(s) = E(|H_s|^\alpha)$.

We begin with a lemma.

3.59 **Lemma.** *Under 3.55, we have $A^Y(u) = a(uH) \cdot A$* (in other words, $A^Y(u)_t = \int_0^t a_s(uH_s)\, dA_s$.

Proof. We write $U \sim V$ if $U - V$ is a local martingale. We recall that $Y^c = H \cdot Z^c$ and that $Z \sim \hat{B} + \sum_{s \leq \cdot} (\Delta Z_s - h(\Delta Z_s))$. If $V = e^{iuY}$, Ito's formula yields

$$V = 1 + iuV_- \cdot Y - \frac{u^2}{2} V_- \cdot \langle Y^c, Y^c \rangle + \sum_{s \leq \cdot} V_{s-}(e^{iu\Delta Y_s} - 1 - iu\Delta Y_s)$$

$$\sim iuHV_- \cdot \hat{B} - \frac{u^2}{2} H^2 V_- \cdot \hat{C} + \sum_{s \leq \cdot} V_{s-}(e^{iuH_s \Delta Z_s} - 1 - iuH_s h(\Delta Z_s))$$

(because $\langle Y^c, Y^c \rangle = H^2 \cdot \langle Z^c, Z^c \rangle = H^2 \cdot \hat{C}$). By definition of $\hat{\nu}$ we obtain

$$V \sim iuHV_- \cdot \hat{B} - \frac{(uH)^2}{2} V_- \cdot \hat{C} + V_-(e^{iuHx} - 1 - iuHh(x)) * \hat{\nu}.$$

Then, using II.2.41 and the definition of $a(u)$, we obtain $V \sim a(Hu) V_- \cdot A$. We deduce from II.2.42 that $a(Hu) V_- \cdot A = V_- \cdot A^Y(u)$ and $A^Y(u) = a(Hu) \cdot A$. $\square$

Proof of Theorem 3.57. Firstly, the function $u \rightsquigarrow na_s(un^{-1/\alpha})$ is the Log of the characteristic function of an infinitely divisible distribution: hence the convergence in 3.56 is uniform over compact subsets, i.e.

3.60 $$\sup_{|u| \leq \theta} |na_s(un^{-1/\alpha}) + |u|^\alpha| \to 0 \text{ for all } \theta > 0.$$

Secondly, 3.54 and 3.59 give

3.61 $$A^n(u) = \sum_{1 \leq p \leq n} a(uH^p n^{-1/\alpha}) \cdot A.$$

Therefore

3.62 $$\begin{aligned} A^n(u)_t = {} & \int_0^t \frac{1}{n} \sum_{1 \leq p \leq n} [na_s(H_s^p un^{-1/\alpha}) + |H_s^p u|^\alpha] \, dA_s \\ & - |u|^\alpha \frac{1}{n} \sum_{1 \leq p \leq n} \int_0^t |H_s^p|^\alpha \, dA_s. \end{aligned}$$

Consider the right-hand side of 3.62: we deduce from 3.60 and from the uniform boundedness of all H^p's that the first term tends to 0; next, the variables $\{\int_0^t |H_s^p|^\alpha \, dA_s\}_{p \geq 1}$ are i.i.d. with mean value $\int_0^t \delta(s) \, dA_s$: hence the law of large numbers implies that the second term converges a. s. to $-|u|^\alpha \int_0^t \delta(s) \, dA_s$. Thus we have 2.37 with $D = \mathbb{R}_+$ and

$$A(u)_t = -|u|^\alpha \int_0^t \delta(s) \, dA_s,$$

and $A(u)$ is the function associated by 1.8 to the characteristics of the PII characterized by 3.58.

In virtue of Theorem 2.36b, it remains to prove that [Sup-β_5] holds. For this, we choose a truncation function h which is odd. We have $a(u) = a(-u)$ by 3.55(i), hence $A^n(u) = A^n(-u)$ by 3.61. But $A^n(u)_t$ is given by formula II.2.45, with $b = B_t^n$, $c = C_t^n$ and $F = \nu^n([0, t] \times \cdot)$, while $A^n(-u)_t$ is given by the same formula, with $b' = -b$, $c' = c$ and $F' = \check{F}$ (the measure symmetrical to F), because $h(x) = -h(-x)$. Therefore the uniqueness in II.2.44 yields $b = b' = 0$; that is, we have $B_t^n = 0$ for all $t \geq 0$, $n \geq 1$. Therefore [Sup-β_5] is satisfied (note that $B = 0$ as well, because $A(-u) = A(u)$. □

3.63 **Remark.** As a matter of fact, the result of [63] is more general for three reasons:

1) they admit a PII Z with fixed times of discontinuity;

2) they admit unbounded H;

3) more importantly, they assume a much weaker condition than 3.55(ii), namely that

$$\sup_{s \leq t} |nA(un^{-1/\alpha})_s + |u|^\alpha A_s| \to 0 \quad \text{for all } u \in \mathbb{R}, t \geq 0.$$

By VII.4.43, this is equivalent to saying that $n^{-1/\alpha}\sum_{1\le p\le n} Z^p \xrightarrow{\mathscr{L}} Z'$, where Z' is a symmetric stable process with $E(\exp iuZ'_t) = \exp -|u|^\alpha A_t$. □

§ 3f. Limit Theorems for Functionals of Markov Processes

(Functional) central limit theorems for additive functionals of recurrent Markov chains or processes are quite old: see for instance the book [12].

Here, we first give a result relative to continuous-time Markov processes. It is far from being optimal, but it is simple enough, and it provides an interesting and non-trivial counter-example to the necessity of conditions (ii) or (iii) of 3.8 when $X^n \xrightarrow{\mathscr{L}} X$ (X = standard Wiener process), when [Sup-β_5] fails.

So we consider a *Markov process* $(\Omega, \mathscr{F}, \mathscr{F}_t, \theta_t, Y_t, P_x)$: we use the standard notation of Blumenthal and Getoor [14], $(\theta_t)_{t\ge 0}$ is the shift semi-group and $Y_{t+s} = Y_t\circ\theta_s$ for all $t, s \ge 0$; the process $Y = (Y_t)_{t\ge 0}$ itself is right-continuous, with values in a topological space E.

3.64 *Hypothesis*: There is an invariant *probability* measure μ, and the invariant σ-field under the semi-group $(\theta_t)_{t\ge 0}$ is P_μ-trivial (where $P_\mu = \int \mu(dx)P_x$) (see for example [19]). □

This is an assumption of ergodicity, or rather, ergodicity within a subclass, for our Markov process.

3.65 **Theorem.** *Assume 3.64, and let f be a bounded Borel function on E which is of the form $f = Ag$, where A is the weak generator and g and g^2 belong to the domain of A* (so in particular g is bounded). *Then the processes*

3.66
$$X^n_t = \frac{1}{\sqrt{n}}\int_0^{nt} f(Y_s)\,ds,$$

under the measure P_μ, converge in law to $\sqrt{\beta}W$, where W is a standard Wiener process, and

3.67
$$\beta = -2\int g(x)Ag(x)\mu(dx).$$

X^n is obviously a semimartingale (under any probability measure) with characteristics $B^n = X^n$, $C^n = 0$, $\nu^n = 0$: then $[\hat\delta_5\text{-}D]$ is met. However, since $X^n \xrightarrow{\mathscr{L}} X = \sqrt{\beta}W$ (under P_μ), we have neither [Sup-β_5], nor $[\hat\gamma_5\text{-}D] = [\gamma_5\text{-}D]$, unless $\beta = 0$. The reason for which this is a counter-example to the necessity of [Sup-β_5] + $[\gamma_5\text{-}D]$ + $[\hat\delta_5\text{-}D]$ for having $X^n \xrightarrow{\mathscr{L}} X$ is simple indeed: although the X^n's converge to a martingale, they themselves are "as far as possible" to being martingales, since they are continuous with finite variation.

We begin with a lemma, well known to those familiar with Markov processes.

3.68 **Lemma.** *Assume that g and g^2 belong to the domain of the infinitesimal generator, and set $\Gamma(g,g) = Ag^2 - 2gAg$. Then there is a locally square-integrable martingale M relative to P_μ, such that*

3.69 $$\text{for all } t \in \mathbb{R}_+,\ M_t = g(Y_t) - g(Y_0) - \int_0^t Ag(Y_s)\,ds \quad P_\mu\text{-a.s.}$$

3.70 $$\langle M, M \rangle_t = \int_0^t \Gamma(g,g)(Y_s)\,ds.$$

In fact, one can prove that if the process Y is *strong Markov*, then $g(Y)$ is P-a.s. càdlàg, so $g(Y_t) - g(Y_0) - \int_0^t Ag(Y_s)\,ds$ itself is a martingale (we will not use this result below). Also, all these claims are valid relative to P_η, where η is an arbitrary initial measure.

Proof. a) Set $\tilde{M}_t = g(Y_t) - g(Y_0) - \int_0^t Ag(Y_s)\,ds$. Then if $(P_t)_{t \geq 0}$ is the transition semi-group of the process Y, the Markov property yields

$$E_\mu(\tilde{M}_{t+s} - \tilde{M}_t | \mathscr{F}_t) = E_\mu(g(Y_{t+s}) - g(Y_t) - \int_t^{t+s} Ag(Y_u)\,du | \mathscr{F}_t)$$

$$= P_s g(Y_t) - g(Y_t) - \int_0^s P_u Ag(Y_t)\,du$$

which equals 0 by Kolmogorov's equation. Moreover $|\tilde{M}_t| \leq 2K + K't$, if $K = \sup|g|$ and $K' = \sup|Ag|$. Therefore $\tilde{M}$ admits a modification M which is a (càdlàg) martingale, with $|M_t| \leq 2K + K't$ as well; hence in particular M is locally square-integrable, and 3.69 holds.

b) We set $F_t = \int_0^t Ag(Y_s)\,ds$ and $G = M + g(Y_0) + F$, so $G_t = g(Y_t)$ P_μ-a.s. Similarly, we associate to $g' = g^2$ the martingale M' by 3.69, and $F'_t = \int_0^t Ag^2(Y_s)\,ds$ and $G' = M' + g^2(Y_0) + F'$. Then $G'_t = g^2(Y_t) = G_t^2$ P_μ-a.s., and since G and G' are càdlàg we have $G' = G^2$ up to a P_μ-evanescent set. Then the integration by parts formula I.4.45 yields

$$G' = G^2 = G_0^2 + 2G_- \cdot M + 2G_- \cdot F + [G, G].$$

Moreover F is continuous with finite variation, hence $[G, G] = [M, M]$, whereas $[M, M] - \langle M, M \rangle$ is a local martingale. The uniqueness of the canonical decomposition $G' = G'_0 + M' + F'$ yields

$$F' = 2G_- \cdot F + \langle M, M \rangle,$$

and thus

$$\langle M, M \rangle_t = F'_t - 2G_- \cdot F_t = \int_0^t Ag^2(Y_s)\,ds - 2\int_0^t G_s Ag(Y_s)\,ds.$$

Since for all s, $G_s = g(Y_s)$ P_μ-a.s., we deduce from Fubini's theorem that $\int_0^t G_s Ag(Y_s)\,ds = \int_0^t g(Y_s) Ag(Y_s)\,ds$ P_μ-a.s., and so we obtain 3.70. $\square$

Proof of Theorem 3.65. We consider the martingale M associated with g in the previous lemma.

Put $M_t^n = M_{nt}/\sqrt{n}$ and $\mathscr{F}_t^n = \mathscr{F}_{nt}$; clearly M^n is a martingale on the basis $\mathscr{B}^n = (\Omega, \mathscr{F}, (\mathscr{F}_t^n), P)$, with a bracket equal to

3.71 $$\langle M^n, M^n \rangle_t = \frac{1}{n} \int_0^{nt} \Gamma(g,g)(Y_s)\, ds.$$

We will now apply 3.24 to the martingales M^n, with $X = \sqrt{\beta} W$ a Wiener process with characteristics $(0, \beta t, 0)$ and where $\beta = \int \Gamma(g,g)(x)\mu(dx)$: note that since $\langle M, M \rangle$ is increasing, the set of all (ω, s) such that $\Gamma(g,g)(Y_s(\omega)) < 0$ is $ds \otimes P_\mu$-negligible; since μ is an invariant measure, it readily follows that $\beta \geq 0$ (as a matter of fact, one can prove that $\Gamma(g,g) \geq 0$ identically). Moreover, $\int Ag^2(x)\mu(dx) = 0$ because μ is invariant, so β is also given by 3.67.

If $K = \sup|g|$ we have $|\Delta M^n| \leq K/\sqrt{n}$, hence the third characteristic of M^n does not charge the set $\{(t,x): t \geq 0,\ |x| > K/\sqrt{n}\}$ and it clearly follows that the sequence (M^n) satisfies 3.23. Then by 3.22 we will obtain $M^n \xrightarrow{\mathscr{L}} X$ if we can prove

3.72 $$\langle M^n, M^n \rangle_t \xrightarrow{P} \beta t \quad \text{for all } t \geq 0.$$

In virtue of 3.64 we can use the continuous-time version of the ergodic theorem, namely that for each bounded random variable V on $(\Omega, \mathscr{F})$ we have

$$\frac{1}{t} \int_0^t (V \circ \theta_s)\, ds \to E_\mu(V), \quad P_\mu\text{-a.s.}$$

as $t \uparrow \infty$. Since $Y_s = Y_0 \circ \theta_s$, we deduce from 3.71 that when $n \uparrow \infty$,

$$\langle M^n, M^n \rangle_t \to t E_\mu[\Gamma(g,g)(Y_0)] = t\beta \quad P_\mu\text{-a.s.},$$

hence 3.72 is met, and thus $M^n \xrightarrow{\mathscr{L}} X$. □

Secondly, we consider the discrete-time case. Although the results are essentially the same, they do present some minor differences.

We start with a discrete-time Markov chain $(\Omega, \mathscr{F}, \mathscr{F}_n, \theta_n, Y_n, P_x)$ taking values in a measurable space $(E, \mathscr{E})$. We denote by Q its transition measure. We assume that:

3.73 *Hypothesis*: There is an invariant *probability* measure μ, and the invariant σ-field under the semi-group $(\theta_n)_{n \in \mathbb{N}}$ is P_μ-trivial. □

3.74 **Theorem.** *Assume 3.73, and let f be a function of the form $f = Qg - g$, where g is bounded and measurable. Then the processes*

3.75 $$X_t^n = \frac{1}{\sqrt{n}} \sum_{1 \leq p \leq [nt]} f(Y_p),$$

under P_μ, converge in law to $\sqrt{\beta} W$, where W is a standard Wiener process, and

3.76 $$\beta = \int \mu(dx)[g^2(x) - (Qg(x))^2] = -\int \mu(dx) f(x)[f(x) + 2g(x)]$$

Proof. We will apply Theorem 3.33 to the array defined by

$$U_k^n = \frac{1}{\sqrt{n}}[Qg(Y_{k-1}) - g(Y_k)], \quad \sigma_t^n = [nt]$$

$$M_t^n = \sum_{1 \le k \le \sigma_t^n} U_k^n = \frac{1}{\sqrt{n}}\left[-g(Y_{[nt]}) + g(Y_0) + \sum_{0 \le p \le [nt]-1} (Qg - g)(Y_p)\right].$$

If $K = \|g\|$ we have $|U_k^n| \le 2K/\sqrt{n}$ and 3.31 is obviously satisfied; we have $E_\mu(U_k^n|\mathscr{F}_{k-1}) = 0$ by the Markov property. We also have

$$\sum_{1 \le k \le \sigma_t^n} (U_k^n)^2 = \frac{1}{n} \sum_{0 \le k \le [nt]-1} [Qg(Y_0) - g(Y_1)]^2 \circ \theta_k$$

which, by the ergodic theorem, converges P_μ-a. s. to

$$tE_\mu([Qg(Y_0) - g(Y_1)]^2) = tE_\mu(g(Y_1)^2 + (Qg(Y_0))^2 - 2g(Y_1)Qg(Y_0)) = t\beta.$$

Hence $[\gamma_6'\text{-}\mathbb{R}_+]$ holds with $C_t = t\beta$, and we deduce that $M^n \xrightarrow{\mathscr{L}} \sqrt{\beta}W$.

Finally, since $f = Qg - g$, we have $|M_t^n - X_t^n| \le 2K/\sqrt{n} \to 0$, hence we deduce the result. □

3.77 **Remarks.** 1) If we strengthen Hypothesis 3.73 into the following: the Markov chain is Harris-positive recurrent, then we have $X^n \xrightarrow{\mathscr{L}} \sqrt{\beta}W$ under any of the measures P_x.

2) One has similar results with functionals of the form

$$X_t^n = \frac{1}{\sqrt{n}} \sum_{1 \le p \le [nt]} f(Y_p, Y_{p+1}, \ldots, Y_{p+q})$$

(they actually reduce to the above, upon considering the E^{q+1}-valued Markov chain $\tilde{Y}_p = (Y_p, Y_{p+1}, \ldots, Y_{p+q})$).

3) Of course, in practical situations the function f is given, and we first have to solve the Poisson equation $Qg - g = f$: there might be an unbounded solution g, or an a. s. solution (see [208]); f itself may be unbounded. In all those cases, the previous theorem remains valid: see [168] or [67]. □

§ 3g. Limit Theorems for Stationary Processes

Central limit theorems for stationary processes are also quite old (see again [12] for example), under various mixing conditions. As we shall see more precisely below, such results contain the theorems about Markov processes given in § 3f as particular cases.

1. Continuous time. According to the established rules for continuous-time stationary processes, we start with a probability space $(\Omega, \mathscr{F}, P)$ endowed with a group $(\theta_t)_{t \in \mathbb{R}}$ of measurable, measure-preserving, shifts (also called a *flow*): every θ_t is a measurable mapping: $(\Omega, \mathscr{F}) \to (\Omega, \mathscr{F})$, and $P \circ \theta_t^{-1} = P$ and $\theta_{t+s} = \theta_t \circ \theta_s$ for all s, $t \in \mathbb{R}$. The next assumption is valid for this subsection, but will be removed in § 5e below.

3.78 *Hypothesis.* The invariant σ-field $\mathscr{I}$ (i.e. the σ-field of all $A \in \mathscr{F}$ such that $\theta_t^{-1}(A) = A$ for all $t \in \mathbb{R}$) is P-trivial. In other words: the flow $(\theta_t)_{t \in \mathbb{R}}$ is *ergodic*. □

Our basic stationary process is a real-valued process $Y = (Y_t)_{t \in \mathbb{R}}$ indexed by $\mathbb{R}$, such that $Y_t \circ \theta_s = Y_{t+s}$ for all s, $t \in \mathbb{R}$, and that $(\omega, t) \rightsquigarrow Y_t(\omega)$ is $\mathscr{F} \otimes \mathscr{R}$-measurable. We set $\mathscr{F}_t = \sigma(Y_s \colon s \in \mathbb{R}, s \le t)$, so that $\mathscr{F}_t = \theta_t^{-1}(\mathscr{F}_0)$.

3.79 **Theorem.** *Assume 3.78, and let $p \in [2, \infty]$ and $q \in [1, 2]$ be conjugate exponents* (i.e., $1/p + 1/q = 1$). *Assume also that*

3.80 $$\begin{cases} \|Y_0\|_p < \infty \\ \displaystyle\int_0^\infty \|E(Y_t | \mathscr{F}_0)\|_q \, dt < \infty \end{cases}$$

Then the processes

3.81 $$X_t^n = \frac{1}{\sqrt{n}} \int_0^{nt} Y_s \, ds, \quad t \ge 0$$

converge finite-dimensionally along $\mathbb{R}_+$ in law to $\sqrt{c}\,W$, where W is a standard Wiener process, and

3.82 $$c = 2 \int_0^\infty E(Y_0 Y_t) \, dt.$$

If $p = 2$, we even have $X^n \xrightarrow{\mathscr{L}} \sqrt{c}\,W$.

Before going to the proof, some comments are in order. Firstly, the measurability of $(\omega, t) \rightsquigarrow Y_t(\omega)$ and the fact that $\|Y_t\|_p = \|Y_0\|_p < \infty$ imply that X_t^n is well-defined and $\mathscr{F}_t$-measurable, and $X_t^n = X_{nt}/\sqrt{n}$, where

3.83 $$X_t = \int_0^t Y_s \, ds.$$

Secondly, the same properties imply that there exists an $\mathscr{F}_0 \otimes \mathscr{R}$-measurable version of $(\omega, t) \rightsquigarrow E(Y_t | \mathscr{F}_0)(\omega)$, and thus $t \rightsquigarrow \|E(Y_t | \mathscr{F}_0)\|_q$ is Borel and the second condition in 3.80 makes sense.

Thirdly, we have

$$|E(Y_0 Y_t)| \le E[|Y_0| |E(Y_t | \mathscr{F}_0)|] \le \|Y_0\|_p \|E(Y_t | \mathscr{F}_0)\|_q$$

and so 3.80 implies that the formula 3.82 defines a real number c; that $c \geq 0$ will be seen in the proof.

Finally, we will see below (Lemma 3.84) that $t \rightsquigarrow \|E(Y_t|\mathscr{F}_0)\|_q$ is non-increasing, so 3.80 yields $\lim_{t\uparrow\infty} \|E(Y_t|\mathscr{F}_0)\|_q = 0$, and a fortiori $\lim_{t\uparrow\infty} E(|E(Y_t|\mathscr{F}_0)|) = 0$. Thus $E(Y_t) = E[E(Y_t|\mathscr{F}_0)] \to 0$ as $t \uparrow \infty$, and since $E(Y_t) = E(Y_0)$ for all t we deduce that $E(Y_t) = 0$. In other words, 3.80 implies that the processes Y and X^n are centered, which is in accordance with the fact that $X^n \xrightarrow{\mathscr{L}(\mathbb{R})} \sqrt{c}\,W$.

3.84 **Lemma.** a) *For all $s, t, h \in \mathbb{R}$ and all integrable random variables ξ, we have $E(\xi \circ \theta_{t+h}|\mathscr{F}_{s+h}) = E(\xi \circ \theta_t|\mathscr{F}_s) \circ \theta_h$ P-a. s.*

b) *If $\xi \in L^q(\Omega, \mathscr{F}, P)$, $t \rightsquigarrow \|E(\xi \circ \theta_t|\mathscr{F}_0)\|_q$ is non-increasing.*

Proof. a) By definition of $\mathscr{F}_t$, we have $A \in \mathscr{F}_s$ if and only if $\theta_h^{-1}(A) \in \mathscr{F}_{s+h}$. For such A we have (by $P = P \circ \theta_h^{-1}$):

$$E(\xi \circ \theta_{t+h} 1_A \circ \theta_h) = E(\xi \circ \theta_t 1_A) = E[E(\xi \circ \theta_t|\mathscr{F}_s) 1_A] = E[E(\xi \circ \theta_t|\mathscr{F}_s) \circ \theta_h 1_A \circ \theta_h)$$

and since $E(\xi \circ \theta_t|\mathscr{F}_s) \circ \theta_h$ is $\mathscr{F}_{s+h}$-measurable, the claim follows.

b) If $s < t$, then $E(\xi \circ \theta_t|\mathscr{F}_s) = E[E(\xi \circ \theta_s|\mathscr{F}_0)|\mathscr{F}_{s-t}) \circ \theta_{t-s}$, hence

$$\|E(\xi \circ \theta_t|\mathscr{F}_0)\|_q = \|E[E(\xi \circ \theta_s|\mathscr{F}_0)|\mathscr{F}_{s-t}]\|_q \leq \|E(\xi \circ \theta_s|\mathscr{F}_0)\|_q. \qquad \square$$

This lemma and the comments following 3.83 show that $(\omega, s) \rightsquigarrow E(Y_s|\mathscr{F}_t)(\omega)$ is $\mathscr{F}_t \otimes \mathscr{R}$-measurable, and $s \rightsquigarrow E(|E(Y_s|\mathscr{F}_t)|)$ is Lebesgue integrable. Hence we define an integrable $\mathscr{F}_t$-measurable random variable by

3.85 $$M_t = \int_0^\infty [E(Y_s|\mathscr{F}_t) - E(Y_s|\mathscr{F}_0)]\, ds.$$

3.86 **Lemma.** *We have $E(M_1^2) = c$* (as defined by 3.82).

Proof. Let $\beta_u = E(Y_u|\mathscr{F}_1) - E(Y_u|\mathscr{F}_0)$ and $N_{s,t} = \int_s^t \beta_u\, du$ for $0 \leq s \leq t \leq \infty$, so that $M_1 = N_{0,\infty}$. β_v is $\mathscr{F}_1$-measurable and $E(\beta_v|\mathscr{F}_0) = 0$, hence

$$\begin{aligned} E(\beta_u \beta_v) &= E[(Y_u - E(Y_u|\mathscr{F}_0))\beta_v] = E(Y_u \beta_v) \\ &= E[Y_u(E(Y_v|\mathscr{F}_1) - E(Y_v|\mathscr{F}_0))] \\ &= E[(Y_u - Y_{u+1})E(Y_v|\mathscr{F}_1)] + E[Y_{u+1}E(Y_v|\mathscr{F}_1)] - E[Y_u E(Y_v|\mathscr{F}_0)]. \end{aligned}$$

Apply 3.84 and $P = P \circ \theta_1^{-1}$ to the last term above:

$$E(\beta_u \beta_v) = E[(Y_u - Y_{u+1})E(Y_v|\mathscr{F}_1)] + E[Y_{u+1}E(Y_v - Y_{v+1}|\mathscr{F}_1)].$$

Thus if $0 \leq s \leq t < \infty$,

$$E(N_{s,t}^2) = \int_s^t du \int_s^t dv\, E(\beta_u \beta_v)$$

$$= \int_s^{s+1} du \int_s^t dv\, E[Y_u E(Y_v|\mathscr{F}_1)] - \int_t^{t+1} du \int_s^t dv\, E[Y_u E(Y_v|\mathscr{F}_1)]$$

$$+ \int_s^t du \int_s^{s+1} dv\, E[Y_v E(Y_{u+1}|\mathscr{F}_1)]$$

$$- \int_s^t du \int_t^{t+1} dv\, E[Y_v E(Y_{u+1}|\mathscr{F}_1)]$$

3.87
$$= \int_s^{s+1} du \int_s^t dv\, E[Y_u E(Y_v + Y_{v+1}|\mathscr{F}_1)]$$

$$- \int_t^{t+1} du \int_s^t dv\, E[Y_u E(Y_v + Y_{v+1}|\mathscr{F}_1)],$$

and we denote by $\gamma_{s,t}$ and $\delta_{s,t}$ the two terms in 3.87.

Firstly, we evaluate $\delta_{s,t}$. Since $\|Y_v\|_p = \|Y_0\|_p$ and $\|E(Y_u|\mathscr{F}_1)\|_q = \|E(Y_{u-1}|\mathscr{F}_0)\|_q$ by 3.84a, we obtain

$$|\delta_{s,t}| \leq \int_t^{t+1} du \int_s^t dv\, |E[E(Y_u|\mathscr{F}_1)(Y_v + Y_{v+1})]|$$

$$\leq \int_t^{t+1} du \int_s^t dv\, \|Y_v + Y_{v+1}\|_p \|E(Y_u|\mathscr{F}_1)\|_q$$

$$\leq 2\|Y_0\|_p (t-s) \int_{t-1}^t \|E(Y_u|\mathscr{F}_0)\|_q\, du.$$

Then 3.84b and 3.80 clearly yield

3.88
$$\lim_{t\uparrow\infty} \delta_{0,t} = 0, \qquad \lim_{s\uparrow\infty} \sup_{t:t\geq s} |\delta_{s,t}| = 0.$$

A similar computation gives:

$$|\gamma_{s,t}| \leq \int_s^{s+1} du \int_s^t dv\, \|Y_u\|_p (\|E(Y_v|\mathscr{F}_1)\|_q + \|E(Y_{v+1}|\mathscr{F}_1)\|_q)$$

$$\leq 2\|Y_0\|_p \int_{s-1}^t \|E(Y_v|\mathscr{F}_0)\|_q\, dv \leq 2\|Y_0\|_p \int_{s-1}^\infty \|E(Y_v|\mathscr{F}_0)\|_q\, dv.$$

So $\gamma_{s,t}$ also satisfies 3.88, and

3.89
$$\lim_{s\uparrow\infty} \sup_{t:t\geq s} E(N_{s,t}^2) = 0.$$

Next, we set $\alpha_u = E(Y_0 Y_u) = E(Y_t Y_{t+u})$. In the definition of $\gamma_{s,t}$ (first term of 3.87) we have $u \leq 1$, so the integrand is $E[Y_u(Y_v + Y_{v+1})] = \alpha_{v-u} + \alpha_{v+1-u}$.

Thus

$$\gamma_{0,t} = \int_0^1 du \int_0^t dv(\alpha_{v-u} + \alpha_{v+1-u}) = 2\int_0^{t-1} \alpha_u\, du + \int_{t-1}^{t+1} (t+1-u)\alpha_u\, du$$

for $t \geq 1$ (because $\alpha_u = \alpha_{-u}$). The last term above is majorized by $2\int_{t-1}^{t+1} |\alpha_u|\, du$, while $u \rightsquigarrow \alpha_u$ is Lebesgue-integrable by 3.80 (because $|\alpha_u| \leq \|Y_0\|_p \|E(Y_u|\mathscr{F}_0)\|_q$), so this last term goes to 0 as $t \uparrow \infty$. We also have $c = 2\int_0^\infty \alpha_u\, du$ by 3.82, so that $\gamma_{0,t} \to c$ as $t \uparrow \infty$. In view of 3.87 and 3.88, we then deduce

3.90 $$E(N_{0,t}^2) \to c \quad \text{as } t \uparrow \infty.$$

If we recall that $M_1 = \lim_{t\uparrow\infty} N_{0,t}$, Fatou's Lemma yields

$$E[(M_1 - N_{0,t})^2] = E\left(\lim_{s\uparrow\infty} N_{t,s}^2\right) \leq \sup_{\{s:s\geq t\}} E(N_{t,s}^2),$$

which goes to 0 as $t \uparrow \infty$ by 3.89. Thus $N_{0,t} \to M_1$ in L^2, and then $E(M_1^2) = c$ follows from 3.90. □

We shall say that a process $V = (V_t)_{t\geq 0}$ is *additive* if for all $s, t \geq 0$ we have $V_{t+s} = V_t + V_s \circ \theta_t$ a.s. This is similar to the notion of "helices", as defined in [37] or [201], except that V is indexed by $\mathbb{R}_+$ only and meets $V_0 = 0$ a.s. Here is an ergodic theorem adapted to our purposes:

3.91 **Lemma.** *Let V be an additive process such that $\sup_{s\leq 1} E(|V_s|) < \infty$. Then under 3.78 we have for all $t \in \mathbb{R}_+$:*

$$\frac{1}{n} V_{nt} \xrightarrow{L^1} tE(V_1) \quad \text{as } n \uparrow \infty.$$

Proof. Let $s \in (0,1]$ and denote by $\mathscr{I}_s$ the θ_s-invariant σ-field: $\mathscr{I}_s = \{A \in \mathscr{F}: \theta_s^{-1}(A) = A\}$. By the additivity of V, and if $p = [nt/s]$,

$$\frac{1}{n} V_{nt} = \frac{p}{n}\frac{1}{p} \sum_{1\leq k\leq p} V_s \circ \theta_{ks} + \frac{1}{n}(V_{nt-ps} \circ \theta_{ps}) \text{ a.s.,}$$

and consider the right-hand side: due to the (discrete-time) ergodic theorem, the first term goes to $\frac{t}{s}E(V_s|\mathscr{I}_s)$ as $n \uparrow \infty$, a.s. and in L^1. The absolute moment of the second term is

$$\frac{1}{n}E(|V_{nt-ps}|) \leq \frac{1}{n} \sup_{0\leq u\leq 1} E(|V_u|),$$

and thus

$$\frac{1}{n} V_{nt} \xrightarrow{L^1} \frac{t}{s}E(V_s|\mathscr{I}_s).$$

Therefore if $W_s = \frac{1}{s} E(V_s|\mathscr{I}_s)$, it follows that $W_s = W_1$ a.s. Hence W_1 is measurable with respect to the completion of the σ-field $\bigcap_{s\in(0,1]}\mathscr{I}_s$, the latter being equal to $\mathscr{I}$. Then 3.78 yields $W_1 = E(W_1) = E(V_1)$. $\square$

Proof of Theorem 3.79. a) If we consider 3.85, it is obvious that $M_s = E(M_t|\mathscr{F}_s)$ for $0 \le s \le t$, so that, except for the right-continuity, M is a martingale. Indeed, $(\mathscr{F}_t)_{t\ge 0}$ is not necessarily càd, and we first have to consider the (right-continuous) filtration $(\mathscr{F}_{t+})_{t\ge 0}$. Then we need to show that in 3.85 we can replace $\mathscr{F}_t$ by $\mathscr{F}_{t+}$. To this end, there are two methods:
1) We can resort to the (rather difficult) paper [37], where it is shown that $\mathscr{F}_t$ and $\mathscr{F}_{t+}$ only differ by P-null sets of the P-completion of $\mathscr{F}$.
2) Or we can replace $\mathscr{F}_t$ by $\mathscr{F}_{t+}$ in the definition 3.85 itself. The proof of 3.86 remains valid (with $\beta_u = E(Y_u|\mathscr{F}_{1+}) - E(Y_u|\mathscr{F}_{0+})$), provided 3.80 is also true with $\mathscr{F}_{0+}$. But by 3.84. $E(Y_t|\mathscr{F}_{0+})$ is the L^p-limit of $E(Y_t|\mathscr{F}_s) = E(Y_{t-s}|\mathscr{F}_0)\circ\theta_s$ as $s\downarrow 0$; thus $\|E(Y_{t-s}|\mathscr{F}_0)\|_q$ decreases (in virtue of 3.84b) to $\|E(Y_t|\mathscr{F}_{0+})\|_q$, and 3.80 obviously yields

$$\int_0^\infty \|E(Y_t|\mathscr{F}_{0+})\|_q\,dt < \infty.$$

In other words, using either one or the other method, we have shown that in 3.85 one can replace $\mathscr{F}_t$ by $\mathscr{F}_{t+}$. Then using the usual regularization procedure (see e.g. I.1.42), we can consider a version of M which is a.s. càdlàg (and so is a martingale in the usual sense).

b) Since 3.84a remains obviously true with $\mathscr{F}_{t+}$ instead of $\mathscr{F}_t$, it is clear from 3.85 (with $\mathscr{F}_{t+}$) that M is an additive process. By construction the process X defined by 3.83 is additive and continuous. Then the formula

3.92 $$Z_t = \int_0^\infty E(Y_s|\mathscr{F}_{0+})\,ds + M_t - X_t = \int_t^\infty E(Y_s|\mathscr{F}_{t+})\,ds$$

defines a càdlàg process Z and $(Z_t - Z_0)_{t\ge 0}$ is additive.

Consider now the quadratic variation $[M,M]$. This is also additive: for this, we can either refer to [201], or use the approximation theorem I.4.47 with subdivisions of the form $\tau_n = \left\{\frac{sm}{n}\right\}_{m\in\mathbb{N}}$, once noticed that because the paths of M are càd, for all $s \ge 0$ we have $M_{s+t} = M_s + M_t\circ\theta_s$ P-a.s. for all $t \ge 0$.

c) We have $X^n = M^n - Z^n + Z_0/\sqrt{n}$, with $M^n_t = M_{nt}/\sqrt{n}$, $Z^n_t = Z_{nt}/\sqrt{n}$. M^n is a martingale (for $(\mathscr{F}_{nt+})_{t\ge 0}$) with quadratic variation $[M^n,M^n]_t = [M,M]_{nt}/n$, and $\sup_{0\le s\le 1} E([M,M]_s) = E(M_1^2) = c$ by 3.86. Hence 3.91 yields

3.93 $$[M^n,M^n]_t \xrightarrow{P} tc \quad \text{for all } t \ge 0.$$

The processes $V(a)_t = \sum_{0\le s\le t}(\Delta M_s)^2 1_{\{|\Delta M_s|>a\}}$ are also obviously additive, so $E(V(a)_N) = NE(V(a)_1)$ for $N\in\mathbb{N}^*$, and

$$E\left(\sum_{s\leq N}(\Delta M_s^n)^2 1_{\{|\Delta M_s^n|>\varepsilon\}}\right)=\frac{1}{n}E(V(\varepsilon\sqrt{n})_{Nn})=NE(V(\varepsilon\sqrt{n})_1).$$

Now, $V(\varepsilon\sqrt{n})_1\leq V(0)_1\in L^1$, and $V(\varepsilon\sqrt{n})_1\downarrow 0$ as $n\uparrow\infty$, hence

3.94 $$E\left(\sum_{s\leq N}(\Delta M_s^n)^2 1_{\{|\Delta M_s^n|>\varepsilon\}}\right)\to 0\quad\text{as } n\uparrow\infty,\text{ for all }\varepsilon>0.$$

Thus (M^n) meets 3.23 (the expected values of the left-hand sides of 3.94 and 3.23 are equal), and in virtue of 3.93, Theorem 3.22 yields

3.95 $$M^n\xrightarrow{\mathscr{L}}\sqrt{c}\,W,$$

where W is a standard Wiener process.

d) It follows easily from 3.84a (for $\mathscr{F}_{t+}$ instead of $\mathscr{F}_t$) and 3.92 that $Z_t=Z_0\circ\theta_t$ a.s., while $Z_0\in L^1$. Hence $E(|Z_t^n|)=E(|Z_0|)/\sqrt{n}\to 0$ as $n\uparrow\infty$. Since $X^n=M^n-Z^n+Z_0/\sqrt{n}$, we deduce $X^n\xrightarrow{\mathscr{L}(\mathbb{R}^+)}\sqrt{c}\,W$ from 3.94.

Assume now that $p=2$. In view of 3.95, in order to obtain $X^n\xrightarrow{\mathscr{L}}\sqrt{c}\,W$ it suffices to prove that $\sup_{s\leq N}|Z_s^n|\xrightarrow{P}0$ for all $N\in\mathbb{N}^*$. Let $\xi=\sup_{0\leq s\leq 1}|Z_s|$, so that $\sup_{s\leq N}|Z_s^n|=\sup_{1\leq k\leq Nn}\xi\circ\theta_k/\sqrt{n}$. Thus it is enough to prove that $\xi\circ\theta_n/\sqrt{n}\to 0$. a.s. or, by Borel-Cantelli Lemma, that $\sum_{n\geq 1}P(\xi\circ\theta_n/\sqrt{n}\geq\varepsilon)<\infty$ for all $\varepsilon>0$. But

$$\sum_{n\geq 1}P\left(\frac{1}{\sqrt{n}}\xi\circ\theta_n\geq\varepsilon\right)=\sum_{n\geq 1}P(\varepsilon^{-2}\xi^2\geq n)\leq\varepsilon^{-2}E(\xi^2)$$

and it remains to prove that $E(\xi^2)<\infty$.

To this effect, we observe that for $0\leq s\leq 1$, $Z_s=E(Z_1+\int_s^1 Y_u\,du|\mathscr{F}_{s+})$. Thus $|Z_s|\leq E(U|\mathscr{F}_{s+})$, where $U=|Z_1|+\int_0^1|Y_u|\,du$, and in virtue of Doob's inequality I.1.43 we have $E(\xi^2)\leq 4E(U^2)$.

On the one hand, $\|Y_u\|_2=\|Y_0\|_2<\infty$, so $\int_0^1 Y_u du\in L^2$. On the other hand, 3.92 yields

$$\|Z_1\|_2\leq\int_1^\infty\|E(Y_s|\mathscr{F}_{1+})\|_2\,ds=\int_0^\infty\|E(Y_s|\mathscr{F}_{0+})\|_2\,ds<\infty$$

by 3.80 (recall that $p=q=2$). Hence $U\in L^2$, and we are finished. □

3.96 **Remark.** This theorem contain 3.65 as a particular case, at least when the state space E of the Markov process is nice enough (a Polish space, for example). For in this case, we can assume that the Markov process is defined on the whole line $\mathbb{R}$, and we have to replace Y in 3.81 by $f(Y)$ in 3.65 (the first assumption in 3.80 is obvious, and the second one comes from the fact that $f=Ag$ with g bounded).

In addition, if we compare the proofs, we observe that they are essentially the same (indeed, the basic martingale M is the same in 3.65 and in 3.85). □

2. *Discrete time.* Now, the space $(\Omega, \mathscr{F}, P)$ is endowed with a bimeasurable bijective mapping θ, with $P \circ \theta^{-1} = P$. Let ξ be a random variable, and $\mathscr{F}_n = \sigma(\xi \circ \theta^k \colon k \in \mathbb{Z}, k \leq n)$.

3.97 **Theorem.** *Assume that θ is ergodic and let $p \in [2, \infty]$ and $q \in [1, 2]$ be conjugate exponents. Assume also that*

3.98
$$\begin{cases} \|\xi\|_p < \infty \\ \sum_{n \geq 1} \|E(\xi \circ \theta^n | \mathscr{F}_0)\|_q < \infty. \end{cases}$$

Then the processes

3.99
$$X_t^n = \frac{1}{\sqrt{n}} \sum_{1 \leq k \leq [nt]} \xi \circ \theta^k, \quad t \geq 0,$$

converge finite-dimensionally along $\mathbb{R}_+$ in law to $\sqrt{c}W$, where W is a standard Wiener process, and

3.100
$$c = E(\xi^2) + 2 \sum_{n \geq 1} E(\xi \xi \circ \theta^n).$$

If $p = 2$, we even have $X^n \xrightarrow{\mathscr{L}} \sqrt{c}W$.

Exactly as in 3.79, the hypothesis 3.98 implies that $E(\xi) = 0$.

Proof. There are two different proofs. One consists in repeating almost word for word the argument of the proof of Theorem 3.79: it is in fact slightly easier, for example Lemma 3.91 or part (a) of the proof of 3.79 are useless in the discrete-time case.

Another proof consists in applying directly 3.79 to the following situation: let λ be the Lebesgue measure on $[0, 1)$, and

$$\tilde{\Omega} = \Omega \times [0, 1), \quad \tilde{\mathscr{F}} = \mathscr{F} \otimes \mathscr{B}([0, 1)), \quad \tilde{P} = P \otimes \lambda$$
$$\left.\begin{array}{l} \tilde{\theta}_t(\omega, x) = (\theta^n(\omega), x - t + n) \\ \tilde{Y}_t(\omega, x) = \xi \circ \theta^n(\omega) \end{array}\right\} \text{if } x + n - 1 \leq t < x + n, \quad n \in \mathbb{Z}.$$

It is straightforward to check that $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}; (\tilde{\theta}_t), (\tilde{Y}_t))$ satisfies the conditions preceding 3.79, including 3.78 if θ is ergodic. It is also clear that $\|\tilde{Y}_t\|_p = \|\xi\|_p$ and that $\tilde{E}(\tilde{Y}_t | \tilde{\mathscr{F}}_0)(\omega, x) = E(\xi \circ \theta^n | \mathscr{F}_0)(\omega)$ on the set $\{x \colon x + n - 1 \leq t < x + n\}$. Hence 3.98 obviously implies 3.80. Using the same sort of argument, one readily obtains that 3.100 and 3.82 (with $\tilde{Y}$) define the same number c.

Finally, define $\tilde{X}^n$ by 3.81, starting with $\tilde{Y}$. Then

$$|\tilde{X}_t^n - X_t^n| \leq \frac{1}{\sqrt{n}} (|\xi| + |\xi| \circ \theta^{[nt]} + |\xi| \circ \theta^{[nt]+1})$$

and the same argument than in part (d) of the proof of 3.79 shows that $\sup_{s\le N}|\tilde{X}^n_s - X^n_s| \xrightarrow{P} 0$ as $n\uparrow\infty$ (because $\xi\in L^2$). Hence, the result follows from 3.79. □

3. *Comparison with the usual mixing conditions.* There exist several coefficients "measuring" the dependence between two σ-fields $\mathscr{A}$ and $\mathscr{B}$ on a probability space $(\Omega,\mathscr{F},P)$, the most usual ones being:

3.101
$$\begin{cases} \varphi(\mathscr{A},\mathscr{B}) = \sup\{|P(B/A) - P(B)|\colon A\in\mathscr{A}, P(A)>0, B\in\mathscr{B}\} \\ \alpha(\mathscr{A},\mathscr{B}) = \sup\{|P(A\cap B) - P(A)P(B)|\colon A\in\mathscr{A}, B\in\mathscr{B}\} \\ \rho(\mathscr{A},\mathscr{B}) = \sup\{|E(XY)|\colon \|X\|_2\le 1, \|Y\|_2\le 1, E(X)=E(Y)=0, \\ \qquad\qquad X \text{ (resp. } Y) \text{ is } \mathscr{A}\text{- (resp. } \mathscr{B}\text{-) measurable}\}. \end{cases}$$

3.102 **Lemma.** *Let X be an integrable, $\mathscr{B}$-measurable random variable. Then*

a) $\|E(X|\mathscr{A}) - E(X)\|_q \le 2^{2-2/p}\varphi(\mathscr{A},\mathscr{B})^{1/p}\|X\|_p$ *if* $\frac{1}{p}+\frac{1}{q}=1$, $p\in[2,\infty]$;

b) $\|E(X|\mathscr{A}) - E(X)\|_q \le 2(2^{1/q}+1)\,\alpha(\mathscr{A},\mathscr{B})^{1/q-1/r}\|X\|_r$ *if* $1\le q\le r\le\infty$;

c) $\|E(X|\mathscr{A}) - E(X)\|_2 \le \rho(\mathscr{A},\mathscr{B})\,\|X - E(X)\|_2$.

Proof. a) Set $\varphi = \varphi(\mathscr{A},\mathscr{B})$. By a standard argument on conditional expectations, one easily deduces from 3.101 that

3.103 $$\text{For all } B\in\mathscr{B},\quad |E(1_B|\mathscr{A}) - P(B)| \le \varphi \quad \text{a.s.}$$

In order to prove the claim, we can assume that $\mathscr{B} = \sigma(X)$, so that there exists a regular version $\eta(\omega,d\omega')$ of the conditional probability $P(\cdot|\mathscr{A})$ on $\mathscr{B}$ (see II.1.2); furthermore, if $\mu(\omega,d\omega') = \eta(\omega,d\omega') - P(d\omega')_{|\mathscr{B}}$, the separability of $\mathscr{B}$ and 3.103 imply that $|\mu|(\omega,\Omega)\le\varphi$ for P-almost all ω (as usual, one considers the Jordan-Hahn decomposition $\mu = \mu_+ - \mu_-$ of the signed measure μ, and $|\mu| = \mu_+ + \mu_-$).

Set $Y = E(X|\mathscr{A}) - E(X) = \int X\,d\mu_+ - \int X\,d\mu_-$, so that

$$|Y|^q \le 2^{q-1}\left(\left|\int X\,d\mu_+\right|^q + \left|\int X\,d\mu_-\right|^q\right).$$

Now, since $\mu_+(\Omega)\le\varphi$ a.s., we have

$$\left|\int X\,d\mu_+\right| \le \mu_+(\Omega)^{1/q}\left(\int |X|^p\,d\mu_+\right)^{1/p} \le \varphi^{1/q}\left(\int |X|^p\,d\mu_+\right)^{1/p} \quad \text{a.s.,}$$

and a similar inequality for $|\int X\,d\mu_-|$. Moreover, since $q\le p$, we have $x^{q/p} + y^{q/p} \le 2^{1-q/p}(x+y)^{q/p}$ and $(x+y)^{q/p}\le x^{q/p}+y^{q/p}$ for all $x, y\ge 0$. Since $|\mu|\le \eta + P_{|\mathscr{B}}$, we obtain:

$$|Y|^q \le \varphi 2^{q-1} 2^{1-q/p} \left(\int |X|^p \, d|\mu| \right)^{q/p},$$

$$\begin{aligned} E(|Y|^q) &\le 2^{q-q/p} \varphi \left\{ E\left[\left(\int |X|^p \, d\eta \right)^{q/p} \right] + E(|X|^p)^{q/p} \right\} \\ &= 2^{q-q/p} \varphi \{ E[E(|X|^p|\mathscr{A})^{q/p}] + E(|X|^p)^{q/p} \} \\ &\le 2^{q-q/p} \varphi \{ E[E(|X|^p|\mathscr{A})]^{q/p} + E(|X|^p)^{q/p} \} \\ &\le 2^{q-q/p} \varphi 2 E(|X|^p)^{q/p} \end{aligned}$$

and the claim follows.

b) Set $\alpha = \alpha(\mathscr{A}, \mathscr{B})$. Here again, a standard argument allows to deduce from 3.101 that for all variables U, V, respectively $\mathscr{A}$- and $\mathscr{B}$-measurable, with $|U| \le 1$, $|V| \le 1$, we have $|E(UV) - E(U)E(V)| \le 4\alpha$. Taking $U = 1$ on the set $\{E(V|\mathscr{A}) > E(V)\}$ and $U = -1$ on the complement, we get

3.104
$$\begin{aligned} E(|E(V|\mathscr{A}) - E(V)|) &= E[U(E(V|\mathscr{A}) - E(V))] \\ &= E(UV) - E(U)E(V) \le 4\alpha. \end{aligned}$$

Let $c > 0$, and $X_1 = X 1_{\{|X| \le c\}}$, $X_2 = X 1_{\{|X| > c\}}$, $Y = E(X|\mathscr{A}) - E(X)$. Then

$$\|Y\|_q \le \|E(X_1|\mathscr{A}) - E(X_1)\|_q + \|E(X_2|\mathscr{A}) - E(X_2)\|_q.$$

3.104 applied to $V = X_1/c$ yields

$$\begin{aligned} \|E(X_1|\mathscr{A}) - E(X_1)\|_q &\le (2c)^{(q-1)/q} E(|E(X_1|\mathscr{A}) - E(X_1)|)^{1/q} \\ &\le (2c)^{(q-1)/q} (4\alpha c)^{1/q}. \end{aligned}$$

Since $|X_2|^r \ge c^{r-q}|X_2|^q$ and $|X_2| \le |X|$ for $r \ge q$, we obtain

$$\|E(X_2|\mathscr{A}) - E(X_2)\|_q \le 2\|X_2\|_q \le 2c^{(q-r)/q} \|X_2\|_r^{r/q} \le 2c^{(q-r)/q} \|X\|_r^{r/q}.$$

Hence

$$\|Y\|_q \le (2c)^{(q-1)/q} (4\alpha c)^{1/q} + 2c^{(q-r)/q} \|X\|_r^{r/q}.$$

Then taking $c = \|X\|_r \alpha^{-1/r}$ gives the claimed inequality.

c) Since $Y = E(X|\mathscr{A}) - E(X)$ is $\mathscr{A}$-measurable and $E(Y) = 0$, we have $\|Y\|_2 = \sup\{|E(YZ)|: \|Z\|_2 \le 1, E(Z) = 0, Z \text{ is } \mathscr{A}\text{-measurable}\}$, and the result is obvious. □

Now we turn back to the discrete-time setting of Theorem 3.97. We set $\mathscr{F}^n = \sigma(\xi \circ \theta^k : k \ge n)$, and

3.105
$$\begin{cases} \varphi_k = \varphi(\mathscr{F}_0, \mathscr{F}^k) \\ \alpha_k = \alpha(\mathscr{F}_0, \mathscr{F}^k) \\ \rho_k = \rho(\mathscr{F}_0, \mathscr{F}^k). \end{cases}$$

If $\varphi_k \to 0$ (resp. $\alpha_k \to 0$, resp. $\rho_k \to 0$), we say that the sequence $(\xi \circ \theta^k)_{k \in \mathbb{Z}}$ is *φ-mixing* (resp. *α-mixing*, resp. *ρ-mixing*).

3.106 **Corollary.** *Let $p \in [2, \infty]$ and assume that $E(\xi) = 0$, that $\|\xi\|_p < \infty$, and that one of the following conditions holds:*
(i) $\sum_{n \geq 0} (\varphi_k)^{(p-1)/p} < \infty$,
(ii) $\sum_{n \geq 0} (\alpha_k)^{(p-2)/p} < \infty$,
(iii) $p = 2$ *and* $\sum_{n \geq 0} \rho_k < \infty$.
Then the conclusions of Theorem 3.97 are valid.

Proof. In view of Lemma 3.102, the assumptions imply that 3.98 holds with $\frac{1}{q} + \frac{1}{p} = 1$ (in case of (ii), use 3.102b with $r = p$). It remains to check the ergodicity. Indeed, θ may be non-ergodic. But if $\mathscr{F}' = \sigma(\xi \circ \theta^k \colon k \in \mathbb{Z})$, then we can consider θ as a bijective bimeasurable measure-preserving transformation on $(\Omega, \mathscr{F}', P)$, and we will see that its invariant σ-field $\mathscr{I}' = \{A \in \mathscr{F}' \colon \theta^{-1}(A) = A\}$ is P-trivial, so that 3.97 applies.

This is in fact a well known property of φ- (resp. α-, resp. ρ-) mixing sequences. Indeed, let $A \in \mathscr{I}'$. Note that $A \in \mathscr{F}_0 \cap \mathscr{F}^k$ for all $k \in \mathbb{Z}$. In view of 3.101, if $P(A) > 0$ we have $|P(A/A) - P(A)| = |1 - P(A)| \leq \varphi_k$, so under (i) we obtain $P(A) = 1$. Similarly, $|P(A \cap A) - P(A)^2| = |P(A) - P(A)^2| \leq \alpha_k$, so under (ii) we have $|P(A) - P(A)^2| = 0$, and $P(A) = 0$ or 1. Finally if $X = 1_A - P(A)$ we get $|P(A) - P(A)^2| = E(X^2) \leq \rho_k$ (because X is $\mathscr{F}_0 \cap \mathscr{F}^k$-measurable), so under (iii) we obtain as before that $P(A) = 0$ or 1. Thus we have seen that in all cases, $\mathscr{I}'$ is P-trivial. □

3.107 **Remark.** According to Ibragimov and Linnik [89], in the corollary above we obtain functional convergence $X^n \xrightarrow{\mathscr{L}} \sqrt{c}\, W$ even if $p > 2$, in case of (i), or in case (ii) is replaced by $\sum_{n \geq 0} (\alpha_k)^{(p-2)/2p} < \infty$. □

Similar results hold of course in the continuous-time setting. With the notation of the beginning of the subsection, we set $\mathscr{F}^t = \sigma(Y_s \colon s \geq t)$, and $\varphi_t = \varphi(\mathscr{F}_0, \mathscr{F}^t)$, $\alpha_t = \alpha(\mathscr{F}_0, \mathscr{F}^t)$, $\rho_t = \rho(\mathscr{F}_0, \mathscr{F}^t)$, and conditions (i), (ii), (iii) of 3.106 are replaced, respectively, by

$$\int_0^\infty \varphi_t^{(p-1)/p}\, dt < \infty, \qquad \int_0^\infty \alpha_t^{(p-2)/p}\, dt < \infty, \qquad \int_0^\infty \rho_t\, dt < \infty.$$

4. Convergence to a General Process with Independent Increments

The chief purpose of this section is to extend Theorem 2.17 to the case where X is a PII with characteristics (B, C, ν), having possibly fixed times of discontinuity.

The setting and notation are as in § 1a: each X^n is a semimartingale with characteristics (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$; the truncation function h is *continuous.*

Our main result is the following, which extends 2.17 and the implication (b) ⇒ (a) of VII.3.13.

4.1 Theorem. *Assume that the following three conditions hold, where D is a dense subset of $\mathbb{R}_+$:*

[Sk-β_5] $\quad B^n \xrightarrow{P} B$ for the Skorokhod topology in $\mathbb{D}(\mathbb{R}^d)$

[γ_5-D] $\quad \tilde{C}^n_t \xrightarrow{P} \tilde{C}_t$ for all $t \in D$

[Sk-$\delta_{5,1}$] $\quad g * \nu^n \xrightarrow{P} g * \nu$ for the Skorokhod topology in $\mathbb{D}(\mathbb{R})$, for all $g \in C_1(\mathbb{R}^d)$.

Then we have $X^n \xrightarrow{\mathscr{L}} X$, and also

[Sk-$\beta\gamma\delta_5$] $\quad (B^n, \tilde{C}^n, g * \nu^n) \xrightarrow{P} (B, \tilde{C}, g * \nu)$ for the Skorokhod topology in $\mathbb{D}(\mathbb{R}^{d+d^2+m})$, for all functions $g\colon \mathbb{R}^d \to \mathbb{R}^m$ whose components belong to $C_2(\mathbb{R}^d)$.

For the meaning of: $B^n \xrightarrow{P} B$, for instance, recall 1.5: we consider B^n as a random variable taking its values in the metric space $\mathbb{D}(\mathbb{R}^d)$.

If $g \geq 0$ then $g * \nu^n$ is increasing and $\Delta(g * \nu^n)_s = \nu^n(\{s\} \times g)$. Hence in virtue of VI.2.15, and upon taking subsequences that allow to consider almost sure convergence (like in the proof of 2.4), we have the following:

4.2 [Sk-$\delta_{5,1}$] ⇔ there is a dense subset D in $\mathbb{R}_+$ such that the following two conditions hold:

$$\begin{cases} [\delta_{5,1}\text{-}D] & g * \nu^n_t \xrightarrow{P} g * \nu_t \quad \text{for all } t \in D,\ g \in C_1(\mathbb{R}^d) \\ [\tilde{\delta}_{5,1}\text{-}D] & \sum_{s \leq t} \nu^n(\{s\} \times g)^2 \xrightarrow{P} \sum_{s \leq t} \nu(\{s\} \times g)^2 \quad \text{for all } t \in D \text{ and } g \in C_1(\mathbb{R}^d). \end{cases}$$

□

In fact, in this chapter we prove Theorem 4.1 under Condition 1.17 only, the proof of the general case being given *in the next chapter* (see IX.3.35 and IX.3.37): this is because, with the tools presently at our disposition, we first need to prove that $X^n \xrightarrow{\mathscr{L}(D)} X$, which is easy enough under 1.17 (see Theorem 1.18), and the proof in this case is given below. Contrarywise, when 1.17 fails for some $t > 0$, a direct proof of $X^n \xrightarrow{\mathscr{L}(D)} X$ under the assumptions of 4.1 is quite intricate: it uses the cumbersome Section VII.4, plus some additional complications (see [105] for a proof along these lines).

Moreover, it is indeed very rarely that one encounters a PII X which does not meet 1.17.

§ 4a. Proof of Theorem 4.1 When the Characteristic Function of X_t Vanishes Almost Nowhere

We begin with two results, independent of whether 1.17 is met or not.

4.3 **Lemma.** *Let D be dense in $\mathbb{R}_+$, and assume* [Sk-β_5], [γ_5-D] *and* [Sk-$\delta_{5,1}$]. *Then* [Sk-$\beta\gamma\delta_5$] *holds.*

Proof. As usual, it suffices to show that for any infinite subsequence (n') one may find a set A with $P(A) = 1$ and a further subsequence (n'') such that for all $\omega \in A$:

$$\begin{cases}(B^{n''}(\omega), C^{n''}(\omega), g * \nu^{n''}(\omega)) \to (B, C, g * \nu) \text{ in } \mathbb{D}(\mathbb{R}^{d+d^2+m}), \text{ for every} \\ \text{function } g\colon \mathbb{R}^d \to \mathbb{R}^m \text{ whose components belong to } C_2(\mathbb{R}^d).\end{cases}$$

In virtue of Corollary VII.3.48 this will be the case if we have

$$\text{4.4} \qquad \omega \in A \Rightarrow \begin{cases} B^{n''}(\omega) \to B \quad \text{in } \mathbb{D}(\mathbb{R}^d) \\ \tilde{C}_t^{n''}(\omega) \to \tilde{C}_t \quad \text{for all } t \in D \\ g * \nu^{n''}(\omega) \to g * \nu \quad \text{in } \mathbb{D}(\mathbb{R}) \text{ for all } g \in C_1(\mathbb{R}^d),\ g \geq 0. \end{cases}$$

Let (n') be an infinite subsequence. We can assume that D and $C_1(\mathbb{R}^d)$ are countable. Then a diagonal argument yields a set A with $P(A) = 1$ and a further subsequence (n'') such that 4.4 holds, and we are finished. □

4.5 **Lemma.** *Under* [Sk-$\beta\gamma\delta_5$] *the sequence* (X^n) *is tight.*

Proof. We will apply Theorem VI.5.10. Condition (i) of VI.5.10 is trivially met because $X_0^n = 0$, while [Sk-β_5] implies Condition VI.5.10iii.

Let $g_q(x) = (q|x| - 1)^+ \wedge 1$. Then [Sk-$\beta\gamma\delta_5$] implies that for each $p > 0$, $G^{n,p} = \sum_{j \leq d} \tilde{C}^{n,jj} + g_p * \nu^n$ converges in measure to $\sum_{j \leq d} \tilde{C}^{jj} + g_p * \nu$: then condition (iv) of VI.5.10 is met with (Cl). Finally, let $\varepsilon > 0$ and $N > 0$ with $\nu(\{N\} \times \mathbb{R}^d) = 0$. There is a $p > 0$ such that $g_p * \nu_N \leq \varepsilon/2$ (because $g_p * \nu_N \downarrow 0$ as $p \downarrow 0$), while $g_p * \nu_N^n \xrightarrow{P} g_p * \nu_N$. Then

$$\lim_n P^n(g_p * \nu_N^n > \varepsilon) = 0.$$

Since $\nu^n([0, N] \times \{|x| > 1/2p\}) \leq g_p * \nu_N^n$, we also have VI.5.10ii. □

Proof of 4.1 when 1.17 is fulfilled for all $t > 0$. Due to the previous lemma, it suffices to show that $X^n \xrightarrow{\mathscr{L}(D)} X$ for $D = D(X) = \{t\colon P(\Delta X_t \neq 0) = 0\}$. Due to Theorem 1.18, it suffices to show that 1.10 holds for all $t \in D$. To see that, consider the proof of 4.3, and pick $\omega \in A$. There is a PII $Z^{n'',\omega}$ whose (deterministic) characteristics are $B^{n''}(\omega)$, $C^{n''}(\omega)$, $\nu^{n''}(\omega)$, so 4.4 and Theorem VII.3.13 yield that $Z^{n'',\omega} \xrightarrow{\mathscr{L}} X$, and in particular $G^{n''}(u)_t(\omega)$, which is the expected value of

$\exp iu \cdot Z_t^{n'',\omega}$, goes to $g(u)_t$ for all $t \in D$. We easily deduce that 1.10 holds for all $t \in D$, and the proof is finished. □

Along the same lines, we prove the following:

4.6 **Theorem.** *Assume that 1.17 holds for all $t > 0$. Let D be a dense subset of $\mathbb{R}_+$, and assume $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D] + [\text{Sk-}\delta_{5,1}]$. Then we have $X^n \xrightarrow{\mathscr{L}(D)} X$.*

Note that $[\text{Sk-}\gamma_{5,1}]$ does not imply $[\delta_{5,1}\text{-}D]$ in general, unless D is included into $\mathbb{R}_+ \backslash J(X) = \{t: P(\Delta X_t \neq 0) = 0\} = \{t: \nu(\{t\} \times \mathbb{R}^d) = 0\}$. Similarly, $X^n \xrightarrow{\mathscr{L}} X$ does not imply $X^n \xrightarrow{\mathscr{L}(D)} X$, unless $D \subset \mathbb{R}_+ \backslash J(X)$. Hence, even when $[\text{Sk-}\delta_{5,1}]$ is met this theorem complements Theorem 4.1 (of course, this distinction does not arise when X has no fixed time of discontinuity).

Proof. If (n') is a subsequence, there is a sub-subsequence (n'') and a set A with $P(A) = 1$, such that for all $\omega \in A$ we have $B_t^{n''}(\omega) \to B_t$, $\tilde{C}_t^{n''}(\omega) \to \tilde{C}_t$, $g * \nu_t^{n''}(\omega) \to g * \nu_t$ for all $t \in D$ (D may be assumed to be countable) and all $g \in C_1(\mathbb{R}^d)$, and also $g * \nu^{n''}(\omega) \to g * \nu$ in $\mathbb{D}(\mathbb{R})$ for all $g \in C_1(\mathbb{R}^d)$ (see the proof of 2.4). Then VII.4.38 implies $G^{n''}(u)_t(\omega) \to g(u)_t$ for all $u \in \mathbb{R}^d$, $t \in D$, $\omega \in A$. Therefore 1.10 holds for all $t \in D$, and Theorem 1.18 yields the result. □

4.7 **Remark.** Even when 1.17 fails, one could prove the following (see [105]) when *D is dense*: under $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D] + [\text{Sk-}\gamma_{5,1}]$, we have $X^n \xrightarrow{\mathscr{L}(D)} X$. □

4.8 **Remark.** Suppose again that 1.17 holds for all $t > 0$. Consider the next condition, which extends condition VII.4.30:

Condition: (a) For each $\varepsilon > 0$ there is a strictly increasing sequence $(t_j(\varepsilon))_{j \geq 1}$ of times, with $\lim_j \uparrow t_j(\varepsilon) = \infty$ and

$$\sup_{s \leq t,\, s \neq t_j(\varepsilon) \text{ for all } j} \nu(\{s\} \times \{|x| > \varepsilon\}) \leq \varepsilon.$$

(b) For each $\varepsilon > 0$ and each $n \geq 1$ there is a strictly increasing sequence $(T_j^n(\varepsilon))_{j \geq 1}$ of predictable times on $\mathscr{B}^n$ with $\lim_j \uparrow T_j^n(\omega) = \infty$ and

(i) if $s, t \in D \cup \{0\}$ with $s < t_j(\varepsilon) \leq t$, then

$$\begin{cases} P^n(s < T_j^n(\varepsilon) \leq t) \to 1 \\ \nu^n(\{T_j^n(\varepsilon)\} \times g) \xrightarrow{P} \nu(\{t_j(\varepsilon)\} \times g) \quad \text{for all } g \in C_1(\mathbb{R}^d) \end{cases}$$

(ii) $\limsup_n P^n(\sup_{s \leq t,\, s \neq T_j^n(\varepsilon)} \nu^n(\{s\} \times \{|x| > \varepsilon\}) > \varepsilon + \eta) = 0$ for all $\varepsilon > 0$, $\eta > 0$, $t \in D$. □

Then, using VII.4.37, and with the same proof than for 4.6, one may show that this condition, plus $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ imply $X^n \xrightarrow{\mathscr{L}(D)} X$, whether D is dense or not. □

4.9 **Remark.** Suppose that 1.17 fails. Again when D is dense, one can prove that the condition 4.8, plus $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ imply that $X^n \xrightarrow{\mathscr{L}(D)} X$ (in fact, under $[\delta_{5,1}\text{-}D]$ and the denseness of D, Condition 4.8 is equivalent to $[\text{Sk-}\gamma_{5,1}]$). However, *when D is not dense*, this result does not seem to be true. □

§ 4b. Convergence of Point Processes

In this subsection we suppose that X, in addition to being a PII, is also a point process: that is, X is an extended Poisson process in the sense of I.3.26. We call A its compensator: by I.3.27 A is non-random and $A_t = E(X_t)$. Conversely, recall also that if X is a point process with deterministic compensator, it is an extended Poisson process. The characteristics of X are again given by 3.35.

4.10 **Theorem.** *We suppose that X is an extended Poisson process with intensity A; we suppose that each X^n is a point process with compensator A^n; let D be a dense subset of $\mathbb{R}_+$.*

a) *Under the two conditions*

4.11D $$A_t^n \xrightarrow{P} A_t \quad \text{for all } t \in D,$$

4.12D $$\sum_{s \le t} (\Delta A_s^n)^2 \xrightarrow{P} \sum_{s \le t} (\Delta A_s)^2 \quad \text{for all } t \in D,$$

we have $X^n \xrightarrow{\mathscr{L}(D)} X$ and $X^n \xrightarrow{\mathscr{L}} X$.

b) *If $A^n \xrightarrow{P} A$ for the Skorokhod topology in $\mathbb{D}(\mathbb{R})$, then $X^n \xrightarrow{\mathscr{L}} X$.*

Recall that 4.11D + 4.12D $\Rightarrow A^n \xrightarrow{P} A$ in $\mathbb{D}(\mathbb{R})$, while conversely $A^n \xrightarrow{P} A$ in $\mathbb{D}(\mathbb{R}) \Rightarrow$ 4.11D + 4.12D if $D = \{t\colon P(\Delta X_t \neq 0) = 0\}$ (see VI.2.15) hence (b) follows from (a). Also recall that when A is continuous (that is, X has no fixed time of discontinuity) we have 4.11D $\Rightarrow$ 4.12D: all this comes from VI.2.15. Therefore Theorem 3.36, when D is dense, is a particular case of the above. Note also that in the situation of the theorem, $X^n \xrightarrow{\mathscr{L}(D)} X$ implies $X^n \xrightarrow{\mathscr{L}} X$ by VI.3.37.

Proof. Take a truncation function h satisfying $h(1) = 0$. Then $B^n = B = 0$, $\tilde{C}^n = \tilde{C} = 0$, and $g * \nu^n = g(1)A^n$; so we have $[\text{Sk-}\beta_5]$ and $[\gamma_5\text{-}D]$. We also have $[\delta_{5,1}\text{-}D]$ by 4.11D, and $[\bar{\delta}_{5,1}\text{-}D]$ (see 4.2) by 4.12D, hence $[\text{Sk-}\delta_{5,1}]$. Finally, since X_t takes its values in $\mathbb{N}$, the set $U_t = \{u\colon g(u)_t = 0\}$ is discrete (or empty). Then 1.17 holds for all $t > 0$. Therefore (a) follows from 4.1 and 4.6. □

4.13 **Remark.** Of course, using 4.1 and 4.6 to prove this theorem is like using a sledge-hammer to crack a nut. The easiest way, see [121], consists in using the Laplace transform $\hat{g}(\lambda)_t = E(\exp -\lambda X_t)$, $\lambda \ge 0$, instead of the Fourier transform $g(u)_t$. However, it should be noted that the idea is always the same: use a version of Theorem 1.9 (this is easier with the Laplace transform when it exists, like here, because $\hat{g}(\lambda)_t$ cannot vanish), and prove that 4.11D + 4.12D imply $\hat{G}^n(\lambda)_t \xrightarrow{P} \hat{g}(\lambda)_t$, where $\hat{G}^n(\lambda)$ is the "Laplace version" of $G^n(u)$. □

§ 4c. Convergence to a Gaussian Martingale

We suppose that each X^n is a 1-dimensional semimartingale, and that X is a *Gaussian 1-dimensional martingale*. Recall that if

4.14 $$\tilde{C}'_t = E(X_t^2),$$

the characteristics of X (see § II.4d) are

4.15 $$\begin{cases} B_t = \sum_{s\leq t} \phi_{\Delta\tilde{C}'_s}(h) \\ C_t = \tilde{C}'_t - \sum_{s\leq t} \Delta\tilde{C}'_s \\ \nu(ds,dx) = \sum_{s>0,\Delta\tilde{C}'_s>0} \varepsilon_s(dt)\otimes\phi_{\Delta\tilde{C}'_s}(dx) \end{cases}$$

where ϕ_a is the normal distribution $\mathcal{N}(0,a)$ with mean 0 and variance $a \geq 0$. Note once more that $\tilde{C}'$ is related to (B, C, ν) by 2.13.

We leave for the reader to translate Theorems 4.1 and 4.6 to this setting; there is no difficulty, since $g(u)_t = \exp -u^2\tilde{C}'_t/2$ does not vanish. Here we concentrate on the "non-classical" situation of Zolotarev's type theorems, as expounded in §§ VII.5c,d.

We recall the notation of these subsections. Set $J^n = \{(\omega,t)\colon \nu^n(\omega;\{t\}\times\mathbb{R}) > 0\}$ and $\nu^{n,c}(\omega;dt,dx) = \nu^n(\omega;dt,dx)1_{(J^n)^c}(\omega,t)$. We suppose that the truncation function h and the function f satisfy the condition VII.5.7. For each (ω,s) we denote by $\Delta^n_s(\omega)$ the nonnegative number that satisfies

4.16 $$\nu^n(\omega;\{s\}\times f) = \phi_{\Delta^n_s(\omega)}(f),$$

and we set $B^{n,c}_t = B^n_t - \sum_{s\leq t}\Delta B^n_s$.

If D is a subset of $\mathbb{R}_+$, we consider the conditions:

$[A_3\text{-}D]$ $\quad \nu^{n,c}([0,t]\times\{|x|>\varepsilon\}) \xrightarrow{P} 0 \quad$ for all $\varepsilon > 0$, $t\in D$

$[B_3\text{-}D]$ $\quad |B^{n,c}_t| + \sum_{s\leq t}|\Delta B^n_s| \xrightarrow{P} 0 \quad$ for all $t \in D$

$[C_3\text{-}D]$ $\quad C^n_t + h^2 * \nu^{n,c}_t + \sum_{s\leq t}\Delta^n_s \xrightarrow{P} \tilde{C}'_t \quad$ for all $t\in D$

$[\tilde{C}_3\text{-}D]$ $\quad \sum_{s\leq t}(\Delta^n_s)^2 \xrightarrow{P} \sum_{s\leq t}(\Delta\tilde{C}'_s)^2 \quad$ for all $t\in D$

$[D_3\text{-}D]$ $\quad \sum_{s\leq t}\int_{|x|>\varepsilon}|\hat{\eta}^n_s(x) - \hat{\phi}_{\Delta^n_s}(x)|\,dx \xrightarrow{P} 0 \quad$ for all $\varepsilon > 0$, $t\in D$, where $\hat{\phi}_a(x) = \phi_a((-\infty, x])$ and $\hat{\eta}_s^n(x) = \eta_s^n((-\infty, x])$ and

$$\eta_s^n(\omega, dx) = v^n(\omega; \{s\} \times dx) + \varepsilon_0(dx)[1 - v^n(\omega; \{s\} \times \mathbb{R})].$$

4.17 **Theorem.** *With the above notation and assumptions, we have*

a) *Under* $[A_3\text{-}D]$, $[B_3\text{-}D]$, $[C_3\text{-}D]$, $[D_3\text{-}D]$, *then* $X^n \xrightarrow{\mathscr{L}(D)} X$.

b) *Under* $[A_3\text{-}D]$, $[B_3\text{-}D]$, $[C_3\text{-}D]$, $[\tilde{C}_3\text{-}D]$, $[D_3\text{-}D]$, *and if* D *is dense in* $\mathbb{R}_+$, *we have* $X^n \xrightarrow{\mathscr{L}} X$.

Proof. a) We can assume that D is at most countable. By 1.18 it suffices to prove that 1.10 holds for all $t \in D$. Take a subsequence (n'); by a diagonal argument there is a set A with $P(A) = 1$ and a further subsequence (n'') such that in $[A_3\text{-}D]$, $[B_3\text{-}D]$, $[C_3\text{-}D]$, $[D_3\text{-}D]$ there is convergence along (n'') for all $\omega \in A$. By VII.5.24 it follows that $G^{n''}(u)_t(\omega) \to g(u)_t$ for all $\omega \in A$, so we deduce 1.10, and the result follows from Theorem 1.9.

b) As above, we extract from (n') a subsequence (n'') and we find a set A with $P(A) = 1$ such that in $[A_3\text{-}D]$, $[B_3\text{-}D]$, $[C_3\text{-}D]$, $[\tilde{C}_3\text{-}D]$, $[D_3\text{-}D]$ there is convergence along (n'') for all $\omega \in A$. Then we deduce from VII.5.26 and VII.3.13 that 4.4 holds. Therefore we have $[\text{Sk-}\beta_5]$ $[\gamma_5\text{-}D]$ and $[\text{Sk-}\delta_{5,1}]$ and the result follows from 4.1. □

The "square-integrable" version is as follows. We assume 2.11 for each n and we define B'^n and $\tilde{C}'^n$ by 2.12 and 2.13. Let $B_t'^{n,c} = B_t'^n - \sum_{s \le t} \Delta B_s'^n$ and $\Delta_s'^n(\omega) = v^n(\omega; \{s\} \times x^2)$. Consider the conditions:

$[A'_3\text{-}D]$ $\quad (x^2 1_{\{|x|>\varepsilon\}}) * v_t^{n,c} \xrightarrow{P} 0 \quad$ for all $\varepsilon > 0, t \in D$

$[B'_3\text{-}D]$ $\quad |B_t'^{n,c}| + \sum_{s \le t} |\Delta B_s'^n| \xrightarrow{P} 0 \quad$ for all $t \in D$

$[C'_3\text{-}D]$ $\quad C_t^n + x^2 * v_t^{n,c} + \sum_{s \le t} \Delta_s'^n \xrightarrow{P} \tilde{C}'_t \quad$ for all $t \in D$

$[\tilde{C}'_3\text{-}D]$ $\quad \sum_{s \le t} (\Delta_s'^n)^2 \xrightarrow{P} \sum_{s \le t} (\Delta \tilde{C}'_s)^2 \quad$ for all $t \in D$

$[D'_3\text{-}D]$ $\quad \sum_{s \le t} \int_{|x|>\varepsilon} |x| \, |\hat{\eta}_s^n(x) - \hat{\phi}_{\Delta_s'^n}(x)| \, dx \xrightarrow{P} 0 \quad$ for all $\varepsilon > 0, t \in D$,

Then, the same proof than above, using VII.5.25 (resp. VII.5.36) instead of VII.5.24 (resp. VII.5.26) yields:

4.18 **Theorem.** *With the above notation and assumptions, we have*:

a) *Under* $[A'_3\text{-}D]$, $[B'_3\text{-}D]$, $[C'_3\text{-}D]$, $[D'_3\text{-}D]$, *then* $X^n \xrightarrow{\mathscr{L}(D)} X$.

b) *Under* $[A'_3\text{-}D]$, $[B'_3\text{-}D]$, $[C'_3\text{-}D]$, $[\tilde{C}'_3\text{-}D]$, $[D'_3\text{-}D]$, *and if* D *is dense in* $\mathbb{R}_+$, *we have* $X^n \xrightarrow{\mathscr{L}} X$.

The version of these theorems for triangular arrays is left to the reader.

5. Convergence to a Mixture of PII's, Stable Convergence and Mixing Convergence

This section is devoted to various improvements or extensions of the previous results. Their interest lies in applications, mainly to statistics. Mathematically speaking, they are very simple (although sometimes tedious) generalizations. The reader might skip the whole section without inconvenience for the next chapter.

§ 5a. Convergence to a Mixture of PII's

The setting is as follows. We have a single probability space $(\Omega', \mathscr{F}', P')$ on which are defined:

5.1 For each $n \in \mathbb{N}^*$, a d-dimensional semimartingale X^n, relative to some filtration $\mathbf{F}^n$, with $X_0^n = 0$ and with characteristics (B^n, C^n, ν^n).

5.2 A σ-field $\mathscr{G}$ satisfying $\mathscr{G} \subset \mathscr{F}_0^n$ for all $n \geq 1$.

5.3 A transition probability $Q(\omega', d\alpha)$ from $(\Omega', \mathscr{G})$ into $(\mathbb{D}(\mathbb{R}^d), \mathbb{D}(\mathbb{R}^d))$ such that for each $\omega' \in \Omega'$ the canonical process X on $\mathbb{D}(\mathbb{R}^d)$ (defined by $X_t(\alpha) = \alpha(t)$) is a PII-semimartingale with characteristics $(B(\omega'), C(\omega'), \nu(\omega'))$. These characteristics are clearly "$\mathscr{G}$-measurable"; we associate $A(u)_t(\omega')$ and $G(u)_t(\omega')$ to them by 1.8, and $\tilde{C}(\omega')$ by 1.3, and we have

5.4 $$G(u)_t(\omega') = \int Q(\omega', d\alpha) e^{iu \cdot \alpha(t)}$$

(*Remark*: the semimartingale property of X is added for simplicity; it could be relaxed without harm!) □

Now, we set

5.5 $\Omega = \Omega' \times \mathbb{D}(\mathbb{R}^d), \quad \mathscr{F} = \mathscr{F}' \otimes \mathscr{D}(\mathbb{R}^d), \quad P(d\omega', d\alpha) = P'(d\omega')Q(\omega', d\alpha)$

and every process or variable or σ-field defined on Ω' (resp. $\mathbb{D}(\mathbb{R}^d)$) is naturally extended to Ω with the same symbol: e.g. $X(\omega', \alpha) = X(\alpha)$, $X^n(\omega', \alpha) = X^n(\omega')$, $\mathscr{G} = \mathscr{G} \otimes \{\emptyset, \mathbb{D}(\mathbb{R}^d)\}$, etc... It is obvious that X^n is still an $\mathbf{F}^n$-semimartingale with the same characteristics (B^n, C^n, ν^n) on Ω than on Ω'. We also denote by $\mathbf{F} = (\mathscr{F}_t)_{t \geq 0}$ the smallest filtration of Ω to which X is adapted and such that $\mathscr{G} \subset \mathscr{F}_0$.

With these assumptions, it is clear that X is, under P, *a "mixture" of PII*'s. It is also a *$\mathscr{G}$-conditional PII* (see II.6.2), which meets *Hypothesis II.6.4*. Finally, in view of II.6.15 (or of II.6.5 as well), X is a *semimartingale* on the basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$, whose characteristics are (B, C, ν).

5.6 **Remarks.** 1) An apparently more general, but indeed equivalent, setting would be to start directly with a probability space $(\Omega, \mathscr{F}, P)$ on which $(X^n, \mathbf{F}^n)$ and $\mathscr{G}$ are defined like in 5.1 and 5.2, and on which is also defined a $\mathscr{G}$-conditional PII-semimartingale.

2) This setting encompasses the situation of § 1a: in the latter case, indeed, we can assume 1.6, and we have 5.1, 5.2, 5.3 if we set $\mathscr{G} = \{\varnothing, \Omega\}$ and $G(u)_t(\omega') = g(u)_t$. □

We begin by stating the main result when X is *quasi-left-continuous*: this amounts to saying that $\nu(\omega', \{t\} \times \mathbb{R}^d) = 0$ identically, or equivalently that for each $Q(\omega', \cdot)$, the PII X has no fixed time of discontinuity.

5.7 **Theorem.** *Assume that 5.1, 5.2, 5.3 hold and that X is quasi-left-continuous.*

a) *Under $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$, plus 2.5 whenever the subset D is not dense in $\mathbb{R}_+$, we have*

5.8 $$E[Yf(X^n_{t_1}, \ldots, X^n_{t_p})] \to E[Yf(X_{t_1}, \ldots, X_{t_p})]$$

for all $t_j \in D$, f continuous bounded on $(\mathbb{R}^d)^p$ and $Y \in b\mathscr{G}$ (= bounded $\mathscr{G}$-measurable). *This of course implies $X^n \xrightarrow{\mathscr{L}(D)} X$.*

b) *Under* [Sup-β_5] $+ [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ *where D is dense in $\mathbb{R}_+$, we have*

5.9 $$E[Yf(X^n)] \to E[Yf(X)]$$

for all f continuous bounded on $\mathbb{D}(\mathbb{R}^d)$ (with the Skorokhod topology) *and $Y \in b\mathscr{G}$. This of course implies $X^n \xrightarrow{\mathscr{L}} X$.*

This theorem is proved at the end of the subsection. Similarly we may get the following results, with 5.8 and 5.9 respectively instead of $X^n \xrightarrow{\mathscr{L}(D)} X$ and $X^n \xrightarrow{\mathscr{L}} X$:

5.10 If X is *quasi-left-continuous*, Theorems 2.14, 2.17, 2.27, 2.29 and 2.30 are valid.

5.11 If X is *continuous*, 3.6 and (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (i) in (b) of 3.8; (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (i) of 3.11; (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) $\Rightarrow$ (i) of 3.12 when X is a local martingale; (ii) $\Leftrightarrow$ (iii) $\Rightarrow$ (i) of 3.24 when X is a local martingale, are all valid.

5.12 If X is a *quasi-left-continuous point process* (i.e. a Cox process), Theorem 3.36 is valid.

Note that in 5.11, however, the "necessary" parts (i) $\Rightarrow$ (ii) in 3.8 or 3.11 or 3.12 or 3.24 *are not true here*: indeed, if $X^n \xrightarrow{\mathscr{L}} X$ then VI.6.6 (or VI.6.1) yields $\hat{C}^n \xrightarrow{\mathscr{L}} C$; when C is deterministic, this amounts to $[\hat{\gamma}_5\text{-}\mathbb{R}_+]$; but in the present situation $[\hat{\gamma}_5\text{-}\mathbb{R}_+]$ has no reason to be true.

When X is not quasi-left-continuous, $G(u)$ may vanish and we introduce the following condition, which extends 1.17:

5.13 $\begin{cases} P \otimes \lambda(U_t) = 0, \text{ where } \lambda \text{ is the Lebesgue measure on } \mathbb{R}^d, \text{ and } U_t \\ \text{is the } \mathscr{G} \otimes \mathbb{R}^d\text{-measurable set } U_t = \{(\omega, u)\colon G(u)_t(\omega) = 0\}. \end{cases}$

5.14 **Theorem.** *Assume 5.1, 5.2 and 5.3. Let D be a dense subset of $\mathbb{R}_+$.*

a) *Under 5.13 for all $t > 0$, $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D] + [\text{Sk-}\delta_{5,1}]$, we have 5.8 for all $t_j \in D$, f continuous bounded on $(\mathbb{R}^d)^p$, $Y \in b\mathscr{G}$, and in particular we have $X^n \xrightarrow{\mathscr{L}(D)} X$.*

b) *Under $[\text{Sk-}\beta_5] + [\gamma_5\text{-}D] + [\text{Sk-}\delta_{5,1}]$, we have 5.9 for all f continuous bounded on $\mathbb{D}(\mathbb{R}^d)$ and $Y \in b\mathscr{G}$, and in particular we have $X^n \xrightarrow{\mathscr{L}} X$.*

Similarly, we also have the following:

5.15 Theorem 4.10 when X is a *point process*, and Theorems 4.17 and 4.18 when X is a *mixture of Gaussian martingales* (i.e., a Gaussian martingale for each measure $Q(\omega', \cdot)$), are valid.

The proofs of 5.7 and 5.14 are based upon the following extension of Theorem 1.9.

5.16 **Theorem.** *Assume 5.1, 5.2, 5.3. Let $D \subset \mathbb{R}_+$ and assume that 5.13 and*

5.17 $$G^n(u)_t \xrightarrow{P} G(u)_t \quad \text{for all } u \in \mathbb{R}^d$$

hold for all $t \in D$. Then we have 5.8 for all $t_j \in D$, f continuous bounded on $(\mathbb{R}^d)^p$ and $Y \in b\mathscr{G}$.

Proof. In 5.8 we may assume that $t_1 < \cdots < t_p$. We prove 5.8 by induction on p, that is, we suppose that it is true for $p-1$. In particular if $u_j \in \mathbb{R}^d$ and $\zeta^n = \exp i \sum_{1 \le j \le p-1} u_j \cdot X^n_{t_j}$, we have

5.18 $$E(Y\zeta^n) \to E(Y\zeta) \quad \text{for all } Y \in b\mathscr{G}.$$

We will prove that for each $Y \in b\mathscr{G}$,

5.19 $$\begin{cases} v^n(u) := E(Y\zeta^n \exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}})) \\ \qquad \to v(u) := E(Y\zeta \exp iu \cdot (X_{t_p} - X_{t_{p-1}})) \end{cases}$$

for λ-almost all $u \in \mathbb{R}^d$, and this will imply 5.8. Because of 5.13, it suffices to prove 5.19 for all u in the set $U = \{u\colon P(G(u)_{t_p} = 0) = 0\}$. In the sequel, we fix u in U.

Let $A = |G(u)_{t_p}|$, and $a > 0$, and

$$v^n_a = E(Y1_{\{A \ge a\}}\zeta^n \exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}}))$$
$$v_a = E(Y1_{\{A \ge a\}}\zeta \exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}})),$$

and $K = \sup|Y|$. We have $|v_a^n - v^n(u)| \le KP(A < a)$, and similarly for $|v_a - v(u)|$. Since $u \in U$, $P(A < a) \downarrow 0$ as $a \downarrow 0$. Therefore it suffices to prove that $v_a^n \to v_a$ as $n \uparrow \infty$, for all $a > 0$, in order to get 5.19.

So in the sequel we fix $a > 0$. Define R^n and S^n (associated with a) like in the proof of 1.9, so that $P(R^n \le t_p, A \ge a) \to 0$ and

5.20 $$P(S^n \le t_p, A \ge a) \to 0.$$

We also define M^n, β^n, γ^n like in 1.9, and $\gamma = G(u)_{t_p}/G(u)_{t_{p-1}}$ (with $0/0 = 0$). We still have

5.21 $$E(\beta^n | \mathscr{F}^n_{t_{p-1}}) = 1.$$

Moreover, $\exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}}) = \beta^n \gamma^n$ on $\{S^n \ge t_p\}$. Hence

$$v_a^n = E(Y 1_{\{A \ge a\}} 1_{\{S^n \le t_p\}} \zeta^n \exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}})) + E(Y 1_{\{A \ge a, S^n > t_p\}} \zeta^n \beta^n \gamma^n)$$

5.22 $$= \begin{cases} E(Y 1_{\{A \ge a, S^n \le t_p\}} \zeta^n \exp iu \cdot (X^n_{t_p} - X^n_{t_{p-1}})) \\ + E(Y 1_{\{A \ge a\}} \zeta^n \beta^n (\gamma^n 1_{\{S^n > t_p\}} - \gamma)) + E(Y 1_{\{A \ge a\}} \zeta^n \gamma) \end{cases}$$

(because of 5.21). Moreover, 5.3 implies that $v_a = E(Y 1_{\{A \ge a\}} \zeta \gamma)$, hence since $|Y| \le K$ and $|\zeta^n| = 1$, we obtain

5.23 $$\begin{cases} |v_a^n - v_a| \le KP(A \ge a, S^n \le t_p) + KE(1_{\{A \ge a\}} |\beta^n| |\gamma^n 1_{\{S^n > t_p\}} - \gamma|) \\ \qquad + |E(Y 1_{\{A \ge a\}} \gamma \zeta^n) - E(Y 1_{\{A \ge a\}} \gamma \zeta)|. \end{cases}$$

Consider the right-hand side of 5.23: on $\{A \ge a\}$ we have $|\beta^n \gamma| \le 2/a^2$, while $|\beta^n \gamma^n| = 1$ on $\{S^n > t_p\}$, and $[\gamma^n 1_{\{S^n > t_p\}} - \gamma] 1_{\{A \ge a\}} \xrightarrow{P} 0$ by 5.17 and 5.20; hence the second term in 5.23 tends to 0 as $n \uparrow \infty$; the first term tends to 0 by 5.20, the third term also tends to 0, because of 5.18 applied to $Y' = Y 1_{\{A \ge a\}} \gamma$ (we have $|Y'| \le K/a$). Therefore $v_a^n \to v_a$, and the theorem is proved. □

Proof of 5.7 *and* 5.14. In order to prove part (a) of each of these theorems we can literally reproduce the proof of 2.4 and 4.6, which give 5.17 for all $t \in D$. Then the previous theorem (instead of 1.18) yields 5.8.

Now we prove (b). In fact, we will prove 5.7b, and 5.14b under the additional assumption 5.13 (exactly as the proof of 4.1 given in § 4a, which requires 1.17; the result in general can easily be deduced from a mild extension of Theorem IX.3.35 of the next chapter).

By linearity, it is enough to prove 5.9 for $Y \ge 0$ with $E(Y) = 1$. Let $a = \sup|Y|$ and consider the new probability measure $\tilde{P}(d\omega) = P(d\omega) Y(\omega)$, so that $\tilde{E}[f(X^n)] = E(Yf(X^n))$ for all bounded measurable functions f on $\mathbb{D}(\mathbb{R}^d)$, where $\tilde{E}$ denotes the expectation with respect to $\tilde{P}$. By the first part of the proof, $X^n \xrightarrow{\mathscr{L}(D)} X$ under $\tilde{P}$, so it remains to prove that, under $\tilde{P}$ again, the sequence (X^n) is tight. If K is a compact subset of $\mathbb{D}(\mathbb{R}^d)$ we have $\tilde{P}(X^n \notin K) \le aP(X^n \notin K)$, so indeed it is sufficient to prove that the sequence (X^n) is tight under the measure P.

For that we will apply Theorems VI.4.18 (for proving 5.7) and VI.5.10 (for 5.14). Condition (i) of these two theorems is trivially met here. We can reproduce the proofs of 2.17 (resp. 4.3) to obtain that [Sup-γ_5] and [Sup-$\delta_{5,2}$] (resp. [Sk-$\beta\gamma\delta_5$]) hold in the case of 5.7 (resp. 5.14). Let $g_p(x) = (p|x| - 1)^+ \wedge 1$, which belongs to $C_2(\mathbb{R}^d)$. Let $\zeta > 0$, $\eta > 0$, $N > 0$ with $\nu(\{N\} \times \mathbb{R}^d) = 0$ a.s.; there exists $p > 0$ such that $P(g_p * \nu_N > \frac{\varepsilon}{2}) \leq \frac{\eta}{2}$ (because $g_p * \nu_N \downarrow 0$ as $p \downarrow 0$), while $g_p * \nu_N^n \xrightarrow{P} g_p * \nu_N$, so for all n large enough we have $P(g_p * \nu_N^n > \varepsilon) \leq \eta$. Since $\nu^n([0, N] \times \{|x| > 1/2p\}) \leq g_p * \nu_N^n$ we deduce that condition (ii) of VI.4.18 (resp. VI.5.10) is met.

In the case of 5.7, [Sup-β_5] + [Sup-γ_5] + [Sup-$\delta_{5,2}$] and the property that $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all t immediately imply that VI.4.18iii is met. Finally in case of 5.14, condition (iii) of VI.5.10 is met in virtue of [Sk-β_5]. Moreover, [Sk-$\beta\gamma\delta_5$] implies that for all $p > 0$, $G^{n,p} = \sum_{j \leq d} \tilde{C}^{n,jj} + g_p * \nu^n$ converge in measure (as $n \uparrow \infty$) to $G^p = \sum_{j \leq d} \tilde{C}^{jj} + g_p * \nu$. Furthermore, by 5.3 the process G^p is predictable with respect to the trivial filtration $\mathscr{G}_t = \mathscr{G}$, which by 5.2 is included in $\mathbf{F}^n$ for all n: therefore condition (iv) of VI.5.10 is met with (C3). So we deduce that in both cases the sequence (X^n) is tight, and the claims are proved. □

§ 5b. More on the Convergence to a Mixture of PII's

The aim of this subsection is to show that under a very mild additional assumption on the setting of § 5a, the results can be considerably strenghtened. As a corollary, it also gives another proof of the results of § 5a, which is not based upon Theorem 5.16.

The setting is as follows (essentially, it is the same as in § 5a, plus the fact that each basis $(\Omega', \mathscr{F}^n_{\infty-}, \mathbf{F}^n, P')$ satisfies II.6.4):

5.24 *Hypothesis.* We assume 5.1, 5.2 and 5.3. We also assume that for each $n \geq 1$ there is a family of separable σ-fields $(\hat{\mathscr{F}}_t^n)_{t \geq 0}$ on Ω' such that $\mathscr{F}^n_{t-} \subset \hat{\mathscr{F}}_t^n \subset \mathscr{F}_t^n$. Finally, we assume that for each $n \geq 1$, there is a regular version $Q^n(\omega', d\omega'')$ of the conditional probability on $(\Omega', \mathscr{F}^n_{\infty-})$ with respect to $\mathscr{G}$. □

Here is the improvement upon Theorems 5.7 and 5.14 (under 5.24). We only give the functional version, leaving to the reader the corresponding statement about finite-dimensional convergence.

5.25 **Theorem.** *Assume 5.24. Suppose that*

a) *when X is quasi-left-continuous, we have* [Sup-β_5], [γ_5-D], [$\delta_{5,1}$-D] *for some dense subset $D \subset \mathbb{R}_+$;*

b) *or, we have* [Sk-β_5] , [γ_5-D], [Sk-$\delta_{5,1}$] *for some dense subset $D \subset \mathbb{R}_+$* (note that (a) ⇒ (b)).

Then, if $\hat{Q}^n(\omega', \cdot)$ is the distribution of X^n under the measure $Q^n(\omega', d\omega'')$, we have

5.26 $$\hat{Q}^n \xrightarrow{P} Q \quad (Q \text{ is defined in 5.3})$$

5.26 means the following: $\hat{Q}^n$ and Q can be considered as random variables on $(\Omega', \mathscr{G})$, taking their values in the Polish space of all probability measures on $\mathbb{D}(\mathbb{R}^d)$, equipped with the weak topology (and with the Skorokhod topology on $\mathbb{D}(\mathbb{R}^d)$). Then 5.26 is just the convergence in probability (for P or P', this is the same here) in this Polish space.

Of course, 5.26 implies 5.9, and in fact is much stronger (5.9 is a sort of convergence for the weak L^1 topology, as we shall see later).

Proof. It suffices to prove that if (n') is an infinite subsequence, there is a further subsequence (n'') such that $\hat{Q}^{n''}(\omega', \cdot) \to Q(\omega', \cdot)$ weakly, for P'-almost all ω'.

Let first Z^n and Z be random variables, taking values in a metric space (E, δ), and suppose that Z^n is $\mathscr{F}^n_{\infty-}$-measurable and Z is $\mathscr{G}$-measurable. Suppose that $Z^n \xrightarrow{P} Z$. Then

$$E[\delta(Z^n, Z) \wedge 1] = \int P'(d\omega') \int Q^n(\omega', d\omega'')\delta(Z^n(\omega''), Z(\omega')) \wedge 1 \to 0.$$

We can find a subsequence (n'') of (n') such that $\int Q^{n''}(\cdot, d\omega'')\delta(Z^n(\omega''), Z(\cdot)) \wedge 1 \to 0$ P-a.s. In other words, if $\xrightarrow{Q^n_{\omega'}}$ denotes the convergence in measure with respect to the sequence $\{Q^n(\omega', \cdot)\}_{n \geq 1}$, towards a function that is $\mathscr{G}$-measurable, we have just shown that $Z^n \xrightarrow{Q^{n''}_{\omega'}} Z(\omega')$ for P-almost all ω'.

Using this auxiliary result, plus a diagonal argument, we can find a subsequence (n'') and a set $A \subset \Omega'$ with $P'(A) = 1$, such that for all $\omega' \in A$ we have

$$\begin{cases} B^{n''} \xrightarrow{Q^{n''}_{\omega'}} B(\omega') & \\ \tilde{C}^{n''}_t \xrightarrow{Q^{n''}_{\omega'}} \tilde{C}_t(\omega') & \text{for all } t \in D \ (D \text{ may be assumed countable}) \\ g * \nu^{n''} \xrightarrow{Q^{n''}_{\omega'}} g * \nu(\omega') & \text{for all } g \in C_1(\mathbb{R}^d). \end{cases}$$

Moreover, we deduce from Corollary II.6.15 (in which the measurability properties of B, C, ν play no rôle; observe that, because of 5.24, the assumptions of this corollary are met) that for all ω' belonging to a P'-full set $A' \subset \Omega'$ each X^n is a $Q^n(\omega', \cdot)$-semimartingale with characteristics (B^n, C^n, ν^n).

Therefore if $\omega' \in A \cap A'$, the sequence $(X^{n''})$ and the limiting process X satisfy all the assumptions of Theorem 4.1, under the measures $Q^{n''}(\omega', \cdot)$ and $Q(\omega', \cdot)$ respectively. Recalling the definition of $\hat{Q}^n$, we deduce that $\hat{Q}^{n''}(\omega', \cdot) \to Q(\omega', \cdot)$ weakly for all $\omega' \in A \cap A'$, and the theorem is proved. $\square$

5.27 **Remarks.** 1) Suppose that 5.24 holds, and that in [Sup-β_5], [γ_5-D], etc., we have almost sure convergence instead of convergence in probability. Then the previous proof shows that 5.26 can be replaced by:

$$Q^n \to Q \quad \text{a.s.}$$

2) Assume 5.24. Not only do the "sufficient" conditions for convergence, in Section 3, give the strong convergence 5.26, but the "necessary" conditions work as well here.

For example, consider Theorem 3.8: we suppose that, in addition to 5.24, X is continuous, with characteristics $(B, C, 0)$, and D is dense. Then under [Sup-β_5], we have the equivalence of 5.26, and 3.8(ii), and 3.8(iii). We leave the details to the reader, the proof being quite similar to that of 5.25. □

§ 5c. Stable Convergence

1. The stable convergence, as well as the mixing convergence to be seen in the next subsection, have been introduced by Renyi [206, 207].

To match with the notation of § 5a, we use *a*-priori strange notation here. Let $(\Omega', \mathscr{F}', P')$ be a probability space endowed with a sequence (Z^n) of random variables, taking their values in a Polish space $(E, \mathscr{E})$.

5.28 **Definition.** Let $\mathscr{G}$ be a sub-σ-field of $\mathscr{F}'$. We say that the sequence (Z^n) *converges $\mathscr{G}$-stably*, or *stably* if $\mathscr{G} = \mathscr{F}'$, if there is a probability measure μ on $(\Omega' \times E, \mathscr{G} \otimes \mathscr{E})$ such that

5.29 $$E[Yf(Z^n)] \to \int_{\Omega' \times E} \mu(d\omega', dx) Y(\omega') f(x)$$

for all $Y \in b\mathscr{G}$, $f \in C(E)$ (= continuous and bounded on E). □

Often one considers only the case when $\mathscr{G} = \mathscr{F}'$. The above definition is apparently stronger than the usual one (when $\mathscr{G} = \mathscr{F}'$), but we shall see in the next proposition that it is indeed the same.

Of course, if (Z^n) converges stably, it also converges in law, and the limiting law is the marginal $\mu(\Omega' \times \cdot)$. In fact, there is a transition probability $Q(\omega', dx)$ from $(\Omega', \mathscr{G})$ on $(E, \mathscr{E})$ such that

5.30 $$\mu(d\omega', dx) = P'(d\omega') Q(\omega', dx) \text{ on } \mathscr{G} \otimes \mathscr{E}.$$

Then it is natural to set

5.31 $$\Omega = \Omega' \times E, \quad \mathscr{F} = \mathscr{F}' \otimes \mathscr{G}, \quad P(d\omega', dx) = P'(d\omega') Q(\omega', dx) \quad \text{on } \mathscr{F}$$

(if $\mathscr{G} = \mathscr{F}'$ we have $P = \mu$; otherwise, P is an extension of μ). We extend Z^n, or any random variable, or any σ-field on Ω' to Ω in the usual fashion; set $Z(\omega', x) = x$ be the "canonical" variable on E. Finally call $\hat{Q}^n(\omega', dx)$ a regular version of the conditional distribution of Z^n with respect to $\mathscr{G}$ (if $\mathscr{G} = \mathscr{F}'$, then $\hat{Q}^n(\omega', dx) = \varepsilon_{Z^n(\omega')}(dx)$). We have

$$E(Yf(Z^n)) = E(Y\hat{Q}^n f), \quad \int \mu(d\omega', dx) Y(\omega') f(x) = E(YQf).$$

Therefore, 5.29 reads as follows:

5.32 (Z^n) converges $\mathscr{G}$-stably (to μ) if and only if $\hat{Q}^n f \to Qf$ weakly in $L^1(\Omega, \mathscr{G}, P)$ for all $f \in C(E)$; we say also: (Z^n) converges $\mathscr{G}$*-stably* to Z (the canonical variable). □

In particular, if (Z^n) converges $\mathscr{G}$-stably and if A is any $\mathscr{G}$-measurable subset of Ω' with positive probability, then the conditional distributions $\mathscr{L}(Z^n/A)$ converge. As will follow from the next result, the converse is also true, and this was the original definition of the stable convergence.

5.33 **Proposition.** *There is equivalence between*

(i) (Z^n) *converges* $\mathscr{G}$*-stably;*

(ii) *for every* $\mathscr{G}$*-measurable random variable* Y *on* Ω', (Y, Z^n) *converges in law;*

(iii) *for every* $\mathscr{G}$*-measurable random variable* Y *on* Ω', (Y, Z^n) *converges* $\mathscr{G}$*-stably;*

(iv) *the sequence* (Z^n) *is tight, and for all* $A \in \mathscr{G}, f \in C(E)$, *the sequence* $E(1_A f(Z^n))$ *converges;*

(v) (when $E = \mathbb{R}^d$) *the sequence* (Z^n) *is tight, and for all* $A \in \mathscr{G}$, $u \in \mathbb{R}^d$, *the sequence* $E(1_A \exp iu \cdot Z^n)$ *converges.*

We begin with an auxiliary lemma on "bi-measures".

5.34 **Lemma.** *Let* L *be a mapping:* $\mathscr{G} \times \mathscr{E} \to [0,1]$ *such that* $L(\Omega', E) = 1$ *and*

(i) *for each* $A \in \mathscr{G}$, $B \rightsquigarrow L(A,B)$ *is a measure on* $(E, \mathscr{E})$

(ii) *for each* $B \in \mathscr{E}$, $A \rightsquigarrow L(A,B)$ *is a measure on* $(\Omega', \mathscr{G})$.

Then, there exists a unique probability measure μ *on* $(\Omega' \times E, \mathscr{G} \otimes \mathscr{E})$ *such that* $\mu(A \times B) = L(A,B)$ *for all* $(A,B) \in \mathscr{G} \times \mathscr{E}$.

Proof. The formula $\mu(A \times B) = L(A,B)$ extends trivially as a finitely additive measure on the algebra $\mathscr{F}^0$ generated by $\mathscr{G} \times \mathscr{E}$ on $\Omega' \times E$. It suffices to prove that $\mu(C_n) \to 0$ if $C_n \in \mathscr{F}^0$ with $C_n \downarrow \phi$.

Each C_n is of the form $C_n = \bigcup_{i \le p_n} A_n^i \times B_n^i$. Let $\varepsilon > 0$, $n \ge 1$, $i \le p_n$. From (i) there is a compact set K_n^i in E such that $K_n^i \subset B_n^i$ and

$$\mu(A_n^i \times K_n^i) = L(A_n^i, K_n^i) \ge L(A_n^i, B_n^i) - \frac{\varepsilon}{p_n} 2^{-n} = \mu(A_n^i \times B_n^i) - \frac{\varepsilon}{p_n} 2^{-n}.$$

Let $C_n' = \bigcap_{m \le n} \bigcup_{i \le p_m} A_m^i \times K_m^i$. We have $C_n' \subset C_n$ and $C_n' \in \mathscr{F}^0$ and since the sequence (C_n) decreases we have

$$\mu(C_n') \ge \mu(C_n) - \sum_{m \le n} \sum_{i \le p_m} \frac{\varepsilon}{p_m} 2^{-m} \ge \mu(C_n) - \varepsilon.$$

Each section $C_n'(\omega') = \{x: (\omega', x) \in C_n'\}$ is compact, and $\bigcap C_n' = \phi$. Hence if $F_n = \{\omega': C_n'(\omega') \ne \phi\}$ we have $\lim_n \downarrow F_n = \phi$. Since

$$\mu(C_n') \le \mu(F_n \times E) = L(F_n, E) \to 0$$

from (ii), we deduce that $\mu(C_n) \downarrow 0$. □

Proof of 5.33. (i) $\Rightarrow$ (iv) is trivial. Conversely, assume (iv). The tightness condition implies that for each $A \in \mathscr{G}$, the sequence of measures $\{\int_A P(d\omega')\hat{Q}^n(\omega', \cdot)\}_{n \geq 1}$ on E is tight; then the second assumption in (iv) implies that this sequence converges weakly to a measure $L(A, \cdot)$ on E. For each nonnegative $f \in C(E)$, the function $A \rightsquigarrow L(A, f)$ is a measure on $(\Omega', \mathscr{G})$ by Vitali-Hahn-Saks Theorem [188]. Therefore it is clear that $A \rightsquigarrow L(A, B)$ is additive for all $B \in \mathscr{E}$; moreover, $L(A, B) \leq L(A, E) = L(A, 1)$, which is a probability measure in A: it easily follows that $A \rightsquigarrow L(A, B)$ is a measure for all $B \in \mathscr{E}$. Then the previous lemma yields a probability measure μ on $(\Omega' \times E, \mathscr{G} \otimes \mathscr{E})$ which satisfies 5.29 by construction for all $Y = 1_A$, $A \in \mathscr{G}$. By linearity and uniform approximation, it also satisfies 5.29 for all $Y \in b\mathscr{G}$, and we have (i).

That (i) $\Leftrightarrow$ (v), when $E = \mathbb{R}^d$, is proved similarly. (i) $\Rightarrow$ (ii) is trivial, because for all f, g continuous bounded we have that $E(f(Y)g(Z^n)) \to \mu(f \otimes g)$ under (i). (ii) $\Rightarrow$ (iii) is also easy, since (Y, Z^n) converges $\mathscr{G}$-stably if and only if (Y', Y, Z^n) converges in law for all $Y' \in b\mathscr{G}$. Finally (iii) $\Rightarrow$ (i) is trivial. □

5.35 **Remark.** If 5.33(i) holds, and if μ is the limit, then for all Y, $\mathscr{G}$-measurable, and all functions $f \in C(\mathbb{R} \times \mathbb{R}^d)$, then $E(f(Y, Z^n)) \to \mu(f)$. In fact one can prove much more (see [110]): we do have $\int P(d\omega')f(\omega', Z^n(\omega')) \to \mu(f)$ for every bounded $\mathscr{G} \otimes \mathscr{E}$-measurable function f, such that each section $f(\omega', \cdot)$ is continuous on E. □

2. Now we turn back to the situation of § 5a. Firstly, assume 5.1, 5.2 and 5.3: we are exactly in the situation of the present subsection, with $Z^n = X^n$ and $Z = X$ and Q which is given by 5.3.

In other words, Theorems 5.7 and 5.14 could be stated as follows:

5.36 Under the conditions of 5.7(a) or 5.14(a), for all $t_j \in D$, the sequence $(X^n_{t_1}, \ldots, X^n_{t_p})$ *converges $\mathscr{G}$-stably to* $(X_{t_1}, \ldots, X_{t_p})$ (here $E = (\mathbb{R}^d)^p$).

Under the conditions of 5.7(b) or 5.14(b), X^n *converge $\mathscr{G}$-stably to* X (here, $E = \mathbb{D}(\mathbb{R}^d)$). □

The notation $\hat{Q}^n$ in this paragraph coincide with the notation $\hat{Q}^n$ of 5.25, when $Z^n = X^n$. In fact the convergence obtained under [Sk-β_5] + [γ_5-D] + [Sk-$\gamma_{5,1}$] is essentially stronger than the $\mathscr{G}$-stable convergence, because we get 5.26 (under 5.24), which is quite stronger than 5.32.

On the other hand, the $\mathscr{G}$-stable convergence under 5.2 is not very much more than the ordinary convergence in law, because $\mathscr{G}$ is essentially "very small".

A more interesting $\mathscr{G}$-stable convergence would be when $\mathscr{G} = \mathscr{F}'$ (so $\mathscr{G}$ is "as big as possible"). Of course, one cannot hope for $\mathscr{F}'$-stable convergence in general; however, a *nesting condition* on the filtrations allows to get $\mathscr{F}'$-stable convergence. More precisely, instead of 5.2 we assume:

5.37 *Hypothesis*: (i) there is a sequence of numbers (α_n), decreasing to 0, such that $\mathscr{F}^n_{\alpha_n} \subset \mathscr{F}^{n+1}_{\alpha_{n+1}}$ for all $n \in \mathbb{N}^*$.

(ii) Moreover, $\mathscr{G} = \mathscr{F}' = \bigvee_n \mathscr{F}^n_{\alpha_n}$ (hence $\mathscr{F}^n_{\infty-} \subset \bigvee_q \mathscr{F}^q_{\alpha_q}$). □

5.38 *Example.* This nesting condition 5.37 is undoubtedly very restrictive. It is however met in a very important case (encountered already in § 3f for instance): suppose that Y is a semimartingale with $Y_0 = 0$, on a stochastic basis $(\Omega', \mathscr{F}', \mathbf{F}, P')$, with $\mathscr{F}' = \mathscr{F}_{\infty-}$. Suppose that $X^n_t = a_n Y_{\gamma_n t}$ where a_n is a norming matrix and γ_n is a sequence of real numbers increasing to $+\infty$ (usually, $\gamma_n = n$); that is, we normalize and change the time on a fixed process.

Then if $\mathbf{F}^n$ is the filtration generated by the process X^n, the nesting condition 5.37 is met: take $\alpha_n = \gamma_n^{-1/2}$. □

5.39 *The discrete case.* We consider now a triangular array scheme, as in 2.22. We assume in addition that $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega', \mathscr{F}', P')$ for all n, to meet Hypothesis 5.1. Suppose also that

5.40 $$\sigma^n_t = [nt].$$

Then the nesting condition goes as follows:

5.41 $$\mathscr{F}' = \bigvee_n \mathscr{G}^n_{\infty-}, \quad \mathscr{G}^n_p \subset \mathscr{G}^{n+1}_p \quad \text{for all } p \geq 0, n \geq 1.$$

(Actually, the continuous-time filtration $\mathbf{F}^n$ associated with the process X^n given by 2.24 is $\mathscr{F}^n_t = \mathscr{G}^n_{[nt]}$, so under 5.41 we have 5.37 with $\alpha_n = 1/\sqrt{n}$). □

Below we do not present stable convergence theorems in full generality. We will content ourselves with one result, which is well-suited for functional convergence, or finite-dimensional convergence along a *dense* subset $D \subset \mathbb{R}_+$. Other results for the discrete case and convergence at one point (i.e. $D = \{1\}$) are given in the book [84]: the method is different from here, and it does not seem that our method would work easily in that case.

5.42 **Theorem.** *Assume 5.1, 5.37 and 5.3. Assume also that X is quasi-left-continuous, and let D be a dense subset of $\mathbb{R}_+$.*

a) *If $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ hold, then for all $t_j \in D$ the sequence $(X^n_{t_1}, \dots, X^n_{t_p})$ converges stably* (i.e.: $\mathscr{F}'$-stably) *to $(X_{t_1}, \dots, X_{t_p})$.*

b) *If $[\text{Sup-}\beta_5] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ hold, then (X^n) converges stably to X.*

We begin with a lemma, which is an easy complement to 1.9 and 5.16:

5.43 **Lemma.** *Assume 5.1, 5.37 and 5.3. Let $D \subset \mathbb{R}_+$ and assume that 5.13 and 5.17 hold for all $t \in D$. Finally, assume that*

5.44 $$[\exp iu \cdot X^n_{\alpha_n \wedge S^n}]/G^n(u)_{\alpha_n \wedge S^n} \xrightarrow{P} 1$$

for all $u \in \mathbb{R}^d$ *and all sequences of stopping times* S^n *such that* $|G^n(u)_{S^n}|$ *is bounded away from* 0 (uniformly in n; each S^n is an $\mathbf{F}^n$-stopping time, and (α_n) is the sequence coming in 5.37). *Then for all* $t_j \in D$, $(X^n_{t_1}, \dots, X^n_{t_p})$ *converges* $\mathscr{F}'$*-stably to* $(X_{t_1}, \dots, X_{t_p})$.

Proof. We need to prove 5.8 for all $Y \in b\mathscr{G}$ (recall that $\mathscr{G} = \mathscr{F}'$). We can reproduce the proof of 5.16 word for word. In 5.16 we used 5.2 at only one point: to derive 5.22; here, the last term in 5.22 should be $E(Y1_{\{A \geq a\}}\zeta^n \beta^n \gamma)$, and it is not equal to $E(Y1_{\{A \geq a\}}\zeta^n \gamma)$ in general because $Z = Y1_{\{A \geq a\}}\zeta$ is $\mathscr{G}$-measurable, but not $\mathscr{F}^n_{t_{p-1}}$-measurable. Therefore we must add a fourth term to the right-hand side of 5.23, namely

$$w^n = |E(Z\zeta^n(\beta^n - 1))|$$

and if we prove that $w^n \to 0$ as $n \uparrow \infty$, the proof will be finished.

If (α_n) is like in 5.37, and using $|\zeta^n| = 1$, we obtain:

5.45 $$w^n \leq |E[E(Z|\mathscr{F}^n_{\alpha_n})\zeta^n(\beta^n - 1)]| + E[|\beta^n - 1||Z - E(Z|\mathscr{F}^n_{\alpha_n})|].$$

We have $|\beta^n| \leq 2/a$ and $|Z| \leq K/a$, and the martingale convergence theorem plus 5.37 yield that $E(Z|\mathscr{F}^n_{\alpha_n}) \to Z$ a.s. Therefore the last term in 5.45 tends to 0.

Consider the first term on the right-hand side of 5.45. If $p \geq 2$ we have $t_{p-1} > 0$, and for all n large enough we have $\alpha_n < t_{p-1}$. Because of 5.21 we deduce that $E[E(Z|\mathscr{F}^n_n)\beta^n(\zeta^n - 1)] = 0$ and the result is proved. If $p = 1$, we have $t_{p-1} = 0$ and this argument breaks down. However we then have $\zeta^n = 1$ and because M^n is a martingale and $\beta^n = M^n_{t_1} = (\exp iu \cdot X^n_{S^n \wedge t_1})/G^n(u)_{S^n \wedge t_1}$ we get

$$E[E(Z|\mathscr{F}^n_{\alpha_n})(\beta^n - 1)] = E[E(Z|\mathscr{F}^n_{\alpha_n})\{[\exp iu \cdot X^n_{S^n \wedge t_1 \wedge \alpha_n}]/G^n(u)_{S^n \wedge t_1 \wedge \alpha_n} - 1\}].$$

For all n large enough we have $\alpha_n < t_1$, while $|Z| \leq K/a$; thus

$$|E[E(Z|\mathscr{F}^n_{\alpha_n})(\beta^n - 1)]| \leq \frac{K}{a} E(|[\exp iu \cdot X^n_{S^n \wedge \alpha_n}]/G^n(u)_{S^n \wedge \alpha_n} - 1|),$$

which tends to 0 by 5.44 (the integrand in the right-hand side above being bounded by $1 + 2/a$). Therefore $w^n \to 0$ in all cases, and the proof is complete. □

Proof of 5.42. a) In the proof of 5.7 we have seen that 5.17 holds, so in view of the previous lemma it remains to prove 5.44. We take the truncation function h to be uniformly continuous, and we set $X'^n = X^n - B^n$, $X' = X - B$, and $G'^n(u)$ and $g'(u)$ be the processes associated to X'^n and X' by 1.8. Using II.2.47, we get $G^n(u) = G'^n(u) \exp iu \cdot B^n$, so it suffices to prove:

5.46 $$(\exp iu \cdot X'^n_{S^n \wedge \alpha_n})/G'^n(u)_{S^n \wedge \alpha_n} \xrightarrow{P} 1.$$

By 2.19 we have $X'^n \xrightarrow{\mathscr{L}} X'$. Since $S^n \wedge \alpha_n \to 0$, we deduce that $X'^n_{S^n \wedge \alpha_n} \xrightarrow{\mathscr{L}} X'_0 = 0$, hence $X'^n_{S^n \wedge \alpha_n} \xrightarrow{P} 0$. Consider also the proof of 2.19: we obtain 2.20, except that $\tilde{C}^{X'}_t$ and $g * \nu^{X'}_t$ depend on ω. Then, from VII.3.4, if $\omega \in A$ the PII $Z^{n'',\omega}$ with characteristics $(B^{X'^{n''}}(\omega), C^{X'^{n''}}(\omega), \nu^{X'^{n''}}(\omega))$ converges in law to X', with characteristics $(B^{X'}(\omega), C^{X'}(\omega), \nu^{X'}(\omega))$; hence, exactly as before, $Z^{n'',\omega}_{S^{n''}(\omega)\wedge\alpha_{n''}}$ converges in law to 0, and we deduce that $G'^{n''}(u)(\omega)_{S^{n''}(\omega)\wedge\alpha_{n''}}$ converges to 1. By the usual argument we deduce that $G'^{n}(u)_{S^n\wedge\alpha_n} \xrightarrow{P} 1$, and this gives us 5.46.

b) By 2.1 we have $X^n \xrightarrow{\mathscr{L}} X$, and thus the sequence (X^n) is tight. Therefore it suffices to reproduce the proof of 5.7(b), in which property 5.2 plays no rôle. □

5.47 **Remark.** Theorem 5.42 remains valid when X is not quasi-left-continuous, provided we replace [$\delta_{5,1}$-D] and [Sup-β_5] by [Sk-$\delta_{5,1}$] and [Sk-β_5]. We leave all details to the reader, just pointing out that the first step is to prove an extension of 2.19. □

3. Now we give a very particular application of stable convergence to random norming. We start with a 1-dimensional local martingale Y on a stochastic basis $(\Omega', \mathscr{F}', \mathbf{F}', P')$, satisfying $Y_0 = 0$ and

5.48 $$|\Delta Y_t| \le K \quad \text{identically}$$

for some constant K (this is just for simplicity). We also suppose that

5.49 $$\frac{1}{n}[Y, Y]_{nt} \xrightarrow{P} t\eta^2$$

where η is a nonnegative variable on $(\Omega', \mathscr{F}', P')$.

5.50 **Theorem.** *Under 5.48 and 5.49, the sequence of processes $X^n_t = Y_{nt}/\sqrt{n}$ converges $\mathscr{F}'$-stably to $X = \eta W$, where W is a standard Wiener process independent of $\mathscr{F}'$.*

The precise construction of X is as follows: let P_0 be the unique measure on $\mathbb{D}(\mathbb{R}^d)$ for which the canonical process W is a standard Wiener process. Then set

$$\Omega = \Omega' \times \mathbb{D}(\mathbb{R}^d), \quad \mathscr{F} = \mathscr{F}' \otimes \mathbb{D}(\mathbb{R}^d), \quad P = P' \otimes P_0$$

and extend X^n and η and W in the usual fashion to Ω. Then $X_t(\omega', \alpha) = \eta(\omega')W_t(\alpha)$. We can also construct X according to 5.3, with $\mathscr{G} = \mathscr{F}'$, provided $Q(\omega', \cdot)$ is the unique measure for which the canonical process X is a Gaussian martingale with $\langle X, X\rangle_t = \eta^2(\omega)t$. These two ways are clearly equivalent.

Proof. We use the second method above for constructing X: we have 5.1 with $\mathscr{F}^n_t = \mathscr{F}'_{nt}$, and 5.3 with $\mathscr{G} = \mathscr{F}'$. In view of 5.38, we have 5.37. Note that $B = 0$, $\nu = 0$ and $C_t(\omega') = \eta^2(\omega')t$. Because of 5.48 we clearly have 3.14; since $[X^n, X^n]_t = [Y, Y]_{nt}/n$, 5.49 yields [$\hat{\gamma}'_5$-$\mathbb{R}_+$].

Now, we have shown in the proof of 3.12 that if $B = 0$ and $\nu = 0$, then 3.14 and [$\hat{\gamma}'_5$-$\mathbb{R}_+$] imply [Sup-β_5], [γ_5-$\mathbb{R}_+$] and [$\delta_{5,1}$-$\mathbb{R}_+$], and this is irrespective to whether C is random or not. Therefore the result follows from 5.42. □

The following corollary gives two version of "random norming":

5.51 **Corollary.** *Assume 5.48, 5.49, and $\eta > 0$ P'-a.s. Then the two sequences of processes $X_t'^n = Y_{nt}/\eta\sqrt{n}$ and $X_t''^n = Y_{nt}/[Y,Y]_n^{1/2}$ converge in law to a standard Wiener process.*

Proof. That $X'^n \xrightarrow{\mathscr{L}} W$ follows from 5.33(ii) and from the fact that $X'^n = f(\eta, X^n)$ and $W = f(\eta, X)$, where f is the continuous function $f(a, \alpha(\cdot)) = \alpha(\cdot)/a$. That $X''^n \xrightarrow{\mathscr{L}} W$ follows from the previous property and from the convergence $[Y,Y]_n^{1/2}/\eta\sqrt{n} \xrightarrow{P} 1$. □

5.52 **Remark.** There exist one-dimensional convergence results of this type (see [3]). There also exists a more sophisticated version where 5.49 is replaced by $\frac{1}{n}[Y,Y]_{nt} \xrightarrow{P} C_t$, C being a continuous process. □

4. We state another very simple application of stable convergence, to *absolutely continuous changes of measures.* Consider first the situation of 5.28, and let $\tilde{P}'$ be another probability measure on $(\Omega', \mathscr{F}')$, which is absolutely continuous with respect to P'; call $Y = \frac{d\tilde{P}'}{dP'}$ the Radon-Nikodym derivative.

Then, if (Z^n) converges $\mathscr{F}'$-stably to Z under P (see 5.32), we obviously have the following: if $\tilde{P}(d\omega', dx) = \tilde{P}'(d\omega')Q(\omega', dx)$, then (Z^n) converges $\mathscr{F}'$-stably to Z under $\tilde{P}$ as well (this is obvious when Y is bounded; if not, one firstly approximate Y by $Y \wedge p$ and then let $p \uparrow \infty$). Henceforth we deduce from 5.42 the following:

5.53 **Theorem.** *Assume 5.1, 5.37 and 5.3; assume also that X is quasi-left-continuous, and let D be a dense subset of $\mathbb{R}_+$; finally, let $\tilde{P}'$ be another probability measure on (Ω', F') with $\tilde{P}' \ll P'$.*

a) *Under* $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ *we have $X^n \xrightarrow{\mathscr{L}(D)} X$ under $\tilde{P}$ as well as under P.*

b) *Under* $[\text{Sup-}\beta_5] + [\gamma_5\text{-}D] + [\delta_{5,1}\text{-}D]$ *we have $X^n \xrightarrow{\mathscr{L}} X$ under $\tilde{P}$ as well as under P.*

Note that under $\tilde{P}$, X still is an $\mathscr{F}'$-PII with the same characteristics (B, C, ν) than under P, but not with the same law. It is also worth noticing that, above, all conditions concern the measure P', while the conclusion is about $\tilde{P}'$: this is very useful when the structure of the X^n's is simple under P', so conditions $[\beta_5\text{-}D]$, $[\gamma_5\text{-}D]$, $[\delta_{5,1}\text{-}D]$ are (relatively) simple to check; then for any measure $\tilde{P}' \ll P'$ we also have convergence in law, and we can identify the limit, no matter how difficult to compute the new characteristics of X^n might be.

§ 5d. Mixing Convergence

5.54 **Definition.** In the situation of 5.28, we say that (Z^n) *converges $\mathscr{G}$-mixing* (or, *mixing*, if $\mathscr{G} = \mathscr{F}'$) if it converges $\mathscr{G}$-stably, and if the limiting measure μ is of the form $\mu(d\omega', dx) = P'(d\omega')\theta(dx)$, where θ is a probability measure on $(E, \mathscr{E})$. $\square$

With the notation of 5.31 and 5.32, the variable Z is then independent from $\mathscr{G}$ under P.

All the previous $\mathscr{G}$-stable convergence results (5.36, 5.42, 5.50) are in fact $\mathscr{G}$-mixing, provided we have the following:

5.55 In 5.3, the measure $Q(\omega', dx) = Q(dx)$ does not depend upon ω', or equivalently the characteristics (B, C, ν) are deterministic, or equivalently the process X is a PII under the measure P defined by 5.5. $\square$

For example, in 5.51, the convergence of X'^n or of X''^n is mixing. Of course, this notion of mixing should not be confused with mixing sequences which appeared in § 3g.

§ 5e. Application to Stationary Processes

In this short subsection we come back to the situation of § 3g: we have a probability space $(\Omega', \mathscr{F}', P')$ with a measure-preserving flow $(\theta_t)_{t \in \mathbb{R}}$ and a measurable process $Y = (Y_t)_{t \in \mathbb{R}}$ satisfying $Y_{t+s} = Y_t \circ \theta_s$ for s, $t \in \mathbb{R}$. To simplify the matter, we also suppose that $\mathscr{F}' = \sigma(Y_s \colon s \in \mathbb{R})$. We denote again the invariant σ-field (see 3.78) by $\mathscr{I}$, and

$$X_t^n = \frac{1}{\sqrt{n}} \int_0^{nt} Y_s \, ds, \quad t \geq 0.$$

Let also $(\Omega'', \mathscr{F}'', P'')$ be another probability space, supporting a standard Wiener process W. We set

$$\Omega = \Omega' \times \Omega'', \quad \mathscr{F} = \mathscr{F}' \otimes \mathscr{F}'', \quad P = P' \otimes P''$$

and we extend (with the same notation) Y, X^n, $\mathscr{I}$ from Ω' to Ω and W from Ω'' to Ω.

5.56 **Theorem.** *Assume that*

5.57 $$\|Y_0\|_2 < \infty, \quad \int_0^\infty \|E'(Y_t | \mathscr{F}'_0)\|_2 \, dt < \infty$$

(i.e. 3.80 with $p = 2$). *Then*

5.58 $$c = 2 \int_0^\infty E'(Y_0 Y_t | \mathscr{I}) \, dt$$

is a nonnegative integrable $\mathscr{I}$-measurable random variable, and X^n converges $\mathscr{F}'$-stably to the process $\sqrt{c}\, W$ (and in particular $X^n \xrightarrow{\mathscr{L}} \sqrt{c}\, W$).

Proof. It suffices to reproduce the proof of 3.79, with the following changes: everywhere, the expectation $E'(\cdot)$ with respect to P' is replaced by the conditional expectation $E'(\cdot|\mathscr{I})$. Then in Lemma 3.86 we obtain $E'(M_1^2|\mathscr{I}) = E(M_1^2|\mathscr{I}) = c$ (note that for any variable defined on Ω' we have $E'(U|\mathscr{I}) = E(U|\mathscr{I})$), and in particular $c \geq 0$ and $E(c) < \infty$. In the conclusion of Lemma 3.91, we naturally obtain that

$$\frac{1}{n} V_{nt} \xrightarrow{L^1} tE(V_1|\mathscr{I}).$$

Then 3.93 and 3.94 still hold. Now, in 3.22 the various equivalences do not use the fact that C is deterministic (except of course for the determination of the law of the limit X), so 3.93 and 3.94 indeed imply the conditions [Sup-β_5], [γ_5-$\mathbb{R}$], [$\delta_{5,1}$-$\mathbb{R}$] for the sequence (M^n), with the limits $B_t = 0$, $C_t = ct$, $\nu = 0$, which are the characteristics of the process $\sqrt{c}W$. Now, by virtue of 5.38, the nesting hypothesis 5.37 is met here, as well of course as 5.1 and 5.3 (because W is independent of $\mathscr{F}'$ by construction). Then 5.42b yields that M^n converges $\mathscr{F}'$-stably to $\sqrt{c}W$.

Since $\sup_{s \leq t}|X_s^n - M_s^n| \xrightarrow{P} 0$ as $n \uparrow \infty$ for all $t < \infty$ by the end of the proof of 3.79, we obtain the result. □

5.59 **Remarks.** 1) If we assume 3.80 with $p \in (2, \infty]$, a similar application of 5.42 would yield that $(X_{t_1}^n, \ldots, X_{t_r}^n)$ converges $\mathscr{F}'$-stably to $(\sqrt{c}W_{t_1}, \ldots, \sqrt{c}W_{t_r})$ for all $0 \leq t_1 < \cdots < t_r$.

2) There is a similar extension of the discrete-time theorem 3.97. We let it to the reader.

3) If we add the hypothesis 3.78, c is deterministic and thus X^n converges *$\mathscr{F}'$-mixing* (in the sense of 5.54) to $\sqrt{c}W$. This is an improvement on Theorem 3.79, which we get for free because we consider convergence of one process, normalized in the sense of 5.38.

4) We are also here in a position to apply Theorem 5.25, because 5.24 is obviously met. So, if we denote by $\hat{Q}^n$ the conditional distribution of X^n relatively to the invariant σ-field $\mathscr{I}$, and if we denote by Q_a the distribution of the Wiener process W^a having $\langle W^a, W^a\rangle_t = at$, then we have

$$\hat{Q}^n(\cdot) \xrightarrow{P} Q_{c(\cdot)}.$$ □

Chapter IX. Convergence to a Semimartingale

Here comes the third—and last—step in our exposition of limit theorems. Not only are the pre-limiting processes X^n arbitrary semimartingales, but the limit process X also is a semimartingale; not quite an arbitrary one, though: since the method is based here on convergence of martingales and on the relations between X and its characteristics, we need these characteristics to indeed characterize the distribution $\mathscr{L}(X)$ of X. So, in most of the chapter, we will assume that $\mathscr{L}(X)$ is the unique solution to the martingale problem associated with the characteristics of X, as introduced in Chapter III.

Section 1 presents the basic technique used in this chapter: essentially that the limit in law of a sequence of martingales is also a martingale.

Then, as usual, our limit theorems will be proved via two steps: tightness of the sequence (X^n), and identification of the limit. Section 2 is devoted to the second problem, while tightness and the proper limit theorems are given in Section 3.

We provide some examples and complements in Section 4: convergence of diffusion processes (possibly with jumps), of pure step Markov processes; and also the famous convergence of normalized empirical distributions to the Brownian bridge. Moreover, exactly as in Section 3 of the previous chapter, we can obtain necessary and sufficient conditions for $X^n \xrightarrow{\mathscr{L}} X$, when X is a continuous semimartingale.

Finally, we present a different problem in Section 5: assuming that $X^n \xrightarrow{\mathscr{L}} X$, we provide conditions for the convergence $H^n \cdot X^n \xrightarrow{\mathscr{L}} H \cdot X$ of stochastic integrals.

1. Limits of Martingales

This section presents an essential tool, which can roughly be explicited as such: the limit in law of a sequence of martingales is a martingale. This result (valid under some assumptions ...) is interesting in itself; however, most of this section is technical.

The setting is as follows: for every $n \in \mathbb{N}^*$ we consider a stochastic basis $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$; E^n denotes the expectation with respect to P^n. All sets, variables, processes, ... with the superscript "n" are defined on $\mathscr{B}^n$, usually without mentioning.

§ 1a. The Bounded Case

Our first proposition is the prototype of the results we are looking for. Although it is not general enough for our purposes, its simplicity makes it worthwhile to state and prove as a first step.

1.1 **Proposition.** *Assume that (M^n) is a sequence of martingales converging in law to a limit process M, and that $|M^n| \le b$ identically for some constant b. Then M is a martingale with respect to the filtration that it generates.*

Proof. a) The limit process is defined on a probability space $(\Omega, \mathscr{F}, P)$ and it generates a filtration $\mathbf{F}^M$. Consider also the space $\mathbb{D} = \mathbb{D}(\mathbb{R})$ with the σ-fields $\mathscr{D} = \mathscr{D}(\mathbb{R})$ and $\mathscr{D}_t = \mathscr{D}_t(\mathbb{R})$ (see VI.1.1). Let $\tilde{P} = \mathscr{L}(M)$ be the law of M, i.e. the image $\tilde{P} = P \circ M^{-1}$ of P by the map $M: \Omega \to \mathbb{D}$. Set $D = \{t: P(\Delta M_t \neq 0) = 0\}$.

b) The function $\alpha \rightsquigarrow \sup_{t \le s} |\alpha(t)|$ is $\tilde{P}$-a.s. continuous on $\mathbb{D}$ if $s \in D$ (see VI.2.4). Hence $\sup_{t \le s} |M_t^n| \xrightarrow{\mathscr{L}} \sup_{t \le s} |M_t|$ and so $\sup_t |M_t| \le b$ P-a.s.

c) Let $t_1 \le t_2$ with $t_1, t_2 \in D$. Let f be a continuous bounded $\mathscr{D}_{t_1-}$-measurable function on $\mathbb{D}$. Then by VI.2.1 the function $\alpha \rightsquigarrow \alpha(t_i) f(\alpha)$ is $\tilde{P}$-a.s. continuous on $\mathbb{D}$, and so

1.2
$$\begin{aligned} E^n[f(M^n) M_{t_i}^n] &= E^n[f(M^n)(M_{t_i}^n \wedge b \vee -b)] \\ &\to E[f(M)(M_{t_i} \wedge b \vee -b)] = E[f(M) M_{t_i}]. \end{aligned}$$

On the other hand, $f(M^n)$ is $\mathscr{F}_{t_1-}^n$-measurable (since f is $\mathscr{D}_{t_1-}$-measurable and M^n is adapted), and M^n is a martingale on $\mathscr{B}^n$, thus $E^n[f(M)(M_{t_2}^n - M_{t_1}^n)] = 0$ and we deduce from 1.2 that

1.3
$$E[f(M)(M_{t_2} - M_{t_1})] = 0.$$

d) A monotone class argument shows that 1.3 holds for all bounded f, which are measurable with respect to the σ-field generated by all continuous bounded $\mathscr{D}_{t_1-}$-measurable functions: in virtue of VI.1.14c, we thus have 1.3 for all bounded $\mathscr{D}_{t_1-}$-measurable functions.

Finally, let $s < t$. There are two sequences $s_n \downarrow\downarrow s$, $t_n \downarrow\downarrow t$ such that $s_n, t_n \in D$. So 1.3 holds for each pair (s_n, t_n) and every bounded $\mathscr{D}_s$-measurable f (because $\mathscr{D}_s \subset \mathscr{D}_{s_n-}$). Letting $n \uparrow \infty$ and using $|M| \le b$, we obtain $E[f(M)(M_t - M_s)] = 0$. But by definition of $\mathbf{F}^M$ we have $\mathscr{F}_s^M = M^{-1}(\mathscr{D}_s)$, so we deduce that $E(M_t - M_s | \mathscr{F}_s) = 0$, and we are finished. □

The above does not cover the local martingales which "characterize" a semimartingale, in the sense of II.2.21, for two reasons:

1) these local martingales are not bounded,

2) they can be (local) martingales with respect to a filtration which is strictly bigger than the filtration which they generate.

Problem (1) will be taken care of in § 1b. As for (2), there are two different generalizations of Proposition 1.1:

1.4 **Proposition.** *Let Y^n be an m-dimensional càdlàg process and let M^n be a martingale, both on $\mathscr{B}^n$. Furthermore, let M be a càdlàg adapted process defined on the canonical space* $(\mathbb{D}(\mathbb{R}^m), \mathscr{D}(\mathbb{R}^m), \mathbf{D}(\mathbb{R}^m))$ (see VI.1.1) *and let D be a dense subset of $\mathbb{R}_+$. Assume that*:

(i) $|M^n| \leq b$ *identically for some constant* b;

(ii) $Y^n \xrightarrow{\mathscr{L}} Y$ *for some limit process* Y, *with law* $\tilde{P} = \mathscr{L}(Y)$;

(iii) *for all* $t \in D$, $\alpha \rightsquigarrow M_t(\alpha)$ *is* $\tilde{P}$-*a.s. continuous on* $\mathbb{D}(\mathbb{R}^m)$;

(iv) $M_t^n - M_t \circ Y^n \xrightarrow{P} 0$ *for all* $t \in D$ (recall that Y^n may be considered as a map from Ω^n into $\mathbb{D}(\mathbb{R}^m)$; recall also the notation VIII.1.5 for $\xrightarrow{P}$);

then the process $M \circ Y$ is a martingale with respect to the filtration generated by Y.

In general, Y is defined on a space $(\Omega, \mathscr{F}, P)$, and then $\tilde{P} = P \circ Y^{-1}$; but quite often Y is indeed the canonical process on $\Omega = \mathbb{D}(\mathbb{R}^m)$, in which case $\tilde{P} = P$ and $M \circ Y = M$.

Proof. The proof is exactly similar to that of 1.1, with the following changes:

a) We consider $\mathbb{D} = \mathbb{D}(\mathbb{R}^m)$, $\mathscr{D} = \mathscr{D}(\mathbb{R}^m)$, $\mathscr{D}_t = \mathscr{D}_t(\mathbb{R}^m)$, and $\tilde{P} = P \circ Y^{-1}$.

b) (ii) and (iii) yield $M_t \circ Y^n \xrightarrow{\mathscr{L}} M_t \circ Y$ for $t \in D$, so (iv) yields $M_t^n \xrightarrow{\mathscr{L}} M_t \circ Y$ and thus $|M_t \circ Y| \leq b$ a.s. because of (i). Since D is dense in $\mathbb{R}_+$ and M is càdlàg, we deduce that $|M \circ Y| \leq b$ identically outside a $\tilde{P}$-null set.

c) Let $t_1 \leq t_2$ with $t_i \in D$; let f be a continuous bounded $\mathscr{D}_{t_1 -}$-measurable function on $\mathbb{D}$. Then $\alpha \rightsquigarrow f(\alpha)[M_{t_i}(\alpha) \wedge b \vee -b]$ is $\tilde{P}$-a.s. continuous (use (ii)), and so

1.5 $$E^n[f(Y^n)((M_{t_i} \circ Y^n) \wedge b \vee -b)] \to E[f(Y)((M_{t_i} \circ Y) \wedge b \vee -b)].$$

On the other hand, (iii) yields

$$(M_{t_i}^n \wedge b \vee -b) - ((M_{t_i} \circ Y^n) \wedge b \vee -b) \xrightarrow{P} 0,$$

and so

1.6 $$E^n[f(Y^n)\{(M_{t_i}^n \wedge b \vee -b) - ((M_{t_i} \circ Y^n) \wedge b \vee -b)\}] \to 0.$$

Putting 1.5 and 1.6 together yields

1.7 $$E^n[f(Y^n)(M_{t_i}^n \wedge b \vee -b)] \to E[f(Y)((M_{t_i} \circ Y) \wedge b \vee -b)].$$

Now we apply (i) and part (b) of the proof, so $M_{t_i}^n \wedge b \vee -b = M_{t_i}^n$, and $(M_{t_i} \circ Y) \wedge b \vee -b = M_{t_i} \circ Y$ P-a.s., and thus

1.8 $$E^n[f(Y^n)M_{t_i}^n] \to E[f(Y)M_{t_i} \circ Y].$$

Thus, since M^n is a martingale on $\mathscr{B}^n$ and since $f(Y^n)$ is $\mathscr{F}^n_{t_1-}$-measurable, we deduce that $E^n[f(Y^n)(M^n_{t_2} - M^n_{t_1})] = 0$, and thus 1.8 implies:

1.9 $$E[f(Y)(M_{t_2} \circ Y - M_{t_1} \circ Y)] = 0.$$

d) Finally, we can reproduce the part (d) of the proof of 1.1 (just replacing M by $M \circ Y$) and we obtain that $E[f(Y)(M_t \circ Y - M_s \circ Y)] = 0$ for all $s \leq t$, f bounded $\mathscr{D}_s$-measurable, hence the claim. □

1.10 **Proposition.** *Let Y^n be an m-dimensional càdlàg process and let M^n be a martingale, both on $\mathscr{B}^n$. Assume that $|M^n| \leq b$ identically for some constant b. Assume also that the $(m+1)$-dimensional processes (Y^n, M^n) converge in law to a limiting process (Y, M). Then M is a martingale with respect to the filtration generated by (Y, M).*

Proof. Again, one could reproduce the proof of 1.1, with $\mathbb{D} = \mathbb{D}(\mathbb{R}^{m+1})$. One can also apply 1.4 to Y'^n, M'^n, M', D', defined by

$$\begin{cases} Y'^n = (Y^n, M^n), \quad M'^n = M^n, \quad D' = \{t\colon P(\Delta M_t = 0, \Delta Y_t = 0) = 1\} \\ Y' = (Y, M), \quad M' = \text{the } (m+1)\text{th coordinate of the canonical process} \\ \qquad\qquad\qquad\qquad \text{on } \mathbb{D}(\mathbb{R}^{m+1}). \end{cases}$$

Then (i) and (ii) hold by hypothesis, (iii) holds because of the definition of D', and (iv) is trivial because $M'^n = M^n = M' \circ Y'^n$. Finally, $M' \circ Y' = M$, hence the claim. □

§ 1b. The Unbounded Case

Now we want to relax the assumption of uniform boundedness on the M^n's in Propositions 1.4 and 1.10. For the first one, we replace boundedness by uniform integrability, while for 1.10 we have a genuine "localization" in time of the result.

1. Before the extension of 1.4, we state a lemma. Let $(Z_i)_{i \in I}$ be a family of random variables; even if they are defined on different probability spaces, we will denote by $E(\cdot)$ the expectation.

1.11 **Lemma.** *The family $(Z_i)_{i \in I}$ is uniformly integrable if and only if $\sup_{i \in I} E(|Z_i| - |Z_i| \wedge b) \to 0$ as $b \uparrow \infty$.*

Proof. The necessity follows from the inequality $E(|Z_i| - |Z_i| \wedge b) \leq E(|Z_i| \times 1_{\{|Z_i| > b\}})$. Conversely, assume that $\rho(b) := \sup_{i \in I} E(|Z_i| - |Z_i| \wedge b)$ goes to 0 as $b \uparrow \infty$. Firstly,

$$\alpha := \sup_{i \in I} E(|Z_i|) \leq b + \rho(b)$$

for all b, and thus $\alpha < \infty$. Secondly, we have

$$E(|Z_i| 1_{\{|Z_i|>b\}}) \leq E((|Z_i| \wedge a) 1_{\{|Z_i|>b\}}) + E(|Z_i| - |Z_i| \wedge a)$$
$$\leq aP(|Z_i| > b) + \rho(a) \leq \frac{a}{b}\alpha + \rho(a).$$

If $\varepsilon > 0$ is given, we choose first a so that $\rho(a) \leq \varepsilon/2$, then b so that $\frac{a}{b}\alpha \leq \varepsilon/2$, so the above is smaller than ε for all $i \in I$: hence the family $(Z_i)_{i \in I}$ is uniformly integrable. □

1.12 **Proposition.** *The result of Proposition 1.4 remains valid if we replace* (i) *by:*
(i′) *the family* $(M_t^n)_{n \in \mathbb{N}^*, t \in \mathbb{R}_+}$ *of random variables is uniformly integrable.*

Proof. The proof goes as for 1.4, except that part (b) of this proof becomes false here. However, the beginning of part (c) remains valid, and we have 1.7 for every $b \in \mathbb{R}_+$. Moreover, the same proof also yields

1.13 $$E^n(|M_t^n| \wedge b) \to E(|M_t \circ Y| \wedge b)$$

for all $t \in D$, $b \in \mathbb{R}_+$. Set

1.14 $$\rho(b) = \sup_{t \geq 0, n \in \mathbb{N}^*} E^n(|M_t^n| - |M_t^n| \wedge b),$$

which, by virtue of 1.11 and (i′) goes to 0 as $b \uparrow \infty$. Applying 1.13 for b and for $b' > b$, and the definition of $\rho(b)$, we obtain for $t \in D$:

$$E(|M_t \circ Y| \wedge b' - |M_t \circ Y| \wedge b) \leq \rho(b).$$

Letting $b' \uparrow \infty$ gives

1.15 $$E(|M_t \circ Y| - |M_t \circ Y| \wedge b) \leq \rho(b)$$

for all $t \in D$. Then, by virtue of Lemma 1.11 again,

1.16 the family $(M_t \circ Y)_{t \in D}$ is uniformly integrable.

Now, it is easy to deduce from 1.7, 1.14, 1.15 and $\lim_{b \uparrow \infty} \rho(b) = 0$ that 1.8 holds. Hence the same argument than in the proof of 1.4 shows that 1.9 holds for all bounded $\mathscr{D}_{t_1 -}$-measurable continuous function on $\mathbb{D} = \mathbb{D}(\mathbb{R}^m)$.

Finally, let $s < t$ and $s_n \downarrow\downarrow s$, $t_n \downarrow\downarrow t$ with s_n, $t_n \in D$; let also f be a bounded $\mathscr{D}_s$-measurable function. Then 1.9 yields

$$E[f(Y)(M_{t_n} \circ Y - M_{s_n} \circ Y)] = 0.$$

Because of the right-continuity of M and of 1.16, $M_{t_n} \circ Y - M_{s_n} \circ Y$ converges to $M_t \circ Y - M_s \circ Y$ in L^1, and so

$$E[f(Y)(M_t \circ Y - M_s \circ Y)] = 0.$$

Since any $\mathscr{F}_s^Y$-measurable variable on Ω is of the form $f(Y)$ for some $\mathscr{D}_s$-measurable function f, the claim follows. □

2. Now we proceed to the localization in Proposition 1.10.

1.17 **Proposition.** *Let Y^n be an m-dimensional càdlàg process and let M^n be a local martingale, both on $\mathscr{B}^n$, and assume that $|\Delta M^n| \leq b$ identically for some constant b. Assume also that the $(m+1)$-dimensional processes (Y^n, M^n) converge in law to a limit process (Y, M). Then M is a local martingale with respect to the filtration generated by (Y, M).*

Proof. Let $\mathbb{D} = \mathbb{D}(\mathbb{R}^{m+1})$ with the filtration $\mathbf{D} = \mathbf{D}(\mathbb{R}^{m+1})$. For simplicity we also set $\tilde{Y}^n = (Y^n, M^n)$ and $\tilde{Y} = (Y, M)$. The process $\tilde{Y}$ is defined on the space $(\Omega, \mathscr{F}, P)$, it generates the filtration $\mathbf{F}^{\tilde{Y}}$, and we call $\tilde{P} = P \circ \tilde{Y}^{-1}$ its law.

According to VI.2.10, we set for $\alpha \in \mathbb{D}$:

$$S_a(\alpha) = \inf(t\colon |\alpha(t)| \geq a \text{ or } |\alpha(t-)| \geq a)$$

$$V(\alpha) = \{a > 0\colon S_a(\alpha) < S_{a+}(\alpha)\}$$

$$V'(\alpha) = \{a > 0\colon \Delta\alpha(S_a(\alpha)) \neq 0 \text{ and } |\alpha(S_a(\alpha)-)| = a\},$$

so by VI.2.12, $\alpha \rightsquigarrow \alpha^{S_a}$ (the "stopped" function at time $S_a(\alpha)$) is continuous for the Skorokhod topology at each point α such that $a \notin V(\alpha) \cup V'(\alpha)$.

Now, $\tilde{V} = \{a > 0\colon \tilde{P}(\alpha\colon a \in V(\alpha)) > 0\}$ is at most countable, as the set of fixed times of discontinuity for the increasing càg process $(S_a)_{a \geq 0}$ (see VI.3.12). Similarly, $\tilde{V}' = \{a > 0\colon \tilde{P}(\alpha\colon a \in V'(\alpha)) > 0\}$ is at most countable: the proof is the same as for $U(X)$ in VI.3.12; more precisely, with notation VI.2.6 we have

$$\tilde{V}' = \bigcup_{n, p \in \mathbb{N}^*} \left\{a\colon \tilde{P}\left(\alpha\colon t^n\left(\alpha, \frac{1}{p}\right) < \infty,\quad \left|\alpha\left(t^n\left(\alpha, \frac{1}{p}\right)-\right)\right| = a\right) > 0\right\}.$$

Let then $a > 0$ be outside $\tilde{V} \cup \tilde{V}'$. We deduce from the above that $\tilde{Y}^n \xrightarrow{\mathscr{L}} \tilde{Y}$ implies $\tilde{Y}^n(a) \xrightarrow{\mathscr{L}} \tilde{Y}(a)$, where $\tilde{Y}^n_t(a)(\omega) = \tilde{Y}^n_{t \wedge S_a(\tilde{Y}^n(\omega))}(\omega)$ and $\tilde{Y}(a)_t(\omega) = \tilde{Y}_{t \wedge S_a(\tilde{Y}(\omega))}(\omega)$. Moreover $|\Delta M^n| \leq b$, so by construction of S_a we have $|M^n(a)| \leq a + b$, where $M^n_t(a) = M^n_{t \wedge S_a(\tilde{Y}^n)}$. Then the same argument than in the proof of 1.4 shows that $|M(a)| \leq a + b$ P-a.s., where $M(a)_t = M_{t \wedge S_a(\tilde{Y})}$. Then, exactly as in the proof of 1.4, we obtain that for all t with $P(\Delta \tilde{Y}_t \neq 0) = 0$ and all continuous bounded $\mathscr{D}_{t-}$-measurable function f on $\mathbb{D}$,

$$\begin{aligned} E^n[f(\tilde{Y}^n(a))M^n_t(a)] &= E^n[f(\tilde{Y}^n(a))(M^n_t(a) \wedge (b + a) \vee -(b + a))] \\ &\to E[f(\tilde{Y}(a))(M_t(a) \wedge (b + a) \vee -(b + a))] = E[f(\tilde{Y}(a))M_t(a)]. \end{aligned}$$

Again, as in 1.4 we deduce that

$$E[f(\tilde{Y}(a))(M_{t \wedge S_a(\tilde{Y})} - M_{s \wedge S_a(\tilde{Y})})] = 0 \tag{1.18}$$

for $s \leq t$ and f bounded $\mathscr{D}_s$-measurable.

Let again f be bounded and $\mathscr{D}_s$-measurable, and $t > s$. If $S_a > s$ we have $\tilde{Y}_r(a) = \tilde{Y}_r$ for all $r \leq s + \varepsilon$, for some $\varepsilon > 0$, while f is $\mathscr{D}^0_{s+\varepsilon}(\mathbb{R}^{m+1})$-measurable; hence we deduce from Lemma III.2.43 that $f(\tilde{Y}(a)) = f(\tilde{Y})$. If $S_a \leq s$, then

$M_{t \wedge S_a} = M_{s \wedge S_a}$. Thus 1.18 gives:

$$E[f(\tilde{Y})(M_{t \wedge S_a(\tilde{Y})} - M_{s \wedge S_a(\tilde{Y})})] = 0.$$

Since $\mathscr{F}_s^{\tilde{Y}} = \tilde{Y}^{-1}(\mathscr{D}_s)$ it follows that the stopped process $M_{\cdot \wedge S_a(\tilde{Y})}$ is a martingale on $(\Omega, \mathscr{F}, \mathbf{F}^{\tilde{Y}}, P)$.

Finally, there is a sequence $a_n \uparrow \infty$ with $a_n \notin \tilde{V} \cup \tilde{V}'$, and $S_{a_n}(\tilde{Y}) \uparrow \infty$ a. s.: therefore M is a local martingale on $(\Omega, \mathscr{F}, \mathbf{F}^{\tilde{Y}}, P)$. □

1.19 **Corollary.** *Let (M^n) a sequence of local martingales which converges in law to a limit process M, and assume that $|\Delta M^n| \leq b$ identically for some constant b. Then M is a local martingale with respect to the filtration which it generates.*

2. Identification of the Limit

The title of the section is slightly misleading, since we study the following: if a sequence (X^n) of semimartingales converges to a limiting process X, is X a semimartingale with some prescribed characteristics?

We also give an application in § 2d, about the existence of solutions to some martingale problems $\mathscr{s}(\sigma(X_0), X | \eta; B, C, \nu)$ in the sense of § III.2, when B, C, ν have some regularity properties.

§ 2a. Introductory Remarks

Let us first delimitate the setting. As before, for every $n \in \mathbb{N}^*$ we have a stochastic basis $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$. There is a d-dimensional semimartingale $X^n = (X^{n,i})_{i \leq d}$ on $\mathscr{B}^n$, with characteristics (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$ (see II.2.16): these are with respect to a fixed *continuous* truncation function h, which is the same for all $n \in \mathbb{N}^*$.

The limit process is a d-dimensional càdlàg process $X = (X^i)_{i \leq d}$ defined on a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$. On this basis, we also consider:

2.1 (i) $B = (B^i)_{i \leq d}$ a predictable process with finite variation over finite intervals and $B_0 = 0$;

(ii) $C = (C^{ij})_{i,j \leq d}$ a continuous adapted process with $C_0 = 0$ and $C_t - C_s$ is a symmetric nonnegative matrix for all $s \leq t$;

(iii) ν a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$, which charges neither $\mathbb{R}_+ \times \{0\}$ nor $\{0\} \times \mathbb{R}^d$, such that $(1 \wedge |x|^2) * \nu_t(\omega) < \infty$ and $\int \nu(\omega; \{t\} \times dx) h(x) = \Delta B_t(\omega)$ and $\nu(\omega; \{t\} \times \mathbb{R}^d) \leq 1$ identically. □

(The same as III.2.3). We also consider, as in III.2.6:

2.2 $$\tilde{C}^{ij} = C^{ij} + (h^i h^j) * \nu - \sum_{s \leq \cdot} \Delta B^i_s \Delta B^j_s.$$

In all this section we will assume that $X^n \xrightarrow{\mathscr{L}} X$, and we are looking for conditions of the type: B^n, C^n, ν^n converge to B, C, ν in a suitable sense, which would imply that X is a semimartingale with characteristics (B, C, ν).

We observe that conditions $[\beta_5\text{-}D]$, $[\gamma_5\text{-}D]$, $[\delta_{5,i}\text{-}D]$ of Chapter VIII make no sense here, because we do not wish B, C, ν to be deterministic, and they are not defined on the same space than B^n, C^n, ν^n.

1. At first glance, the natural extension of, say, $[\beta_5\text{-}D]$, could be

2.3 $$B^n_t \xrightarrow{\mathscr{L}} B_t \quad \text{for all } t \in D,$$

and similar conditions on $\tilde{C}^n$ and ν^n.

It turns out that these conditions are not quite enough, but in the same spirit we will prove the following in § 2c:

2.4 **Theorem.** *Assume that* $\mathbf{F}$ *is the filtration generated by* X, *and that the following conditions hold*:

2.5 $$\begin{cases} [X\beta\gamma\text{-}\mathscr{L}] & (X^n, B^n, \tilde{C}^n) \xrightarrow{\mathscr{L}} (X, B, \tilde{C}) \\ [X\delta_1\text{-}\mathscr{L}] & (X^n, g * \nu^n) \xrightarrow{\mathscr{L}} (X, g * \nu) \quad \text{for all } g \in C_1(\mathbb{R}^d). \end{cases}$$

Then X *is a semimartingale on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ *with characteristics* (B, C, ν).

(Recall that $C_1(\mathbb{R}^d)$ is defined in VII.2.7; in $[X\beta\gamma\text{-}\mathscr{L}]$ we ask for the weak convergence of the laws of the $(d + d + d^2)$-dimensional processes $(X^n, B^n, \tilde{C}^n)$, for the Skorokhod topology in $\mathbb{D}(\mathbb{R}^{d+d+d^2})$, and similarly in $[X\delta_1\text{-}\mathscr{L}]$).

As we shall see later, this theorem is useful in some contexts. However, it suffers from an important limitation: it supposes that one knows the joint law of X with (B, C, ν) under P, while in most applications we just know that X^n converges in law to some limiting process, whose law P is essentially unknown! this is why we wish for conditions depending on the P^n's, but not on P (as $[\beta_5\text{-}D]$, ... in Chapter VIII).

2. To this end, we assume the following:

2.6 *Hypothesis.* $(\Omega, \mathscr{F}, \mathbf{F})$ is the canonical space $(\mathbb{D}(\mathbb{R}^d), \mathscr{D}(\mathbb{R}^d), \mathbf{D}(\mathbb{R}^d))$ of VI.1.1, and X is the canonical process: $X_t(\alpha) = \alpha(t)$. □

One can consider X^n as a map: $\Omega^n \to \Omega = \mathbb{D}(\mathbb{R}^d)$, so $B \circ X^n$, etc... makes sense. Then our second "natural" generalization of $[\beta_5\text{-}D]$, ... is the following set of conditions, where D denotes a subset of $\mathbb{R}_+$.

$$
2.7 \qquad \begin{cases} [\beta_7\text{-}D] & B_t^n - B_t \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D. \\ [\gamma_7\text{-}D] & \tilde{C}_t^n - \tilde{C}_t \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D. \\ [\delta_{7,i}\text{-}D] & g * \nu_t^n - (g * \nu_t) \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D, g \in C_i(\mathbb{R}^d). \end{cases}
$$

(when B, C, ν are deterministic, these conditions clearly reduce to $[\beta_5\text{-}D]$, $[\gamma_5\text{-}D]$, $[\delta_{5,i}\text{-}D]$).

These conditions may appear a little strange at first glance. To see better what they mean, let us make the connection with a version of an old result, known as Trotter-Kato's Theorem: let $(P_t^n)_{t \geq 0}$ for $n \in \mathbb{N}^*$ and $(P_t)_{t \geq 0}$ be Markov semi-groups on $\mathbb{R}^d$; call A^n and A their infinitesimal generators. Then if $A^n \to A$ in a suitable sense, the semi-groups converge in the following sense: $P_t^n f \to P_t f$ for all t and all "nice" functions f.

It readily follows that if these nice functions include $C(\mathbb{R}^d)$, the space of all bounded continuous functions on $\mathbb{R}^d$, and if moreover (P_t) is Feller, i.e. $P_t f \in C(\mathbb{R}^d)$ for $f \in C(\mathbb{R}^d)$, then finite-dimensional convergence in law along $\mathbb{R}_+$ holds for the corresponding Markov processes X^n and X, provided $X_0^n \xrightarrow{\mathcal{L}} X_0$.

Let us be more specific, assuming that X^n and X are homogeneous (continuous) diffusion processes with coefficients b^n, c^n (resp. b, c): see § III.2c. The generator A^n operates on C^2 functions as

$$
A^n f(x) = \sum_{i \leq d} b^{n,i}(x) D_i f(x) + \frac{1}{2} \sum_{i,j \leq d} c^{n,ij}(x) D_{ij} f(x)
$$

and likewise for A. Then "$A^n \to A$" amounts to saying that

$$
2.8 \qquad b^n \to b, \quad c^n \to c \text{ uniformly,}
$$

while the Feller property of the limiting semi-group (P_t) more or less amounts to

$$
2.9 \qquad b, c \text{ are continuous function on } \mathbb{R}^d.
$$

Therefore under 2.8 and 2.9 and $X_0^n \xrightarrow{\mathcal{L}} X_0$, we obtain $X^n \xrightarrow{\mathcal{L}(\mathbb{R}_+)} X$ by using Trotter-Kato's Theorem.

But the characteristics of X^n and X are

$$
2.10 \qquad \begin{cases} B_t^n = \int_0^t b^n(X_s^n)\, ds, & C_t^n = \int_0^t c^n(X_s^n)\, ds, \quad \nu^n = 0 \\ B_t = \int_0^t b(X_s)\, ds, & C_t = \int_0^t c(X_s)\, ds, \quad \nu = 0 \end{cases}
$$

and so one easily sees that 2.8 implies $[\beta_7\text{-}\mathbb{R}_+]$, $[\gamma_7\text{-}\mathbb{R}_+]$, $[\delta_{7,i}\text{-}\mathbb{R}_+]$.

So in a sense the main results of this chapter, which are based on conditions like 2.7, generalize theorems of the type of Trotter-Kato (see Yoshida [250]) to the non-necessarily Markovian case (and to functional convergence), along the lines developed in the book [233] of Stroock and Varadhan.

§ 2b. Identification of the Limit: The Main Result

This subsection is devoted to proving the next result:

2.11 **Theorem.** *Assume 2.6. Assume that the sequence $\mathscr{L}(X^n)$ weakly converges to a limit P* (a probability measure on $(\Omega, \mathscr{F})$). *Let D be a dense subset of $\mathbb{R}_+$, which is contained in $\mathbb{R}_+ \backslash J(X)$* (where as usual $J(X) = \{t > 0: P(\Delta X_t \neq 0) > 0\}$ is the set of fixed times of discontinuity of X under P). *Moreover, assume that*

(i) $[\beta_7\text{-}D]$, $[\gamma_7\text{-}D]$, $[\delta_{7,1}\text{-}D]$ *hold* (see 2.7);

(ii) *Majoration condition*: $\sup_{\alpha \in \Omega} |\tilde{C}_t(\alpha)| < \infty$, $\sup_{\alpha \in \Omega} |g * \nu_t(\alpha)| < \infty$ for all $t \in \mathbb{R}_+$, $g \in C_1(\mathbb{R}^d)$;

(iii) *Continuity condition*: for all $t \in \mathbb{R}_+$, $g \in C_1(\mathbb{R}^d)$ the functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\alpha)$, $g * \nu_t(\alpha)$ are P-almost surely Skorokhod continuous.

Then X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν).

Proof. a) We set

$$X'_t = X_t - \sum_{s \leq t} [\Delta X_s - h(\Delta X_s)], \qquad V_t = X'_t - B_t - X_0$$

$$X'^n_t = X^n_t - \sum_{s \leq t} [\Delta X^n_s - h(\Delta X^n_s)], \quad V^n_t = X'^n_t - B^n_t - X^n_0$$

(so $X'^n = X' \circ X^n$), and also for $i, j \leq d$ and $g \in C_1(\mathbb{R}^d)$:

$$Z^{ij} = V^i V^j - \tilde{C}^{ij}, \qquad N^g_t = \sum_{s \leq t} g(\Delta X_s) - g * \nu_t$$

$$Z^{n,ij} = V^{n,i} V^{n,j} - \tilde{C}^{n,ij}, \quad N^{n,g}_t = \sum_{s \leq t} g(\Delta X^n_s) - g * \nu^n_t.$$

By virtue of II.2.21, for every n the processes V^n, Z^n, $N^{n,g}$ are local martingales on $\mathscr{B}^n$, and we need to prove that V, Z, N^g are local martingales on $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$.

b) Firstly we prove that the i^{th} component V^i of V is a martingale. It is enough to prove that $M_t = V^i_{t \wedge T}$ is a martingale, for every fixed $T \in D$. By (ii) there is a constant K such that $\tilde{C}^{ii}_T(\alpha) \leq K$ for all α, and we define the stopping time

$$T_n = \inf(t: \tilde{C}^{n,ii}_t \geq K + 1).$$

We will apply Proposition 1.12 to $Y^n = X^n$, $M^n_t = V^{n,i}_{t \wedge T_n \wedge T}$, $Y = X$ and M. We know that M^n is a local martingale on $\mathscr{B}^n$, and $X^n \xrightarrow{\mathscr{L}} X$ (with $\mathscr{L}(X) = P$) by hypothesis, so 1.4(ii) is met. Since h is continuous, VI.2.3 and VI.2.8 yield that $\alpha \rightsquigarrow X'_t(\alpha)$ is continuous at each point α such that $\Delta\alpha(t) = 0$; then one easily deduces from (iii) and from the inclusion $D \subset \mathbb{R}_+ \backslash J(X)$ that $\alpha \rightsquigarrow M_t(\alpha)$ is P-a.s. continuous for all $t \in D$, and thus 1.14(iii) is met.

$[\gamma_7\text{-}D]$ yields $\tilde{C}^{n,ii}_T - \tilde{C}^{ii}_T \circ X^n \xrightarrow{P} 0$; since $\tilde{C}^{ii}_T \circ X^n \leq K$, we deduce that $P^n(\tilde{C}^{n,ii}_T \geq K + 1) \to 0$ as $n \uparrow \infty$, and thus

2.12 $$P^n(T_n < T) \to 0 \quad \text{as } n \uparrow \infty.$$

Since $X'^n = X' \circ X^n$, we have $M_t^n - M_t \circ X^n = B_{t\wedge T}^i \circ X^n - B_{t\wedge T}^{n,i}$ for all $t \geq 0$ on the set $\{T_n \geq T\}$. Then in virtue of 2.12 and $[\beta_7\text{-}D]$, we get

$$M_t^n - M_t \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D,$$

and so 1.4(iv) is met.

Finally it remains to prove 1.12(i′) for the sequence (M^n). Doob's inequality I.1.43 and I.4.6 yield

$$E^n\left(\sup_t |M_t^n|^2\right) \leq 4E^n[(M_\infty^n)^2] = 4E^n(\tilde{C}_{T_n}^{n,ii}),$$

which is bounded by a constant by definition of T_n, and because the jumps of $\tilde{C}^n$ are bounded by another constant, depending only on the truncation function h. Then 1.12(i′) is obviously met, and we are finished.

c) Secondly we prove that Z^{ij} is a martingale. Here again it is enough to prove that $M_t = Z_{t\wedge T}^{ij}$ is so, for every fixed $T \in D$. We follow the same route as in (b), with the following changes: we choose the constant K such that $\tilde{C}_T^{ii}(\alpha) + \tilde{C}_T^{jj}(\alpha) \leq K$ identically, and

$$T_n = \inf(t: \tilde{C}_t^{n,ii} + \tilde{C}_t^{n,jj} \geq K + 1)$$

and $M_t^n = Z_{t\wedge T_n \wedge T}^{n,ij}$, $Y^n = X^n$, $Y = X$. Then 1.4(ii) is met, and we have seen above that $\alpha \rightsquigarrow V_{t\wedge T}^i(\alpha)$ and $\alpha \rightsquigarrow V_{t\wedge T}^j(\alpha)$ are P-a.s. continuous when $t \in D$; then (iii) yields that $\alpha \rightsquigarrow M_t(\alpha)$ is P-a.s. continuous for $t \in D$, and 1.4(iii) is met.

Using $[\gamma_7\text{-}D]$, we deduce 2.12 as above. A simple computation shows that for all $t \in D$,

$$\begin{aligned} M_t^n - M_t \circ X^n = {} & V_{t\wedge T\wedge T_n}^{n,i}(V_{t\wedge T\wedge T_n}^{n,j} - V_{t\wedge T}^j \circ X^n) \\ & + (V_{t\wedge T}^j \circ X^n)(V_{t\wedge T\wedge T_n}^{n,i} - V_{t\wedge T}^i \circ X^n) - \tilde{C}_{t\wedge T\wedge T_n}^{n,ij} + \tilde{C}_{t\wedge T}^{ij} \circ X^n \end{aligned}$$

and we have seen above that $(V_{t\wedge T}^{n,i}|P^n)$ is uniformly integrable (in n) and that $V_{t\wedge T\wedge T_n}^n - V_{t\wedge T} \circ X^n \xrightarrow{P} 0$. Then one easily deduces from $[\gamma_7\text{-}D]$ and 2.12 that $M_t^n - M_t \circ X^n \xrightarrow{P} 0$ for all $t \in D$, and 1.4(iv) is met.

It remains to prove 1.12(i′). As already seen, $\tilde{C}_{T_n}^{n,ii}$ and $\tilde{C}_{T_n}^{n,jj}$ are bounded by a constant K' depending only on K and on the truncation function h, hence Lemma VII.3.34 yields

$$E\left(\sup_t |V_{t\wedge T_n}^{n,i}|^4\right) \leq K'', \quad E\left(\sup_t |V_{t\wedge T_n}^{n,j}|^4\right) \leq K''$$

for another constant K'' depending only on K'. Since $Z^{n,ij} = V^{n,i}V^{n,j} - \tilde{C}^{n,ij}$ we easily deduce that the family $\{Z_{t\wedge T_n}^{n,ij}\}_{t\geq 0, n\geq 1}$ is uniformly integrable, and we are finished.

d) It remains to prove that $M_t = N_{t\wedge T}^g$ is a martingale on $\mathscr{B}$ for all $g \in C_1(\mathbb{R}^d)$, $T \in D$. Again we follow the same route. There is a constant K such that $g * \nu_T(\alpha) \leq K$, and we set

$$T_n = \inf(t: g * \nu_t^n \geq K + 1).$$

We apply Proposition 1.12 once more, to $Y^n = X^n$, $Y = X$, $M_t^n = N_{t\wedge T_n\wedge T}^{n,g}$ and M as above. We have 1.4(ii), and 1.4(iii) is deduced from (iii) and from the fact that $\alpha \rightsquigarrow \sum_{s\le t} g(\Delta\alpha(s))$ is continuous at each point α such that $\Delta\alpha(t) = 0$ (see VI.2.8). Then $[\delta_{7,1}\text{-}D]$ yields 2.12, and on the set $\{T_n \ge T\}$ we have

$$M_t^n - M_t \circ X^n = (g * \nu_{t\wedge T}) \circ X^n - g * \nu_{t\wedge T}^n,$$

so $[\delta_{7,1}\text{-}D]$ again yields 1.4(iv).

Finally M^n is also the stochastic integral of $g 1_{[\![0, T_n\wedge T]\!]}$ with respect to $\mu^{X^n} - \nu^n$, where μ^{X^n} is the random measure associated with the jumps of X^n, so by virtue of II.1.31 and II.1.33 we obtain

$$\langle M^n, M^n\rangle_t \le g^2 * \nu_{t\wedge T_n}^n,$$

which is bounded by a constant K', by definition of T_n and because g itself is bounded and $\nu^n(\{t\} \times \mathbb{R}^d) \le 1$ identically. Then, as in (b), we obtain

$$E^n\left(\sup_t |M_t^n|^2\right) \le 4E^n(g^2 * \nu_{t\wedge T_n}^n) \le 4K'$$

and thus 1.12(i′) is met. □

2.13 **Remark.** The continuity condition (iii) in 2.11 is obviously "minimal" for our argument, but it is *a*-priori difficult to check because it supposes more or less that P is known. It is obviously fulfilled (for any P) if we have:

2.14 The functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\alpha)$, $g * \nu_t(\alpha)$ are continuous on $\mathbb{D}(\mathbb{R}^d)$ for all $t \in \mathbb{R}_+$, $g \in C_1(\mathbb{R}^d)$. □

2.15 **Remark.** The following condition is perhaps more appropriate than 2.14 in the present context:

2.16 The maps: $\alpha \rightsquigarrow B^i(\alpha)$, $\tilde{C}^{ij}(\alpha)$, $g * \nu(\alpha)$ from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R})$ are continuous for the Skorokhod topology, for all $g \in C_1(\mathbb{R}^d)$.

Note that 2.16 does not imply 2.11(iii). But if we set

$$\hat{J}(\alpha) = \{t > 0\colon \nu(\alpha; \{t\} \times \mathbb{R}^d) > 0\}$$

and since $B_\cdot(\alpha)$ and $\tilde{C}_\cdot(\alpha)$ and $g * \nu_\cdot(\alpha)$ are continuous in time at each $t \in \mathbb{R}_+ \backslash \hat{J}(\alpha)$, the conclusion of Theorem 2.11 remains valid if we replace (iii) by 2.16, provided

2.17 $$D \subset \mathbb{R}_+ \backslash \{t\colon P(\Delta X_t \ne 0 \text{ or } \nu(\cdot; \{t\} \times \mathbb{R}^d) > 0) > 0\}.$$ □

2.18 **Remark.** After a close examination of the previous proof one readily sees that instead of (ii) one can impose the following conditions on the characteristics $\tilde{C}^n$ and ν^n:

2.19 For every $t \in \mathbb{R}_+$ and every $n \in \mathbb{N}^*$ there is a stopping time T_t^n on $\mathscr{B}^n$ such that:

(i) $P^n(T_t^n < t) \to 0$ as $n \uparrow \infty$;

(ii) $\sup_{n\in\mathbb{N}^*,\omega\in\Omega^n}|\tilde{C}^n_{t\wedge T_t^n}(\omega)| < \infty$, $\sup_{n\in\mathbb{N}^*,\omega\in\Omega^n}|g * \nu^n_{t\wedge T_t^n}(\omega)| < \infty$ for all $g\in C_1(\mathbb{R}^d)$. □

The uniform majoration 2.11(ii) is quite stringent, and we would like to weaken it as much as possible! It is indeed possible to replace it by:

$$2.20 \quad \text{for all } t > 0,\, g\in C_1(\mathbb{R}^d),\, a > 0, \begin{cases} \sup\limits_{\alpha:\sup_{s\le t}|\alpha(s)|\le a} |\tilde{C}_t(\alpha)| < \infty \\ \sup\limits_{\alpha:\sup_{s\le t}|\alpha(s)|\le a} |g * \nu_t(\alpha)| < \infty, \end{cases}$$

provided we replace (i) and (iii) by "local" conditions as well. For example in the case of continuous diffusion processes (see § 2a), 2.11(ii) amounts to saying that b and c are bounded functions, while 2.20 is fulfilled when they are only locally bounded on $\mathbb{R}^d$ (which is automatic when they are continuous!).

We do not state this "local" theorem here, but we shall see later a local version for our main limit theorem (see 3.39 below).

§ 2c. Identification of the Limit Via Convergence of the Characteristics

Here we prove Theorem 2.4, and we derive some easy consequences.

Proof of Theorem 2.4. We use the notation X', V, Z, N^g and X'^n, V^n, Z^n, $N^{n,g}$ introduced in the proof of 2.11. We know that $V^{n,i}$, $Z^{n,ij}$, $N^{n,g}$ for $g\in C_1(\mathbb{R}^d)$ are local martingales on $\mathscr{B}^n$, and their jumps are bounded by a constant depending on the truncation function h only for V^n and Z^n, and on g only for $N^{n,g}$. Moreover, V, Z, N^g are **F**-adapted and **F** is the filtration generated by X. Hence if

$$(1) \qquad (X^n, V^{n,i}) \xrightarrow{\mathscr{L}} (X, V^i), \quad (X^n, Z^{n,ij}) \xrightarrow{\mathscr{L}} (X, Z^{ij}), \quad (X^n, N^{n,g}) \xrightarrow{\mathscr{L}} (X, N^g)$$

we deduce from 1.17 that V, Z, N^g are local martingales on $(\Omega, \mathscr{F}, \mathbf{F}, P)$, and II.2.21 yields the claim.

Observe that if f is a continuous functions: $\mathbb{R}^m \to \mathbb{R}^{m'}$, then the map $\alpha \rightsquigarrow f(\alpha)$ is continuous from $\mathbb{D}(\mathbb{R}^m)$ into $\mathbb{D}(\mathbb{R}^{m'})$ (this readily follows, e.g., from the characterization VI.1.14a of the Skorokhod topology).

We denote by (α, β, γ) the points of $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d^2})$, and if $\alpha\in\mathbb{D}(\mathbb{R}^d)$ we set $x'(\alpha)_t = \alpha(t) - \sum_{s\le t}[\Delta\alpha(s) - h(\Delta\alpha(s))]$. Then VI.2.8 yields that $(\alpha,\beta,\gamma)\rightsquigarrow(\alpha, x'(\alpha), \beta, \gamma)$ is continuous from $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d^2})$ into $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d^2})$. We deduce that the following maps are continuous:

$$(\alpha,\beta,\gamma)\rightsquigarrow(\alpha, x'^i(\alpha) - \alpha^i(0) - \beta^i)$$

$$(\alpha,\beta,\gamma)\rightsquigarrow(\alpha, (x'^i(\alpha) - \alpha^i(0) - \alpha^i)(x'^j(\alpha) - \alpha^j(0) - \beta^j) - \gamma^{ij})$$

from $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d^2})$ into $\mathbb{D}(\mathbb{R}^d \times \mathbb{R})$. Therefore $[X\beta\gamma\text{-}\mathscr{L}]$ implies the first two convergences in (1).

Similarly if $g \in C_1(\mathbb{R}^d)$ and $\alpha \in \mathbb{D}(\mathbb{R}^d)$ we set $\hat{\alpha}^g(t) = \sum_{s \le t} g(\Delta\alpha(s))$. VI.2.8 yields that $(\alpha, \beta) \rightsquigarrow (\alpha, \hat{\alpha}^g, \beta)$ is continuous from $\mathbb{D}(\mathbb{R}^d \times \mathbb{R})$ in $\mathbb{D}(\mathbb{R}^d \times \mathbb{R} \times \mathbb{R})$, and so $(\alpha, \beta) \rightsquigarrow (\alpha, \hat{\alpha}^g - \beta)$ is continuous from $\mathbb{D}(\mathbb{R}^d \times \mathbb{R})$ into itself. Therefore the last convergence in (1) follows from $[X\delta\text{-}\mathscr{L}]$. □

2.21 **Remark.** If $\mathbf{F}$ is the filtration generated by (X, Y), where Y is an auxiliary càdlàg m-dimensional process, if each basis $\mathscr{B}^n$ is endowed with an adapted càdlàg m-dimensional process Y^n, and if

$$(X^n, Y^n, B^n, \tilde{C}^n) \xrightarrow{\mathscr{L}} (X, Y, B, \tilde{C})$$

$$(X^n, Y^n, g * \nu^n) \xrightarrow{\mathscr{L}} (X, Y, g * \nu)$$

for all $g \in C_1(\mathbb{R}^d)$, the conclusion of 2.4 remains obviously valid. □

We will end this subsection with a result which is half-way between Theorems 2.4 and 2.11.

2.22 **Theorem.** *Assume 2.6. Assume that $\mathscr{L}(X^n)$ weakly converge to a limit P, and*

(i) *If δ_m denotes a distance on $\mathbb{D}(\mathbb{R}^m)$ compatible with the Skorokhod topology, then*

$$[\text{Sk-}X\beta\gamma] \quad \delta_{d+d+d^2}((X^n, B^n, \tilde{C}^n), (X, B, \tilde{C}) \circ X^n) \xrightarrow{P} 0$$

$$[\text{Sk-}X\delta] \quad \delta_{d+1}((X^n, g * \nu^n), (X, g * \nu) \circ X^n) \xrightarrow{P} 0 \quad \text{for all } g \in C_1(\mathbb{R}^d);$$

(ii) *Continuity condition*: The functions $\alpha \rightsquigarrow (\alpha, B(\alpha), \tilde{C}(\alpha))$ from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d^2})$ and $\alpha \rightsquigarrow (\alpha, g * \nu(\alpha))$ from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^{d+1})$ for all $g \in C_1(\mathbb{R}^d)$, are P-almost surely continuous.

Then X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν).

Proof. We will check that the assumptions of Theorem 2.4 are met. Firstly $\mathbf{F}$ is the filtration generated by X, by 2.6. Since $X^n \xrightarrow{\mathscr{L}} X$ (where we implicitely suppose that $\mathscr{L}(X) = P$), (ii) yields

$$(X, B, \tilde{C}) \circ X^n \xrightarrow{\mathscr{L}} (X, B, \tilde{C}) \circ X = (X, B, \tilde{C})$$

$$(X, g * \nu) \circ X^n \xrightarrow{\mathscr{L}} (X, g * \nu) \circ X = (X, g * \nu).$$

Then $[X\beta\gamma\text{-}\mathscr{L}]$ and $[X\delta\text{-}\mathscr{L}]$ readily follow from (i). □

2.23 **Remark.** Note that $X \circ X^n = X^n$, and recall that the Skorokhod topology is coarser than the local uniform topology, and also that the local uniform topology on $\mathbb{D}(\mathbb{R}^m)$ is the m-fold product of the local uniform topology on $\mathbb{D}(\mathbb{R})$. Hence the conditions

[Sup-β_7] $\quad \sup_{t \le N} |B_t^n - B_t \circ X^n| \xrightarrow{P} 0$ for all $N > 0$;

[Sup-γ_7] $\quad \sup_{t \le N} |\tilde{C}_t^n - \tilde{C}_t \circ X^n| \xrightarrow{P} 0$ for all $N > 0$;

[Sup-$\delta_{7,1}$] $\quad \sup_{t \le N} |g * \nu_t^n - (g * \nu_t) \circ X^n| \xrightarrow{P} 0$ for all $N > 0, g \in C_1(\mathbb{R}^d)$;

imply [Sk-$X\beta\gamma$] and [Sk-$X\delta$]. □

§ 2d. Application: Existence of Solutions to Some Martingale Problems

This subsection has the flavor of a topic which is barely touched upon in this book: stochastic differential equations. It will not be used in the sequel. Let

2.24 $(\Omega, \mathscr{F}, \mathbf{F})$ and X are as in 2.6; the space $(\Omega, \mathscr{F}, \mathbf{F})$ is endowed with a triplet (B, C, ν) meeting 2.1, and $\tilde{C}$ is given by 2.2; η is a probability measure on $\mathbb{R}^d$. □

Our aim is to prove existence of solutions to the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$ in the sense of III.2.4, i.e. of probability measures P such that X is a semimartingale with characteristics (B, C, ν) and initial distribution $\mathscr{L}(X_0|P) = \eta$ on $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

Firstly, we list a series of assumptions on (B, C, ν).

2.25 *Factorization property*: There is a *deterministic* increasing càdlàg function $t \rightsquigarrow A_t$ such that

$B_t(\alpha) = \int_0^t b_s(\alpha)\, dA_s$ where $b = (b^i)_{i \le d}$ is a predictable d-dimensional process;

$C_t(\alpha) = \int_0^t c_s(\alpha)\, dA_s$ where $c = (c^{ij})_{i,j \le d}$ is a predictable process with values in the set of all $d \times d$ symmetric non-negative matrices;

$\nu(\alpha; dt, dx) = dA_t\, K_t(\alpha, dx)$ where $(\alpha, t) \rightsquigarrow K_t(\alpha, dx)$ is a positive predictable kernel from $\Omega \times \mathbb{R}_+$ into $\mathbb{R}^d$. □

2.26 *Majoration property*: $\sup_{\alpha,t} |b_t(\alpha)| < \infty$ and $\sup_{\alpha,t} |c_t(\alpha)| < \infty$ and $\sup_{\alpha,t} \int K_t(\alpha, dx)(|x|^2 \wedge 1) < \infty$. □

2.27 *Skorokhod continuity property*: The functions $\alpha \rightsquigarrow B_t(\alpha)$, $\alpha \rightsquigarrow \tilde{C}_t(\alpha)$, $\alpha \rightsquigarrow g * \nu_t(\alpha)$ are continuous for the Skorokhod topology on $\Omega = \mathbb{D}(\mathbb{R}^d)$ for all $t \ge 0$, $g \in C_1(\mathbb{R}^d)$. □

2.28 *Local uniform continuity property*: For all $t \geq 0$, $g \in C_1(\mathbb{R}^d)$ and all Skorokhod compact subsets K of $\mathbb{D}(\mathbb{R}^d)$, the functions $\alpha \rightsquigarrow b_t(\alpha)$, $\alpha \rightsquigarrow c_t^{ij}(\alpha) + \int K_t(\alpha, dx) h^i(x) h^j(x)$, $\alpha \rightsquigarrow \int K_t(\alpha, dx) g(x)$ are uniformly continuous on K, equipped with the local uniform topology (see VI.1.2). $\square$

2.29 **Remark.** Property 2.28 may seem strange. Since the Skorokhod topology is coarser than the local uniform topology, it is weaker than if we had imposed continuity with respect to the Skorokhod topology.

We chose to introduce the "weak" condition 2.28, in spite of its complicated formulation, in order to encompass the *following example*:

Assume that b, c, K have the form $b_t(\alpha) = \bar{b}(t, \alpha(t))$, $c_t(\alpha) = \bar{c}(t, \alpha(t))$, $K_t(\alpha, dx) = \bar{K}_t(\alpha(t), dx)$, where $\bar{b}$ and $\bar{c}$ are functions on $\mathbb{R}_+ \times \mathbb{R}^d$ and where $\bar{K}_t(x, dy)$ is a kernel from $\mathbb{R}_+ \times \mathbb{R}^d$ into $\mathbb{R}^d$. Then if

2.30 (i) $t \rightsquigarrow A_t$ is continuous,

(ii) $\bar{b}(t, \cdot)$, $\bar{c}^{ij}(t, \cdot) + \int \bar{K}_t(\cdot, dx) h^i(x) h^j(x)$, $\int \bar{K}_t(\cdot, dx) g(x)$ are continuous on $\mathbb{R}^d$ for all $t \geq 0$, $g \in C_1(\mathbb{R}^d)$,

then *2.27 and 2.28 are met* (for the former, use VI.2.3 and Lebesgue convergence theorem). $\square$

2.31 **Theorem.** *Assume 2.24, 2.25, 2.26, 2.27 and 2.28, and also*

2.32
$$\lim_{a \uparrow \infty} \sup_{\alpha} K_t(\alpha; \{x: |x| > a\}) = 0 \quad \text{for all } t \geq 0.$$

Then there exists at least one probability measure P on $(\Omega, \mathscr{F})$ such that X is a semimartingale with characteristics (B, C, ν) and initial distribution $\mathscr{L}(X_0) = \eta$ on $(\Omega, \mathscr{F}, \mathbf{F}, P)$.

By virtue of Remark 2.29 we deduce the following corollary, which can be compared to Theorem III.2.34 (no uniqueness in the next result!)

2.33 **Corollary.** *Let $\bar{b}$: $\mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\bar{c}$: $\mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be Borel functions, with $\bar{c}(s, x)$ being symmetric nonnegative; let $\bar{K}_t(x, dy)$ be a Borel kernel from $\mathbb{R}_+ \times \mathbb{R}^d$ into $\mathbb{R}^d$. Assume that 2.30(ii) holds, that $\bar{b}$, $\bar{c}$, $\int \bar{K}_t(x, dy)\,(|y|^2 \wedge 1)$ are bounded, and that*

$$\lim_{a \uparrow \infty} \sup_{x} K_t(x, \{y: |y| > a\}) = 0 \quad \text{for all } t \geq 0.$$

Then for every probability measure η on $\mathbb{R}^d$ there is (at least) a diffusion process X with initial distribution η and with characteristics:

$$B_t = \int_0^t \bar{b}_s(X_s)\, ds, \quad C_t = \int_0^t \bar{c}_s(X_s)\, ds, \quad \nu(dt, dx) = dt\, \bar{K}_t(X_t, dx).$$

Now we proceed to the proof of 2.31, through several steps. Firstly, we construct approximating solutions to the martingale problem. For each $n \in \mathbb{N}^*$

we consider a finite subdivision $0 = t(n,0) < t(n,1) < \cdots < t(n,k_n)$, such that

2.34 $$\begin{cases} \theta^n := \sup\limits_{1 \le p \le k_n} [t(n,p) - t(n,p-1)] \to 0 \\ t(n,k_n) \to \infty \end{cases}$$

as $n \uparrow \infty$. With every $\alpha \in \Omega = \mathbb{D}(\mathbb{R}^d)$ we associate the stopped function α^s at s: $\alpha^s(t) = \alpha(t \wedge s)$, and also the function α^{s-} "stopped strictly before s": $\alpha^{s-}(t) = \alpha(t)$ if $t < s$ and $\alpha^{s-}(t) = \alpha(s-)$ if $t \geq s$.

Next, we set:

2.35 $$\begin{cases} b_0^n(\alpha) = b_0(\alpha), \quad c_0^n(\alpha) = c_0(\alpha), \quad K_0^n(\alpha, \cdot) = K_0(\alpha, \cdot) \\ b_t^n(\alpha) = b_t(\alpha^{t(n,p)}), \quad c_t^n(\alpha) = c_t(\alpha^{t(n,p)}), \quad K_t^n(\alpha, \cdot) = K_t(\alpha^{t(n,p)}, \cdot) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } t(n,p) < t \le t(n,p+1) \\ b_t^n(\alpha) = 0, \quad c_t^n(\alpha) = 0, \quad K_t^n(\alpha, \cdot) = 0 \quad \text{if } t > t(n,k_n) \\ B_t^n = \int_0^t b_s^n \, dA_s, \quad C_t^n = \int_0^t c_s^n \, dA_s, \quad \nu^n(dt,dx) = dA_t \, K_t^n(\cdot, dx). \end{cases}$$

The idea is then to construct a measure P^n for which X has (B^n, C^n, ν^n) for characteristics and η for initial distribution. The technical details are very similar to those of § III.2d, so we will only indicate the scheme of the proof.

Let $\mathscr{F}_t^0 = \mathscr{D}_t^0(\mathbb{R}^d) = \sigma(X_s: s \le t)$, and $t(n,k_n+1) = \infty$. We will construct by induction on $p \in \{0, 1, \ldots, k_n + 1\}$ a measure $P^{n,p}$ such that $\mathscr{L}(X_0 | P^{n,p}) = \eta$ and that X is a semimartingale for $P^{n,p}$ with characteristics

2.36 $$B^{n,p} = (B^n)^{t(n,p)}, \quad C^{n,p} = (C^n)^{t(n,p)}, \quad \nu^{n,p} = (\nu^n)^{t(n,p)},$$

i.e. the characteristics 2.35 stopped at time $t(n,p)$.

We start with $P^{n,0}$, which is the unique measure such that $\mathscr{L}(X_0 | P^{n,0}) = \eta$ and $P^{n,0}(X_t = X_0 \text{ for all } t) = 1$. So the induction hypothesis is obviously met for $p = 0$.

Suppose that the induction hypothesis is satisfied for some p, with the measure $P^{n,p}$. In view of 2.35, the terms

2.37 $$\begin{cases} B'_s = B^n_{(t(n,p)+s) \wedge t(n,p+1)} - B^n_{t(n,p)} \\ C'_s = C^n_{(t(n,p)+s) \wedge t(n,p+1)} - C^n_{t(n,p)} \\ \nu' \text{ defined by } \nu'([0,s] \times G) = \nu^n((t(n,p), t(n,p+1) \wedge (t(n,p)+s)] \times G) \end{cases}$$

are $\mathscr{F}^0_{t(n,p)}$-measurable. Thus by virtue of II.5.2b, for each $\alpha_0 \in \Omega$ there is a probability measure Q_{α_0} on $(\Omega, \mathscr{F})$ for which X is a PII-semimartingale with the deterministic characteristics $(B'(\alpha_0),\ C'(\alpha_0),\ \nu'(\alpha_0))$. Furthermore one easily deduces from II.4.16 that $\alpha \rightsquigarrow Q_\alpha(G)$ is $\mathscr{F}^0_{t(n,p)}$-measurable for all $G \in \mathscr{F}$. Thus, using the same method as for III.2.47 (note that here we have natural shifts θ_t on Ω), one readily obtain a probability measure $P^{n,p+1}$ on $(\Omega, \mathscr{F})$ such that:

1) $P^{n,p+1} = P^{n,p}$ on the σ-field $\mathscr{F}^0_{t(n,p)}$,

2) the $P^{n,p+1}$-conditional law of the process $s \rightsquigarrow X_{t(n,p)+s} - X_{t(n,p)}$ with respect to $\mathscr{F}^0_{t(n,p)}$ is $\alpha \rightsquigarrow Q_\alpha$.

Then, in view of 2.36 and 2.37, an adaptation of III.2.48 shows that under $P^{n,p+1}$, X is a semimartingale with characteristics $(B^{n,p+1}, C^{n,p+1}, \nu^{n,p+1})$ and of course $\mathscr{L}(X_0|P^{n,p+1}) = \eta$.

It remains to put $P^n = P^{n,k_n+1}$, and we have proved the

2.38 Lemma. *X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P^n)$ with characteristics (B^n, C^n, ν^n) and initial distribution $\mathscr{L}(X_0|P^n) = \eta$.*

(One could prove that P^n is indeed the unique solution to the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B^n, C^n, \nu^n))$.

Our second step is:

2.39 Lemma. *The sequence (P^n) is tight.*

Proof. We will apply Theorem VI.5.10. Firstly, since $\mathscr{L}(X_0|P^n) = \eta$, condition VI.5.10(i) is met, while VI.5.10(ii) readily follows from 2.25, 2.26 and 2.32.

2.26 also yields that the increasing process $\sum_{i\le d} \mathrm{Var}(B^{n,i})$ is strongly majorized by γA for some constant γ. Then VI.3.36 implies that the sequence $\mathscr{L}(B^n|P^n)$ is tight, and VI.5.10(iii) is met.

Finally, recall that $g_a(x) = (a|x| - 1)^+ \wedge 1$ belongs to $C_1(\mathbb{R}^d)$ for all $a \in \mathbb{Q}_+$. Then 2.26 yields a constant γ_a such that the increasing function $\gamma_a A$ strongly majorizes $\sum_{i\le d} \tilde{C}^{n,ii} + g_a * \nu^n$, so VI.5.10(iv) is met with (C1). $\square$

Upon taking a subsequence, we can therefore suppose that P^n weakly converges to some limit P. It is then evident that $\mathscr{L}(X_0|P) = \eta$ and so it remains to prove that the conditions of Theorem 2.11 are fulfilled, with $\mathscr{B}^n = (\Omega, \mathscr{F}, \mathbf{F}, P^n)$ and $X^n = X$.

Indeed, 2.11(ii, iii) follow from 2.25, 2.26, 2.27. So we are left to prove 2.11(i). We begin with some notation and auxiliary result:

2.40
$$s_n = \sup(t(n,p): p \ge 0, t(n,p) < s) \quad \text{for } s \in \mathbb{R}_+.$$

2.41 Lemma. *There is a constant γ such that for all $s > 0$, $\eta > 0$, $n \in \mathbb{N}^*$, and if $D(n,\eta,s) = \{\alpha: |\Delta\alpha(r)| \le \eta$ for all $r \in (s_n, s)\}$, then*

$$P^n(D(n,\eta,s)^c) \le \frac{\gamma}{\eta^2 \wedge 1}(A_{s-} - A_{s_n}).$$

Proof. Let μ be the random measure associated with the jumps of X (see II.1.16), so ν^n is the P^n-compensator of μ. We have $1_{D(n,\eta,s)^c} \le \mu((s_n, s) \times \{x: |x| > \eta\})$, and so

$$P^n(D(n,\eta,s)^c) \le E^n[\mu((s_n,s) \times \{x: |x| > \eta\})] = E^n[\nu^n((s_n,s) \times \{x: |x| > \eta\})]$$
$$= E^n\left[\int_{(s_n,s)} dA_r K_r^n(\cdot, \{x: |x| > \eta\})\right] \le \frac{\gamma}{\eta^2 \wedge 1}(A_{s-} - A_{s_n})$$

(where $\gamma = \sup_{t,\alpha} \int K_t(\alpha, dx)(|x|^2 \wedge 1)$), if we use 2.25 and Bienaymé-Tchebycheff inequality. □

Set

2.42 k is one of the following predictable processes:
(i) $k_t(\alpha) = b_t^i(\alpha)$,
(ii) $k_t(\alpha) = c_t^{ij}(\alpha) + \int K_t(\alpha, dx)h^i(x)h^j(x) - \Delta A_t b_t^i(\alpha)b_t^j(\alpha)$,
(iii) $k_t(\alpha) = \int K_t(\alpha, dx)g(x)$ for some $g \in C_1(\mathbb{R}^d)$,

and, according to 2.35,

2.43 $$k_t^n(\alpha) = \begin{cases} k_0(\alpha) & \text{if } t = 0 \\ k_t(\alpha^{t_n}) & \text{if } 0 < t \le t(n,k_n) \\ 0 & \text{if } t > t(n,k_n). \end{cases}$$

2.44 **Lemma.** *We have $E^n(|k_t^n - k_t|) \to 0$ as $n \uparrow \infty$.*

Proof. Since $P^n \to P$ weakly, if $\varepsilon > 0$ there is a compact subset K of $\mathbb{D}(\mathbb{R}^d)$ for the Skorokhod topology, such that $P^n(K^c) \le \varepsilon$ for all n.

We use the notation VI.1.2 for the sup-norm, and VI.1.8 for $w'_N(\alpha,\theta)$. We trivially have $\|\alpha^{t-}\|_N \le \|\alpha\|_N$ and $\|\alpha^{t_n}\|_N \le \|\alpha\|_N$, and $w'_N(\alpha^{t-},\theta) \le w'_N(\alpha,\theta)$ and $w'_N(\alpha^{t_n},\theta) \le w'_N(\alpha,\theta)$: so if $\tilde{K}$ is the set of all functions α, α^{t-}, α^{t_n}, where α ranges through K and $n \in \mathbb{N}^*$, then $\tilde{K}$ is relatively compact (see the characterization VI.1.14b). Since $K \subset \tilde{K}$, we have

2.45 $$P^n(\tilde{K}^c) \le \varepsilon \quad \text{for all } n \in \mathbb{N}^*.$$

Next, we observe that k_t has the continuity property 2.28. Hence there exists $\eta > 0$ (depending on t) such that $\alpha, \beta \in \tilde{K}$ and $\|\alpha - \beta\|_t \le \eta$ imply $|k_t(\alpha) - k_t(\beta)| \le \varepsilon$ (observe that because of III.2.43, $k_t(\alpha) = k_t(\alpha^{t-})$). Then 2.43 implies

2.46 $$\alpha \in \tilde{K}, t \le t(n,k_n) \quad \|\alpha^{t_n} - \alpha^{t-}\|_\infty \le \eta \Rightarrow |k_t^n(\alpha) - k_t(\alpha)| \le \varepsilon.$$

Moreover, $\tilde{K}$ being relatively compact, there exists $\theta > 0$ such that

2.47 $$\alpha \in \tilde{K} \Rightarrow w'_t(\alpha,\theta) \le \frac{\eta}{4}.$$

Now, if $w'_t(\alpha,\theta) \le \frac{\eta}{4}$ there is a subdivision $0 = r_0 < \cdots < r_p = t$ such that $|\alpha(u) - \alpha(v)| \le \frac{\eta}{3}$ for $r_i \le u$, $v < r_{i+1}$, and $r_{i+1} - r_i \ge \theta$ for $i \le p-2$. Hence if

$t(n,k_n) \geq t$ and $\theta_n < \theta$ (see 2.34), there is at most one point $r_i (= r_{p-1})$ in the interval (t_n, t); moreover $\|\alpha^{t_n} - \alpha^{t-}\|_\infty \leq \eta$ unless $r_{p-1} \in (t_n, t)$ and $|\Delta(r_{p-1})| > \frac{\eta}{3}$. Furthermore there is a constant γ' such that $|k_t(\alpha)| \leq \gamma'$ for all α, by 2.26. Hence 2.46 and 2.47 yield

$$\alpha \in \tilde{K}, t(n,k_n) \geq t, \theta_n < \theta \Rightarrow |k_t^n(\alpha) - k_t(\alpha)| \leq \varepsilon + 2\gamma' 1_{D(n,\eta/3,t)^c}.$$

For all n large enough we have $t(n,k_n) \geq t$ and $\theta_n < \theta$. Hence 2.41 and 2.45 yield

$$\begin{aligned} E^n(|k_t^n - k_t|) &\leq \varepsilon + 2\gamma' P^n(\tilde{K}^c) + 2\gamma' P\left[D\left(n, \frac{\eta}{3}, t\right)^c\right] \\ &\leq \varepsilon + 2\gamma'\varepsilon + \frac{2\gamma\gamma'}{(\frac{\eta}{3})^2 \wedge 1}(A_{t-} - A_{t_n}). \end{aligned}$$

Since $A_{t-} - A_{t_n} \to 0$ as $n \uparrow \infty$ (because $t_n \to t$ and $t_n < t$) and $\varepsilon > 0$ is arbitrary, we obtain the claim. □

Finally, in virtue of 2.35 and 2.25 we have

$$|B_t^{n,i} - B_t^i \circ X^n| \leq \int_0^t |k_s^n - k_s| \, dA_s \quad \text{with } k \text{ as in 2.42(i)}$$

$$|\tilde{C}_t^{n,ij} - \tilde{C}_t^{ij} \circ X^n| \leq \int_0^t |k_s^n - k_s| \, dA_s \quad \text{with } k \text{ as in 2.42(ii)}$$

$$|g * \nu_t^n - (g * \nu_t) \circ X^n| \leq \int_0^t |k_s^n - k_s| \, dA_s \quad \text{with } k \text{ as in 2.42(iii)}.$$

Moreover, in all cases $k^n - k$ is uniformly bounded; therefore Lemma 2.44 and Lebesgue convergence theorem yield [β_7-$\mathbb{R}_+$] and [γ_7-$\mathbb{R}_+$] and [$\delta_{7,1}$-$\mathbb{R}_+$] and the proof of Theorem 2.31 is complete.

3. Limit Theorems for Semimartingales

We at last proceed to the heart of our subject. The setting is the same as in the previous section. For each $n \in \mathbb{N}^*$, $\mathscr{B}^n$ is a stochastic basic endowed with a semimartingale $X^n = (X^{n,i})_{i \leq d}$, with characteristics (B^n, C^n, ν^n) and $\tilde{C}^n$ (the truncation function h is *continuous* and fixed throughout). $(\Omega, \mathscr{F}, \mathbf{F})$ is the canonical space $\mathbb{D}(\mathbb{R}^d)$ with the canonical process X (see 2.6) and it is equipped with the triplet (B, C, ν) and $\tilde{C}$ as in 2.1 and 2.2. We also consider:

3.1 $$\begin{cases} \eta^n = \mathscr{L}(X_0^n), \text{ the initial distribution of } X^n \text{ (a measure on } \mathbb{R}^d), \\ \eta = \text{a probability measure on } \mathbb{R}^d \text{ (to be the initial distribution of our limiting process).} \end{cases}$$

The convergence conditions on the characteristics will be of the following type:

—either $[\beta_7\text{-}D]$, $[\gamma_7\text{-}D]$, $[\delta_{7,1}\text{-}D]$, as in 2.7;

—or $[\text{Sup-}\beta_7]$, $[\text{Sup-}\gamma_7]$, $[\text{Sup-}\delta_{7,1}]$, as in 2.23;

—or (with δ_m denoting a distance on $\mathbb{D}(\mathbb{R}^m)$ compatible with the Skorokhod topology):

3.2
$$\begin{cases} [\text{Sk-}\beta_7] & \delta_d(B^n, B\circ X^n) \xrightarrow{P} 0 \\ [\text{Sk-}\gamma_7] & \delta_{d^2}(\tilde{C}^n, \tilde{C}\circ X^n) \xrightarrow{P} 0 \\ [\text{Sk-}\delta_{7,1}] & \delta_1(g*\nu^n, (g*\nu)\circ X^n) \xrightarrow{P} 0 \quad \text{for all } g\in C_1(\mathbb{R}^d) \\ [\text{Sk-}\beta\gamma\delta_7] & \delta_{d+d^2+m}((B^n, \tilde{C}^n, g*\nu^n), (B, \tilde{C}, g*\nu)\circ X^n) \xrightarrow{P} 0 \text{ for all} \\ & \text{functions } g\colon \mathbb{R}^d \to \mathbb{R}^m \text{ whose components are in } C_1(\mathbb{R}^d). \end{cases}$$

At this stage, we have criteria for identifying the characteristics of the limits of the sequence $\mathscr{L}(X^n)$. It remains to prove tightness for the sequence (X^n) (in § 3a), and then to assume the uniqueness property for the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$, and we will deduce the convergence result.

§ 3a. Tightness of the Sequence (X^n)

1. We begin with two lemmas.

3.3 **Lemma.** *For every $n\in\mathbb{N}^*$, let $G^n = (G^{n,i})_{i\le m}$ be a càdlàg m-dimensional process on $\mathscr{B}^n$ with $G^n_0 = 0$. Let $G = (G^i)_{i\le m}$ be a càdlàg m-dimensional process with finite variation and $G_0 = 0$, defined on $(\Omega, \mathscr{F})$, and such that $\sum_{i\le m}\mathrm{Var}(G^i)$ is strongly majorized by an increasing càd deterministic function F.*

a) *If $\sup_{s\le t}|G^n_s - G_s\circ X^n| \xrightarrow{P} 0$ for all $t\ge 0$ and if F is continuous, then the sequence $\{G^n\}$ is C-tight.*

b) *If $\delta_m(G^n, G\circ X^n) \xrightarrow{P} 0$, where δ_m is a distance on $\mathbb{D}(\mathbb{R}^m)$ compatible with the Skorokhod topology, then the sequence $\{G^n\}$ is tight.*

Observe that the first assumption in (a) implies the assumption in (b).

Proof. We begin with (b). Since $\sum_{i\le m}\mathrm{Var}(G^i\circ X^n) \prec F$ by hypothesis, VI.3.35 and VI.3.36 imply that the sequence $\{G\circ X^n\}$ is tight. Moreover, if a subsequence $\{G\circ X^{n_k}\}$ converges in law to a process Y, the assumption in (b) clearly implies that $G^{n_k} \xrightarrow{\mathscr{L}} Y$ as well. These two properties yield that any subsequence of $\{G^n\}$ admits a further subsequence which converges, and so the sequence $\{G^n\}$ itself is tight.

Now we turn to (a). From what precedes, the sequence $\{G^n\}$ is tight and it has the same limit points than the sequence $\{G\circ X^n\}$. F being continuous, $\{G\circ X^n\}$ is C-tight (see VI.3.35), and thus so is $\{G^n\}$. □

3.4 **Lemma.** *For every* $n \in \mathbb{N}^*$, *let* G^n, H^n *be càdlàg processes on* $\mathscr{B}^n$ *with* $G_0^n = H_0^n = 0$, *and* G^n *with finite variation, and* H^n *increasing, and such that* $\mathrm{Var}(G^n) \prec H^n$. *On* $(\Omega, \mathscr{F})$ *let* G, H *be two càdlàg processes with* $G_0 = H_0 = 0$ *and* G *with finite variation and* H *increasing, such that* $\mathrm{Var}(G) \prec H \prec F$, *where* F *is a continuous deterministic increasing function. Then if*

3.5
$$\left.\begin{array}{l} G_t^n - G_t \circ X^n \xrightarrow{P} 0 \\ H_t^n - H_t \circ X^n \xrightarrow{P} 0 \end{array}\right\} \quad \text{for all } t \in D$$

where D *is a dense subset on* $\mathbb{R}_+$, *then* $\sup_{s \leq t} |G_s^n - G_s \circ X^n| \xrightarrow{P} 0$ *for all* $t \geq 0$.

(This result is similar to VI.3.37a).

Proof. Let $\varepsilon > 0$ and $t_0 = 0 < t_1 \cdots < t_p < \cdots$ with $t_i \in D$ for $i \geq 1$, and $F_{t_{i+1}} - F_{t_i} \leq \varepsilon$, and $\lim_p t_p = \infty$ (recall that F is continuous). Then if $t_i \leq s \leq t_{i+1}$,

$$\begin{aligned} |G_s^n - G_s \circ X^n| &\leq |G_s^n - G_{t_i}^n| + |G_{t_i}^n - G_{t_i} \circ X^n| + |G_{t_i} \circ X^n - G_s \circ X^n| \\ &\leq H_{t_{i+1}}^n - H_{t_i}^n + |G_{t_i}^n - G_{t_i} \circ X^n| + \varepsilon \\ &\leq |H_{t_{i+1}}^n - H_{t_{i+1}} \circ X^n| + |H_{t_i}^n - H_{t_i} \circ X^n| + |G_{t_i}^n - G_{t_i} \circ X^n| + 2\varepsilon \end{aligned}$$

and so

$$\sup_{s \leq t} |G_s^n - G_s^n \circ X^n| \leq 2\varepsilon + \sup_{i: t_i \leq t} \{2|H_{t_i}^n - H_{t_i} \circ X^n| + |G_{t_i}^n - G_{t_i} \circ X^n|\}.$$

Then 3.5 yields

$$P^n\left(\sup_{s \leq t} |G_s^n - G_s \circ X^n| > 3\varepsilon\right) \to 0,$$

and since $\varepsilon > 0$ is arbitrary, we get our claim. □

2. Next we prove the tightness criterion in the case where the limiting process is quasi-left continuous (i.e., $\nu(\{t\} \times \mathbb{R}^d) = 0$ for all t).

3.6 *Strong majoration hypothesis* (I). There is a *continuous* and *deterministic* increasing function $t \rightsquigarrow F_t$ which strongly majorizes the functions $\sum_{i \leq d} \mathrm{Var}(B^i(\alpha))$ and $\sum_{i \leq d} C^{ii}(\alpha) + (|x|^2 \wedge 1) * \nu(\alpha)$ for all $\alpha \in \Omega$. □

3.7 *Condition on the big jumps.* For all $t \geq 0$ we have:

$$\lim_{a \uparrow \infty} \sup_{\alpha \in \Omega} \nu(\alpha; [0, t] \times \{x: |x| > a\}) = 0. \qquad \square$$

If we apply Lemma 3.4 to $G^n = H^n = g * \nu^n$ for $g \in C_1(\mathbb{R}^d)$, and to $G^n = H^n = \tilde{C}^{n,ii}$, and to $G^n = \tilde{C}^{n,ij}$ and $H^n = \tilde{C}^{n,ii} + \tilde{C}^{n,jj}$, we obtain:

3.8 **Lemma.** *Under the strong majoration condition 3.6, for every dense subset* D *of* $\mathbb{R}_+$ *we have* $[\gamma_7\text{-}D] \Leftrightarrow [\mathrm{Sup}\text{-}\gamma_7]$ *and* $[\delta_{7,1}\text{-}D] \Leftrightarrow [\mathrm{Sup}\text{-}\delta_{7,1}]$.

3.9 **Theorem.** *Assume 3.6 and 3.7, and that the sequence $\{X_0^n\}_{n\in\mathbb{N}^*}$ is tight* (in $\mathbb{R}^d$). *Then under* [Sup-β_7] + [γ_7-D] + [$\delta_{7,1}$-D] *for some dense subset D of $\mathbb{R}_+$, the sequence $\{X^n\}$ is tight.*

Proof. We will apply Theorem VI.4.18, whose condition (i) is met by hypothesis.

For every rational b the function $g_b(x) = (b|x| - 1)^+ \wedge 1$ belongs to $C_1(\mathbb{R}^d)$, and we clearly have

3.10 $$g_b(x) \leq (b^2 \vee 1)(|x|^2 \wedge 1).$$

Let $t \in D$, $\varepsilon > 0$, $\eta > 0$. In view of 3.7, there exists $a \in \mathbb{Q}_+$ such that $g_{2/a} * \nu_t(\alpha) \leq \varepsilon/2$ for all $\alpha \in \Omega$. [$\delta_{7,1}$-D] implies that for all n large enough, $P^n(|g_{2/a} * \nu_t^n - (g_{2/a} * \nu_t) \circ X^n| \geq \varepsilon/2) \leq \eta$. Since $\nu^n([0,t] \times \{x: |x| > a\}) \leq g_{2/a} * \nu_t^n$, we deduce

$$P^n(\nu^n([0,t] \times \{x: |x| > a\}) \geq \varepsilon) \leq \eta.$$

Since $\varepsilon > 0$ and $\eta > 0$ and $t \in D$ are arbitrary, VI.4.18(ii) follows.

[Sup-β_7], plus 3.6 and Lemma 3.3a applied to $G^n = B^n$, yield that the sequence $\{B^n\}$ is C-tight. Secondly, we have [Sup-γ_7] by 3.8, and 3.6 yields a constant γ such that $\tilde{C}^{ii}(\alpha) \prec \gamma F$ for all $\alpha \in \Omega$; then 3.3a implies that each sequence $\{\tilde{C}^{n,ii}\}_{n\in\mathbb{N}^*}$ is C-tight, and thus $\{\tilde{C}^n\}$ is C-tight by VI.3.36b. Thirdly, we have [Sup-$\delta_{7,1}$], and 3.6 and 3.10 give $g_b * \nu(\alpha) \prec (|b|^2 \wedge 1)F$; therefore 3.3a implies that each sequence $\{g_b * \nu^n\}_{n\in\mathbb{N}}$ is C-tight: hence VI.4.18(iii) is met, and we are finished. □

3. Now we turn to the "general" case. The end of this subsection may be omitted by a reader interested only in quasi-left continuous processes. We have to replace 3.6 by the following *weaker* assumption:

3.11 *Strong majoration condition* (II). There are *deterministic* increasing càdlàg functions $t \rightsquigarrow F_t$, F_t^g such that for all $\alpha \in \Omega$, $g \in C_1(\mathbb{R}^d)$, $\sum_{i\leq d} \mathrm{Var}(B^i(\alpha)) \prec F$ and $\sum_{i\leq d} \tilde{C}^{ii}(\alpha) \prec F$ and $g * \nu(\alpha) \prec F^g$. □

3.12 **Remarks.** 1) This condition allows ν to have $\nu(\alpha; \{t\} \times \mathbb{R}^d) > 0$, but only for a countable set of times t which *does not depend on* α! so this condition is still very stringent, and this fact drastically limits the interest of what follows.

2) Note that Hypothesis 3.6, even with the continuous function F replaced by a càdlàg function, remains stronger than 3.11.

3) Note also that, in the setting of § 2d, conditions 2.25 and 2.26 imply 3.11. □

3.13 **Lemma.** *Assume 3.11. If* $D = \{t: \Delta F_t = 0$ and $\Delta F_t^g = 0$ for all $g \in C_1(\mathbb{R}^d)\}$ *is the set of continuity points of all functions F and F^g, we have* [Sk-β_7] $\Rightarrow$ [β_7-D], [Sk-γ_7] $\Rightarrow$ [γ_7-D], [Sk-$\delta_{7,1}$] $\Rightarrow$ [$\delta_{7,1}$-D].

Observe that, if $J = \{t\colon \Delta F_t^{g_a} > 0 \text{ for some } a \in \mathbb{Q}_+\}$, under 3.11 the processes $g * \nu(\alpha)$ are still strongly majorized by the increasing functions $F'^g = 1_J \cdot F^g + (F^g)^c$, while J is at most countable: so in the statement above it is always possible to choose the F^g's so that the complement of D is at most countable.

Proof. We will prove only [Sk-β_7] $\Rightarrow$ [β_7-D], the other claims being proved similarly. So we assume [Sk-β_7], and we pick $t \in D$. We have to prove that $B_t^n - B_t \circ X^n \xrightarrow{P} 0$. For this it suffices to prove that from any subsequence (n') one may extract a further subsequence (n'') such that $B_t^{n''} - B_t \circ X^{n''} \xrightarrow{P} 0$.

Upon taking the product space $\prod_{n \in \mathbb{N}^*} \mathscr{B}^n$, we can always assume that all processes are defined on the same basis. So, upon taking a subsequence, still denoted by n, we can assume that $\delta_d(B^n, B \circ X^n) \to 0$ a. s.

Now we fix ω such that $\delta_d(B^n(\omega), B \circ X^n(\omega)) \to 0$, and we will prove that $B_t^n(\omega) - B_t \circ X^n(\omega) \to 0$. Again, it suffices to prove that from any subsequence one may extract a further subsequence having this property. Using VI.3.35 and VI.3.36 (for the "deterministic" process $B \circ X^n(\omega)$), and 3.11, we observe that the sequence $\{B \circ X^n(\omega)\}_{n \in \mathbb{N}^*}$ is relatively compact in $\mathbb{D}(\mathbb{R}^d)$. Hence, upon taking a subsequence, still denoted by n, we can assume that $B \circ X^n(\omega) \to \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ for some limit α, and $\delta_d(B^n(\omega), B \circ X^n(\omega)) \to 0$ implies $B^n(\omega) \to \alpha$ as well. Hence if $\Delta\alpha(t) = 0$ we have $B_t^n(\omega) \to \alpha(t)$ and $B_t \circ X^n(\omega) \to \alpha(t)$, and thus $B_t^n(\omega) - B_t \circ X^n(\omega) \to 0$.

Thus it remains to prove that $\Delta\alpha(t) = 0$. By VI.2.1 there is a sequence $t_n \to t$ with $\Delta B_{t_n} \circ X^n(\omega) \to \Delta\alpha(t)$. But $|\Delta B_{t_n} \circ X^n(\omega)| \le \Delta F_{t_n}$, which goes to 0 as $n \uparrow \infty$ because $\Delta F_t = 0$: hence $\Delta\alpha(t) = 0$ and we are finished. □

3.14 **Lemma.** *Under 3.11 we have* [Sk-β_7] + [Sk-γ_7] + [Sk-$\delta_{7,1}$] $\Leftrightarrow$ [Sk-$\beta\gamma\delta_7$].

Proof. Only the implication $\Rightarrow$ needs proving, and so we assume [Sk-β_7] and [Sk-γ_7] and [Sk-$\delta_{7,1}$]. We can, and will, assume that $C_1(\mathbb{R}^d)$ is countable.

a) The beginning of the proof is very similar to that of 3.13. The same "subsequence principle" holds, and we can assume that all processes are defined on the same probability space. Up to taking a subsequence, we can also assume that

3.15
$$\begin{cases} \delta_d(B^n, B \circ X^n) \to 0 \\ \delta_{d^2}(\tilde{C}^n, \tilde{C} \circ X^n) \to 0 \\ \delta_1(g * \nu^n, (g * \nu) \circ X^n) \to 0 \quad \text{for all } g \in C_1(\mathbb{R}^d) \end{cases}$$

outside a null set.

Now we fix a point such that 3.15 holds, say ω. Using 3.11 and VI.3.35 and VI.3.36, we obtain that the sequences of functions $\{B \circ X^n(\omega)\}$, $\{\tilde{C} \circ X^n(\omega)\}$, $\{g * \nu \circ X^n(\omega)\}$ are relatively compact in $\mathbb{D}(\mathbb{R}^d)$, $\mathbb{D}(\mathbb{R}^{d^2})$, $\mathbb{D}(\mathbb{R})$ respectively. Hence from any subsequence we may extract a further subsequence (n') such that (use 3.15):

3.16 $$\begin{cases} B \circ X^{n'}(\omega) \to \beta, \quad B^{n'}(\omega) \to \beta \quad \text{in } \mathbb{D}(\mathbb{R}^d) \\ \tilde{C} \circ X^{n'}(\omega) \to \gamma, \quad \tilde{C}^{n'}(\omega) \to \gamma \quad \text{in } \mathbb{D}(\mathbb{R}^{d^2}) \\ (g * \nu) \circ X^{n'}(\omega) \to \delta^g, \quad g * \nu^{n'}(\omega) \to \delta^g \quad \text{in } \mathbb{D}(\mathbb{R}) \quad \text{for all } g \in C_1(\mathbb{R}^d) \end{cases}$$

and clearly $\gamma(t) - \gamma(s)$ is a nonnegative symmetric matrix if $s \leq t$.

b) Consider now 3.16, for our fixed ω: it says that $\{B^{n'}(\omega)\}$ meets condition [Sk-β_3] of Chapter VII with β instead of B, and $\{\tilde{C}^{n'}(\omega)\}$ meets [Sk-γ_3] with γ instead of $\tilde{C}$, and $\{\nu^{n'}(\omega)\}$ meets [Sk-$\delta_{3,1}$] except that we do not know whether δ^g is obtained by integrating g against a measure.

However, a close look at the proof of Lemma VII.3.42 shows that this last property plays no rôle, provided we replace $g * \nu$ by δ^g and $\nu(\{t\} \times g)$ by $\Delta\delta^g(t)$ and $J(X)$ by $\{t: \Delta\delta^g(t) \neq 0$ for at least one $g \in C_1(\mathbb{R}^d)\}$, and provided also that in (i) we only wish the property for $g \in C_1(\mathbb{R}^d)$. So, due to this lemma, for every $t > 0$ we obtain a sequence $t_{n'} \to t$ with

3.17 $$\begin{cases} \Delta(g * \nu^{n'})_{t_{n'}}(\omega) = \nu^{n'}(\omega; \{t_{n'}\} \times g) \to \Delta\delta^g(t) \quad \text{for } g \in C_1(\mathbb{R}^d) \\ \lim_{\eta \downarrow 0} \limsup_{n'} \nu^{n'}(\omega; ([t-\eta, t+\eta] \backslash \{t_{n'}\}) \times \{|x| > \varepsilon\}) = 0 \quad \text{for } \varepsilon > 0. \end{cases}$$

If $\Delta\beta(t) = 0$, then $\Delta B^{n'}_{t_{n'}}(\omega) \to 0$ by 3.16. If $\Delta\beta(t) \neq 0$ there is a sequence $s_{n'} \to t$ such that $\Delta B^{n'}_{s_{n'}}(\omega) \to \Delta\beta(t)$. Since $\Delta B^{n'}_{s_{n'}}(\omega) = \nu^{n'}(\omega; \{s_{n'}\} \times h)$, we easily deduce from the second property 3.17 and from $\Delta\beta(t) \neq 0$ that $s_{n'} = t_{n'}$ for all n' large enough, and thus $\Delta B^{n'}_{t_{n'}}(\omega) \to \Delta\beta(t)$. We prove similarly that $\Delta\tilde{C}^{n'}_{t_{n'}}(\omega) \to \Delta\gamma(t)$.

Hence the conclusion of Corollary VII.3.46 is valid here. Then if we reproduce Part (c) of the proof of VII.3.48, we obtain

3.18 $$(B^{n'}(\omega), \tilde{C}^{n'}(\omega), g^1 * \nu^{n'}(\omega), \ldots, g^m * \nu^{n'}(\omega)) \to (\beta, \gamma, \delta^{g^1}, \ldots, \delta^{g^m})$$

in $\mathbb{D}(\mathbb{R}^{d+d^2+m})$, for all $g^i \in C_1(\mathbb{R}^d)$.

c) The same proof, based also on 3.16, shows that 3.18 holds with $B^{n'}$, $\tilde{C}^{n'}$, $\nu^{n'}$ replaced by $B \circ X^{n'}$, $\tilde{C} \circ X^{n'}$, $\nu \circ X^{n'}$. So we deduce

3.19 $$\delta_{d+d^2+m}[(B^{n'}(\omega), \tilde{C}^{n'}(\omega), g^1 * \nu^{n'}(\omega), \ldots, g^m * \nu^{n'}(\omega)), \\ (B, \tilde{C}, g^1 * \nu, \ldots, g^m * \nu) \circ X^{n'}(\omega)] \to 0.$$

By the subsequence principle, we indeed have 3.19 for any sequence (n) and any point ω such that 3.15 holds. Another application of the subsequence principle shows that [Sk-$\beta\gamma\delta_7$] holds. □

3.20 **Theorem.** *Assume 3.11 and 3.7, and that the sequence* $\{X_0^n\}_{n \in \mathbb{N}^*}$ *is tight* (in $\mathbb{R}^d$). *Then under* [Sk-β_7] + [Sk-γ_7] + [Sk-$\delta_{7,1}$], *the sequence* $\{X^n\}$ *is tight.*

Proof. We will apply Theorem VI.5.10, whose condition (i) is met by hypothesis. By Lemma 3.13 (and the comments which follow this lemma) we have [$\delta_{7,1}$-D] for a set D whose complement in $\mathbb{R}_+$ is at most countable. Then the same proof than in 3.9 yields that VI.5.10(ii) is met. Moreover 3.3b and [Sk-β_7] show that the sequence $\{B^n\}$ is tight, hence VI.5.10(iii).

It remains to prove VI.5.10(iv), for which we will use Condition (C2). Let $G^{n,a} = \sum_{i \le d} \tilde{C}^{n,ii} + g_a * v^n$, where $g_a(x) = (a|x| - 1)^+ \wedge 1$ belongs to $C_1(\mathbb{R}^d)$ for $a \in \mathbb{Q}_+$. Set $G^a = \sum_{i \le d} \tilde{C}^{ii} + g_a * v$.

Let (n') be a subsequence. Since $G^a \circ X^n \prec \bar{F}^a := F + F^{g_a}$ and since $\bar{F}^a$ is deterministic, VI.3.35 implies that the sequence $\{G^a \circ X^n\}_{n \in \mathbb{N}^*}$ is tight. So there is a subsequence (n'') of (n') with $G^a \circ X^{n''} \xrightarrow{\mathscr{L}} \bar{G}^a$ for all $a \in \mathbb{Q}_+$, as $n'' \uparrow \infty$. Moreover $G^a_t \circ X^n - G^a_s \circ X^n \le \bar{F}^a_t - \bar{F}^a_s$ for all n, so $\bar{G}^a_t - \bar{G}^a_s \le \bar{F}^a_t - \bar{F}^a_s$ a.s. for all $s \le t$: in other words, $\bar{G}^a \prec \bar{F}^a$ a.s.

Furthermore we have [Sk-$\beta\gamma\delta_7$] by 3.14. Hence for all $a \in \mathbb{Q}_+$

$$\delta_1(G^{n,p}, G^p \circ X^n) \xrightarrow{P} 0 \quad \text{as } n \uparrow \infty,$$

and we deduce that $G^{n'',a} \xrightarrow{\mathscr{L}} \bar{G}^a$ as $n'' \uparrow \infty$. In other words, all sequences $(G^{n'',a})_{n''}$ satisfy Condition (C2) of VI.5.4. Hence VI.5.10 implies that the sequence $\{X^{n''}\}$ is tight.

So, from any subsequence (n') one can extract a further subsequence (n'') for which $\{X^{n''}\}$ is tight: this clearly implies that the sequence $\{X^n\}$ itself is tight, and we are finished. □

§ 3b. Limit Theorems: The Bounded Case

1. Here again we begin with the "quasi-left continuous" case for the limiting process. The main theorem is as follows (the setting is as before § 3a):

3.21 **Theorem.** *Let D be a dense subset of* $\mathbb{R}_+$, *and assume*:

(i) *The strong majoration hypothesis* (I): *see 3.6.*

(ii) *The condition 3.7 on big jumps.*

(iii) *Uniqueness*: there is a unique probability measure P on $(\Omega, \mathscr{F})$ such that the canonical process X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, v) and initial distribution η (in other words, $\mathscr{s}(\sigma(X_0), X|\eta; B, C, v) = \{P\}$).

(iv) *Continuity condition*: For all $t \in D$, $g \in C_1(\mathbb{R}^d)$, the functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\omega)$, $g * v_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$.

(v) $\eta^n \to \eta$ *weakly* (on $\mathbb{R}^d$).

(vi) [Sup-β_7], [γ_7-D], [$\delta_{7,1}$-D] *hold.*

Then the laws $\mathscr{L}(X^n)$ weakly converge to P.

Proof. a) Firstly we apply Theorem 3.9: because of (i, ii, v, vi) we obtain that the sequence $\{X^n\}$ is tight, and it remains to prove that P is the only limit point of the sequence $\{\mathscr{L}(X^n)\}$.

So we can assume that $\mathscr{L}(X^n) \to P'$ weakly, for some measure P'. Let $J'(X) = \{t > 0: P'(\Delta X_t \ne 0) > 0\}$, and assume that $D' = D \cap (\mathbb{R}_+ \backslash J'(X))$ is still dense in $\mathbb{R}_+$ (this is the case, e.g., when $D = \mathbb{R}_+$). We clearly have (i) $\Rightarrow$ 2.11(ii), while

(vi) and (iv) imply 2.11(i, iii) with D' instead of D. Hence Theorem 2.11 yields that X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν). Since $\mathscr{L}(X_0^n) \to \mathscr{L}(X_0|P')$ we deduce from (v) that $\mathscr{L}(X_0|P') = \eta$, and thus the uniqueness condition (iii) yields $P' = P$.

b) It remains to prove that D' is dense in $\mathbb{R}_+$, and for this it is enough to prove that $J'(X) = \emptyset$. Let $t > 0$, $\varepsilon > 0$ with $\varepsilon \in \mathbb{Q}$. In view of 3.6 there exist $s, s' \in D$ with $s < t < s'$ and $g_{2/\varepsilon} * \nu_{s'}(\alpha) - g_{2/\varepsilon} * \nu_s(\alpha) \leq \varepsilon$ for all $\alpha \in \mathbb{D}(\mathbb{R}^d)$. There also exist $r, r' \in \mathbb{R}_+ \backslash J'(X)$ with $s \leq r < t < r' \leq s'$. An easy consequence of VI.2.4, plus $\mathscr{L}(X^n) \to P'$, yield $\mathscr{L}(\sup_{r<u\leq r'} |\Delta X_u^n|) \to \mathscr{L}(\sup_{r<u\leq r'} |\Delta X_u| \,|P')$ weakly, and thus

3.22
$$\begin{aligned} P'(|\Delta X_t| > \varepsilon) &\leq P'\Big(\sup_{r<u\leq r'} |\Delta X_u| > \varepsilon\Big) \\ &\leq \limsup_n P^n\Big(\sup_{r<u\leq r'} |\Delta X_u^n| > \varepsilon\Big) \\ &\leq \limsup_n P^n\Big(\sup_{s<u\leq s'} |\Delta X_u^n| > \varepsilon\Big) \\ &\leq \limsup_n P^n\Big(\sum_{s<u\leq s'} g_{2/\varepsilon}(\Delta X_u^n) \geq 1\Big). \end{aligned}$$

By definition of ν^n, the predictable increasing process $1_{(s,\infty)} g_{2/\varepsilon} * \nu^n$ is the P^n-compensator of $\sum_{s<u\leq \cdot} g_{2/\varepsilon}(\Delta X_u^n)$. So Lenglart's inequality I.3.31 and 3.22 yield

3.23
$$P'(|\Delta X_t| > \varepsilon) \leq 2\varepsilon + \limsup_n P^n(g_{2/\varepsilon} * \nu_{s'}^n - g_{2/\varepsilon} * \nu_s^n \geq 2\varepsilon).$$

Now, $[\delta_{7,1}\text{-}D]$ and the property $s, s' \in D$ yield $|g_{2/\varepsilon} * \nu_s^n - (g_{2/\varepsilon} * \nu_s) \circ X^n| \xrightarrow{P} 0$, and similarly for s'. Since $|g_{2/\varepsilon} * \nu_{s'} - g_{2/\varepsilon} * \nu_s| \circ X^n \leq \varepsilon$ identically, we deduce from 3.23 that $P'(|\Delta X_t| > \varepsilon) \leq 2\varepsilon$. Since ε is arbitrary in $\mathbb{Q}_+ \backslash \{0\}$, it follows that $P'(|\Delta X_t| > 0) = 0$, and so $t \notin J'(X)$. This being true for all $t > 0$, we obtain $J'(X) = \emptyset$, and we are finished. □

2. Now we state the "square-integrable" version of the previous theorem. We assume that each X^n is a locally square-integrable semimartingale, which amounts to saying that

3.24
$$|x|^2 * \nu_t^n < \infty \quad \text{for all } t \in \mathbb{R}_+$$

(see II.2.27). And, according to II.2.29, we can define the first and modified second characteristics "without truncation" as

3.25
$$\begin{cases} B'^n = B^n + (x - h(x)) * \nu^n \\ \tilde{C}_t'^{n,ij} = C_t^{n,ij} + (x^i x^j) * \nu_t^n - \sum_{s\leq t} \Delta B_s'^{n,i} \Delta B_s'^{n,j}. \end{cases}$$

Similarly, we assume that $|x|^2 * \nu_t < \infty$ for all t, and we define B' and $\tilde{C}'$ by 3.25, with (B, C, ν).

As said before, the choice of $C_1(\mathbb{R}^d)$ is arbitrary, as long as it meets VII.2.7. So we assume that

3.26 $C_1(\mathbb{R}^d)$ contains the positive and negative parts of the following functions (where $a \in \mathbb{Q}_+$, recall that $g_a(x) = (a|x| - 1)^+ \wedge 1$):

$$\begin{cases} g_a^i(x) = (x^i - h^i(x))(1 - g_a(x)) \\ g_a^{ij}(x) = (x^i x^j - h^i(x)h^j(x))(1 - g_a(x)). \end{cases} \qquad \square$$

3.27 **Theorem.** *Assume that $C_1(\mathbb{R}^d)$ meets 3.26, and let D be a dense subset of $\mathbb{R}_+$. Assume also that v^n and v satisfy 3.24, and*

(i) *The strong majoration hypothesis* (III): There is a *continuous* and *deterministic* increasing càdlàg function F which strongly majorizes the functions $\sum_{i \le d} \operatorname{Var}(B'^i(\alpha))$ and $\sum_{i \le d} \tilde{C}'^{ii}(\alpha)$.

(ii) *The condition on big jumps*: for all $t \in \mathbb{R}_+$,

3.28 $$\lim_{a \uparrow \infty} \sup_{\alpha \in \Omega} |x|^2 1_{\{|x|>a\}} * v_t(\alpha) = 0.$$

(iii) *The uniqueness condition 3.21 (iii).*

(iv) *The continuity condition*: for all $t \in D$, $g \in C_1(\mathbb{R}^d)$ the functions $\alpha \rightsquigarrow B_t'(\omega)$, $\tilde{C}_t'(\omega)$, $g * v_t(\omega)$ are Skorokhod continuous on $\mathbb{D}(\mathbb{R}^d)$.

(v) *$\eta^n \to \eta$ weakly.*

(vi) *$[\delta_{7,1}\text{-}D]$ and the following three conditions hold*:

$$[\text{Sup-}\beta_7'] \quad \sup_{s \le t} |B_s'^n - B_s' \circ X^n| \xrightarrow{P} 0 \quad \text{for all } t \ge 0.$$

$$[\gamma_7'\text{-}D] \quad \tilde{C}_t'^n - \tilde{C}_t' \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D,$$

3.29 $$\lim_{a \uparrow \infty} \limsup_n P^n(|x|^2 1_{\{|x|>a\}} * v_t^n > \varepsilon) = 0 \quad \text{for all } t \in \mathbb{R}_+, \varepsilon > 0.$$

Then the laws $\mathscr{L}(X^n)$ weakly converge to P.

Proof. We will prove that all conditions of 3.21 are met. This is clear for 3.21 (iii, v), and 3.21 (ii) easily follows from 3.28.

Since $|x - h(x)| \le A|x|^2$ for some constant A, 3.25 yields $\operatorname{Var}(B^i) \prec \operatorname{Var}(B'^i) + A|x|^2 * v$, and so 3.21 (i) easily follows from (i).

We have $|x - h(x)| g_a(x) \le A|x|^2 1_{\{|x|>1/a\}}$ and $|x - h(x)|^2 g_a(x) \le A|x|^2 1_{\{|x|>1/a\}}$ for a constant A. Therefore 3.25 and the fact that B and B' are continuous in t (because F is so) yield, with notation 3.26:

3.30 $$\left.\begin{aligned} B^i &= B'^i - g_a^i * v + R^{i,a} \\ \tilde{C}^{ij} &= \tilde{C}'^{ij} - g_a^{ij} * v + R^{ij,a} \end{aligned}\right\} \quad \text{with } |R^{i,a}|, |R^{ij,a}| \le A|x|^2 1_{\{|x|>1/a\}} * v.$$

Hence 3.28 and (iv) easily yield that for $t \in D$, the functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\alpha)$ are Skorokhod-continuous, and 3.21 (iv) is fulfilled.

In view of 3.25, we also have 3.30 for B^n and B'^n, while by Lemma 3.8, [Sup-$\delta_{7,1}$] holds. Thus it readily follows from [Sup-β'_7] and from 3.28 and 3.29 that [Sup-β_7] holds as well.

Usually $\tilde{C}^n$ and $\tilde{C}'^n$ do not satisfy 3.30, because B^n and B'^n are discontinuous. However,

3.31
$$\begin{cases} \tilde{C}^{n,ij} = \tilde{C}'^{n,ij} - g_a^{ij} * \nu^n + R^{n,ij,a} + \sum_{s \le \cdot} \gamma_s^{n,ij}, \quad \text{where} \\ |R^{n,ij,a}| \le A|x|^2 1_{\{|x|>1/a\}} * \nu^n \\ \gamma_s^{n,ij} = \Delta B_s'^{n,i} \Delta B_s'^{n,j} - \Delta B_s^{n,i} \Delta B_s^{n,j}. \end{cases}$$

Assume for a while that

3.32
$$\sum_{s \le t} |\gamma_s^{n,ij}| \xrightarrow{P} 0 \quad \text{as } n \uparrow \infty, \text{for all } t \ge 0.$$

Then 3.30 and 3.31, plus 3.28, 3.29, [γ'_7-D] and [$\delta_{7,1}$-D], readily imply [γ_7-D]: thus we have 3.21(vi), and Theorem 3.21 gives the result.

It remains to prove 3.32. We have

3.33
$$\begin{aligned} \sum_{s \le t} |\gamma^{n,ij}| &\le \sum_{s \le t} \{|\Delta B_s^{n,i} - \Delta B_s'^{n,i}||\Delta B_s'^{n,j}| + |\Delta B_s^{n,j} - \Delta B_s'^{n,j}||\Delta B_s^{n,i}|\} \\ &\le \left(\sum_{s \le t} |\Delta B_s^{n,i} - \Delta B_s'^{n,i}| \right) \left(\sup_{s \le t} |\Delta B_s'^{n,j}| \right) \\ &\quad + \left(\sum_{s \le t} |\Delta B_s^{n,j} - \Delta B_s'^{n,j}| \right) \left(\sup_{s \le t} |\Delta B_s^{n,i}| \right). \end{aligned}$$

Moreover,

$$\sum_{s \le t} |\Delta B_s^{n,i} - \Delta B_s'^{n,i}| \le |g_a^i| * \nu_t^n + A|x|^2 1_{\{|x|>1/a\}} * \nu_t^n,$$

and $\sup_{s \le t} ||g_a^i| * \nu_s^n - (|g_a^i| * \nu_s) \circ X^n| \xrightarrow{P} 0$ by [Sup-$\delta_{7,1}$], while by (i) there is a constant γ_a such that $|g_a^i| * \nu_s \le \gamma_a F_s$ identically. Therefore, in view of 3.29 we deduce that

3.34
$$\lim_{N \uparrow \infty} \sup_n P^n \left(\sum_{s \le t} |\Delta B_s^{n,i} - \Delta B_s'^{n,i}| > N \right) = 0,$$

and the same holds with j instead of i. On the other hand, [Sup-β_7] implies $\sup_{s \le t} |\Delta B_s^{n,i} - \Delta B_s^i \circ X^n| \xrightarrow{P} 0$ and $\Delta B_s^i = 0$, so $\sup_{s \le t} |\Delta B_s^{n,i}| \xrightarrow{P} 0$, and similarly $\sup_{s \le t} |\Delta B_s'^{n,j}| \xrightarrow{P} 0$ because of [Sup-β'_7]. This, together with 3.33 and 3.34, readily implies 3.32, and we are finished. □

3. Now we turn to the general case.

3.35 **Theorem.** *Assume the following*:

(i) *The strong majoration hypothesis* (II): *see 3.11.*

(ii) *The condition on the big jumps 3.7.*

(iii) *Uniqueness* (the same as 3.21 iii): there is a unique probability measure P such that X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν) and initial distribution η.

(iv) *Continuity condition*: If D is a dense subset of $\mathbb{R}_+$ contained in the set of times t such that $\Delta F_t = 0$ and $\Delta F_t^g = 0$ for all $g \in C_1(\mathbb{R}^d)$ (where F, F^g appear in 3.11: see the comments after 3.13), the functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\alpha)$, $g * \nu_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$ for all $t \in D$, $g \in C_1(\mathbb{R}^d)$.

(v) $\eta^n \to \eta$ *weakly.*

(vi) $[\text{Sk-}\beta_7]$, $[\text{Sk-}\gamma_7]$, $[\text{Sk-}\delta_{7,1}]$ *hold.*

Then the laws $\mathscr{L}(X^n)$ *weakly converge to* P.

Proof. The proof is similar to that of 3.21, except that we use 3.20 instead of 3.9. Firstly, (i), (ii), (v) and (vi) imply that the sequence $\{X^n\}$ is tight (by 3.20), and it remains to prove that P is the only limit point of the sequence $\{\mathscr{L}(X^n)\}$.

We can assume that $\mathscr{L}(X^n) \to P'$ weakly. Let D be as in (iv). Then 3.13 and (vi) imply that 2.11(i) is met. 2.11(ii) easily follows from (i), and 2.11(iii) is exactly (iv). Moreover if $t \in D$ we have $\Delta F_t^g = 0$ for all $g \in C_1(\mathbb{R}^d)$, so Part (b) of the proof of 3.21 works here (recall that $[\delta_{7,1}\text{-}D]$ holds) and we deduce that $P'(\Delta X_t \neq 0) = 0$. In other words, D is contained in the set $\{t: P'(\Delta X_t \neq 0) = 0\}$ and so Theorem 2.11 applies: since (v) immediately yields that $\mathscr{L}(X_0|P') = \eta$, we deduce that $P' = P$ from the uniqueness (iv). □

3.36 **Remark.** The continuity condition (iv) is met in particular when 2.16 holds, i.e. when $\alpha \rightsquigarrow B^i(\alpha)$, $\tilde{C}^{ij}(\alpha)$, $g * \nu(\alpha)$ are Skorokhod-continuous from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R})$. □

3.37 **Remark.** Suppose that the characteristics (B, C, ν) are *deterministic*. Then (i), (ii), (iv) are trivially fulfilled, and (iii) follows from III.2.16. Moreover, with the notation of the previous chapter, we have $[\text{Sk-}\beta_7] = [\text{Sk-}\beta_5]$, $[\text{Sk-}\gamma_7] = [\text{Sk-}\gamma_5]$, $[\text{Sk-}\delta_{7,1}] = [\text{Sk-}\delta_{5,1}]$, and for any dense subset D of $\mathbb{R}_+$ it is proved in VIII.4.3 that

$$[\text{Sk-}\delta_{5,1}] + [\gamma_5\text{-}D] \Rightarrow [\text{Sk-}\gamma_5].$$

Thus, *Theorem VIII.4.1 is a particular case of Theorem 3.35.* □

§ 3c. Limit Theorems: The Locally Bounded Case

For simplicity, we will only consider the case where the limiting process is quasi-left continuous.

For every $a \geq 0$ we define on $\Omega = \mathbb{D}(\mathbb{R}^d)$ and on each Ω^n:

3.38 $$\begin{cases} S_a(\alpha) = \inf(t: |\alpha(t)| \geq a \text{ or } |\alpha(t-)| \geq a) \\ S_a^n = S_a \circ X^n = \inf(t: |X_t^n| \geq a \text{ or } |X_{t-}^n| \geq a). \end{cases}$$

3.39 **Theorem.** *Let D be a dense subset of $\mathbb{R}_+$, and assume*:

(i) *The local strong majoration hypothesis*: for all $a > 0$ there is an increasing *continuous* and *deterministic* function $F(a)$ such that the stopped processes $\sum_{i \le d} \mathrm{Var}(B^i)^{S_a}$ and $(\sum_{i \le d} C^{ii} + (|x|^2 \wedge 1) * \nu)^{S_a}$ are strongly majorized by $F(a)$.

(ii) *The local condition on big jumps*: for all $a > 0$, $t > 0$,

$$\lim_{b \uparrow \infty} \sup_{\alpha \in \Omega} \nu(\alpha; [0, t \wedge S_a(\alpha)] \times \{x: |x| > b\}) = 0.$$

(iii) *Local uniqueness* (see III.2.37) *for the martingale problem* $\mathscr{s}(\sigma(X_0), X | \eta;$ $B, C, \nu)$; *we denote by P the unique solution to this problem.*

(iv) *Continuity condition*: for all $t \in D$, $g \in C_1(\mathbb{R}^d)$ the functions $\alpha \rightsquigarrow B_t(\alpha)$, $\tilde{C}_t(\alpha)$, $g * \nu_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$.

(v) $\eta^n \to \eta$ *weakly.*

(vi) *The following three conditions hold*:

$$[\text{Sup-}\beta_{\text{loc}}] \quad \sup_{s \le t} |B^n_{s \wedge S^n_a} - (B_{s \wedge S_a}) \circ X^n| \xrightarrow{P} 0 \quad \text{for all } t > 0, a > 0;$$

$$[\gamma_{\text{loc}}\text{-}D] \quad \tilde{C}^n_{t \wedge S^n_a} - (\tilde{C}_{t \wedge S_a}) \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D, a > 0;$$

$$[\delta_{\text{loc}}\text{-}D] \quad g * \nu^n_{t \wedge S^n_a} - (g * \nu_{t \wedge S_a}) \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D, a > 0, g \in C_1(\mathbb{R}^d).$$

Then the laws $\mathscr{L}(X^n)$ weakly converge to P.

3.40 **Remark.** With the (important) exception of local uniqueness, the conditions of 3.39 are weaker than those of 3.21: this is obvious for (i) and (ii), and (iv), (v) are the same. Moreover 3.6 and 3.21(vi) yield [Sup-β_7], [Sup-γ_7], [Sup-$\delta_{7,1}$]; since $S^n_a = S_a \circ X^n$, the three conditions of (vi) above easily follow. □

Proof. a) To simplify the notation, for every $a > 0$ we write $X^n(a) = (X^n)^{S^n_a}$, $B^n(a) = (B^n)^{S^n_a}$, $C^n(a) = (C^n)^{S^n_a}$, $\tilde{C}^n(a) = (\tilde{C}^n)^{S^n_a}$, $\nu^n(a) = (\nu^n)^{S^n_a}$ for the processes and random measures stopped at time S^n_a, and similarly we write $B(a) = B^{S_a}$, $C(a) = C^{S_a}$, $\tilde{C}(a) = \tilde{C}^{S_a}$, $\nu(a) = \nu^{S_a}$.

$X^n(a)$ is a semimartingale on $\mathscr{B}^n$ with characteristics $(B^n(a), C^n(a), \nu^n(a))$. On the other hand, (i) and (ii) imply that the triplet $(B(a), C(a), \nu(a))$ satisfies 3.6 and 3.7. Furthermore II.2.43d yields that $B_{t \wedge S_a} \circ X^n = B_{t \wedge S_a} \circ X^n(a)$, and similarly for $\tilde{C}$ and ν. Hence (vi) gives

3.41
$$\begin{cases} \sup_{s \le t} |B^n(a)_s - B(a)_s \circ X^n(a)| \xrightarrow{P} 0 \quad \text{for all } t \ge 0 \\ \tilde{C}^n(a)_t - \tilde{C}(a)_t \circ X^n(a) \xrightarrow{P} 0 \quad \text{for all } t \ge 0 \\ g * \nu^n(a)_t - (g * \nu(a)_t) \circ X^n(a) \xrightarrow{P} 0 \quad \text{for all } t \ge 0, g \in C_1(\mathbb{R}^d) \end{cases}$$

(we have convergence for all $t \ge 0$, and even local uniform convergence, in the two last lines of 3.41 because of Lemma 3.8). Then these properties, plus (v), allow to deduce from Theorem 3.9 that for every $a > 0$ the sequence $\{X^n(a)\}_{n \in \mathbb{N}^*}$ is tight.

b) The second step consists in proving:

3.42 The functions $\alpha \rightsquigarrow B(a)_t(\alpha)$, $\tilde{C}(a)_t(\alpha)$, $g * \nu(a)_t(\alpha)$ are continuous for all $t \geq 0$, $g \in C_1(\mathbb{R}^d)$, at each point of continuity of the function $\alpha \rightsquigarrow S_a(\alpha)$.

We will only prove the claim for $B(a)$, the others being similar. Let $t \geq 0$ and α be a continuity point of $S_a(\cdot)$. There clearly exists $b \geq a$ such that α is also a continuity point of $S_b(\cdot)$, and $S_b(\alpha) > t$. Let $\alpha_n \to \alpha$ and $\varepsilon > 0$. We have $S_a(\alpha_n) \to S_a(\alpha)$ and $S_b(\alpha_n) \to S_b(\alpha)$, so $S_b(\alpha_n) > t$ for all n large enough. There exist $s \in D$, $\eta > 0$ with $s < t \wedge S_a(\alpha) < s + \eta$ and $F(b)_{s+\eta} - F(b)_s \leq \varepsilon$. Then (i) yields $|B(a)_{t \wedge S_a}(\alpha) - B_s(\alpha)| \leq \varepsilon$ and $|B(a)_{t \wedge S_a}(\alpha_n) - B_s(\alpha_n)| \leq \varepsilon$ provided $t \wedge S_a(\alpha_n) \in (s, s+\eta)$ and $S_b(\alpha_n) > t$, which are true for all n large enough. Moreover (iv) yields $B_s(\alpha_n) \to B_s(\alpha)$: since $\varepsilon > 0$ is arbitrary we deduce the claim.

c) Consider a subsequence of $\{X^n\}$. In virtue of (*a*), there is a further subsequence, still denoted by (n), such that for each $p \in \mathbb{N}^*$ the sequence $\mathscr{L}(X^n(p))$ weakly converges to a limit Q^p as $n \uparrow \infty$.

We introduce the sets $V(\alpha)$, $V'(\alpha)$ as in VI.2.10 (or as in the proof of 1.17). We have seen in the proof of 1.17 that there are only countably many $a > 0$ such that $Q^p(\alpha: a \in V(\alpha) \cup V'(\alpha)) > 0$. So we choose $a_p \in [p-1, p]$ such that $Q^p(\alpha: a_p \in V(\alpha) \cup V'(\alpha)) = 0$, and due to VI.2.11 and VI.2.12 and to the convergence $\mathscr{L}(X^n(p)) \to Q^p$, we obtain

3.43
$$\begin{cases} \alpha \rightsquigarrow S_{a_p}(\alpha) & \text{is } Q^p\text{-a.s. continuous,} \\ \mathscr{L}(X^n(p), X^n(a_p)) \to \mathscr{L}[(X, X^{S_{a_p}}) | Q^p]. \end{cases}$$

In particular, $S_{a_p} \circ X^{S_{a_p}} = S_{a_p}$ is Q^p-a.s. continuous. Hence if we denote by $\tilde{Q}^p$ the law of $X^{S_{a_p}}$ on the space $(\Omega, \mathscr{F}, Q^p)$, we deduce from 3.43 that

3.44
$$\begin{cases} \alpha \rightsquigarrow S_{a_p}(\alpha) & \text{is } \tilde{Q}^p\text{-a.s. continuous,} \\ \mathscr{L}(X^n(a_p)) \to \tilde{Q}^p. \end{cases}$$

d) Next, we will apply Theorem 2.11 to the sequence $\{X^n(a_p)\}_{n \in \mathbb{N}^*}$ and the measure $\tilde{Q}^p$ (instead of P in 2.11), and the triplet $(B(a_p), C(a_p), \nu(a_p))$ instead of (B, C, ν). 2.11(ii) is a simple consequence of (i), and 2.11(iii) follows from 3.42 and 3.44. Furthermore 2.11(i), with the set $\{t > 0: \tilde{Q}^p(\Delta X_t \neq 0) = 0\}$ instead of D, is nothing else than 3.41. Hence by 2.11, X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, \tilde{Q}^p)$ with characteristics $(B(a_p), C(a_p), \nu(a_p))$, and $\mathscr{L}(X_0 | \tilde{Q}^p) = \eta$ is obvious from (v). Since $X^{S_{a_p}} = X$ $\tilde{Q}^p$-a.s. (because the characteristics are constant after S_{a_p}), we deduce $\tilde{Q}^p \in \mathscr{s}(\sigma(X_0), X^{S_{a_p}} | \eta; B(a_p), C(a_p), \nu(a_p))$.

By VI.2.10, S_{a_p} is a strict stopping time, relative to $\mathscr{F}_t^0 = \mathscr{D}_t^0(\mathbb{R}^d)$. Hence the local uniqueness (iii) implies

3.45
$$\tilde{Q}^p = P \text{ in restriction to } (\Omega, \mathscr{F}^0_{S_{a_p}}).$$

e) In this step we will prove that the sequence $\{X^n\}$ is tight, for which we will use Theorem VI.3.21. Let $N \in \mathbb{N}^*$, $\varepsilon > 0$, $\eta > 0$. There exists $p \in \mathbb{N}^*$ such that $P(S_{a_p} \leq N + 1) \leq \varepsilon$ (recall that $a_p \in [p-1, p]$), and 3.45 implies that $\tilde{Q}^p(S_{a_p} \leq N + 1) \leq \varepsilon$ (recall that S_{a_p} is $\mathscr{F}^0_{S_{a_p}}$-measurable). Since $S_{a_p} \circ X^n = S_{a_p} \circ X^n(a_p)$, 3.44 yields

3.46 $$n \geq n_0 \Rightarrow P^n(S^n_{a_p} \leq N) \leq 2\varepsilon$$

for some $n_0 \in \mathbb{N}^*$. Moreover 3.44 also implies tightness of the sequence $\{X^n(a_p)\}_{n \in \mathbb{N}^*}$, so by VI.3.21 there exist $K > 0$, $\theta > 0$, $n'_0 \geq n_0$ with

3.47 $$n \geq n'_0 \Rightarrow \begin{cases} P^n\left(\sup_{t \leq N} |X^n(a_p)_t| > K\right) \leq \varepsilon \\ P^n(w'_N(X^n(a_p), \theta) > \eta) \leq \varepsilon. \end{cases}$$

Now, $\sup_{t \leq N} |X^n(a_p)_t| = \sup_{t \leq N} |X^n_t|$ and $w'_N(X^n(a_p), \theta) = w'_N(X^n, \theta)$ on the set $\{S^n_{a_p} \geq N\}$. Therefore 3.46 and 3.47 give

$$n \geq n'_0 \Rightarrow \begin{cases} P^n\left(\sup_{t \leq N} |X^n_t| > K\right) \leq 3\varepsilon \\ P^n(w'_N(X^n, \theta) > \eta) \leq 3\varepsilon, \end{cases}$$

and another application of VI.3.21 yields tightness for $\{X^n\}$.

f) Upon taking a subsequence, still denoted by (n), we can then assume that $\mathscr{L}(X^n) \to P'$ weakly for some measure P'. Let ψ be an $\mathscr{F}_t$-measurable, continuous and bounded function on $\mathbb{D}(\mathbb{R}^d)$, with $|\psi| \leq 1$. Then $E^n(\psi(X^n)) \to E_{P'}(\psi)$. Moreover 3.44 yields $E^n[\psi(X^n(a_p))] \to E_{\tilde{Q}^p}(\psi)$. But 3.45 and the definition of $X^n(a_p)$ give

$$|E_{\tilde{Q}^p}(\psi) - E_P(\psi)| \leq 2P(S_{a_p} \leq t)$$

$$|E^n(\psi(X^n)) - E^n[\psi(X^n(a_p))]| \leq 2P^n(S^n_{a_p} \leq t)$$

and thus

$$|E_{P'}(\psi) - E_P(\psi)| \leq 2P(S_{a_p} \leq t) + 2 \limsup_n P^n(S^n_{a_p} \leq t).$$

But $\lim_{p \uparrow \infty} P(S_{a_p} \leq t) = 0$, and 3.46 (in which N and ε are arbitrary) implies that $\lim_{p \uparrow \infty} \limsup_n P^n(S^n_{a_p} \leq t) = 0$: hence $E_{P'}(\psi) = E_P(\psi)$ for all $\mathscr{F}_t$-measurable continuous and bounded functions ψ. Since $t \geq 0$ is arbitrary, we deduce from VI.1.14 that $P' = P$.

g) At this stage, we have proved that any subsequence (n') contains a further subsequence (n'') such that $\mathscr{L}(X^{n''}) \to P$. This is enough to insure that $\mathscr{L}(X^n) \to P$. □

Finally, we state the square-integrable version: recall the notation 3.25.

3.48 **Theorem.** *Assume that $C_1(\mathbb{R}^d)$ meets 3.26 and that ν^n and ν satisfy 3.24. Let D be a dense subset of $\mathbb{R}_+$, and assume*:

(i) *The local strong majoration hypothesis*: for all $a \geq 0$ there is an increasing *continuous* and *deterministic* function $F(a)$ such that the stopped processes $\sum_{i \leq d} \mathrm{Var}(B'^i)^{S_a}$ and $(C^{ii})^{S_a}$ and $(|x|^2 * \nu)^{S_a}$ are strongly majorized by $F(a)$.

(ii) *The local condition on big jumps*: for all $a \geq 0$, $t \geq 0$,

$$\lim_{b \uparrow \infty} \sup_{\alpha \in \Omega} |x|^2 1_{\{|x| > b\}} * \nu_{t \wedge S_a}(\alpha) = 0.$$

(iii) *Local uniqueness* (see III.2.37) *for the martingale problem* $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$; *we denote by P the unique solution to this problem.*

(iv) *Continuity condition*: for all $t \in D$, $g \in C_1(\mathbb{R}^d)$ the functions $\alpha \rightsquigarrow B'_t(\alpha)$, $\tilde{C}'_t(\alpha)$, $g * \nu_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$.

(v) $\eta^n \to \eta$ *weakly.*

(vi) $[\delta_{\text{loc}}\text{-}D]$ *and the following three conditions hold*:

$$[\text{Sup-}\beta'_{\text{loc}}] \quad \sup_{s \le t} |B'^n_{t \wedge S^n_a} - (B'_{t \wedge S_a}) \circ X^n| \xrightarrow{P} 0 \quad \text{for all } t > 0, a > 0;$$

$$[\gamma'_{\text{loc}}\text{-}D] \quad \tilde{C}'^n_{t \wedge S^n_a} - (\tilde{C}'_{t \wedge S_a}) \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \in D, a > 0;$$

3.49 $$\lim_{b \uparrow \infty} \limsup_n P^n(|x|^2 1_{\{|x|>b\}} * \nu^n_{t \wedge S^n_a} > \varepsilon) = 0 \quad \text{for all } t > 0, a > 0, \varepsilon > 0.$$

Then the laws $\mathscr{L}(X^n)$ *weakly converge to P.*

Proof. It suffices to reproduce the proof of Theorem 3.27, to the effect that the above hypotheses imply all hypotheses of Theorem 3.39. □

3.50 **Remark.** Upon a close examination of the proof of 3.39, one sees that in $[\text{Sup-}\beta_{\text{loc}}]$, $[\gamma_{\text{loc}}\text{-}D]$, $[\delta_{\text{loc}}\text{-}D]$, and also in $[\text{Sup-}\beta'_{\text{loc}}]$ and $[\gamma'_{\text{loc}}\text{-}D]$, it would be enough to assume the properties for all $a \in A$, where A is a subset of $\mathbb{R}_+$ with a countable complement. □

3.51 **Remark.** The localization of Theorem 3.35 does not seem so easy. See however Pagès [192] for some results in this direction. □

4. Applications

§ 4a. Convergence of Diffusion Processes with Jumps

We suppose that $(\Omega, \mathscr{F}, \mathbf{F})$ is the canonical space, with the canonical process X: see 2.6. We want X to be a *diffusion process with jumps*, in the sense of III.2.18, and for simplicity we only consider the homogeneous case. That is,

4.1 $$\begin{cases} B_t = \int_0^t b(X_s)\,ds, & b \text{ Borel function: } \mathbb{R}^d \to \mathbb{R}^d \\ C_t = \int_0^t c(X_s)\,ds, & c \text{ Borel function: } \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d \text{ with values in} \\ & \text{the set of symmetric nonnegative matrices,} \\ \nu(dt, dx) = dt K(X_t, dx), & K \text{ Borel kernel from } \mathbb{R}^d \text{ into itself, with} \\ & \int K(x, dy)(|y|^2 \wedge 1) < \infty, \end{cases}$$

and so we have

4.2 $$\tilde{C}_t = \int_0^t \tilde{c}(X_s)\,ds, \quad \text{with } \tilde{c}^{ij}(x) = c^{ij}(x) + \int K(x,dy)h^i(x)h^j(x).$$

We will assume the following hypothesis (conditions for it to be met can be found in § III.2c):

4.3 *Uniqueness-measurability hypothesis*: (i) for each $x \in \mathbb{R}^d$ the martingale problem $\mathscr{s}(\sigma(X_0), X|\varepsilon_x; B, C, \nu)$ has a unique solution P_x;
(ii) $x \rightsquigarrow P_x(A)$ is Borel for all $A \in \mathscr{F}$. □

4.4 **Lemma.** *Under 4.3, local uniqueness holds for the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$, for every initial distribution η on $\mathbb{R}^d$.*

Proof. a) We suppose first that $\eta = \varepsilon_x$ for some $x \in \mathbb{R}^d$. We have natural shifts θ_t defined on Ω by $\alpha \circ \theta_t(s) = \alpha(t+s)$, and in view of 4.1,

$$B_s \circ \theta_t = B_{s+t} - B_t, \quad C_s \circ \theta_t = C_{s+t} - C_t,$$
$$\nu(\theta_t\omega; [0,s] \times A) = \nu(\omega; [t,t+s] \times A).$$

That is, we have III.2.39 with $p_t B = B$ and $p_t C = C$ and $p_t \nu = \nu$. Then Theorem III.2.40 and 4.3 obviously yield the result.

b) Now let η be any probability measure on $\mathbb{R}^d$, and let P', P'' be two solutions to the stopped problem $\mathscr{s}(\sigma(X_0), X^T|\eta; B^T, C^T, \nu^T)$, where T is a strict stopping time. If $\mathscr{H} = \sigma(X_0)$, Hypothesis II.6.4 is met by P' and P'', and there are transition measures $P'_x(d\omega)$ and $P''_x(d\omega)$ from $\mathbb{R}^d$ into $(\Omega, \mathscr{F})$, such that $P'_{X_0(\omega)}(d\omega')$ and $P''_{X_0(\omega)}(d\omega')$ are regular versions of the P'- and P''-conditional distributions with respect to $\mathscr{H}$.

We can then apply II.6.15: for η-almost all x, P'_x and P''_x are also solution to $\mathscr{s}(\sigma(X_0), X^T|\varepsilon_x; B^T, C^T, \nu^T)$. For these x's (a) implies that $P'_x = P''_x$ on $(\Omega, \mathscr{F}^0_T)$; since $P'(A) = \int \eta(dx) P'_x(A)$ and similarly for P'' by definition of the conditional distributions, we deduce $P' = P''$ on $(\Omega, \mathscr{F}^0_T)$, and the claim is proved. □

4.5 **Remark.** Under 4.3, and with the shifts θ_t introduced above, it can be easily shown that $(\Omega, \mathscr{F}, \mathscr{F}_t, \theta_t, X_t, P_x)$ is a Markov process in the sense of Blumenthal and Getoor [14].

Moreover if f is a function of class C^2 on $\mathbb{R}^d$, and if we set

4.6 $$\begin{aligned} Af(x) = {} & \sum_{i \le d} b^i(x) D_i f(x) + \frac{1}{2} \sum_{i,j \le d} b^{ij}(x) D_{ij} f(x) \\ & + \int K(x,dy)\Big[f(x+y) - f(x) - \sum_{i \le d} h^i(x) D_i f(x)\Big], \end{aligned}$$

II.2.42 implies that $f(X_t) - f(X_0) - \int_0^t Af(X_s)\,ds$ is a local martingale for all measures P_x (and even all measures $P_\eta = \int \eta(dx)P_x$). Hence A is the *extended generator* of the Markov process (see e.g. [98] for more details). $\square$

As for the X^n's, we will also assume that they have the same form, that is their characteristics are

4.7
$$B_t^n = \int_0^t b^n(X_s^n)\,ds, \quad C_t^n = \int_0^t c^n(X_s^n)\,ds, \quad \nu^n(dt,dx) = dt\,K^n(X_t^n, dx)$$
$$\tilde{C}_t^n = \int_0^t \tilde{c}^n(X_s^n)\,ds$$

with b^n, c^n, K^n as in 4.1, and $\tilde{c}^n$ as in 4.2. However we do not assume that 4.3 holds for X^n, which thus is not necessarily a Markov process.

4.8 **Theorem.** *Assume that (b,c,K) satisfy 4.3 and*

4.9
$$\lim_{b\uparrow\infty} \sup_{x:|x|\le a} K(x,\{y;|y|>b\}) = 0 \quad \text{for all } a>0;$$

4.10 $x \rightsquigarrow b(x), \tilde{c}(x), \int K(x,dy)g(y)$ are continuous on $\mathbb{R}^d$, for $g \in C_1(\mathbb{R}^d)$.

Assume also the following:

4.11
$$b^n \to b, \quad \tilde{c}^n \to \tilde{c}, \quad \int K^n(\cdot,dy)g(y) \to \int K(\cdot,dy)g(y)$$
locally uniformly on $\mathbb{R}^d$, for all $g \in C_1(\mathbb{R}^d)$;

4.12 $\eta^n \to \eta$ weakly, where $\eta^n = \mathscr{L}(X_0^n | P^n)$ is the initial distribution of X_0^n.

Then the laws $\mathscr{L}(X^n)$ weakly converge to $P = \int \eta(dx)P_x$.

Proof. We will check that all hypotheses of 3.39 are fulfilled.

Firstly, there exists a function $g \in C_1(\mathbb{R}^d)$ such that $|x|^2 \wedge 1 \le \sum_{i\le d} |h^i(x)|^2 + g(x)$, so 4.10 yields that $x \rightsquigarrow b(x)$ and $x \rightsquigarrow \sum_{i\le d} c^{ii}(x) + \int K(x,dy)\,(|y|^2 \wedge 1)$ are locally bounded. Hence, by 4.1, we see that 3.39(i) holds. 4.9 obviously implies 3.39(ii), and 3.39(iii) follows from Lemma 4.4. 3.39(iv) is clearly implied by 4.10, and 3.39(v) is nothing else than 4.12.

Finally, we have

$$B_{t\wedge S_a^n}^n - (B_{t\wedge S_a}) \circ X^n = \int_0^{t\wedge S_a^n} [b^n(X_s^n) - b(X_s^n)]\,ds$$

and similarly for $\tilde{C}^n$ and $g * \nu^n$. Then we readily deduce 3.39(vi) from 4.11, and we are finished. $\square$

4.13 **Remark.** Define A^n by 4.6, with b^n, c^n, K^n; then if X^n is a Markov process, A^n is its extended generator, and 4.11 is equivalent to:

4.14 $$A^n f \to Af \text{ locally uniformly for all } C^2 \text{ functions } f \text{ on } \mathbb{R}^d.$$

So we have got our promised extension of Trotter-Kato Theorem. □

Next, we give the square-integrable version. We assume that $K(x, \cdot)$ and $K^n(x, \cdot)$ integrate $|y|^2$, so we can set

$$b'^i(x) = b^i(x) + \int K(x,dy)(y^i - h^i(y))$$

$$\tilde{c}'^{ij}(x) = c^{ij}(x) + \int K(x,dy)y^i y^j,$$

and similarly for $b'^n = (b'^{n,i})_{i \le d}$ and $\tilde{c}'^n = (\tilde{c}'^{n,ij})_{i,j \le d}$.

4.15 **Theorem.** *Assume that (b,c,K) satisfies 4.3 and*

4.16 $$\lim_{b \uparrow \infty} \sup_{x:|x| \le a} \int K(x,dy)|y|^2 1_{\{|y|>b\}} = 0 \quad \text{for all } a > 0;$$

4.17 $$x \rightsquigarrow b'(x), \quad \tilde{c}'(x), \quad \int K(x,dy)g(y) \text{ are continuous on } \mathbb{R}^d, \quad \text{for } g \in C_1(\mathbb{R}^d)$$

Assume also that K^n integrates $|y|^2$, and that 4.12 and

4.18 $$b'^n \to b', \quad \tilde{c}'^n \to \tilde{c}', \quad \int K^n(\cdot,dy)g(y) \to \int K(\cdot,dy)g(y)$$
$$\text{locally uniformly on } \mathbb{R}^d \text{ for all } g \in C_1(\mathbb{R}^d)$$

hold. Then the laws $\mathscr{L}(X^n)$ weakly converge to $P = \int \eta(dx)P_x$.

Proof. It is the same as for 4.8, except that we use Theorem 3.48 instead of Theorem 3.39. □

§ 4b. Convergence of Step Markov Processes to Diffusions

Here we give one example, among many, of convergence of step (or, pure-jump) processes to a continuous diffusion.

The setting is as in § 4a, and in addition we suppose that each X^n is a pure step Markov process: this means that its generator has the form

4.19 $$A^n f(x) = \int K^n(x,dy)[f(x+y) - f(x)]$$

where K^n is a *finite* transition measure on $\mathbb{R}^d$. If we plug 4.19 into 4.6, we obtain for the other coefficients b^n and c^n:

$$b^n(x) = \int K^n(x, dy)h(y), \quad c^n(x) = 0, \quad \tilde{c}^{n,ij}(x) = \int K^n(x, dy)h^i(x)h^j(x).$$

In order to slightly simplify the matter, we will assume that all $K^n(x, \cdot)$ integrate $|y|^2$, so instead of b^n and $\tilde{c}^n$ we consider:

4.20 $$b'^n(x) = \int K^n(x, dy)y, \quad \tilde{c}'^{n,ij}(x) = \int K^n(x, dy)y^i y^j,$$

and 4.15 gives:

4.21 **Theorem.** *Assume that 4.3 holds, with $K = 0$ (so the limiting process will be a continuous diffusion process), and that $b = b'$ and $\tilde{c} = \tilde{c}' = c$ are continuous functions on $\mathbb{R}^d$. Assume also that X^n is as described above, and that*
(i) *$b'^n \to b$, $\tilde{c}'^n \to c$ locally uniformly;*
(ii) $\sup_{x:|x|\le a} \int K^n(x, dy)|y|^2 1_{\{|y|>\varepsilon\}} \to 0$ *as $n \uparrow \infty$, for all $\varepsilon > 0$;*
(iii) *$\eta^n \to \eta$ weakly on $\mathbb{R}^d$.*
Then the laws $\mathscr{L}(X^n)$ weakly converge to $P = \int \eta(dx)P_x$, the law of the diffusion process with coefficients b and c and initial distribution η.

Now we specialize this result. Firstly we assume

4.22 $$\eta^n = \eta = \varepsilon_x \quad \text{for some } x \in \mathbb{R}^d.$$

Secondly we assume 4.21(ii) and

4.23 $$\begin{cases} b'^n \to b & \text{locally uniformly, where } b \text{ is a Lipschitz function;} \\ \tilde{c}'^n \to 0 & \text{locally uniformly.} \end{cases}$$

Then all the hypotheses of 4.21 are met. Furthermore, P_x is then the "law" of the deterministic diffusion $dX_t = b(X_t)\,dt$. In other words, if $x_t(x)$ denotes the unique solution of the following ordinary d-dimensional differential equation

4.24 $$dx_t(x) = b(x_t(x))\,dt, \quad x_0(x) = x,$$

and since the Skorokhod convergence coincides with the local uniform convergence when the limit is continuous, we deduce

4.25 $$\sup_{s\le t} |X_s^n - x_s(x)| \xrightarrow{P} 0 \quad \text{for all } t \ge 0.$$

This is a rather simple-minded result. But now we can evaluate the rate of convergence in 4.25, via another application of Theorem 4.21:

4.26 **Theorem.** *Assume 4.22 and 4.23. Let (a_n) be a sequence of positive numbers converging to $+\infty$ and such that*

(i) $a_n^2 \tilde{c}'^n$ *converges locally uniformly to a continuous function* $\hat{c}$;

(ii) $\lim_n \sup_{x:|x|\leq a} a_n^2 \int K^n(x, dy)|y|^2 1_{\{|y|>\varepsilon/a_n\}} = 0$ *for all* $a, \varepsilon > 0$.

Then the processes

4.27 $$Y_t^n = a_n(X_t^n - X_0^n - \int_0^t b'^n(X_s^n)\,ds)$$

converge in law to a continuous PII Y *with characteristics* $(0, \hat{C}(x), 0)$, *where* $\hat{C}(x)_t = \int_0^t \hat{c}(x_s(x))\,ds$ (so Y is also a Gaussian martingale).

Proof. Note that (ii) implies 4.21 (ii), so 4.25 holds. Y^n is a locally square-integrable semimartingale on $\mathscr{B}^n$, with $Y_0^n = 0$, and its characteristics B'^{Y^n}, C^{Y^n}, ν^{Y^n} and modified second characteristic $\tilde{C}'^{Y^n}$, relative to the "truncation function" $h(x) = x$, obviously are

$$B'^{Y^n} = 0, \quad C^{Y^n} = 0,$$

$$g * \nu_t^{Y^n} = \int_0^t ds \int K^n(X_s^n, dy) g(a_n y)$$

$$\tilde{C}_t'^{Y^n, ij} = a_n^2 \int_0^t ds \int K^n(X_s^n, dy) y^i y^j = a_n^2 \int_0^t \tilde{c}'^{n, ij}(X_s^n)\,ds.$$

Let also Y be the PII with characteristics $(0, \hat{C}(x), 0)$. We will apply Theorem VIII.2.18: firstly [Sup-β_5'] is trivially met. (ii) obviously implies VIII.2.15 and [$\delta_{5,1}$-$\mathbb{R}_+$] (recall that $\nu = 0$ here, and that $g(x) = 0$ for all $|x|$ small enough, for $g \in C_1(\mathbb{R}^d)$). Moreover,

$$\tilde{C}_t'^{Y^n} - \hat{C}(x)_t = \int_0^t [a_n^2 \tilde{c}^n(X_s^n) - \hat{c}(X_s^n)]\,ds + \int_0^t [\hat{c}(X_s^n) - \hat{c}(x_s(x))]\,ds.$$

The first term in the right-hand side above converges in measure to 0, because of (i) and of 4.25, which implies $\lim_{A\uparrow\infty} \limsup_n P^n(\sup_{s\leq t}|X_s^n| > A) = 0$. The second term also goes to 0, because of 4.25 and because $\hat{c}$ is continuous. Thus [γ_5'-$\mathbb{R}_+$] holds, and we are finished. □

4.28 **Corollary.** *In addition to the assumptions of 4.26, assume that* $a_n(b'^n - b) \to 0$ *locally uniformly. Then the processes* $a_n(X_t^n - x_t(x))$ *also converge in law to the same PII* Y *with characteristics* $(0, \hat{C}(x), 0)$.

This gives a rate of convergence in 4.25.

Proof. We have

$$Y_t^n = a_n(X_t^n - x_t(x)) + \int_0^t a_n[b(x_s(x)) - b(X_s^n)]\,ds + \int_0^t a_n[b(X_s^n) - b'^n(X_s^n)]\,ds$$

and the same argument than in 4.26 (using 4.25) shows that the last two terms above go to 0 in measure locally uniformly in t as $n \uparrow \infty$. Since $Y^n \xrightarrow{\mathscr{L}} Y$ by 4.26, we deduce the claim. □

§ 4c. Empirical Distributions and Brownian Bridge

As another corollary of our main result, we prove here the well-known convergence of normalized empirical distributions to the Brownian bridge.

Let $(Z_i)_{i\in\mathbb{N}^*}$ be i.i.d. random variables, *uniformly distributed* on $(0, 1]$. According to II.3.31 and II.3.35, we set for $t \in [0, 1]$:

4.29 $$X_t^n = \frac{1}{n} \sum_{1 \le i \le n} 1_{\{Z_i \le t\}}, \quad V_t^n = \sqrt{n}(X_t^n - t).$$

Note that these processes are naturally indexed by $[0, 1]$ only, and $V_0^n = V_1^n = 0$ by construction.

There are several characterizations of the Brownian bridge, the simplest one consisting in saying that it is a centered Gaussian process with covariance function $C(s, t) = s(1 - t)$ for $0 \le s \le t \le 1$.

We will use here another characterization: the (standard) Brownian bridge is the 1-dimensional non-homogeneous diffusion process indexed by $[0, 1]$, which starts at 0 and has the following coefficients (see III.2.18):

4.30 $$b(s, x) = \frac{-x}{1 - s}, \quad c(s, x) = 1.$$

4.31 **Theorem.** *The processes V^n converge in law to the Brownian bridge.*

Proof. a) The characteristics of the Brownian bridge are

$$B_t = -\int_0^t \frac{X_s}{1 - s} ds, \quad C_t = t, \quad \nu = 0,$$

and B satisfies no majoration hypothesis, because of the possible explosion at time 1.

So we fix $T \in (0, 1)$, and we stop the characteristics at T: that is, we consider

$$B(T)_t = -\int_0^{t \wedge T} \frac{X_s}{1 - s} ds, \quad C(T)_t = t \wedge T, \quad \nu(T) = 0.$$

We will also consider $V^n(T)_t = V_{t \wedge T}^n$. The characteristics $(B(T), C(T), \nu(T))$ obviously meet 3.39(i, ii, iv), and local uniqueness 3.39(iii) for the martingale problem $\mathscr{s}(\sigma(X_0), X | \varepsilon_0; B(T), C(T), \nu(T))$ is easily deduced from the results of §§ III.2c, d. We obviously have 3.39(v). Finally, II.3.37 gives the characteristics of $V^n(T)$, relative to the filtration generated by V^n:

$$B^n(T)_t = -\int_0^{t \wedge T} V^n(T)_s \frac{1}{1 - s} ds$$

$$\tilde{C}^n(T)_t = \int_0^{t \wedge T} \left[1 - \frac{1}{(1 - s)\sqrt{n}} V^n(T)_s\right] ds$$

$$g * \nu^n(T)_t = \int_0^{t \wedge T} \left[n - \frac{\sqrt{n}}{1 - s} V^n(T)_s\right] g\left(\frac{1}{\sqrt{n}}\right) ds.$$

Then $g * \nu^n(T)_t \xrightarrow{P} 0$ as $n \uparrow \infty$ for all $t \geq 0$, $g \in C_1(\mathbb{R}^d)$, because $g(1/\sqrt{n}) = 0$ for n large enough. Moreover, $B^n(T) - B(T) \circ V^n(T) = 0$. Finally

4.32 $$\tilde{C}^n(T)_t - \tilde{C}(T)_t \circ V^n(T) = -\int_0^{t \wedge T} \frac{1}{(1-s)\sqrt{n}} V^n(T)_s \, ds.$$

But a simple computation shows that $E[(V^n(T)_s)^2] = s(1-s)$ for $s \leq T$, so 4.32 goes to 0 in L^2 and a-fortiori in measure for all $t \in \mathbb{R}_+$. Therefore Theorem 3.39 gives:

4.33 For all $T \in (0,1)$, $V^n(T)$ converges in law to the Brownian bridge stopped at time T.

b) Let $V_t'^n = V_{1-t}^n$ for $t \in [0,1]$. Clearly the process V'^n has the same distribution as V^n. Hence if $T \in (0,1)$, the processes $V'^n(T)_t = V_{t \wedge T}'^n = V_{1-t\wedge T}^n$ also converge in law to the Brownian bridge stopped at T. Now we extend V^n and the Brownian bridge V to $\mathbb{R}_+$ by putting $V_t^n = V_t = 0$ for $t > 1$. We clearly have for $N > 1$:

$$\sup_{s \leq N} |V_s^n| \leq \sup_{s \leq N} \left| V^n\left(\frac{2}{3}\right)_s \right| + \sup_{s \leq N} \left| V'^n\left(\frac{2}{3}\right)_s \right|$$

$$w_N'(V^n, \theta) \leq w_N'\left(V^n\left(\frac{2}{3}\right), \theta\right) + w_N'\left(V'^n\left(\frac{2}{3}\right), \theta\right) \quad \text{if } \theta < \frac{1}{2}.$$

Then a double application of VI.3.21 yields tightness for $\{V^n\}$.

c) Finally, let $0 \leq t_1 < t_2 < \cdots < t_p$, with $t_{i-1} < 1 \leq t_i$ for some i. Applying 4.33 for $T = t_{i-1}$ shows in particular that $(V_{t_1}^n, \ldots, V_{t_{i-1}}^n) \xrightarrow{\mathcal{L}} (V_{t_1}, \ldots, V_{t_{i-1}})$. Moreover $V_{t_i}^n = \cdots = V_{t_p}^n = V_{t_i} = \cdots = V_{t_p} = 0$, so we deduce that

$$(V_{t_1}^n, \ldots, V_{t_p}^n) \xrightarrow{\mathcal{L}} (V_{t_1}, \ldots, V_{t_p}).$$

In other words, $V^n \xrightarrow{\mathcal{L}(\mathbb{R}_+)} V$. Since $\{V^n\}$ is tight, we deduce the claim. □

§ 4d. Convergence to a Continuous Semimartingale: Necessary and Sufficient Conditions

Here we specialize the results of Section 3 to the case where the limiting process X is continuous. Similarly to what happens in §§ VIII.3a, b, we can here obtain necessary and sufficient conditions for convergence.

The setting is the same as in Section 3, with the additional assumption

4.34 $$\nu = 0, \quad \text{so } \tilde{C} = C.$$

1. Let us recall notation II.2.4 and II.2.5:

4.35 $$X^n(h) = X^n - \sum_{s \leq \cdot} [\Delta X_s^n - h(\Delta X_s^n)]$$
$$M^n = X^n(h) - X_0^n - B^n,$$

and we set

4.36 $$\hat{C}^n = (\hat{C}^{n,ij})_{i,j \le d}, \quad \hat{C}^{n,ij} = [M^{n,i}, M^{n,j}].$$

We introduce the following conditions:

4.37 $$\begin{cases} [\hat{\gamma}_7\text{-}\mathbb{R}_+] & \hat{C}^n_t - C_t \circ X^n \xrightarrow{P} 0 \quad \text{for all } t \ge 0 \\ [\hat{\delta}_7] & \nu^n([0,t] \times \{x: |x| > \varepsilon\}) \xrightarrow{P} 0 \quad \text{for all } t \ge 0, \varepsilon > 0; \end{cases}$$

4.38 *Majoration condition*: $\sup_{\alpha \in \Omega} C^{ii}_t(\alpha) < \infty$ for all $t \in \mathbb{R}_+$ (the same as 2.11(ii) when $\nu = 0$). □

4.39 *Strong majoration condition*: for all $\alpha \in \mathbb{D}(\mathbb{R}^d)$ the functions $\mathrm{Var}(B^i(\alpha))$ and $C^{ii}(\alpha)$ are strongly majorized by an increasing continuous (deterministic) function F (the same as 3.6 when $\nu = 0$). □

4.40 *Pointwise continuity*: $\alpha \rightsquigarrow B_t(\alpha)$, $C_t(\alpha)$ are Skorokhod-continuous on $\mathbb{D}(\mathbb{R}^d)$ for all $t \in \mathbb{R}_+$ (the same as 2.11(iii) when $\nu = 0$). □

4.41 *Functional continuity*: $\alpha \rightsquigarrow B(\alpha)$, $C(\alpha)$ are Skorokhod-continuous from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^d)$ and $\mathbb{D}(\mathbb{R}^{d^2})$. □

4.42 **Lemma.** a) *4.39 implies 4.38.*
b) *4.39 and 4.40 imply 4.41.*

Proof. (a) is obvious. VI.3.36 implies that the families $\{B(\alpha)\}_{\alpha \in \Omega}$ and $\{C(\alpha)\}_{\alpha \in \Omega}$ are relatively compact in $\mathbb{D}(\mathbb{R}^d)$ and $\mathbb{D}(\mathbb{R}^{d^2})$ respectively, as soon as 4.39 holds. Then the claim in (b) is obvious. □

4.43 **Lemma.** a) *If X^n converges in law to a continuous process, $[\hat{\delta}_7]$ holds.*
b) *If 4.34 and 4.38 and $[\hat{\delta}_7]$ hold, then $[\gamma_7\text{-}\mathbb{R}_+] \Leftrightarrow [\hat{\gamma}_7\text{-}\mathbb{R}_+]$.*

Proof. a) The hypothesis implies that $\sup_{s \le t} |\Delta X^n_s| \xrightarrow{P} 0$ by VI.3.26, which is turn yields $[\hat{\delta}_7]$ by VI.4.22.

b) The proof is the same as for VIII.3.6a: in this, the deterministic character of C is used only to prove (1) or (2), and these properties follow here from the majoration condition 4.38 and from $[\gamma_7\text{-}\mathbb{R}_+]$ for (1), from $[\hat{\gamma}_7\text{-}\mathbb{R}_+]$ for (2). □

Here is the main result, which extends Theorem VIII.3.8. Recall that $\eta^n = \mathcal{L}(X^n_0)$.

4.44 **Theorem.** a) *Assume the strong majoration 4.39, the pointwise continuity 4.40, and also that the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, 0)$ has a unique solution, say P. If moreover $[\text{Sup-}\beta_7]$ holds, there is equivalence between:*

(i) $\mathscr{L}(X^n) \to P$;
(ii) $[\hat{\delta}_7]$, $[\hat{\gamma}_7\text{-}\mathbb{R}_+]$, $\eta^n \to \eta$ *hold*;
(iii) $[\hat{\delta}_7]$, $[\gamma_7\text{-}\mathbb{R}_+]$, $\eta^n \to \eta$ *hold.*

b) *Assume the functional continuity 4.41 and* [Sup-β_7]. *If* $\mathscr{L}(X^n) \to P$ *weakly and if the canonical process* X *on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ *is a semimartingale with characteristics* $(B, C, 0)$ *and initial distribution* η, *then* $[\hat{\delta}_7]$ *and* $[\hat{\gamma}_7\text{-}\mathbb{R}_+]$ *and* $\eta^n \to \eta$ *hold. If moreover the majoration 4.38 holds, we have* $[\gamma_7\text{-}\mathbb{R}_+]$ *as well.*

Proof. a) The implication (iii) $\Rightarrow$ (i) is a particular case of 3.21 (if $\nu = 0$, $[\hat{\delta}_7] \Leftrightarrow [\delta_{7,1}\text{-}\mathbb{R}_+]$ and 3.7 is obvious). (ii) $\Rightarrow$ (iii) follows from 4.43, and (i) $\Rightarrow$ (ii) will follow from (b), because of 4.42b.

b) Accordingly to 4.35, we define $X(h) = X - \sum_{s \le \cdot} [\Delta X_s - h(\Delta X_s)]$ and $M = X(h) - X_0 - B$. Then $\alpha \rightsquigarrow X(h)(\alpha)$ is continuous from $\mathbb{D}(\mathbb{R}^d)$ into itself by VI.2.8, and VI.3.33 (or VI.2.2) plus 4.41 and the continuity in t of B_t and C_t imply that $\alpha \rightsquigarrow (X(h), B, C)(\alpha)$ is continuous from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^{d+d+d^2})$. Then $\mathscr{L}(X^n) \to P$ yields

$$\mathscr{L}(X^n(h), B \circ X^n, C \circ X^n) \to \mathscr{L}[(X(h), B, C) | P]$$

(note that $X^n(h) = X(h) \circ X^n$). Therefore

$$\mathscr{L}(X^n(h) - B \circ X^n - X_0^n, C \circ X^n) \to \mathscr{L}[(X(h) - B - X_0, C) | P] = \mathscr{L}[(M, C) | P].$$

Then [Sup-β_7] clearly implies that $\mathscr{L}(X^n(h) - B \circ X^n - X_0^n, C \circ X^n)$ and $\mathscr{L}(X^n(h) - B^n - X_0^n, C \circ X^n)$ have the same limits, hence

4.45 $$\mathscr{L}(M^n, C \circ X^n) \to \mathscr{L}[(M, C) | P].$$

Moreover there is a constant A depending only on the truncation function h, such that $|\Delta M^n| \le A$ identically. So VI.6.29 allows to deduce from 4.45 that

4.46 $$\mathscr{L}(\hat{C}^n, C \circ X^n) \to \mathscr{L}[(C, C) | P]$$

because $\hat{C}^{n,ij} = [M^{n,i}, M^{n,j}]$, and under P we have $[M^i, M^j] = \langle M^i, M^j \rangle = C^{ij}$ (recall that X and M are P-a.s. continuous). Now, 4.46 clearly implies that $\mathscr{L}(\hat{C}^n - C \circ X^n)$ converges to the law of the process $C - C = 0$, and thus $[\hat{\gamma}_7\text{-}\mathbb{R}_+]$ holds.

That $[\hat{\delta}_7]$ follows from 4.43a, and $\eta^n \to \eta = \mathscr{L}(X_0 | P)$ is trivial. Finally, under 4.38 we deduce $[\gamma_7\text{-}\mathbb{R}_+]$ from 4.43b. □

2. When the limiting process X is a *continuous local martingale* (i.e. $B = 0$ above) we have genuine necessary and sufficient conditions for convergence, as in § VIII.3b whose results can all be generalized to our present situation. We will only state here, and without proof, the extension of VIII.3.12 (in view of 4.44, the proof is similar).

4.47 **Theorem.** *Assume that* $B = 0$ *and* $\nu = 0$, *and also that each* X^n *is a local martingale, with* $X_0^n = 0$ (the latter is just for simplicity). *Assume also that*

4.48 $$\lim_{a\uparrow\infty} \limsup_n P^n(|x| 1_{\{|x|>a\}} * \nu_t^n > \varepsilon) = 0 \text{ for all } \varepsilon > 0, t > 0.$$

a) *Under 4.39 and 4.40, and if P is the unique solution to the martingale problem* $\mathscr{s}(\sigma(X_0), X | \varepsilon_0; 0, C, 0)$, *there is equivalence between*:

(i) $\mathscr{L}(X^n) \to P$;

(ii) $[\hat{\delta}_7] + [\hat{\gamma}_7\text{-}\mathbb{R}_+]$;

(iii) $[\hat{\delta}_7] + [\gamma_7\text{-}\mathbb{R}_+]$;

(iv) $[X^{n,i}, X^{n,j}]_t - C_t^{ij} \circ X^n \xrightarrow{P} 0$ *for all* $t \geq 0, i, j \leq d$.

b) *Assume 4.41 and that* $\mathscr{L}(X^n) \to P$, *and that under P the canonical process X is a continuous local martingale with* $X_0 = 0$ *a. s. and* $\langle X^i, X^j \rangle = C^{ij}$. *Then* (ii) *and* (iv) *above hold. If moreover 4.38 holds, we also have* (iii).

5. Convergence of Stochastic Integrals

Here we will consider another sort of limit theorems. We suppose that a sequence (X^n) of semimartingales converges in law to a semimartingale X. We also have a sequence (H^n) of locally bounded predictable processes which converges in a suitable sense to a locally bounded predictable process H. Do we have convergence in law of the stochastic integrals $H^n \cdot X^n$ toward $H \cdot X$?

In general, the answer is no: consider the deterministic case, where $X_t^n(\omega) = x^n(t)$ converges in the Skorokhod sense (or even uniformly) to a limit $X_t(\omega) = x(t)$, and let $H_t^n(\omega) = H_t(\omega) = h(t)$ be a bounded function. Then of course it is not true in general that $\int_0^t h(s)\, dx^n(s) \to \int_0^t h(s)\, dx(s)$, unless we make either a continuity assumption on h, or the assumption that $x^n \to x$ in variation.

Here we make no continuity assumption on the H^n's, but we strengthen the convergence of the X^n's. Since these do not have finite variation, we cannot use the convergence in law for the variation distance, but we will replace this by the convergence in variation of the characteristics.

§ 5a. Characteristics of Stochastic Integrals

In this subsection we make calculations that have nothing to do with limit theorems. Let $X = (X^i)_{i \leq d}$ be a semimartingale on $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν) and modified second characteristic $\tilde{C}$, relative to some truncation function $h \in \mathscr{C}_t^d$.

Let $H = (H^{ij})_{i \leq m, j \leq d}$ be an $m \times d$-dimensional *predictable* process on $\mathscr{B}$, which for simplicity we suppose *locally bounded*. Let $Y = (Y^i)_{i \leq m}$ be

5.1 $$Y^i = \sum_{j \le d} H^{ij} \cdot X^j \quad (\text{denoted by } Y = H \cdot X)$$

be the stochastic integral. We wish to describe the characteristics of the $(d+m)$-dimensional semimartingale $Z = (X, Y)$. For this, we fix a truncation function $h' \in \mathscr{C}_t^{d+m}$. Hx will stand for the process $Hx = ((Hx)^i = \sum_{j \le d} H^{ij} x^j)_{i \le m}$, and (x, Hx) for the $(d+m)$-dimensional vector with projections x and Hx on the first d- and last m-dimensional subspaces. We define the following quantities (by convention we set $+\infty$ if the right-hand side is not well defined):

5.2 $$\begin{cases} B'^i = \begin{cases} B^i + [h'^i(x, Hx) - h^i(x)] * \nu & \text{if } i \le d \\ \sum_{j \le d} H^{i-d,j} \cdot B^j + \left[h'^i(x, Hx) - \sum_{j \le d} H^{i-d,j} h^j(x) \right] * \nu & \text{if } d < i \le d+m \end{cases} \\ C'^{ij} = \begin{cases} C^{ij} & \text{if } i, j \le d \\ \sum_{k \le d} H^{i-d,k} \cdot C^{kj} & \text{if } j \le d < i \le d+m \\ \sum_{k \le d} H^{j-d,k} \cdot C^{ik} & \text{if } i \le d < j \le d+m \\ \sum_{k,l \le d} (H^{i-d,k} H^{j-d,l}) \cdot C^{kl} & \text{if } d < i, j \le d+m \end{cases} \\ \nu' \text{ defined by } 1_G * \nu' = 1_G(x, Hx) * \nu \quad \text{for all } G \in \mathscr{R}^{d+m}. \end{cases}$$

5.3 **Proposition.** *If $Y = H \cdot X$, the characteristics of $Z = (X, Y)$ are the terms (B', C', ν') defined by 5.2.*

Proof. This is a tedious, but simple, calculation. Firstly, the continuous martingale parts are related by $Y^c = H \cdot X^c$, and I.4.41 yields

$$\langle Y^{c,i}, X^{c,j} \rangle = \sum_{k \le d} H^{ik} \cdot \langle X^{c,k}, X^{c,j} \rangle$$

$$\langle Y^{c,i}, X^{c,j} \rangle = \sum_{k,l \le d} (H^{i,k} H^{j,l}) \cdot \langle X^{c,k}, X^{c,l} \rangle,$$

while $C^{ij} = \langle X^{c,i}, X^{c,j} \rangle$: hence the second characteristic C' is given by 5.2.

Secondly, we have $\Delta Y = H \Delta X$, and so $\Delta Z = (\Delta X, H \Delta X)$. Thus if μ^X and μ^Z denote the random measures associated with the jumps of X and Z,

5.4 $$1_G * \mu_t^Z = \int_0^t \int_{\mathbb{R}^d} 1_G(x, Hx) \mu^X(ds \times dx)$$

for all $G \in \mathbb{R}^{d+m}$. This relation carries over to the compensators, and thus the third characteristic ν' of Z is given by 5.2.

Finally, by definition of B there is a local martingale M with

$$X = X_0 + M + B + (x - h(x)) * \mu^X$$

and thus

$$Y = H \cdot M + H \cdot B + (Hx - Hh(x)) * \mu^X.$$

Now, using the notation II.2.4, we obtain by 5.4 (and with obvious notation):

$$\begin{aligned} Z(h') &:= Z - (z - h'(z)) * \mu^Z \\ &= \begin{pmatrix} X_0 \\ 0 \end{pmatrix} + \begin{pmatrix} M \\ H \cdot M \end{pmatrix} + \begin{pmatrix} B \\ H \cdot B \end{pmatrix} + \begin{pmatrix} (x - h(x)) * \mu^X \\ (Hx - Hh(x)) * \mu^X \end{pmatrix} \\ &\quad - \left[\begin{pmatrix} x \\ Hx \end{pmatrix} - h'(x, Hx) \right] * \mu^X \end{aligned}$$

5.5
$$= \begin{pmatrix} X_0 \\ 0 \end{pmatrix} + \begin{pmatrix} M \\ H \cdot M \end{pmatrix} + \begin{pmatrix} B \\ H \cdot B \end{pmatrix} + \left[h'(x, Hx) - \begin{pmatrix} h(x) \\ Hh(x) \end{pmatrix} \right] * \mu^X.$$

Moreover $Z(h')$ is a special semimartingale and B and $H \cdot B$ have locally integrable variation by I.3.10, because they are predictable with finite variation. Hence I.4.23 yields that the last term in 5.5 also has locally integrable variation, and its compensator is the same integral, but with respect to ν. By definition of the first characteristic B' of Z, we obtain

$$B' = \begin{pmatrix} B \\ H \cdot B \end{pmatrix} + \left[h'(x, Hx) - \begin{pmatrix} h(x) \\ Hh(x) \end{pmatrix} \right] * \nu,$$

which is the expression given by 5.2. □

5.6 **Proposition.** *Let Y' be an m-dimensional semimartingale on $\mathscr{B}$, with $Y'_0 = 0$, and suppose that the $(d + m)$-dimensional semimartingale $Z' = (X, Y')$ admits (B', C', ν'), as given by 5.2, for its characteristics. Then $Y' = H \cdot X$.*

Proof. a) Set $Y = H \cdot X$. We first prove that the continuous martingale parts Y^c and Y'^c coincide. Since $Y^{c,i} = \sum_{j \le d} H^{ij} \cdot X^{c,j}$, by using the fact that C' is the second characteristic of both $Z = (X, Y)$ and $Z' = (X, Y')$, and I.4.41, we obtain:

$$\begin{aligned} \langle Y^{c,i} - Y'^{c,i}, Y^{c,i} - Y'^{c,i} \rangle &= \langle Y^{c,i}, Y^{c,i} \rangle + \langle Y'^{c,i}, Y'^{c,i} \rangle - 2 \langle Y^{c,i}, Y'^{c,i} \rangle \\ &= C'^{d+i,d+i} + C'^{d+i,d+i} - 2 \sum_{j \le d} H^{ij} \cdot \langle X^{c,j}, Y'^{c,i} \rangle \\ &= 2C'^{d+i,d+i} - 2 \sum_{j \le d} H^{ij} \cdot C'^{j,d+i} \\ &= 2C'^{d+i,d+i} - 2 \sum_{j \le d} H^{ij} \cdot \left[\sum_{k \le d} H^{ik} \cdot C^{jk} \right], \end{aligned}$$

which equals 0 by 5.2. Hence, since $Y^c_0 = Y'^c_0 = 0$, I.4.6 yields that $Y^{c,i} - Y'^{c,i} = 0$ a. s., and the claim is proved.

b) Let $G = \{(\omega, t, x, y) \in \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^m \colon y \neq H_t(\omega) x\}$, which is $\mathscr{P} \otimes \mathscr{R}^{d+m}$-measurable. By 5.2 we have $1_G * \nu'_\infty = 0$, so $E(1_G * \mu^{Z'}_\infty) = 0$ where $\mu^{Z'}$ is the random measure associated with the jumps of Z'. Hence

5.7 $\Delta Y' = H\Delta X$, and so $\mu^{Z'} = \mu^Z$ up to a null set (because $\Delta Y = H\Delta X$ by definition of Y).

c) Finally, the canonical representation of semimartingales II.2.35 allows to write (recalling that $Y_0 = Y'_0 = 0$):

$$Y^i = Y^{c,i} + h'^{d+i} * (\mu^Z - \nu') + (z^{d+i} - h'^{d+i}(z)) * \mu^Z + B'^{d+i}$$

$$Y'^i = Y'^{c,i} + h'^{d+i} * (\mu^{Z'} - \nu') + (z^{d+i} - h'^{d+i}(z)) * \mu^{Z'} + B'^{d+i},$$

where we have used that B' is the first characteristic of Z (by 5.3) and of Z' (by hypothesis). In view of 5.7 and of (a), we deduce that $Y' = Y$ up to a P-null set, and that finishes the proof. □

For further reference, we also compute the modified second characteristic $\tilde{C}'$ of Z. Applying 5.2 and 2.2, we obtain

5.8
$$\begin{aligned}
\tilde{C}'^{ij}_t &= \tilde{C}^{ij}_t + [h'^i h'^j(x, Hx) - h^i h^j(x)] * \nu_t \\
&\quad + \sum_{s \le t} [\Delta B^i_s \Delta B^j_s - \Delta B'^i_s \Delta B'^j_s] \quad \text{if } i, j \le d \\
&= \sum_{k \le d} H^{i-d,k} \cdot \tilde{C}^{kj}_t + \left[h'^i h'^j(x, Hx) - \sum_{k \le d} H^{i-d,k}(h^k h^j)(x) \right] * \nu_t \\
&\quad + \sum_{s \le t} \left[\sum_{k \le d} H^{i-d,k}_s \Delta B^k_s \Delta B^j_s - \Delta B'^i_s \Delta B'^j_s \right] \quad \text{if } j \le d < i \le d + m \\
&= \tilde{C}'^{ji}_t \quad \text{if } i \le d < j \le d + m \\
&= \sum_{k,l \le d} (H^{i-d,k} H^{j-d,l}) \cdot \tilde{C}^{kl}_t \\
&\quad + \left[h'^i h'^j(x, Hx) - \sum_{k,l \le d} H^{i-d,k} H^{j-d,l}(h^k h^l)(x) \right] * \nu_t \\
&\quad + \sum_{s \le t} \left[\sum_{k,l \le d} H^{i-d,k}_s H^{j-d,l}_s \Delta B^k_s \Delta B^l_s - \Delta B'^i_s \Delta B'^j_s \right] \quad \text{if } d < i, j \le d + m.
\end{aligned}$$

§ 5b. Statement of the Results

1. Let us first describe the general setting of this section. For every $n \in \mathbb{N}^*$ the stochastic basis $\mathscr{B}^n$ is endowed with a semimartingale $X^n = (X^{n,i})_{i \le d}$ with characteristics (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$, relatively to a given *continuous* truncation function $h \in \mathscr{C}^d_t$.

There is also a predictable process $H^n = (H^{n,ij})_{i \le m, j \le d}$, which for simplicity we suppose locally bounded, and according to 5.1 we define $Y^n = (Y^{n,i})_{i \le m}$ by

$Y^n = H^n \cdot X^n$. We also put $Z^n = (X^n, Y^n)$, which is an $(m + d)$-dimensional semimartingale.

Secondly, we consider the canonical space $\Omega = \mathbb{D}(\mathbb{R}^d)$ with $\mathscr{F} = \mathscr{D}(\mathbb{R}^d)$ and $\mathbf{F} = \mathbf{D}(\mathbb{R}^d)$ and the canonical process X (see 2.6). On Ω we are given a triplet (B, C, ν) satisfying 2.1, and we define $\tilde{C}$ by 2.2. We also have a locally bounded predictable process $H = (H^{ij})_{i \le m, j \le d}$ on Ω.

Our aim is to find conditions insuring that, knowing $\mathscr{L}(X^n) \to P$ weakly for some measure P on $(\Omega, \mathscr{F})$, then $Z^n = (X^n, H^n \cdot X^n) \xrightarrow{\mathscr{L}} (X, H \cdot X)$. For this we need a reinforcement of the previous conditions of Sections 2 and 3, concerning both the convergence of B^n, C^n, ν^n to B, C, ν and the continuity properties of B, C, ν. So we put

5.9
$$\begin{cases} [\text{Var-}\beta] & \operatorname{Var}(B^{n,i} - B^i \circ X^n)_t \xrightarrow{P} 0 \quad \text{for all } t \ge 0, i \le d \\ [\text{Var-}\gamma] & \operatorname{Var}(\tilde{C}^{n,ij} - \tilde{C}^{ij} \circ X^n)_t \xrightarrow{P} 0 \quad \text{for all } t \ge 0, i,j \le d \\ [\text{Var-}\delta] & \operatorname{Var}(g * \nu^n - (g * \nu) \circ X^n)_t \xrightarrow{P} 0 \quad \text{for all } t \ge 0, g \in C_1(\mathbb{R}^d), \end{cases}$$

and

5.10 *Continuity condition at a point* $\alpha \in \mathbb{D}(\mathbb{R}^d)$: for every sequence (α_n) converging to α in $\mathbb{D}(\mathbb{R}^d)$ and all $t \ge 0$, $g \in C_1(\mathbb{R}^d)$,

$$\operatorname{Var}(B^i(\alpha_n) - B^i(\alpha))_t \to 0, \quad \operatorname{Var}(\tilde{C}^{ij}(\alpha_n) - \tilde{C}^{ij}(\alpha))_t \to 0$$
$$\operatorname{Var}(g * \nu(\alpha_n) - g * \nu(\alpha))_t \to 0. \qquad \square$$

For simplicity, we will also assume the following property of $C_1(\mathbb{R}^d)$, which does not prevent it to be chosen countable:

5.11 $\qquad C_1(\mathbb{R}^d)$ is stable by multiplication.

2. There are essentially two types of results.

5.12 **Theorem.** *Assume that the truncation function h is continuous with compact support, that (B, C, ν) meets 3.7 and 3.11 and the continuity condition 5.10 at every point $\alpha \in \mathbb{D}(\mathbb{R}^d)$, and also that there is a unique measure P solution to the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$, where η is a probability measure on $\mathbb{R}^d$.*

Moreover, assume that [Var-β], [Var-γ], [Var-δ] *hold, and that $\eta^n := \mathscr{L}(X_0^n) \to \eta$ weakly. Finally, assume the following on H^n and H:*

5.13 $\qquad H_t^n - H_t \circ X^n \xrightarrow{P} 0 \quad$ for all $t \ge 0$;

5.14 $\qquad$ for all $t \ge 0$, $\sup_{\alpha \in \Omega, s \le t} |H_s(\alpha)| < \infty$ and $\alpha \rightsquigarrow H_t(\alpha)$ is Skorokhod continuous on $\mathbb{D}(\mathbb{R}^d)$;

5.15 for all $t \ge 0$ there is a constant K_t with $P^n(\sup_{s \le t} |H_s^n| > K_t) \to 0$ as $n \uparrow \infty$.

Then the laws $\mathscr{L}(X^n, H^n \cdot X^n)$ weakly converge to the law $\mathscr{L}(X, H \cdot X)$ of the $(d + m)$-dimensional process $(X, H \cdot X)$ on $(\Omega, \mathscr{F}, P)$.

This result is, in principle, a direct corollary of Theorem 3.35: it suffices to prove that the sequence (Z^n) and the characteristics (B', C', ν') defined by 5.2 satisfy all the conditions of this theorem. Consequently, there exists a "local" version of the previous result in the quasi-left continuous case (based on Theorem 3.39 instead of 3.35).

For these reasons, Theorem 5.11 has nothing new in it. The next result, on the opposite, is more interesting, although it only concerns the case where the limit is quasi-left continuous; in it, we will only suppose that the processes X^n converge in law to a limit process X, and it needs no uniqueness assumption as for the martingale problem $\mathscr{s}(\sigma(X_0), X|\eta; B, C, \nu)$ (the proofs are provided in § 5c below).

5.16 **Theorem.** *Assume that the truncation function h is continuous with compact support, that the sequence $\mathscr{L}(X^n)$ weakly converges to a limit* $\mathbf{P}$*, that X is a quasi-left continuous semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with characteristics (B, C, ν), and that the continuity condition 5.10 holds for P-almost all α.*

Moreover, assume that [Var-β], [Var-γ], [Var-δ] *hold, as well as 5.14 and*

5.17
$$\sup_{s \leq t} |H^n_s - H_s \circ X^n| \xrightarrow{P} 0 \text{ for all } t \geq 0.$$

Then the sequence $\mathscr{L}(X^n, H^n \cdot X^n)$ weakly converges to the law of the $(d+m)$-dimensional process $(X, H \cdot X)$ under P.

Conditions 5.9 and 5.10 look very stringent. In many cases, though, they are naturally fulfilled. For instance, suppose that (B, C, ν) is given by 4.1, and also that each X^n is a diffusion process with jumps, with characteristics given by 4.7. Then

5.18 **Corollary.** *Assume further that 4.10, 4.11, 5.14, and 5.17 hold, and that $\mathscr{L}(X^n)$ weakly converges to a limit P under which X is a semimartingale with characteristics (B, C, ν). Then $\mathscr{L}(X^n, H^n \cdot X^n)$ weakly converges to the law of $(X, H \cdot X)$ under P.*

Proof. In view of 4.1 and 4.7, it is obvious that 4.11 and the convergence $\mathscr{L}(X^n) \to P$ (resp. 4.10) imply [Var-β] + [Var-γ] + [Var-δ] (resp. 5.10 for all α). □

5.19 **Remark.** The convergence in variation in 5.9 and 5.10 is certainly too much for our needs. Let us consider 5.10 for instance. It will be apparent from the proof that it is enough to ask:

(i) $\sup_{s \leq t} |u \cdot B^i(\alpha_n)_s - u \cdot B^i(\alpha)_s| \to 0$

(ii) $\sup_{s \leq t} |u \cdot \tilde{C}^{ij}(\alpha_n)_s - u \cdot \tilde{C}^{ij}(\alpha)_s| \to 0$

(iii) $\sup_{s \leq t} |v * \nu(\alpha_n)_s - v * \nu(\alpha)_s| \to 0$

(iv) $\sup_n \mathrm{Var}(B^i(\alpha_n)) < \infty$

for all Borel functions $u\colon \mathbb{R}_+ \to [0,1]$ and $v\colon \mathbb{R}_+ \times \mathbb{R}^d \to [0,1]$, with $v(s,x)$ continuous in x for all s.

One could weaken [Var-β], [Var-γ], [Var-δ] accordingly. □

§ 5c. The Proofs

We proceed through several steps.

1. Our first observation intends to show that one may assume H and H^n to be uniformly bounded. Note that 5.14 plus 5.17 certainly imply 5.15. Since the convergence $\mathscr{L}(X^n, H^n \cdot X^n) \to \mathscr{L}(X, H \cdot X)$ is a "local" property, it suffices to prove the result for H^n and H replaced by $H^n 1_{[0,T]}$ and $H 1_{[0,T]}$, for all $T \in \mathbb{R}_+$. So it is not a restriction in 5.14 and 5.15 to assume that

5.20
$$\begin{cases} \sup\limits_{\alpha \in \Omega, t \geq 0} |H_t(\alpha)| < \infty, \\ P^n\left(\sup\limits_{s \leq t} |H_s^n| > K\right) \to 0 \end{cases}$$

for a constant K not depending on t.

Moreover, let $H_t'^n = H_t^n 1_{\{|H_t^n| \leq K\}}$. We have $H^n \cdot X_s^n = H'^n \cdot X_s^n$ for all $s \leq t$ on the set $\{\sup_{s \leq t} |H_s^n| \leq K\}$. So in view of 5.20 the sequence $\mathscr{L}(X^n, H^n \cdot X^n)$ converges to a limit $\tilde{P}$ if and only if the sequence $\mathscr{L}(X^n, H'^n \cdot X^n)$ converges to the same limit $\tilde{P}$. Furthermore 5.13 (resp. 5.17) yields $H_t'^n - H_t \circ X^n \xrightarrow{P} 0$ (resp. $\sup_{s \leq t} |H_s'^n - H_s \circ X^n| \xrightarrow{P} 0$), because of 5.20 again. Therefore we can replace H^n by H'^n, or in other words we can assume

5.21
$$|H^n| \leq K, \quad |H| \leq K \text{ identically}$$

for some constant K.

2. In the next lemma, ρ^n and ρ'^n are positive random measures on $\mathscr{B}^n$ which integrate $(|x|^2 \wedge 1) 1_{[0,t]}$ for all t. We denote by $\mathscr{K}$ the compact subset of all $u \in \mathbb{R}^m \otimes \mathbb{R}^d$ such that $|u| \leq K$, where K is as in 5.21. We consider two sequences of processes K^n and K'^n with values in $\mathscr{K}$.

5.22 **Lemma.** *Assume that* (i), *and either* (ii) *or* (ii′) *below, hold*:

(i) $\mathrm{Var}(g * \rho^n - g * \rho'^n)_t \xrightarrow{P} 0$ *for all* $t \geq 0$, $g \in C_1(\mathbb{R}^d)$.

(ii) (1) *For each* $g \in C_1(\mathbb{R}^d)$ *there is a càdlag increasing* (deterministic) *function* F^g *which strongly majorizes all* $g * \rho'^n(\omega)$;

(2) $K_t^n - K_t'^n \xrightarrow{P} 0$ *for all* $t \geq 0$.

(ii′) (1) $\lim_{N \uparrow \infty} \limsup_n P^n(g * \rho_t'^n > N) = 0$ *for all* $t \geq 0$, $g \in C_1(\mathbb{R}^d)$;

(2) $\sup_{s \leq t} |K_s^n - K_s'^n| \xrightarrow{P} 0$ *for all* $t \geq 0$.

Then if

5.23
$$f : \mathbb{R}^d \times \mathscr{K} \to \mathbb{R} \text{ is continuous, has a limit at infinity, and } f(x,u) = 0 \text{ for all } |x| \leq \varepsilon, u \in \mathscr{K} \text{ for some } \varepsilon > 0,$$

we have

5.24
$$\mathrm{Var}[f(x, K^n) * \rho^n - f(x, K'^n) * \rho'^n]_t \xrightarrow{P} 0 \text{ for all } t \geq 0.$$

Proof. a) Let $\bar{C}$ be the space of all continuous bounded functions on the compact set $(\mathbb{R}^d \cup \{\infty\}) \times \mathscr{K}$, where $\mathbb{R}^d \cup \{\infty\}$ is the one-point compactification of $\mathbb{R}^d$. Recall that $g_a(x) = (a|x| - 1)^+ \wedge 1$ belongs to $C_1(\mathbb{R}^d)$ if $a \in \mathbb{Q}_+$. A function f meets 5.23 if and only if it has the form $f(x,u) = g_a(x)\bar{f}(x,u)$ for some $a \in \mathbb{Q}_+$ and $\bar{f} \in \bar{C}$. So we fix $a \in \mathbb{Q}_+$, and it suffices to prove that for $\bar{f} \in \bar{C}$, the function $f = g_a \bar{f}$ meets 5.24.

b) We first show that this is true if $\bar{f}(x,u) = \bar{g}(x)k(u)$, where $\bar{g}$ is the extension to $\mathbb{R}^d \cup \{\infty\}$ of a function $g \in C_1(\mathbb{R}^d)$, and k is a continuous function on $\mathscr{K}$. In this case we have $f(x,K^n) * \rho^n = k(K^n) \cdot [gg_a * \rho^n]$, and similarly for $f(x,K'^n) * \rho'^n$, hence

$$
\begin{aligned}
&\operatorname{Var}[f(x,K^n) * \rho^n - f(x,K'^n) * \rho'^n]_t \\
\text{5.25} \qquad &\le |k(K^n)| \cdot \operatorname{Var}[gg_a * \rho^n - gg_a * \rho'^n]_t + |k(K^n) - k(K'^n)| \cdot (gg_a * \rho'^n)_t \\
&\le \|k\| \operatorname{Var}(gg_a * \rho^n - gg_a * \rho'^n)_t + \|g\|\, |k(K^n) - k(K'^n)| \cdot (g_a * \rho'^n)_t.
\end{aligned}
$$

The first term in 5.25 goes to 0 in measure by (i). Under (ii), the second term is smaller than $\|g\|\, |k(K^n) - k(K'^n)| \cdot F_t^{g_a}$, which goes to 0 in measure by Lebesgue convergence theorem. Since k is uniformly continuous, for $\varepsilon > 0$ there exists $\eta(\varepsilon) > 0$ such that $|u - u'| \le \eta(\varepsilon) \Rightarrow |k(u) - k(u')| \le \varepsilon$, and so

$$
\begin{aligned}
P^n(|k(K^n) - k(K'^n)| \cdot (g_a * \rho'^n)_t > \varepsilon) &\le P^n\left(\sup_{s \le t} |K_s^n - K_s'^n| > \eta\left(\frac{\varepsilon}{N}\right)\right) \\
&\quad + P^n(g_a * \rho_t'^n > N).
\end{aligned}
$$

Under (ii′), this expression can be made as small as one wishes for n big. So in all cases 5.25 goes to 0 in measure, and our claim is proved.

c) By linearity, 5.24 is also true for all $\bar{f}$ belonging to the linear space $\mathscr{E}$ spanned by the functions $\bar{g}(x)k(u)$ ($g \in C_1(\mathbb{R}^d)$, k continuous on $\mathscr{K}$). In view of 5.11, $\mathscr{E}$ is an algebra, and because of the convergence-determining property of $C_1(\mathbb{R}^d)$ it certainly separates the points of $(\mathbb{R}^d \cup \{\infty\}) \times \mathscr{K}$. Hence the Stone-Weierstrass Theorem yields that $\mathscr{E}$ is dense in $\bar{C}$ for the uniform convergence.

d) Let $\bar{f} \in \bar{C}$, and $\bar{f}_q \in \mathscr{E}$ with $\|\bar{f} - \bar{f}_q\| \le 1/q$. Then if $f_q = g_a \bar{f}_q$ and $f = g_a \bar{f}$ we have

$$
\operatorname{Var}[f(x,K^n) * \rho^n - f_q(x,K^n) * \rho^n] \le |f(x,K^n) - f_q(x,K^n)| * \rho^n \le \frac{1}{q} g_a * \rho^n,
$$

and similarly for ρ'^n. Thus

$$
\begin{aligned}
&\operatorname{Var}[f(x,K^n) * \rho^n - f(x,K'^n) * \rho'^n]_t \\
\text{5.26} \qquad &\le \operatorname{Var}[f_q(x,K^n) * \rho^n - f_q(x,K'^n) * \rho'^n]_t + \frac{1}{q} g_a * \rho_t^n + \frac{1}{q} g_a * \rho_t'^n.
\end{aligned}
$$

Note that (ii.1) $\Rightarrow$ (ii′.1), and that (i) and (ii′.1) immediately imply that ρ^n also satisfies (ii′.1). We easily deduce that $\lim_{q \uparrow \infty} \limsup_n P^n(\frac{1}{q} g_a * \rho_t'^n > \varepsilon) = 0$ for all

$\varepsilon > 0$, and similarly for ρ^n. Hence 5.26 and the fact that each f_q meets 5.24 imply that f also satisfies 5.24. □

This lemma will be applied twice. Its first application concerns the continuity of the characteristics (B', C', ν') and $\tilde{C}'$ defined in 5.2 and 5.8, and in which we have chosen a continuous truncation function $h' \in \mathscr{C}_t^{d+m}$ with compact support.

5.27 **Proposition.** *Assume that 5.10 holds for α, that $|H| \leq K$ and that $H_t(\cdot)$ is continuous at point α for all t. Then if $\alpha_n \to \alpha$ we have for all $t \in \mathbb{R}_+$, $g' \in C_2(\mathbb{R}^{d+m})$:*

5.28
$$\begin{cases} \operatorname{Var}[B'^i(\alpha_n) - B'^i(\alpha)]_t \to 0 \\ \operatorname{Var}[\tilde{C}'^{ij}(\alpha_n) - \tilde{C}'^{ij}(\alpha)]_t \to 0 \\ \operatorname{Var}[g' * \nu'(\alpha_n) - g' * \nu'(\alpha)]_t \to 0 \end{cases}$$

(where $C_2'(\mathbb{R}^{d+m})$ is the set of all continuous functions on $\mathbb{R}^{d+m}$ having a limit at infinity and vanishing on a neighbourhood of 0).

Proof. a) We will apply 5.22 to the deterministic measures $\rho^n(\omega) = \nu(\alpha_n)$, $\rho'^n(\omega) = \nu(\alpha)$, and $K^n(\omega) = H(\alpha_n)$, $K'^n(\omega) = H(\alpha)$. Then 5.22(i, ii.2) hold by hypothesis, and 5.22(ii.1) holds with $F^g = g * \nu(\alpha)$. So for every f meeting 5.23 we have

5.29
$$\operatorname{Var}[f(x, H(\alpha_n)) * \nu(\alpha_n) - f(x, H(\alpha)) * \nu(\alpha)]_t \to 0.$$

b) If $g' \in C_2'(\mathbb{R}^{d+m})$ the function $f(x,u) = g'(x, ux)$ clearly has 5.23, and $g' * \nu' = f(x, H) * \nu$ by 5.2, hence the last convergence in 5.28 follows at once from 5.29.

c) In view of 5.2, B'^i is a sum of terms of the form (1) B^i, and (2) $H^{i-d,k} \cdot B^k$, and (3) $f(x,H) * \nu$ with $f(x,u) = h'^i(x,ux) - h^i(x)$ or $f(x,u) = h'^i(x,ux) - \sum_{k \leq d} u^{i-d,k} h^k(x)$. Due to the properties of the truncation functions h and h', these functions f meet 5.23 and so the convergence in variation of the terms of type (3) follows from 5.29. By hypothesis $B^i(\alpha_n) \to B^i(\alpha)$ in variation on each interval $[0, t]$. Finally

$$\operatorname{Var}(H^{i-d,k}(\alpha_n) \cdot B^k(\alpha_n) - H^{i-d,k}(\alpha) \cdot B^k(\alpha))_t$$

5.30
$$\leq K \operatorname{Var}(B^k(\alpha_n) - B^k(\alpha))_t + |H^{i-d,k}(\alpha_n) - H^{i-d,k}(\alpha)| \cdot \operatorname{Var}(B^k(\alpha))_t$$

also tends to 0 (use Lebesgue convergence theorem for the last term), and thus the first convergence in 5.28 follows.

d) Finally we consider 5.8: $\tilde{C}'^{ij}$ is a sum of terms of the form $\tilde{C}^{ij}$, $H^{i-d,k} \cdot \tilde{C}^{kj}$, $(H^{i-d,k} H^{j-d,l}) \cdot \tilde{C}^{kl}$ (for which the convergence in variation is proved as in (c)), of terms of the form $f(x,H) * \nu$ with $f(x,u) = h'^i h'^j(x,ux) - h^i h^j(x)$, or $= h'^i h'^j(x,ux) - \sum_{k \leq d} u^{i-d,k}(h^k h^j)(x)$, or $= h'^i h'^j(x,ux) - \sum_{k,l \leq d} u^{i-d,k} u^{j-d,l}(h^k h^l)(x)$, which all converge in variation because of 5.29, and finally of terms of the form $\sum_{s \leq \cdot} \Delta U_s \Delta V_s$ where U and V are two processes among B^i, B'^i, $\sum_{k \leq d} H^{i-d,k} \cdot B^k$. Hence in all cases, $\operatorname{Var}[U(\alpha_n) - U(\alpha)]_t \to 0$ and the same for V (because of (c)), and it is easy to deduce that

$$\operatorname{Var}\left(\sum_{s\leq\cdot} \Delta U_s(\alpha_n)\Delta V_s(\alpha_n) - \sum_{s\leq\cdot} \Delta U_s(\alpha)\Delta V_s(\alpha)\right)_t \to 0.$$

This finishes to prove the convergence in 5.28. $\square$

The second application of Lemma 5.22 concerns the convergence of the characteristics (B'^n, C'^n, ν'^n) and $\tilde{C}'^n$ of $Z^n = (X^n, H^n \cdot X^n)$, as defined by 5.2 and 5.8 again (with the same truncation function h' as above) toward (B', C', ν') and $\tilde{C}'$.

5.31 **Proposition.** *Assume 5.21 and*

(i) *either 3.11 and 5.13 hold,*

(i′) *or $\mathscr{L}(X^n) \to P$ weakly and 5.10 holds for P-almost all α, and 5.17 holds.*

Then [Var-β], [Var-γ], [Var-δ] *yield*

[Var-β]′ $\operatorname{Var}(B'^{n,i} - B'^i \circ X^n)_t \xrightarrow{P} 0$ for all $t \geq 0, i \leq d + m$

[Var-γ]′ $\operatorname{Var}(\tilde{C}'^{n,ij} - \tilde{C}'^{ij} \circ X^n)_t \xrightarrow{P} 0$ for all $t \geq 0, i, j \leq d + m$

[Var-δ]′ $\operatorname{Var}(g' * \nu'^n - (g' * \nu') \circ X^n)_t \xrightarrow{P} 0$ for all $t \geq 0$, $g \in C'_2(\mathbb{R}^{d+m})$.

Proof. a) We will apply 5.22 to the measures $\rho^n = \nu^n$, $\rho'^n = \nu \circ X^n$ and to the processes $K^n = H^n$, $K'^n = H \circ X^n$. 5.22(i) is implied by [Var-δ], and 5.22(ii) by (i) above. Under (i′) we have 5.22(ii′.2); furthermore $\mathscr{L}(X^n) \to P$ and the fact that 5.10 holds P-a.s. imply that $\mathscr{L}((g * \nu_t) \circ X^n) \to \mathscr{L}(g * \nu_t | P)$ for all $g \in C_1(\mathbb{R}^d)$. Since $g * \nu_t < \infty$, it is easy to deduce that $\rho'^n = \nu \circ X^n$ satisfies 5.22(ii′.1).

Therefore in all cases we can apply 5.22, and if f has 5.23 we obtain

5.32 $$\operatorname{Var}[f(x, H^n) * \nu^n - (f(x, H) * \nu) \circ X^n]_t \xrightarrow{P} 0.$$

b) At this stage, we can essentially reproduce the proof of 5.27. For getting [Var-δ]′ we apply 5.32 to $f(x, u) = g'(x, ux)$. For [Var-β]′ we obtain the convergence in variation (in measure) for the terms of the form (1) (resp. (3)) in part (c) of the proof of 5.27 by hypothesis (resp. by 5.32). As for the terms of the form (2), we have to replace 5.30 by

$$\operatorname{Var}[H^{n,i-d,k} \cdot B^{n,k} - (H^{i-d,k} \cdot B^k) \circ X^n]_t$$

5.33 $$\leq K \operatorname{Var}(B^{n,k} - B^k \circ X^n)_t + |H^{n,i-d,k} - H^{i-d,k} \circ X^n| \cdot \operatorname{Var}(B^k \circ X^n)_t.$$

The first term in 5.33 goes to 0 in measure by [Var-β]. Under (i), the second term in 5.33 is smaller than (recall that $\operatorname{Var}(B^k) \prec F$):

$$\int_0^t |H_s^{n,i-d,k} - H_s^{i-d,k} \circ X^n| \, dF_s,$$

which goes to 0 in measure by Lebesgue convergence theorem. Under (i′), the second term in 5.33 is smaller than

5.34 $$\left(\sup_{s\leq t} |H_s^n - H_s \circ X^n|\right) \operatorname{Var}(B^k \circ X^n)_t$$

and $\mathscr{L}(X^n) \to P$ and 5.10 yield that $\mathscr{L}(\mathrm{Var}(B^k \circ X^n)) \to \mathscr{L}(\mathrm{Var}(B^k)_t | P)$, while $\mathrm{Var}(B^k)_t < \infty$; hence $\lim_{N \uparrow \infty} \limsup_n P^n(\mathrm{Var}(B^k \circ X^n)_t > N) = 0$ and we deduce from 5.17 that 5.34 goes to 0 in measure as $n \uparrow \infty$. Therefore in all cases 5.33 goes to 0 in measure, and we deduce [Var-β]'.

It remains to prove [Var-γ]'. Convergence in variation for the parts of $\tilde{C}'^{n,ij}$ which are integrals with respect to $\tilde{C}^{n,kl}$ or with respect to v^n (see (d) in the proof of 5.22) is proved as above for [Var-β]'. It remains to consider the parts having the form $\sum_{s \le t} \Delta U_s^n \Delta V_s^n$ where U^n and V^n are of the form $B^{n,i}$, or $B'^{n,i}$, or $\sum_{k \le d} H^{n,i-d,k} \cdot B^{n,k}$. From what precedes we have $\mathrm{Var}(U^n - U \circ X^n)_t \xrightarrow{P} 0$ and $\mathrm{Var}(V^n - V \circ X^n)_t \xrightarrow{P} 0$ for all these U^n, V^n. It readily follows that

5.35 $$\mathrm{Var}\left\{\sum_{s \le \cdot} \Delta U_s^n \Delta V_s^n - \left(\sum_{s \le \cdot} \Delta U_s \Delta V_s\right) \circ X^n\right\}_t \xrightarrow{P} 0$$

(one can for example take subsequences such that $\mathrm{Var}(U^{n'} - U \circ X^{n'})_t \to 0$ and $\mathrm{Var}(V^{n'} - V \circ X^{n'})_t \to 0$ a.s., and then deduce that 5.35 holds a.s. as well for these subsequences). This finishes the proof of [Var-γ]'. □

3. Now it remains to apply the results of Sections 2 and 3. To this end, we consider the canonical space $\Omega' = \mathbb{D}(\mathbb{R}^d \times \mathbb{R}^m)$ with the canonical filtration $\mathbf{F}'$ and σ-field $\mathscr{F}'$, and the canonical process Z. We write $Z = (X, Y)$ where $X = (X^i)_{i \le d} = (Z^i)_{i \le d}$ and $Y = (Y^i)_{i \le m} = (Z^{d+i})_{i \le m}$.

B', C', v', H', $\tilde{C}'$ have a natural extension to Ω': $B' \circ Z = B' \circ X$, etc... Then we have $(B', \tilde{C}', v') \circ Z^n = (B', \tilde{C}', v') \circ X^n$ on $\mathscr{B}^n$.

Proof of Theorem 5.12. As said before, it is not a restriction to assume 5.21. We will apply Theorem 3.35 to Z^n. Since (B, C, v) meet 3.7 and 3.11, the boundedness of H and the explicit forms 5.2 and 5.8 readily imply that (B', C', v') also meet 3.7 and 3.11.

Proposition 5.31 gives [Var-β]', [Var-γ]', [Var-δ]', which are stronger than [Sk-β_7]', [Sk-γ_7]', [Sk-δ_7]' analogous to 3.2, but with Z^n, B'^n, $\tilde{C}'^n$, v'^n and B', $\tilde{C}'$, v' instead of X^n, B^n, $\tilde{C}^n$, v^n and B, $\tilde{C}$, v.

Let $\eta' = \eta \otimes \varepsilon_0$ on $\mathbb{R}^d \times \mathbb{R}^m$. Then $\mathscr{L}(Z_0^n) \to \eta'$ weakly because $(H^n \cdot X^n)_0 = 0$. For all $t \ge 0$, $g' \in C_2(\mathbb{R}^{d+m})$, Proposition 5.27 yields that $\alpha \leadsto B'_t(\alpha)$, $\tilde{C}'_t(\alpha)$, $g' * v'_t(\alpha)$ are continuous on $\mathbb{D}(\mathbb{R}^d)$, and so their extensions to $\mathbb{D}(\mathbb{R}^{d+m})$ are also obviously continuous for the Skorokhod topology.

Finally, let $P' \in \mathscr{S}(\sigma(Z_0), Z | \eta'; B', C', v')$. In virtue of Proposition 5.6, we have $Y = H \cdot X$ on $(\Omega', \mathscr{F}', \mathbf{F}', P')$, and of course X is a semimartingale on this space with characteristics (B, C, v) and initial distribution η. Since $(B, C, v) \circ Z = (B, C, v) \circ X$ we easily deduce that $\mathscr{L}(X | P')$ is a solution to the martingale problem $\mathscr{S}(\sigma(X_0), X | \eta; B, C, v)$. Then the uniqueness assumption yields $\mathscr{L}(X | P') = P$. Moreover $H_t \circ Z = H_t \circ X$: then the law $\mathscr{L}((X, H \cdot X) | P)$ is exactly P', and we conclude that P' is the unique solution to $\mathscr{S}(\sigma(Z_0), Z | \eta'; B', C', v')$.

Hence all hypotheses of Theorem 3.35 are met by Z^n, $(B'^n, \tilde{C}'^n, v'^n)$ and $(B', \tilde{C}', v')$. Therefore $\mathscr{L}(Z^n) \to P'$, and the claim is proved. □

Proof of Theorem 5.16. Again we can suppose that 5.21 holds.

a) Our first aim is to prove that the sequence $\{Z^n\}$ is tight, and for this we will use Theorem VI.4.18. Recall that $\mathscr{L}(X^n) \to P$, so the sequence $(Z_0^n = (X_0^n, 0))$ is obviously tight in $\mathbb{R}^{d+m}$.

Since 5.10 holds P-a. s., we deduce from 5.27 that $\mathscr{L}(B' \circ X^n)$ converges to the the law of B' on $(\Omega, \mathscr{F}, P)$, while 5.31 yields that $\sup_{s \le t} |B_s'^n - B_s' \circ X^n| \xrightarrow{P} 0$ for all t, so $\mathscr{L}(B'^n)$ also converges to the law of B' on $(\Omega, \mathscr{F}, P)$. Since X is quasi-left continuous under P, B_t' is P-a. s. continuous in t, and thus the sequence B'^n is C-tight. One proves similarly that the sequences $\{\tilde{C}'^n\}$ and $\{g' * \nu'^n\}$ for all $g' \in C_2(\mathbb{R}^{d+m})$ are C-tight.

Finally, the necessary part in VI-4.18 implies that ν^n meets VI.4.19. Since $|H^n| \le K$, we readily deduce from 5.2 that ν'^n meets VI.4.19 as well. Therefore the sufficient part of VI.4.18 applies to the sequence $\{Z^n\}$ which thus is tight.

b) It remains to prove that if a subsequence, still denoted $\tilde{P}^n = \mathscr{L}(Z^n)$, weakly converges to a limit $\tilde{P}$ on $\mathbb{D}(\mathbb{R}^{d+m})$, then $\tilde{P} = \mathscr{L}((X, H \cdot X)|P)$. For this we will apply Theorem 2.22. Firstly, 5.31 implies that [Sup-β_7], [Sup-γ_7], [Sup-$\delta_{7,1}$] are met by B'^n, $\tilde{C}'^n$, ν'^n, and so 2.22(i) is met (see Remark 2.23).

Secondly, 5.10 holds P-a. s., so 5.27 yields in particular that $\alpha \rightsquigarrow B'(\alpha)$, $\tilde{C}'(\alpha)$ $g' * \nu'(\alpha)$ (for $g' \in C_2(\mathbb{R}^{d+m})$) are continuous for the Skorokhod topology from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^{d+m})$, $\mathbb{D}(\mathbb{R}^{(d+m)^2})$, $\mathbb{D}(\mathbb{R})$ respectively, at P-almost all points α. Moreover, since X is quasi-left continuous under P, the functions $t \rightsquigarrow B_t'(\alpha)$, $\tilde{C}_t'(\alpha)$, $g' * \nu_t'(\alpha)$ are continuous, for P-almost all α. Then we deduce from VI.2.2 that, P-a. s., the functions $\alpha \rightsquigarrow (\alpha, B'(\alpha), \tilde{C}'(\alpha))$ and $\alpha \rightsquigarrow (\alpha, g' * \nu'(\alpha))$ are continuous from $\mathbb{D}(\mathbb{R}^d)$ into $\mathbb{D}(\mathbb{R}^{d+(d+m)+(d+m)^2})$ and $\mathbb{D}(\mathbb{R}^{d+1})$. It readily follows that their extensions to $\mathbb{D}(\mathbb{R}^{d+m})$ are $\tilde{P}$-a. s. continuous, provided

5.36 $$\mathscr{L}(X|\tilde{P}) = P.$$

But $\mathscr{L}(Z^n) \to \tilde{P}$, so $\mathscr{L}(X^n) \to \mathscr{L}(X|\tilde{P})$. Since $\mathscr{L}(X^n) \to P$ by hypothesis, 5.36 holds. Hence 2.22(ii) is met by (B', C', ν') $\tilde{P}$-a. s.

Therefore Theorem 2.22 applies: under $\tilde{P}$, Z is a semimartingale with characteristics (B', C', ν'). Furthermore $Z_0 = (X_0, 0)$ $\tilde{P}$-a. s. is obvious. Hence 5.6 implies that $Y = H \cdot X$ on $(\Omega', \mathscr{F}', \mathbf{F}', P)$, which in turn implies that $\mathscr{L}(Z = (X, Y) = (X, H \cdot X)|\tilde{P}) = \mathscr{L}((X, H \cdot X)|P)$ (as in the previous proof), and we are finished. □

6. Stability for Stochastic Differential Equation

Here we give a very useful result in practice, concerning the so-called "stability" for stochastic differential equations. This result hinges upon the convergence of stochastic integrals proved in § VI.6.c and on the P-UT property. We begin with some auxiliary results which have interest on their own.

§ 6a. Auxiliary Results

The first auxiliary result is an extension of Doob's inequality.

6.1 Proposition. *For all locally square-integrable martingales M and stopping times T,*

6.2 $$E(\sup_{s<T}(M_s - M_0)^2) \leq 4\,(E([M,M]_{T-}) + E(\langle M,M\rangle_{T-}))\,.$$

Proof. By localization it is enough to prove the result when M is a square-integrable martingale, with furthermore $M_0 = 0$.

Let μ be the measure associated with the jumps of M, and ν its compensator, and $J = \{(\omega, s) : \nu(\omega; \{s\} \times \mathbb{R}) > 0\}$. Set $A = \Delta M_T 1_{[T,\infty[}$ and let B be the compensator of A and $N = A - B$. Combining I.2.31 and I.3.21 yields that B is continuous. Then set $U = 1_J \cdot N$ and $V = 1_{J^c} \cdot N$. We have $U = 1_J \cdot A - 1_J \cdot B$ and, since B is continuous and J is at most countable in time, $1_J \cdot B = 0$: therefore $U = 1_J \cdot A$ and $V = 1_{J^c} \cdot A - B$. At this point, we set $N' = M - N$. If $t < T$ then $A_t = 1_{J^c} \cdot A_t = 0$, hence $M_t = N'_t - B_t = N'_t + V_t$ and applying I.1.43 and I.4.50 to $N' + V$ yields

$$E(\sup_{s<T} M_s^2) \leq E(\sup_{s\leq T}(N'_s + V_s)^2) \leq 4E([N' + V, N' + V]_T).$$

Now $\Delta N'_t = \Delta M_t$ if $t < T$ and $\Delta N'_T = 0$, hence $[N', N']_T = [M, M]_{T-}$; we also have $\Delta V_t = 0$ if $t \neq T$, hence $[N' + V, N' + V] = [N', N'] + [V, V]$; finally $[V, V]_T \leq (1_{J^c} \cdot [M, M])_T$. Therefore

$$\begin{aligned} E(\sup_{s<T} M_s^2) &\leq 4E([M,M]_{T-}) + 4E((1_{J^c} \cdot [M,M])_T) \\ &\leq 4E([M,M]_{T-}) + 4E((1_{J^c} \cdot \langle M,M\rangle_T). \end{aligned}$$

Since $1_{J^c} \cdot \langle M, M\rangle$ is continuous we have $E((1_{J^c} \cdot \langle M, M\rangle)_T) \leq E(\langle M, M\rangle_{T-})$. □

The second result is a Gronwall type inequality:

6.3 Lemma. *Let A and B be two increasing processes, with A adapted having $E(A_\infty) < \infty$, and B having $B_\infty \leq K$ identically for some constant K. Suppose that for each stopping time T we have*

6.4 $$E(A_{T-}) \leq \alpha + E(A_- \cdot B_{T-}).$$

for some constant α. Then $E(A_\infty) \leq \alpha e^K$.

Proof. Set $C_t = \inf(s : B_s \geq t)$. Each C_t is a stopping time, so (i) and I.3.12(1) yields

$$E(A_{(C_t)-}) \leq \alpha + E\left(\int_0^\infty A_{s-} 1_{\{s<C_t\}} dB_s\right) = \alpha + E\left(\int_0^\infty A_{(C_u)-} 1_{\{C_u<C_t\}} du\right)$$
$$\leq \alpha + E\left(\int_0^t A_{(C_u)-} du\right) = \alpha + \int_0^t E(A_{(C_u)-}) du.$$

Then the usual Gronwall lemma yields $E(A_{C_t-}) \leq \alpha e^t$. Since $C_K = \infty$ by our assumption on B, we obtain the result by letting $t = K$. $\square$

Finally we give a complement on stable convergence in law (see § VIII.5c): we have a probability space $(\Omega, \mathscr{F}, P)$ endowed with a sequence (Z_n) of variables taking their values in a Polish space $(E, \mathscr{E})$, and another E–valued variable Z. Recall that (Z_n) converges $\mathscr{F}$–stably to Z if $E[Yf(Z_n))] \to E[Yf(Z)]$ for every bounded $\mathscr{F}$–measurable Y and every continuous bounded function f on E.

6.5 **Lemma.** *If (Z_n) converges $\mathscr{F}$–stably to Z, then $Z_n \xrightarrow{P} Z$.*

Proof. Take $q \in \mathbb{N}^\star$. We can find a compact K in E such that $P(Z \notin K) \leq \frac{1}{q}$. Since E is Polish, we can find finitely many points $x_1, \dots, x_p$ in K such that any x in K satisfies $d(x, x_i) < \frac{1}{q}$ for some i (d is a metric on E). If $d(Z_n, Z) \geq \frac{3}{q}$ and $Z \in K$, there exists i such that $d(x_i, Z) < \frac{1}{q}$ and $d(x_i, Z_n) > \frac{2}{q}$, hence $f_i(Z_n) \geq 1$ where $f_i(x) = q[(d(x_i, x) - \frac{1}{q})^+ \bigwedge 1]$. Therefore

$$P(d(Z, Z_n) \geq \frac{3}{q}) \leq P(Z \notin K) + \sum_{i=1}^p E[1_{\{d(x_i, Z)<1/q\}} f_i(Z_n)].$$

Observe that each f_i is continuous and bounded. Hence the last sum above converges to $\sum_{i=1}^p E[1_{\{d(x_i,Z)<1/q\}} f_i(Z)]$ as $n \to \infty$, and this limit equals 0: hence $\limsup_n P(d(Z, Z_n) \geq 3/q) \leq 1/q$, which gives the result. $\square$

§ 6b. Stochastic Differential Equations

We are not going to give a complete theory of stochastic differential equations (SDE in short) driven by semimartingales. Some elements have been given in Section III.2c, without proofs. Here we consider a slightly different sort of SDEs, for which we refer for example to the book [286] of Protter.

The driving term is a d–dimensional semimartingale X, given on a stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$, and the solution will be a q–dimensional process Y. There are two more ingredients: an "initial condition" which is in fact a q–dimensional càdlàg adapted process Z, and a "coefficient" which is a function $f : \mathbb{R}^q \to \mathbb{R}^q \otimes \mathbb{R}^d$, and the equation is

6.6 $$Y = Z + f(Y_-) \cdot X$$

(or, componentwise, as $Y^i = Z^i + \sum_{j=1}^d f(Y_-)^{ij} \cdot X^j$ for $i = 1, \dots, q$).

Consider the following condition:

6.7 *Local Lipschitz and linear growth condition:* there are constants C and C_i for $i = 1, 2, \ldots$, such that $|f(x)| \leq C(1 + |x|)$ for all $x \in \mathbb{R}^d$ and that $|f(x) - f(y)| \leq C_i|x - y|$ for all $x, y \in \mathbb{R}^q$ with $|x|, |y| \leq i$. □

Under 6.7 it is well known that Equation 6.6 has a unique solution, the uniqueness meaning that any two solutions agree outside a P–null set. Furthermore the solution can be approximated by "explicit" functions of the processes X and Z, so in fact we can write $Y = g(X, Z)$ P–a.s., where g is a Borel function (of course not explicit) from $\mathbb{D}(\mathbb{R}^{d+q})$ into $\mathbb{D}(\mathbb{R}^q)$: therefore if we have another basis $\mathscr{B}'$ on which are defined a càdlàg adapted process Z' and a semimartingale X', and if further the pair (X, Z) on $\mathscr{B}$ and the pair (X', Z') on $\mathscr{B}'$ have the same law, then $Y' = g(X', Z')$ is indeed the solution of the equation $Y' = Z' + f(Y'_-) \cdot X'$ on $\mathscr{B}'$.

§ 6c. Stability

The setting is as follows: first we have Equation 6.6 on the basis $\mathscr{B}$ (with X, Z and the solution Y, and f satisfying 6.7). Next, for each $n \in \mathbb{N}$ we have a stochastic basis $\mathscr{B}^n = (\Omega^n, \mathscr{F}^n, \mathbf{F}^n, P^n)$ endowed with a d–dimensional semimartingale X^n and a q–dimensional càdlàg adapted process Z^n. We also have functions $f_n : \mathbb{R}^q \to \mathbb{R}^q \otimes \mathbb{R}^d$, satisfying 6.7, and such that each equation

6.8 $$Y^n = Z^n + f_n(Y^n_-) \cdot X^n$$

admits a (P^n–a.s. unique) solution Y^n.

Our aim is to prove the following result:

6.9 **Theorem.** *Assume that the functions f_n satisfy 6.7 with constants C and C_i that do not depend on n, and also that $f_n \to f$ pointwise as $n \to \infty$. Assume also that the sequence (X^n) is P-UT. Then, with Y^n being the unique solution of 6.8,*

(a) If $(X^n, Z^n) \xrightarrow{\mathscr{L}} (X, Z)$, then $(X^n, Z^n, Y^n) \xrightarrow{\mathscr{L}} (X, Z, Y)$.

(b) If further $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ for all n (the filtrations $\mathbf{F}^n$ may be different one from the other, and from $\mathbf{F}$), and if $(X^n, Z^n) \xrightarrow{P} (X, Z)$ (convergence in measure, for the Skorokhod topology), *then* $(X^n, Z^n, Y^n) \xrightarrow{P} (X, Z, Y)$.

This result is not the best possible: one could consider "non-homogeneous" coefficients $f_n(s, x)$ depending (continuously) on time; one could replace the coefficients f_n depending only on the "current" value of the process by coefficients depending on the whole past of the solution, etc... But Theorem 6.9 remains reasonably simple and covers most applications.

The proof is divided into a number of steps. To ease the notation, write $U^n = (X^n, Z^n)$ and $V^n = (X^n, Z^n, Y^n)$, as well as $U = (X, Z)$ and $V = (X, Z, Y)$.

(a) *Localization-1.* We suppose here that Theorem 6.9 holds when in addition the processes U^n and U are uniformly bounded by a constant K, and we deduce that it holds in general.

For each $p \geq 1$ let $g_p : \mathbb{R}^{d+q} \to \mathbb{R}^{d+q}$ be a C^2 bounded function, with $g_p(x) = x$ when $|x| \leq p$. Set $U^n(p) = (X^n(p), Z^n(p)) := g_p(U^n)$ and $U(p) = (X(p), Z(p)) := g_p(U)$. The processes $U^n(p)$ and $U(p)$ are uniformly bounded for each fixed p; further since g_p is C^2, the processes $X^n(p)$ and $X(p)$ are semimartingales, and $Z^n(p)$ and $Z(p)$ are càdlàg adapted, so the equations

6.10
$$Y^n(p) = Z^n(p) + f_n(Y^n(p)_-) \cdot X^n(p), \qquad Y(p) = Z(p) + f(Y(p)_-) \cdot X(p),$$

have unique solutions $Y^n(p)$ and $Y(p)$, and we write $V^n(p) = (U^n(p), Y^n(p))$ and $V(p) = (U(p), Y(p))$. Since g_p is continuous, we have $U^n(p) \xrightarrow{\mathscr{L}} U(p)$ in case (a) (resp. $U^n(p) \xrightarrow{P} U(p)$ in case (b)). So our assumption yields that for any given p we have

6.11 $\quad V^n(p) \xrightarrow{\mathscr{L}} V(p) \quad$ in case (a) $\quad$ (resp. $\quad V^n(p) \xrightarrow{P} V(p) \quad$ in case (b)).

In both cases the sequence (U^n) is tight, so Theorem VI.3.21 yields

6.12
$$\lim_p \sup_n P^n(T_p^n \leq N) = 0$$

for all $N > 0$, if $T_p^n = \inf(t : |U_t^n| \geq p)$. But $U^n = U^n(p)$ on the interval $[\![0, T_p^n[\![$, so obviously $Y^n = Y^n(p)$ on the same interval. Similarly $Y = Y(p)$ on $[\![0, T_p]\!]$, where $T_p = \inf(t : |U_t| \geq p)$. Hence for all $N > 0$ and $\varepsilon > 0$ there exists p_0 such that

6.13
$$\begin{aligned} &\inf_n P^n(V^n(p)_t = V_t^n \ \ \forall t \in [0, N]) \geq 1 - \varepsilon, \\ &P(V(p)_t = V_t \ \ \forall t \in [0, N]) \geq 1 - \varepsilon. \end{aligned}$$

This and 6.11 obviously imply the desired result for the original sequence (X^n, Z^n, Y^n) in both cases (a) and (b).

(b) *Some notation.* In view of the first step we assume from now on that U^n and U are bounded by a constant, say A. We take a truncation function for X^n and X which coincides with the identity on the set $\{x \in \mathbb{R}^d : |x| \leq 2A\}$. Since $|\Delta X^n| \leq 2A$, we have

6.14
$$X^n = X_0^n + B^n + M^n,$$

where B^n is the first characteristic, and M^n is a locally square-integrable martingale with $\mathbb{R}^d \otimes \mathbb{R}^d$–valued quadratic variation $[M^n, M^n]$ and angle bracket $\widetilde{C}^n = \langle M^n, M^n \rangle$. We also consider the following adapted increasing processes

6.15
$$G^n = \sum_{i=1}^d \left(\mathrm{Var}(B^{n,i} + [M^{n,i}, M^{n,i}] + \langle M^{n,i}, M^{n,i} \rangle\right).$$

Note that $\Delta G^n \leq A' = A + 2A^2$. In a similar fashion we associate the process G with X.

(c) *Localization-2* . We suppose here that Theorem 6.9 holds when in addition not only the processes U^n and U are bounded by the constant A, but also G^n and G are bounded by a constant K, and we deduce that it holds when only the U^n's and U are bounded (i.e. in the setting of Step (b) above).

Let $T_p^n = \inf(t : G_t^n \geq p)$, so $G_{T_p^n}^n \leq p + A'$. Set $X^n(p)_t = X_{t\wedge T_p^n}^n$ and $Z^n(p)_t = Z_{t\wedge T_p^n}^n$; similarly we associate T_p, $X(p)$ and $Z(p)$ with X, Z and G. The solutions of 6.10 are $Y^n(p)_t = Y_{t\wedge T_p^n}^n$ and $Y(p)_t = Y_{t\wedge T_p}$, and our assumption yields 6.11 because $G^n(p)_\infty \leq K = p+A'$ and $G(p)_\infty \leq K$ (with obvious notation). But by VI.6.15 the P-UT property of the sequence (X^n) implies 6.12, so we have 6.13 and we conclude as in the previous step.

(d) *Some estimates-1.* In view of Step (c) we assume now that $|U^n| \leq A$, $|U| \leq A$ and $G_\infty^n \leq K$ identically for some finite A, K.

Consider any $\mathbb{R}^q \times \mathbb{R}^d$–valued predictable and locally bounded process $H^n = (H^{n,ij})$ and any stopping time S^n on $\mathscr{B}^n$. We have by 6.14

$$E^n(\sup_{s<S^n} |H^n \cdot X_s^n|^2) \leq 2E^n(\sup_{s<S^n} |H^n \cdot B_s^n|^2) + 2E^n(\sup_{s<S^n} |H^n \cdot M_s^n|^2).$$

By 6.15 and $G_\infty^n \leq K$ and Hölder inequality we have

$$\sup_{s<ST^n} |H^n \cdot B_s^n|^2 \leq \sum_{i=1}^q \sum_{i=1}^d K(H^{n,ij})^2 \cdot G_{S^n-}^n \leq qdK\ |H^n|^2 \cdot G_{S^n-}^n.$$

By 6.15 again and Proposition 6.1, we also get

$$E^n(\sup_{s<S^n} |H^n \cdot M_s^n|^2 = \sum_{i=1}^q \sum_{i=1}^d E^n(\sup_{s<S^n} |H^{n,ij} \cdot M_s^{n,j}|^2) \leq 8qd\ E^n(|H^n|^2 \cdot G_{S^n-}^n).$$

Therfore we obtain, with $K' = 2qd(K + 8)$:

6.16 $$E^n(\sup_{s<S^n} |H^n \cdot X_s^n|^2) \leq K'E^n(|H^n|^2 \cdot G_{S^n-}^n).$$

(e) *Some estimates-2.* Now we put $S_z^n = \inf(t : |Y_t^n| \geq z)$. Recall the constant C in 6.7, which works for all functions f_n. Since $|Z^n| \leq A$ and $G_\infty^n \leq K$, we deduce from 6.8 and 6.16 that, as soon as $z > A$ and with $D_t^n = \sup_{s\leq t} |Y_s^n|^2$,

$$E^n(D_{(S_z^n\wedge S^n)-}^n) \leq 2A^2 + 2K'CK + 2K'C\ E^n(D_-^n \cdot G_{(S_z^n\wedge S^n)-}^n)$$

for any stopping time S^n on $\mathscr{B}$. Then Lemma 6.3 and the property $D_{(S_z^n)-}^n \leq z$ give $E^n(D_{(S_z^n)-}^n) \leq K'' := (2A^2 + 2K'CK)\ e^{2K'CK}$. Now, $D_{(S_z^n)-}^n$ increases to D_∞^n as $z \uparrow \infty$, so we deduce that for all n,

6.17 $$E^n(\sup_s |Y_s^n|^2) = E^n(D_\infty^n) \leq K''.$$

(f) *Some estimates-3.* The sequence (U^n) converges in law to U^∞, so by Proposition VI.7.3(b) we can associate with it some numbers $a(n,p,i) \in (\frac{1}{2p}, \frac{1}{p}]$ and $\varrho_p^N > 0$ (where p and N are integers bigger than 1), so that the stopping times $T_i^{n,p}$ defined by VI.7.4 with U^n in place of X^n satisfy VI.7.1 and VI.7.2 (with U^n in place of X^n as well). We set

$$X_t^{n,p} = X^n_{T_i^{n,p}}, \quad Z_t^{n,p} = Z^n_{T_i^{n,p}}, \quad Y_t^{n,p} = Y^n_{T_i^{n,p}}, \qquad \text{if } T_i^{n,p} \le t < T_{i+1}^{n,p}.$$

If $T_i^{n,p} \le t < T_{i+1}^{n,p}$ we have

$$Y_t^n - Y_t^{n,p} = Z_t^n - Z_t^{n,p} + \int_{T_i^{n,p}}^t [f_n(Y_{s-}^n) - f_n(Y_{s-}^{n,p})]dX_s^n + f_n(Y^n_{T_i^{n,p}})(X_t^n - X_t^{n,p}).$$

We have $\left|\int_{T_i^{n,p}}^t H_{s-}^n dX_s^n\right| \le 2\sup_{s\le t}\left|\int_0^s H_{s-}^n dX_s^n\right|$. By VI.7.4, $|U^n - U^{n,p}| \le \frac{1}{p}$. Therefore $D_t^{n,p} = \sup_{s\le t}|Y_s^n - Y_s^{n,p}|^2$ satisfies (use 6.7):

$$D_t^{n,p} \le \frac{3}{p^2} + 12 \sup_{s\le t}\left[(f_n(Y_-^n) - f_n(Y_-^{n,p}))\cdot X_s^n\right]^2 + \frac{6C^2}{p^2}(1 + D_t^n).$$

By 6.7 again we have $|f_n(Y_-^n) - f_n(Y_-^{n,p})|^2 \le C_i^2 D_-^{n,p}$ on $[\![0, S_i^n]\!]$ (recall the notation S_z^n of (e)). We then deduce from 6.16 and 6.17 that, for any stopping time S^n on $\mathscr{B}$,

$$E^n(D^{n,p}_{(S_z^n \wedge S^n)-}) \le \frac{3}{p^2}(1 + 2C^2(1 + K'')) + 12K'C_i^2\, E^n(D_-^{n,p}\cdot G^n_{S_i^n z \wedge S^n)-}).$$

Another application of Lemma 6.3 gives

6.18 $$E^n(D^{n,p}_{(S_i^n)-}) \le \frac{3}{p^1}(1 + 2C^2(1 + K''))\, e^{12K'C_i^2}.$$

(g) *Tightness of* (V^n). We are now ready to prove the tigthness of the sequence $(V^n = (U^n, Y^n))$, by using Proposition VI.7.3(a): it is clearly enough to prove that, separately, the two sequences (U^n) and (Y^n) satisfy VI.3.21(i) and VI.7.2, with the stopping times $T_i^{n,p}$ defined above. This is obvious for (U^n) by VI.7.3(b). That (Y^n) satisfies VI.3.21(i) follows from 6.17. Finally $\sup_i w(Y^n, [T_i^{n,p}, T_{i-1}^{n,p}) \cap [0,N]) \le (D^{n,p}_{(S_j^n)-})^{1/2}$ on the set $\{S_j^n > N\}$. Since $\{S_j^n \le N\} \subset \{D_\infty^n \ge j^2\}$, we get from 6.17 and 6.18:

$$\sup_n P^n(\sup_i w(Y^n, [T_i^{n,p}, T_{i-1}^{n,p}) \cap [0,N]) \ge \varepsilon)$$
$$\le \frac{3}{p^2\varepsilon^2}(1 + C^2(1 + K''))\, e^{12K'C_j^2} + \frac{K''}{j^2}$$

for all $j \ge 1$ and $\varepsilon > 0$. By choosing j first and p next, we see that the above goes to 0 as $p \to \infty$, which proves that (Y^n) satisfies VI.7.2: therefore the sequence (V^n) is tight.

(h) *Proof of 6.9(a).* In view of (g), it is enough to prove that for any subsequence of (V^n) which converges in law, the limit is V. So it is no restriction to assume that the sequence (V^n) itself converges in law to a limit $V' = (X', Z', Y')$ which is defined on some stochastic basis $\mathscr{B}' = (\Omega', \mathscr{F}', \mathbf{F}, P')$.

The pointwise convergence $f_n \to f$ and the fact that the f_n's are locally Lipschitz, uniformly in n, yield that $f_n \to f$ uniformly on each compact subset of $\mathbb{R}^q$. Since the Y^n's satisfy VI.3.21(i) we deduce that $f_n(Y^n) - f(Y^n) \xrightarrow{P} 0$, uniformly on each finite interval $[0, N]$: hence $(X^n, Z^n, Y^n, f_n(Y^n)) \xrightarrow{\mathscr{L}} (X', Z', Y', f(Y'))$. Since the sequence (X^n) is P-UT, we can apply Theorem VI.6.22(b) to obtain $(X^n, Z^n, Y^n, f_n(Y^n_-) \cdot X^n) \xrightarrow{\mathscr{L}} (X', Z', Y', f(Y'_-) \cdot X')$. But since $Y^n - Z^n - f_n(Y^n_-) \cdot X^n = 0$, we deduce that $Y' - Z' - f(Y'_-) \cdot X' = 0$ as well: in other words, Y' is the solution on $\mathscr{B}'$ of Equation 6.6.

Now, as recalled at the end of §6c, we can write $Y = g(X, Z)$ P–a.s. for some Borel map g from $\mathbb{D}(\mathbb{R}^{d+q})$ into $\mathbb{D}(\mathbb{R}^q)$, and we also have $Y' = g(X', Z')$ P'–a.s. Then clearly V on $\mathscr{B}'$ and V on $\mathscr{B}^\infty$ have the same law, and we are finished.

(i) *Proof of 6.9(b).* We suppose now that $(\Omega^n, \mathscr{F}^n, P^n) = (\Omega, \mathscr{F}, P)$ for all n, and that $U^n \xrightarrow{P} U$.

Let $B \in \mathscr{F}$ with $P(B) > 0$, and consider the measure $P_B(A) = \frac{P(A \cap B)}{P(B)}$. Since P_B is absolutely continuous w.r.t. P, each X^n is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}^n, P_B)$ (see III.3.13), and X is a semimartingale on $(\Omega, \mathscr{F}, \mathbf{F}, P_B)$. Since we even have $P_B \le \frac{1}{P(B)} P$, it is obvious from Definition VI.6.1 that the sequence (X^n) is P-UT w.r.t. P_B as well. Furthermore $f_n(Y^n_-) \cdot X^n_t$ is obtained as the limit in probability of Riemann sums (see I.4.44) which do not depend on the underlying measure: so clearly the stochastic integral $f_n(Y^n_-) \cdot X^n$ under P is also a version of the stochastic integral w.r.t. P_B, and similarly for $f(Y_-) \cdot X$. Therefore Y^n (resp. Y) is the solution of 6.8 (resp. 6.6) under P_B as well as under P. Further, since $U^n \xrightarrow{P} U$, we also have $U^n \xrightarrow{\mathscr{L}} U$ under P_B.

Therefore all of (h) remains true here, for P and for P_B as well, so $V^n \xrightarrow{\mathscr{L}} V$ under P_B. Thus for any continuous bounded function Φ on $\mathbb{D}(\mathbb{R}^{d+q+q})$ we have (E_B being the expectation w.r.t. P_B):

$$E_B(\Phi(V^n)) \to E_B(\Phi(V^\infty)),$$

which trivially yields

$$E(1_B \Phi(V^n)) \to E(1_B \Phi(V^\infty)).$$

This being true for all $B \in \mathscr{F}$, the equivalence (i)$\Leftrightarrow$(iv) in VIII.5.33 yields that (V^n) converges $\mathscr{F}$–stably to V, and we conclude by Lemma 6.5. $\square$

7. Stable Convergence to a Progressive Conditional Continuous PII

Here we give first a general result on the stable convergence in law for a sequence of semimartingales towards a "progressive conditional PII" in the sense of Section II.7. The main application of this general result is stable convergence in law for discretized processes, given in § 7b below: this has consequences for studying the Euler scheme approximation of the solution of an SDE, or higher order schemes as well; it also has a number of statistical applications, for diffusion processes observed at discrete times on a given regular grid.

§ 7a. A General Result

Since we are going to speak about stable convergence, we need everything to be defined on the same probability space. So we are given a stochastic basis $\mathscr{B} = (\Omega, \mathscr{F}, \mathbf{F}, P)$, which is endowed with a "reference" d–dimensional continuous local martingale Z, and $\mathscr{M}_b$ denotes the set of all bounded martingales on $\mathscr{B}$. Recall that for a q–dimensional locally square-integrable martingale M we denote by $\langle M, M\rangle$ the q^2–dimensional process $(\langle M^i, M^j\rangle)_{1\leq i,j\leq q}$: it takes its values in the set $\mathscr{S}_q$ of $q \times q$ symmetric nonnegative matrices and is an increasing process in this set.

Next, for each integer n we are given a filtration $\mathbf{F}^n = (\mathscr{F}_t^n)_{t\geq 0}$ on $(\Omega, \mathscr{F})$, which may differ from $\mathbf{F}$, and this sequence of filtrations has the following property:

7.1 For each n there is a d–dimensional square-integrable $\mathbf{F}^n$–martingale $Z(n)$ and, for each $N \in \mathscr{M}_b$, there is a bounded $\mathbf{F}^n$–martingale $N(n)$, such that for all $t \geq 0$:

(i) $\sup_{n,t,\omega}|N(n)_t(\omega)| < \infty$,

(ii) $\langle Z(n), Z(n)\rangle_t \xrightarrow{P} \langle Z, Z\rangle_t$

(the brackets on the left above are taken relatively to the filtration $\mathbf{F}^n$, of course). Further, for any finite family $(N^1, .., N^m)$ in $\mathscr{M}_b$, we have the following convergence for the Skorokhod topology on $\mathbb{D}(\mathbb{R}^{d+m})$:

(iii) $(Z(n), N^1(n), .., N^m(n)) \xrightarrow{P} (Z, N^1, .., N^m)$. □

In practice we encounter two situations: first, $\mathbf{F}^n = \mathbf{F}$ for all n, so 7.1 is obvious with $Z(n) = Z$ and $N(n) = N$. Second, $\mathscr{F}_t^n = \mathscr{F}_{[nt]/n}$, a situation looked at in § 7b.

Next, for each n we have a q–dimensional semimartingale X^n on $\mathscr{B}^n = (\Omega, \mathscr{F}, \mathbf{F}^n, P)$, with $X_0^n = 0$, and we denote by (B^n, C^n, ν^n) and $\widetilde{C}^n$ its characteristics and second modified characteristic, relative to a given *continuous* truncation function h on $\mathbb{R}^q$. Recall that

7.2 $$X^n = B^n + M^n + \check{X}^n, \quad \text{where} \quad \check{X}^n_t = \sum_{s\le t}(\Delta X^n_s - h(\Delta X^n_s)),$$

where M^n is an $\mathbf{F}^n$–local martingale with bounded jumps. The result writes as follows, with the notation of VIII.2.1, VIII.2.2 and VIII.3.5 (although the limits below are random, the conditions still make sense):

7.3 **Theorem.** *Assume 7.1, and let $\widetilde{C}$, G and B be continuous adapted processes on $\mathscr{B}$, null at* 0 *and taking values in $\mathscr{S}_q$, $\mathbb{R}^q \otimes \mathbb{R}^d$ and $\mathbb{R}^q$ respectively.*

a) Suppose that, for some countable subset D of $\mathbb{R}_+$, we have

[Sup-β_5] $$\sup_{s\le t} |B^n_s - B_s| \xrightarrow{P} 0 \qquad \text{for all } t > 0,$$

[γ_5-D] $$\widetilde{C}^n_t \xrightarrow{P} C_t \qquad \text{for all } t \in D,$$

[$\hat{\delta}_5$] $$\nu^n([0,t] \times \{|x| > \varepsilon\}) \xrightarrow{P} 0 \qquad \text{for all } t > 0, \varepsilon > 0.$$

Suppose also that for all $N \in \mathscr{M}_b$ orthogonal to all components of Z, we have (the bracket in the left sides below being relative to the filtration $\mathbf{F}^n$):

7.4 $$G^n_t := \langle M^n, Z(n)\rangle_t \xrightarrow{P} G_t \qquad \textit{for all } t \in D,$$

7.5 $$V(N)^n_t := \langle M^n, N(n)\rangle_t \xrightarrow{P} 0 \qquad \textit{for all } t \in D.$$

Then there is a very good extension $\widetilde{\mathscr{B}}$ of $\mathscr{B}$ (see II.7.1) and a continuous Z–biased $\mathscr{F}$–progressive conditional martingale PII X' on this extension (see II.7.8) with

7.6 $$\langle X', X'\rangle = \widetilde{C}, \qquad \langle X', Z\rangle = G,$$

such that (X^n) converges $\mathscr{F}$–stably to $X = B + X'$.

b) Assume further that $d\langle Z^i, Z^i\rangle_t \ll dt$ and $dF^{ii}_t \ll dt$. There are predictable processes u, v, w on $\mathscr{B}$, with values in $\mathbb{R}^q \otimes \mathbb{R}^d$, $\mathbb{R}^d \otimes \mathbb{R}^d$ and $\mathbb{R}^q \otimes \mathbb{R}^q$ respectively, such that with matrix notation (u^* is the transpose of u):

$$\langle Z, Z\rangle_t = \int_0^t u_s u^*_s ds, \quad G_t = \int_0^t u_s v_s v^*_s ds,$$

$$\widetilde{C}_t = \int_0^t (u_s v_s v^*_s u^*_s + w_s w^*_s) ds,$$

and the limit of X^n can be realized on the canonical q–dimensional Wiener extension of $\mathscr{B}$, with the canonical Wiener process W (see before II.7.12), as

$$X = B + u \cdot M + w \cdot W.$$

Proof. We divide the proof in a number of steps.

1) Let $H^n = \langle Z(n), Z(n)\rangle$ and $H = \langle Z, Z\rangle$, and set $K^n = \begin{pmatrix} H^n & G^{n*} \\ G^n & \widetilde{C}^n \end{pmatrix}$ and $K = \begin{pmatrix} H & G^* \\ G & \widetilde{C} \end{pmatrix}$. The processes K and K^n take their values in $\mathscr{S}_{d+q}$. By 7.1,

[γ_5-D] and 7.4, we have $K^n_t \xrightarrow{P} K_t$ for all $t \in D$, while K^n is a nondecreasing process in $\mathscr{S}^+_{d+q}$. So there is a version of K which is also a nondecreasing process in $\mathscr{S}^+_{d+q}$. Further K is continuous in time, so exactly as in the proof of VIII.2.17 we have indeed

$$[\text{Sup-}\gamma] \qquad \sup_{s\leq t}|K^n_s - K_s| \xrightarrow{P} 0 \qquad \text{for all } t \geq 0.$$

We can apply Proposition II.7.11: there is a very good extension $\widetilde{\mathscr{B}}$ of $\mathscr{B}$, on which we can define a continuous Z–biased $\mathscr{F}$–progressive conditional martingale PII X' satisfying 7.6.

Further $K = k \cdot A$ for some continuous adapted increasing process A and some predictable $\mathscr{S}^+_{d+q}$–valued process k on $\mathscr{B}$, and as seen in the proof of Proposition II.7.9 we have $k = zz^*$ with z given by II.7.10: under the additional assumptions of (b), we can take $A_t = t$, so the last claim of (b) follows from (a) and from Proposition II.7.12.

2) In this step we prove that 7.5 can be strenghtened as such:

$$[\text{Sup-}\gamma^N] \qquad \sup_{s\leq t}|V(N)^n_s| \xrightarrow{P} 0 \qquad \text{for all } t \geq 0.$$

In view of 7.5 it suffices to prove that

$$\textbf{7.7} \qquad \begin{array}{c} \forall\varepsilon,\ \forall\eta > 0,\ \forall T > 0,\ \exists\theta > 0,\ \exists n_0 < \infty,\ \forall n \geq n_0 \\ \Rightarrow \quad P(w_T(V(N)^n, \theta) > \eta) \leq \varepsilon, \end{array}$$

where $w_T(\alpha, \theta)$ is the modulus defined in VI.1.4. The fact that the predictable quadratic variation of $(M^n, N(n))$ is increasing for the strong order on $\mathscr{S}^+_{q+1}$ easily yields that for all $u > 0$ we have $2|V(N)^n_t - V(N)^n_s| \leq \frac{1}{u}|\widetilde{C}^n_t - \widetilde{C}^n_s| + u(\langle N(n), N(n)\rangle_t - \langle N(n), N(n)\rangle_s)$ if $s < t$, so $2w_T(V(N)^n, \theta) \leq w_T(\widetilde{C}^n, \theta)/u + u\langle N(n), N(n)\rangle_T$. Hence

$$P(w_T(V(N)^n, \theta) > \eta) \leq P(w_T(\widetilde{C}^n, \theta) > u\eta) + \frac{u}{\eta} E(N(n)^2_T).$$

But 7.1(i) implies $E(N(n)^2_T) \leq C$ for some constant C, and 7.7 holds with $\widetilde{C}^n$ instead of $V(N)^n$ by [Sup-γ]. Since $u > 0$ is arbitrary we deduce 7.7, and thus [Sup-γ^N] holds.

3) In this step we prove that if the sequence (M^n) converges $\mathscr{F}$–stably in law to X', as defined in the theorem, then (X^n) converges $\mathscr{F}$–stably to $X = B + X'$.

Observe first that $\sup_{s\leq t}|\check{X}^n_s| \xrightarrow{P} 0$ by VI-4.22 and [$\hat{\delta}_5$], so by [Sup-β_5] the sequence $(B^n + \check{X}^n)$ converges to B in probability, locally uniformly in time. Assuming that (M^n) converges $\mathscr{F}$–stably in law to X', by VIII.5.53 the pair (Y, M^n) converges in law to (Y, X') for any $\mathscr{F}$–measurable random variable on Ω. By Slutzky Theorem we also have convergence in law of the triple $(Y, M^n, B^n + \check{X}^n)$ to (Y, X', B). Since X' is continuous it follows that (Y, X^n) converges in law to (Y, X), and applying again VIII.5.53 we get the result.

Therefore it remains to prove the following: denoting by $\widetilde{E}$ the expectation w.r.t. $\widetilde{P}$, for any bounded random variable Y on $\mathscr{B}$ and any continuous bounded function Φ on $\mathbb{D}(\mathbb{R}^q)$, we have

7.8 $$E[Y\Phi(M^n)] \to \widetilde{E}[Y\Phi(X')].$$

4) Here we prove 7.8 under the additional assumption that $\mathscr{F}$ is separable: there is a sequence of bounded variables $(Y_m)_{m\in\mathbb{N}}$ which is dense in $L^2(\Omega, \mathscr{F}, P)$. We set $N_t^m = E(Y_m|\mathscr{F}_t)$, so $N^m \in \mathscr{M}_b$, and we have two important properties:

7.9 Let $N \in \mathscr{M}_b$. Its terminal variable N_∞ is the limit in L^2 of a subsequence Y_{m_k}. By Doob's inequality we deduce that $\sup_t |N_t - N_t^{m_k}| \to 0$ in L^2.

7.10 If $(\mathscr{G}_t)_{t\geq 0}$ is the smallest filtration w.r.t. which all N^m's are adapted, then $\mathscr{F}_t = \mathscr{G}_t$ up to P–null sets. Indeed let $A \in \mathscr{F}_t$; there is a sequence Y_{m_k} converging to 1_A in L^2, so $N_t^{m_k} = E(Y_{m_k}|\mathscr{F}_t)$ is $\mathscr{G}_t$–measurable and converges in L^2 to $E(1_A|\mathscr{F}_t) = 1_A$.

Let us introduce some more notation. First, $\mathscr{N} = (N^m)_{m\in\mathbb{N}}$ and $\mathscr{N}(n) = (N^m(n))_{m\in\mathbb{N}}$ (recall Property 7.1) can be considered as processes with paths in $\mathbb{D}(\mathbb{R}^\mathbb{N})$: here $\mathbb{R}^\mathbb{N}$ is a Polish space for the product topology, and all of Sections VI-1,2,3 can be readily extended to this situation. Then 7.1(iii) and [Sup-γ] yield

7.11 $$(Z(n), \mathscr{N}(n), K^n) \xrightarrow{P} (Z, \mathscr{N}, K) \quad \text{in } \mathbb{D}(\mathbb{R}^d \times \mathbb{R}^\mathbb{N} \times \mathbb{R}^{(d+q)^2}).$$

On the other hand, VI-4.18 and VI-4.22 and [Sup-γ] and [$\hat{\delta}_5$] imply that the sequence (M^n) is C-tight. It follows from 7.12 and VI.3.33 that the sequence $(M^n, Z(n), \mathscr{N}(n))$ is tight and that any limiting process $(\hat{M}, \hat{Z}, \hat{\mathscr{N}})$ has $\mathscr{L}(\hat{Z}, \hat{\mathscr{N}}) = \mathscr{L}(Z, \mathscr{N})$.

Choose now any subsequence, indexed by n', such that $(M^{n'}, Z(n'), \mathscr{N}(n'))$ converges in law. From what precedes one can realize the limit as such: consider the canonical space $(\Omega', \mathscr{F}', \mathbf{F}')$ of all continuous functions from $\mathbb{R}_+$ into $\mathbb{R}^q$, with the canonical process M', and define $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbf{F}})$ by II.7.2. Since $\mathscr{F} = \sigma(Y_m : m \in \mathbb{N})$ up to P–null sets, there is a probability measure $\widetilde{P}$ on $(\widetilde{\Omega}, \widetilde{\mathscr{F}})$ whose Ω–marginal is P, and such that the laws of $(M^{n'}, Z(n'), \mathscr{N}(n'))$ converge to the law of $(M', Z, \mathscr{N})$ under $\widetilde{P}$.

Therefore we have an extension $\widetilde{\mathscr{B}}$ of $\mathscr{B}$ (the existence of a disintegration of $\widetilde{P}$ as in II.7.2 is obvious, due to the definition of $(\Omega', \mathscr{F}')$), and up to $\widetilde{P}$–null sets the filtrations $\mathbf{F}$ and $\widetilde{\mathbf{F}}$ are generated by $(Z, \mathscr{N})$ and $(M', Z, \mathscr{N})$ respectively (use 7.10).

Set $U^n = (Z(n), M^n)$ and $U = (Z, M')$. By contruction, all components of $U^n, \mathscr{N}(n), U^{n,i}U^{n,j} - K^{n,ij}$ are $\mathbf{F}^n$–local martingales with uniformly bounded jumps. Then IX-1.17 (applied to processes with countably many components, which does not change the proof) yields that all components of U, $\mathscr{N}$ and

$U^iU^j - K^{ij}$ are $\widetilde{\mathbf{F}}$–local martingales under $\widetilde{P}$. This implies first that on our extension we have

$$\widetilde{C} = \langle M', M'\rangle, \qquad G = \langle M', Z\rangle$$

(since K is continuous increasing in $\mathscr{S}_{d+q}$), and second that all N^m are $\widetilde{\mathbf{F}}$–martingales. Then 7.9 yields that all elements of $\mathscr{M}_b$ are $\widetilde{\mathbf{F}}$–martingales, hence our extension is very good by Lemma II.7.3.

Let now $N \in \mathscr{M}_b$ be orthogonal to Z. We could have included N in the sequence (N^m): what precedes remains valid, with the same limit, for a suitable subsequence (n'') of (n'). Moreover $M^nN(n) - V(N)^n$ is an $\mathbf{F}^n$–local martingale with bounded jumps. So [Sup-γ^N] implies that $(M^{n''}, \mathscr{N}(n''), N(n''), V(N)^{n''}) \xrightarrow{\mathscr{L}} (M', \mathscr{N}, \underset{\sim}{N}, 0)$. The same argument as above yields that $M'N$ is a local martingale on $\mathscr{B}$, so M' is othogonal to all elements of $\mathscr{M}_b$ which are orthogonal to Z.

Therefore M' satisfies (i) of Proposition II.7.9: hence M' is an Z–biased continuous $\mathscr{F}$–progressive conditional Gaussian PII, whose law under Q_ω, which is Q_ω itself, is determined by the processes Z, $\widetilde{C}$, G, and in particular it does not depend on the subsequence (n') chosen above. In other words all convergent subsequence of $(M^n, \mathscr{N}(n))$ have the same limit $(M', \mathscr{N})$ in law, with the same measure $\widetilde{P}$, and thus the original sequence $(M^n, \mathscr{N}(n))$ converges in law to $(M', \mathscr{N})$. In particular if Φ is a bounded continuous function on $\mathbb{D}(\mathbb{R}^q)$ which is $\mathscr{D}_t$–measurable for some $t \geq 0$, and since $N(n)^m$ is a component of $\mathscr{N}(n)$ bounded uniformly in n, we get

$$E(\Phi(N^n)N(n)^m_t) \to \widetilde{E}(\Phi(M')N^m_t).$$

Now 7.1(i),(iii) yields that $N(n)^m_t \to N^m_t$ in L^2, hence

$$E(\Phi(M^n)Y_m) = E(\Phi(M^n)N^m_t) \to \widetilde{E}(\Phi(M')N^m_t) = \widetilde{E}(\Phi(M')Y_m).$$

Since any bounded $\mathscr{F}$–measurable variable Y is the L^2–limit of a subsequence of (Y_m), one readily deduces that 7.8 holds.

Now let Y as above and Φ bounded continuous on $\mathbb{D}(\mathbb{R}^q)$. For $p \in \mathbb{N}^*$ set $\Phi_p(\alpha) = \Phi(\alpha_p)$, where $\alpha_p(t) = \alpha(t)g_p(t)$ and $g_p(t) = 1 \wedge (p-t)^+$. Then Φ_p is bounded and continuous and $\mathscr{D}_{p+1}$–measurable, and $\delta(M^n, g_pM^n) \leq 2^{-p}$ for the distance δ defined in VI.1.26. Since $\Phi_p(M^n) = \Phi(g_pM^n)$, one deduces from 7.8 for each Φ_p that 7.8 holds also for Φ, exactly like at the end of Step (3). So we are finished for the present step.

5) It remains to remove the separability assumption on $\mathscr{F}$. Denote by $\mathscr{H}$ the σ–field generated by the random variables $(Z_t, K_t, B_t, X^n_t : t \geq 0, n \geq 1)$, and let $\mathscr{G}$ be any separable σ–field containing $\mathscr{H}$. Let $(Y_m)_{m\in\mathbb{N}}$ be a dense sequence of bounded variables in $L^2(\Omega, \mathscr{G}, P)$, and $N^m_t = E(Y_m|\mathscr{G}_t)$, and set $(\mathscr{G}_t)_{t\geq 0}$ for the filtration generated by the processes $(N^m)_{m\in\mathbb{N}}$.

We have $E(Y_m|\mathscr{F}_t) = E(Y_m|\mathscr{G}_t)$ for all m, so by a density argument $E(Y|\mathscr{F}_t) = E(Y|\mathscr{G}_t)$ for all $Y \in L^2(\Omega, \mathscr{G}, P)$: this implies that any $(\mathscr{G}_t)$–martingale is an **F**–martingale, and in particular each N^m is in $\mathscr{M}_b$, and also that every **F**–adapted and $\mathscr{G}$–measurable process (like K, B and M) is $(\mathscr{G}_t)$–adapted. Thus Z is a $(\mathscr{G}_t)$–local martingale. Finally, any bounded $(\mathscr{G}_t)$–martingale which is orthogonal w.r.t. $(\mathscr{G}_t)$ to Z is also orthogonal to Z w.r.t. **F**.

In other words, 7.1 is satisfied by $(\mathscr{G}_t)$ and the same filtrations $\mathbf{F}^n$ and processes $Z(n)$, $N(n)$, and all assumptions in our theorem are satisfied as well with $(\mathscr{G}_t)$ instead of **F**. We can thus apply Step 4 with the same space $(\Omega', \mathscr{F}', \mathbf{F})$ and the same process M', and $\widetilde{\Omega} = \Omega \times \Omega'$, $\widetilde{\mathscr{G}} = \mathscr{G} \otimes \mathscr{F}'$, $\widetilde{\mathscr{G}}_t = \bigcap_{s>t} \mathscr{G}_s \otimes \mathscr{F}'_s$. We have a transition probability $Q_{\mathscr{G},\omega}(d\omega')$ from $(\Omega, \mathscr{G})$ into $(\Omega', \mathscr{F}')$, such that if $\widetilde{P}_{\mathscr{G}}(d\omega, d\omega') = P_{\mathscr{G}}(d\omega) Q_{\mathscr{G},\omega}(d\omega')$ (where $P_{\mathscr{G}}$ is the restriction of P to $\mathscr{G}$), then

7.12 $$E_{\mathscr{G}}(\Phi(M^n)Y) \to \widetilde{E}_{\mathscr{G}}(\Phi(M')Y)$$

for all bounded continuous functions Φ on $\mathbb{D}(\mathbb{R}^q)$ and all bounded $\mathscr{G}$–measurable variables Y.

Further, $Q_{\mathscr{G},\omega}$ only depends on Z, F, G and so is indeed a transition from $(\Omega, \mathscr{H})$ into $(\Omega', \mathscr{F}')$ not depending on $\mathscr{G}$ and written Q_ω.

It remains to define $\widetilde{\mathscr{B}}$ by II.7.2: since $\omega \rightsquigarrow Q_\omega(A)$ is $\mathscr{F}_t$–measurable for $A \in \mathscr{F}'_t$ it is a very good extension of $\mathscr{B}$. Furthermore $E_{\mathscr{G}}(\Phi(M^n)Y) = E(\Phi(M^n)Y)$ and $\widetilde{E}_{\mathscr{G}}(\Phi(M')Y) = \widetilde{E}(\Phi(M')Y)$ for all bounded $\mathscr{G}$–measurable Y: hence 7.12 yields 7.8 for all such Y. Since any $\mathscr{F}$–measurable variable Y is also $\mathscr{G}$–measurable for some separable σ–field $\mathscr{G}$ containing $\mathscr{H}$, we deduce that 7.8 holds for all bounded $\mathscr{F}$–measurable Y, and we are done. □

As all limit theorems, the previous one admits a square-integrable version: assume that each X^n is a locally square integrable semimartingale (relative to $\mathbf{F}^n$; see II.2.27), that is we have $|x|^2 \star \nu^n_t < \infty$ for all $t > 0$, or equivalently its canonical decomposition $X^n = X^n_0 + B'^n + M'^n$ has that M'^n is a locally square-integrable martingale. Then we consider the processes

$$\widetilde{C}'^{n,ij}_t = \langle M'^{n,i}, M'^{n,j} \rangle_t = \widetilde{C}^{n,ij}_t + (x^i x^j) \star \nu^n_t - \sum_{s \le t} \Delta B'^{n,i}_s \Delta B'^{n,j}_s.$$

Then by essentially reproducing the proof of VIII.2.14, we deduce from 7.3 that:

7.13 **Theorem.** *Assume 7.1, and let $\widetilde{C}$, G and B be continuous adapted processes on $\widetilde{\mathscr{B}}$, null at 0 and taking values in $\mathscr{S}_q$, $\mathbb{R}^q \otimes \mathbb{R}^d$ and $\mathbb{R}^q$ respectively. Suppose that, for some countable subset D of $\mathbb{R}_+$, we have*

[Sup-β'_5] $$\sup_{s \le t} |B'^n_s - B_s| \xrightarrow{P} 0 \qquad \text{for all } t > 0,$$

[γ'_5-D] $$\widetilde{C}'^n_t \xrightarrow{P} C_t \qquad \text{for all } t \in D,$$

7.14 $$(|x|^2 1_{\{|x|>\varepsilon\}}) \star \nu^n_t \xrightarrow{P} 0 \qquad \text{for all } t > 0,\ \varepsilon > 0,$$

and also that for all $N \in \mathscr{M}_b$ orthogonal to all components of Z, we have:

7.15 $$G'^n_t := \langle M'^n, Z(n)\rangle_t \xrightarrow{P} G_t \qquad \text{for all } t \in D,$$

7.16 $$V(N)'^n_t := \langle M'^n, N(n)\rangle_t \xrightarrow{P} 0 \qquad \text{for all } t \in D.$$

Then all results of Theorem 7.3 hold true.

§ 7b. Convergence of Discretized Processes

Here we specialize the previous results to the case when the filtration $\mathbf{F}^n$ is the "discretized" filtration $\mathscr{F}^n_t = \mathscr{F}_{[nt]/n}$. For every càdlàg process X we use the following notation:

7.17 $$X^n_t = X_{[nt]/n}, \qquad \Delta^n_i X = X_{i/n} - X_{(i-1)/n} = \Delta X^n_{i/n}.$$

Here again we have a continuous d–dimensional local martingale Z on the stochastic basis $\mathscr{B}$. We have a *continuous* truncation function h on $\mathbb{R}^q$, and we also fix another *continuous* truncation function h' on $\mathbb{R}^d$. We also consider for each n an $\mathbf{F}^n$–semimartingale, i.e. a process of the form

7.18 $$X^n_t = \sum_{i=1}^{[nt]} \chi^n_i,$$

where each χ^n_i is $\mathscr{F}_{i/n}$–measurable.

7.19 **Theorem.** *Let $\widetilde{C}$, G and B be continuous adapted processes on $\mathscr{B}$, null at 0 and taking values in $\mathscr{S}_q$, $\mathbb{R}^q \otimes \mathbb{R}^d$ and $\mathbb{R}^q$ respectively. If for all $t > 0$, $\varepsilon > 0$, $1 \le j, k \le q$, $1 \le l \le d$ and all $N \in \mathscr{M}_b$ which are orthogonal to all components of Z we have*

7.20 $$\sup_t \Big| \sum_{i=1}^{[nt]} E(h(\chi^n_i)|\mathscr{F}_{\frac{i-1}{n}}) - B_t \Big| \xrightarrow{P} 0,$$

7.21 $$\sum_{i=1}^{[nt]} \Big(E(h(\chi^n_i)^j h(\chi^n_i)^k|\mathscr{F}_{\frac{i-1}{n}}) - E(h(\chi^n_i)^j|\mathscr{F}_{\frac{i-1}{n}}) E(h(\chi^n_i)^k|\mathscr{F}_{\frac{i-1}{n}}) \Big) \xrightarrow{P} \widetilde{C}^{jk}_t,$$

7.22 $$\sum_{i=1}^{[nt]} \Big(E(h(\chi^n_i)^j h'(\Delta^n_i Z)^l|\mathscr{F}_{\frac{i-1}{n}}) - E(h(\chi^n_i)^j|\mathscr{F}_{\frac{i-1}{n}}) E(h'(\Delta^n_i Z)^l|\mathscr{F}_{\frac{i-1}{n}}) \Big) \xrightarrow{P} G^{jl}_t,$$

7.23 $$\sum_{i=1}^{[nt]} P(|\chi^n_i| > \varepsilon|\mathscr{F}_{\frac{i-1}{n}}) \xrightarrow{P} 0,$$

7.26 $$\sum_{i=1}^{[nt]} E(h(\chi_i^n)\Delta_i^n N|\mathscr{F}_{\frac{i-1}{n}}) \xrightarrow{P} 0,$$

then all results of Theorem 7.3 hold true.

Proof. We will prove that the assumptions of Theorem 7.3 are in force.

First we check 7.1: we take $N(n) = N^n$, as defined in 7.17, for all $N \in \mathscr{M}_b$, so 7.1(i) is obvious. Note also that that if $N^1, \dots, N^m$ are in $\mathscr{M}_b$, then by Proposition VI.7.5

7.25 $$(Z^n, N(n)^1, \dots, N(n)^m) \xrightarrow{P} (Z, N^1, \dots, N^m) \qquad \text{in } \mathbb{D}(\mathbb{R}^{d+m}).$$

Next, we define $Z(n)$ as

$$Z(n)_t = \sum_{i=1}^{[nt]} \left(h'(\Delta_i^n Z) - E(h'(\Delta_i^n Z)|\mathscr{F}_{\frac{i-1}{n}})\right),$$

so $Z^n - Z(n) = A^n + A'^n$, where $A_t^n = \sum_{i=1}^{[nt]} E(h'(\Delta_i^n Z)|\mathscr{F}_{\frac{i-1}{n}})$ and $A'^n_t = \sum_{i=1}^{[nt]}(\Delta_i^n Z - h'(\Delta_i^n Z)$. Then 7.1(ii) follows from combining the results (1.15) and (2.12) in [269], because Z is continuous. These results also yield $\sup_t|A_t^n| \xrightarrow{P} 0$, and for all $t > 0$, $\varepsilon > 0$:

$$\sum_{i=1}^{[nt]} P(|\Delta_i^n Z| > \varepsilon|\mathscr{F}_{\frac{i-1}{n}}) \xrightarrow{P} 0.$$

This and VI-4.22, together with $h'(x) = x$ for $|x|$ small enough, imply $\sup_t|A'^n_t| \xrightarrow{P} 0$, so finally $\sup_t|Z_t^n - Z(n)_t| \xrightarrow{P} 0$ and 7.1(iii) follows from 7.25: we thus have 7.1.

The decomposition 7.2 of X^n has $B_t^n = \sum_{i=1}^{[nt]} E(h(\chi_i^n)|\mathscr{F}_{\frac{i-1}{n}})$ and $M_t^n = \sum_{i=1}^{[nt]} (h(\chi_i^n) - E(h(\chi_i^n)|\mathscr{F}_{\frac{i-1}{n}}))$. Hence 7.20, 7.21, 7.22, 7.23 and 7.24 are respectively [Sup-β_5], [γ_5-$\mathbb{R}_+$], 7.4, $\hat{\delta}_5$] and 7.5, so we are finished. □

Finally, we could state the discrete version of Theorem 7.13. Let us rather specialize a little bit more, by supposing that Z is square-integrable and that each χ_i^n is square-integrable. Then we get:

7.28 **Theorem.** *Assume that Z is a continuous d–dimensional local martingale such that $E(|Z_t|^2) < \infty$ for all t, and that each χ_i^n is square-integrable. Let $\tilde{C}$, G and B be continuous adapted processes on $\mathscr{B}$, null at* 0 *and taking values in $\mathscr{S}_q$, $\mathbb{R}^q \otimes \mathbb{R}^d$ and $\mathbb{R}^q$ respectively. If for all $t > 0$, $\varepsilon > 0$, $1 \le j, k \le q$, $1 \le l \le d$ and all $N \in \mathscr{M}_b$ which are orthogonal to all components of Z we have*

7.27
$$\sup_t \Big| \sum_{i=1}^{[nt]} E(\chi_i^n | \mathscr{F}_{\frac{i-1}{n}}) - B_t \Big| \xrightarrow{P} 0,$$

7.28
$$\sum_{i=1}^{[nt]} \Big(E(\chi_i^{n,j} \chi_i^{n,k} | \mathscr{F}_{\frac{i-1}{n}}) - E(\chi_i^{n,j} | \mathscr{F}_{\frac{i-1}{n}}) E(\chi_i^{n,k} | \mathscr{F}_{\frac{i-1}{n}}) \Big) \xrightarrow{P} \widetilde{C}_t^{jk},$$

7.29
$$\sum_{i=1}^{[nt]} \Big(E(\chi_i^{n,j} \Delta_i^n Z^l | \mathscr{F}_{\frac{i-1}{n}}) \Big) \xrightarrow{P} G_t^{jl},$$

7.30
$$\sum_{i=1}^{[nt]} E(|\chi_i^n|^2 1_{\{|\chi_i^n| > \varepsilon\}} | \mathscr{F}_{\frac{i-1}{n}}) \xrightarrow{P} 0,$$

7.31
$$\sum_{i=1}^{[nt]} E(\chi_i^n \Delta_i^n N | \mathscr{F}_{\frac{i-1}{n}}) \xrightarrow{P} 0,$$

then all results of Theorem 7.3 hold true.

Chapter X. Limit Theorems, Density Processes and Contiguity

Let us roughly describe the problems which will retain our attention in this last chapter.

These problems are connected with contiguity and convergence of processes, so the setting is the same than in Chapter V: for each $n \in \mathbb{N}^*$ we have a filtered space $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ endowed with two probability measures P^n and P'^n. Set $Q^n = (P^n + P'^n)/2$, and denote by z^n and z'^n the density processes of P^n and P'^n with respect to Q^n. Exactly as before V.2.5, we introduce the "density process of P'^n with respect to P^n" by

1.1 $$Z_t^n = \frac{z_t'^n}{z_t^n} \quad \left(\text{with } \frac{2}{0} = +\infty\right)$$

(when $P'^n \overset{\text{loc}}{\ll} P^n$, Z^n takes only finite values, and is the ordinary density process; in all cases, Z^n is a P^n-supermartingale).

As usual, P_t^n and $P_t'^n$ denote the restrictions of P^n and P'^n to $\mathscr{F}_t^n$, and we always will assume

1.2 $$P_0^n = P_0'^n \quad \text{for all } n \in \mathbb{N}^* \quad (\text{hence } z_0^n = z_0'^n = Z_0^n = 1).$$

This assumption is met in almost all applications, and it avoids a lot of trivial complications in the statements of the results. We also assume that $\mathscr{F}^n = \mathscr{F}_{\infty-}^n = \bigvee_t \mathscr{F}_t^n$.

There are essentially two subjects entered upon in this chapter:

1) The first one concerns the convergence of the sequence Z^n when we have "local contiguity", i.e. $(P_t'^n) \lhd (P_t^n)$ for all $t \in \mathbb{R}_+$. Of course one has to distinguish between convergence of $\mathscr{L}(Z^n|P^n)$ (the law of Z^n under P^n) and of $\mathscr{L}(Z^n|P'^n)$, which we also write as

1.3 $$\begin{cases} Z^n \xrightarrow{\mathscr{L}(P^n)} Z \\ Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' \end{cases}$$

(no reason in general for having $\mathscr{L}(Z) = \mathscr{L}(Z')$!)

In Section 1 we obtain necessary and sufficient conditions, when the limiting process Z is a continuous martingale which never vanishes. In Section 2 we examine the case where $Z = e^X$, X being a process with independent increments. Such results are motivated by statistical applications, especially when Z has the form $Z = \exp\left(M - \frac{C}{2}\right)$, with M a continuous Gaussian martingale, $M_0 = 0$, $C_t = E(M_t^2)$: in this case, 1.3 says that in the limit one can replace the simple hypothesis P^n and its alternative P'^n by a Gaussian experiment, i.e. a simple statistical model where the Log-likelihood is Gaussian.

2) The second subject concerns the following: we suppose that each space $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ is endowed with a càdlàg adapted process X^n. Then, assuming that $(P_t'^n) \lhd (P_t^n)$ for all t and that $X^n \xrightarrow{\mathscr{L}(P^n)} X$, under which conditions does the sequence $\mathscr{L}(X^n | P'^n)$ also converge in law? This problem is studied in Section 3.

1. Convergence of the Density Processes to a Continuous Process

§ 1a. Introduction, Statement of the Main Results

1. In order to gain some insight on the problem at hand, it is perhaps useful to begin with a (partial) extension of LeCam's third lemma V.1.13, relevant to stochastic processes.

1.4 Proposition. *Assume that $\mathscr{L}(Z^n | P^n) \to Q$, where Q is a probability measure on the canonical space $(\Omega, \mathscr{F}, \mathbf{F}) = (\mathbb{D}(\mathbb{R}), \mathscr{D}(\mathbb{R}), \mathbf{D}(\mathbb{R}))$.*

a) *There is equivalence between:* (i) $(P_t'^n) \lhd (P_t^n)$ *for all* $t \in \mathbb{R}_+$.

(ii) $\mathscr{L}(Z^n | P'^n) \to Q'$, *where Q' is another probability measure on $(\Omega, \mathscr{F})$.*

b) *If* (i) *or* (ii) *holds, we have $Q' \overset{\text{loc}}{\ll} Q$, and the canonical process Z on Ω is a Q-martingale, and is the density process of Q' with respect to Q. If moreover $Q(Z_t > 0) = 1$ we have $(P_t^n) \lhd (P_t'^n)$.*

Proof. Let $D = \{t > 0\colon Q(\Delta Z_t \neq 0) = 0\}$.

Assume first that (ii) holds, and let $D' = \{t > 0\colon Q'(\Delta Z_t \neq 0) = 0\}$. If $t \in D'$ we have $\mathscr{L}(Z_t^n | P_t'^n) \to \mathscr{L}(Z_t | Q')$, so V.1.13 (implication (ii) $\Rightarrow$ (i)) implies that $(P_t'^n) \lhd (P_t^n)$, and a-fortiori $(P_s'^n) \lhd (P_s^n)$ for all $s \leq t$. Hence (i) holds.

Secondly, assume that (i) holds. If $t \in D$, the stopping map $\alpha \rightsquigarrow \alpha^t(s) = \alpha(t \wedge s)$ is P-a.s. Skorokhod-continuous on Ω. Then

$$\mathscr{L}[(Z_t^n, (Z^n)^t) | P_t^n] \to \mathscr{L}[(Z_t, Z^t) | Q_t]$$

and a-fortiori $\mathscr{L}(Z_t^n | P_t^n) \to \mathscr{L}(Z_t | Q_t)$. Since $(P_t'^n) \lhd (P_t^n)$ by hypothesis, V.1.13 yields $\mathscr{L}[(Z_t^n, (Z^n)^t | P_t'^n] \to \overline{\eta}_t'$ weakly, where $\overline{\eta}_t'$ is the probability measure on $\mathbb{R}_+ \times \Omega$ characterized by

$$\overline{\eta}_t'(f) = E_Q[Z_t f(Z_t, Z^t)].$$

Hence in particular, $E_Q(Z_t) = 1$ and

1.5 $$\mathscr{L}((Z^n)^t|P'^n) \to Q_t'^0 := Z_t \cdot Q_t.$$

Now, 1.5 implies that the sequences $\{\mathscr{L}((Z^n)^t|P'^n)\}_{n\in\mathbb{N}^*}$ are tight for all $t \in D$, and so the tightness criterion VI.3.21 implies that the sequence $\{\mathscr{L}(Z^n|P'^n)\}$ itself is tight. If Q' is a limit point of this sequence, we easily deduce from 1.5 again that $Q'_t := Q'|\mathscr{F}_t = Q_t'^0$. Therefore this limit point Q' is unique, and (ii) holds.

Furthermore, $Q' \overset{\text{loc}}{\ll} Q$ and the density process of Q' with respect to Q is Z, by 1.5 again, and in particular Z is a Q-martingale. Finally, the last claim in (b) follows from V.1.15. □

2. The previous proposition shows that, when 1.3 holds, Z can a-priori be any nonnegative martingale with $Z_0 = 1$. However, the interest of contiguous sequences of measures lies in statistics, and according to the discussion by LeCam [145] one is mainly interested in "asymptotically Gaussian experiments", or mixtures of these in the sense of § VIII.5a, essentially because they are (almost) the only cases which are well understood from the statistical point of view.

The measures P^n and P'^n may of course converge. But if they do not, nevertheless in the setting of 1.4 the distributions of the sufficient statistics Z^n converge, and so "in the limit" we replace P^n and P'^n by Q and Q', the respective limits of $\mathscr{L}(Z^n|P^n)$ and $\mathscr{L}(Z^n|P'^n)$.

In the setting of stochastic processes, Gaussian experiment means that the Log-likelihood (the logarithm of the density process) is a (continuous) Gaussian process of the form $\operatorname{Log} Z = M - C/2$, where M is a continuous Gaussian martingale with $M_0 = 0$ and $\langle M, M\rangle = C$.

Observe also that, in virtue of Theorem III.5.35, the process Z is then the relative density process of two laws of continuous processes with independent increments.

3. The conditions for having 1.3 will be expressed in terms of various predictable processes on $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$, which we introduce now. As a matter of fact, we put here some more notation than strictly needed for the present section, in order to prepare for Section 2:

1.6 $H^n = h^n(\frac{1}{2}; P^n, P'^n)$ is any version of the Hellinger process of order $\frac{1}{2}$ between P^n and P'^n (see IV.1.24). □

1.7 Let $a \in (1, \infty)$. The function $x \rightsquigarrow \varphi_a(x) = (1-x)1_{[0,1/a]}(x)$ meets IV.1.40. Then we set $I^n(a) = \imath(\varphi_a; P^n, P'^n) + \imath(\varphi_a; P'^n, P^n)$, where $\imath(\varphi_a; P^n, P'^n)$ and $\imath(\varphi_a; P'^n, P^n)$ are any versions of the processes defined in IV.1.46. (Note that we also have $I^n(a) = \imath(\psi_a; P^n, P'^n) + h(0; P'^n, P^n)$, where $\psi_a(x) = |x-1|\,1_{\{1/a<x<a\}^c}$). □

1.8 Let $g: \mathbb{R} \to \mathbb{R}$ be a Borel bounded function vanishing on a neighbourhood of 0. Then

$$\varphi_g(x) = \begin{cases} xg(-\operatorname{Log} x) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

also meets IV.1.40, and we denote by $G^n(g)$ any version of the process $\imath(\varphi_g; P^n, P'^n)$, as defined in IV.1.46. □

Explicit versions for these processes are given in IV.1.36 and IV.1.49. Let us recall them. We denote by $\langle z^{n,c}, z^{n,c} \rangle$ and ν^{z^n} the second and third characteristics of the martingale z^n, relatively to Q^n. Since $z^n + z'^n = 2$, we have $1/z^n_- + 1/z'^n_- = 2/z^n_- z'^n_-$, and ν^{z^n} only charges the set

1.9 $$\{(\omega, t, x): z^n_{t-}(\omega) > 0, z'^n_{t-}(\omega) > 0, -z^n_{t-}(\omega) \leq x \leq z'^n_{t-}(\omega)\}$$

by IV.1.33a. Hence we obtain the following versions for H^n, $I^n(a)$, $G^n(g)$:

1.10 $$\begin{cases} H^n = \dfrac{1}{2(z^n_- z'^n_-)^2} \cdot \langle z^{n,c}, z^{n,c} \rangle + \dfrac{1}{2}\left(\sqrt{1 + \dfrac{x}{z^n_-}} - \sqrt{1 - \dfrac{x}{z'^n_-}}\right)^2 * \nu^{z^n} \\[2ex] I^n(a) = \dfrac{2|x|}{z^n_- z'^n_-} 1_{\{1/a < (1 + x/z^n_-)/(1 - x/z'^n_-) < a\}^c} * \nu^{z^n} \\[2ex] G^n(g) = \left(1 + \dfrac{x}{z^n_-}\right) g \circ \operatorname{Log} \dfrac{1 - x/z'^n_-}{1 + x/z^n_-} 1_{\{x \neq z'^n_-, x \neq -z^n_-\}} * \nu^{z^n}. \end{cases}$$

1.11 **Remark.** If $P'^n \overset{\text{loc}}{\ll} P^n$, Z^n is a P^n-martingale and we denote by $\langle Z^{n,c}, Z^{n,c} \rangle$ and ν^{Z^n, P^n} its second and third characteristics under P^n. Then the following also gives versions of H^n, $I^n(a)$, $G^n(g)$ (cf IV.1.39 and IV.1.50):

$$\begin{cases} H^n = \dfrac{1}{8(Z^n_-)^2} \cdot \langle Z^{n,c}, Z^{n,c} \rangle + \dfrac{1}{2}\left(1 - \sqrt{1 + \dfrac{x}{Z^n_-}}\right)^2 * \nu^{Z^n, P^n} \\[2ex] I^n(a) = \dfrac{|x|}{Z^n_-} 1_{\{1/a < 1 + x/Z^n_- < a\}^c} * \nu^{Z^n, P^n} \\[2ex] G^n(g) = g \circ \operatorname{Log}\left(1 + \dfrac{x}{Z^n_-}\right) 1_{\{x \neq -Z^n_-\}} * \nu^{Z^n, P^n}. \end{cases}$$ □

1.12 **Theorem.** *Let $t \rightsquigarrow C_t$ be a non-decreasing continuous function with $C_0 = 0$. Let M be a continuous martingale with $M_0 = 0$ and $\langle M, M \rangle_t = C_t$, on some stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$* (so M is Gaussian). *The following statements* (i) *and* (ii) *are equivalent*:

(i) $Z^n \xrightarrow{\mathscr{L}(P^n)} Z = e^{M - C/2}$.

(ii) *There is a dense subset D in $\mathbb{R}_+$ such that the following holds*:

$[L\text{-}D]$ $\quad I^n(1 + \varepsilon)_t \xrightarrow{P^n} 0$ for all $t \in D$, $\varepsilon > 0$.

$[H\text{-}D]$ $\quad H^n_t \xrightarrow{P^n} \frac{1}{8} C_t$ for all $t \in D$.

In this case, we have $[L\text{-}\mathbb{R}_+]$ *and* $[H\text{-}\mathbb{R}_+]$, *and* $Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' = e^{M+C/2}$, *and* $(P_t'^n) \lhd (P_t^n)$ *and* $(P_t^n) \lhd (P_t'^n)$ *for all* $t \in \mathbb{R}_+$.

$[L\text{-}D]$ is a sort of Lindeberg condition: see the version 1.10 for $I^n(1+\varepsilon)$. Observe that Z^n is P^n-a.s. finite-valued, so (i) is the usual convergence in $\mathbb{D}(\mathbb{R})$. However, Z^n may take the value $+\infty$ with a positive P'^n-probability, so $Z^n \xrightarrow{\mathscr{L}(P'^n)} Z'$ means convergence in $\mathbb{D}(\bar{\mathbb{R}})$ (or $\mathbb{D}([0,\infty])$).

We will see in § 1e another formulation of the same theorem, in which $[L\text{-}D]$ is replaced by a condition on Hellinger processes.

1.13 **Remark.** Contrarily to the appearances, this result is completely symmetrical between (P^n) and (P'^n):

a) Under (ii) we have $(P_t'^n) \lhd (P_t^n)$, and so

1.14 $$I^n(1+\varepsilon)_t \xrightarrow{P'^n} 0, \quad H_t^n \xrightarrow{P'^n} \tfrac{1}{8} C_t \quad \text{for all } t \in D, \varepsilon > 0.$$

b) Conversely if 1.14 holds, we can apply the theorem after exchanging (P^n) and (P'^n). Then $(P_t^n) \lhd (P_t'^n)$ and $[L\text{-}D]$, $[H\text{-}D]$ are deduced as above.

c) If $Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' = e^{M+C/2}$, then $1/Z^n \xrightarrow{\mathscr{L}(P'^n)} 1/Z' = e^{M'-C/2}$ with $M' = -M$, and obviously $\mathscr{L}(M') = \mathscr{L}(M)$, while $1/Z^n$ is the density process of P^n with respect to P'^n. □

1.15 **Remark.** The limiting function $C/8$ in $[H\text{-}D]$ is also a Hellinger process. Indeed, let $Q = \mathscr{L}(M)$ and $Q' = \mathscr{L}(M-C)$. Then IV.4.24 shows that $C/8 = h(\frac{1}{2}; Q, Q')$. We also have $C/8 = h(\frac{1}{2}; P, P')$ where $P = \mathscr{L}(Z)$, $P' = \mathscr{L}(Z')$, because P and P' are the images of Q and Q' by the map $\alpha \rightsquigarrow e^{\alpha - C/2}$ from $\mathbb{D}(\mathbb{R})$ into itself, and this map preserves the filtrations. □

There is also a finite-dimensional version for the implication (ii) ⇒ (i), but of course no necessary condition for finite-dimensional convergence in general.

1.16 **Theorem.** *Let C and M be as in* 1.12, *and let D be a subset of* $\mathbb{R}_+$. *If* $[L\text{-}D]$ *and* $[H\text{-}D]$ *hold, then*

$$Z^n \xrightarrow{\mathscr{L}(D|P^n)} Z = e^{M-C/2}, \quad Z^n \xrightarrow{\mathscr{L}(D|P'^n)} Z' = e^{M+C/2}$$

(finite-dimensional convergence along D), *and* $(P_t'^n) \lhd (P_t^n)$ *and* $(P_t^n) \lhd (P_t'^n)$ *for all t such that* $[t,\infty) \cap D \neq \emptyset$.

We finish the subsection with the "discrete-time" version of Theorem 1.12. Here we assume that $(\Omega^n, \mathscr{F}^n, \mathbf{G}^n = (\mathscr{G}_p^n)_{p \in \mathbb{N}})$ are discrete-time bases with two measures P^n, P'^n, and $Q^n = (P^n + P'^n)/2$; as usual, $z^n = (z_p^n)_{p \in \mathbb{N}}$ and $z'^n = (z_p'^n)_{p \in \mathbb{N}}$ are the density processes of P^n and P'^n with respect to Q^n, and for simplicity we assume here also 1.2, so $z_0^n = z_0'^n = 1$. Finally we set

$$Z_t^n = \frac{z_{[nt]}'^n}{z_{[nt]}^n} \quad \left(\text{with } \frac{2}{0} = +\infty\right).$$

and P_p^n, $P_p'^n$ are the restrictions of P^n and P'^n to $\mathscr{G}_p^n$.

1.17 **Theorem.** *Let C and M be as in 1.12, and let D be a dense subset of $\mathbb{R}_+$. There is equivalence between*

(i) $Z^n \xrightarrow{\mathscr{L}(P^n)} Z = e^{M-C/2}$.

(ii) *The following two conditions hold, for all $t \in D$, $\varepsilon > 0$*

$$\sum_{1 \le p \le [nt]} \frac{1}{z^n_{p-1} z'^n_{p-1}} E_{Q^n}(|z^n_p - z'^n_p| 1_{\{1/(1+\varepsilon) < z'^n_{p-1} z^n_p / z^n_{p-1} z'^n_p < 1+\varepsilon\}^c} | \mathscr{G}^n_{p-1}) \xrightarrow{P^n} 0,$$

$$\sum_{1 \le p \le [nt]} E_{Q^n}((\sqrt{z^n_p / z^n_{p-1}} - \sqrt{z'^n_p / z'^n_{p-1}})^2 | \mathscr{G}^n_{p-1}) \xrightarrow{P^n} \frac{1}{4} C_t.$$

In this case, we also have $Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' = e^{M+C/2}$, and $(P'^n_{[nt]}) \lhd (P^n_{[nt]})$ and $(P^n_{[nt]}) \lhd (P'^n_{[nt]})$ for all $t \in \mathbb{R}_+$.

Proof. We introduce the continuous-time filtration $\mathbf{F}^n = (\mathscr{F}^n_t = \mathscr{G}^n_{[nt]})$, so that Z^n is exactly the density process given by 1.1. Then, in view of IV.1.64 and IV.1.65, the conditions in (ii) are exactly [L-D] and [H-D], and it suffices to apply Theorem 1.12. $\square$

§ 1b. An Auxiliary Computation

Our aim in this subsection is to compute the characteristics of "Log Z^n" with respect to P^n. Of course, Log Z^n may take the values $+\infty$ and $-\infty$, so one needs to be careful.

Let us fix $n \in \mathbb{N}^*$, and set

1.19 $$\begin{cases} T^n = \inf\left(t: z^n_t < \frac{1}{n} \text{ or } z'^n_t < \frac{1}{n}\right), & \text{so } [\![0, T^n]\!] \subset \{z^n_- > 0, z'^n_- > 0\} \\ A^n = \{(\omega, t, x): t \le T^n(\omega), \quad -z^n_{t-}(\omega) \le x \le z'^n_{t-}(\omega)\}. \end{cases}$$

1.20 $f_n: \mathbb{R} \to [1/2n^2, \infty)$ is a C^2 function with $f_n(x) = x$ for $x \ge 1/n^2$.

1.21 $$Y^n = \operatorname{Log} \frac{f_n[(z'^n)^{T^n}]}{f_n[(z^n)^{T^n}]}, \quad \text{so } Z^n = e^{Y^n} \quad \text{on } [\![0, T^n[\![.$$

1.22 $$V^n(\omega, t, x) = \begin{cases} \operatorname{Log} \dfrac{z^n_{t-}(\omega) f_n(z'^n_{t-}(\omega) - x)}{z'^n_{t-}(\omega) f_n(z^n_{t-}(\omega) + x)} & \text{if } t \le T^n(\omega) \\ 0 & \text{if } t > T^n(\omega) \end{cases}$$

1.23 **Proposition.** *Let $h \in \mathscr{C}^1_t$ be a truncation function. The semimartingale Y^n admits the following characteristics (B^n, C^n, ν^n) and modified second characteristic $\tilde{C}^n$, with respect to P^n, and relatively to the truncation function h:*

$$B^n = -\frac{2}{(z_-^n z_-'^n)^2}\cdot\langle z^{n,c}, z^{n,c}\rangle^{T^n} + \left[h(V^n)\left(1+\frac{x}{z_-^n}\right) + \frac{2x}{z_-^n z_-'^n}\right]1_{A^n} * \nu^{z^n},$$

$$C^n = \frac{4}{(z_-^n z_-'^n)^2}\cdot\langle z^{n,c}, z^{n,c}\rangle^{T^n},$$

$$g * \nu^n = \left(1+\frac{x}{z_-^n}\right)g(V^n)1_{A^n} * \nu^{z^n},$$

$$\tilde{C}^n = \frac{4}{(z_-^n z_-'^n)^2}\cdot\langle z^{n,c}, z^{n,c}\rangle^{T^n} + \left(1+\frac{x}{z_-^n}\right)h(V^n)^2 1_{A^n} * \nu^{z^n} - \sum_{s\leq\cdot}(\Delta B_s^n)^2.$$

Proof. a) We denote by μ^{z^n} and μ^{Y^n} the random measures associated with the jumps of z^n and Y^n (see II.1.16).

We firstly compute the P^n-characteristics of z^n, denoted by $(B^{n,P}, C^{n,P}, \nu^{n,P})$, and relative to a truncation function $h' \in \mathscr{C}_t^1$ which has $h'(x) = x$ for $|x| \leq 2$. To this end, we apply Girsanov's Theorem III.3.24: firstly, $C^{n,P} = \langle z^{n,c}, z^{n,c}\rangle$. Secondly, by definition of μ^{z^n}, we have $z_t^n = z_{t-}^n(1 + x/z_{t-}^n)$ $\mu^{z^n}(dt,dx)$-a.e., so $M_{\mu^{z^n}}^{Q^n}(z^n|\tilde{\mathscr{P}}^n) = z_-^n(1 + x/z_-^n)$ and therefore $\nu^{n,P} = (1 + x/z_-^n)\cdot\nu^{z^n}$. Thirdly, $\langle z^{n,c}, z^{n,c}\rangle = \dfrac{z_-^n}{z_-^n}\cdot\langle z^{n,c}, z^{n,c}\rangle$ on $\{z_-^n > 0\}$, so we obtain

$$B^{n,P} = \frac{1}{z_-^n}\cdot\langle z^{n,c}, z^{n,c}\rangle + \frac{x^2}{z_-^n} * \nu^{z^n}$$

(since $|\Delta z^n| \leq 2$ we have $x = h'(x)$ ν^{z^n}-a.e.). Hence the canonical representation II.2.34 of z^n, relatively to P^n, is

1.24
$$\begin{cases} z^n = 1 + M^n + x*(\mu^{z^n} - \nu^{n,P}) + (1/z_-^n)\cdot\langle z^{n,c}, z^{n,c}\rangle + (x^2/z_-^n)*\nu^{z^n}, \\ \text{where (a) } M^n \text{ is a continuous local martingale with} \\ \qquad \langle M^n, M^n\rangle = \langle z^{n,c}, z^{n,c}\rangle \\ \text{(b) } \nu^{n,P} = \left(1+\dfrac{x}{z_-^n}\right)\cdot\nu^{z^n}. \end{cases}$$

b) It remains to apply Ito's formula to 1.21, since Y^n is a C^2 function of z^n and z'^n. Observing that $z'^n = 2 - z^n$, that the stopped measures $(\mu^{z^n})^{T^n}$ and $(\nu^{z^n})^{T^n}$ charge A^n only, and that $f_n((z^n)^{T^n}) = z^n$ and $f_n((z'^n)^{T^n}) = z'^n$ on $[\![0, T^n[\![$, we obtain (relatively to P^n):

$$\begin{aligned} Y^n = &-\frac{2}{z_-^n z_-'^n}\cdot(M^n)^{T^N} - \frac{2x}{z_-^n z_-'^n}1_{A^n}*(\mu^{z^n} - \nu^{n,P}) - \frac{2}{(z_-^n)^2 z_-'^n}\cdot\langle z^{n,c}, z^{n,c}\rangle^{T^n} \\ &- \frac{2x^2}{(z_-^n)^2 z_-'^n}1_{A^n}*\nu^{z^n} + \frac{1}{2}\left(\frac{1}{(z_-^n)^2} - \frac{1}{(z_-'^n)^2}\right)\cdot\langle M^n, M^n\rangle^{T^n} \\ &+ \left(V^n + \frac{2x}{z_-^n z_-'^n}\right)1_{A^n}*\mu^{z^n}. \end{aligned}$$

Using 1.24 and the property $\frac{1}{2}(y^{-2}-y'^{-2})-2y^{-2}y'^{-1} = -2y^{-2}y'^{-2}$ if $y+y' = 2$, we get

1.25
$$\begin{aligned} Y^n = &-\frac{2}{z^n_- z'^n_-}\cdot (M^n)^{T^n} + h(V^n)1_{A^n} * (\mu^{z^n} - \nu^{n,P}) + (V^n - h(V^n))1_{A^n} * \mu^{z^n} \\ &- \frac{2}{(z^n_- z'^n_-)^2}\cdot \langle z^{n,c}, z^{n,c}\rangle^{T^n} \\ &+ \left[\left(h(V^n) + \frac{2x}{z^n_- z'^n_-}\right)\left(1 + \frac{x}{z^n_-}\right) - \frac{2x^2}{(z^n_-)^2 z'^n_-}\right] 1_{A^n} * \nu^{z^n}. \end{aligned}$$

We deduce that the continuous martingale part of Y^n is the first term in the right-hand side of 1.25, so by 1.24a we deduce the claimed form for C^n. Secondly, a simple computation based on 1.21 shows that

1.26
$$\Delta Y^n_t = 1_{A^n}(t, \Delta z^n_t)\, V^n(t, \Delta z^n_t),$$

and so

$$g * \mu^{Y^n} = g(V^n)1_{A^n} * \mu^{z^n}.$$

This relation carries over to compensators: so $g * \nu^n = g(V^n)1_{A^n} * \nu^{n,P}$ and in virtue of 1.24b we deduce the claimed form for ν^n.

Moreover 1.26 implies $\sum_{s\leq \cdot}[\Delta Y^n_s - h(\Delta Y^n_s)] = (V^n - h(V^n))1_{A^n} * \mu^{z^n}$, so if we compare to 1.25 we deduce that the first characteristic B^n is exactly the sum of the last two terms in 1.25, which again is the claimed form for B^n. Finally, the formula giving $\tilde{C}^n$ is just an application of II.2.18. □

We define a predictable function α^n on $\Omega^n \times \mathbb{R}_+ \times \mathbb{R}$ by

1.27
$$\begin{aligned} \alpha^n(\omega, t, x) &= \left[\frac{1 - x/z'^n_{t-}(\omega)}{1 + x/z^n_{t-}(\omega)} - 1\right] 1_{A^n}(\omega, t, x) \\ &= \frac{-2x}{z^n_{t-}(\omega) z'^n_{t-}(\omega)(1 + x/z^n_{t-}(\omega))} 1_{A^n}(\omega, t, x) \end{aligned}$$

(with $\alpha^n = +\infty$ if $x + z^n_{t-} = 0$). Then

1.28
$$\begin{cases} -1 \leq \alpha^n \leq +\infty \\ \alpha^n(t,x) = +\infty \Leftrightarrow x = -z^n_{t-} \text{ and } t \leq T^n \\ \alpha^n(t,x) = -1 \Leftrightarrow x = z'^n_{t-} \text{ and } t \leq T^n. \end{cases}$$

As mentioned in Chapter IV, all versions of $h(\frac{1}{2}; P^n, P'^n)$ and $\imath(\psi; P^n, P'^n)$ coincide on the set $\{z^n_- > 0, z'^n_- > 0\}$, which contains $[\![0, T^n]\!]$. Hence the processes $H^n, I^n(a), G^n(g)$ are given by 1.10 on the interval $[\![0, T^n]\!]$. Therefore we immediately deduce from 1.23 that

1.29 $B^n = -4(H^n)^{T^n} + \bar{B}^n$, where

$$\bar{B}^n = \left(1 + \frac{x}{z_-^n}\right)[h(V^n) - \alpha^n + 2(1 - \sqrt{1 + \alpha^n})^2]1_{A^n}1_{\{\alpha^n < \infty\}} * v^{z^n}$$

$$+ \left(1 - \frac{x}{z_-'^n}\right)1_{A^n}1_{\{\alpha^n = \infty\}} * v^{z^n}.$$

1.30 $\tilde{C}^n = 8(H^n)^{T^n} + \bar{C}^n$, where

$$\bar{C}^n = \left(1 + \frac{x}{z_-^n}\right)[h(V^n)^2 - 4(1 - \sqrt{1 + \alpha^n})^2]1_{A^n}1_{\{\alpha^n < \infty\}} * v^{z^n}$$

$$- 4\left(1 - \frac{x}{z_-'^n}\right)1_{A^n}1_{\{\alpha^n = \infty\}} * v^{z^n} - \sum_{s \leq \cdot} (\Delta B_s^n)^2.$$

1.31 $g * v^n = G^n(g)^{T^n} + \overline{g * v^n}$, where

$$\overline{g * v^n} = \left(1 + \frac{x}{z_-^n}\right)[g(V^n) - g \circ \operatorname{Log}(1 + \alpha^n)]1_{\{-1 < \alpha^n < \infty\}}1_{A^n} * v^{z^n}$$

$$+ \left(1 + \frac{x}{z_-^n}\right)g(V^n)1_{A^n}1_{\{\alpha^n = -1\}} * v^{z^n}.$$

1.32 **Lemma.** *We have* $V^n = \operatorname{Log}(1 + \alpha^n)$ *on the set* $\tilde{A}^n = A^n \cap \left\{-1 + \frac{1}{n} \leq \alpha^n \leq n - 1\right\}$.

Proof. A simple computation shows that

$$z_-'^n - x = 2\frac{z_-^n(1 + \alpha^n)}{\alpha^n z_-'^n + 2}, \quad z_-^n + x = \frac{2(2 - z_-'^n)}{\alpha^n z_-'^n + 2}.$$

Since $\frac{1}{n} \leq z_-^n \leq 2$ and $\frac{1}{n} \leq z_-'^n \leq 2$ on $\tilde{A}^n$, we easily deduce that on this set,

$$1/n^2 \leq z_-'^n - x \leq 2, \quad 1/n^2 \leq z_-^n + x \leq 2.$$

Then $f_n(z_-'^n - x) = z_-'^n - x$ and $f_n(z_-^n + x) = z_-^n + x$ on this set by 1.20. We then deduce the result from the definition 1.22. □

1.33 **Lemma.** *There is a constant* $\beta > 0$ *and, for every* $\varepsilon > 0$ *small enough, a constant* γ_ε*, such that for* $n \geq 2$ (recall that $A \prec B$ means that the process $B - A$ is non-decreasing):

1.34 $$\operatorname{Var}(\bar{B}^n \prec \varepsilon\beta H^n + \beta I^n(1 + \varepsilon),$$

1.35 $$\operatorname{Var}(\overline{C}^n)_t \le \varepsilon\beta H_t^n + 4\beta I^n(1+\varepsilon)_t + \left(\varepsilon + \beta\gamma_\varepsilon \sup_{s\le t} \Delta I^n(e^\varepsilon)_s\right)[(4+\beta)H_t^n + \beta I^n(1+\varepsilon)_t],$$

1.36 $$1_{\{|x|>\varepsilon\}} * \nu^n \prec \gamma_\varepsilon I^n(e^\varepsilon).$$

Proof. In view of 1.10 and 1.27, we have

1.37 $$I^n(a)^{T^n} = \left(1 + \frac{x}{z_-^n}\right)|\alpha^n| 1_{\{-1+(1/a)<\alpha^n<a-1\}^c} 1_{\Lambda^n} 1_{\{\alpha^n<\infty\}} * \nu^{z^n} + \left(1 - \frac{x}{z_-'^n}\right) 1_{\Lambda^n} 1_{\{\alpha^n=\infty\}} * \nu^{z^n}.$$

Since h is bounded and $h(x) = x$ for $|x|$ small enough, and since $V^n = \operatorname{Log}(1+\alpha^n)$ by 1.32 when $|\alpha^n| \le 1/2$ (for $n \ge 2$), on the set Λ^n, there exists a constant $\beta \ge 1$ such that $|h| \le \beta$ and, on Λ^n:

1.38 $$\begin{cases} |h(V^n) - \alpha^n + 2(1 - \sqrt{1+\alpha^n})^2| \le \beta[|\alpha^n|^3 \wedge |\alpha^n|], \\ |h(V^n)^2 - 4(1 - \sqrt{1+\alpha^n})^2| \le \beta[|\alpha^n|^3 \wedge |\alpha^n|], \\ \left(1 + \frac{x}{z_-^n}\right)(\alpha^n)^2 \le \beta\frac{1}{2}(\sqrt{1 + x/z_-^n} - \sqrt{1 - x/z_-'^n})^2 \quad \text{if } |\alpha^n| \le 1. \end{cases}$$

Then 1.34 readily follows from 1.37, 1.38, 1.10, and from the explicit form 1.29 for $\overline{B}^n$, provided $\varepsilon \in (0,1)$.

Secondly, $|V^n| > \varepsilon$ implies either $\alpha^n < -1 + e^{-\varepsilon}$ or $\alpha^n > e^\varepsilon - 1$ on Λ^n, provided $n \ge 2$ and $\varepsilon \in (0, \operatorname{Log} 2]$ (see 1.32 again). Hence, in view of 1.23 and 1.27,

$$1_{\{|x|>\varepsilon\}} * \nu^n \prec \gamma_\varepsilon \left(1 + \frac{x}{z_-^n}\right) |\alpha^n| 1_{\lambda^n} 1_{\{\alpha^n<\infty\}\cap\{-1+e^{-\varepsilon}<\alpha^n<e^\varepsilon-1\}^c} * \nu^{z^n}$$

Then if $\gamma_\varepsilon = 1/(1-e^{-\varepsilon})$, we have $1 \le \gamma_\varepsilon|\alpha^n|$ on $\{-1+e^{-\varepsilon} < \alpha^n < e^\varepsilon - 1\}^c$. Hence

$$1_{\{|x|>\varepsilon\}} * \nu^n \prec \gamma_\varepsilon \left(1 + \frac{x}{z_-^n}\right)|\alpha^n| 1_{\Lambda^n} 1_{\{\alpha^n<\infty\}} * \nu^{z^n}$$

and 1.36 follows from 1.37 when $\varepsilon \in (0, \operatorname{Log} 2]$.

Finally, it follows from 1.37, 1.38 and 1.10 that the variation of the sum of the first two terms in $\overline{C}^n$ is strongly majorized by $\varepsilon\beta H^n + 4\beta I^n(1+\varepsilon)$ (exactly as for 1.34). Moreover if $\varepsilon > 0$ is small enough for having $h(x) = x$ for $|x| \le \varepsilon$, we have $|\Delta B_s^n| \le \varepsilon + \beta\Delta(1_{\{|x|>\varepsilon\}} * \nu^n)_s$ (recall that $|h| \le \beta$). In view of 1.36, we get

1.39 $$|\Delta B_s^n| \le \varepsilon + \beta\gamma_\varepsilon \Delta I^n(e^\varepsilon)_s.$$

On the other hand, $\operatorname{Var}(B^n) \prec \operatorname{Var}(\overline{B}^n) + 4H^n$ by 1.29, so 1.39 and 1.34 yield

$$\sum_{s\le t} |\Delta B_s^n|^2 \le \left[\varepsilon + \beta\gamma_\varepsilon\, sup_{s\le t}\, \Delta I^n(e^\varepsilon)_s\right] [\varepsilon\beta H_t^n + 2\beta I^n(1+\varepsilon)_t + 4H_t^n],$$

and 1.35 follows. □

Finally, to the end of applying the criterion V.2.3 for contiguity, we need to evaluate the processes $i(1_{[0,\beta]}; P^n, P'^n)$ and $i(1_{[0,\beta]}; P'^n, P^n)$ which occur in V.2.3 (and are defined by IV.1.46 with $\psi = 1_{[0,\beta]}$).

1.40 **Lemma.** *There are versions $i^n(\beta)$ of $i(1_{[0,\beta]}; P^n, P'^n)$ and $i'^n(\beta)$ of $i(1_{[0,\beta]}; P'^n, P^n)$ which satisfy for all $a > 1$:*

1.41 $$i^n\left(\frac{1}{a}\right) \le \frac{a}{a-1} I^n(a), \quad i'^n\left(\frac{1}{a}\right) \le \frac{a}{a-1} I^n(a),$$

1.42 $$I^n(b) \le i^n\left(\frac{1}{a}\right) + i'^n\left(\frac{1}{a}\right) + a 1_{\{|x| > \operatorname{Log} b\}} * \nu^n \text{ on } [\![0, T^n]\!], \text{ for } 1 < b \le a \le n.$$

Proof. We will consider the "strict" version of $i^n(\beta)$, i.e. the only one which satisfies $i^n(\beta) = 1_{\Gamma''^n} \cdot i^n(\beta)$ where $\Gamma''^n = \{z^n_- > 0, z'^n_- > 0\}$, and similarly for $i'^n(\beta)$.

a) In view of IV.1.49 (or V.2.2), we have

1.43 $$\begin{aligned} i^n\left(\frac{1}{a}\right) &= \left(1 - \frac{x}{z'^n_-}\right) 1_{\{1 + x/z^n_- \le (1 - x/z'^n_-)/a\}} * \nu^{z^n} \\ &\le \frac{a}{a-1} \left| \left(1 - \frac{x}{z'^n_-}\right) - \left(1 + \frac{x}{z^n_-}\right) \right| 1_{\{1 + x/z^n_- \le (1 - x/z'^n_-)/a\}} * \nu^{z^n} \\ &\le \frac{a}{a-1} \frac{2|x|}{z^n_- z'^n_-} 1_{\{1/a < (1 + x/z^n_-)/(1 - x/z'^n_-) < a\}^c} * \nu^{z^n} \le \frac{a}{a-1} I^n(a) \end{aligned}$$

by 1.10, so the first inequality in 1.41 holds. Since $\Delta z'^n = -\Delta z^n$, we have $g(x) * \nu^{z'^n} = g(-x) * \nu^{z^n}$, and thus

1.44 $$i'^n\left(\frac{1}{a}\right) = \left(1 + \frac{x}{z^n_-}\right) 1_{\{1 - x/z'^n_- \le (1 + x/z^n_-)/a\}} * \nu^{z^n}$$

and the same computation as above shows the second inequality 1.41.

b) Using again 1.10, which holds for every version of $I^n(b)$ on $[\![0, T^n]\!]$, and if $1 < b \le a \le n$, we can write on the set $[\![0, T^n]\!]$:

$$\begin{aligned} I^n(b) &= \left| \left(1 + \frac{x}{z^n_-}\right) - \left(1 - \frac{x}{z'^n_-}\right) \right| 1_{\{1 + \alpha^n \le (1/b) \text{ or } 1 + \alpha^n \ge b\}} * \nu^{z^n} \\ &\le \left(1 - \frac{x}{z'^n_-}\right) 1_{\{1 + \alpha^n \ge a\}} * \nu^{z^n} + \left(1 + \frac{x}{z^n_-}\right) 1_{\{1 + \alpha^n \le (1/a)\}} * \nu^{z^n} \\ &\quad + a\left(1 + \frac{x}{z^n_-}\right) 1_{\{|V^n| > \operatorname{Log} b\}} * \nu^{z^n} \end{aligned}$$

$\left(\text{recall that } \operatorname{Log}(1 + \alpha^n) = V^n \text{ if } \frac{1}{a} \le 1 + \alpha^n \le a, \text{ when } a \le n, \text{ by } 1.32\right)$. Then we deduce 1.42 from 1.43, 1.44 and 1.23. □

§ 1c. Proofs of Theorems 1.12 and 1.16

We begin by proving 1.16, and the implication (ii) ⇒ (i) in 1.12.

1.45 **Lemma.** *Under* $[L\text{-}D] + [H\text{-}D]$, *we have* $(P_t^n) \lhd (P_t'^n)$ *for all* t *such that* $[t, \infty) \cap D \neq \emptyset$, *and*

1.46 $$P^n(T^n \leq t) \to 0 \quad \text{as } n \uparrow \infty \quad \text{for all } t \in D.$$

Proof. In view of 1.41, $[L\text{-}D]$ implies for all $t \in D$:

$$\lim_{a \uparrow \infty} \limsup_n P^n\left(\imath'^n\left(\frac{1}{a}\right)_t > \eta\right) = 0 \quad \text{for all } \eta > 0.$$

Moreover, $[H\text{-}D]$ implies that the sequence $(H_t^n | P^n)$ is $\mathbb{R}$-tight, while $P_0^n = P_0'^n$ by hypothesis. Hence Theorem V.2.3, applied with $T^n = t$ and (P'^n, P^n) instead of (P^n, P'^n), yields $(P_t^n) \lhd (P_t'^n)$ for all $t \in D$, and thus also for all t such that $[t, \infty) \cap D \neq \emptyset$.

Finally, V.1.17 and V.1.19 imply that the sequences $(\sup_{s \leq t}(1/z_s^n) | P^n)$ and $(\sup_{s \leq t}(1/z_s'^n) | P^n)$ are $\mathbb{R}$-tight for $t \in D$, which clearly implies 1.46. □

Let C, M be as in 1.12, and set $X = M - C/2$, which is a continuous PII with characteristics (B, C, ν), with $B = -C/2$ and $\nu = 0$.

1.47 **Lemma.** *Under* $[L\text{-}D] + [H\text{-}D]$ *we have* $Y^n \xrightarrow{\mathscr{L}(D|P^n)} X$, *and* $Y^n \xrightarrow{\mathscr{L}(P^n)} X$ *when* D *is dense in* $\mathbb{R}_+$.

Proof. It is a simple application of Theorems VIII.3.6 and VIII.3.8. Indeed, consider the conditions

$[\beta_5\text{-}D]$ $\quad B_t^n \xrightarrow{P^n} B_t = -\frac{1}{2}C_t \quad$ for all $t \in D$,

$[\text{Sup-}\beta_5]$ $\quad \sup_{s \leq t} |B_s^n - B_s| \xrightarrow{P^n} 0 \quad$ for all $t \in \mathbb{R}_+$,

$[\gamma_5\text{-}D]$ $\quad \tilde{C}_t^n \xrightarrow{P^n} C_t \quad$ for all $t \in D$,

$[\hat{\delta}_5\text{-}D]$ $\quad 1_{\{|x| > \varepsilon\}} * \nu_t^n \xrightarrow{P^n} 0 \quad$ for all $t \in D$, $\varepsilon > 0$.

Then it suffices to prove that $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$, and also $[\text{Sup-}\beta_5]$ when D is dense, hold.

In view of $[L\text{-}D]$ and 1.36, we obviously have $[\hat{\delta}_5\text{-}D]$. 1.34 and 1.35 and $[L\text{-}D] + [H\text{-}D]$ also clearly imply

1.48 $$\operatorname{Var}(\bar{B}^n)_t \xrightarrow{P^n} 0, \quad \operatorname{Var}(\bar{C}^n)_t \xrightarrow{P^n} 0 \quad \text{for all } t \in D$$

(because in 1.34 and 1.35, $\varepsilon > 0$ is arbitrarily small, and $\sup_{s \leq t} \Delta I^n(e^\varepsilon)_s \leq I^n(e^\varepsilon)_t$). Then 1.29 and 1.30, plus $[H\text{-}D]$ and 1.46, obviously yields $[\beta_5\text{-}D]$ and $[\gamma_5\text{-}D]$.

Finally, assume that D is dense in $\mathbb{R}_+$. Since H^n and C are increasing and C is continuous, the usual argument (see e.g. VIII.2.17) shows that $\sup_{s\le t}|H^n_s - C_s/8| \xrightarrow{P^n} 0$ for all $t\in\mathbb{R}_+$. Then 1.48 and 1.29 imply [Sup-β_5]. $\square$

Recalling that $Z^n = e^{Y^n}$ on $[\![0, T^n[\![$ (see 1.21), and using again 1.46, we immediately deduce from 1.47 the implication (ii) $\Rightarrow$ (i) of 1.12, and the finite-dimensional convergence $Z^n \xrightarrow{\mathscr{L}(D|P^n)} Z = e^{M-C/2}$ under $[L\text{-}D] + [H\text{-}D]$ when D is arbitrary in $\mathbb{R}_+$.

Now, it is obvious that $E(Z_t) = 1$ for all t. So, under $[L\text{-}D] + [H\text{-}D]$, we deduce from V.1.12 that $(P_t'^n) \lhd (P_t^n)$ for all $t\in D$ (and thus for all t such that $[t,\infty)\cap D \neq \emptyset$). Then, according to the argument in Remark 1.13, we have 1.14: this allows to exchange P^n and P'^n, and thus the density process $1/Z^n$ of P^n with respect to P'^n converges in law (along D, if D is not dense), relatively to P'^n, toward $Z = e^{M-C/2}$. Therefore Z^n converges in law, relatively to P'^n, toward $1/Z = e^{-M+C/2}$. Since $\mathscr{L}(M) = \mathscr{L}(-M)$, $1/Z$ and $Z' = e^{M+C/2}$ have the same distribution, and this finishes to prove 1.16, and the last claims in 1.12.

It remains to prove that (i) $\Rightarrow$ (ii), with $D = \mathbb{R}_+$, in 1.12.

1.49 **Lemma.** *Assume* 1.12(i).

a) $(P_t^n) \lhd (P_t'^n)$ *and* $(P_t'^n) \lhd (P_t^n)$ *and* $P^n(T^n \le t) \to 0$ *for all* $t\in\mathbb{R}_+$.

b) *We have* $[\gamma_5\text{-}\mathbb{R}_+]$ *and* $[\hat{\delta}_5\text{-}\mathbb{R}_+]$ (with the notation of the proof of 1.47).

Proof. a) We have $Z_t^n \xrightarrow{\mathscr{L}(P^n)} Z_t$ and $E(Z_t) = 1$ and $Z_t > 0$ for all t, so V.1.12 and V.1.15 imply the first two claims. That $P^n(T^n \le t) \to 0$ then follows, exactly like in the end of the proof of 1.45.

b) Since 1.46 holds and $Z^n = e^{Y^n}$ on $[\![0, T^n[\![$ (see 1.21), the hypothesis cleary yields

1.50 $$Y^n \xrightarrow{\mathscr{L}(P^n)} X = M - \frac{C}{2}.$$

Then $[\hat{\delta}_5\text{-}\mathbb{R}_+]$ and also

1.51 $$\sup_{s\le t}|\Delta Y_s^n| \xrightarrow{P^n} 0$$

follow from VIII.3.8 and VIII.3.5 and from the continuity of X.

There exists a constant $\beta' > 0$ such that

$$\left(1+\frac{x}{Z_-^n}\right)|\alpha^n| 1_{\{\alpha^n<\infty\}} + 2\left(1-\frac{x}{Z_-'^n}\right)1_{\{\alpha^n=\infty\}}$$
$$\le \beta'\frac{1}{2}\left(\sqrt{1+\frac{x}{Z_-^n}} - \sqrt{1-\frac{x}{Z_-'^n}}\right)^2 \quad \text{if } |\alpha^n| > 1.$$

Then, using 1.38 and 1.29 and 1.10,

$$\operatorname{Var}(\overline{B}^n) \le \beta(1+\beta')H^n, \quad \operatorname{Var}(B^n) \le (4+\beta+\beta\beta')H^n.$$

By $(P_t^n) \lhd (P_t'^n)$ and V.2.3, the sequence $(H_t^n | P^n)$ is $\mathbb{R}$-tight, and thus so is the sequence $(\mathrm{Var}(B^n)_t | P^n)$ for every $t \in \mathbb{R}_+$. This, plus 1.50, allow to apply Theorem VI.6.26, and we have convergence of the quadratic variations:

1.52 $$[Y^n, Y^n] \xrightarrow{\mathscr{L}(P^n)} [X, X] = C.$$

We have the decomposition $Y^n = M^n + B^n - \check{Y}^n$, where $\check{Y}_t^n = \sum_{s \le t} [\Delta Y_s^n - h(\Delta Y_s^n)]$. We have seen before 1.39 that $|\Delta B_t^n| \le \varepsilon + \beta \Delta(1_{\{|x| > \varepsilon\}} * \nu^n)_t$ for all $\varepsilon > 0$ small enough, hence $[\hat{\delta}_5\text{-}\mathbb{R}_+]$ yields

$$\sup_{s \le t} |\Delta B_s^n| \xrightarrow{P^n} 0.$$

Moreover $P^n(\sup_{s \le t} |\check{Y}_s^n| > 0) \to 0$ by 1.51, so $\sup_{s \le t} |\Delta M_s^n| \xrightarrow{P^n} 0$ holds as well. Furthermore, on the set $\{\sup_{s \le t} |\check{Y}_s^n| = 0\}$ we have $[Y^n, Y^n] = [M^n, M^n] + [B^n, B^n] + 2[M^n, B^n]$, and so on this set,

$$\sup_{s \le t} |[Y^n, Y^n]_s - [M^n, M^n]_s| \le \mathrm{Var}(B^n)_t \sup_{s \le t} (|\Delta B_s^n| + 2|\Delta M_s^n|).$$

Since $(\mathrm{Var}(B^n)_t | P^n)$ is $\mathbb{R}$-tight, we deduce from what precedes that

1.53 $$\sup_{s \le t} |[Y^n, Y^n]_s - [M^n, M^n]_s| \xrightarrow{P^n} 0,$$

and so 1.52 yields in particular that $[M^n, M^n]_t \xrightarrow{P^n} C_t$ for all $t \in \mathbb{R}_+$. But then, it follows from $[\hat{\delta}_5\text{-}\mathbb{R}_+]$ and VIII.3.6a that $[\gamma_5\text{-}\mathbb{R}_+]$ holds. □

Proof of (i) ⇒ (ii) *in 1.12.* By 1.42, we have for $1 + \varepsilon \le a \le n$:

1.54 $$P^n(I^n(1+\varepsilon)_t > \eta) \le P^n\left(\imath^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right) + P^n\left(\imath'^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right) + P^n(T^n \le t) + P^n\left(1_{\{|x| > \mathrm{Log}(1+\varepsilon)\}} * \nu_t^n > \frac{\eta}{3a}\right).$$

Now, 1.49(i) and V.2.3 imply

1.55 $$\begin{cases} \lim_{a \uparrow \infty} \limsup_n P^n\left(\imath'^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right) = 0 \\ \lim_{a \uparrow \infty} \limsup_n P'^n\left(\imath^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right) = 0. \end{cases}$$

Let $F(n, a) = \left\{\imath^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right\}$. Observe that $F(n, b) \subset F(n, a)$ if $a \le b$. 1.55 implies the existence of two sequences $a_k \uparrow \infty$ and $\{n_k\} \subset \mathbb{N}$ such that $P'^n(F(n, a_k)) \le 1/k$ for all $n \ge n_k$. If $\lim_{a \uparrow \infty} \limsup_n P^n(F(n, a)) = \theta > 0$, we have $\limsup_n P^n(F(n, a)) \ge \theta$ for all a, and so there exists $n'_k \ge n_k$ such that $P^{n'_k}(F(n'_k, a_k)) \ge \theta/2$ for all k. Since $P'^{n'_k}(F(n'_k, a_k)) \le 1/k$, this would contradicts the property $(P_t^n) \lhd (P_t'^n)$ proved in

1.49. Hence

1.56
$$\lim_{a\uparrow\infty}\limsup_n P^n\left(\imath^n\left(\frac{1}{a}\right)_t > \frac{\eta}{3}\right) = 0.$$

Then we easily deduce from this, and from 1.54, 1.55 and 1.49 that $[L\text{-}\mathbb{R}_+]$ holds. Moreover, as already seen in the proof of 1.49, the sequence $(H_t^n|P^n)$ is $\mathbb{R}$-tight for all $t\in\mathbb{R}_+$. Using 1.35, where $\varepsilon > 0$ is arbitrarily small, and $[L\text{-}\mathbb{R}_+]$, we see that $\mathrm{Var}(\bar{C}^n)_t \xrightarrow{P^n} 0$ for all t. Then, in view of 1.30, $[H\text{-}\mathbb{R}_+]$ is an immediate consequence of 1.49b. □

§ 1d. Convergence to the Exponential of a Continuous Martingale

1. In this subsection we consider a natural extension of Theorem 1.12: namely the case where the limiting process Z has the form $Z = e^{M-C/2}$, M being a continuous martingale with $M_0 = 0$ and $C = \langle M,M\rangle$ (equivalently, $Z = \mathscr{E}(M)$ is the Doléans-Dade exponential of M).

To this end, we will apply the necessary and sufficient conditions established in § IX.4d, and we begin by listing a series of specifications on the limiting process.

1.57 **Hypotheses.** a) $(\Omega,\mathscr{F},\mathbf{F}) = (\mathbb{D}(\mathbb{R}),\mathscr{D}(\mathbb{R}),\mathbf{D}(\mathbb{R}))$ is the canonical space, with the canonical process denoted by X.

b) C is an adapted continuous increasing process with $C_0 = 0$, defined on $(\Omega,\mathscr{F},\mathbf{F})$.

c) There is a continuous increasing function $t\rightsquigarrow F_t$ with $F_0 = 0$, such that $C(\alpha)\prec F$ (i.e., $F - C(\alpha)$ is increasing) for all $\alpha\in\Omega$.

d) $\alpha\rightsquigarrow C_t(\alpha)$ is Skorokhod-continuous for all $t\geq 0$.

e) There is a unique probability measure P on $(\Omega,\mathscr{F})$ under which $M = X + C/2$ is a continuous martingale with $M_0 = 0$ and $\langle M,M\rangle = C$: in other words, the martingale problem $\mathscr{s}\left(\sigma(X_0),X|\varepsilon_0;-\frac{C}{2},C,0\right)$ has P as its unique solution. □

1.58 **Lemma.** *Under* 1.57, *the martingale problem* $\mathscr{s}\left(\sigma(X_0),X|\varepsilon_0;\frac{C}{2},C,0\right)$ *has a unique solution P'. Moreover $P'\overset{\text{loc}}{\ll}P$ and the density process of P' with respect to P is $Z = e^X$, which thus is a P-martingale.*

Proof. a) We first prove that Z is a P-martingale. Observe that $Z = \mathscr{E}(M)$, where $M = X + C/2$ is a P-local martingale. Furthermore $\langle 2M,2M\rangle = 4C$, so $\mathscr{E}(2M) = e^{2M-2C} = Z^2e^{-C}$. Since $\mathscr{E}(2M)$ is also a nonnegative local martingale, and hence a supermartingale, we deduce from 1.57c that

$$\sup_{s\leq t} E(Z_s^2) = \sup_{s\leq t} E(e^{C_t}\mathscr{E}(2M)_t) \leq e^{F_t} < \infty.$$

Therefore the family $(Z_s|P)_{s\le t}$ is uniformly integrable for all t, and it follows that the local martingale Z is a martingale (see I.1.47).

b) For each $n\in\mathbb{N}^*$ let $P'^n(d\omega)=Z_n(\omega)P(d\omega)$, which by (a) defines a probability measure P'^n on $(\Omega,\mathscr{F})$. If $A\in\mathscr{F}$ we have for all $a>0$:

$$\sup_n P'^n(A)\le \sup_n\{E_P(1_A Z_n 1_{\{Z_n\le a\}})+E_P(Z_n 1_{\{Z_n>a\}})\}\le aP(A)+\frac{1}{a}.$$

In particular, with the notation of Chapter VI, and if $\frac{1}{a}<\varepsilon$:

$$\sup_n P'^n\left(\sup_{s\le N}|X_s|\ge K\right)\le aP\left(\sup_{s\le N}|X_s|\ge K\right)+\frac{1}{a}\le\varepsilon\quad\text{for } K \text{ large enough;}$$

$$\sup_n P'^n(w'_N(X,\theta)>\delta)\le aP(w'_N(X,\theta)>\delta)+\frac{1}{a}\le\varepsilon\quad\text{for } \theta>0 \text{ small enough.}$$

Hence by VI.3.21 the sequence (P'^n) is tight. Moreover $P'^n(A)=E_P(1_A Z_t)$ for all $n\ge t$, when $A\in\mathscr{F}_t$: we easily deduce that the sequence (P'^n) has a unique limit point P', with $P'\overset{\text{loc}}{\ll}P$, and that Z is the density process of P' with respect to P.

c) Now we prove that $P'\in\mathscr{s}\left(\sigma(X_0),X|\varepsilon_0;\frac{C}{2},C,0\right)$: this is a straightforward application of Girsanov's Theorem III.3.24, once noticed that $Z=\mathscr{E}(M)$ and $\langle M,M\rangle=C$, so $(1/Z_-)\cdot\langle Z,M\rangle=C$ (relatively to P).

d) It remains to prove the uniqueness of the solution P' to $\mathscr{s}\Big(\sigma(X_0),X|\varepsilon_0;$ $\frac{C}{2},C,0\Big)$. Let indeed $\tilde{P}'$ be another solution, and $M'=X-C/2$, and $Z'=\mathscr{E}(M')$. One proves as in (a) that Z' is a $\tilde{P}'$-martingale, then as in (b) that there is a probability measure $\tilde{P}$ such that $\tilde{P}\overset{\text{loc}}{\ll}\tilde{P}'$ and that Z' is the density process of $\tilde{P}$ with respect to $\tilde{P}'$. Finally, as in (c), we obtain $\tilde{P}\in\mathscr{s}\left(\sigma(X_0),X|\varepsilon_0;-\frac{C}{2},C,0\right)$. Therefore 1.57e yields $\tilde{P}=P$: so $\tilde{P}'\overset{\text{loc}}{\ll}P$ and the density process of $\tilde{P}'$ with respect to P is $1/Z'$. Since $Z'=e^{M'-C/2}$ we have $1/Z'=e^X=Z$, and thus $\tilde{P}'=P'$. □

The other notation and assumptions about $\Omega^n,\dots$, are the same than in the rest of the section. The process $\operatorname{Log}\left(Z^n\vee\frac{1}{2n}\right)$ is P^n-a. s. with paths in $\mathbb{D}(\mathbb{R})$, so the process $C\circ\operatorname{Log}\left(Z^n\vee\frac{1}{2n}\right)$ is P^n-a. s. well-defined.

1.59 **Theorem.** *Assume* 1.57. *There is equivalence between*:

(i) $\mathscr{L}(Z^n|P^n)\to\mathscr{L}(e^X|P)$ (P is defined in 1.57e).

(ii) *The following two conditions hold*:

$[L\text{-}\mathbb{R}_+]$ $\quad I^n(1+\varepsilon)_t \xrightarrow{P^n} 0 \quad$ for all $t \geq 0, \varepsilon > 0$.

$[H_1\text{-}\mathbb{R}_+]$ $\quad H_t^n - \frac{1}{8} C_t \circ \operatorname{Log}\left(Z^n \vee \frac{1}{2n}\right) \xrightarrow{P^n} 0 \quad$ for all $t \geq 0$.

In this case, we also have $\mathscr{L}(Z^n | P'^n) \to \mathscr{L}(e^X | P')$, *where* P' *is the unique probability measure on* $\mathbb{D}(\mathbb{R})$ *under which* $X - C/2$ *is a continuous martingale with* $X_0 = 0$ *a.s. and* $\left\langle X - \frac{C}{2}, X - \frac{C}{2} \right\rangle = X$ (see 1.58), *and* $(P_t^n) \lhd (P_t'^n)$ *and* $(P_t'^n) \lhd (P_t^n)$ *for all* $t \in \mathbb{R}_+$.

Observe that 1.12 is a particular case of 1.59: indeed, if C is deterministic, 1.57 is fulfilled and $[H_1\text{-}\mathbb{R}_+] = [H\text{-}\mathbb{R}_+] \Leftrightarrow [H\text{-}D]$ and $[L\text{-}\mathbb{R}_+] \Leftrightarrow [L\text{-}D]$ for any dense subset D of $\mathbb{R}_+$. Observe also that there is no finite-dimensional analogue to 1.59, as 1.16.

Here again, we will see in § 1e another formulation in terms of Hellinger processes only.

Proof. The proof is very similar to that of Theorem 1.12, and below we just mention the necessary changes to be made.

1) The conclusions of Lemma 1.45 are valid under $[L\text{-}\mathbb{R}_+]$ and $[H_1\text{-}\mathbb{R}_+]$: the only point which is not strictly identical concerns $\mathbb{R}$-tightness of the sequence $(H_t^n | P^n)$. Here, in view of 1.57c and $[H_1\text{-}\mathbb{R}_+]$, we have $P^n(H_t^n > F_t + \varepsilon) \to 0$ as $n \uparrow \infty$ for all $\varepsilon > 0$, and the $\mathbb{R}$-tightness follows.

2) Under $[L\text{-}\mathbb{R}_+]$ and $[H_1\text{-}\mathbb{R}_+]$, we have $\mathscr{L}(Y^n | P^n) \to P$. To see that, we apply Theorem IX.4.44, in which the strong majoration IX.4.39, the pointwise continuity IV.4.40 and the uniqueness of the solution to the martingale problem $\mathscr{s}\left(\sigma(X_0), X | \varepsilon_0; -\frac{C}{2}, C, 0\right)$ hold by 1.57 (we have $B = -C/2$). So it suffices to prove the following:

$[\text{Sup-}\beta_7]$ $\quad \sup_{s \leq t} \left| B_s^n - \frac{1}{2} C_s \circ Y^n \right| \xrightarrow{P^n} 0 \quad$ for $t \geq 0$

$[\gamma_7\text{-}\mathbb{R}_+]$ $\quad \tilde{C}_t^n - C_t \circ Y^n \xrightarrow{P^n} 0 \quad$ for $t \geq 0$

$[\hat{\delta}_7]$ $\quad 1_{\{|x| > \varepsilon\}} * \nu_t^n \xrightarrow{P^n} 0 \quad$ for $t \geq 0, \varepsilon > 0$.

Due to 1.36 and $[L\text{-}\mathbb{R}_+]$, $[\hat{\delta}_7]$ is obvious, and 1.48 holds because of 1.34, 1.35, $[L\text{-}\mathbb{R}_+]$ and of the $\mathbb{R}$-tightness of $(H_t^n | P^n)$.

We have $Y^n = \operatorname{Log} Z^n$ and $Z^n \geq 1/2n$ on $[\![0, T^n[\![$, so $C_t \circ Y^n = C_t \circ \operatorname{Log}\left(Z^n \vee \frac{1}{2n}\right)$ if $t < T^n$. Therefore 1.46 and $[H_1\text{-}\mathbb{R}_+]$ give for all $t \geq 0$:

1.60 $$(H^n)_t^{T^n} - C_t \circ Y^n \xrightarrow{P^n} 0.$$

So $[\gamma_7\text{-}\mathbb{R}_+]$ follows from 1.60 and 1.30 and 1.48. Moreover, if we apply Lemma IX.3.4 and the strong majoration 1.57c, we deduce from 1.60 that

$$\sup_{s\leq t}|(H^n)_s^{T^n} - C_s\circ Y^n| \xrightarrow{P^n} 0$$

for all $t\geq 0$. Thus $[\text{Sup-}\beta_7]$ follows from 1.48 again and from 1.29.

3) Applying again $Z^n = e^{Y^n}$ on $[\![0,T^n[\![$ and 1.46, we deduce from 2) that $\mathscr{L}(Z^n|P^n)\to\mathscr{L}(e^X|P)$ under $[L\text{-}\mathbb{R}_+] + [H_1\text{-}\mathbb{R}_+]$, so we have proved the implication (ii) $\Rightarrow$ (i).

4) Conversely, assume that (i): $\mathscr{L}(Z^n|P^n)\to\mathscr{L}(e^X|P)$, holds. Firstly, we know by 1.58 that $E_p(e^{X_t}) = 1$, while $e^{X_t}>0$ and (i) yields $\mathscr{L}(Z_t^n|P^n)\to\mathscr{L}(e^{X_t}|P)$. Then V.1.12 and V.1.14 imply $(P_t^n)\lhd(P_t'^n)$ and $(P_t'^n)\lhd(P_t^n)$ for all $t\in\mathbb{R}_+$, and $P^n(T^n\leq t)\to 0$ (i.e. 1.46) follows as in 1.45.

We have just proved 1.49a. That $[\hat\delta_7]$ holds is proved as in 1.49. We also prove as in 1.49 that 1.53 holds, and that $(\text{Var}(B^n)_t|P^n)$ is $\mathbb{R}$-tight for all $t\in\mathbb{R}_+$. Therefore, VI.6.26 implies

$$\mathscr{L}(Y^n,[Y^n,Y^n])|P^n)\to\mathscr{L}[(X,C)|P].$$

Since $C_t(\cdot)$ is Skorokhod-continuous, we deduce that

$$\mathscr{L}\{(C_t\circ Y^n,[Y^n,Y^n]_t)|P^n\}\to\mathscr{L}[(C_t,C_t)|P],$$

and in particular $\mathscr{L}(C_t\circ Y^n - [Y^n,Y^n]_t|P^n)\to\mathscr{L}(C_t - C_t|P) = \varepsilon_0$. In other words, $[Y^n,Y^n]_t - C_t\circ Y^n \xrightarrow{P^n} 0$. Then, in view of 1.53, we deduce that

$$[M^n,M^n]_t - C_t\circ Y^n \xrightarrow{P^n} 0.$$

With the notation IX.4.37, this is exactly $[\hat\gamma_7\text{-}\mathbb{R}_+]$. Since $[\hat\delta_7]$ holds, we deduce from IX.4.43 that $[\gamma_7\text{-}\mathbb{R}_+]$ holds.

At this point, one proves exactly as for 1.12 that $[L\text{-}\mathbb{R}_+]$ holds. Finally it follows from 1.35 and $[L\text{-}\mathbb{R}_+]$ that $\text{Var}(\bar C^n)_t \xrightarrow{P^n} 0$ for all $t\geq 0$, and in view of 1.30, $[H_1\text{-}\mathbb{R}_+]$ is a consequence of $[\gamma_7\text{-}\mathbb{R}_+]$.

5) It remains to prove that (i) implies $\mathscr{L}(Z^n|P'^n)\to\mathscr{L}(e^X|P')$. For this, we apply 1.4: if $\tilde P = \mathscr{L}(e^X|P)$ (so $P = \mathscr{L}(\text{Log}\,X|\tilde P)$) we have that $\mathscr{L}(Z^n|P'^n)\to\tilde P'$, with $\tilde P' \overset{\text{loc}}{\ll} \tilde P$ and where the density process of $\tilde P'$ with respect to $\tilde P$ is the canonical process. This amounts to saying that if $Q = \mathscr{L}(\text{Log}\,X|\tilde P')$, we have $Q\overset{\text{loc}}{\ll} P$ and the density process of Q with respect to P is $Z = e^X$. Hence we deduce from 1.58 that $Q = P'$, and the claim follows. □

§ 1e. Convergence in Terms of Hellinger Processes

Here we wish to replace the "Lindeberg" condition $[L\text{-}D]$ appearing in 1.12 or 1.59 by conditions on Hellinger processes. So, for every $n\in\mathbb{N}^*$, $\alpha\in(0,1)$, we consider:

1.61 $h^n(\alpha) = h^n(\alpha; P^n, P'^n)$ is any version of the Hellinger process of order α between P^n and P'^n. □

Recall that, with the same notation as in 1.10, a version of $h^n(\alpha)$ is

1.62 $$h^n(\alpha) = \frac{2\alpha(1-\alpha)}{(z^n_- z'^n_-)^2} \cdot \langle z^{n,c}, z^{n,c} \rangle + \varphi_\alpha(1 + x/z^n_-, 1 - x/z'^n_-) * \nu^{z^n},$$

where $\varphi_\alpha(u, v) = \alpha u + (1-\alpha)v - u^\alpha v^{1-\alpha}$ (and $h^n(\frac{1}{2}) = H^n$).

Next, in Theorems 1.12 and 1.59 there is a continuous increasing process C (deterministic in 1.12); in view of Remark 1.15, $C/8$ is the Hellinger process of order 1/2 between two probability measures whose relative density processes are continuous. But in virtue of IV.1.24, in this case we have $h(\alpha) = 4\alpha(1-\alpha)h(1/2)$. So here it is natural to set

1.63 $$h(\alpha) = \frac{\alpha(1-\alpha)}{2} C$$

(and so $h(1/2) = C/8$).

Firstly we give another version for Theorems 1.12 and 1.16.

1.64 **Theorem.** *Let $t \rightsquigarrow C_t$ be a non-decreasing continuous function with $C_0 = 0$, and $h(\alpha)$ be defined by 1.63. Let M be a continuous martingale with $M_0 = 0$ and $\langle M, M \rangle_t = C_t$, and let us introduce the conditions*

$[h^\alpha\text{-}D]$ $\quad h^n(\alpha)_t \xrightarrow{P^n} h(\alpha)_t$ for all $t \in D$.

a) *If D is a dense subset of $\mathbb{R}_+$, there is equivalence between*
(i) $Z^n \xrightarrow{\mathscr{L}(P^n)} Z = e^{M - C/2}$ (implying $Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' = e^{M + C/2}$ by 1.12).
(ii) $[h^\alpha\text{-}D]$ *holds for all* $\alpha \in (0,1)$;
(iii) $[h^\alpha\text{-}D]$ *holds for* $\alpha = \frac{1}{2}$, $\alpha = \beta$, $\alpha = 1 - \beta$, *for some* $\beta \in (0, \frac{1}{2})$.

b) *Assume that $[h^\alpha\text{-}D]$ holds for $\alpha = \frac{1}{2}$, $\alpha = \beta$, $\alpha = 1 - \beta$ for some $\beta \in (0, \frac{1}{2})$. Then $Z^n \xrightarrow{\mathscr{L}(D|P^n)} Z$ and $Z^n \xrightarrow{\mathscr{L}(D|P'^n)} Z'$.*

One can remark that Condition (iii) is symmetrical between P^n and P'^n, because $h(\beta; P^n, P'^n) = h(1-\beta, P'^n, P^n)$.

In a similar fashion, we can complement Theorem 1.59 as such:

1.65 **Theorem.** *Assume* 1.57, *define* $h(\alpha)$ *by* 1.63, *and introduce the conditions*:

$[h_1^\alpha\text{-}D]$ $\quad h^n(\alpha)_t - h(\alpha)_t \circ \mathrm{Log}\left(Z^n \wedge \frac{1}{2^n}\right) \xrightarrow{P^n} 0$ for all $t \in D$.

Then there is equivalence between
(i) $\mathscr{L}(Z^n | P^n) \to \mathscr{L}(e^X | P)$;
(ii) $[h_1^\alpha\text{-}\mathbb{R}_+]$ *holds for all* $\alpha \in (0,1)$;
(iii) $[h_1^\alpha\text{-}\mathbb{R}_+]$ *holds for* $\alpha = \frac{1}{2}$, $\alpha = \beta$, $\alpha = 1 - \beta$ *for some* $\beta \in (0, \frac{1}{2})$.

Proof of 1.64 and 1.65. a) Note that $[H\text{-}D] = [h^{1/2}\text{-}D]$ and $[H_1\text{-}D] = [h_1^{1/2}\text{-}D]$. Then in view of Theorems 1.12, 1.16 and 1.59 it is clearly enough to prove the following: assuming $[H\text{-}D]$ (resp. $[H_1\text{-}D]$), then for every $\beta \in (0, \frac{1}{2})$ we have

1.66 $$\begin{cases} [L\text{-}D] \Rightarrow [h^\alpha\text{-}D] \quad (\text{resp. } [h_1^\alpha\text{-}D]) \text{ for all } \alpha \in (0,1); \\ [h^\beta\text{-}D] + [h^{1-\beta}\text{-}D] \quad (\text{resp. } [h_1^\beta\text{-}D] + [h_1^{1-\beta}\text{-}D]) \Rightarrow [L\text{-}D]. \end{cases}$$

Set $\psi_\alpha = \varphi_\alpha - 4\alpha(1-\alpha)\varphi_{1/2}$ and $\theta_a(u,v) = |u - v| 1_{\{1/a < u/v < a\}^c}$. 1.10 and 1.62 yield

$$I^n(a) = \theta_a(1 + x/z_-^n, 1 - x/z_-'^n) * \nu^{z^n}$$

$$h^n(\alpha) = 4\alpha(1-\alpha)H^n + \psi_\alpha(1 + x/z_-^n, 1 - x/z_-'^n) * \nu^{z^n},$$

and $h(\alpha) = 4\alpha(1-\alpha)h(1/2)$ by 1.63. Hence, under $[H\text{-}D]$ (resp. $[H_1\text{-}D]$),

1.67 $$\begin{cases} [L\text{-}D] \Leftrightarrow \theta_a(1 + x/z_-^n, 1 - x/z_-'^n) * \nu_t^{z^n} \xrightarrow{P^n} 0 \quad \text{for all } a > 1, t \in D; \\ [h^\alpha\text{-}D] \quad (\text{resp. } [h_1^\alpha\text{-}D]) \Leftrightarrow \psi_\alpha(1 + x/z_-^n, 1 - x/z_-'^n) * \nu_t^{z^n} \xrightarrow{P^n} 0 \quad \text{for all } t \in D. \end{cases}$$

b) Now, we presently prove the first implication in 1.66. Note that $\varphi_\alpha(u,0) = \alpha u$, and $\varphi_\alpha(u,v) = v\bar\varphi_\alpha(u/v)$ for $v \neq 0$, with $\bar\varphi_\alpha(u) = \alpha u + 1 - \alpha - u^\alpha$. We have $\bar\varphi_\alpha(u) \sim \frac{\alpha(1-\alpha)}{2}(u-1)^2$ as $u \to 1$. Hence $\bar\varphi_\alpha(u) - 4\alpha(1-\alpha)\bar\varphi_{1/2}(u) = o(|u-1|^2)$ as $u \to 1$, and so for every $\varepsilon > 0$ there exists $a > 1$ such that $|\bar\varphi_\alpha(u) - 4\alpha(1-\alpha)\bar\varphi_{1/2}(u)| \leq \varepsilon\bar\varphi_{1/2}(u)$ for $1/a < u < a$. Thus

$$\frac{1}{a} < \frac{u}{v} < a \Rightarrow |\psi_\alpha(u,v)| \leq \varepsilon\varphi_{1/2}(u,v).$$

On the other hand, it is obvious that there exists a constant K_a (depending on α) such that $|\psi_\alpha(u,v)| \leq K_a|u - v|$ if $\frac{u}{v} \leq \frac{1}{a}$ or $\frac{u}{v} \geq a$. Then $|\psi_\alpha| \leq \varepsilon\varphi_{1/2} + K_a\theta_a$, and

$$|\psi_\alpha(1 + x/z_-^n, 1 - x/z_-'^n) * \nu_t^{z^n}| \leq K_a I^n(a)_t + \varepsilon H_t^n.$$

Since $\varepsilon > 0$ is arbitrary, it is then clear from 1.67 that $[H\text{-}D]$ (resp. $[H_1\text{-}D]$) and $[L\text{-}D]$ imply $[h^\alpha\text{-}D]$ (resp. $[h_1^\alpha\text{-}D]$) (in case of $[H_1\text{-}D]$, recall that we have 1.57c, and so $P(H_t^n > F_t/8 + 1) \to 0$ as $n \uparrow \infty$).

c) Now we prove the second implication in 1.66. Let $\beta \in (0, 1/2)$, and set $\rho_\beta = \psi_\beta + \psi_{1-\beta}$. We have $\rho_\beta(u,0) = (1-2\beta)^2 u$, and $\rho_\beta(u,v) = v\bar\rho_\beta(u)$ for $v > 0$, with

$$\bar\rho_\beta(u) = (1-2\beta)^2(u+1) - u^\beta - u^{1-\beta} + 8\beta(1-\beta)u^{1/2}.$$

It is a routine computation to obtain that $\bar\rho_\beta(u) \geq 0$, and $\bar\rho_\beta(u) = 0$ if and only if $u = 1$. Moreover $\bar\rho_\beta(u)/u \to (1-2\beta)^2$ as $u \uparrow \infty$. So if $a > 1$ is fixed, there is a constant K (depending on a and β) such that $|u - 1| \leq K\bar\rho_\beta(u)$ for $u \leq 1/a$ or $u \geq a$. We can take $K \geq (1-2\beta)^{-2}$, and then we get

$$\theta_a(u,v) \leq K\rho_\beta(u,v)$$

$$0 \leq I^n(a)_t \leq K(\psi_\beta + \psi_{1-\beta})(1 + x/z_-^n, 1 - x/z_-'^n) * \nu_t^{z^n}.$$

Thus, using 1.67, we readily obtain the second implication in 1.66. □

2. Convergence of the Log-Likelihood to a Process with Independent Increments

§ 2a. Introduction, Statement of the Results

1. As said before, the most important case for statistical applications is by far when the Log-likelihood is approximately Gaussian. However, a more general situation has recently drawn much interest, especially in sequential analysis. The setting is as follows:

2.1 **Definition.** A (homogeneous) *exponential family of stochastic processes* is a filtered space $(\Omega, \mathscr{F}, \mathbf{F})$ endowed with a family $(P_\theta)_{\theta \in \Theta}$ of probability measures, indexed with an open subset Θ of $\mathbb{R}^d$, such that $P_\theta \overset{\text{loc}}{\ll} Q$ for all $\theta \in \Theta$, where Q is a reference (probability) measure, and such that the density process Z^θ of P_θ' with respect to Q has the form

$$Z_t^\theta = \exp(\theta \cdot X_t - t\varphi(\theta)), \tag{2.2}$$

where $X = (X^i)_{i \leq d}$ is an adapted càdlàg d-dimensional process with $X_0 = 0$, and $\theta \rightsquigarrow \varphi(\theta)$ is an arbitrary function on Θ. □

In this case, the process X is a sufficient statistic. It turns out that the structure of X is, necessarily, very particular:

2.3 **Proposition.** *Let $(\Omega, \mathscr{F}, \mathbf{F}, (P_\theta)_{\theta \in \Theta})$ be a homogeneous exponential family with X, φ as in 2.2. Then under Q and under each P_θ, the process X is a process with stationary independent increments, and φ is a "Lévy-Khintchine" function of the form VII.2.3 obtained by formally substituting t with iu.*

Proof. Any measure P_{θ_0} may serve as Q: we just have to replace Z^θ in 2.2 by

$$Z_t^{\theta_0, \theta} = \exp\{(\theta - \theta_0) \cdot X_t - t(\varphi(\theta) - \varphi(\theta_0))\}$$

and Θ by $\Theta - \theta_0$. So it is enough to prove the claims for Q.

By hypothesis, Z^θ is a positive Q-martingale. Then $E_Q(Z_{t+s}^\theta / Z_t^\theta | \mathscr{F}_t) = 1$ for s, $t \geq 0$, which gives

$$E_Q(e^{\theta \cdot (X_{t+s} - X_t)} | \mathscr{F}_t) = e^{s\varphi(\theta)} = E_Q(e^{\theta \cdot X_s}), \quad \theta \in \Theta. \tag{2.4}$$

Since the moment-generating function $\theta \rightsquigarrow E(e^{\theta \cdot Y})$, defined on an open subset Θ of $\mathbb{R}^d$, completely determines the law of the $\mathbb{R}^d$-valued random variable Y, 2.4 implies that the conditional law of $X_{t+s} - X_t$ with respect to $\mathscr{F}_t$, under Q, is the same than the a-priori law of X_s: this is just saying that X is a PIIS. Moreover, 2.4 again shows than φ is the logarithm of the Laplace transform of an infinitely divisible distribution, and so it has the form VII.2.3. □

This result explains why it has some interest to examine cases when the likelihood function is approximately the exponential of a PII.

2. The next step consists in looking at the properties of the limiting process $Z = e^X$ when X is a PII. Of course, X cannot be arbitrary here, because Z has to be a martingale, and, for simplicity, we will only consider the case where X has no fixed time of discontinuity.

2.5 **Proposition.** *Let X be the canonical process on $(\Omega, \mathscr{F}, \mathbf{F}) = (\mathbb{D}(\mathbb{R}), \mathscr{D}(\mathbb{R}), \mathbf{D}(\mathbb{R}))$. Let P be the unique probability measure on $(\Omega, \mathscr{F})$ under which X is a PII with characteristics (B, C, ν) with respect to some truncation function $h \in \mathscr{C}_t^1$, and assume that X has no fixed time of discontinuity under P* ($\Leftrightarrow \nu(\{t\} \times \mathbb{R}) = 0$ for all t).

a) *$Z = e^X$ is a local martingale on $(\Omega, \mathscr{F}, \mathbf{F}, P)$ if and only if*

2.6 $$\begin{cases} (e^x 1_{\{|x| \geq 1\}}) * \nu_t < \infty & \text{for all } t \in \mathbb{R}_+ \\ B + \frac{1}{2}C + (e^x - 1 - h(x)) * \nu = 0. \end{cases}$$

b) *Assuming 2.6, let P' be the probability measure on $(\Omega, \mathscr{F})$ under which X is a PII with characteristics (B', C', ν') given by*

2.7 $$B' = B + C + h(x)(e^x - 1) * \nu, \quad C' = C, \quad \nu'(dt, dx) = e^x \nu(dt, dx).$$

Then: (i) $P \overset{\text{loc}}{\ll} P'$ *and* $P' \overset{\text{loc}}{\ll} P$.

(ii) *The density process of P' with respect to P is $Z = e^X$* (which thus is a P-martingale, and not only a P-local martingale).

(iii) *The Hellinger process of order* $1/2$ *between P and P' is*

2.8 $$H := h(\tfrac{1}{2}; P, P') = \tfrac{1}{8}C + \tfrac{1}{2}(1 - \sqrt{e^x})^2 * \nu.$$

This result is partly analogous to Lemma 1.58.

Proof. a) The following canonical representation for X holds (see II.2.34):

$$X = X^c + h(x) * (\mu - \nu) + B + (x - h(x)) * \mu,$$

where μ is the random measure associated with the jumps of X. Ito's formula yields for $Z = e^X$:

$$Z = 1 + Z_- \cdot X^c + Z_- h(x) * (\mu - \nu) + Z_- \cdot B + Z_-(x - h(x)) * \mu + \tfrac{1}{2} Z_- \cdot C$$
$$+ Z_-(e^x - 1 - x) * \mu.$$

If Z is a local martingale, the process $Z_-(x - h(x)) * \mu + Z_-(e^x - 1 - x) * \mu$ must have locally integrable variation (see I.4.23), hence the increasing process $Z_- |e^x - 1 - h(x)| * \mu$ is locally integrable, and so is $Z_- |e^x - 1 - h(x)| * \nu$, and in particular the first property in 2.6 holds.

Moreover, Z has then the following representation:

2.9 $$\begin{aligned} Z = 1 &+ Z_- \cdot X^c + Z_-(e^x - 1) * (\mu - \nu) + Z_- \cdot B + \tfrac{1}{2} Z_- \cdot C \\ &+ Z_-(e^x - 1 - h(x)) * \nu \end{aligned}$$

and thus the sum of the last three terms in 2.9 should vanish: so the second property in 2.6 also holds.

Conversely if 2.6 holds we have 2.9, which becomes

2.10 $$Z = 1 + Z_- \cdot X^c + Z_-(e^x - 1) * (\mu - \nu).$$

Hence Z is a local martingale.

b) We assume now 2.6, and so in particular B' in 2.7 is a well-defined continuous function with finite variation on finite intervals, and $|x|^2 \wedge 1 * \nu'_t < \infty$: hence the measure P' exists.

In order to obtain that $P' \overset{\text{loc}}{\ll} P$ it suffices to apply Theorem IV.4.32 and Remark IV.4.37: in this theorem, (i) and (iii) and (vi) are trivially met, and (ii) holds with $Y = e^x$, and (iv) holds because of 2.6, and (v) is met by 2.7 with $A = C$, $\beta = c^{11} = 1$, and (vii) is met because

$$C_t + (e^{x/2} - 1)^2 * \nu_t < \infty \quad \text{for all } t < \infty,$$

which again is true because of 2.6.

Hence we have $P' \overset{\text{loc}}{\ll} P$. Furthermore, 2.10 yields $Z = \mathscr{E}(N)$, where $N = X^c + (e^x - 1) * (\mu - \nu)$: so by Theorem III.5.11, Z is the density process of P' with respect to P. In particular, since $Z > 0$ by construction, $P \overset{\text{loc}}{\ll} P'$ follows.

Finally, in order to obtain (iii) we apply Theorem IV.4.24: a simple computation shows that $H = h(1/2; P, P')$ is given by 2.8 (take $\lambda = \nu$, $U = 1$, $U' = e^x$, $\Sigma = \Omega \times \mathbb{R}_+$, $\tilde{B} = A = C$, $\tilde{\beta} = c = 1$, so $\tau = \infty$ and $\tau' = \infty$). □

2.11 **Remark.** If $Q = \mathscr{L}(Z|P)$ and $Q' = \mathscr{L}(Z|P')$, then H is also the Hellinger process $h(1/2; Q, Q')$. □

3. Now we can state our convergence result. The setting is the same than in Section 1. Recall that H^n, $I^n(a)$, $G^n(g)$ are defined in 1.6, 1.7, 1.8, and admit versions given by 1.10.

2.12 **Theorem.** *Let X, X' be two PII with characteristics (B, C, ν) and (B', C', ν') satisfying $\nu(\{t\} \times \mathbb{R}) = 0$ for all t, and 2.6 and 2.7. Define H by 2.8, and let D be a subset of $\mathbb{R}_+$.*

a) *Assume that*

2.13D $$\sup_{s \le t} \Delta G^n(1_{\{|x|>\varepsilon\}})_s \xrightarrow{P^n} 0 \quad \text{for all } t \in D,\ \varepsilon > 0;$$

[H-D] $$H^n_t \xrightarrow{P^n} H_t \quad \text{for all } t \in D;$$

[δ-D] $$G^n(g)_t \xrightarrow{P^n} g * \nu_t \quad \text{for all } t \in D \text{ and all continuous bounded functions } g \text{ on } \mathbb{R} \text{ vanishing around } 0;$$

[L_∞-D] $$\lim_{a \uparrow \infty} \limsup_n P^n(I^n(a)_t > \eta) = 0 \quad \text{for all } t \in D,\ \eta > 0.$$

Then $(P_t^n) \lhd (P_t'^n)$ *and* $(P_t'^n) \lhd (P_t^n)$ *for all* t *such that* $[t, \infty) \cap D \neq \emptyset$, *and we have the following finite-dimensional convergences*:

2.14
$$\begin{cases} Z^n \xrightarrow{\mathscr{L}(D|P^n)} Z = e^X \\ Z^n \xrightarrow{\mathscr{L}(D|P'^n)} Z' = e^{X'}. \end{cases}$$

b) *If* D *is dense in* $\mathbb{R}_+$, $[H\text{-}D] + [\delta\text{-}D] + [L_\infty\text{-}D]$ *imply* $(P_t^n) \lhd (P_t'^n)$ *and* $(P_t'^n) \lhd (P_t^n)$ *for all* $t \in \mathbb{R}_+$, *and the following functional convergences*:

2.15
$$\begin{cases} Z^n \xrightarrow{\mathscr{L}(P^n)} Z = e^X \\ Z^n \xrightarrow{\mathscr{L}(P'^n)} Z' = e^{X'}. \end{cases}$$

Observe that the condition $[H\text{-}D]$ above is the same than in 1.12, when $\nu = 0$. In fact, 1.16 and the implication (ii) $\to$ (i) of 1.12 are particular cases of this theorem, but here we have no necessary condition for having 2.15.

§ 2b. The Proof of Theorem 2.12

Our first remark consists in observing that all computations of § 1b are valid here, and we use the same set of notation (as $Y^n, V^n, \alpha^n, \ldots$). Moreover, in 1.20 we can always choose f_n to be increasing and such that $f_n(x) \geq x$ for all x. Then, 1.32 is reinforced as such:

2.16
$$\text{On the set } \Lambda^n, \begin{cases} \exp V^n \leq \dfrac{1}{n} & \text{if } \alpha^n \leq -1 + 1/n \\ \exp V^n = 1 + \alpha^n & \text{if } -1 + 1/n \leq \alpha^n \leq n - 1 \\ \exp V^n \leq 1 + \alpha^n & \text{if } \alpha^n \geq n - 1. \end{cases}$$

2.17 **Lemma.** *Under* $[H\text{-}D]$ *and* $[L_\infty\text{-}D]$ *we have* $(P_t^n) \lhd (P_t'^n)$ *for all* t *such that* $[t, \infty) \cap D \neq \emptyset$, *and* 1.46 *holds* (i.e. $P^n(T^n \leq t) \to 0$ for all $t \in D$).

Proof. 1.41 and $[L_\infty\text{-}D]$ imply

$$\lim_{a \uparrow \infty} \limsup_n P^n\left(\check{\imath}'^n\left(\frac{1}{a}\right)_t > \eta \right) = 0 \quad \text{for all } \eta > 0,$$

and $[H\text{-}D]$ implies that the sequence $(H_t^n | P^n)$ is $\mathbb{R}$-tight, for $t \in D$. Then one concludes as in Lemma 1.45. □

2.18 **Lemma.** *Under* $[L_\infty\text{-}D]$ *we have* $\mathrm{Var}(\overline{g * \nu^n})_t \xrightarrow{P^n} 0$ *for all* $t \in D$ (where $\overline{g * \nu^n}$ is defined in 1.31) *for every function* g *such that* $\theta := \sup_x \dfrac{|g(x)|}{1 \vee e^x} < \infty$.

Proof. Due to 2.16 and to the definition of $\overline{g * \nu^n}$, we have

$$\begin{aligned}
&\mathrm{Var}(\overline{g * \nu^n})_t \\
&\quad \leq 2\theta\left(1 + \frac{x}{z_-^n}\right) 1_{\{\alpha^n \leq -1+1/n\}} 1_{A^n} * \nu^{z^n} + 2\theta\left(1 + \frac{x}{z_-^n}\right)(1 + \alpha^n) 1_{\{\alpha^n \geq n-1\}} 1_{A^n} * \nu^{z^n} \\
&\quad \leq 2\theta\left\{\left(1 + \frac{x}{z_-^n}\right) 1_{\{\alpha^n \leq -1+1/n\}} * \nu^{z^n} + \left(1 - \frac{x}{z_-'^n}\right) 1_{\{\alpha^n \geq n-1\}} * \nu^{z^n}\right\},
\end{aligned}$$

which equals $2\theta\left[\imath^n\left(\frac{1}{n}\right) + \imath'^n\left(\frac{1}{n}\right)\right]$ by 1.43 and 1.44. Then 1.41 yields $\mathrm{Var}(\overline{g * \nu^n})$ $\leq 4\theta \dfrac{a}{a-1} I^n(a)$ for $n \geq a$ and thus the claim follows from $[L_\infty\text{-}D]$. □

2.19 **Corollary.** *Under* $[L_\infty\text{-}D] + [H\text{-}D]$ *and 2.13D we have*

2.20D $$\sup_{s \leq t} \nu^n(\{s\} \times \{|x| > \varepsilon\}) \xrightarrow{P^n} 0 \quad \text{for all } t \in D, \varepsilon > 0.$$

Proof. Let $g(x) = (2|x|/\varepsilon - 1)^+ \wedge 2$. By 1.31,

$$\sup_{s \leq t} \nu^n(\{s\} \times \{|x| > \varepsilon\}) \leq 1_{\{T^n \leq t\}} + \sup_{s \leq t} \Delta G^n(1_{\{|x| > \varepsilon\}})_s + |\overline{g * \nu_t^n}|.$$

Then 2.20D follows from putting together 1.46 (valid by 2.17), 2.18 and 2.13D. □

2.21 **Corollary.** *Under* $[L_\infty\text{-}D] + [H\text{-}D] + [\delta\text{-}D]$ *we have* $g * \nu_t^n \xrightarrow{P^n} g * \nu_t$ *for all* $t \in D$ *and all continuous functions* g *such that*

2.22 $$|g(x)| \leq \theta(1 \vee e^x), \quad |x| \leq 1 \Rightarrow |g(x)| \leq \theta |x|^3$$

for some constant θ.

Proof. Let $0 < \varepsilon \leq 1/2$, $a > 1$, and let f_ε, f_a' be two continuous functions with

$$0 \leq f_\varepsilon, f_a' \leq 1; \quad f_\varepsilon(x) = \begin{cases} 0 & \text{if } |x| \leq \frac{\varepsilon}{2} \\ 1 & \text{if } |x| \geq \varepsilon \end{cases} \quad f_a'(x) = \begin{cases} 0 & \text{if } x \leq a \\ 1 & \text{if } x \geq 2a, \end{cases}$$

and put $g_a' = g f_a'$, $g_\varepsilon = (1 - f_\varepsilon) g$, $g_{a\varepsilon} = (1 - f_a') f_\varepsilon g$, so that

$$g = g_a' + g_\varepsilon + g_{a\varepsilon}.$$

a) $g_{a\varepsilon}$ is continuous, bounded, null on a neighbourhood of 0. 1.31 yields

$$g_{a\varepsilon} * \nu_t^n = G^n(g_{a\varepsilon})_{t \wedge T^n} + \overline{g_{a\varepsilon} * \nu_t^n}.$$

Thus $[\delta\text{-}D]$ and 2.18 and 1.46 imply

2.23 $$g_{a\varepsilon} * \nu_t^n \xrightarrow{P^n} g_{a\varepsilon} * \nu_t \quad \text{for all } t \in D, a > 1, \varepsilon \in (0, \tfrac{1}{2}].$$

b) We have $|g'_a(x)| \leq \theta e^x 1_{\{x>a\}}$, so by 1.10

$$|G^n(g'_a)_{t\wedge T^n}| \leq \theta\left(1+\frac{x}{z^n_-}\right)\frac{1-x/z'^n_-}{1+x/z^n_-}1_{\{(1-x/z'^n_-)/(1+x/z^n_-)>e^a\}}1_{A^n} * \nu^{z^n}_t$$

$$= \theta\iota'^n(e^{-a})_{t\wedge T^n} \leq \theta\frac{e^a}{e^a-1}I^n(e^a)_{t\wedge T^n}$$

(use 1.43 and 1.41). Then we deduce from $[L_\infty\text{-}D]$ that

$$\lim_{a\uparrow\infty}\limsup_n P^n(|G^n(g'_a)_{t\wedge T^n}| > \eta) = 0$$

for all $\eta > 0$. Moreover $\overline{g'_a * \nu^n_t} \xrightarrow{P^n} 0$ by 2.18. Hence we deduce from 1.31:

2.24 $$\lim_{a\uparrow\infty}\limsup_n P^n(|g'_a * \nu^n_t| > \eta) = 0 \quad \text{for all } t\in D,\ \eta > 0.$$

c) By construction, $|g_\varepsilon(x)| \leq \theta\varepsilon|x|^2 1_{\{|x|\leq\varepsilon\}}$, so 1.23 yields

$$|g_\varepsilon * \nu^n_t| \leq \theta\varepsilon\left(1+\frac{x}{z^n_-}\right)|V^n|^2 1_{\{|V^n|\leq\varepsilon\}}1_{A^n} * \nu^{z^n}.$$

Now, 2.16 implies that for $n\geq 2$ and on the set $\{|V^n|\leq 1/2\}\cap A^n$, we have $V^n = \mathrm{Log}(1+\alpha^n)$, so $|\alpha^n| < 0.7$ and $|V^n|^2 \leq 4|\alpha^n|^2$. Hence 1.38 and 1.10 give $|g_\varepsilon * \nu^n_t| \leq 4\theta\varepsilon\beta H^n_t$. Moreover, we have seen in the proof of 2.17 that the sequence $(H^n_t|P^n)$ is $\mathbb{R}$-tight for $t\in D$. Then we easily deduce that

2.25 $$\lim_{\varepsilon\downarrow 0}\limsup_n P^n(|g_\varepsilon * \nu^n_t| > \eta) = 0 \quad \text{for all } t\in D,\ \varepsilon > 0.$$

d) Since $|x|^2 \wedge 1 * \nu_t < \infty$ and $e^x 1_{\{x>1\}} * \nu_t < \infty$ by 2.6, we have

$$\lim_{a\uparrow\infty} g'_a * \nu_t = 0, \quad \lim_{\varepsilon\downarrow 0} g_\varepsilon * \nu_t = 0$$

by Lebesgue convergence theorem. Putting this together with 2.23, 2.24 and 2.25 finishes to prove that $g * \nu^n_t \xrightarrow{P^n} g * \nu_t$ for all $t\in D$. □

Proof of 2.12. (i) The first step consists in proving the first convergence in 2.14 (resp. 2.15), and for this it is enough to show that $Y^n \xrightarrow{\mathscr{L}(D|P^n)} X$ (resp. $Y^n \xrightarrow{\mathscr{L}(P^n)} X$), because $Z = e^X$, and $Z^n = e^{Y^n}$ on $[\![0, T^n[\![$, and 1.46 holds.

To this end, we use Theorem VIII.2.4a (resp. VIII.2.18). 2.20-D is exactly VIII.2.5, and 2.21 implies

$[\delta_{5,1}\text{-}D]$ $\quad g * \nu^n_t \xrightarrow{P^n} g * \nu_t \quad$ for all $t\in D$, $g\in C_1(\mathbb{R})$.

So it remains to prove

$[\beta_5\text{-}D]$ $\quad B^n_t \xrightarrow{P^n} B_t \quad$ for all $t\in D$

(resp. $[\mathrm{Sup}\text{-}\beta_5]$ $\sup_{s\leq t}|B^n_s - B_s| \xrightarrow{P^n} 0$ for all $t\geq 0$), and

$[\gamma_5\text{-}D]$ $\quad \tilde{C}^n_t \xrightarrow{P^n} \tilde{C}_t \quad$ for all $t\in D$,

where $\tilde{C}_t = C_t + h^2 * \nu_t$ (of course, we choose a continuous truncation function h).

(ii) Let us define

$$A^n = \left(1 + \frac{x}{z_-^n}\right) 1_{\{\alpha^n = -1\}} 1_{A^n} * \nu^{z^n} = \frac{2}{z_-^n} 1_{\{x = z_-'^n\}} 1_{A^n} * \nu^{z^n}$$

$$A'^n = \left(1 - \frac{x}{z_-'^n}\right) 1_{\{\alpha^n = \infty\}} 1_{A^n} * \nu^{z^n} = \frac{2}{z_-'^n} 1_{\{x = -z_-^n\}} 1_{A^n} * \nu^{z^n},$$

which are two increasing processes, with $A^n + A'^n \leq I^n(a)$ for every $a > 1$. Set also

$$f(x) = e^x - 2(1 - \sqrt{e^x})^2$$

$$g(x) = 2(1 - \sqrt{e^x})^2 + h(x) - e^x + 1.$$

A (tedious) calculation, using 1.23 and 1.10 and 1.31, shows that

2.26 $$B^n = -4(H^n)^{T^n} + g * \nu^n + \overline{f * \nu^n} + 2A^n - A'^n.$$

In view of $[L_\infty\text{-}D]$ and of $A^n + A'^n \leq I^n(a)$, we have

2.27 $$A_t^n \xrightarrow{P^n} 0, \quad A_t'^n \xrightarrow{P^n} 0 \quad \text{for all } t \in D.$$

The function f meets $\sup_x |f(x)|/(1 \vee e^x) < \infty$, so 2.18 implies

2.28 $$\mathrm{Var}(\overline{f * \nu^n})_t \xrightarrow{P^n} 0 \quad \text{for all } t \in D.$$

g meets 2.22, as well as g^+ and g^-. Hence 2.21 and $[H\text{-}D]$ and 1.46 imply for $t \in D$:

2.29 $$\begin{cases} (H^n)_t^{T^n} \xrightarrow{P^n} H_t, & g * \nu_t \xrightarrow{P^n} g * \nu_t \\ g^+ * \nu_t^n \xrightarrow{P^n} g^+ * \nu_t, & g^- * \nu_t^n \xrightarrow{P^n} g^- * \nu_t. \end{cases}$$

Since $B = -4H + g * \nu$ by 2.6 and 2.8, we deduce from 2.27, 2.28, 2.29 and from the explicit form 2.26 that $[\beta_5\text{-}D]$ holds. Moreover if D is dense the first, third and fourth convergences in 2.29 are uniform over every finite time interval (because the left-hand side processes are increasing and the right-hand sides are increasing and continuous), hence so is the second convergence (by difference between the third and the fourth ones). This, plus 2.27 and 2.28, clearly imply $[\text{Sup-}\beta_5]$.

(iii) Set $f'(x) = 4(1 - \sqrt{e^x})^2$ and $g'(x) = h^2(x) - 4(1 - \sqrt{e^x})^2$. A calculation using 1.23 and 1.10 and 1.31 again shows

2.30 $$\tilde{C}^n = 8(H^n)^{T^n} + g' * \nu^n + \overline{f' * \nu^n} - 4A^n - 4A'^n - \sum_{s \leq \cdot} (\Delta B_s^n)^2.$$

We have seen before 1.39 that $|\Delta B_s^n| \leq \varepsilon + \beta \nu^n(\{s\} \times \{|x| > \varepsilon\})$, hence 2.20-$D$ yields $\sup_{s \leq t} |\Delta B_s^n| \xrightarrow{P^n} 0$ for all $t \in D$. Moreover 2.26 gives

$$\mathrm{Var}(B^n)_t \leq 4H_t^n + |g| * \nu_t^n + \mathrm{Var}(\overline{f * \nu^n})_t + 2A_t^n + A_t'^n$$

and 2.27, 2.28, 2.29 imply that the sequence $(\mathrm{Var}(B^n)_t | P^n)$ is $\mathbb{R}$-tight for all $t \in D$. Since

$$\sum_{s \leq t} |\Delta B_s^n|^2 \leq \mathrm{Var}(B^n)_t \sup_{s \leq t} |\Delta B_s^n|$$

we deduce that $\sum_{s\le t}|\Delta B_s^n|^2 \xrightarrow{P^n} 0$ for $t\in D$. Moreover, g' and f' meet the requirements of 2.21 and 2.18, respectively. Then 2.18, 2.21, 2.29 and 2.27 allow to deduce from the explicit form 2.30 that

$$\tilde{C}_t^n \xrightarrow{P^n} 8H_t + g' * \nu_t \quad \text{for all } t\in D.$$

Since $\tilde{C} = 8H + g' * \nu$ by 2.8, we have $[\gamma_5\text{-}D]$.

(iv) At this point, we have proved the first convergences in 2.14 and 2.15. We can of course assume that X is the canonical process, with the measure P, as in 2.5: then 2.5(b, ii) implies that $E_P(Z_t) = 1$ for all t, and thus V.1.12 yields $(P_t'^n) \lhd (P_t^n)$ for all $t\in D$.

Finally, let $t\in D$, and in case (a) (resp. (b)) let f be a continuous bounded function on $\mathbb{R}^d$ (resp. on $\Omega = \mathbb{D}(\mathbb{R})$, which is $\mathscr{D}(\mathbb{R})_t$-measurable), and $U^n = f(Z_{t_1}^n,\dots,Z_{t_d}^n)$ where $t_i\in D$, $t_i\le t$ (resp. $U^n = f(Z^n)$), and $U = f(\exp X_{t_1},\dots,\exp X_{t_d})$ (resp. $U = f(e^X)$). Then $\mathscr{L}[(Z_t^n, U^n)|P_t^n] \to \bar\eta$ where $\bar\eta$ is the measure characterized by $\bar\eta(g) = E_P[g(e^{X_t}, U)]$ (this follows from the first convergence in 2.14, resp. in 2.15). Thus V.1.13 implies that $\mathscr{L}[(Z_t^n, U^n)|P_t'^n] \to \bar\eta'$, with $\bar\eta'(g) = E_P(e^{X_t}g(e^{X_t}, U))$. So, if P' is the measure described in 2.5, i.e. the law of the PII X', we deduce that

$$E_{P'^n}(U^n) \to E_{P'}(X).$$

This is exactly saying that $Z^n \xrightarrow{\mathscr{L}(D,P'^n)} e^{X'}$ (resp. $Z^n \xrightarrow{\mathscr{L}(P'^n)} e^{X'}$), and the proof is complete. □

§ 2c. Example: Point Processes

To illustrate the previous theorem, we consider the case when P^n and P'^n are laws of point processes. More precisely, let N^n be a point process on Ω^n, and assume that $\mathbf{F}^n$ is the filtration generated by N^n. We denote by A^n and A'^n the compensators of N^n under P^n and P'^n. For simplicity, we also assume the following:

2.31 A^n and A'^n are continuous (this amounts to saying that N^n is quasi-left continuous under P^n and P'^n). □

Let $\bar{A}^n$ be any continuous adapted process which strongly majorizes A^n and A'^n (e.g. $\bar{A}^n = A^n + A'^n$), and let y^n, y'^n be two nonnegative predictable processes such that

2.32 $$A^n = y^n\cdot \bar{A}^n,\quad A'^n = y'^n\cdot \bar{A}^n$$

up to a Q^n-null set.

In the limit, we want Z to be the relative density of two Poisson processes which are locally equivalent. In view of III.5.45, this amounts to assuming:

2.33 Let N be a Poisson process on some basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ with (continuous deterministic) compensator A. Let y be a Borel function: $\mathbb{R}_+ \to (0, \infty)$ such that $A'_t = \int_0^t y_s \, dA_s < \infty$ for all $t \in \mathbb{R}_+$, and

$$Z_t = \exp\left(A'_t - A_t + \int_0^t (\operatorname{Log} y_s) \, dN_s\right). \qquad \square$$

Z is then a martingale, and is the density process of the law of the Poisson process with compensator A' with respect to P. This corresponds to the situation of 2.5, with the characteristics (B, C, ν) of the PII X given by

2.34 $\quad C = 0, \quad g * \nu_t = (g \circ \operatorname{Log} y) \cdot A_t, \quad B_t = (h \circ \operatorname{Log} y + 1 - e^y) \cdot A_t.$

Then, under these assumptions, 2.12 reads as follows (we only give the functional convergence under P^n; the other statements are left to the reader):

2.35 **Theorem.** *Assume that*

(i) $(\sqrt{y^n} - \sqrt{y'^n})^2 \cdot \bar{A}^n_t \xrightarrow{P^n} (1 - \sqrt{y})^2 \cdot A_t$ for all $t \geq 0$;

(ii) $k(y'^n / y^n) 1_{\{y^n > 0, y'^n > 0\}} \cdot A^n_t \xrightarrow{P^n} k(y) \cdot A_t$ for all $t \geq 0$ and all continuous bounded functions k on $(0, \infty)$ which vanish on a neighbourhood of 1;

(iii) $\lim_{a \uparrow \infty} \limsup_n P^n(1_{\{y^n \geq a y'^n\}} \cdot A^n_t + 1_{\{y'^n \geq a y^n\}} \cdot A'^n_t > \eta) = 0$ for all $t \geq 0$, $\eta > 0$.

Then $Z^n \xrightarrow{\mathscr{L}(P^n)} Z$, *where Z is given by 2.33.*

Proof. By IV.4.2, we have (recall 2.31):

$$H^n = \tfrac{1}{2}(\sqrt{y^n} - \sqrt{y'^n})^2 \cdot \bar{A}^n$$
$$I^n(a) = |y^n - y'^n| \, 1_{\{y^n \geq a y'^n \text{ or } y'^n \geq a y^n\}} \cdot \bar{A}^n$$
$$G^n(g) = [g \circ \operatorname{Log}(y'^n / y^n) 1_{\{y^n > 0, y'^n > 0\}}] \cdot A^n.$$

We also have $H = \frac{1}{2}(1 - \sqrt{y})^2 \cdot A$ by 2.8 and 2.34. Then (i) $= [H\text{-}\mathbb{R}_+]$ and (ii) $= [\delta\text{-}\mathbb{R}_+]$ (with $k(x) = g \circ \operatorname{Log} x$), and (iii) is clearly equivalent to $[L_\infty\text{-}\mathbb{R}_+]$ because for $a \geq 2$,

$$\tfrac{1}{2} y^n 1_{\{y^n \geq a y'^n\}} \cdot \bar{A}^n \leq |y^n - y'^n| \, 1_{\{y^n \geq a y'^n\}} \cdot \bar{A}^n \leq y^n 1_{\{y^n \geq a y'^n\}} \cdot \bar{A}^n$$

and similar inequalities with (y'^n, y^n) instead of (y^n, y'^n). $\square$

2.36 **Remark.** Under (ii), the condition (i) is equivalent to

(i') $\quad (1 - \sqrt{y'^n / y^n})^2 \cdot A^n_t \xrightarrow{P^n} (1 - \sqrt{y})^2 \cdot A_t \quad$ for all $t \geq 0$. $\square$

3. The Statistical Invariance Principle

In this last section we focus on the following question: assuming that each space $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$ is endowed with a d-dimensional càdlàg process X^n, and that

$\mathscr{L}(X^n|P^n)$ weakly converges, is it true that $\mathscr{L}(X^n|P'^n)$ also weakly converges, when $(P_t'^n) \lhd (P_t^n)$ for all t?

In general, the answer is negative. However, we have that the sequence $\mathscr{L}(X^n|P'^n)$ is tight, although it can have many limit points. Moreover, if we strengthen the hypothesis up to the convergence of the sequence $\mathscr{L}[(X^n, Z^n)|P^n]$, then $\mathscr{L}(X^n|P'^n)$ also converges (this is another extension of LeCam's third lemma).

Finally, in § 3b we examine the case where X^n is a semimartingale and converges in law, under P^n, to a continuous Gaussian martingale and moreover meets the conditions of Chapter VIII (as $[\hat{\delta}_5]$ and $[\gamma_5\text{-}\mathbb{R}_+]$): the results are stronger, and deserve the name of "statistical" invariance principle, which says that the convergence (or, "invariance principle") is preserved under contiguous changes of measures.

§ 3a. General Results

As said before, for each n we have a d-dimensional càdlàg adapted process $X^n = (X^{n,i})_{i \le d}$ on the basis $(\Omega^n, \mathscr{F}^n, \mathbf{F}^n)$. The other notation is as in the rest of the chapter (see before Section 1).

3.1 **Theorem.** *Assume that $(P_t'^n) \lhd (P_t^n)$ for all $t \in \mathbb{R}_+$. If the sequence $\{\mathscr{L}(X^n|P^n)\}$ weakly converges, or more generally if it is tight, then the sequence $\{\mathscr{L}(X^n|P'^n)\}$ also is tight.*

Proof. In view of VI.3.21, the tightness assumption implies that for all $N \in \mathbb{N}^*$, $\delta > 0$,

3.2
$$\begin{cases} \lim_{K \uparrow \infty} \limsup_n P^n\left(\sup_{t \le N} |X_t^n| > K\right) = 0 \\ \lim_{\theta \downarrow 0} \limsup_n P^n(w_N'(X^n, \theta) > \delta) = 0 \end{cases}$$

Now, $(P_N'^n) \lhd (P_N^n)$, so we can replace P^n by P'^n in 3.2, because $\sup_{t \le N} |X_t^n|$ and $w_N'(X^n, \theta)$ are $\mathscr{F}_N^n$-measurable, exactly as 1.56 was deduced from 1.55. Another application of VI.3.21 then yields the claim. $\square$

Next comes the extension of LeCam's third lemma (this is also a direct extension of the last claim in 1.4).

3.3 **Theorem.** *Assume that $(P_t'^n) \lhd (P_t^n)$ for all $t \in \mathbb{R}_+$, and that $\mathscr{L}[(X^n, Z^n)|P^n]$ weakly converges to a probability measure P on the canonical space $(\Omega, \mathscr{F}, \mathbf{F}) = (\mathbb{D}(\mathbb{R}^{d+1}), \mathscr{D}(\mathbb{R}^{d+1}), \mathbf{D}(\mathbb{R}^{d+1}))$. Then $\mathscr{L}[(X^n, Z^n)|P'^n]$ weakly converges to a probability measure P', which satisfies $P' \overset{\text{loc}}{\ll} P$, and the density process of P' with respect to P is the last component, say Z, of the canonical process on Ω.*

In particular, it follows that $\mathscr{L}(X^n|P'^n)$ converges to a probability measure on $\mathbb{D}(\mathbb{R}^d)$, which is also locally absolutely continuous with respect to the limit of $\mathscr{L}(X^n|P^n)$.

Proof. One could reproduce the proof of 1.4. Using 3.1, there also is a slightly simpler proof: indeed, 3.1 applied to the $(d+1)$-dimensional process (X^n, Z^n) yields that $\{\mathscr{L}((X^n, Z^n)|P'^n)\}$ is tight. Let P' be a limit point, and consider a subsequence (n_k) such that

3.4 $$\mathscr{L}[(X^{n_k}, Z^{n_k})|P'^{n_k}] \to P'.$$

Let f be a bounded Skorokhod-continuous function on $\Omega = \mathbb{D}(\mathbb{R}^{d+1})$, which is $\mathscr{F}_t$-measurable for some $t \in \mathbb{R}_+$. Let $s > t$, which is not a fixed time of discontinuity for the canonical process Y on Ω under P. V.1.14 gives

3.5 $$E_{P'^{n_k}}[f(X^{n_k}, Z^{n_k})] = E_{P'^{n_k}}[f(X^{n_k}, Z^{n_k}) 1_{\{Z_s^{n_k} = \infty\}}] + E_{P^{n_k}}[f(X^{n_k}, Z^{n_k}) Z_s^{n_k}].$$

By V.1.11, $P'^{n_k}(Z_s^{n_k} = \infty) \to 0$, and the sequence $\{Z_s^{n_k} f(X^{n_k}, Z^{n_k})|P^{n_k}\}$ is $\mathbb{R}$-tight. Then, using 3.4 and the assumption $\mathscr{L}((X^n, Z^n)|P^n) \to P$, and passing to the limit in 3.5 as $k \uparrow \infty$, we get

$$E_{P'}[f(Y)] = E_P[f(Y) Z_s].$$

This clearly implies that $P' \overset{\text{loc}}{\ll} P$ and that Z is the density process of P' with respect to P. In particular, P' is unique, and so is the limit of the sequence $\mathscr{L}[(X^n, Z^n)|P'^n]$. □

If we combine this theorem with the results of the previous chapters, we obtain a number of interesting applications. As an example, we now give one of these.

Let $V = (V^i)_{i \le d+1}$ be a $(d+1)$-dimensional continuous Gaussian martingale on some space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbf{F}}, \bar{P})$, with $V_0 = 0$ and $C_t^{ij} = E_{\bar{P}}(V_t^i V_t^j)$. For simplicity of notation we write $U = (V^i)_{i \le d}$, $M = V^{d+1}$ and $C^0 = C^{d+1,d+1}$. Let also $B = (B^i)_{i \le d}$ be a continuous function: $\mathbb{R}_+ \to \mathbb{R}^d$ with $B_0 = 0$.

3.6 **Corollary.** *Assume that $(P_t'^n) \lhd (P_t^n)$ for all $t \in \mathbb{R}_+$. Assume also that*

$$(X^n, Z^n) \xrightarrow{\mathscr{L}(P^n)} (U + B, e^{M - C^0/2}).$$

Then, if $B'^i = B^i + C^{i,d+1}$ for $i \le d$, we have

$$(X^n, Z^n) \xrightarrow{\mathscr{L}(P'^n)} (U + B', e^{M + C^0/2}).$$

In particular, $X^n \xrightarrow{\mathscr{L}(P'^n)} U + B'$, which is a continuous PII.

Proof. By hypothesis, $\mathscr{L}[(X^n, Z^n)|P^n] \to P = \mathscr{L}(U + B, e^{M - C^0/2})$. Thus 3.3 implies that $\mathscr{L}[(X^n, Z^n)|P'^n] \to P'$, where $P' \overset{\text{loc}}{\ll} P$ and the density process of P' with respect to P is the $(d+1)^{\text{th}}$ coordinate Z of the canonical process. Denote by $X = (X^i)_{i \le d}$ the d first coordinates of the canonical process on $\Omega = \mathbb{D}(\mathbb{R}^{d+1})$, and set $W^i = X^i - B^i$ for $i \le d$, $W^{d+1} = (\operatorname{Log} Z) + C^0/2$. Then $P = \mathscr{L}(U + B, e^{M - C^0/2})$ implies that $W = (W^i)_{i \le d+1}$ is a continuous Gaussian martingale with second characteristic $C = (C^{ij})_{i,j \le d+1}$, relatively to P.

In particular, $\langle W^i, Z \rangle = (1/Z_-) \cdot C^{i,d+1}$ under P. So Girsanov's Theorem III.3.24 implies that W is a semimartingale under P', with the characteristics $(\bar{B}, C, 0)$, and $\bar{B}^i = C^{i,d+1}$, and in particular $\bar{B}^{d+1} = C^0$. Hence under P', $W - \bar{B}$ is a continuous Gaussian martingale having the same law as V under $\bar{P}$. Therefore $P' = \mathscr{L}(U + B', e^{M + C^0/2})$, and the claim is proved. □

§ 3b. Convergence to a Gaussian Martingale

Here we fix a truncation function $h \in \mathscr{C}_t^d$, and we assume the following:

3.7 **Hypothesis.** For each $n \in \mathbb{N}^*$ the process X^n is a semimartingale with respect to P^n and to P'^n, with respective characteritics (B^n, C^n, ν^n) and (B'^n, C'^n, ν'^n) and modified second characteristics $\tilde{C}^n$ and $\tilde{C}'^n$ (all relative to the same truncation function h). □

Let also M be a d-dimensional continuous Gaussian martingale defined on some space $(\Omega, \mathscr{F}, \mathbf{F}, P)$, with $M_0 = 0$ and $C_t^{ij} = E_P(M_t^i M_t^j)$. So M has the characteristics $(0, C, 0)$. Recall once more the following notation, where D is a subset of $\mathbb{R}_+$:

$[\hat{\delta}_5\text{-}D]$ $\quad 1_{\{|x| > \varepsilon\}} * \nu_t^n \xrightarrow{P^n} 0 \quad$ for all $t \in D$, $\varepsilon > 0$

$[\gamma_5\text{-}D]$ $\quad \tilde{C}_t^n \xrightarrow{P^n} C_t \quad$ for all $t \in D$.

We have proved in VIII.2.19 that

3.8 (i) Under $[\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ we have $X^n - B^n \xrightarrow{\mathscr{L}(D|P^n)} M$ (finite-dimensional convergence along D);

(ii) If moreover D is dense in $\mathbb{R}_+$, then $X^n - B^n \xrightarrow{\mathscr{L}(P^n)} M$. (Apply VIII.2.19 with $X = M$, so $B = 0$, $\nu = 0$, and $[\hat{\delta}_5\text{-}D] \Leftrightarrow [\delta_{5,1}\text{-}D]$). □

3.9 **Theorem.** *Assume that $(P_t'^n) \lhd (P_t^n)$ for all $t \in \mathbb{R}_+$, and* 3.7.

a) *If $[\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ holds, we have $X^n - B'^n \xrightarrow{\mathscr{L}(D|P'^n)} M$.*

b) *If $[\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ holds and if D is dense in $\mathbb{R}_+$, then $X^n - B'^n \xrightarrow{\mathscr{L}(P'^n)} M$.*

3.10 **Remark.** Recall also the following notation

$[\beta_5\text{-}D]$ $\quad B_t^n \xrightarrow{P^n} 0 \quad$ for $t \in D$

$[\text{Sup-}\beta_5]$ $\quad \sup_{s \le t} |B_s^n| \xrightarrow{P^n} 0$ for $t \ge 0$

(corresponding to the limiting process M with first characteristic $B = 0$). Then if $[\beta_5\text{-}D] + [\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ (resp. $[\text{Sup-}\beta_5] + [\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ and D is dense) holds, we have $X^n \xrightarrow{\mathscr{L}(D|P^n)} M$ (resp. $X^n \xrightarrow{\mathscr{L}(P^n)} M$).

However, even if $(P_t'^n) \lhd (P_t^n)$ for all $t \geq 0$, the above conditions do not necessarily imply that $X^n \xrightarrow{\mathscr{L}(D|P'^n)} M$ (resp. $X^n \xrightarrow{\mathscr{L}(P'^n)} M$). We now present a counter-example to these properties. □

3.11 **Counter-example.** Suppose that under P^n (resp. P'^n) the 1-dimensional random variables $(\chi_k^n)_{k\geq 1}$ are i.i.d., with a normal distribution with mean 0 (resp. $(-1)^n/n$) and variance $1/n$. Let

$$X_t^n = \sum_{1\leq k\leq [nt]} \chi_k^n$$

and suppose that $\mathbf{F}^n$ is the filtration generated by X^n.

The Hellinger integral of order α between the laws $\rho = \mathscr{N}\left(0, \frac{1}{n}\right)$ and $\rho'_\pm = \mathscr{N}\left(\pm\frac{1}{n}, \frac{1}{n}\right)$ is

$$\begin{aligned} H(\alpha; \rho, \rho'_\pm) &= \frac{\sqrt{n}}{\sqrt{2\pi}} \int \exp\left[-\frac{nx^2}{2} - \frac{n(1-\alpha)}{2}\left[\left(x \pm \frac{1}{n}\right)^2 - x^2\right]\right] dx \\ &= \exp - \frac{\alpha(1-\alpha)}{2n}. \end{aligned}$$

Hence, in virtue of IV.1.73 we have

$$H(\alpha; P_t^n, P_t'^n) = \exp - \frac{[nt]}{n}\frac{\alpha(1-\alpha)}{2}$$

and thus $\lim_{\alpha\downarrow 0} \lim_n H(\alpha; P_t^n, P_t'^n) = 1$, and V.1.6 yields $(P_t'^n) \lhd (P_t^n)$.

Now, under P^n, X^n is a PII which has the same law as $\frac{1}{\sqrt{n}} \sum_{1\leq k\leq [nt]} \xi_k$, where the ξ_k's are i.i.d. with law $\mathscr{N}(0, 1)$. Hence by Donsker's Theorem (see VII.3.11), $X^n \xrightarrow{\mathscr{L}(P^n)} W$, where W is a standard Wiener process. In view of the necessary conditions in VII.3.4, $[\text{Sup-}\beta_5] + [\gamma_5\text{-}D] + [\hat{\delta}_5\text{-}D]$ hold with $D = \mathbb{R}_+$ and $C_t = t$ (a direct verification is also easy!) Moreover, if the truncation function h is odd, II.3.20 (which gives the characteristics of the PII X^n under P'^n) yields:

$$\begin{aligned} B_t'^n &= [nt]\frac{\sqrt{n}}{\sqrt{2\pi}} \int h(x)\left[\exp -\frac{n}{2}(x - (-1)^n/n)^2\right] dx \\ &= (-1)^n a_n \frac{[nt]}{n} \end{aligned}$$

where a_n is a sequence which clearly converges to 1 as $n \uparrow \infty$.

Then, 3.9 yields that $X^n - B'^n \xrightarrow{\mathscr{L}(P'^n)} W$ (again, a direct verification is immediate), but the sequence (B'^n) has two limit points $B'_t = t$ and $B''_t = -t$. Hence the sequence $\mathscr{L}(X^n|P'^n)$ does not converge, but it is tight and also admits two limit points, namely the laws of $W_t + t$ and of $W_t - t$. □

Proof of 3.9. In virtue of VIII.2.19, it suffices to prove:

$[\hat{\delta}_5\text{-}D]'$ $\quad 1_{\{|x|>\varepsilon\}} * \nu'^n_t \xrightarrow{P'^n} 0$ $\quad$ for all $t \in D$, $\varepsilon > 0$;

$[\gamma_5\text{-}D]'$ $\quad \tilde{C}'^n_t \xrightarrow{P'^n} C_t$ $\quad$ for all $t \in D$.

(i) We have seen in VIII.3.5 that $[\hat{\delta}_5\text{-}D]$ is equivalent to $\sup_{s\leq t}|\Delta X^n_s| \xrightarrow{P^n} 0$ for $t \in D$. Since $(P'^n_t) \lhd (P^n_t)$ we also have $\sup_{s\leq t}|\Delta X^n_s| \xrightarrow{P'^n} 0$ for $t \in D$, which in turn implies $[\hat{\delta}_5\text{-}D]'$.

$[\gamma_5\text{-}D]$ implies $\tilde{C}^n_t \xrightarrow{P'^n} C_t$, again because $(P'^n_t) \lhd (P^n_t)$, so it remains to prove:

3.12 $$\tilde{C}'^n_t - \tilde{C}^n_t \xrightarrow{P'^n} 0 \quad \text{for } t \in D.$$

(ii) Since $Q^n = (P^n + P'^n)/2$, X^n is a Q^n-semimartingale by III.3.40. Moreover, if C^{Q^n} denotes the second characteristic of X^n under Q^n, III.3.24 implies that $C^n_t = C^{Q^n}_t$ P^n-a. s. and $C'^n_t = C^{Q^n}_t$ P'^n-a. s. Hence if T^n is defined by 1.19, we have $C^n_t = C'^n_t$ Q^n-a. s. (and so P^n- and P'^n-a. s.) on the set $\{T^n > t\}$. Now, the same proof than in 1.45, using once more $(P'^n_t) \lhd (P^n_t)$, shows that $P'^n(T^n \leq t) \to 0$. Therefore $C^n_t - C'^n_t \xrightarrow{P'^n} 0$ for all $t \in \mathbb{R}_+$; then, in view of the explicit form II.2.18 for $\tilde{C}^n$ and $\tilde{C}'^n$, 3.12 will follow from

3.13 $$\begin{cases} A^{n,ij}_t \xrightarrow{P'^n} 0 \quad \text{for } t \in D, \text{ where} \\ A^{n,ij}_t = (h^i h^j) * \nu'^n_t - (h^i h^j) * \nu^n_t - \sum_{s\leq t} [\Delta B'^{n,i}_s \Delta B'^{n,j}_s - \Delta B^{n,i}_s \Delta B^{n,j}_s]. \end{cases}$$

Moreover, let $\lambda^n = \nu^n + \nu'^n$ and let U^n, U'^n be two predictable functions on $\Omega^n \times \mathbb{R}_+ \times \mathbb{R}^d$ such that $\nu^n = U^n \cdot \lambda^n$ and $\nu'^n = U'^n \cdot \lambda^n$ up to a Q^n-null set. For every function W on $\Omega^n \times \mathbb{R}_+ \times \mathbb{R}^d$ we use the notation

$$\hat{W}_t = \int \lambda^n(\{t\} \times dx) W(t,x) \quad (= +\infty \text{ if the integral diverges}),$$

so in particular

3.14 $$\begin{cases} a^n_t := \nu^n(\{t\} \times \mathbb{R}^d) = (\widehat{U^n})_t, & a'^n_t := \nu'^n(\{t\} \times \mathbb{R}^d) = (\widehat{U'^n})_t \\ \Delta B^{n,i} = \widehat{h^i U^n}, & \Delta B'^{n,i} = \widehat{h^i U'^n} \end{cases}$$

(see II.2.14). Hence 3.13 gives, for i, j fixed:

3.15 $$\begin{cases} A^{n,ij} = V^n + W^n - F^n, \quad \text{where} \\ V^n = (h^i - \widehat{h^i U^n})(h^j - \widehat{h^j U^n})(U'^n - U^n) * \lambda^n \\ W^n_t = \sum_{s\leq t} (a^n_s - a'^n_s)(\widehat{h^i U^n})_s (\widehat{h^j U^n})_s \\ F^n_t = \sum_{s\leq t} (\widehat{h^i(U'^n - U^n)})_s (\widehat{h^j(U'^n - U^n)})_s \end{cases}$$

(iii) We also set

3.16 $$K_t^n = (\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_t^n + \sum_{s \le t} (\sqrt{1 - a_s^n} - \sqrt{1 - a_s'^n})^2.$$

Now, we apply IV.3.39: the Hellinger process $H^n = h(\frac{1}{2}; P^n, P'^n)$ strongly majorizes the process $h^0(\frac{1}{2})$ given by IV.3.12, and we deduce that $K^n \le 2H^n$. Since $(P_t'^n) \lhd (P_t^n)$, we deduce from V.2.3 that

3.17 the sequence $(K_t^n | P'^n)$ is $\mathbb{R}$-tight for all $t \in \mathbb{R}_+$.

For $r = 1, 2$ let us also define

$$m_t^{n,r} = \sup_{s \le t} \widehat{|h|^r(U^n + U'^n)}_s.$$

Since $h(x) = x$ for $|x|$ small and $\theta := \sup|h(x)| < \infty$, we have for $\varepsilon > 0$ small enough

$$m_t^{n,r} \le 2\varepsilon^r + \theta^r \sup_{s \le t} \widehat{[(U^n + U'^n)1_{\{|x|>\varepsilon\}}]}_s$$

3.18 $$\le 2\varepsilon^r + \theta^r (U^n + U'^n) 1_{\{|x|>\varepsilon\}} * \lambda_t^n = 2\varepsilon^r + \theta^r 1_{\{|x|>\varepsilon\}} * (\nu^n + \nu'^n)_t.$$

$[\hat{\delta}_5\text{-}D]$ and $(P_t'^n) \lhd (P_t^n)$ imply that $1_{\{|x|>\varepsilon\}} * \nu_t^n \xrightarrow{P'^n} 0$ for $t \in D$. This, plus $[\hat{\delta}_5\text{-}D]'$, shows that the last term in 3.18 goes to 0 in P'^n-measure for all $\varepsilon > 0$ as $n \uparrow \infty$. Therefore, $\varepsilon > 0$ being arbitrary,

3.19 $$m_t^{n,r} \xrightarrow{P'^n} 0 \quad \text{for all } t \in D, r = 1, 2.$$

(iv) Let us first consider the processes F^n in 3.15. We have

$$F_t^n = \sum_{s \le t} \widehat{[h^i(\sqrt{U^n} + \sqrt{U'^n})(\sqrt{U^n} - \sqrt{U'^n})]}_s \times \widehat{[h^j(\sqrt{U'^n} + \sqrt{U^n})(\sqrt{U^n} - \sqrt{U'^n})]}_s.$$

Hence Hölder's inequality yields

$$\begin{aligned} |F_t^n| &\le \sum_{s \le t} \{\widehat{|h^i|^2(\sqrt{U^n} + \sqrt{U'^n})^2}\}_s^{1/2} \{\widehat{|h^j|^2(\sqrt{U^n} + \sqrt{U'^n})^2}\}_s^{1/2} \\ &\quad \times \{\widehat{(\sqrt{U^n} - \sqrt{U'^n})^2}\}_s \\ &\le 2 \sum_{s \le t} \{\widehat{|h|^2(U^n + U'^n)}\}_s \{\widehat{(\sqrt{U^n} - \sqrt{U'^n})^2}\}_s \\ &\le 2 m_t^{n,2} \sum_{s \le t} \{\widehat{(\sqrt{U^n} - \sqrt{U'^n})^2}\}_s \\ &\le 2 m_t^{n,2} (\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_t^n \le 2 m_t^{n,2} K_t^n. \end{aligned}$$

Hence 3.17 and 3.19 imply

3.20 $$F_t^n \xrightarrow{P'^n} 0 \quad \text{for all } t \in D.$$

(v) Next, we examine the processes V^n in 3.15. We have

$$V^n = (h^i - \widehat{h^i U^n})(h^j - \widehat{h^j U^n})(\sqrt{U^n} + \sqrt{U'^n})(\sqrt{U^n} - \sqrt{U'^n}) * \lambda^n.$$

Then Hölder's inequality gives

$$|V_t^n| \le \{(h^i - \widehat{h^i U^n})^2 (h^j - \widehat{h^j U^n})^2 (\sqrt{U^n} + \sqrt{U'^n})^2 * \lambda_t^n\}^{1/2}$$
$$\times \{(\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda_t^n\}^{1/2}$$

3.21 $$\le 2(V_t'^n)^{1/2}(K_t^n)^{1/2},$$

where

$$V'^n = (h^i - \widehat{h^i U^n})^2 (h^j - \widehat{h^j U^n})^2 (U^n + U'^n) * \lambda^n$$
$$\le \varepsilon^2 (h^j - \widehat{h^j U^n})^2 (U^n + U'^n) * \lambda^n + 4\theta^2 1_{\{|h^i - \widehat{h^i U^n}| > \varepsilon\}} (U^n + U'^n) * \lambda^n$$

(recall that $|h| \le \theta$, so $|h^j - \widehat{h^j U^n}| \le 2\theta$). Since

3.22 $$|x + y| \le 3x + 2(\sqrt{x} - \sqrt{y})^2 \quad \text{for } x, y \ge 0,$$

it follows that

3.23 $$V'^n \le \varepsilon^2 \{3(h^j - \widehat{h^j U^n})^2 U^n * \lambda^n + 8\theta^2 (\sqrt{U^n} - \sqrt{U'^n})^2 * \lambda^n\}$$
$$+ 4\theta^2 1_{\{|h^i - \widehat{h^i U^n}| > \varepsilon\}} * (v^n + v'^n).$$

Using again the definition II.2.18 of $\tilde{C}^n$, and 3.14, we see that

3.24 $$\tilde{C}_t^{n,jj} = C_t^{n,jj} + (h^j - \widehat{h^j U^n})^2 U^n * \lambda_t^n + \sum_{s \le t} (1 - a_s^n)(\widehat{h^j U^n})_s^2.$$

Moreover if $\varepsilon > 0$ is small enough and if $m_t^{n,1} \le \varepsilon/2$, then $|h^i - \widehat{h^i U^n}|(\omega, s, x) > \varepsilon$ implies $|x| > \varepsilon/2$ if $s \le t$. Hence 3.23 yields

3.25 $$V_t'^n \le 3\varepsilon^2 \tilde{C}_t^{n,jj} + 8\varepsilon^2\theta^2 K_t^n + 4\theta^2 1_{\{|x| > \varepsilon/2\}} * (v^n + v'^n)_t \quad \text{on } \left\{m_t^{n,1} \le \frac{\varepsilon}{2}\right\}.$$

Now, [γ_5-D] clearly implies that the sequence $(\tilde{C}_t^{n,jj}|P^n)$ is $\mathbb{R}$-tight; since $(P_t'^n) \lhd (P_t^n)$ we deduce that $(\tilde{C}_t^{n,jj}|P'^n)$ also is $\mathbb{R}$-tight (as for the implication 1.55 $\Rightarrow$ 1.56), and we have seen in (iii) that the last term in 3.25 goes to 0 in P'^n-measure for all $\varepsilon > 0, t \in D$. Since $\varepsilon > 0$ is arbitrary in 3.25, we then deduce from 3.17 and 3.19 that

$$V_t'^n \xrightarrow{P'^n} 0 \quad \text{for all } t \in D.$$

Therefore, using again 3.17, it follows from 3.21 that

3.26 $$V_t^n \xrightarrow{P'^n} 0 \quad \text{for all } t \in D.$$

(vi) It remains to consider the processes W^n in 3.15. Since $a - a' = (\sqrt{1-a} + \sqrt{1-a'})(\sqrt{1-a} - \sqrt{1-a'})$, Hölder's inequality applied to the definition of W^n gives

$$|W_t^n| \le \left\{\sum_{s\le t} (\widehat{h^iU^n})_s^2 (\widehat{h^jU^n})_s^2 (\sqrt{1-a_s^n} + \sqrt{1-a_s'^n})^2\right\}^{1/2}$$

$$\times \left\{\sum_{s\le t} (\sqrt{1-a_s^n} - \sqrt{1-a_s'^n})^2\right\}^{1/2}$$

3.27 $$\le \sqrt{2}(W_t'^n)^{1/2}(K_t^n)^{1/2},$$

where

$$W_t'^n = \sum_{s\le t} (\widehat{h^iU^n})_s^2 (\widehat{h^jU^n})_s^2 (1 - a_s^n + 1 - a_s'^n)$$

$$\le (m_t^{n,1})^2 \sum_{s\le t} (\widehat{h^jU^n})_s^2 (1 - a_s^n + 1 - a_s'^n)$$

$$\le (m_t^{n,1})^2 \left\{3 \sum_{s\le t} (\widehat{h^jU^n})_s^2 (1 - a_s^n) + 2\theta^2 \sum_{s\le t} (\sqrt{1-a_s^n} - \sqrt{1-a_s'^n})^2\right\}$$

(we have used 3.22 and $|\widehat{h^jU^n}| \le \theta$ for the last inequality). Using 3.24 and 3.16, we obtain

$$W_t'^n \le (m_t^{n,1})^2 [3\tilde{C}_t^{n,jj} + 2\theta^2 K_t^n].$$

We have already seen that the sequences $(\tilde{C}_t^{n,jj}|P'^n)$ and $(H_t^n|P'^n)$ are $\mathbb{R}$-tight, so we deduce from 3.19 that $W_t'^n \xrightarrow{P'^n} 0$ for all $t \in D$. This, plus 3.27 and 3.17 again, yields

$$W_t^n \xrightarrow{P'^n} 0 \quad \text{for all } t \in D.$$

Finally, putting this together with 3.26 and 3.20 and 3.15, we obtain that $A_t^{n,ij} \xrightarrow{P'^n} 0$ for all $t \in D$, and 3.12 follows: hence the proof is complete. □

Bibliographical Comments

Chapter I

After the introduction of Kolmogorov's axiomatic, an important step in probability theory consisted in introducing a filtration on the probability space: this has led to a lot of additional structure and so to deeper results. It started with Doob [43] and developed, initially for Markov processes and later for more general applications, by the French school (in a wide sense!) and primarily by P.A. Meyer.

The initial exposition of the material of this chapter has changed a lot through the years: for example the so-called "accessible σ-field" has virtually disappeared, the notion of "natural" increasing process has been recognized to be the same as "predictable", etc... and of course many proofs have been considerably simplified. For a complete history of the evolution of the subject, we refer to the book [36] of Dellacherie and Meyer.

The theory of stochastic integration can now be considered as a part of the "general theory", although it began earlier with the works of Wiener [246, 247], and especially Ito [91, 92, 93]. It was developed for square-integrable martingales by Courrège [31] and Kunita and Watanabe [136], local martingales and semimartingales by Doléans-Dade and Meyer [39] (where the notion of semimartingale appears for the first time), and by Meyer [183] for the most general (reasonnable) integrals with respect to a local martingale. Here we essentially follow the same route as in Doléans-Dade and Meyer [39], except that we use "Yen's Lemma" (Proposition 4.17): see Meyer [184], Jacod and Mémin [106].

Now, there exist other ways of constructing stochastic integrals, which in a sense are shorter or more efficient, but which necessitate more, or other, prerequisites: one is based upon Bichteler-Dellacherie-Mokobodzki's Theorem (see Bichteler [10], Dellacherie [35], Kussmaul [139], Métivier and Pellaumail [181]). Another one is based upon Burkhölder-Davis-Gundy's inequalities (see Meyer [183]). A third one uses Theorem 4.56, which characterizes the jumps of a local martingale and which is due to Chou [29] and Lépingle [147] (see Jacod [98]).

We want also to especially mention two notions which are used all over this book: one is the domination property (§ 3c) introduced by Lenglart [146], while the form 3.32 is due to Rebolledo [202]. The other is the exponential of a semimartingale (§ 4f), introduced in the real-valued case by Doléans-Dade [38].

Chapter II

Random measures and the associated integrals have been introduced by Ito [93] and thoroughly studied by Skorokhod [225] in the case where the measure is Poisson, i.e. with deterministic compensator. More general integer-valued random measures with their compensators have been considered by S. Watanabe [244] under the name of "Lévy systems" for Markov processes (the characterization of the Lévy system as a compensator is due to Benveniste and Jacod [7]). Without reference to an underlying Markov process, they were introduced by Grigelionis [70] when the compensator is absolutely continuous (in time) with respect to the Lebesgue measure, and by Jacod [94] in the general case. The corresponding stochastic integrals appear in Jacod [95] (see also Grigelionis [75] for the quasi-left continuous case).

The idea of the characteristics of a semimartingale goes back to Ito [93], who studied "locally infinitely divisible processes" in the Markov case (it was also more or less implicit in various works of Kolmogorov, Lévy, Feller, or more recently in Watanabe [244] and Grigelionis [69]). This idea has been applied to general processes by Grigelionis [70, 71] (when the characteristics are "absolutely continuous" in time) and by Jacod and Mémin [106] in the general case (see Jacod [97] for the multi-dimensional version). The notion of modified second characteristic arose from limit theorems, in Liptser and Shiryaev [158] for the first time. Theorem 2.21 has been used in various disguises nearly since the introduction of the characteristics, and special cases (as Lévy's characterization of Wiener process, or Watanabe's characterization of Poisson process [244]) are much older. The content of § 2b is presented here for the first time.

The canonical representation for semimartingales (§ 2c) is taken from Jacod [95]; Theorem 2.42 is due in this form to Grigelionis and Mikulevicius [77], but for diffusion processes it goes back to Dynkin [47] and Stroock and Varadhan [232] (for diffusions with jumps, see also Komatsu [133], Makhno [169], Stroock [231]). Theorems 2.47 and 2.49 appeared in Jacod, Kłopotowski and Mémin [105] and the "Laplace transform" version for point processes in Kabanov, Liptser and Shiryaev [121], but again it extends older results: see Stroock and Varadhan [232]; it can also be viewed as an extension of the fundamental Wald's identity for sums of random numbers of independent random variables.

The content of § 3b is new, and relates to Kolmogorov's three series theorem (and its "conditional" version). The "one-point" point process in § 3c is the example described by Dellacherie [33].

Theorem 4.5 is essentially due to S. Watanabe [244] and its generalization 4.8 is taken from Jacod [94]. Let us emphazise here that there is an extensive literature on Poisson measures, or more general random measures, where the time plays no specific rôle: see e.g. a fundamental result of Kingman [127], very close to 4.8, or Kallenberg [125]. There also is a theory which we do not touch here, about stationary point processes or measures: see Neveu [189] or Kerstan, Matthes and Mecke [126].

The structure of PII's goes back essentially to Lévy and Doob, who were primarily (but not exclusively) interested in PII without fixed times of discontinuity, that is, PII which are continuous in probability. In particular, formula 4.16 in this case is due to Lévy. Doob [43] also studied general PII's (even non-càdlàg ones) and his book implicitely contains 4.16, including when the PII is not a semimartingale (Theorem 5.2). The characterization 4.14 of PII's that are semimartingales is taken from Jacod [98]. The result according to which a semimartingale is a PII if and only if its characteristics are deterministic is a "well-known" result formalized by Grigelionis [76].

The content of Section 5 is borrowed to Jacod [101], but the idea of the proof of 5.1, as presented here, comes from Lobo [167].

Conditional PII's have appeared early in the literature, at least in many particular cases (like Cox' processes), and usually in connection with applications. The general characterization theorem 6.6 is due to Grigelionis [74].

The results of § 7a are taken from Jacod [270]; it should be noted that similar results for discontinuous conditional PIIS would be very useful, but only partial results in this direction are available in Jacod [271]. Some earlier work on this subject has also been done by Traki [236], [294]. For some related (but different in spirit) problems one can also consult Grigelionis [266] and Ocone [285].

The results of § 8a have been proved by many different authors in special circumstances, but a systematic approach is provided by Kallsen and Shiryaev [276]. In § 8b multiplicative decompositions go back to Ito and Watanabe [267] and Meyer [284], and Theorem 8.19 in this form is due to Yœurp and Yor in an unpublished paper and is also given in Jacod [98]. The end of this paragraph also appears in different form in the literature, but we follow here Eberlein and Jacod [261] and Kallsen and Shiryaev [276].

Chapter III

The first examples of what is now called a martingale problem are the Lévy characterization of Wiener process (example 1.4) and the Watanabe characterization of Poisson point processes (example 1.5). The essential step in the formalization of martingale problems was taken by Stroock and Varadhan [232] for proving existence of weak solutions to stochastic differential equations related to diffusions (see § 3c). The general formulation can be found in Yor [249] and in Jacod and Yor [112].

The study of point processes via a martingale approach was initiated by Watanabe [244] and Papangelou [194], while Brémaud [20, 21] was the first to undergo a systematic study. The uniqueness theorems 1.21 and 1.26 are in Kabanov, Liptser and Shiryaev [118] and Jacod [94], as well as the explicit form of the compensator in Theorem 1.33, which indeed is an easy extension of Dellacherie's results [33] for the "one-point" point process (II.3.26).

Martingale problems associated with the characteristics of a semimartingale were introduced in Jacod [96]. The results presented in § 2c give a summary of many results scattered through the literature, and they are due to many different authors; for an historical account (and also many further results) we refer to the books [157] of Liptser and Shiryaev, [233] of Stroock and Varadhan, [61] of Gikhman and Skorokhod, [98] of Jacod. The equivalence between solution-measures and solutions to a martingale problem (Theorem 2.26) is due to Stroock and Varadhan [232] for diffusions and generalized diffusions, an the general formulation comes from El-Karoui and Lepeltier [49] and Jacod [98]. Theorem 2.32 is a "classical" result, proved in various contexts by many different authors. Theorem 2.33 is essentially due to Yamada and Watanabe [248]. Theorem 2.34 is the basic result of Stroock and Varadhan [232] (continuous case) and Stroock [231], Komatsu [133], Makhno [169]. The notion of local uniqueness (§ 2d) appeared in Jacod and Mémin [106], but it should be emphazised that this sort of method goes back a long way for Markov processes: see in particular Courrège and Priouret [32] for Lemmas 2.43 and 2.44.

Our first "Girsanov's Theorem" 3.11 is an extension due to Van Schuppen and Wong [240] of the well known result of Girsanov [64] (and also Cameron and Martin [27] for a deterministic drift) concerning the case where the local martingale is a Wiener process. The other Girsanov's Theorems 3.13 and 3.24 are due to Jacod [94] and Jacod and Mémin [106]; Lemma 3.31 is essentially due (in the case of possibly discontinuous locally square-integrable martingales) to Kunita and Watanabe [136]. Theorem 3.40 is taken from Kolomietch [132].

The construction of stochastic integrals with respect to multi-dimensional continuous local martingales is due to Galtchouk [56], who more generally considered locally square-integrable martingales. § 4b comes from Jacod [95]. All the notions and results of §§ 4c,d come from Jacod [96] and Jacod and Yor [112], following an idea of Dellacherie [34] (see also Grigelionis [76] and Kabanov [115] for 4.34, and Liptser [155] for diffusion processes). Theorem 4.37 is taken from Jacod [94].

Section 5 is a revised version of Jacod and Mémin [106] and Kabanov, Liptser and Shiryaev [120]. A version of 4.35 can be found in Skorokhod [224] in the quasi-left continuous case. Theorem 5.38 subsumes results due to many authors (Kailath, Zakai, Liptser, Shiryaev, Yershov, Wong, ...): see Liptser and Shiryaev [157] for detailed comments. A partial version of Theorem 5.45 was given by Brémaud [20].

The first to observe that multi-dimensional integrals do not amount to just adding 1–dimensional integrands has been Galtchouk in [96], in the case of square-integrable martingales. The construction of the stochastic integral for the most general predictable integrands, w.r.t. to a semimartingale is due to

Jacod in [98] for the 1–dimensional case and in [268] in the multidimensional case: most results and also the method for § 6c are taken from [268], although the reader can also consult Stricker [292]. Example 6.21 is taken from Emery [263]. Emery introduced the topology which bears his name in [262], and Proposition 6.26 is due to Mémin [281]. For non-predictable (optional) stochastic integrals, see for example Jacod [98]. There is also the so-called "Skorokhod stochastic integral" of non-adapted integrands, but the resulting process is obviously no longer a semimartingale, so this theory does not fit at all in the present framework. The characterization of integrable processes w.r.t. a symmetric stable process is in Rosiński and Woyczinski [288] and also in Kallsen and Shyriaev [276].

σ–martingales have been introduced by Chou [257] and studied by Emery [263], to whom Theorem 6.41 is due. The name itself is due to Delbaen and Schachermayer (see [258], [259]), who used this notion for a formulation of the a general version of the fundamental theorem of asset pricing. Proposition 6.35 is found in Kabanov [274] and also in Goll and Kallsen [265]. The fact that σ–martingales that are bounded from below are local martingales is due to Ansel and Stricker [255], and the identity between local martingales and martingale transforms in the discrete time case may be found in Jacod and Shiryaev [273]. Finally, all of Section 7 on cumulant processes and Esscher's transform is borrowed to Kallsen and Shiryaev [276]. Many examples of application of Esscher's transform are found in financial mathematics, see e.g. Shyriaev [290].

Chapter IV

Kakutani [124] was the first to exploit the "Kakutani-Hellinger" distance for absolute continuity-singularity problems, but the corresponding integrals were introduced in analysis by E. Hellinger; their relations with the variation metric are well known (see Chapter V). In connection with stochastic processes they have been used for example by Liese [150, 151, 152] and Newman [190]. Hellinger processes and their explicit form 1.34 were introduced by Liptser and Shiryaev [162] for studying contiguity, while the presentation of these as compensators of suitable processes appears in Mémin and Shiryaev [179] and Jacod [102]. The material in § 1d is new and could presumably lead to many other applications: for instance taking $\psi(x) = \operatorname{Log} x$ would give a "Kullback process" associated with Kullback information as Hellinger processes are with Hellinger integrals.

The problem of finding necessary and sufficient conditions for absolute continuity of measures associated with processes has a long story, starting with Kakutani. The problem was firstly solved in the "independent case", for Gaussian processes and for diffusion processes: Hájek [80], Feldman [53], Ibragimov and Rozanov [90] (for more historical details, see the bibliography of Liptser and

Shiryaev [157]). The development of stochastic calculus allowed to solve the problem in general, assuming local absolute continuity (see Jacod and Mémin [106] for the quasi-left continuous case, Kabanov, Liptser and Shiryaev [120] for the general case, and also Engelbert and Shiryaev [50] in the discrete-time setting). The general case (without local absolute continuity) comes from Jacod [102]. Another generalization, not involving semimartingales, can be found in Fernique [55]. The criteria developped in the present chapter lead to "deterministic" criteria for PII's which naturally generalize Kakutani's Theorem 2.38, and to conditions on the coefficients for diffusion processes. The proofs given here are new and seem much simpler than the previous ones. For an elementary exposition of the discrete-time case, one can also consult the text-book [222] of Shiryaev.

Now about Sections 3 and 4: the first explicit computations of Hellinger processes in terms of characteristics of semimartingales are in Liptser and Shiryaev [165] and Mémin and Shiryaev [179] for PII's (Theorem 4.32). The presentation and the general results of Section 3 are new. Theorem 4.6 appeared in Kabanov, Liptser and Shiryaev [118], and 4.16 in Kabanov, Liptser and Shiryaev [119]; apart from a different presentation, Theorem 4.23 comes from Liptser and Shiryaev [157]. § 4c is essentially borrowed from Mémin and Shiryaev [179], but the case of PII's without fixed time of discontinuity already is in Skorokhod [224]; see also Newman [190].

Chapter V

The notions of contiguity and entire separation are due to LeCam [142]. An extensive account on contiguity and its statistical applications (especially for the independent and the Markov case) may be found in Roussas [217]). Some of the basic equivalences in Lemmas 1.6 and 1.13 appear in Hall and Loynes [85], Liptser, Pukelsheim and Shiryaev [156], Eagleson and Mémin [48], Hájek and Šidak [82], Jacod [102], Greenwood and Shiryaev [68].

In the discrete-time setting, the first general contiguity result (case of independent variables) is due to Oosterhof and Van Zwet [191], and the general criterion appeared in Liptser, Pukelsheim and Shiryaev [156] and Eagleson and Mémin [48] (the latter assumes local absolute continuity). The continuous-time problem was solved by Liptser and Shiryaev [162] (with a different method) and Jacod [102] (the criteria given here, though, are slightly different).

The relations between Hellinger integrals of various order and the variation metric can be found in Kraft [134] or Matusita [172]; see also Vajda [237]. Proposition 4.16 is just an exercise on multiplicative decompositions of nonnegative supermartingales; its corollary 4.19 was given by Kabanov, Liptser and Shiryaev [123], as well as the estimates of 4.21. The discrete-time version of Theorem 4.31 is in Vostrikova [241]. Essentially all the results of § 4c are due to Kabanov, Liptser and Shiryaev [116, 117, 122], T. Brown [26], Valkeila [238], and Mémin [176]. The case of diffusion processes (§ 3d) has been investigated by Liese [151].

Chapter VI

The first examples of weak convergence are due to Kolmogorov [130], Erdös and Kac [51], Donsker [40] and Maruyama [171].

The basic facts of the chapter, weak convergence and properties of the Skorokhod (J_1) topology, originate in the works of Prokhorov [200] and Skorokhod [223], and they also appear in Billingsley [12]. In these references, the authors consider processes indexed by [0, 1], but in many instances it is more natural to consider processes indexed by $\mathbb{R}_+$. For this purpose, the Skorokhod topology was extended by Stone [230] and Lindvall [154], and here we essentially follow Lindvall's method. The metric δ' of Remark 1.27 has been described by Skorokhod [223]; Kolmogorov [131] showed that the space $\mathbb{D}$ with the associated topology is topologically complete, and the metric δ of 1.26 for which it is complete was exhibited by Prokhorov [199].

It should be emphazised that Prokhorov's Theorem 3.5 has two parts:

(1) all relatively compact sequences of measures are tight,

(2) all tight sequences are relatively compact.

For (2) we only need a metric separable space, and of course (2) is the most useful of the two statements. However, (1) requires completeness, and we also use (1) in this book (in Section 6 for example).

The results of Section 2 are essentially "well-known", and scattered through the literature. See e.g. Billingsley [12], Aldous [2], Whitt [245] , Pagès [192]. § 2b is taken from Jacod and Mémin [107].

Aldous' criterion was introduced in [1]. Theorem 4.13 is due to Rebolledo [202], and 4.18 is a modernized version of results in Liptser and Shiryaev [158] and Jacod and Mémin [107] (see also Lebedev [141]; other results belonging to the same circle of ideas can be found in Billingsley [13] and Grigelionis [73]).

Section 5 is based upon Jacod, Mémin and Métivier [111], with an amelioration due to Pagès [193] (condition C5). Section 6 has its origin in Liptser and Shiryaev [159], and the general case comes from Jacod [100].

The condition P-UT has been introduced, under the (slightly misleading) name UT, by Jakubowski, Mémin and Pagès in [113] in order to obtain a stability result for stochastic integrals (Theorem 6.22). As said in 6.2, This condition is strongly related with the Bichteler-Dellacherie-Mokobodski characterization of semimartingales. The various criteria given in § 6a can be found in various papers by Kurtz and Protter [277], [278], [279] and Mémin and Słomiński [282], and also Stricker [293]. Theorem 6.26 has its origin in Liptser and Shiryaev [159] and the general case comes from Jacod [100] (as well as Remark 6.28), except that the P-UT condition is replaced by a condition expressed in terms of the characteristics (and which turns out to be equivalent to

P-UT). Proposition 7.3 is taken from Słomiński [291], while Proposition 7.5 comes from Jacod and Protter [272].

Chapter VII

Most of the material (including the proofs) of Section 2 is borrowed from the book [65] of Gnedenko and Kolmogorov and, more indirectly, from Lévy [149]. Only the presentation differs, through two significant changes: firstly, we use a truncation function h that is continuous, instead of the usual truncation "at 1"; this allows for substantial simplifications, both in the formulation of the results and in their proofs. Secondly, in the lemmas on characteristic functions (§ 2b) the centering of the variables is around the truncated mean instead of the median. The finite-dimensional convergence (§ 2d) is of course a simple consequence of Gnedenko and Kolmogorov's results. A formulation similar to 2.52 may be found in Skorokhod [224].

The sufficient condition in 3.4 is due to Liptser and Shiryaev [158], and to Jacod, Kłopotowski and Mémin [105] in 3.13. The necessary part comes from Jacod [101].

The ideas of §§ 4a,b (convergence of non-infinitesimal triangular arrays, etc...) are taken from Jacod, Kłopotowski and Mémin [105]. § 4c is new, but the idea originates in the paper [114] of Jakubowski and Słominski.

The sufficient part of 5.2 was given by Lindeberg [153], the necessary part by Feller [54]. Theorem 5.18 is firstly due, with a slightly different formulation, to Zolotarev [251, 252, 253], and for the formulation presented here to Rotar [214, 215, 216]. Theorem 5.9 is a mild generalization of these and it appears, as well as the material in §§ 5c, d in Liptser and Shiryaev [163].

Chapter VIII

The basic idea that underlies Theorem 1.9, in its present form, comes from Kabanov, Liptser and Shiryaev [121], and the formulation itself appears in Jacod, Kłopotowski and Mémin [105]. But the first "general" convergence results for sums of dependent random variables are due to Bernstein [8] and Lévy [148], and the idea of considering conditional expectations and convergence in measure in the conditions for convergence of rowwise dependent triangular arrays originates in Dvoretsky [46].

Many authors have proved various versions of the theorems presented in Section 2, essentially (but not exclusively) for triangular arrays, and either for finite-dimensional convergence or for functional convergence (usually when the limiting process is a Wiener process, in which case the result is also called "invariance principle"). Let us quote for example Billingsley [11], Borovkov [16, 17], B. Brown [22], B. Brown and Eagleson [23], Durrett and Resnick [45], Gänssler, Strobel and Stute [58], P. Hall [83], Kłopotowski [128, 129], McLeish [174, 175], Rootzen [209, 210], Rosén [212], Scott [220], etc... Books (partly) devoted to this subject include Ibragimov and Linnik [89], Hall and Heyde [84], and to a lesser extent Billingsley [12] and Ibragimov and Has'minski [88]. The forms 2.4 and 2.17 are taken from Jacod and Mémin [108], Liptser and Shiryaev [158, 160], Jacod, Kłopotowski and Mémin [105] for the most general version. Theorem 2.20 is due to Jakubowski and Słominski [114].

The content of §§ 3a,b,c provides unification for a lot of results in the literature, especially concerning triangular arrays of martingale differences (McLeish [174, 175], Scott [220], B. Brown [22], etc...). It also contains the "necessary" part due to Gänssler and Hausler [57] and Rootzen [211] for triangular arrays, and to Liptser and Shiryaev [159, 161] in general (see also Rebolledo [205]).

Theorem 3.36 is essentially due to T. Brown [24], the method is taken from Kabanov, Liptser and Shiryaev [121]. Proposition 3.40 also is due to T. Brown [25]. Theorem 3.43 was first proved by Mémin. Theorem 3.54 is a particular case of a result due to Giné and Marcus [63] (a close look at [63] shows indeed that, although the authors do not speak about characteristics, the basic steps of the proof are the same as here): theorems of such type really belong to the theory of "central limit theorem in Banach spaces" (although $\mathbb{D}(\mathbb{R})$ is not even a topological vector space!). It should be emphazised that our approach does not seem to provide with a very powerful method to solve this type of problems.

Theorem 3.65 (continuous-time version) is borrowed to Touati [235]. The discrete-time version 3.74 is essentially due to Gordin and Lišic [67]. (see also Bhattacharya [9]). This type of theorems shows indeed that our convergence conditions sometimes cannot be applied directly: one has first to transform the semimartingales of interest into other semimartingales to which our theorems apply, plus some remainder terms which we can control.

The content of § 3g (and § 5e as well) is intended to give an idea of a vast subject, initiated by Rosenblatt [213], and pursued by many authors, e.g. Rozanov and Volkonski [218], Statulevicius [227], Serfling [221], Gordin [66], McLeish [173, 174] (who introduced the concept of "mixingale" to unify martingales and mixing processes), etc... For more bibliographical information, and also for many other variants of the theorems, see the books [89] of Ibragimov and Linnik, and [84] of Hall and Heyde. 3.102a is due to Serfling [221], 3.102b is due to McLeish [173]. Results very similar to Theorems 3.79 or 3.97 may be found in Chikin [28] and in Dürr and Goldstein [44].

Theorem 4.1 comes from Jacod, Kłopotowski and Mémin [105], and Theorem 4.10 from Kabanov, Liptser and Shiryaev [121]. § 4c is due to Liptser and Shiryaev [163].

Convergence of triangular arrays to a mixture of infinitely divisible laws is a rather old subject: see the history in the book [84] of Hall and Heyde (see also Kłopotowski [129]) and, from the statistical point of view, in the book [5] of Basawa and Scott. In the present functional setting, § 5a is taken from Jacod, Kłopotowski and Mémin [105], and § 5b is new (see also Grigelionis and Mikulevicius [78] and Rootzen [210]).

Stable convergence has been introduced by Renyi [207], but it also appears in various disguises in control theory (Schäl [219]), Markov processes (Baxter and Chacon [6]), stochastic differential equations (Jacod and Mémin [109]). Here we follow the exposition of Aldous and Eagleson [3]; see also Hall and Heyde [84]. Lemma 5.34 is due to Morando [186] (see also Dellacherie and Meyer [36]). The nesting condition 5.37 appears in McLeish [175] and Hall and Heyde [84] for the discrete-time case, in Feigin [52] for the continuous-time; Theorem 5.42 is due to Feigin [52]. Theorem 5.50 and Corollary 5.51 may be found in Aldous and Eagleson [3] and Durrett and Resnick [45]. The idea of Theorem 5.53 belongs to Renyi [206, 207], as well as the notion of mixing convergence (§ 5d).

Chapter IX

The ideas underlying the martingale method presented here originate in the work [232] of Stroock and Varadhan, which e.g. contains the essentials of Section 1. In fact, these authors show the existence of a weak solution to a stochastic differential equation driven by a Wiener process (or, equivalently, to a martingale problem of the type of § 2d, with $K = 0$) by the very same method used here for Theorem 2.31. An account on the results of Sections 1, 2, 3 in the case where the limiting process is a continuous diffusion process may be found in the book [233] of Stroock and Varadhan; see also Gikhman and Skorokhod [62], and Borovkov [15, 18] for similar results with a different method, and Yoshida [250] for the Trotter-Kato's Theorem refered to in § 2a.

Rebolledo [204] was the first to introduce a variant to conditions 2.7 (although theorems with arbitrary pre-limiting processes have been considered previously by various authors: see e.g. Gikhman [60] and Morkvenas [187]), with a "general" limiting process. These conditions were exploited by Grigelionis and Mikulevicius [77], who proved Theorem 2.11 (including under the weaker assumptions of Remark 2.18) and a version of Theorem 2.31. This last theorem has also been proved, in a more general context, by Traki [236]. Theorems 2.4 and 2.22 were proved by Pagès [192].

The first version of Theorem 3.21 (with a limiting process which is continuous, but not necessarily a diffusion) is due to Liptser and Shiryaev [164] see also Kabanov, Liptser and Shiryaev [122] for point processes. Theorems 3.9 and 3.21, and the local case 3.39, are due to Jacod [103]; theorem 3.35 is borrowed to Pagès [192], who also has a "local" theorem (see 3.51).

In fact, all the results presented in Sections 2 and 3 relate more or less explicitely to stability and convergence results in the theory of stochastic differential equations: see the afore-mentioned books [233] and [62] for example, or Jacod and Mémin [109] for a "general" point of view of this subject.

The content of § 4a is due to various authors: see again the books [233] of Stroock and Varadhan and [62] of Gikhman and Skorokhod, and Jacod [103]. The diffusion approximation of pure-step Markov processes (§ 4b) is due to Kurtz [137, 138] and Allain [4], but there are many other results in various situations, which more or less relate to §§ 4a,b: see for example Papanicolaou, Stroock and Varadhan [195]. Convergence of the empirical distributions to the Brownian bridge is an old result due to Doob [42] and Donsker [41], and initially proved using finite-dimensional convergence methods. § 4d is new.

The content of § 5a is due to Jacod [99], and conditions similar to 5.9 appear for the first time in Jacod and Mémin [109] for studying stability of solutions to stochastic differential equations. Theorems 5.12 and 5.16 are due to Pagès [193]; other results, with weaker assumptions of convergence on the characteristics, but continuity assumptions (in time) on the integrands, may be found in Mamatov [170]. In Jakubowski, Mémin and Pagès [113] one may find a different (and in a sense much stronger) sort of results, without any assumption on the characteristics (but left-continuity assumptions on the integrands).

In connection with what precedes, and especially with the necessary and sufficient conditions of § 4d, we ought to mention here that attempts to adapt the notion of weak convergence in the aim of obtaining necessary conditions more frequently have been undertaken by various authors: Aldous [2], and Helland [86, 87] have introduced a stronger form of convergence (called "extended weak convergence") which essentially amounts to saying that X^n goes to X if $X^n \xrightarrow{\mathscr{L}} X$ and $M^n(\phi) \xrightarrow{\mathscr{L}} M(\phi)$, where $M^n(\phi)_t$ is the conditional expectation of $\phi(X^n)$ with respect to $\mathscr{F}_t^n$ and ϕ runs throughout all bounded Skorokhod-continuous functions. Then these authors proved that if X^n goes to X and X is a martingale, then X^n is close to being a martingale (a property which is completely false for the ordinary weak convergence). Aldous [2] proved that if X^n and X are PII and $X^n \xrightarrow{\mathscr{L}} X$, then the extended convergence also takes place. Jakubowski and Słominski [114] and Kubilius and Mikulevicius [135] have proved that the conditions of Chapter VIII imply the extended weak convergence, when X is a PII without fixed time of discontinuity; Kubilius and Mikulevicius [135] also proved several necessary and sufficient conditions for convergence when the limiting process is a PII. Helland [87] examined the case where the limiting process is a diffusion process.

In § 6a, the first result 6.1 is due to Métivier and Pellaumail [283] and the second one 6.3 to Mackevicius [280]; both were introduced for studying stochastic differential equations. Lemma 6.5 comes from Jacod and Mémin [110]. But the central part of Section 6 has been derived by Słomiński in [291], although Kurtz and Protter [277], [278], [279] have proved a number of more general statements. All of Section 7 comes from Jacod [270].

Chapter X

The motivations of this chapter are statistical, and originate in the work of LeCam [142], who was the first to give a precise meaning to "asymptotically Gaussian experiments". LeCam discussed this notion and conditions to achieve it in various papers (see e.g. [144]) and his book [145]; see also Hájek [81] or Ibragimov and Has'minski [88], while Kutoyants [140] provides a discussion which concerns the case of continuous-time stochastic processes.

Of course, in statistical applications, one is concerned with general parametric models rather than looking at "simple hypotheses": for each n there is a family $(P_\theta^n)_{\theta \in \Theta_n}$ of measures, and thus a family $(Z^{n,\theta})_{\theta \in \Theta_n}$ of likelihood processes with respect to some reference measure Q^n. Then one looks at the limit $Z_t^{n,\theta} \xrightarrow{\mathscr{L}} Z_t$, either as finite-dimensional in (θ, t), or as functional in θ for a given value of t, or functional in (θ, t).

The main result of asymptotic normality 1.12 is new in this form, but it has been proved in the discrete-time case (i.e. when each pre-limiting filtration $\mathbf{F}^n$ is a discrete-time filtration) by Greenwood and Shiryaev [68], and in continuous-time by Vostrikova [242] under a mild additional assumption: this paper also contains the finite-dimensional (in θ) convergence result refered to above, while the functional convergence in (θ, t) is proved in Vostrikova [243]. § 1d is new.

The statistical models called here "exponential families of stochastic processes", which naturally extend the classical exponential statistical models, have been considered by various authors: LeCam [143], Stefanov [228, 229], and also Sörensen [226] (see a complete bibliography in this last paper). They encompass a lot of particular cases (as models for branching processes) and are used mainly for sequential analysis (optimal stopping, etc...). Proposition 2.5 is just an exercise; Theorem 2.12 is new, but Mémin [178] has proved a closely related result (where the conditions are expressed in terms of the processes Z^n themselves), and his method is different (and simpler, but it only gives the functional convergence). See also Taraskin [234] for a result closely related to 2.12.

§ 3a contains simple variations about LeCam's third Lemma (see e.g. Grigelionis and Mikulevicius [79] for results in this direction). § 3b is new, but the same results when the processes X^n are discrete-time are due to Greenwood and Shiryaev [68].

References

1 D. Aldous: Stopping times and tightness. Ann. Probab. 6 (1978) 335–340
2 D. Aldous: Weak convergence of stochastic processes, for processes viewed in the Strasbourg manner. Preprint (1978)
3 D. Aldous, G.K. Eagleson: On mixing and stability of limit theorems. Ann. Probab. 6 (1978) 325–331
4 M.-F. Allain: Approximation par un processus de diffusion des oscillations, autour d'une valeur moyenne, d'une suite de processus de Markov de saut pur. C.R. Acad. Sci. Paris, Ser. A-B, 282 (1976) 891–894
5 I.W. Basawa, D.J. Scott: Asymptotic optimal inference for some non-ergodic models. Lect. Notes in Statistics 17. Springer, Berlin Heidelberg New York (1983)
6 J.R. Baxter, R.V. Chacon: Compactness of stopping times. Z. Wahrsch. Verw. Geb. 40 (1977) 169–182
7 A. Benveniste, J. Jacod: Systèmes de Lévy des processus de Markov. Invent. Math. 21 (1973) 183–198
8 S.N. Bernstein: Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. Math. Ann. 97 (1926) 1–59
9 R.N. Bhattacharya: On the functional central limit theorem and the law of the iterated logarithm for Markov processes. Z. Wahrsch. Verw. Geb. 60 (1982) 185–201
10 K. Bichteler: Stochastic integration theory and L^p theory of semimartingales. Ann. Probab. 9 (1981) 49–89
11 P. Billingsley: The Lindeberg-Lévy theorem for martingales. Proc. Am. Math. Soc. 12 (1961) 788–792
12 P. Billingsley: Convergence of probability measures. Wiley and Sons, New York (1968)
13 P. Billingsley: Conditional distributions and tightness. Ann. Probab. 2 (1974) 480–485
14 R.M. Blumenthal, R.K. Getoor: Markov processes and potential theory. Academic Press, New York (1968)
15 A.A. Borovkov: Theorems on the convergence to Markov diffusion processes. Z. Wahrsch. Verw. Geb. 18 (1970) 47–76
16 A.A. Borovkov: The convergence of distributions of functionals of stochastic processes. Russ. Math. Surveys 27, 1, (1972) 1–42
17 A.A. Borovkov: Convergence of measures and random processes. Russ. Math. Surveys 31, 2 (1976) 1–69
18 A.A. Borovkov: Asymptotic methods in queueing theory. Nauka, Moscow (1980)
19 L. Breiman: Probability. Addison-Wesley, Reading (1968)
20 P. Brémaud: A martingale approach to point processes. PhD Thesis, El. Res. Lab. Berkeley, M-345 (1972)
21 P. Brémaud: Point processes and Queues. Springer, Berlin Heidelberg New York (1981)
22 B. Brown: Martingale central limit theorems. Ann. Math. Statistics 42 (1971) 59–66
23 B. Brown, G.K. Eagleson: Martingale convergence to infinitely divisible laws with finite variances. Trans. Am. Math. Soc. 162 (1971) 449–453
24 T. Brown: A martingale approach to the Poisson convergence of simple point processes. Ann. Probab. 6 (1978) 615–628

25 T. Brown: Compensators and Cox convergence. Math. Proc. Camb. Philos. Soc. 90 (1981) 305–319
26 T. Brown: Some Poisson approximations using compensators. Ann. Probab. 11 (1983) 726–744
27 R.H. Cameron, W.T. Martin: Transformation of Wiener integrals by non-linear transformations. Trans. Am. Math. Soc. 66 (1949) 253–283
28 D.O. Chikin: A functional limit theorem for stationary processes: a martingale approach. Teor. Veroyatnost. i Primenen. 34 (1989), 731–741; translation: Theory Probab. Appl. 34 (1990) 668–678
29 C.S. Chou: Le processus des sauts d'une martingale locale. Séminaire de Proba. XI. Lecture Notes in Mathematics 581, 356–361. Springer, Berlin Heidelberg New York (1977)
30 Y.S. Chow, H. Teicher: Probability theory; independence, interchangeability, martingales. Springer, Berlin Heidelberg New York (1978)
31 P. Courrège: Intégrale stochastique par rapport à une martingale de carré intégrable. Séminaire Brelot-Choquet-Deny (Théorie du potentiel) 7 (1963)
32 P. Courrège, P. Priouret: Temps d'arrêt d'une fonction aléatoire. Publ. Inst. Stat. Univ. Paris 14 (1965) 254–274
33 C. Dellacherie: Capacités et processus stochastiques. Springer, Berlin Heidelberg New York (1972)
34 C. Dellacherie: Intégrales stochastiques par rapport aux processus de Wiener et de Poisson. Séminaire de Proba. VIII. Lecture Notes in Mathematics 381, 25–26 (1974). Correction: Séminaire de Proba. IX. Lecture Notes in Mathematics 465, p. 494 (1975). Springer, Berlin Heidelberg New York
35 C. Dellacherie: Un survol de la théorie de l'intégrale stochastique. Stochastic Processes Appl. 10 (1980) 115–144
36 C. Dellacherie, P.A. Meyer: Probabilités et potentiel, I (1976), II (1982). Hermann: Paris
37 J. De Sam Lazaro, P.A. Meyer: Méthodes de martingales et théorie des flots. Séminaire de Proba. IX. Lecture Notes in Mathematics 465, 1–96. Springer, Berlin Heidelberg New York (1975)
38 C. Doléans-Dade: Quelques applications de la formule de changement de variable pour les semimartingales. Z. Wahrsch. Verw. Geb. 16 (1970) 181–194
39 C. Doléans-Dade, P.A. Meyer: Intégrales stochastiques par rapport aux martingales locales. Séminaire de Proba. IV. Lecture Notes in Mathematics 124, 77–107. Springer, Berlin Heidelberg New York (1970)
40 M. Donsker: An invariance principle for certain probability limit theorems. Mem. Am. Math. Soc. 6 (1951)
41 M. Donsker: Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov Theorems. Ann. Math. Statistics 23 (1952) 277–281
42 J.L. Doob: Heuristic approach to the Kolmogorov-Smirnov Theorems. Ann. Math. Statistics 20 (1949) 393–403
43 J.L. Doob: Stochastic processes. Wiley and Sons, New York (1954)
44 D. Dürr, S. Goldstein: Remarks on the central limit theorem for weakly dependent random variables. Proc. BIBOS Conf. Bielefeld (Albeverio, Blanchard, Streit, eds.). Lecture Notes in Mathematics 1158, 104–118. Springer, Berlin Heidelberg New York (1986)
45 R. Durrett, S.I. Resnick: Functional limit theorems for dependent random variables. Ann. Probab. 6 (1978) 829–846
46 A. Dvoretsky: Asymptotical normality of sums of dependent random variables. Proc. 6th Berkeley Symp., Univ. of Calif. Press (1972) 513–535
47 E.V. Dynkin: Markov processes I. Springer-Verlag (1965): Berlin
48 G.K. Eagleson, J. Mémin: Sur la contiguité de deux suites de mesures, généralisation d'un théorème de Kabanov-Liptser-Shiryaev. Séminaire de Proba. XVI. Lecture Notes in Mathematics 920, 319–337. Springer, Berlin Heidelberg New York (1982)
49 N. El Karoui, J.P. Lepeltier: Représentation des processus ponctuels multivariés à l'aide d'un processus de Poisson. Z. Wahrsch. Verw. Geb. 39 (1977) 111–134
50 H.J. Engelbert, A.N. Shiryaev: On absolute continuity and singularity of probability measures. Banach Cent. Publ. 6 (1979)

51 P. Erdös, M. Kač: On the number of positive sums of independent random variables. Bull. Am. Math. Soc. 53 (1947) 1011–1021

52 P.D. Feigin: Stable convergence of semimartingales. Stoch. Proc. and Appl. 19 (1985) 125–134

53 J. Feldman: Equivalence and perpendicularity of Gaussian measures. Pac. J. Math. 8 (1958) 699–708

54 W. Feller: Uber den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. Math. Z. 40 (1935) 521–559

55 X. Fernique: Comparaison de mesures gaussiennes et de mesures produit. Ann. Inst. Henri Poincaré (Probab. Stat.) 20 (1984) 165–175

56 L. Galtchouk: The structure of a class of martingales. Proc. School-seminar (Druskininkai), Ac. Sci. Lit. SSR-I, 7–32 (1975): Vilnius

57 P. Gänssler, E. Häusler: Remarks on the functional central limit theorem for martingales. Z. Wahrsch. Verw. Geb. 50 (1979) 237–243

58 P. Gänssler, J. Strobel, W. Stute: On central limit theorems for martingale triangular arrays. Acta Math. Acad. Sci. Hung. 31 (1978) 205–216

59 R.K. Getoor: On the construction of kernels. Séminaire de Proba. IX. Lecture Notes in Mathematics 465, 443–463. Springer, Berlin Heidelberg New York (1979)

60 I.I. Gikhman: Limit theorems for sequences of series of random variables. Rep. Mejd. Cb. (1974) 2, 37–47

61 I.I. Gikhman, A.V. Skorokhod: Theory of stochastic processes, I (1974), II (1975), III (1979). Springer, Berlin Heidelberg New York

62 I.I. Gikhman, A.V. Skorokhod: Stochastic differential equations and their applications. Kiev (1982)

63 E. Giné, M.B. Marcus: The central limit theorem for stochastic integrals with respect to a Lévy process. Ann. Probab. 11 (1983) 53–77

64 I.V. Girsanov: On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. Theory Probab. Appl. 5 (1960) 285–301

65 B.W. Gnedenko, A.N. Kolmogorov: Limit distributions for sums of independent random variables. Addison-Wesley, New York (1954)

66 M.I. Gordin: The central limit theorem for stationary processes. Dokl. Akad. Nauk SSSR 188 (1969) 739–741

67 M.I. Gordin, B.A. Lifšic: Central limit theorem for stationary Markov processes. Dokl. Akad. Nauk SSSR 239 (1978) 766–767

68 P. Greenwood, A.N. Shiryaev: Contiguity and the statistical invariance principle. Gordon and Breach, London (1985)

69 B. Grigelionis: On Markov property of stochastic processes. Litovsk. Mat. Sb. 8 (1968) 3, 489–502

70 B. Grigelionis: On the representation of integer-valued measures by means of stochastic integrals with respect to Poisson measure. Litovsk. Mat. Sb. 11 (1971) 93–108

71 B. Grigelionis: On the absolute continuity of measures corresponding to stochastic processes. Litovsk. Mat. Sb. 11 (1971) 783–794

72 B. Grigelionis: On non-linear filtering theory and absolute continuity of measures, corresponding to stochastic processes. Proc. 2d Japan-USSR Symp. Lecture Notes in Mathematics 330, 80–94. Springer, Berlin Heidelberg New York (1973)

73 B. Grigelionis: On relative compactness of sets of probability measures in $D_{[0,\infty)}(\mathbb{R})$. Litovsk. Math. Sb. 13 (1973) 4, 83–96

74 B. Grigelionis: The characterization of stochastic processes with conditionally independent increments. Litovsk. Mat. Sb. 15 (1975) 53–60

75 B. Grigelionis: Stochastic point processes and martingales. Litovsk. Mat. Sb. 15 (1975) 3, 101–114

76 B. Grigelionis: Martingale characterization of stochastic processes with independent increments. Litovsk. Mat. Sb. 17 (1977) 75–86

77 B. Grigelionis, R. Mikulevicius: On weak convergence of semimartingales. Litovsk. Mat. Sb. 21 (1981) 3, 9–24

78 B. Grigelionis, R. Mikulevicius: On stably weak convergence of semimartingales and of point processes. Theory Probab. Appl. 28 (1983) 337–350

79 B. Grigelionis, R. Mikulevicius: On contiguity and weak convergence of probability measures. Proc. 5th Japan-USSR Symp., Lecture Notes in Mathematics 1021, 177–194. Springer, Berlin Heidelberg New York (1983)

80 J. Hájek: On a property of normal distribution of an arbitrary stochastic process. Czech. Math. J. 8 (1958) 610–618

81 J. Hájek: A characterization of limiting distributions of regular estimates. Z. Wahrsch. Verw. Geb. 14 (1970) 324–330

82 J. Hájek, Z. Šidak: Theory of rank tests. Academic Press (1967): New York

83 P. Hall: Martingale invariance principles. Ann. Probab. 5 (1977) 875–887

84 P. Hall, C. Heyde: Martingale limit theory and its applications. Academic Press, New York (1980)

85 W.J. Hall, R.M. Loynes: On the concept of contiguity. Ann. Probab. 5 (1977) 278–282

86 I. Helland: On weak convergence to Brownian motion. Z. Wahrsch. Verw. Geb. 52 (1980) 251–265

87 I. Helland: Minimal conditions for weak convergence to a diffusion process on the line. Ann. Probab. 9 (1981) 429–452

88 I.A. Ibragimov, R.Z. Has'minski: Statistical estimation: asymptotic theory. Springer, Berlin Heidelberg New York (1981)

89 I.A. Ibragimov, Yu.V. Linnik: Independent and stationary sequences of random variables. Walters-Noordhoff, Groningen (1977)

90 I.A. Ibragimov, Yu.A. Rozanov: Gaussian random processes. Springer, Berlin Heidelberg New York (1978)

91 K. Ito: Stochastic integrals. Proc. Imp. Acad. Tokyo 20 (1944) 519–524

92 K. Ito: On a formula concerning stochastic integrals. Nagoya Math. J. 3 (1951) 55–65

93 K. Ito: On stochastic differential equations. Mem. Am. Math. Soc. 4 (1951)

94 J. Jacod: Multivariate point processes: predictable projection, Radon-Nikodym derivative, representation of martingales. Z. Wahrsch. Verw. Geb. 31 (1975) 235–253

95 J. Jacod: Un théorème de représentation pour les martingales discontinues. Z. Wahrsch. Verw. Geb. 34 (1976) 225–244

96 J. Jacod: A general theorem of representation for martingales. Proc. Symp. Pure Math. 31 (1977) 37–53

97 J. Jacod: Local characteristics and absolute continuity conditions for d-dimensional semimartingales. Banach Cent. Publ. 5 (1979) 133–143

98 J. Jacod: Calcul stochastique et problèmes de martingales. Lecture Notes in Mathematics 714. Springer, Berlin Heidelberg New York (1979)

99 J. Jacod: Weak and strong solutions of stochastic differential equations. Stochastics 3 (1980) 171–191

100 J. Jacod: Convergence en loi de semimartingales et variation quadratique. Séminaire de Proba. XV. Lecture Notes in Mathematics 850, 547–560. Springer, Berlin Heidelberg New York (1981)

101 J. Jacod: Processus à accroissements indépendants: une condition nécessaire et suffisante de convergence en loi. Z. Wahrsch. Verw. Geb. 63 (1983) 109–136

102 J. Jacod: Processus de Hellinger, absolue continuité, contiguité. Séminaire de Proba. de Rennes 1983, Univ. de Rennes (1984)

103 J. Jacod: Théorèmes limite pour les processus. Ecole d'été de St-Flour XIII, 1983. Lecture Notes in Mathematics 1117. Springer, Berlin Heidelberg New York (1985)

104 J. Jacod: Sur la convergence des processus ponctuels. To appear (1986)

105 J. Jacod, A. Kłopotowski, J. Mémin: Théorème de la limite centrale et convergence fonctionnelle vers un processus à accroissements indépendants. Ann. Inst. Henri Poincaré (B) 18 (1982) 1–45

106 J. Jacod, J. Mémin: Caractéristiques locales et conditions de continuité absolue pour les semimartingales. Z. Wahrsch. Verw. Geb. 35 (1976) 1–37

107 J. Jacod, J. Mémin: Un nouveau critère de compacité relative pour une suite de processus. Séminaire de Proba. de Rennes 1979, Univ. de Rennes (1980)

108 J. Jacod, J. Mémin: Sur la convergence des semimartingales vers un processus à accroissements indépendants. Séminaire de Proba. XIV. Lecture Notes in Mathematics 784, 227–248. Springer, Berlin Heidelberg New York (1980)

109 J. Jacod, J. Mémin: Weak and strong solutions of stochastic differential equations: existence and stability. In "Stochastic Integrals", ed. by D. Williams, Proc. LMS Symp. Durham 1980. Lecture Notes in Mathematics 851, 169–212. Springer, Berlin Heidelberg New York (1981)

110 J. Jacod, J. Mémin: Sur un type de convergence intermédiaire entre convergence en loi et convergence en probabilité. Séminaire Proba. XV. Lecture Notes in Mathematics 850 (1981) 529–546. Corrections: Séminaire de Proba. XVII. Lecture Notes in Mathematics 986 (1963) 505–511. Springer, Berlin Heidelberg New York

111 J. Jacod, J. Mémin, M. Métivier: On tightness and stopping times. Stochastic Processes Appl. 14 (1983) 109–146

112 J. Jacod, M. Yor: Etude des solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales. Z. Wahrsch. Verw. Geb. 38 (1977) 83–125

113 A. Jakubowski, J. Mémin, G. Pagès: Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorokhod. Probab. Theory and Related Fields 81 (1989) 111–137

114 A. Jakubowski, L. Słominski: Extended convergence to continuous in probability processes with independent increments. Probab. Theo. Rel. Fields 72 (1986) 55–82

115 Yu. Kabanov: Representation of functionals of Wiener and Poisson processes in the form of stochastic integrals. Theory Probab. Appl. 18 (1973) 362–365

116 Yu. Kabanov: An estimate of the variation distance between probability measures. Theory Probab. Appl. 30 (1985) 386–390

117 Yu. Kabanov, R.S. Liptser: On convergence in variation of the distributions of multivariate point processes. Z. Wahrsch. Verw. Geb. 63 (1983) 475–485

118 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: Martingale methods in the theory of point processes. Proc. School-seminar (Druskininkai), Ac. Sci. Lit. SSR, II, 269–354 (1975): Vilnius

119 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: Criteria of absolute continuity of measures corresponding to multivariate point processes. Proc. 3d Japan-USSR Symp. Lecture Notes in Mathematics 550, 232–252. Springer, Berlin Heidelberg New York (1976)

120 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: Absolute continuity and singularity of locally absolutely continuous probability distributions. Math. USSR Sb. 35 (1979) 631–680 (Part I), 36 (1980) 31–58 (Part II) (English transl.)

121 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: Some limit theorems for simple point processes (martingale approach). Stochastics 3 (1981) 203–216

122 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: Weak and strong convergence of the distributions of counting processes. Theory Probab. Appl. 28 (1983) 303–336

123 Yu. Kabanov, R.S. Liptser, A.N. Shiryaev: On the variation distance for probability measures defined on a filtered space. Probab. Theo. Rel Fields 71 (1986) 19–36

124 S. Kakutani: On equivalence of infinite product measures. Ann. Math. 49 (1948) 214–224

125 O. Kallenberg: Random measures. Akademie Verlag (1975): Berlin

126 J. Kerstan, K. Matthes, J. Mecke: Unbegrenzt teilbare Punktprozesse. Akademie-Verlag, Berlin (1974)

127 J.F.C. Kingman: Completely random measures. Pac. J. Math. 21 (1967) 59–78

128 A. Kłopotowski: Limit theorems for sums of dependent random vectors in $\mathbb{R}^d$. Diss. Math. *CLI* (1977) 1–62

129 A. Kłopotowski: Mixtures of infinitely divisible distributions as limit laws for sums of dependent random variables. Z. Wahrsch. Verw. Geb. 51 (1980) 101–115

130 A.N. Kolmogorov: Eine Verallgemeinerung des Laplace-Liapounoffschen Satzes. Izv. Akad. Nauk SSSR (1931) 959–962

131 A.N. Kolmogorov: On the Skorokhod convergence. Theory Probab. Appl. (1956) 215–222

132 E.I. Kolomietch: Relations between triplets of local characteristics of semimartingales. Usp. Mat. Nauk 39 (1984) 163–164

133 T. Komatsu: Markov processes associated with certain integro-differential operators. Osaka J. Math. 10 (1973) 271–303

134 C. Kraft: Some conditions for consistency and uniform consistency of statistical procedures. Univ. of Calif. Publ. Statistics 2 (1955) 125–141

135 K. Kubilius, R. Mikulevicius: On necessary and sufficient conditions for the convergence of semimartingales. Proc. 5th Japan-USSR Symp. Lecture Notes in Mathematics 1021, 338–351. Springer, Berlin Heidelberg New York (1983)

136 H. Kunita, S. Watanabe: On square-integrable martingales. Nagoya J. Math. 30 (1967) 209–245

137 T.G. Kurtz: Limit theorems for sequences of jump Markov processes approximating ordinary differential equations. J. Appl. Probab. 8 (1971) 344–356

138 T.G. Kurtz: Semigroups of conditional shifts and approximation of Markov processes. Ann. Probab. 3 (1975) 618–642

139 A.U. Kussmaul: Stochastic integration and generalized martingales. Pitman, London (1977)

140 Yu.A. Kutoyants: Parameter estimation for stochastic processes. Heldermann-Verlag, Berlin (1984)

141 V.A. Lebedev: On the weak compactness of distribution families of general semimartingales. Theory Probab. Appl. 27 (1982) 14–23

142 L. Le Cam: Locally asymptotically normal families of distributions. Univ. of Calif. Publ. Stat. 3 (1960) 27–98

143 L. Le Cam: A reduction theorem for certain sequential experiments, II. Ann. Stat. 7 (1979) 847–859

144 L. Le Cam: Sur l'approximation de familles de mesures par des familles gaussiennes. Ann. Inst. Henri Poincaré (probab. stat.) 21 (1985) 225–288

145 L. Le Cam: Asymptotic methods in statistical decision theory. Springer, Berlin Heidelberg New York (1986)

146 E. Lenglart: Relation de domination entre deux processus. Ann. Inst. Henri Poincaré (B) 13 (1977) 171–179

147 D. Lépingle: Sur la représentation des sauts d'une martingale. Séminaire Proba. XI. Lecture Notes in Mathematics 581, 418–434. Springer, Berlin Heidelberg New York (1977)

148 P. Lévy: Propriétés asymptotiques des sommes de variables aléatoires enchainées. Bull. Sci. Math. 59 (1935) 84–96 and 108–128

149 P. Lévy: Théorie de l'addition des variables aléatoires. Gauthiers-Villars, Paris (1948)

150 F. Liese: Hellinger integrals of Gaussian processes with independent increments. Stochastics 6 (1982) 81–96

151 F. Liese: Hellinger integrals of diffusion processes. Statistics 17 (1986) 63–78

152 F. Liese: An estimation of Hellinger integrals of point processes. (1984) To appear

153 J.W. Lindeberg: Eine neue Herleitung des Exponentialgesetzes in der Wahrscheislichkeitsrechnung. Math. Z. 15 (1922) 211–225

154 T. Lindvall: Weak convergence of probability measures and random functions in the function space $D[0, \infty)$. J. Appl. Probab. 10 (1973) 109–121

155 R.S. Liptser: On a representation of local martingales. Theory Probab. Appl. 21 (1976) 718–726

156 R.S. Liptser, F. Pukelsheim, A.N. Shiryaev: Necessary and sufficient conditions for contiguity and entire separation of probability measures. Russian Math. Surveys 37 (1982) 107–136

157 R.S. Liptser, A.N. Shiryaev: Statistics of stochastic processes. Springer, Berlin Heidelberg New York (1977)

158 R.S. Liptser, A.N. Shiryaev: A functional central limit theorem for semimartingales. Theory Probab. Appl. 25 (1980) 667–688

159 R.S. Liptser, A.N. Shiryaev: On necessary and sufficient conditions in the functional central limit theorem for semimartingales. Theory Probab. Appl. 26 (1981) 130–135

160 R.S. Liptser, A.N. Shiryaev: Weak convergence of semimartingales to stochastically continuous processes with independent and conditionally independent increments. USSR Math Sb. 116, 3 (1981) 331–358 (English transl.: 44, 3, 299–323)

161 R.S. Liptser, A.N. Shiryaev: On a problem of necessary and sufficient conditions in the functional central limit theorem for local martingales. Z. Wahrsch. Verw. Geb. 59 (1982) 311–318

162 R.S. Liptser, A.N. Shiryaev: On the problem of "predictable" criteria of contiguity. Proc. 5th Japan-USSR Symp. Lecture Notes in Mathematics 1021, 384–418. Springer, Berlin Heidelberg New York (1983)
163 R.S. Liptser, A.N. Shiryaev: On the invariance principle for semimartingales: the nonclassical case. Theory Probab. Appl. 28 (1983) 1–34
164 R.S. Liptser, A.N. Shiryaev: Weak convergence of a sequence of semimartingales to a process of diffusion type. USSR Math. Sb. 121, 2 (1983) 176–200 (English transl.: 49, 171–195)
165 R.S. Liptser, A.N. Shiryaev: On contiguity of probability measures corresponding to semimartingales. Analysis Mathematicae 11 (1985) 93–124
166 R.S. Liptser, A.N. Shiryaev: Theory of martingales. Nauka, Moscow (1986)
167 J. Lobo: Processus à accroissements indépendants et méthode des semimartingales. Thèse de 3ième cycle, Univ. Paris-6 (1985)
168 N. Maigret: Théorèmes de limite centrale pour une chaîne de Markov récurrente Harris positive. Ann. Inst. Henri Poincaré (B) 14 (1978) 425–440
169 S. Makhno: On the existence and uniqueness of solutions of stochastic differential equations. Sb. Teor. Sluch. Proc. 3 (1975) 72–80
170 Kh.M. Mamatov: Weak convergence of stochastic integrals with respect to semimartingales. (Russian) Uspekhi Mat. Nauk 41 (1986) no. 5(251) 187–188
171 G. Maruyama: Continuous time processes and stochastic equations. Rend. Circ. Math. Palermo 4 (1955) 1–43
172 K. Matusita: Decision rules based on the distance for problems of fit, time samples, and estimation. Ann. Math. Statistics 26 (1955) 631–640
173 D.L. McLeish: A maximal inequality and dependent strong laws. Ann. Probab. 3 (1975) 829–839
174 D.L. McLeish: Invariance principles for dependent variables. Z. Wahrsch. Verw. Geb. 32 (1975) 165–178
175 D.L. McLeish: An extended martingale principle. Ann. Probab. 6 (1978) 144–150
176 J. Mémin: Distance en variation et conditions de contiguité pour des lois de processus ponctuels. Séminaire de Proba. de Rennes 1981. Univ. de Rennes (1982)
177 J. Mémin: Sur la contiguité relative de deux suites de processus. Séminaire Proba. XVII. Lecture Notes in Mathematics 986, 371–376. Springer, Berlin Heidelberg New York (1983)
178 J. Mémin: Théorèmes limite fonctionnels pour les processus de vraisemblance (cadre asymptotiquement a-normal). Séminaire de Proba. de Rennes 1985. Univ. de Rennes (1986)
179 J. Mémin, A.N. Shiryaev: Distance de Hellinger-Kakutani des lois correspondant à deux processus à accroissements indépendants. Z. Wahrsch. Verw. Geb. 70 (1985) 67–90
180 M. Métivier: Semimartingales: a course on stochastic integration. De Gruyter, Berlin (1982)
181 M. Métivier, J. Pellaumail: Stochastic integration. Academic Press, New York (1980)
182 P.A. Meyer: Intégrales stochastiques I–IV. Séminaire de Proba. I. Lecture Notes in Mathematics 39, 72–162. Springer, Berlin Heidelberg New York (1967)
183 P.A. Meyer: Un cours sur les intégrales stochastiques. Séminaire Proba. X. Lecture Notes in Mathematics 511, 245–400. Springer, Berlin Heidelberg New York (1976)
184 P.A. Meyer: Le théorème fondamental sur les martingales locales. Séminaire Proba. XI. Lecture Notes in Mathematics 581, 463–464. Springer, Berlin Heidelberg New York (1977)
185 P.A. Meyer, W.A. Zheng: Tightness criteria for laws of semimartingales. Ann. Inst. Henri Poincaré (Probab. Stat.) 20 (1984) 353–372
186 Ph. Morando: Mesures aléatoires. Séminaire Proba. III. Lecture Notes in Mathematics 88, 190–229. Springer, Berlin Heidelberg New York (1969)
187 R. Morkvenas: On the weak convergence of stochastic processes to the solution of martingale problem. Litovsk. Mat. Sb. 15 (1975) 2, 67–75
188 J. Neveu: Bases mathématiques du calcul des probabilités. Masson, Paris (1964)
189 J. Neveu: Processus ponctuels. Ecole d'été de St-Flour VI, 1976. Lecture Notes in Mathematics 598. Springer, Berlin Heidelberg New York (1977)
190 C.M. Newman: The inner product of path space measures corresponding to random processes with random increments. Bull. Am. Math. Soc. 78, 2 (1972) 268–272

191 J. Oosterhoff, W.R. Van Zwet: A note on contiguity and Hellinger distance. In: Contributions to Statistics, ed. J. Jureckova. Reidel, Dordrecht (1979)

192 G. Pagès: Théorèmes limite pour les semimartingales. Thèse de 3ième cycle, Univ. Paris-6 (1985)

193 G. Pagès: Un théorème de convergence fonctionnel pour les intégrales stochastiques. Séminaire de Proba. XX. Lecture Notes in Mathematics 1204, 572–611, Springer, Berlin Heidelberg New York (1986)

194 F. Papangelou: Integrability of expected increments of point processes and a related change of scale. Trans. Am. Math. Soc. 165 (1972) 483–506

195 G.C. Papanicolaou, D.W. Stroock, S.R.S. Varadhan: Martingale approach to some limit theorems. Duke turbulence conf. 1976, Paper no 6, Duke Univ. Math. Series III (1977)

196 K.R. Parthasarathy: Probability measures on metric spaces. Academic Press, New York (1967)

197 V.V. Petrov: Sums of independent random variables. Springer, Berlin Heidelberg New York (1975)

198 D. Pollard: Convergence of stochastic processes. Springer, Berlin Heidelberg New York (1984)

199 Yu.V. Prokhorov: Convergence of stochastic processes and limit theorems in probability theory. Theory Probab. Appl. 1 (1956) 157–214

200 Yu.V. Prokhorov: Probability distribution in functional spaces. Usp. Mat. Nauk 8 (1953) 165–167

201 P. Protter: Semimartingales and measure-preserving flows. Ann. Inst. Henri Poincaré (Probab. Stat.) 22 (1986) 127–147

202 R. Rebolledo: La méthode des martingales appliquée à la convergence en loi des processus. Mém. Soc. Math. Fr. 62 (1979)

203 R. Rebolledo: Central limit theorems for local martingales. Z. Wahrsch. Verw. Geb. 51 (1980) 269–286

204 R. Rebolledo: Sur l'existence de solutions à certains problèmes de martingales. C.R. Acad. Sci. Paris, Ser. A 290 (1980) 843–846

205 R. Rebolledo: The central limit theorem for semimartingales, necessary and sufficient conditions. Preprint (1980)

206 A. Rényi: On mixing sequences of sets. Acta Math. Acad. Sci. Hung. 9 (1958) 215–228

207 A. Rényi: On stable sequences of events. Sankhya, Ser A, 25 (1963) 293–302

208 D. Revuz: Markov chains. North-Holland, Amsterdam (1975)

209 H. Rootzen: On the functional limit theorem for martingales. Z. Wahrsch. Verw. Geb. 38 (1977) 199–210

210 H. Rootzen: A note on convergence to mixtures of normal distributions. Z. Wahrsch. Verw. Geb. 38 (1977) 211–216

211 H. Rootzen: On the functional limit theorem for martingales, II. Z. Wahrsch. Verw. Geb. 51 (1981) 79–93

212 B. Rosén: On the central limit theorem for sums of dependent random variables. Z. Wahrsch. Verw. Geb. 7 (1967) 48–82

213 M. Rosenblatt: A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA 42 (1956) 43–47

214 V.I. Rotar: On extensions of Lindeberg-Feller Theorem. Mat. Zametki 18 (1975) 129–135

215 V.I. Rotar: Some remarks on summing independent variables in the nonclassical case. Theory Probab. Appl. 21 (1976) 130–137

216 V.I. Rotar: On summation of independent random variables: non-classical conditions. Usp. Mat. Nauk 37 (1982) 137–156

217 G. Roussas: Contiguity of probability measures. Cambridge Univ. Press, London (1972)

218 Yu.A. Rozanov, V.A. Volkonski: Some limit theorems for random functions I. Theory Probab. Appl. 4 (1959) 178–197

219 M. Schäl: Conditions for optimality in dynamic programming and for the limit of n-stages optimal policies to be optimal. Z. Wahrsch. Verw. Geb. 32 (1975) 179–196

220 D.J. Scott: Central limit theorems for martingales with stationary increments using Skorokhod representation theorem. Adv. Appl. Probab. 5 (1975) 119–137

221 R.J. Serfling: Contributions to central limit theory for dependent variables. Ann. Math. Stat. 39 (1968) 1158–1175
222 A.N. Shiryaev: Probability. Springer, Berlin Heidelberg New York (1984)
223 A.V. Skorokhod: Limit theorems for stochastic processes. Theory Probab. Appl. 1 (1956) 261–290
224 A.V. Skorokhod: Random processes with independent increments. Nauka, Moscow (1964)
225 A.V. Skorokhod: Studies in the theory of random processes. Addison-Wesley, Reading (1965)
226 M. Sørensen: On sequential maximum likelihood estimates for exponential families of stochastic processes. Int. Stat. Review 54 (1986) 191–210
227 V.A. Statulevicius: Some new results for sums of weakly dependent random variables. Theory Probab. Appl. 5 (1960) 233–234
228 V.T. Stefanov: Explicit solutions in the first passage problem for exponential type processes. Preprint (1987)
229 V.T. Stefanov: Efficient sequential estimation in exponential-type processes. Ann. Statist. 14 (1986) 1606–1611
230 C. Stone: Weak convergence of stochastic processes defined on a semifinite time interval. Proc. Am. Math. Soc. 14 (1963) 694–696
231 D.W. Stroock: Diffusion processes associated with Lévy generators. Z. Wahrsch. Verw. Geb. 32 (1975) 209–244
232 D.W. Stroock, S.R.S. Varadhan: Diffusion processes with continuous coefficients I, II. Commun. Pure Appl. Math. 22 (1969) 345–400, 479–530
233 D.W. Stroock, S.R.S. Varadhan: Multidimensional diffusion processes. Springer, Berlin Heidelberg New York (1979)
234 A.F. Taraskin: On the behavior of the likelihood ratio of semimartingales. Theory Probab. Appl. 29 (1984) 452–464
235 A. Touati: Théorèmes de limite centrale fonctionnelle pour les processus de Markov. Ann. Inst. Henri Poincaré (B) 19 (1983) 43–55
236 M. Traki: Existence de solutions d'un problème de martingales. C.R. Acad. Sci. Paris, Sér. A, 297 (1983) 353–356
237 I. Vajda: On the f-divergence and singularity of probability measures. Period. Math. Hung. 2 (1972) 223–234
238 E. Valkeila: Studies in distributional properties of counting processes. Thesis, Dept Math., Univ. of Helsinki (1984)
239 E. Valkeila, L.Yu. Vostrikova: An integral representation for the Hellinger process. Math. Scand. 58 (1986) 239–254
240 J.H. Van Schuppen, E. Wong: Transformations of local martingales under a change of law. Ann. Probab. 2 (1974) 879–888
241 L.Yu. Vostrikova: On necessary and sufficient conditions for convergence of probability measures in variation. Stochastic Processes Appl. 18 (1984) 99–112
242 L.Yu. Vostrikova: Functional limit theorems for the likelihood ratio processes. To appear in: Ann. Univ. Sci. Budap. R. Eötvös Nom. (1987)
243 L.Yu. Vostrikova: On the weak convergence of likelihood ratio processes of general statistical parametric models. Stochastics 23 (1988) 277–298
244 S. Watanabe: On discontinuous additive functionals and Lévy measures of a Markov process. Jap. J. Math. 34 (1964) 53–79
245 W. Whitt: Some useful functions for functional limit theorems. Math. Oper. Res. 5 (1980) 67–85
246 N. Wiener: Generalized harmonic analysis. Acta Math. 55 (1930) 117–258
247 N. Wiener: Extrapolation, interpolation and smoothing of stationary time series. Wiley and Sons, New York (1949)
248 T. Yamada, S. Watanabe: On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto Univ. 11 (1971) 156–167
249 M. Yor: Représentation intégrale des martingales de carré intégrable. C.R. Acad. Sci. Paris, Sér. A-B 282 (1976) 899–901

250 K. Yoshida: Functional analysis. Springer, Berlin Heidelberg New York (1966)
251 V.M. Zolotarev: Théorèmes limites pour les sommes de variables aléatoires indépendantes qui ne sont pas infinitésimales. C.R. Acad. Sci. Paris, Sér. A, 264 (1967) 799–800
252 V.M. Zolotarev: A generalization of the Lindeberg-Feller Theorem. Theory Probab. Appl. 12 (1967) 608–618
253 V.M. Zolotarev: Théorèmes limites généraux pour les sommes de variables aléatoires indépendantes. C.R. Acad. Sci. Paris, Sér. A, 270 (1970) 899–902
254 V.M. Zolotarev: Modern theory of summation of independent random variables. Nauka, Moscow (1986)
255 J-P. Ansel, C. Stricker: Couverture des actifs contingents et prix maximum. Ann. Inst. H. Poincaré Probab. Statist. 30 (1994), 303–315
256 J. Bertoin: Lévy processes. Cambridge University Press, Cambridge (1996)
257 C. Chou: Caractérisation d'une classe de semimartingales. Séminaire de Proba. XIII, Lecture Notes in Mathematics 721, 250–252. Springer, Berlin (1979)
258 F. Delbaen, W. Schachermayer: A general version of the fundamental theorem of asset pricing. Math. Ann. 300 (1994) 463–520
259 F. Delbaen, W. Schachermayer: The fundamental theorem of asset pricing for unbounded stochastic processes. Math. Ann. 312 (1998) 215–250
260 S.N. Ethier, T.G. Kurtz: Markov processes. Characterization and convergence. John Wiley & Sons, New York (1986)
261 E. Eberlein, J. Jacod; On the range of option prices. Finance and Stochastics 1 (1997), 131–140
262 M. Emery: Une topologie sur l'espace des semimartingales. Séminaire de Proba. XIII, Lecture Notes in Mathematics 721, 260–281. Springer, Berlin (1979)
263 M. Emery: Compensation de processus V.F. non localement intégrables. Séminaire de Proba. XIV, Lecture Notes in Mathematics 784, 152–160. Springer, Berlin (1980)
264 F. Esscher: On the probability function in the collective theory of risk. Skandinavisk Aktuarietidskrift 15 (1932) 175–195
265 T. Goll, J. Kallsen: A complete explicit solution to the log-optimal portfolio problem. Preprint (2001)
266 B. Grigelionis: The characterization of stochastic processes with conditionally independent increments. Litovk. Math. Sb. 15 (1975), 53–60
267 K. Ito, S. Watanabe: Transformation of Markov processes by multiplicative functionals. Ann. Institut Fourier 15 (1965), 15–30
268 J. Jacod: Intégrales stochastiques par rapport à une semimartingale vectorielle et changements de filtration. Séminaire de Proba. XIV, Lecture Notes in Mathematics 784, 161–172. Springer, Berlin (1980)
269 J. Jacod: Une généralisation des semimartingales: les processus admettant un processus à accroissements indépendants tangent. Séminaire de Proba. XVIII, Lecture Notes in Mathematics 1059, 91–118. Springer, Berlin (1984)
270 J. Jacod: On continuous Gaussian martingales and stable convergence in law. Séminaire de Proba. XXXI, Lecture Notes in Mathematics 1655, 232–246. Springer, Berlin (1997)
271 J. Jacod: On processes with conditional independent increments and stable convergence in law. To appear in Séminaire de Proba. XXXVI
272 J. Jacod, P. Protter (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. Annals Probab. 26, 267–307
273 J. Jacod, A.N. Shiryaev: Local martingales and the fundamental asset pricing theory. Finance and Stochastics 2 (1998), 259–273
274 Yu. Kabanov: On tha FTAP of Kreps–Delbaen–Schachermayer. Statistics and Control of Stoch. Processes (R. Sh. Liptser Festchrift). World Scientific Publ., Singapore (1997) 191–203
275 J. Kallsen, A.N. Shiryaev: Time change representation of stochastic integrals. Theory Probab. Appl. 46, 3 (2000) 579–585
276 J. Kallsen, A.N. Shiryaev: The cumulant processes and Esscher's change of measure. Finance and Stochastics (to appear) (2001)
277 T.G. Kurtz, P. Protter: Weak limit theorems for stochastic integrals and stochastic differential equations. Annals Probab. 19 (1991), 1035–1070

278 T.G. Kurtz, P. Protter: Wong-Zakai corrections, random evolutions, and simulation schemes for SDEs. in “Stochastic Analysis”, Proc. Honor Moshe Zakai, Academic Press (1991) 331–346
279 Kurtz T.G., Protter P. (1996): Weak convergence of stochastic integrals and differential equations, 1995 CIME School in Probability. Lect. Notes in Math. 1627: 1–41, Springer-Verlag, Berlin
280 V. Mackevicius: S^p stability of symmetric stochastic differential equations with discontinuous driving semimartingales. Ann. Inst. Henri Poincaré B, 23 (1987), 575–592
281 J. Mémin: Espaces de semimartingales et changements de probabilité. Z. für Wahrsch. 52 (1980), 9–40
282 J. Mémin, L. Słomiński: Condition UT et stabilité en loi des solutions d'équations différentielles stochastiques. Séminaire de Proba. XXV, Lecture Notes in Mathematics 1485, 162–177. Springer, Berlin (1991)
283 M. Métivier, J. Pellaumail: On a stopped Doob's inequality and general stochastic equations. Annals Probab. 8 (1980), 96–114
284 P.A. Meyer: Multiplicative decompositions of positive supermartingales. In “Markoff Processes and Potential Theory”, J. Chover ed., Wiley and Sons: New York (1967)
285 D.L. Ocone: A symmetry characterization of conditionally independent increment martingales. Barcelona Seminar on Stochastic Analysis 1991, Progr. Probab. 32, Birkhäuser, Basel (1993)
286 P. Protter: Stochastic integration and differential equations. A new approach. Springer-Verlag, Berlin, (1990)
287 D. Revuz, M. Yor: Continuous martingales and Brownian motion. Third edition. Springer-Verlag, Berlin (1999)
288 J. Rosiński, W.A. Woyczinski: On Ito stochastic integration with respect to p–stable motion: Inner clock, integrability of sample paths, double and multiple integrals. Annals Probab. 14 (1986), 271–286
289 K-I Sato: Lévy processes and infinitely divisible distributions. Cambridge University Press, Cambridge (1999)
290 A.N. Shiryaev: Essentials of Stochastic Finance. Facts, Models, Theory. World Scientific Publ., Singapore (1999)
291 L. Słomiński: Stability of strong solutions of stochastic differential equations. Stoch. Processes and their Appl. 31 (1989), 173–202
292 C. Stricker: Quelques remarques sur la topologie des semimartingales. Applications aux intégrales stochastiques. Séminaire de Proba. XV, Lecture Notes in Mathematics 850, 499–522. Springer, Berlin (1981)
293 C. Stricker: Lois de semimartingales et critères de compacité. Séminaire de Proba. XXIX, Lecture Notes in Mathematics 1123, 209–217. Springer, Berlin (1985)
294 M. Traki: Solutions faibles d'équations différentielles stochastiques et problèmes de martingales. Thèse Université de Rennes-1 (1985)

Index of Symbols

Classes of Processes

Other Symbols

Index of Terminology

Index of Topics

Markov processes

Point processes and multivariate point processes

Processes with independent increments

Index of Conditions for Limit Theorems

Grundlehren der mathematischen Wissenschaften

A Series of Comprehensive Studies in Mathematics

A Selection

221. Mumford: Algebraic Geometry I: Complex Projective Varieties
222. Lang: Introduction to Modular Forms
223. Bergh/Löfström: Interpolation Spaces. An Introduction
224. Gilbarg/Trudinger: Elliptic Partial Differential Equations of Second Order
225. Schütte: Proof Theory
226. Karoubi: K-Theory. An Introduction
227. Grauert/Remmert: Theorie der Steinschen Räume
228. Segal/Kunze: Integrals and Operators
229. Hasse: Number Theory
230. Klingenberg: Lectures on Closed Geodesics
231. Lang: Elliptic Curves. Diophantine Analysis
232. Gihman/Skorohod: The Theory of Stochastic Processes III
233. Stroock/Varadhan: Multidimensional Diffusion Processes
234. Aigner: Combinatorial Theory
235. Dynkin/Yushkevich: Controlled Markov Processes
236. Grauert/Remmert: Theory of Stein Spaces
237. Köthe: Topological Vector Spaces II
238. Graham/McGehee: Essays in Commutative Harmonic Analysis
239. Elliott: Probabilistic Number Theory I
240. Elliott: Probabilistic Number Theory II
241. Rudin: Function Theory in the Unit Ball of C^n
242. Huppert/Blackburn: Finite Groups II
243. Huppert/Blackburn: Finite Groups III
244. Kubert/Lang: Modular Units
245. Cornfeld/Fomin/Sinai: Ergodic Theory
246. Naimark/Stern: Theory of Group Representations
247. Suzuki: Group Theory I
248. Suzuki: Group Theory II
249. Chung: Lectures from Markov Processes to Brownian Motion
250. Arnold: Geometrical Methods in the Theory of Ordinary Differential Equations
251. Chow/Hale: Methods of Bifurcation Theory
252. Aubin: Nonlinear Analysis on Manifolds. Monge-Ampère Equations
253. Dwork: Lectures on ρ-adic Differential Equations
254. Freitag: Siegelsche Modulfunktionen
255. Lang: Complex Multiplication
256. Hörmander: The Analysis of Linear Partial Differential Operators I
257. Hörmander: The Analysis of Linear Partial Differential Operators II
258. Smoller: Shock Waves and Reaction-Diffusion Equations
259. Duren: Univalent Functions
260. Freidlin/Wentzell: Random Perturbations of Dynamical Systems
261. Bosch/Güntzer/Remmert: Non Archimedian Analysis – A System Approach to Rigid Analytic Geometry
262. Doob: Classical Potential Theory and Its Probabilistic Counterpart
263. Krasnosel'skiĭ/Zabreĭko: Geometrical Methods of Nonlinear Analysis
264. Aubin/Cellina: Differential Inclusions
265. Grauert/Remmert: Coherent Analytic Sheaves
266. de Rham: Differentiable Manifolds
267. Arbarello/Cornalba/Griffiths/Harris: Geometry of Algebraic Curves, Vol. I
268. Arbarello/Cornalba/Griffiths/Harris: Geometry of Algebraic Curves, Vol. II
269. Schapira: Microdifferential Systems in the Complex Domain
270. Scharlau: Quadratic and Hermitian Forms

271. Ellis: Entropy, Large Deviations, and Statistical Mechanics
272. Elliott: Arithmetic Functions and Integer Products
273. Nikol'skiĭ: Treatise on the Shift Operator
274. Hörmander: The Analysis of Linear Partial Differential Operators III
275. Hörmander: The Analysis of Linear Partial Differential Operators IV
276. Liggett: Interacting Particle Systems
277. Fulton/Lang: Riemann-Roch Algebra
278. Barr/Wells: Toposes, Triples and Theories
279. Bishop/Bridges: Constructive Analysis
280. Neukirch: Class Field Theory
281. Chandrasekharan: Elliptic Functions
282. Lelong/Gruman: Entire Functions of Several Complex Variables
283. Kodaira: Complex Manifolds and Deformation of Complex Structures
284. Finn: Equilibrium Capillary Surfaces
285. Burago/Zalgaller: Geometric Inequalities
286. Andrianaov: Quadratic Forms and Hecke Operators
287. Maskit: Kleinian Groups
288. Jacod/Shiryaev: Limit Theorems for Stochastic Processes
289. Manin: Gauge Field Theory and Complex Geometry
290. Conway/Sloane: Sphere Packings, Lattices and Groups
291. Hahn/O'Meara: The Classical Groups and K-Theory
292. Kashiwara/Schapira: Sheaves on Manifolds
293. Revuz/Yor: Continuous Martingales and Brownian Motion
294. Knus: Quadratic and Hermitian Forms over Rings
295. Dierkes/Hildebrandt/Küster/Wohlrab: Minimal Surfaces I
296. Dierkes/Hildebrandt/Küster/Wohlrab: Minimal Surfaces II
297. Pastur/Figotin: Spectra of Random and Almost-Periodic Operators
298. Berline/Getzler/Vergne: Heat Kernels and Dirac Operators
299. Pommerenke: Boundary Behaviour of Conformal Maps
300. Orlik/Terao: Arrangements of Hyperplanes
301. Loday: Cyclic Homology
302. Lange/Birkenhake: Complex Abelian Varieties
303. DeVore/Lorentz: Constructive Approximation
304. Lorentz/v. Golitschek/Makovoz: Construcitve Approximation. Advanced Problems
305. Hiriart-Urruty/Lemaréchal: Convex Analysis and Minimization Algorithms I. Fundamentals
306. Hiriart-Urruty/Lemaréchal: Convex Analysis and Minimization Algorithms II. Advanced Theory and Bundle Methods
307. Schwarz: Quantum Field Theory and Topology
308. Schwarz: Topology for Physicists
309. Adem/Milgram: Cohomology of Finite Groups
310. Giaquinta/Hildebrandt: Calculus of Variations I: The Lagrangian Formalism
311. Giaquinta/Hildebrandt: Calculus of Variations II: The Hamiltonian Formalism
312. Chung/Zhao: From Brownian Motion to Schrödinger's Equation
313. Malliavin: Stochastic Analysis
314. Adams/Hedberg: Function Spaces and Potential Theory
315. Bürgisser/Clausen/Shokrollahi: Algebraic Complexity Theory
316. Saff/Totik: Logarithmic Potentials with External Fields
317. Rockafellar/Wets: Variational Analysis
318. Kobayashi: Hyperbolic Complex Spaces
319. Bridson/Haefliger: Metric Spaces of Non-Positive Curvature
320. Kipnis/Landim: Scaling Limits of Interacting Particle Systems
321. Grimmett: Percolation
322. Neukirch: Algebraic Number Theory
323. Neukirch/Schmidt/Wingberg: Cohomology of Number Fields
324. Liggett: Stochastic Interacting Systems: Contact, Voter and Exclusion Processes
325. Dafermos: Hyperbolic Conservation Laws in Continuum Physics
326. Waldschmidt: Diophantine Approximation on Linear Algebraic Groups