Semimartingales and Stochastic Differential Equations

by

Philip Protter

Technical Report #85-25

Department of Statistics
Purdue University

September 1985

Part of this work was supported by NSF Grant #MCS-8301073

Table of Contents

Semimartingales and Stochastic Differential Equations

Semimartingales and Stochastic Differential Equations

I. INTRODUCTION

1. Introduction

The general theory of stochastic integration has recently been enjoying a great deal of interest. No doubt this is partly due to newly discovered applications in applied fields (such as filtering theory, control theory, the theory of continuous trading in economics, and statistical communication theory) as well as to applications in theoretical mathematics. An impediment however to the diffusion of the subject is its high barrier to entry: traditionally one must master a large part of the abstract and technically difficult "general theory of processes" in order to learn even the elementary theory. This is the approach taken by recent pedagogic treatments ([5], [11], [17], [23], [28], and [43]) which to varying degrees follow the historical development of the subject. (The small book by Letta [27] is a qualified exception).

The approach of these notes is quite different. Recent developments have made possible an introduction to both stochastic integration and stochastic differential equations that has virtually no technical prerequisites. We have systematically tried to keep our approach on as technically simple a level as possible. Our treatment has three principal advantages: (i) it is rapid and direct: the reader learns about semimartingales and stochastic integrals immediately; (ii) it is intuitive: for example, the restriction of the space of integrands to $\mathbb{L}$ allows the integral to be expressed as the limit of sums; and (iii) a number of traditionally difficult results (such as III.2.2, III.2.3, III.2.4,...) become startlingly simple and natural. While a disadvantage may be a loss of generality, with some work (Chapter IV) one can develop enough "general theory" to extend our approach to the general case. We do this in Chapter V.

2. A Brief History

The Wiener process is a mathematical model of Brownian motion and it has a.s. continuous paths of infinite variation on all compacts. This of course precludes the possibility of considering the process path by path as a Riemann-Stieltjes differential. Nevertheless the concept of "dW" had long had an intuitive meaning before K. Itô ([21], 1944) found a way to make it mathematically rigorous. Itô used his integral to construct diffusions as solutions of stochastic differential equations:

$$X_t = X_0 + \int_0^t \sigma(s,X_s)dW_s + \int_0^t b(s,X_s)ds \ .$$

This gave a probabilistic method of studying diffusions which even today is still a primary tool. J. Doob ([16],1953) realized that Itô's integral was really a martingale integral and he proposed a general martingale integral; but he needed to be able to decompose a submartingale X into a sum $X = M + A$, with M a martingale and A an increasing process. P. A. Meyer ([31], 1963) found the right conditions under which this could be done, and later K. Itô and S. Watanabe ([22], 1965) introduced the notion of local martingale, extending the Doob-Meyer decomposition to arbitrary submartingales: any submartingale X has a decomposition $X = M \pm A$ where M is a local martingale and A is an increasing process. In 1967 H. Kunita and S. Watanabe [24], building on earlier work of Ph. Courrege ([7], 1963), developed an elegant theory complete with a change of variables formula. But it was P. A. Meyer ([32], 1967) who made the crucial realization that if one restricted the space of integrands to predictable processes then the jumps of the stochastic integral process behaved nicely. Later C. Doléans-Dade and P. A. Meyer ([14], 1970) removed the extraneous but simplifying assumption of quasi-left-continuity of the underlying filtration, an assumption whose origins came from connections with Markov process theory.

The theory was now essentially complete, but in 1976 P. A. Meyer published a "course"on stochastic integration [34]. This seminal work proved many new and fascinating results about stochastic integration (including, for example, the existence of semimartingale local time) and it led to an explosion of interest in the subject. Meyer emphasized the centrality of semimartingales in his course, and this together with the work of M. Métivier and J. Pellaumail ([30], [37]) eventually led to the theorem of K. Bichteler ([1], 1981) and C. Dellacherie ([10], 1980), discovered independently. This theorem showed, in essence, that semimartingales are the most general reasonable stochastic differentials. Indeed this theorem is the starting point of these notes, if not the pedagogical basis.

The theory of stochastic differential equations lagged behind that of stochastic integration, the first existence and uniqueness results being those of P. Protter ([38], [39], 1977) and C. Doléans-Dade ([13], 1976) as well as C. Doléans-Dade and P. A. Meyer ([15], 1977). The question of stability was resolved by P. Protter [41] and then it was elegantly handled by M. Emery [18], [19], and [20], who also developed the semimartingale topology. M. Métivier and J. Pellaumail used a different approach to stochastic differential equations ([29], 1980), based on a profound martingale inequality. They also independently discovered the semimartingale topology. The last step was to return the theory of stochastic differential equations to its Markov process birthplace, and this was done by E. Çinlar, J. Jacod, P. Protter, and M. Sharpe [6] (see also [40]).

3. The Contents

We give here a brief description of these notes. In Chapter II we establish notational conventions and prove a few preliminary results. We begin Chapter III by defining semimartingales as "good integrators" instead of the traditional definition (e.g. [11] or [28]); our approach here was inspired by C. Dellacherie [10] and we follow in large part an article of E. Lenglart [26]. This new approach is not widespread: the only similar treatment we know of is that of G. Letta [27]. We treat stochastic integration with respect to semimartingales for processes in $\mathbb{L}$ and prove many of their well known properties: the Kunita-Watanabe inequality, behavior of the jumps, integration by parts, Itô's formula, the Fisk-Stratonovich integral, exponential semimartingales, etcetera.

In Chapter IV we "pay the price" for the simplicity of Chapter III. Our goal is to prove the theorem of K. Bichteler and C. Dellacherie, and to do this we must develop some of the "general theory". We develop the minimum that we need, however, and we have tried to keep the proofs on as technically simple a level as possible. Nevertheless if the reader is willing to stipulate Theorems IV.1.2 and IV.1.3 stated in the introduction to the chapter, then Chapter IV can be omitted on a first reading without loss for Chapters V, VI, and VII.

In Chapter V we extend the space of integrands from $\mathbb{L}$ to predictable processes. While ostensibly this larger space of integrands is not needed, for example, to prove Itô's lemma or to treat stochastic differential equations, it is needed to consider semimartingale local time, which we present in paragraph four.

Chapter VI is devoted to stochastic differential equations with semimartingale driving terms. We establish general existence and uniqueness results (though not the most general (cf [23]), again because of a desire to keep to a low level of technicality). The semimartingale topology is presented and we use it to study the stability of solutions of stochastic differential equations.

Stochastic differential equations arose from a desire to study Markov processes (i.e. diffusions) and in Chapter VII we return them to a Markov framework. The theory of Markov processes is the most technically overwhelming branch of Probability theory, but we have presented the necessary notational conventions in paragraph two. Our approach follows that of [6] but we have simplified it since we are considering a special situation. Theorem IV.3.12 extends the classical result of K. Itô: a solution of an appropriate stochastic differential equation is strong Markov if the semimartingale driving term has independent increments (cf [40]). If the increments are only conditionally independent, then a weaker result holds (VI.3.5).

4. Acknowledgements

I wish to thank the Chilean Winter School of Probability and Statistics, and Professor Rolando Rebolledo in particular, for inviting me to participate. While preparing these lectures I enjoyed the hospitality of the Université de Provence in Marseille, France. I was further supported by the National Science Foundation while working on these notes. I am grateful for the patient cooperation of my typist, Ms. Eleanor Gerns. Finally, I wish to thank Eimear Goggin and Richard Stockbridge for having alerted me to numerous mistakes in a first version of these notes.

II. PRELIMINARIES AND NOTATIONAL CONVENTIONS

1. General Definitions and Notations

In Chapters III through VI we will always assume given an underlying filtered probability space: $(\Omega, \underline{\underline{F}}, \underline{\underline{F}}_t, P)$, where $(\Omega, \underline{\underline{F}}, P)$ is a complete probability space, and $\underline{\underline{F}}_t$ is an increasing family of σ-algebras ($\underline{\underline{F}}_s \subseteq \underline{\underline{F}}_t$ if $s \leq t$). Moreover, we assume that $\underline{\underline{F}}_0$ contains all the P-null sets of $\underline{\underline{F}}$, and that $(\underline{\underline{F}}_t)_{t \geq 0}$ is right continuous: that is, $\underline{\underline{F}}_t = \bigcap_{u>t} \underline{\underline{F}}_u$. These are known as the "usual hypotheses". Chapter VII involves connections to Markov processes, and since we use the theory of Dynkin realizations of Markov processes, we use the notations of Bluemthal and Getoor [2], and the underlying assumptions are slightly different and are presented there. Indeed all new notations in Chapter VII are explained there; the reader need not be familiar with [2].

A process X is said to be adapted if X_t is $\underline{\underline{F}}_t$ - measurable for each $t \geq 0$. A process X is cadlag if it has paths which are right continous with left limits, a.s. ["cadlag" is an acronym from the French "continue à droite, limites à gauche"]. The space of cadlag adapted processes is denoted $\mathbb{D}$; an adapted process X is in $\mathbb{L}$ if it has left continuous paths with right limits. For a process X in $\mathbb{D}$ we will often be interested in its jumps, and we let $\Delta X_t = X_t - X_{t-}$, the jump at time t, where X_{t-} denotes $\lim_{\substack{s \to t \\ s<t}} X_s$, the left limit at t.

A stopping time T is a random variable $T: \Omega \to \overline{\mathbb{R}}_+ = [0,\infty]$ with the additional property that $\{T \leq t\} \in \underline{\underline{F}}_t$, each $t > 0$. We let $s \wedge t$ = minimum (s,t), $s \vee t$ = maximum (s,t), and we adopt the convention $X^T = X_{t \wedge T} 1_{\{T>0\}} = X_t 1_{\{t<T\}} + X_T 1_{\{t \geq T>0\}}$, where 1_A is the indicator function of a set A. We also write $X^{T-} = X_t 1_{\{t<T\}} + X_{T-} 1_{\{t \geq T>0\}}$. We assume the reader is familiar with the notions of martingale and submartingale, as well as Lebesgue-Stieltjes integration.

(1.1) Definition. An adapted process X is a local martingale if there exists a sequence of stopping times T^n increasing to ∞ a.s. such that X^{T^n} is a uniformly integrable martingale for each n.

The stopping time σ-algebra $\underline{\underline{F}}_T$ is defined to be: $\underline{\underline{F}}_T = \{\Lambda \in \underline{\underline{F}}: \Lambda \cap \{T \leq t\} \in \underline{\underline{F}}_t\}$. It is easy to check that $\underline{\underline{F}}_T = \sigma\{H_T: H \in \mathbb{D}\}$, which is perhaps a more intuitive characterization. A class of processes contained in $\mathbb{L}$ that is used throughout these notes is:

(1.2) Definition. A process H is said to be simple predictable if H has a representation:

$$H_t = H_0 1_{\{0\}} + \sum_{i=1}^{n} H_i 1_{]T_i, T_{i+1}]}(t)$$

where $0 = T_0 \leq T_1 \leq \ldots \leq T_{n+1} < \infty$ is a finite sequence of stopping times, and where $H_i \in \underline{\underline{F}}_{T_i}$, $|H_i| < \infty$ a.s., each i, $0 \leq i \leq n$. The collection of simple predictable processes is denoted $\underline{\underline{S}}$.

We can topologize $\underline{\underline{S}}$ by uniform convergence (uniform in (t,ω)), and we denote $\underline{\underline{S}}$ with this topology by $\underline{\underline{S}}_u$. We let L^0 denote the space of finite random variables, topologized by convergence in probability.

2. An Elementary Result

A very deep result is that if H is a progressively measurable process*, then $T(\omega) = \inf\{t: H_t(\omega) \in \Lambda\}$, for Λ a Borel set in $\mathbb{R}$, is a stopping time (cf. [8, p. 51], or [9] for a direct proof). We do not here have need of such a general result, and we can content ourselves with the following theorems.

* that is, the function $(s,\omega) \to X_s(\omega)$ from $[0,t] \times \Omega$ into $\mathbb{R}$ is $\underline{\underline{B}}[0,t] \otimes \underline{\underline{F}}_t / \underline{\underline{B}}$-measurable, each $t > 0$.

(2.1) Theorem. If either $X \in \mathbb{D}$, or $X \in \mathbb{L}$; if Λ is an open subset of $\mathbb{R}$; and if

$$T(\omega) = \inf\{t: X_t(\omega) \in \Lambda\},$$

then T is a stopping time.

Proof. Since $\underline{\underline{F}}_t$ is right continuous, it suffices to show $\{T < t\} \in \underline{\underline{F}}_t$. But

$$\{T < t\} = \bigcup_{s \in \mathbb{Q} \cap [0,t[} \{X_s \in \Lambda\},$$

and this is in $\underline{\underline{F}}_t$. □

(2.2) Theorem. If $X \in \mathbb{D}$; if Λ is a closed subset of $\mathbb{R}$; and if

$$T(\omega) = \inf\{t: X_t(\omega) \in \Lambda \text{ or } X_{t-}(\omega) \in \Lambda\},$$

then T is a stopping time.

Proof. Let $A_n = \{x: d(x,\Lambda) < 1/n\}$. Then A_n is an open set, and the result follows from

$$\{T \leq t\} = \{X_t \in \Lambda \text{ or } X_{t-} \in \Lambda\} \cup \{\bigcap_n \bigcup_{s \in \mathbb{Q} \cap [0,t[} \{X_s \in A_n\}\}. \quad \square$$

We remark that we use the convention $\inf \phi = +\infty$. We will freely use elementary facts about stopping times, such as: if S and T are stopping times, then:

$S \wedge T$, $S \vee T$, $S+T$ are stopping times;

$S \leq T$ a.s. implies $\mathcal{F}_S \subseteq \mathcal{F}_T$

$\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S \wedge T}$.

3. Caveats

We wish to alert the educated reader to two deviations from the norm. Our definition of a semimartingale (III.1.1) is not the traditional one but is, of course, equivalent to it. The traditional definition (a process X is a semimartingale if X has a decomposition X = M + A where M is a local martingale and A is adapted, right continuous, and has paths of finite variation on compacts) is close to what we call a decomposable process (IV.1.1) and is shown to be equivalent to it (IV.1.2 and IV.4.8).

We have systematically avoided the classification of stopping times, predictable projections, and dual predictable projections (compensations), in order to keep these notes on as technically simple a level as possible. For this reason we are led to the resurrection of the anachronistic notion of naturality in Chapter IV (IV.2.3). Of course the process A in definition IV.2.3 is natural if and only if it is predictably measurable, but the proof of such a result is not necessary to our development and we have not included it.

III. SEMIMARTINGALES AND STOCHASTIC INTEGRALS

For much of this chapter, we follow the article of E. Lenglart [26], which was in turn inspired by the article of C. Dellacherie [10]. See also [42].

1. Introduction to Semimartingales

Given a process X, we define a linear mapping $I_X\colon \underline{\underline{S}} \to L^0$ as follows:

$I_X(H) = H_0X_0 + \sum_{i=1}^{n} H_i(X_{T_{i+1}} - X_{T_i})$, where $H \in \underline{\underline{S}}$ has the representation

$$H_t = H_0 1_{\{0\}} + \sum_{i=1}^{n} H_i 1_{]T_i, T_{i+1}]}$$

(1.1) Definition: A process X is a total semimartingale if X is cadlag, adapted, and $I_X\colon \underline{\underline{S}}_u \to L^0$ is continuous.

A process X is called a semimartingale if, for each $t \in [0,\infty[$, X^t is a total semimartingale.

2. Stability Properties of Semimartingales

We state a sequence of theorems giving some of the stability results which are particularly simple with our approach.

(2.1) Theorem. The set of (total) semimartingales is a vector space.

Proof. This is immediate from the definition. □

(2.2) Theorem. If Q is a probability and absolutely continuous with respect to P, then every P-semimartingale (total) X is a Q-semimartingale (total).

Proof. Convergence in P-probability implies convergence in Q-probability. Thus the theorem follows from the definition of X. □

(2.3) Theorem. Let $(P_k)_{k\geq 1}$ be a sequence of probabilities such that X is a

P_k-semimartingale for each k. Let $R = \sum_{k=1}^{\infty} \lambda_k P_k$, where $\lambda_k \geq 0$, each k, and $\sum_{k=1}^{\infty} \lambda_k = 1$. Then X is a semimartingale under R as well.

Proof. Suppose $H^n \in \underline{\underline{S}}$ such that H^n converges to H under R. Then H^n converges to H for all P_k with $\lambda_k > 0$. Therefore $I_X(H^n)$ converges to $I_X(H)$ in P_k-probability for all such k as well. This then implies $I_X(H^n)$ converges to $I_X(H)$ under R. □

(2.4) Theorem. Let X be a semimartingale for the filtration $(\mathcal{F}_t)_{t \geq 0}$. Let $(\mathcal{G}_t)_{t \geq 0}$ be a subfiltration of $(\mathcal{F}_t)_{t \geq 0}$, such that X is adapted to the $\mathcal{G}$-filtration. Then X is a $\mathcal{G}$-semimartingale.

Proof. Since $\underline{\underline{S}}(\mathcal{G})$ is contained in $\underline{\underline{S}}(\mathcal{F})$, this again follows from the definition. □

(2.5) Theorem. Let X be a cadlag, adapted process; let (T_n) be a sequence of positive r.v. increasing to ∞ a.s.; and let (X^n) be a sequence of semimartingales such that, for each n, $X^{T_n -} = (X^n)^{T_n -}$. Then X is a semimartingale.

Proof. We wish to show X^t is a total semimartingale, each $t > 0$. Define

$R_n = T_n 1_{(T_n \leq t)} + \infty 1_{(T_n > t)}$. Then $P\{|I_X(H)| \geq c\} \leq P\{|I_{X^n}(H)| \geq c\} + P(R_n < \infty)$.

But $P(R_n < \infty) = P(T_n \leq t)$, and since T_n increases to ∞ a.s., $P(T_n \leq t) \to 0$ as $n \to \infty$. Thus if H^k tends to 0 in $\underline{\underline{S}}_u$, given $\varepsilon > 0$, we choose n so that $P(R_n < \infty) < \varepsilon/2$, and then choose k so large that $P\{|I_{X^n}(H^k)| \geq c\} < \varepsilon/2$. Thus, for k large enough, $P\{|I_{X^t}(H^k)| \geq c\} < \varepsilon$. □

(2.6) Corollary. Let X be a process. If there exists a sequence (T_n) of stopping times increasing to ∞ a.s., such that X^{T_n} is a semimartingale, each n, then X is also a semimartingale.

3. Elementary Examples of Semimartingales

(3.1) Proposition: Each adapted process with cadlag paths of finite variation on compacts (of finite total variation) is a semimartingale (a total semimartingale).

Proof. It suffices to observe that $|I_X(H)| \leq ||H||_u \int_0^\infty |dX_s|$, where $\int_0^\infty |dX_s|$ denotes the Lebesque-Stieltjes total variation. □

(3.2) Proposition: Each square integrable martingale with cadlag paths is a semimartingale.

Proof. Let X be a square integrable martingale with $X_0 = 0$, and let $H \in \underline{\underline{S}}$. It suffices to observe that $E\{(I_X(H))^2\}$

$$= E\{(\sum_{i=0}^{n} H_i(X_{T_{i+1}} - X_{T_i}))^2\}$$

$$= E\{\sum_{i=0}^{n} H_i^2(X_{T_{i+1}} - X_{T_i})^2\}$$

$$\leq ||H||_u^2 \, E\{\sum_{i=0}^{n} (X_{T_{i+1}} - X_{T_i})^2\}$$

$$= ||H||_u^2 \, E\{\sum_{i=0}^{n} (X^2_{T_{i+1}} - X^2_{T_i})\} = ||H||_u^2 \, E\{X^2_{T_{n+1}}\}$$

$$\leq ||H||_u^2 \, E(X^2_\infty). \quad \square$$

(3.3) Corollary. Each cadlag, locally square integrable martingale is a semimartingale.

Proof. Apply Proposition (3.2) together with Theorem (2.5). □

(3.4) Corollary. A local martingale with continuous paths is a semimartingale.

Proof. Without loss of generality we assume $X_0 = 0$. Let $R_p = \inf\{t: |X_t| \geq p\}$.

It is easy to check that X^{R_p} is a bounded (and hence square-integrable) martingale for each p. Also R_p increases to ∞ a.s. as p increases to ∞. Thus X is a locally square-integrable martingale and hence a semimartingale by Corollary (3.3). □

Note that, in particular, the Wiener process (i.e., Brownian motion), which is a martingale with continuous paths, is a semimartingale.

(3.5) Definition. We will say an adapted process X with right continuous paths is decomposable if it can be decomposed $X_t = X_0 + M_t + A_t$, where $M_0 = A_0 = 0$, M is a locally square integrable local martingale, and A has paths of finite variation on compacts.

(3.6) Corollary. *A decomposable process is a semimartingale.*

Proof. This is Proposition (3.1), Corollary (3.3), and the result that semimartingales form a vector space. □

Let X be a process with stationary and independent increments. It is well known that any such process X has a modification with cadlag paths. We therefore assume X has cadlag paths. Define $J_t = \sum_{s \leq t} \Delta X_s 1_{\{|\Delta X_s| \geq 1\}}$, and we set $Y_t = X_t - J_t$. ($\Delta X_t = X_t - X_{t-}$). Then the Lévy theory (e.g. [1]) of such processes tells us that both J and Y are again processes with stationary and independent increments, and that Y has a finite mean. The stationarity implies the function $t \to E(Y_t)$ is affine. Thus $X_t = \{Y_t - E(Y_t)\} + \{E(Y_t) + J_t\}$, where $Y_t - E(Y_t)$ is a locally bounded martingale and where $E(Y_t) + J_t$ has paths of finite variation on compacts. Thus X is a semimartingale by Corollary (3.6). □

A process with stationary and independent increments is a time homogeneous strong Markov process and also, as we have just seen, a semimartingale. All such Markov processes that are semimartingales have been characterized ([6]).

4. Stochastic Integrals

Let $\mathbb{D}$ represent all adapted processes with cadlag paths. On $\mathbb{D}$ we put the topology of uniform convergence on compacts in probability, abbreviated "ucp." That is, for a process $Y \in \mathbb{D}$, let $Y_t^* = \sup_{s \le t} |Y_s|$. Then Y^n converges to Y in ucp if $(Y^n - Y)_t^*$ converges to 0 in probability for every $t \geq 0$. Note that this topology is metrizable: the metric $d(Y,Z) = \sum_{n=1}^{\infty} \frac{1}{2^n} E\{\min(1,(Y-Z)_n^*\}$ defines a topology compatible with ucp and thus makes $(\mathbb{D},\text{ucp})$ into a complete metric space.

Let $\mathbb{L}$ denote all adapted processes in $\mathbb{L}$ with left continuous paths a.s. Also $b\mathbb{L}$ denotes all processes in $\mathbb{L}$ with uniformly bounded paths.

(4.1) Theorem. The space $\underline{\underline{S}}$ is dense in $\mathbb{L}$ under the ucp topology.

Proof. Let $Y \in \mathbb{L}$. Let $T_n = \inf\{t: |Y_t| > n\}$. Then T_n is a stopping time and $Y^n = Y^{T^n} 1_{\{T^n > 0\}}$ are in $b\,\mathbb{L}$ and converge to Y in ucp. Thus $b\,\mathbb{L}$ is dense in $\mathbb{L}$. Thus without loss we now assume $Y \in b\,\mathbb{L}$. Define Z by $Z_t = \lim_{\substack{u \to t \\ u > t}} Y_u$. Then $Z \in \mathbb{D}$ is the cadlag modification of Y. For $\varepsilon > 0$, define

$$T_0^\varepsilon = 0$$

$$T_{n+1}^\varepsilon = \inf\{t: t > T_n^\varepsilon \text{ and } |Z_t - Z_{T_n^\varepsilon}| > \varepsilon\}.$$

Since Z is cadlag, the T_n^ε are stopping times increasing to ∞ a.s. an n increases.

Let $Z_n^\varepsilon = \sum_n Z_{T_n^\varepsilon} 1_{[T_n^\varepsilon, T_{n+1}^\varepsilon[}$,

each $\varepsilon > 0$. Then Z_n^ε are bounded and converge uniformly to Z as ε tends to 0. Let $U^\varepsilon = Z_0 1_{\{0\}} + \sum_n Z_{T_n^\varepsilon} 1_{]T_n^\varepsilon, T_{n+1}^\varepsilon]}$, and the preceding implies U^ε converges uniformly on compacts to $Z_- = Y$.

Finally, define

$$Y^{n,\varepsilon} = Y_0 1_{\{0\}} + \sum_{i=1}^{n} Y_{T_i^\varepsilon} 1_{]T_i^\varepsilon \wedge n,\, T_{i+1}^\varepsilon \wedge n]}$$

and this can be made arbitrarily close to $Y \in b\mathbb{L}$ by taking ε small enough and n large enough. □

(4.2) <u>Definition</u>: For $H \in \underline{\underline{S}}$ and X a cadlag process, <u>define the (linear) mapping</u> $J_X: \underline{\underline{S}} \to \mathbb{D}$ by:

$$J_X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i (X^{T_{i+1}} - X^{T_i})$$

for H with representation

$$H = H_0 + \sum_{i=1}^{n} H_i 1_{]T_i, T_{i+1}]} \quad ,$$

$H_i \in \mathcal{F}_{T_i}$ and $0 = T_0 < T_1 < \ldots < T_{n+1} < \infty$ stopping times.

(4.3) <u>Definition</u>: For $H \in \underline{\underline{S}}$ and X an adapted cadlag process, we call $J_X(H)$ <u>the stochastic integral of</u> H <u>with respect to</u> X.

We use interchangeably three notations for the stochastic integral:

(4.4) $J_X(H) = \int H_s dX_s = H \cdot X.$

(4.5) <u>Theorem</u>. <u>Let</u> X <u>be a semimartingale</u>. <u>Then the mapping</u> $J_X: \underline{\underline{S}}_{ucp} \to \mathbb{D}_{ucp}$ <u>is continuous</u>.

<u>Proof</u>. $\underline{\underline{S}}_{ucp}$ denotes the space $\underline{\underline{S}}$ endowed with the ucp topology. Since we are only dealing with convergence on compact sets, without loss of generality we take X to be a total semimartingale. It suffices to show that if H^k converges to 0 (ucp), then one can extract a subsequence k_n such that $J_X(H^{k_n})$ converges to 0 ucp. First suppose H^k tends to 0 uniformly and is uniformly bounded. Let

$$T^k = \inf\{t: |(H^k\cdot X)_t| \geq \delta\}.$$

Then $P\{(H^k\cdot X)^* \geq \delta\} = P\{(H^k\cdot X)^*_{t\wedge T^k} \geq \delta\}$

$$\leq P\{(H^k\cdot X^t)^*_{T^k} \geq \delta\}$$

$$= P\{|H^k 1_{[0,T^k]} \cdot X^t| \geq \delta\}$$

$$= P\{|J_{X^t}(H^k 1_{[0,T^k]})| \geq \delta\}$$

which tends to 0 by the definition of total semimartingale (1.1).

Therefore we have for $\varepsilon > 0$, $t > 0$, that there exists a c such that if $H \in \underline{\underline{S}}$ with $||H||_u < c$, then $P\{J_X(H)^*_t > \delta\} < \varepsilon/2$. For fixed t we can find k_n such that $P(R_n < \infty) < \varepsilon/2$ where

$$R_n = \inf\{s: |H^{k_n}_{s\wedge t}| > c\}.$$

Moreover by the left continuity one has $|(H^{k_n})^{R_n}| \leq c$. Thus:

$$P\{(H^{k_n}\cdot X)^*_t \geq \delta\} \leq P\{(H^{k_n}\cdot X)^*_{R_n} > \delta\} + P\{R_n < \infty\}$$

$$\leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

and the continuity is established. □

We have seen that when X is a semimartingale, the integration operator J_X is continuous on $\underline{\underline{S}}_{ucp}$(4.5), and also that $\underline{\underline{S}}_{ucp}$ is dense in $\mathbb{L}_{ucp}$ (4.1), hence we extend the integration operator J_X from $\underline{\underline{S}}$ to $\mathbb{L}$ by continuity.

(4.6) <u>Definition</u>: Let X be a semimartingale. The continuous linear mapping $J_X: \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ obtained as the extension of $J_X: \underline{\underline{S}} \to \mathbb{D}$ is called the <u>stochastic integral</u>.

5. Properties of the Stochastic Integral

Throughout this paragraph X will denote a semimartingale and H will denote an element of $\mathbb{L}$. The stochastic integral defined in paragraph 4 will also be denoted $J_X(H) = H\cdot X = \int H_s dX_s$.

(5.1) Theorem. Let T be a stopping time. Then $(H\cdot X)^T = H1_{[0,T]}\cdot X = H\cdot(X^T)$.

(5.2) Theorem. The jump process $\Delta(H\cdot X)_s$ is indistinguishable from $H_s(\Delta X_s)$.

Proofs: Both properties are clear when $H\in \underline{\underline{S}}$, and thus they follow when $H\in\mathbb{L}$ by passing to the limit (convergence in ucp). □

Let Q denote another probability law, and let $H_Q\cdot X$ denote the stochastic integral of H with respect to X computed under the law Q.

(5.3) Theorem. Let $Q \ll P$. Then $H_Q\cdot X$ is Q-indistinguishable from $H_P\cdot X$.

Proof. Note that by Theorem (2.2), X is known to be a Q-semimartingale. The theorem is clear if $H\in \underline{\underline{S}}$, and it follows for $H\in\mathbb{L}$ by passage to the limit in the ucp topology, since convergence in P-probability implies convergence in Q-probability. □

(5.4) Corollary. Let P_k be a sequence of probabilities such that X is a P_k-semimartingale for each k. Let $R = \sum_{k=1}^{\infty} \lambda_k P_k$ where $\lambda_k \geq 0$, each k, and $\sum_{k=1}^{\infty} \lambda_k = 1$. Then $H_R\cdot X = H_{P_k}\cdot X$, P_k a.s., for all k such that $\lambda_k > 0$.

Proof. If $\lambda_k > 0$ then $P_k \ll R$, and the result follows from (5.3). Note that by (2.3) we know that X is an R-semimartingale. □

(5.5) Corollary. Let P and Q be any probabilities and suppose X is a semimartingale relative to both P and Q. Then there exists a process $H\cdot X$ which is a version of both $H_P\cdot X$ and $H_Q\cdot X$.

Proof. Let $R = \frac{P+Q}{2}$. Then $H_R \cdot X$ is such a process by Corollary (5.4). □

(5.6) Theorem. *Let $(\mathcal{G}_t)_{t\geq 0}$ be another filtration such that H is in both $\mathbb{L}(\mathcal{G})$ and $\mathbb{L}(\mathcal{F})$, and such that X is also a $\mathcal{G}$-semimartingale. Then $H_{\mathcal{G}} \cdot X = H_{\mathcal{F}} \cdot X$.*

Proof. $\mathbb{L}(\mathcal{G})$ denotes left continuous processes adapted to the filtration $(\mathcal{G}_t)_{t\geq 0}$. Since $H \in \mathbb{L}(\mathcal{G}) \cap \mathbb{L}(\mathcal{F})$, we can find $H^n \in \underline{\underline{S}}(\mathcal{G}) \cap \underline{\underline{S}}(\mathcal{F})$ converging to H in ucp, and since the result is clear for $H^n \in \underline{\underline{S}}$, the full result follows by passing the limit. □

(5.7) Lemma. *If the semimartingale X has paths of finite variation on compacts, then $H \cdot X$ is indistinguishable from the Lebesgue-Stieltjes integral, computed path by path.*

Proof. The result is evident for $H \in \underline{\underline{S}}$. Let $H^n \in \underline{\underline{S}}$ converge to H in ucp. Then there exists a subsequence n_k such that $\lim_{n_k \to \infty} (H^{n_k} - H)^*_t = 0$ a.s., and the result follows by interchanging limits, justified by the uniform a.s. convergence. □

(5.8) Theorem. *Let X, $\overline{X}$ be two semimartingales, and let $H, \overline{H} \in \mathbb{L}$. Let $A = \{\omega: H.(\omega) = \overline{H}.(\omega)$ and $X.(\omega) = \overline{X}.(\omega)\}$, and let $B = \{\omega: t \to X_t(\omega)$ is of finite variation on compacts$\}$. Then $H \cdot X = \overline{H} \cdot \overline{X}$ on A, and $H \cdot X$ is equal to a path by path Lebesgue-Stieltjes integral on B.*

Proof. Without loss of generality we assume $P(A) > 0$. Define a new probability law Q by $Q(\Lambda) = P(\Lambda|A)$. Then under Q we have that H and $\overline{H}$ as well as X and $\overline{X}$ are indistinguishable. Thus $H_Q \cdot X = \overline{H}_Q \cdot \overline{X}$, and hence $H \cdot X = \overline{H} \cdot \overline{X}$ P- a.s. on A by (5.3), since $Q \ll P$.

As for the second assertion, if $B = \Omega$ the result is merely Lemma (5.7). Define R by $R(\Lambda) = P(\Lambda|B)$, assuming without loss that $P(B) > 0$. Then $R \ll P$ and $B = \Omega$, R- a.s. Hence $H_R \cdot X$ equals the Lebesgue-Stieljes integral R- a.s. by (5.7), and the result follows by (5.3). □

The preceding theorem and following corollary are known as the local behavior of the integral.

(5.9) Corollary. With the notations of (5.8), let S,T be two stopping times with $S < T$. Define

$$C = \{\omega: H_t(\omega) = \overline{H}_t(\omega);\ X_t(\omega) = \overline{X}_t(\omega);\ S(\omega) < t < T(\omega)\}$$

$$D = \{\omega: t \to X_t(\omega) \text{ is of finite variation on } S(\omega) \le t \le T(\omega)\}\ .$$

Then $H\cdot X^T - H\cdot X^S = \overline{H}\cdot\overline{X}^T - \overline{H}\cdot\overline{X}^S$ on C and $H\cdot X^T - H\cdot X^S$ equals a path by path Lebesgue-Stieltjes integral on D.

Proof. Let $Y_t = X_t - X_{t\wedge S}$. Then $H\cdot Y = H\cdot X - H\cdot X^S$, and Y does not charge the set [0,S] , which is evident, or which - alternately - can be viewed as an easy consequence of (5.8). One now applies (5.8) to Y to obtain the result. □

(5.10) Theorem (Associativity). The stochastic integral process $Y = H\cdot X$ is itself a semimartingale, and for $G \in \mathbb{L}$ we have

$$G\cdot Y = G\cdot(H\cdot X) = (GH)\cdot X.$$

Proof. Suppose we know $Y = H\cdot X$ is a semimartingale. Then $G\cdot Y = J_Y(G)$. If G, H are in $\underline{\underline{S}}$, then it is clear that $J_Y(G) = J_X(GH)$. The associativity then extends to $\mathbb{L}$ by continuity.

It remains to show that $Y = H\cdot X$ is a semimartingale. Let (H^n) be in $\underline{\underline{S}}$ converging in ucp to H. Then $H^n\cdot X$ converges to $H\cdot X$ in ucp. Thus there exists a subsequence (n_k) such that $H^{n_k}\cdot X$ converges a.s. to $H\cdot X$.

Let $G \in \underline{\underline{S}}$ and let $Z^{n_k} = H^{n_k}\cdot X$, $Z = H\cdot X$. The Z^{n_k} are semimartingales converging pointwise to the process Z. For $G \in \underline{\underline{S}}$, $J_Z(G)$ is defined for any process Z; so we have

$$J_Z(G) = G\cdot Z = \lim_{n_k\to\infty} G\cdot Z^{n_k}$$

$$= \lim_{n_k\to\infty} G\cdot(H^{n_k}\cdot X)$$

$$= \lim_{n_k\to\infty} (GH^{n_k})\cdot X$$

which equals $\lim_{n_k\to\infty} J_X(GH^{n_k}) = J_X(GH)$, since X is a semimartingale. Therefore $J_Z(G) = J_X(GH)$ for $G \in \underline{\underline{S}}$.

Let G_n converge to G in $\underline{\underline{S}}_u$. Then G_nH converges to GH in $\mathbb{L}_{ucp}$, and since X is a semimartingale, $\lim_{n\to\infty} J_Z(G_n) = \lim_{n\to\infty} J_X(G_nH) = J_X(GH) = J_Z(G)$. This implies Z^t is a total semimartingale, and so $Z = H\cdot X$ is a semimartingale. □

(5.11) Theorem. *Let X be a locally square integrable local martingale, and let $H \in \mathbb{L}$. Then the stochastic integral $H\cdot X$ is also a locally square integrable local martingale.*

Proof. We have seen that a locally square integrable martingale is a semimartingale (3.3), so we can formulate $H\cdot X$. Without loss of generality, assume $X_0 = 0$. Also, if T^k increases to ∞ a.s. and $(H\cdot X)^{T_k}$ is a locally square integrable martingale for each k, it is simple to check that $H\cdot X$ itself is one. Thus without loss we assume X is a square integrable martingale. By stopping H, we may further assume H is bounded, by ℓ. Let $H^n \in \underline{\underline{S}}$ be such that H^n converges to H in ucp. We can then modify H^n, call it $\tilde{H}^n$, such that $\tilde{H}^n$ is bounded by ℓ, $\tilde{H}^n \in \underline{\underline{S}}$, and $\tilde{H}^n$ converges uniformly to H in probability in [0,t]. Then

$$E\{(\tilde{H}^n\cdot X)_t^2\} = E\{\sum_{i=1}^{k_n} \tilde{H}_i^n (X_t^{T_{i+1}} - X_t^{T_i})^2\}$$

$$\leq \ell^2 E\{\sum_{i=1}^{k_n} (X_{T_{i+1}}^2 - X_{T_i}^2)\}$$

$$\leq \ell^2 E(X_\infty^2),$$

and hence $(\tilde{H}^n\cdot X)_t$ are uniformly bounded in L^2 and thus uniformly integrable. Passing to the limit then shows both that $H\cdot X$ is a martingale and that it is square integrable. □

The next result (Theorem 5.13) shows that the stochastic integral in our framework (thanks to the smooth paths of the integrands) has a simple interpretation as a limit of sums.

(5.12) Definition. Let σ denote a sequence (finite or infinite) of stopping times: $0 = T_0 \leq T_1 \leq \ldots \leq T_i \leq \ldots$ σ is called a random partition.

We say a sequence of random partitions $\sigma^n = T_0^n \leq T_1^n \leq \ldots \leq T_i^n \leq \ldots$ tends to the identity if $\sup_k T_k^n < \infty$ a.s., and

(i) $\lim_n \sup_k T_k^n = \infty$ a.s.

(ii) $||\sigma_n|| = \sup_k (T_{k+1}^n - T_k^n)$ converges to 0 a.s.

For Y a process and σ a random partition, let $Y^\sigma = Y_0 1_{\{0\}} + \sum_i Y_{T_i} 1_{]T_i,T_{i+1}]}$.

It is easy to check that

$$\int Y_s^\sigma dX_s = Y_0 X_0 + \sum_i Y_{T_i}(X^{T_{i+1}} - X^{T_i}),$$

for any semimartingale X, any optional process Y.

(5.13) Theorem. Let X be a semimartingale, and let Y be an adapted process with paths that either have right limits and are left continuous, or are cadlag. Let (σ_n) be a sequence of random partitions tending to the identity. Then the processes

$Y_0X_0 + \sum_i Y_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})$ <u>tend to the stochastic integral</u> $(Y_-)\cdot X$ <u>in ucp</u>.

<u>Proof</u>. (The notation Y_- means the process whose value at s is given by $(Y_-)_s = \lim_{\substack{u\to s \\ u<s}} Y_u$; also, $(Y_-)_0 = Y_0$, by convention). We prove the theorem for the case where Y is cadlag, the other case being analogous. Y cadlag implies $Y_- \in \mathbb{L}$. Let $Y^k \in \underline{S}$ such that Y^k converges to Y_- (ucp). We have:

$$\int (Y_- - Y^{\sigma_n})_s dX_s = \int (Y_- - Y^k)_s dX_s + \int (Y^k - (Y^k)^{\sigma_n})_s dX_s + \int ((Y^k)^{\sigma_n} - Y^{\sigma_n})_s dX_s \ .$$

The first term on the right side equals $J_X(Y_- - Y^k)$, and since J_X is continuous in $\mathbb{L}_{ucp}$ and since $Y_- - Y^k \to 0$, we have $\int (Y_- - Y^k)_s dX_s \to 0$(ucp). The same reasoning applies to the third term, for fixed n, as $k \to \infty$. Indeed, the convergence to 0 of $[(Y^k)^{\sigma_n} - Y^{\sigma_n}]$ as $k \to \infty$ is uniform in n.

It remains to consider the middle term on the right side above. For fixed k and ω, the integrand $(Y_s^k(\omega) - (Y^k)_s^{\sigma_n}(\omega))$ converges to 0 uniformly on compacts; moreover since the Y^k are simple predictable we can write the stochastic integrals in closed form, and since X is right continuous the integrals (for fixed (k,ω)) $\int (Y^k - (Y^k)^{\sigma_n})_s dX_s$ tend to 0 as $n \to \infty$. Thus one merely chooses k so large that the first and third terms are small, and then for fixed k, the middle term can be made small for large enough n. □

6. The Quadratic Variation of a Semimartingale.

The reason for the name "quadratic variation" is justified by Theorem (6.2-ii). This seemingly innocuous process has a central role not just in the stochastic calculus, but also in modern martingale theory (the $\underline{\mathcal{H}}^p$ spaces, BMO, and the Burkholder-Davis-Gundy type of inequalities).

(6.1) Definition. Let X be a semimartingale. The quadratic variation process of X, denoted $[X,X] = ([X,X]_t)_{t\geq 0}$, is defined by:

$$[X,X] = X^2 - 2\int X_- \, dX \quad (X_{0-} = 0).$$

The next theorem gives some elementary properties of $[X,X]$. (X is assumed to be a given semimartingale throughout this paragraph.)

(6.2) Theorem. The quadratic variation process of X is a cadlag, increasing, adapted process. Moreover it satisfies:

(i) $[X,X]_0 = X_0^2$ and $\Delta[X,X] = (\Delta X)^2$

(ii) If σ_n is a sequence of random partitions tending to the identity, then

$$X_0^2 + \sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \longrightarrow [X,X]$$

with convergence in ucp, where σ_n is the sequence $0 = T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots$, where T_i^n are stopping times.

Proof. X is cadlag, adapted, and so also is $\int X_- dX$ by its definition; thus $[X,X]$ is cadlag, adapted as well. Recall the property of the stochastic integral: $\Delta(X_- \cdot X) = X_- \Delta X$.

Then
$$(\Delta X)_s^2 = (X_s - X_{s-})^2 = X_s^2 - 2X_sX_{s-} + X_{s-}^2$$
$$= X_s^2 - X_{s-}^2 + 2X_{s-}(X_{s-} - X_s).$$
$$= \Delta(X^2)_s - 2X_{s-}(\Delta X_s),$$

from which part (i) follows.

For part (ii), by replacing X with $\tilde{X} = X - X_0$, we may assume $X_0 = 0$. Let $R_n = \sup_i T_i^n$. Then $R_n < \infty$ a.s., and thus by telescoping series:

$$(X^2)^{R_n} = \sum_i \{(X^2)^{T_{i+1}^n} - (X^2)^{T_i^n}\}$$

converges ucp to X^2. Moreover, the series $\sum_i X_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})$ converges in ucp to

$\int X_- dX$ by Theorem (5.13), since X is cadlag. Since $b^2 - a^2 - 2a(b-a) = (b-a)^2$, and since $X_{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n}) = X^{T_i^n}(X^{T_{i+1}^n} - X^{T_i^n})$, we can combine the two series convergences above to obtain the result. Finally, note that if $s < t$, then the approximating sums in part (ii) include more terms (all nonnegative), so it is clear that $[X,X]$ is nondecreasing. (Note that, a priori, one only has $[X,X]_s \leq [X,X]_t$ a.s., with the null set depending on s and t; it is the property that $[X,X]$ has cadlag paths that allows one to eliminate the dependence of the null set on s and t.) □

(6.3) Definition. Let X and Y be two semimartingales. The bracket product of X and Y is defined by

$$[X,Y] = \frac{1}{2}\{[X+Y,\ X+Y] - [X,X] - [Y,Y]\}$$

(6.4) Proposition. *The bracket product $[X,Y]$ of two semimartingales has paths of finite variation on compacts, and it is also a semimartingale.*

Proof. By definition, $[X,Y]$ is the difference of two increasing processes, hence its paths are of finite variation. Moreover, the paths are clearly cadlag, and the process is adapted. Hence by (3.1) it is a semimartingale. □

(6.5) Theorem (Integration by Parts). *Let X,Y be two semimartingales. Then:*

$$XY = \int X_- dY + \int Y_- dX + [X,Y].$$

Proof. By the definitions we have

$$[X,Y] = \frac{1}{2}\{(X+Y)^2 - 2\int (X_- + Y_-)d(X+Y) - X^2 - Y^2 + 2\int X_- dX + 2\int Y_- dY\},$$

and the result follows from the bilinearity of the stochastic integral $\int H dX$ in (H,X). □

(6.6) Corollary. *All semimartingales on a given filtered probability space form an algebra.*

Proof. We have already seen that the stochastic integral is a semimartingale (5.10) and that the bracket process $[X,Y]$ is a semimartingale (6.4). The result then follows from the integration by parts formula, since semimartingales clearly form a vector space. □

A theorem analogous to Theorem (6.2) holds for $[X,Y]$ as well as $[X,X]$. Its proof is also analogous, so we give the theorem without proof.

(6.7) Theorem. *Let X and Y be two semimartingales. Then the bracket process $[X,Y]$ satisfies:*

(i) $[X,Y]_0 = X_0Y_0$; $\Delta[X,Y] = \Delta X \Delta Y$;

(ii) *If σ_n is a sequence of random partitions tending to the identity, then*

$$[X,Y] = X_0Y_0 + \lim \sum_i (X^{T_{i+1}^n} - X^{T_i^n})(Y^{T_{i+1}^n} - Y^{T_i^n}),$$

where convergence is in ucp, *and where σ_n is the sequence $0 = T_0^n \leq T_1^n \leq \cdots \leq T_i^n \leq \cdots$ with T_i^n stopping times.*

We next record a real analysis theorem from the Lebesgue-Stieltjes theory of integration. We do not give its proof.*

(6.8) Theorem. *Let α, β, γ be functions mapping $[0,\infty[$ to $\mathbb{R}$ with $\alpha(0) = \beta(0) = \gamma(0) = 0$; Suppose α, β, γ are all right continuous, α is of finite variation, and β and γ are each increasing. Suppose further that for all s,t with $s \leq t$, we have*

$$\left|\int_s^t d\alpha_u\right| \leq \left(\int_s^t d\beta_u\right)^{\frac{1}{2}} \left(\int_s^t d\gamma_u\right)^{\frac{1}{2}}.$$

Then for any measurable functions f,g we have

$$\int_s^t |fg|\,|d\alpha| \leq \left(\int f^2 d\beta\right)^{\frac{1}{2}} \left(\int g^2 d\gamma\right)^{\frac{1}{2}}.$$

In particular, the measure $d\alpha$ is absolutely continuous with respect to both $d\beta$ and

---

*For a proof, the reader can consult, e.g., [26, p. 263].

$d\gamma$.

Note that $|d\alpha|$ denotes the total variation measure corresponding to the measure $d\alpha$, the Lebesgue-Stieltjes signed measure induced by α. We use this theorem to prove an important inequality concerning the quadratic variation and bracket processes.

(6.9) Theorem (The Kunita-Watanabe Inequality). Let X and Y be two semimartingales, and let H and K be two measurable processes. Then one has a.s.

$$\int_0^\infty |H_s||K_s||d[X,Y]_s| \leq \left(\int_0^\infty H_s^2 d[X,X]_s\right)^{\frac{1}{2}} \left(\int_0^\infty K_s^2 d[Y,Y]_s\right)^{\frac{1}{2}}$$

Proof. By Theorem (6.8) we only need to show that there exists a null set N, such that for $\omega \notin N$, and (s,t) with $s \leq t$, we have:

$$(*) \quad \left|\int_s^t d[X,Y]_u\right| \leq \left(\int_s^t d[X,X]_u\right)^{\frac{1}{2}} \left(\int_s^t d[Y,Y]_u\right)^{\frac{1}{2}} .$$

Let N be the null set such that if $\omega \notin N$, then $0 \leq \int_s^t d[X+rY, X+rY]_u$, for every r, s, t; $s \leq t$, with r, s, t all rational numbers. Then

$$0 \leq [X+rY, X+rY]_t - [X+rY, X+rY]_s$$

$$= r^2([Y,Y]_t - [Y,Y]_s) + 2r([X,Y]_t - [X,Y]_s) + ([X,X]_t - [X,X]_s) .$$

The right side being positive for all rational r, it must be positive for all real r by continuity. Thus the discriminant of this quadratic equation in r must be nonnegative, which gives us exactly the inequality (*). Since we have, then, the inequality for all rational (s,t), it must hold for all real (s,t), by the right continuity of the paths of the processes. □

By studying the process [X,X], we will prove some pretty results from the martingale theory, which are especially simple in our context [(6.12)-(6.14)]. Since [X,X] (recall X is a semimartingale) has right continuous paths, and since $\Delta[X,X] = (\Delta X)^2$, we can decompose [X,X] path by path into its continuous part and its pure jump part. We write this as:

$$[X,X]_t = X_0^2 + [X,X]_t^c + \sum_{s\le t} (\Delta X_s)^2 .$$

(6.10) Definition. A semimartingale X will be called quadratic pure jump if $[X,X]^c = 0$. Of course, if X is quadratic pure jump, then $[X,X]_t = X_0^2 + \sum_{s\le t}(\Delta X_s)^2$.

(6.11) Theorem. *If X is adapted, cadlag, with paths of finite variation on compacts, then X is a quadratic pure jump semimartingale.*

Proof. We have already seen that such an X is a semimartingale (3.1), and that the stochastic integral with respect to X is nothing more than a pathwise Lebesgue-Stieltjes integral (5.7). The integration by parts formula for Lebesgue-Stieltjes differentials applied to X times itself yields: $X^2 = \int X_- dX + \int XdX$, computed path by path. The semimartingale integration by parts formula (6.5), on the other hand, yields: $X^2 = 2\int X_- dX + [X,X]$.
Moreover $\int XdX = \int(X_- + \Delta X)dX = \int X_- dX + \int \Delta X dX$, and $\int_0^t X_- dX_s + \int_0^t \Delta X dX_s = \int_0^t X_- dX_s + \sum_{s\le t}(\Delta X_s)^2$.

Thus equating the two formulas, we deduce $[X,X]_t = \sum_{s\le t}(\Delta X_s)^2$, whence the theorem. □

Note in particular that if X is adapted with continuous paths of finite variation, then $[X,X]_t = X_0^2$, all $t \ge 0$.

(6.12) Theorem. *Let X be a local martingale with continuous paths that are not everywhere constant. Then [X,X] is not the constant process X_0^2.*

Proof. Note that a continuous local martingale is a semimartingale (3.4). We have $X^2 - [X,X] = 2\int X_- dX$, and by the martingale preservation property (5.11) we have that $2\int X_- dX$ is a local martingale. Moreover $\Delta 2\int X_- dX = 2(X_-)(\Delta X)$, and since X is continuous, $\Delta X = 0$, and thus $2\int X_- dX$ is a continuous local martingale, hence locally square integrable. Thus $X^2 - [X,X]$ is a locally square integrable martingale.

Suppose that $[X,X]$ were the constant process X_0^2. Then $X^2 - [X,X] + X_0^2 = X^2$ would be a locally square integrable martingale, hence so would $\tilde{X}^2 = X^2 - X_0^2$; by optimal stopping, we may assume that $\tilde{X}^2$ is square integrable (and $\tilde{X}_0 = 0$). Then $E(\tilde{X}_\infty^2) = E(\tilde{X}_0^2) = 0$, which implies $\tilde{X}_\infty = 0$ a.s. Thus $X^2 - X_0^2$ is trivial; that is, $E(X_\infty | \mathcal{F}_t) = X_t = X_0$ is constant, which is a contradiction. □

(6.13) Corollary. Let X be a continuous local martingale. If X has paths of finite variation on a compact set, then those paths are a.s. constant on that set.

Proof. If X has paths of finite variation then we must have $[X,X]$ be constant by Theorem (6.11); then X must be constant by Theorem (6.12). The corollary now follows from (5.9). □

(6.14) Corollary. Let X and Y be two locally square-integrable martingales. Then $[X,Y]$ is the unique adapted cadlag process A with paths of finite variation on compacts satisfying the two properties:

(i) $XY - A$ is a local martingale;

(ii) $\Delta A = \Delta X \Delta Y$.

Proof. Integration by parts yields:

$$XY = \int X_- dY + \int Y_- dX + [X,Y];$$

but the martingale preservation property tells us that both stochastic integrals are local martingales. Thus $XY - [X,Y]$ is a local martingale. Property (ii) is simply an application of Theorem (6.7). Thus it remains to show uniqueness. Suppose A,B both satisfy properties (i) and (ii). Then $A-B = (XY-B) - (XY-A)$, the difference of two local martingales which is again a local martingale. Moreover,

$$\Delta(A-B) = \Delta A - \Delta B = \Delta X \Delta Y - \Delta X \Delta Y = 0.$$

Thus A-B is a continuous local martingale, $A_0 - B_0 = 0$, and it has paths of finite variation on compacts. Thus $A_t - B_t - A_0 - B_0 = 0$ by (6.13) and we have uniqueness. □

[Corollary (6.14) is true as well for X,Y local martingales.]

(6.15) Theorem. Let X be a quadratic pure jump semimartingale. Then for any semimartingale Y we have:

$$[X,Y]_t = X_0Y_0 + \sum_{0<s\leq t} \Delta X_s \Delta Y_s .$$

Proof. The Kunita-Watanabe Theorem (6.9) tells us $d[X,Y]_s$ is a.s. absolutely continuous with respect to $d[X,X]$ (path by path). Thus $[X,X]^c = 0$ implies $[X,Y]^c = 0$, and hence [X,Y] is the sum of its jumps, and the result follows by (6.7). □

(6.16) Theorem. Let X and Y be two semimartingales, and let $H,K \in \mathbb{L}$. Then

$$[H\cdot X,\ K\cdot Y]_t = \int_0^t H_sK_s d[X,Y]_s$$

and, in particular,

$$[H\cdot X,\ H\cdot X]_t = \int_0^t H_s^2 d[X,X]_s$$

Proof. It suffices to establish the following result (*) $[H\cdot X,Y]_t = \int_0^t H_s d[X,Y]_s$, and then apply it again, by the symmetry of the form $[\cdot\ ,\ \cdot\ ]$.

First suppose H is the indicator of a stochastic interval. That is, $H = 1_{]0,T]}$, where T is a stopping time. Establishing (*) is equivalent in this case to showing $[X^T,Y] = [X,Y]^T$, a result that is an obvious consequence of (6.7), which approximates [X,Y] by sums.

Next suppose $H_t = U1_{]S,T]}$, where S,T are stopping times, $S \leq T$ a.s., and $U \in \mathcal{F}_S$. Then $\int H_s dX_s = U(X^T - X^S)$, and it is easy to check that in this case

$$[H\cdot X,Y] = U\{[X^T,Y] - [X^S,Y]\}$$
$$= U\{[X,Y]^T - [X,Y]^S\} = \int H_s d[X,Y]_s .$$

The result now follows for $H \in \underline{\underline{S}}$ by linearity. Finally, suppose $H \in \mathbb{L}$ and let H^n be a sequence in $\underline{\underline{S}}$ converging in ucp to H. Let $Z^n = H^n \cdot X$; $Z = H \cdot X$. We know Z^n, Z are all semimartingales. We have $\int H_s^n d[X,Y]_s = [Z^n,Y]$, since $H^n \in \underline{\underline{S}}$, and using integration parts:

$$[Z^n,Y] = YZ^n - \int Y_- dZ^n - \int Z_-^n dY$$

$$= YZ^n - \int Y_- H^n dX - \int Z_-^n dY.$$

By the definition of the stochastic integral, we know $Z^n \to Z$ in ucp, and since $H^n \to H$(ucp), letting $n \to \infty$ we have

$$\lim_{n\to\infty} [Z^n,Y] = YZ - \int Y_- HdX - \int Z_- dY$$

$$= YZ - \int Y_- dZ - \int Z_- dY$$

$$= [Z,Y],$$

again by integration by parts. Since $\lim_{n\to\infty} \int H_s^n d[X,Y]_s = \int H_s d[X,Y]_s$, we have $[Z,Y] = [H \cdot X,Y] = \int H_s d[X,Y]_s$, and the proof is complete. □

(6.17) Theorem. Let H be a cadlag, adapted process, and let X,Y be two semimartingales. Let σ_n be a sequence of random partitions tending to the identity. Then

$$\sum H_{T_i^n} (X^{T_{i+1}^n} - X^{T_i^n}) (Y^{T_{i+1}^n} - Y^{T_i^n})$$

converges in ucp to $\int H_{s-} d[X,Y]_s$ $(H_{0-} = H_0)$.

Here $\sigma_n = (0 \le T_0^n \le T_1^n \le \dots \le T_i^n \le \dots)$.

Proof. [X,Y] is cadlag, adapted, and has paths of finite variation; hence it is a semimartingale. The theorem then follows as a corollary of (5.13) [approximating the stochastic integral by sums], and (6.7) [approximating the quadratic variation process [X,Y] by sums]. □

EXAMPLE. Let W_t be a Wiener process (i.e., Brownian motion). It is easy to check That $W_t^2 - t$ is a continuous martingale. Thus by (6.14) we have $[W,W]_t = t$, hence $[H\cdot W,H\cdot W]_t = \int_0^t H_s^2\,ds$. By the martingale preservation property, $\int H_s dW_s$ is also a continuous local martingale, with $H\cdot W_0 = 0$. Using approximations by sums it is easy to check that $E(\int_0^t H_s dW_s) = 0$, $0 \le t \le \infty$, and $E\{(\int_0^t H_s dW_s)^2\} = E\{[H\cdot W,H\cdot W]_t\}$

$$= E\{\int H_s^2 ds\}.$$

It was this last equality: $E\{(\int_0^t H_s dW_s)^2\} = E\{\int_0^t H_s^2 ds\}$ that was crucial in K. Itô's original treatment of a stochastic integral.

7. Itô's Formula; Change of Variables.

Let V be a process with continuous paths of finite variation on compacts. Then if $f \in \mathcal{C}^1$ (possessing at least one continuous derivative) it is well known that f(V) is again a process with paths of finite variation. Moreover, the formula $f(V_t) = f(V_0) + \int_0^t f'(V_s)dV_s$ holds (path by path). This is usually called the change of variables formula. Less well known, but an easy extension of the above, is to the case where V is only right continuous, but still has paths of finite variation on compacts. The change of variables formula then takes the form:

$$f(V_t) - f(V_0) = \int_0^t f'(V_{s-})dV_s + \sum_{0<s\le t} \{f(V_s) - f(V_{s-}) - f'(V_{s-})\Delta V_s\} .$$

We wish to establish a formula analogous to the above, but for the stochastic integral; that is, when the process is a semimartingale. The formula is different in this case: one must add an extra term!

(7.1) Theorem (Itô's Formula). Let X be a semimartingale and let f be a $\mathcal{C}^2$ real function. Then f(X) is again a semimartingale, and the following formula holds:

(7.2) $f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X,X]^c_s$

$$+ \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\}$$

Proof. First note that the jump part of the stochastic integral $\int f''(X_{s-})d[X,X]_s$ is given by $\sum_{s\leq t} f''(X_{s-})(\Delta X_s)^2$, and this is a convergent series. By adding and subtracting $\frac{1}{2}$ of this series, we can rewrite Itô's formula in the equivalent form:

(7.3) $f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X,X]_s$

$$+ \sum_{s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2} f''(X_{s-})(\Delta X_s)^2\}$$

which is perhaps less obviously a generalization of the "classical" case, but notationally simpler to prove. The proof rests, of course, on Taylor's theorem:

$$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2} f''(x)(y-x)^2 + R(x,y)$$

where $R(x,y) \leq r(|y-x|)(y-x)^2$, such that $r: \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function with $\lim_{u\downarrow 0} r(u) = 0$.

Proof for the continuous case: We first restrict our attention to a continuous semimartingale X, since the proof is less complicated but nevertheless gives the basic idea. We fix a $t > 0$, and let σ_n be a sequence of random partitions of $[0,t]$ tending to the identity $[\sigma_n = (0 = T_0^n \leq T_1^n \leq \dots \leq T_{k_n}^n = t)]$.

Then

(7.4) $f(X_t) - f(X_0) = \sum_{i=0}^{k_n} \{f(X_{T_{i+1}^n}) - f(X_{T_i^n})\} = \sum_i f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})$

$$+ \frac{1}{2}\sum_i f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2 + \sum_i R(X_{T_i^n}, X_{T_{i+1}^n})$$

The first sum converges in probability to the stochastic integral $\int_0^t f'(X_{s-})\,dX_s$

(5.13); the second sum converges in probability to $\frac{1}{2}\int_0^t f''(X_s)d[X,X]_s$ (6.2). It remains to consider the third sum: $\sum_i R(X_{T_i^n}, X_{T_{i+1}^n})$. But this sum is majorized, in absolute value, by $\sup_i r(|X_{T_{i+1}^n} - X_{T_i^n}|)\{\sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2\}$, and since $\sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2$ converges in probability to $[X,X]_t$ (6.2), the last term will tend to 0 if $\lim_{n\to\infty}\sup_i r(|X_{T_{i+1}^n} - X_{T_i^n}|) = 0$. However $s \to X_s(\omega)$ is a continuous function on $[0,t]$, each fixed ω, and hence uniformly continuous. Since $\lim_{n\to\infty}\sup_i |T_{i+1}^n - T_i^n| = 0$ by hypothesis, we have the result. Thus, in the continuous case, $f(X_t) - f(X_0) = \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X,X_s]$, for each t, a.s. The continuity of the paths then permits us to remove the dependence of the null set on t, giving the complete result in the continuous case.

<u>Proof for the general case:</u> X is now given as a right continuous semimartingale. Once again we have a representation as in (7.4), but we need a closer analysis. For any $t > 0$ we have $\sum_{0<s\le t} (\Delta X_s)^2 \le [X,X]_t < \infty$ a.s., hence $\sum_{0\le s\le t} (\Delta X_s)^2$ is convergent. Given $\varepsilon > 0$ let A be a subset of $\mathbb{R}_+ \times \Omega$ such that $\sum_{s\in A} (\Delta X_s)^2 \le \varepsilon^2$, and let $B = \{(s,\omega): (\Delta X_s)^2 > 0, (s,\omega) \notin A\}$. Then we can rewrite (7.4) as follows:

(7.5) $$f(X_t) - f(X_0) = \sum_{i=0}^{k_n}\{f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})\} + \frac{1}{2}\{\sum_i f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2\}$$

$$+\sum_i 1_{\{B\cap]T_i^n,T_{i+1}^n]\neq\phi\}}\{f(X_{T_{i+1}^n}) - f(X_{T_i^n}) - f'(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})$$

$$- \frac{1}{2} f''(X_{T_i^n})(X_{T_{i+1}^n} - X_{T_i^n})^2\}$$

$$+ \sum_i 1_{\{B \cap]T_i^n, T_{i+1}^n] = \phi\}} R(X_{T_i^n}, X_{T_{i+1}^n}) .$$

As in the continuous case, the first two sums on the right side of (7.5) converge respectively to $\int_0^t f'(X_{s-})dX_s$ and $\frac{1}{2}\int_0^t f''(X_{s-})d[X,X]_s$ by (5.13) and (6.2). The third sum converges to:

$$(7.6) \quad \sum_{\substack{s \in B \\ |\Delta X_s| > 0}} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\, \Delta X_s - \frac{1}{2} f''(X_{s-})(\Delta X_s)^2\},$$

and it remains to consider the fourth and last sum on the right side of equation (7.5). Since $\limsup_n |T_{i+1}^n - T_i^n| = 0$, for large enough n we have $|X_{T_{i+1}^n} - X_{T_i^n}| \leq 2\varepsilon$ when $B \cap]T_i^n, T_{i+1}^n] = \phi$. But then $R(x,y) \leq r(|y-x|)(y-x)^2$, hence we can majorize

$$(7.7) \quad \sum_i 1_{\{B \cap]T_i^n, T_{i+1}^n] = \phi\}} R(X_{T_i^n}, X_{T_{i+1}^n})$$

by $r(2\varepsilon) \sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2$; since $\sum_i (X_{T_{i+1}^n} - X_{T_i^n})^2$ converges to $[X,X]_t < \infty$, as n tends to ∞ and as ε tends to 0, we have that $r(2\varepsilon)$ tends to 0 and thus the sums (7.7) tend to 0 with ε. Moreover the sums (7.6) clearly tend, as ε tends to 0, to:

$$(7.8) \quad \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2} f''(X_{s-})(\Delta X_s)^2\}$$

provided this series is absolutely convergent.

Let $V_k = \inf \{t > 0: |X_t| \geq k\}$, with $X_0 = 0$. By first establishing (7.3) for $X1_{[0,V_k[}$, which is a semimartingale since it is the product of two semimartingales (6.6), it suffices to consider semimartingales taking their values in intervals of the form $[-k,k]$. For f restricted to $[-k,k]$ we have $|f(y) - f(x) - (y-x)f'(x)| \leq C(y-x)^2$. Then

$$\sum_{0<s\leq t} |f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s| \leq C \sum_{0<s\leq t} (\Delta X_s)^2 \leq C[X,X]_t < \infty ,$$

and

$$\sum_{0<s\leq t} |f''(X_{s-})|(\Delta X_s)^2 \leq d \sum_{0<s\leq t} (\Delta X_s)^2 \leq d[X,X]_t < \infty \text{ a.s.}$$ Thus the sum (7.8) is absolutely convergent and this completes the proof. □

Theorem (7.1) has a multidimensional analog. We omit the proof.

(7.9) Theorem. Let $\underset{\sim}{X} = (X^1,\dots,X^n)$ be an n-tuple of semimartingales, and let $f: \mathbb{R}^n \to \mathbb{R}$ have second order continuous partial derivatives. Then $f(\underset{\sim}{X})$ is a semimartingale and the following formula holds:

$$f(\underset{\sim}{X}_t) - f(\underset{\sim}{X}_0) = \sum_{i=1}^{n} \int_0^t \frac{\partial f}{\partial x_i}(\underset{\sim}{X}_{s-})dX_s^i + \frac{1}{2} \sum_{1\leq i,j\leq n} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(\underset{\sim}{X}_{s-})d[X^i,X^j]_s^c$$

$$+ \sum_{0<s\leq t} \{f(\underset{\sim}{X}_s) - f(\underset{\sim}{X}_{s-}) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underset{\sim}{X}_{s-})\Delta X_s^i\} .$$

The stochastic integral calculus, as revealed by Theorems (7.1) and (7.9), is different from the classical Lebesgue-Stieltjes calculus. By restricting the class of integrands to semimartingales made left continuous (instead of $\mathbb{L}$), one can define a stochastic integral that obeys the traditional rules of the Lebesgue-Stieltjes calculus.

(7.10) Definition. Let X,Y be semimartingales. Define the Fisk-Stratonovich integral of Y with respect to X, denoted $\int_0^t Y_{s-} \circ dX_s$, by:

$$\int_0^t Y_{s-} \circ dX_s \equiv \int_0^t Y_{s-}\, dX_s + \frac{1}{2} [Y,X]_t^c .$$

Note that we have defined the Fisk-Stratonovich integral in terms of the semimartingale integral. With some work one can slightly enlarge the domain of the definition (cf. [34 , p. 360]).

(7.11) Theorem. Let X be a semimartingale and let f be $\mathcal{C}^3$. Then

$$f(X_t) - f(X_0) = \int_0^t f'(X_{s-}) \circ dX_s + \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\} .$$

Proof. Note that f' is $\mathcal{C}^2$, so that f'(X) is a semimartingale by (7.1) and in the domain of the F - S integral. By (7.1) and the definition, it suffices to establish $\frac{1}{2}[f'(X),X]^c = \frac{1}{2}\int_0^t f''(X_{s-})d[X,X]_s^c$. However

$$f'(X_t) - f'(X_0) = \int_0^t f''(X_{s-})dX_s + \frac{1}{2}\int_0^t f^{(3)}(X_{s-})d[X,X]_s .$$

Thus

(7.12) $[f'(X),X]^c = [f''(X_-)\cdot X,X]^c + [\frac{1}{2} f^{(3)}(X_-)\cdot[X,X],X]^c$.

The first term on the right side of (7.12) is $\int_0^t f''(X_{s-})d[X,X]_s^c$ by (6.16); the second term can easily be seen, as a consequence of (6.2) and the fact that [X,X] has paths of finite variation, to be $(\sum_{0<s\leq\cdot} f^{(3)}(X_{s-})(\Delta X_s)^3)^c$; that is, zero, and the theorem is proved. □

Note that if X is a semimartingale with continuous paths, then Theorem (7.11) reduces to the classical Riemann-Stieltjes formula: $f(X_t) - f(X_0) = \int_0^t f'(X_s) \circ dX_s$; this is, of course, the main attraction of the Fisk-Stratonovich integral.

As an application of the change of variables formula, we next present an investigation of a simple, yet important and non-trivial, stochastic differential equation. We treat it, of course, in integral form.

(7.13) Theorem. Let X be a semimartingale. Then there exists a (unique) semimartingale Z that satisfies the equation: $Z_t = 1 + \int_0^t Z_{s-}dX_s$; Z is given by

$$Z_t = \exp\left(X_t - \frac{1}{2}[X,X]_t\right) \prod_{0<s\leq t} (1+\Delta X_s) \exp\left(-\Delta X_s + \frac{1}{2}(\Delta X_s)^2\right)$$

where the infinite product converges.

Proof. We will not prove the uniqueness here, since it is a trivial consequence of the general theory to be established later. Note that the formula for Z_t is equivalent to the formula:

$$Z_t = \exp\left(X_t - \frac{1}{2}[X,X]_t^c\right) \prod_{s \leq t} (1+\Delta X_s) \exp(-\Delta X_s),$$

and since $X_t - \frac{1}{2}[X,X]_t^c$ is a semimartingale, exp (x) is $\mathcal{C}^2$, we need only show that $\prod_{s \leq t} (1+\Delta X_s) \exp(-\Delta X_s)$ is cadlag, adapted, and of finite variation and it will be a semimartingale, too; thus Z will be a semimartingale. The product is clearly cadlag, adapted; it thus suffices to show the product converges and is of finite variation.

Since X has cadlag paths, there are only a finite number of s such that $|\Delta X_s| \geq 1/2$ on each compact interval (fixed ω). Thus it suffices to show $V_t = \prod_{0<s\leq t} (1+\Delta X_s 1_{\{|\Delta X_s|<1/2\}}) \exp(-\Delta X_s 1_{\{|\Delta X_s|<1/2\}})$ converges and is of finite variation. Let $U_s = \Delta X_s 1_{\{|\Delta X_s| \leq 1/2\}}$. Then we have $\log V_t = \sum_{s \leq t} \{\log(1+U_s) - U_s\}$, which is an absolutely convergent series a.s., since $\sum_{0<s\leq t} (U_s)^2 \leq [X,X]_t < \infty$ a.s., because $\log(1+x) - x \leq x^2$ when $|x| < 1/2$. Thus $\log(V_t)$ is a process with paths of finite variation, and hence so also is $\exp(\log V_t) = V_t$.

To show that Z is a solution, we set $K_t = X_t - \frac{1}{2}[X,X]_t^c$, and let $f(x,y) = ye^x$. Then $Z_t = f(K_t, S_t)$, where $S_t = \prod_{0\leq s\leq t} (1+\Delta X_s) \exp(-\Delta X_s)$.

By the change of variables formula we have

$$(7.14)\quad Z_t - 1 = \int_0^t Z_{s-}\, dK_s + \int_0^t e^{K_{s-}}\, dS_s + \frac{1}{2}\int_0^t Z_{s-}\, d[K,K]_s^c$$

$$+ \sum_{0<s\leq t} (Z_s - Z_{s-} - Z_{s-}\Delta K_s - e^{K_{s-}}\Delta S_s)$$

$$(7.15)\quad = \int_0^t Z_{s-}\, dX_s - \frac{1}{2}\int_0^t Z_{s-}\, d[X,X]_s^c + \int_0^t e^{K_{s-}} dS_s + \frac{1}{2}\int_0^t Z_{s-}\, d[X,X]_s^c$$

$$+ \sum_{0\leq s\leq t} (Z_s - Z_{s-} - Z_{s-}\Delta K_s - e^{K_{s-}}\Delta S_s) \ .$$

Note that S, being the exponential of a pure jump process, is again a pure jump process; hence $\int_0^t e^{K_{s-}} dS_s = \sum_{0\leq s\leq t} e^{K_{s-}}\Delta S_s$; also $Z_s = Z_{s-}(1+\Delta X_s)$, and $Z_{s-}\Delta K_s = Z_{s-}\Delta X_s$, so the last sum on the right side of equation (7.15) becomes:

$$\sum_{0<s\leq t} (Z_{s-}(1+\Delta X_s) - Z_{s-} - Z_{s-}\Delta X_s - e^{K_{s-}}\Delta S_s) = \sum_{0\leq s\leq t} - e^{K_{s-}}\Delta S_s \ .$$

Thus equation (7.15) simplifies due to cancellation:

$$Z_t - 1 = \int_0^t Z_{s-} dX_s \ ,$$ and we have the result. □

(7.16) <u>Definition</u>. For a semimartingale X, <u>the stochastic exponential of</u> X, written $\mathcal{E}(X)$, is the (unique) semimartingale Z that is a solution of: $Z_t = 1 + \int_0^t Z_{s-}\, dX_s$.

Of course, the previous theorem gives us $\mathcal{E}(X)$ in closed form. We also have the following pretty result.

(7.17) <u>Theorem.</u> <u>Let</u> X <u>and</u> Y <u>be two semimartingales</u>. <u>Then</u> $\mathcal{E}(X)\,\mathcal{E}(Y) = \mathcal{E}(X + Y + [X,Y])$.

<u>Proof</u>. Let $U_t = \mathcal{E}(X)_t$ and $V_t = \mathcal{E}(Y)_t$. Then the integration by parts formula gives that $U_tV_t - 1 = \int_0^t U_{s-}dV_s + \int_0^t V_{s-}dU_s + [U,V]_t$.

Using that U and V are exponentials, this is equivalent to:

$$= \int_0^t U_{s-}V_{s-}\, dY_s + \int_0^t U_{s-}V_{s-} dX_s + \int_0^t U_{s-}V_{s-} d[X,Y]_s \;;$$

letting $W_t = U_t V_t$, we deduce:

$$W_t = 1 + \int_0^t W_{s-} d(X + Y + [X,Y])_s,$$ and so $W_t = \mathcal{E}(X + Y + [X,Y])$, which was to be shown. □

IV. SEMIMARTINGALES AND DECOMPOSABLE PROCESSES

1. Introduction

The usual definition of a semimartingale as found, for example, in any of [11], [17], [23], [38], or [34] is that of a process X that can be decomposed $X = M+A$ where M is a local martingale and A is a right continuous, adapted process with paths of finite variation on compacts. This differs slightly (although it is equivalent) to our definition of a decomposable process (III .3.5) which we repeat here:

(1.1) Definition. An adapted process X with right continuous paths is decomposable if it can be decomposed $X_t = X_0 + M_t + A_t$, where $M_0 = A_0 = 0$, M is a locally square integrable local martingale, and A is an adapted process having right continuous paths of finite variation on compacts.

The primary purpose of this chapter is to show that a process is a semimartingale if and only if it is a decomposable process. This will allow us to extend the stochastic integral to more general integrands and will be crucial in putting a useful topology on the space of semimartingales. (Note that we have already seen (III.3.7) that a decomposable process is a semimartingale). We will show in passing the little known result that a process satisfying the traditional definition of a semimartingale is, in fact, decomposable. Indeed, the results of this chapter that are essential to the rest of these notes can be summarized by the following two theorems, whose proofs are contained within paragraphs two through five.

(1.2) Theorem. Let X be a cadlag, adapted process. The following are equivalent:

(i) X is a semimartingale (in the sense of III.1.1);

(ii) X has a decomposition $X = X_0 + M + A$ where M is a local martingale and A has paths of finite variation on compacts;

(iii) for any $\varepsilon > 0$, X has a decomposition $X = X_0 + N + B$ where N is a local martingale with jumps bounded by ε, and B has paths of finite variation on compacts;

(iv) X is decomposable.

(1.3) Theorem. Let X be a semimartingale. If X has a decomposition $X = X_0 + M + A$ with A natural, then such a decomposition is unique.

In Theorem (1.2), that (iii) implies (iv) is elementary. That (ii) implies (iii) is Corollary (4.7); that (iv) implies (i) is III.3.6. Thus the heart of the theorem is to show (i) $\rightarrow$ (ii) , which is essentially the theorem of K. Bichteler and C. Dellacherie, itself using the Doob-Meyer decomposition theorem, Rao's theorem on quasimartingales, and the Girsanov-Lenglart theorem on changes of probability laws.

We have tried to present this succession of deep theorems in the most direct and elementary manner possible. This contrasts with their usual treatment, which is customarily part of a general pedagogic presentation of the "general theory of processes". In particular, we have systematically avoided presenting "the classification of stopping times". We believe this will keep the treatment on as brief and technically simple a level as is possible; the cost comes in that we were unable to give a complete proof of Lemma (4.5), although we believe it is fairly intuitive nevertheless, and proofs of the facts used can easily be found in the literature. The reader who is willing to stipulate the truth of Theorems (1.2) and (1.3) could skip this chapter without loss for the subsequent chapters.

2. The Doob-Meyer Decompositions

We begin with a definition.

(2.1) Definition. A right continuous process X is a potential if it is a nonnegative supermartingale such that $\lim_{t\to\infty} E(X_t) = 0$.

The next theorem is known as the Doob decomposition [16].

(2.2) Theorem. A potential $(X_n)_{n\in\mathbb{N}}$ has a decomposition $X_n = M_n - A_n$, where $A_{n+1} \geq A_n$ a.s., $A_0 = 0$, $A_n \in \mathcal{F}_{n-1}$, and $M_n = E(A_\infty|\mathcal{F}_n)$. Such a decomposition is unique.

Proof. Let $M_0 = X_0$ and $A_0 = 0$. Define $M_1 = M_0+(X_1 - E(X_1|\mathcal{F}_0))$; $A_1 = X_0 - E(X_1|\mathcal{F}_0)$. Define M_n, A_n inductively as follows:

$$M_n = M_{n-1} + (X_n - E(X_n|\mathcal{F}_{n-1}))$$
$$A_n = A_{n-1} + (X_{n-1} - E(X_n|\mathcal{F}_{n-1})).$$

It is then simple to check that M_n and A_n so defined satisfy the hypotheses.

Next suppose $X_n = N_n - B_n$ is another such representation. Then $M_n - N_n = A_n - B_n$ and in particular $M_1 - N_1 = A_1 - B_1 \in \mathcal{F}_0$; thus $M_1 - N_1 = E(M_1 - N_1|\mathcal{F}_0) = M_0 - N_0 = X_0 - X_0 = 0$, hence $M_1 = N_1$. Continuing inductively shows $M_n = N_n$, all n. □

We next wish to extend Theorem (2.2) to continuous time potentials. First note that if $E(A_\infty) < \infty$ in (2.2), then it is simple to check that $A_{n+1} \in \mathcal{F}_n$ if and only if for all bounded martingales Y we have

$$E \sum_{k=1}^{\infty} Y_{k-1}(A_k - A_{k-1}) = E\{Y_\infty A_\infty\} .$$

The analogous condition in the continuous parameter case would be that if Y is a bounded positive martingale, then

$$E\{\int_0^\infty Y_{s-} dA_s\} = E\{Y_\infty A_\infty\}.$$

(2.3) Definition. Let A be an adapted, right continuous, increasing process such that $A_0 = 0$ with $E(A_\infty) < \infty$. Then A is natural if $E\{\int_0^\infty Y_{s-} dA_s\} = E\{Y_\infty A_\infty\}$ for any bounded, positive martingale Y.

The next theorem is known as the Riesz decomposition, due to its potential theory analogue.

(2.4) Theorem. Let $(X_t)_{t \in \mathbb{R}_+}$ be a positive supermartingale. Then there exists a unique decomposition of X into a martingale and a potential Z such that X = M + Z.

Proof. As is well known $\lim_{t\to\infty} X_t = Y$ exists a.s., and moreover $Y \in L^1$. Let $M_t = E(Y|\mathfrak{F}_t)$, and let $Z_t = X_t - M_t$. One easily verifies that Z is a potential.

Let X = N + W be another such decomposition. Then N - M = Z - W is a martingale with $Z_\infty - W_\infty = 0$. Therefore Z = W and N = M. □

The next theorem is our first version of the Doob-Meyer decomposition theorem.

(2.5) Theorem. Let X be a potential such that the collection $\mathcal{H} = \{X_T;$ T a stopping time$\}$ is uniformly integrable. Then X has a decomposition X = M - A, where M is a martingale and A is a right continuous, increasing process, $A_0 = 0$, and A is natural. Such a decomposition is unique.

Proof. For each $n \in \mathbb{N}$, define $Y_i = (X_{i/2^n})_{i\in\mathbb{N}}$. Each process $(Y_i)_{i\in\mathbb{N}}$ is a discrete potential. By the Doob decomposition (2.2), there exists A_i^n such that $Y_i = E\{A_\infty^n|\mathfrak{F}_{i/2^n}\} - A^n_{i/2^n}$ where $A^n_{i/2^n} \in \mathfrak{F}_{i-1/2^n}$ and $A^n_\infty = \lim_{i\to\infty} A^n_{i/2^n}$. Suppose we know $(A^n_\infty)_{n\in\mathbb{N}}$ is a uniformly integrable collection. Then by the Dunford-Pettis theorem there exists a r.v. A_∞ and a subsequence n_k such that $A_\infty^{n_k}$ tends to A_∞ in $\sigma(L^1, L^\infty)$. Let M_t be the right continuous version of $E(A_\infty|\mathfrak{F}_t)$. Then for $r \le s$ dyadic rationals $A_r^n \le A_s^n$ a.s., hence $E(A_\infty^n|\mathfrak{F}_r) - X_r \le E(A_\infty^n|\mathfrak{F}_s) - X_s$, and it follows that $M_r - X_r \le M_s - X_s$ a.s. Therefore $A_t = M_t - X_t$ is right continuous and a.s. increasing on the dyadic rationals; hence we can take A right continuous, everywhere increasing, and $\lim_{t\to\infty} A_t = A_\infty$, since $\lim_{t\to\infty} X_t = 0$.

Now let N be a bounded positive martingale. By the dominated convergence theorem we have

$$E\{\int_0^\infty N_{s-}dA_s\} = \lim \sum_{i=0}^\infty E\{N_{i/2^n}(A_{i+1/2^n} - A_{i/2^n})\}$$

and since $N_{i/2^n} \in \mathfrak{F}_{i/2^n}$, this becomes:

$$= \lim \sum_i E\{N_{i/2^n} E(A_{i+1/2^n} - A_{i/2^n} | \mathfrak{F}_{i/2^n})\}$$

$$= \lim \sum_i E\{N_{i/2^n} E(X_{i+1/2^n} - X_{i/2^n} | \mathfrak{F}_{i/2^n})\}$$

$$= \lim \sum_i E\{N_{i/2^n}(A^n_{i+1/2^n} - A^n_{i/2^n})\} .$$

Moreover $E(N_{i/2^n} A^n_{i+1/2^n}) = E(N_{i+1/2^n} A^n_{i+1/2^n})$, hence the above becomes:

$$= \lim \sum_i E\{N_{i+1/2^n} A^n_{i+1/2^n} - N_{i/2^n} A^n_{i/2^n}\}$$ and we have a telescoping sum. Thus:

$$= \lim E(A^n_\infty N_\infty),$$

and taking the limit along the subsequence n_k shows that the limit is $E(A_\infty N_\infty)$, and therefore A is natural.

As for uniqueness, let $X = L + B$ be another such decomposition. We have seen that the potential X determines A^n_∞ and that this converges weakly to A_∞; by the uniqueness of weak limits we have $A_\infty = B_\infty$. But then since $X_t = E(A_\infty | \mathfrak{F}_t) - A_t$ $= E(B_\infty | \mathfrak{F}_t) - B_t = E(B_\infty | \mathfrak{F}_t) - A_t$, we have that $L = M$ and $A = B$.

It remains only to show that the collection (A^n_∞) is uniformly integrable. For each $\lambda > 0$ define the stopping-times

$$T^\lambda_n = \inf\{i/2^n: A^n_{i+1/2^n} > \lambda\} .$$

Then $A^n_\infty > \lambda$ if and only if $T^\lambda_n < \infty$. Moreover $X_{T^\lambda_n} = E(A^n_\infty | \mathfrak{F}_{T^\lambda_n}) - A^n_{T^\lambda_n}$.
Therefore

$$(2.6)\quad E\{A_\infty^n 1_{(A_\infty^n>\lambda)}\} = E\{A^n_{T_n^\lambda} 1_{(T_n^\lambda<\infty)}\} + E\{X_{T_n^\lambda} 1_{(T_n^\lambda<\infty)}\}$$

$$\leq \lambda P(A_\infty^n > \lambda) + E\{X_{T_n^\lambda} 1_{(T_n^\lambda<\infty)}\} ,$$

and

$$E\{(A_\infty^n - \lambda)1_{(A_\infty^n \geq 2\lambda)}\} \leq E\{X^n_{T_n^\lambda} 1_{(T_\lambda^n<\infty)}\}$$

which implies

$$2\lambda P(A_\infty^n > 2\lambda) \leq 2E\{X_{T_n^\lambda} 1_{(T_n^\lambda<\infty)}\} .$$

Replacing λ by 2λ in (2.6) yields

$$E\{A_\infty^n 1_{(A_\infty^n>2\lambda)}\} \leq 2\lambda P(A_\infty^n>2\lambda) + E\{X_{T_n^{2\lambda}} 1_{(T_n^{2\lambda}<\infty)}\}$$

$$\leq 2E(X_{T_n^\lambda} 1_{(T_n^\lambda<\infty)}) + E(X_{T_n^{2\lambda}} 1_{(T_n^{2\lambda}<\infty)})$$

and this tends to 0 uniformly in n as λ tends to ∞ by the hypothesis that $\mathcal{H} = \{X_T;\ T$ a stopping time$\}$ is uniformly integrable. □

(2.7) Corollary. Let X be a positive supermartingale, and suppose $\{X_T;\ T$ a stopping time$\}$ is uniformly integrable. Then X has a unique decomposition $X = M - A$ where M is a martingale and A is a right continuous, increasing, natural process with $A_0 = 0$.

Proof. This is a combination of (2.4) and (2.5). □

The next theorem, which can also be considered a Doob-Meyer decomposition theorem, is due to K. Itô and S. Watanabe [22]. It exchanges the uniform integrability for a weakening of the conclusion that M be a martingale to that of M being a local martingale.

(2.8) Theorem. Let X be a nonnegative supermartingale. Then X has a decomposition

$X = M - A$ <u>where</u> M <u>is a local martingale and</u> A <u>is a right continuous, integrable, increasing natural process with</u> $A_0 = 0$. <u>Such a decomposition is unique</u>.

<u>Proof</u>. Since X is a supermartingale, $P(X_t < \infty) = 1$. Define the stopping times $R_n = \inf\{t: X_t \geq n\}$, and let $T_n = R_n \wedge n$. Then T_n increases to ∞ a.s. Let $X_t^n = X_{t\wedge T_n}$; then X^n is a positive supermartingale verifying the hypotheses of (2.7). Let $X^n = M^n - A^n$ be its decomposition, each n. Since M^{n+1} is a uniformly integrable martingale so also is $(M^{n+1}_{t\wedge T_n})_{t\geq 0}$, and the uniqueness of the decompositions shows that $M^{n+1}_{t\wedge T_n} = M^n_{t\wedge T_n}$, and $A^{n+1}_{t\wedge T_n} = A^n_{t\wedge T_n}$. Thus we can define

$$\left.\begin{array}{l} M_t = M_t^n \\ A_t = A_t^n \end{array}\right\} t \leq T_n .$$

Clearly M is a local martingale, and one easily checks that A is natural. □

3. <u>Quasimartingales</u>

Let X be a right continuous, adapted process, defined on $[0,\infty]$. We will call τ a <u>partition</u> of $[0,\infty]$ if τ consists of points $(t_0, t_1, \ldots, t_{n+1})$ such that $0 = t_0 < t_1 < \ldots < t_{n+1} = \infty$, where n is finite. Assume that $X_{t_i} \in L^1$ for each $t_i \in \tau$ and set

$$C(X,\tau) = \sum_{i=0}^{n} |E\{X_{t_i} - X_{t_{i+1}} | \mathcal{F}_{t_i}\}|$$

$$\mathrm{Var}_\tau(X) = E\{C(X,\tau)\} .$$

(3.1) <u>Definition</u>. A right continuous, adapted process X is a <u>quasimartingale</u> on $[0,\infty]$ if $E\{|X_t|\} < \infty$ for each t, and if $\mathrm{Var}(X) = \sup_\tau \mathrm{Var}_\tau(X)$ is finite, where the sup is taken over all finite partitions of $[0,\infty]$.

(3.2) Theorem. Let X be a process indexed by $[0,\infty[$. Then X is a quasimartingale if and only if X has a decomposition X = Y - Z where Y and Z are each positive right continuous supermartingales, where $X_\infty \equiv 0$.

Proof. For given $s \geq 0$, let $\sum(s)$ denote the set of subdivisions of $[s,\infty]$. For each $\tau \in \sum(s)$, set

$$Y_s^\tau = E\{C(X,\tau)^+ \mid \mathfrak{F}_s\}; \; Z_s^\tau = E\{C(X,\tau)^- \mid \mathfrak{F}_s\}$$

where $C(X,\tau)^+$ denotes $\sum_{t_i \in \tau} E\{X_{t_i} - X_{t_{i+1}} \mid \mathfrak{F}_{t_i}\}^+$, and analogously for $C(X,\tau)^-$. Also let $\prec$ denote the ordering of set containment. Suppose $\sigma,\tau \in \sum(s)$ with $\sigma \prec \tau$. We claim $Y_s^\sigma \leq Y_s^\tau$ a.s. To see this let $\sigma = (t_0,\dots,t_n)$, and it suffices to consider what happens upon adding a subdivision point t before t_0, after t_n or between t_i and t_{i+1}. The first two situations being clear, let us consider the third. Set

$$A = E\{X_{t_i} - X_t \mid \mathfrak{F}_{t_i}\} \; ; \; B = E\{X_t - X_{t_{i+1}} \mid \mathfrak{F}_t\}$$

$$C = E\{X_{t_i} - X_{t_{i+1}} \mid \mathfrak{F}_{t_i}\} \; ;$$

then $C = A + E\{B \mid \mathfrak{F}_{t_i}\}$, hence

$$C^+ \leq A^+ + E\{B \mid \mathfrak{F}_{t_i}\}^+$$

$$\leq A^+ + E\{B^+ \mid \mathfrak{F}_{t_i}\}$$, by Jensen's inequality. Therefore

$$E\{C^+ \mid \mathfrak{F}_s\} \leq E\{A^+ \mid \mathfrak{F}_s\} + E\{B^+ \mid \mathfrak{F}_s\}$$

and we conclude $Y_s^\sigma \leq Y_s^\tau$. Since $E(Y_s^\tau)$ is bounded by Var (X), taking limits in L^1 along the directed ordered set $\sum(s)$ we define

$$\hat{Y}_s = \lim_\tau Y_s^\tau \; ,$$

and we can define $\hat{Z}_s$ analogously. Taking a subdivision with $t_0 = s$ and $t_{n+1} = \infty$ we see $Y_s^\tau - Z_s^\tau = E\{C^+ - C^- | \mathcal{F}_s\} = X_s$, and we deduce that $\hat{Y}_s - \hat{Z}_s = X_s$. Moreover if $s < t$ it is easily checked that $\hat{Y}_s \geq E\{\hat{Y}_t | \mathcal{F}_s\}$ and $\hat{Z}_s \geq E\{\hat{Z}_t | \mathcal{F}_s\}$. Define the right continuous processes $Y_t \equiv \hat{Y}_{t+}$, $Z_t \equiv \hat{Z}_{t+}$, with the right limits taken through the rationals. Then Y and Z are positive supermartingales and $Y_s - Z_s = X_s$. The converse is an easy consequence of (2.8). □

(3.3) Theorem. A quasimartingale X has a unique decomposition $X = M + A$, where M is a local martingale and A is a natural process with paths of finite variation on compacts and $A_0 = 0$.

Proof. This theorem is a combination of (3.2) and (2.8). □

4. Special Semimartingales

We begin with a useful theorem that is a slight generalization of III.6.13, since it does not assume path continuity. If A is a process with paths of finite variation on compacts, we let $\int_0^t |dA_s|$ denote the path by path total variation process. We say that A is of integrable total variation if $\int_0^\infty |dA_s| \in L^1$.

(4.1) Theorem. Let A be a right continuous process of integrable total variation which is natural. Suppose also that $A_0 = 0$ and that A is a martingale. Then $A \equiv 0$.

Proof. Let T be a finite stopping time and let H be any bounded, nonnegative martingale. Then $E\{\int_0^T H_{s-} dA_s\} = 0$, as is easily seen by approximating sums and the dominated convergence theorem, since $\int_0^T |dA_s| \in L^1$ and $E(A_T) = 0$. Using the naturality of A, $E\{H_T A_T\} = E\{\int_0^T H_{s-} dA_s\} = 0$, and letting $H_t = E\{1_{\{A_T > 0\}} | \mathcal{F}_t\}$ then shows that $P(A_T > 0) = 0$. Since $E(A_T) = 0$, we conclude $A_t \equiv 0$ a.s., hence $A \equiv 0$. □

(4.2) Definition. Let X be decomposable. If X has a decomposition $X = M + A$ where M is a local martingale and A is a natural process with paths of finite variation on compacts, then X is said to be special.

(4.3) Theorem. If X is a special decomposable process, then its decomposition $X = M + A$ with A natural is unique.

Proof. Let $X = N + B$ be another such decomposition. Then $M - N = B - A$, hence $B - A$ is a local martingale. That $B - A = 0$ and hence $B = A$ is a simple consequence of (4.1). □

(4.4) Definition. If X is a special decomposable process, then the unique decomposition $X = M + A$ with A natural is called the canonical decomposition.

We next present a lemma which is used in the proof of Theorem (4.6), a theorem that P.A. Meyer has dubbed the "fundamental lemma of local martingales". Unfortunately we must use a result from the "general theory of processes" in its proof, thus not making these notes self-contained. We refer the reader to any of the excellent sources [8], [11], [23], or [28] where the details can be found. Also we use an asterisk (*) in the proof to indicate those results taken from the general theory. In the proof we use the notion of predictable stopping time: a stopping time T is predictable if there exists a sequence of stopping times S^n increasing a.s. to T and such that $S^n < T$ on $\{T > 0\}$, each n.

(4.5) Lemma. Let A be a natural process with paths of finite variation on compacts, and $\beta > 0$. Let $T = \inf\{t: |\Delta A_t| \geq \beta\}$. If H is a uniformly integrable martingale, then $D_t = \Delta H_T 1_{\{t \geq T\}}$ is also a martingale.

Proof. The natural process A is predictably measurable (*), and therefore T is a predictable stopping time (*). Let R be any other stopping time. Then

$$E\{D_R\} = E\{\Delta H_T 1_{\{T \geq R\}}\} = E\{E\{\Delta H_T | \mathcal{F}_{T-}\} 1_{\{T \geq R\}}\} ,$$

since $\{T \geq R\} \in \mathcal{F}_{T-}$, as is easily verified, where $\mathcal{F}_{T-} = \mathcal{F}_0 \vee \sigma\{\Lambda \cap \{t < T\}$; all $\Lambda \in \mathcal{F}_t$, $t > 0\}$. Moreover, $E\{\Delta H_T | \mathcal{F}_{T-}\} = 0$, since H is a martingale and T is predictable (*). Therefore $E\{D_R\} = 0$ for all stopping times R. Let $\Gamma \in \mathcal{F}_t$ and set

$$R(\omega) = \begin{cases} t & \omega \in \Gamma \\ \infty & \omega \notin \Gamma \end{cases} .$$

Such an R is a stopping time, and hence $E\{D_t 1_\Gamma\} + E\{D_\infty 1_{\Gamma^c}\} = 0$. Since also $E\{D_\infty 1_\Gamma\} + E\{D_\infty 1_{\Gamma^c}\} = 0$, we conclude $E\{D_t 1_\Gamma\} = E\{D_\infty 1_\Gamma\}$, whence $D_t = E(D_\infty | \mathcal{F}_t)$, and D is a martingale. □

(4.6) <u>Theorem</u>. <u>Let</u> M <u>be a local martingale and let</u> $\beta > 0$ <u>be given</u>. <u>Then</u> M <u>has a decomposition</u> $M = N + A$ <u>where</u> N <u>has jumps bounded by</u> β (i.e. $|\Delta N| \leq \beta$) <u>and</u> A <u>is a local martingale with paths of finite variation on compacts</u>.

<u>Proof</u>. Define the stopping times

$$S_1 = \inf \{t: |\Delta M_t| \geq \beta\}$$

$$\vdots$$

$$S_n = \inf \{t > S_{n-1}: |\Delta M_t| \geq \beta\}$$

and $C_t = \sum_{n=1}^{\infty} \Delta M_{S_n} 1_{\{t \geq S_n\}}$.

Then C_t is a local quasimartingale and hence by (3.3) we can decompose C uniquely as: $C_t = N_t + \tilde{C}_t$ where N is a local martingale and $\tilde{C}_t$ is natural with paths of finite variation on compacts. Let $H_t = M_t - C_t + \tilde{C}_t$, a local martingale. Then $|\Delta H_s| \leq |\Delta M_s - \Delta C_s| + |\Delta\tilde{C}_s| \leq \beta + |\Delta\tilde{C}_s|$.

Next define $U_1 = \inf \{t: |\Delta\tilde{C}_t| \geq \beta\}$

$$\vdots$$

$$U_n = \inf \{t > U_{n-1}: |\Delta\tilde{C}_t| \geq \beta\} .$$

Set $D_t^p = \Delta H_{U_p} 1_{\{t \geq U_p\}}$. Then by (4.5) we have D_t^p is a local martingale. Let

$D_t = \sum_p D_t^p$, and D_t^p is then a local martingale for each p and thus D itself is a local martingale. The decomposition $M = \{H - D\} + \{C - \tilde{C} + D\}$ is then the decomposition sought. □

(4.7) Corollary. Let X be an adapted, cadlag process. If X has a decomposition $X = M + A$ with M a local martingale and A a process with paths of finite variation on compacts, then for any $\beta > 0$ X also has a decomposition $X = N + B$ where N is a local martingale with jumps bounded by β.

(4.8) Corollary. Let X be an adapted, cadlag process. If X has a decomposition $X = M + A$ with M a local martingale and A a process with paths of finite variation on compacts, then X is decomposable.

Proof. Corollary (4.7) is an immediate consequence of (4.6). By (4.7) we can write $X = N + B$, where N is a local martingale with bounded jumps. Let β be a bound for the jumps. We wish to show that N is locally square integrable. It is easy to see that it suffices to show that N is locally bounded. Let $T_n = \inf\{t: |N_t| \geq n\}$. Then $|N_{t \wedge T_n}| \leq n + \beta$, and since T_n is a stopping time, N is locally bounded. □

(4.9) Corollary. If X is an adapted, cadlag process with a decomposition $X = M + A$ where M is a local martingale and A is a process with paths of finite variation on compacts, then X is a semimartingale.

Proof. This is a combination of (4.8) and III.3.6. □

(4.10) Theorem. A quasimartingale is a semimartingale.

Proof. This is a combination of (3.3) and (4.9). □

We conclude this paragraph with a useful extension of III.5.11.

(4.11) Theorem. Let M be a local martingale and let $H \in \mathbb{L}$. Then the stochastic integral $H \cdot M$ is again a local martingale.

Proof. The local martingale M is a semimartingale by (4.9), so $H\cdot M$ is defined. By (4.6) we decompose $M = N + A$ where N is a locally bounded (hence locally square integrable) martingale, and A has paths of finite variation on compacts. Since H is left continuous, by stopping at a time T we may assume, with loss of generality, that H is bounded, M is uniformly integrable, N is bounded, and $\int_0^t |dA_s| \in L^1$, each $t > 0$. We also assume without loss that $M_0 - N_0 = A_0 = 0$. We know $H\cdot N$ is a local martingale by III.5.11, thus we need show only that $H\cdot A$ is a local martingale.

Let σ_n be a sequence of random partitions of $[0,t]$ tending to the identity. Then $\sum H_{T_i^n}(A^{T_i^{n+1}} - A^{T_i^n})$ tends to $(H\cdot A)_t$ in ucp, where σ_n is the sequence $0 = T_0^n \leq T_1^n \leq \ldots \leq T_i^n \leq \ldots$ (III.5.13). Let (n_k) be a subsequence such that the sums converge uniformly a.s. on $[0,t]$.

$$\text{Then } E\{\int_0^t H_u dA_u \mid \mathcal{F}_s\} = E\{\lim_{n_k} \sum_i H_{T_i^{n_k}} (A_t^{T_{i+1}^{n_k}} - A_t^{T_i^{n_k}}) \mid \mathcal{F}_s\}$$

$$= \lim_{n_k} E\{\sum_i H_{T_i^{n_k}} (A_t^{T_{i+1}^{n_k}} - A_t^{T_i^{n_k}}) \mid \mathcal{F}_s\}$$

$$= \lim_{n_k} \{E \sum_i H_{T_i^{n_k}} (A_s^{T_{i+1}^{n_k}} - A_s^{T_i^{n_k}})\}$$

by Lebesgue's dominated convergence theorem. Since the last limit above equals $(H\cdot A)_s$, we conclude that $H\cdot A$ is indeed a local martingale. □

5. A Semimartingale is Decomposable: the Theorem of Bichteler and Dellacherie.

We have seen that a decomposable process is a semimartingale (III.3.6). We have also seen that the "usual" definition of a semimartingale: $X = M + A$ with M a local martingale and A a process with paths of finite variation on compacts, is a decomposable process (4.8), hence also a semimartingale (4.9). In this paragraph we will establish the converse: that a semimartingale is decomposable. It is this theorem that has inspired the entire pedagogic approach presented here. If X is a decomposable process, we will call $X = M + A$ a decomposition of X if M is a local martingale and A is an adapted, right continuous process with paths of finite variation on compacts. We begin with a theorem which is often referred to as Girsanov's theorem.

(5.1) Theorem. Let Q be a probability law equivalent to P (i.e. $Q<<P$ and $P<<Q$). If X is decomposable under P with a decomposition $X = M + A$, then X is decomposable under Q and has a decomposition $X = N + B$, where N is a Q-local martingale given by:

$$N_t = M_t - \int_0^t \frac{1}{Z_s}\, d[Z,M]_s$$

where $Z_t = E_P\{\frac{dQ}{dP}|\mathcal{F}_t\}$, the cadlag version.

Proof. Note that we know X is a P-semimartingale, and thus III.2.2 implies that X is a Q-semimartingale.

We first observe that a process Y is a Q-local martingale if and only if YZ is a P-local martingale. Since M and Z are both P-local martingales, we know they are P-semimartingales by (4.9), hence $\int Z_-\, dM + \int M_- dZ$ is a local martingale by (4.11). Therefore using the integration by parts formula (III.6.5) we have that

(5.2) $ZM - [Z,M] = \int Z_-\, dM + \int M_-\, dZ$

is a P-local martingale. Since Z is a version of $E_P\{\frac{dQ}{dP}|\mathcal{F}_t\}$, we have $1/Z$ is a cadlag version of $E_Q\{\frac{dP}{dQ}|\mathcal{F}_t\}$, and hence 1/Z is a Q-martingale. Multiplying by 1/Z in (5.2) yields:

$$(5.3)\quad \frac{1}{Z}\{ZM - [Z,M]\} = M - (\frac{1}{Z})[Z,M]$$

and since Z times the right side of (5.3) is a P-local martingale, we deduce $M - (\frac{1}{Z})[Z,M]$ is a Q-local martingale. We now use integration by parts (under Q):

$$(5.4)\quad (\frac{1}{Z})[Z,M] = \int \frac{1}{Z_-} d[Z,M] + \int [Z,M]_- d(\frac{1}{Z}) + [[Z,M], \frac{1}{Z}] .$$

Let $N = \int [Z,M]_- d(\frac{1}{Z})$. Since $\frac{1}{Z}$ is a Q-local martingale, so also is N by (4.11). Thus (5.4) becomes:

$$\begin{aligned}(\frac{1}{Z_t})[Z,M]_t &= \int_0^t \frac{1}{Z_-} d[Z,M] + N_t + [[Z,M], \frac{1}{Z}]_t \\ &= \int_0^t \frac{1}{Z_{s-}} d[Z,M]_s + N_t + \sum_{0<s\leq t} \Delta(\frac{1}{Z_s})\Delta[Z,M]_s \\ &= \int_0^t \frac{1}{Z_s} d[Z,M]_s + N_t .\end{aligned}$$

Subtracting our two Q-local martingales yields:

$$\begin{aligned}&N - \{M - (\frac{1}{Z})[Z,M]\} \\ &\quad = N - \{M - \int \frac{1}{Z} d[Z,M] - N\} \\ &\quad = \{-M + \int \frac{1}{Z} d[Z,M]\}\end{aligned}$$

which, being the difference of two Q-local martingales, is itself one. This establishes the theorem. □

We note that in the proof we strongly used that Q is equivalent to P. Lenglart has established a generalization of this theorem where one only need assume Q<<P (cf [25]).

Our next theorem was proven independently by K. Bichteler [1] and C. Dellacherie [10]. Our approach is that of Dellacherie and others (Mokobodzki, Letta, Yan).

(5.6) Theorem. A right continuous adapted process $X = (X_t)_{0 \leq t < \infty}$ is a semimartingale if and only if it can be written $X = M + A$, where M is a local martingale, and A is a right continuous process with paths of finite variation on compacts. (That is, X is a semimartingale if and only if it is a decomposable process).

Proof. We have already seen the sufficiency. This is the content of Corollary (4.9).

We turn to the proof of necessity. Choose an arbitrary $t > 0$. If suffices to show X is decomposable on $[0,t]$. By a homeomorphism, then, we will show X is a decomposable process on $[0,\infty]$, but for an equivalent probability Q. The result then follows by an application of Girsanov's Theorem (5.1).
Moreover we have seen in this chapter that it will suffice to show that X is a quasimartingale under the new probability Q, since this implies it is a Q-decomposable process (3.3).

Since we are assuming X is a total semimartingale, we know that (by definition) for $H \in \underline{\underline{S}}$ such that $H_t = H_0 1_{]0,T_1]} + \dots + H_{n-1} 1_{]T_{n-1},\infty]}$; the mapping I_X given by:

$$I_X(H) = (H \cdot X)_\infty = H_0(X_{T_1} - X_0) + \dots + H_{n-1}(X_\infty - X_{T_{n-1}}) ,$$

from $\underline{\underline{S}}_u$ to L^0, is continuous. Let $\mathcal{B} = \{H \in \underline{\underline{S}}$ such that H has a representation as above and such that $|H| \leq 1\}$.

Let $\beta = I_X(\mathcal{B})$, the image of $\mathcal{B}$ under I_X. It will now suffice to find a probability Q equivalent to P such that $X_t \in L^1(dQ)$, all t, and such that $\sup_{U \in \beta} E_Q(U) = c < \infty$.

The reason this suffices is that if we take, for a given $0 = t_0 < t_1 < \dots < t_n = \infty$, the random variables $H_0 = \text{sign}\,(E_Q\{X_{t_1} - X_0 | \mathcal{F}_0\})$,

$H_1 = \text{sign}\,(E_Q\{X_{t_2} - X_{t_1} | \mathfrak{F}_{t_1}\})$, ..., we have that for this $H \in \mathfrak{B}$, $E_Q(I_X(H)) =$ $E_Q\{|E_Q\{X_{t_0} - X_{t_1} | \mathfrak{F}_0\}| + \ldots + |E_Q\{X_{t_n} - X_{t_{n+1}} | \mathfrak{F}_{t_n}\}|\}$. Since this partition τ was arbitrary, we have $\text{Var}\,(X) = \sup_\tau \text{Var}_\tau(X) \leq \sup_{U \in \beta} E_Q(U) = c < \infty$, and so X is a semimartingale.

We next make several observations:

(i) β is convex, since it is the image under I_X of a convex subset of $\underline{\underline{S}}$;

(ii) β is bounded in L^0;

Note that the above, which follow from the hypothesis that X is a semimartingale, can be expressed as:

(iii) for any sequence Y_n of elements of β, and for any sequence λ_n of scalars such that $\lim_n \lambda_n = 0$, we have $\lim_n \lambda_n Y_n = 0$ in L^0.

We continue with two lemmas that show how the semimartingale hypothesis on X is used:

(5.7) __Lemma__: __Property__ (iii) __above holds if and only if__ $\lim_{c\to\infty} \sup_{Y\in\beta} P(|Y|>c) = 0$

__Proof__. Suppose $\lim_{c\to\infty} \sup\{\sup_{Y\in\beta} P(|Y|>c)\} > 0$. Then there exists a sequence c_n tending to ∞, $Y_n \in \beta$, and $a > 0$ such that $P(|Y_n|>c_n) \geq a$, all n. If we take $\lambda_n = \frac{1}{c_n}$, we contradict (iii).

Next suppose $\lim_{c\to\infty} \sup_{Y\in\beta} P(|Y|>c) = 0$. Then for each a, $P(|\lambda_n Y_n|>a)$ $= P(|Y_n|>a/\lambda_n) \leq \sup_{Y\in\beta} P(|Y|>a/\lambda_n)$, which tends to 0 as n tends to infinity. □

(5.8) __Lemma__. __The random variable essential__ $\underset{t}{\text{sup}}\, X_t$ __is finite a.s.__

Proof. Let D be a countable dense set (e.g., the rationals) such that $\sup_{t\in D}|X_t| = \operatorname{ess\,sup}_t |X_t|$, using the right continuity of the paths of X. Suppose $P(\sup_{t\in D}|X_t| = \infty) > a > 0$ for some a. Let $\hat{D}$ be a finite subset of D and let $T_n = m\wedge \inf\{t\in D : |X_t|>n\}$. Then for each n, if we choose $\hat{D}$ and m big enough, we can take T_n such that $P(|X_{T_n}|>n) > a$.

Let $H_n = 1_{(0<t\leq T_n)}\in \mathcal{B}$, so that $X_{T_n} = I_X(H_n)\in\mathcal{B}$. But then $\limsup_n P(|X_{T_n}|>c) \geq a$, which contradicts that X is a semimartingale by (5.7). □

(5.9) Lemma. There exists a law Q equivalent to P such that $X_t\in L^1(dQ)$, all t.

Proof. Let $Y = \operatorname{ess\,sup}_t |X_t|$. Then $Y < \infty$ a.s. Let $A_m = \{m\leq Y<m+1\}$, and set $Z = \sum_{m=0}^{\infty} 2^{-m} 1_{A_m}$. Then Z is bounded, strictly positive, and $YZ\in L^1(dP)$. Define Q by $dQ = E_p(Z)^{-1}Z dP$ and then

$$E_Q\{\int_0^\infty |X_t|e^{-t}dt\} \leq E_Q\{\int_0^\infty Ye^{-t}dt\} = E_Q\{Y\int_0^\infty e^{-t}dt\} = E_Q(Y) < \infty;$$

hence $E_Q\{|X_T|\} < \infty$, each $t > 0$. □

Observe that $\beta\in L^1(dQ)$, and since Q is equivalent to P, if we can show X has a decomposition $X = M + A$ as in (5.6), then it does under P as well by Theorem (5.1). Thus without loss of generality we can assume that $\beta \subseteq L^1(dP)$. The theorem now follows from the following lemma, which follows Yan [47].

(5.10) Lemma. Let β be a subset of $L^1(dP)$, $0\in\beta$, that is bounded in probability: that is, for any $\varepsilon > 0$ there exists a $c > 0$ such that $P(\zeta > c) \leq \varepsilon$, for any $\zeta\in\beta$. Then there exists a probability Q equivalent to P, with a bounded density, such that $\sup_{U\in\beta} E_Q(U) < \infty$.

Proof. First note that the hypotheses imply that $\beta \subset L^1(dQ)$. What we must show is that there exists a bounded random variable Z, such that $P(Z > 0) = 1$, and such that $\sup_{\zeta \in \beta} E_P(Z\zeta) < \infty$.

Let $A \in \mathcal{F}$ such that $P(A) > 0$. Then there exists a constant d such that $P(\zeta > d) \leq P(A)/2$, for all $\zeta \in \beta$. Using this constant d, let $\epsilon = 2d$, and we have that $0 \leq c1_A \notin \beta$, and moreover if B_+ denotes all bounded, positive r.v., then $c1_A$ is not in the $L^1(dP)$ closure of $\beta - B_+$, denoted $\overline{\beta - B_+}$. That is, $c1_A \notin \overline{\beta - B_+}$. Since the dual of L^1 is L^∞ , and $\beta - B_+$ is convex, by the Hahn-Banach theorem there exists a bounded random variable Y such that

(5.11) $\sup_{\zeta \in \beta, \eta \in B_+} E\{Y(\zeta - \eta)\} < cE\{Y1_A\}$. Replacing η by $a\eta$ and letting a tend to ∞ shows that $Y \geq 0$ a.s., since otherwise the expectation on the left side above would get arbitrarily large. Next suppose $\eta = 0$. Then (5.11) implies that

$$\sup_{\zeta \in \beta} E[Y\zeta] \leq c\, E\{Y1_A\} < +\infty.$$

Now set $\mathcal{H} = \{Y \in B_+ : \sup_{\zeta \in \beta} E[Y\zeta] < \infty\}$. Since $0 \in B_+$, we know $\mathcal{H}$ is not empty. Let $\mathcal{A}$ = {all sets of the form $\{Z = 0\}$, $Z \in \mathcal{H}$ }. We wish to show that there exists a $Z \in \mathcal{H}$ such that $P(\{Z = 0\}) = \inf_{A \in \mathcal{A}} P(A)$. Suppose, then, that Z_n is a sequence of elements of $\mathcal{H}$. Let $c_n = \sup_{\zeta \in \beta} E[Z_n\zeta]$ and $d_n = ||Z_n||_{L^\infty}$. (Since $0 \in \beta$, we have $c_n \geq 0$). Choose b_n such that $\sum b_n c_n < \infty$ and $\sum b_n d_n < \infty$, and set $Z = \sum b_n Z_n$. Then clearly $Z \in \mathcal{H}$. Moreover, $\{Z = 0\} = \bigcap_n \{Z_n = 0\}$. Thus $\mathcal{A}$ is stable under countable intersections, and so there exists a Z such that $P(\{Z = 0\}) = \inf_{A \in \mathcal{A}} P(A)$.

We now wish to show $Z > 0$ a.s. Suppose not. That is, suppose $P(\{Z = 0\}) > 0$. Let Y satisfy (5.11) (we have seen that there exists such a Y and that it hence is in $\mathcal{H}$). Further we take for our set A in (5.11) the set $A = \{Z = 0\}$, for which we are assuming $P(A) > 0$. Since $0 \in \beta$ and $0 \in B_+$, we have from (5.9) that

(5.12) $0 < E\{Y1_A\} = E\{Y1_{\{Z = 0\}}\}$.

Since each of Y and Z are in $\mathcal{H}$, their sum is in $\mathcal{H}$ as well. But then
$P\{Y + Z = 0\} = P\{Z = 0\} - P(\{Z = 0\} \cap \{Y > 0\}) < P(\{Z = 0\})$, by (5.12).
This, then, is a contradiction, since $P(\{Z = 0\})$ is minimal for $Z \in \mathcal{H}$. Therefore we conclude $Z > 0$ a.s., and since $Z \in B_+$, it is bounded as well, and the lemma is proved; thus also, Theorem (5.6) is proved. □

V. STOCHASTIC INTEGRATION WITH PREDICTABLE INTEGRANDS AND SEMIMARTINGALE LOCAL TIME

1. Introduction.

In Chapter III we treated stochastic integration for semimartingales in a relatively non-technical manner. We were able to do so by limiting the integrands to the space $\mathbb{L}$. This is, of course, sufficient to prove a change of variables formula, and it is also sufficient in many applications, such as the study of stochastic differential equations. Extending the integral to more general integrands, analogous to a Lebesgue-type integral, requires deep results: namely, Theorems IV.1.2 and IV.1.3. Extending the semimartingale integral to such integrands is essential to many applications (such as martingale representation) and gives semimartingale local time a natural interpretation through the Meyer-Tanaka formula (4.24), as we shall see.

In this chapter we first define stochastic integration for bounded, predictable integrands and $\mathcal{H}^2$-semimartingales (2.10). We then extend this definition to arbitrary semimartingales (2.20) and to locally bounded predictable integrands (2.21). In paragraph three we establish some measurability results used in paragraph four in the development of local time. A generalization (4.14) of the change of variables formula of chapter three (III.7.1) is given using local time.

2. Stochastic Integration for Predictable Integrands

Let X be a semimartingale. Then by IV.5.6 we know X is decomposable. Let $X = X_0 + N + B$ be any decomposition where $N_0 = B_0 = 0$. Define:

(2.1) $$j_2(N,B) = \left|\left|[N,N]_\infty^{\frac{1}{2}} + |X_0| + \int_0^\infty |dB_s|\right|\right|_{L^2}.$$

(2.2) Definition. The space $\mathcal{H}^2$ of semimartingales consists of all semimartingales X for which there is a decomposition $X = X_0 + N + B$ such that $j_2(N,B) < \infty$.

It is simple to check that if X is an $\mathcal{H}^2$-semimartingale, then X is also a quasimartingale, and hence by IV.3.3 we know that X is a special semimartingale. We can now next define:

(2.3) Definition. Let X be a special semimartingale with $X_0 = 0$ and with canonical decomposition $X = \bar{N} + \bar{B}$. Then the $\mathcal{H}^2$-norm of X is defined to be:

$$||X||_{\mathcal{H}^2} = ||[\bar{N},\bar{N}]_\infty^{\frac{1}{2}}||_{L^2} + ||\int_0^\infty |d\bar{B}_s| \;||_{L^2}$$

It can be shown that X has finite $\mathcal{H}_2$-norm if and only if there exists a decomposition $X = N + B$ where $j_2(N,B) < \infty$, with j_2 as given in (2.1).

(2.4) Theorem. The space of $\mathcal{H}^2$ semimartingales is a Banach space.

Proof. The space is clearly a normed linear space. Since $||\bar{N}||_{L^2} = ||[\bar{N},\bar{N}]_\infty^{\frac{1}{2}}||_{L^2}$, it follows from Doob's maximal quadratic inequality that the space of square integrable martingales is complete. As for $\bar{B}$, let $(\bar{B}^n)$ be a sequence such that $\sum_n ||\bar{B}^n|_\infty||_{L^2} < \infty$, where $|\bar{B}^n|_\infty = \int_0^\infty |d\bar{B}_s^n|$. Then the series $\sum_n \bar{B}^n$ converges to a limit $\bar{B}$ and $\lim_{m\to\infty} \sum_{n\geq m} \int_0^\infty |d\bar{B}_s^n| = 0$ in L^1 and is dominated in L^2 by $\sum \int_0^\infty |d\bar{B}_s^n|$, hence it tends to 0 in L^2 as well. Thus $\sum \bar{B}^n$ converges to $\bar{B}$ in $L^2(dP)$, and completeness follows. □

For simplicity, in this paragraph we henceforth assume that all semimartingales X have the property $X_0 = 0$.

Let $b\mathbb{L}$ denote the bounded processes in $\mathbb{L}$. For $H \in b\mathbb{L}$ and $X \in \mathcal{H}^2$ we have $H \cdot X$ is in $\mathcal{H}^2$ as well. Moreover if X has the canonical decomposition $X = \bar{N} + \bar{B}$, then $H \cdot X$ has the canonical decomposition $H \cdot X = H \cdot \bar{N} + H \cdot \bar{B}$, and moreover:

$$(2.5) \qquad ||H \cdot X||_{L^2} = ||[H \cdot \bar{N}, H \cdot \bar{N}]_\infty^{\frac{1}{2}}||_{L^2} + ||\int_0^\infty |d(H \cdot \bar{B})_s| \, ||_{L^2}$$

$$= ||(\int_0^\infty H_s^2 d[\bar{N},\bar{N}]_s)^{\frac{1}{2}}||_{L^2} + ||\int_0^\infty |H_s||d\bar{B}_s| \, ||_{L^2} .$$

Since the Lebesgue-Stieltjes path by path integrals $\int H_s^2 d[\bar{N},\bar{N}]_s$, $\int H_s d\bar{B}_s$ make sense for any $H \in b\mathcal{P}$ (the bounded, predictable processes) as well as $H \in \mathbb{L}$, we can use property (2.5) to extend our class of integrands.

(2.6) Definition. Given an $\mathcal{H}^2$-semimartingale X with canonical decomposition $X = \bar{N} + \bar{B}$, and processes $H, J \in b\mathcal{P}$, define $d_X(H,J) =$

$$||(\int_0^\infty (H_s - J_s)^2 d[\bar{N},\bar{N}]_s)^{\frac{1}{2}}||_{L^2} + ||\int_0^\infty |H_s - J_s||d\bar{B}_s| \, ||_{L^2}.$$

(2.7) Theorem. $b\mathbb{L}$ is dense in $b\mathcal{P}$ under the "distance" $d_X(\cdot,\cdot)$.

Proof. Let $\mathcal{H} = \{H \in b\mathcal{P}$: for any $\varepsilon > 0$, there exists a $J \in b\mathbb{L}$ such that $d_X(H,J) < \varepsilon\}$. Then $\mathcal{H}$ contains $b\mathbb{L}$ and the constants. Moreover if $H^n \in \mathcal{H}$ and increasing to H with H bounded, then by the dominated convergence theorem for $n > N$, $d_X(H,H^n) < \delta$. Since $H^n \in \mathcal{H}$, there exists a $J^n \in b\mathbb{L}$ such that $d_X(H^n,J^n) < \gamma$. Therefore for $n > N$, there exists a $J^n \in b\mathbb{L}$ such that $d_X(H,J^n) < \varepsilon$ by appropriate choices of δ and γ. An application of the monotone class theorem yields the result. □

(2.8) Theorem. Given a semimartingale X in $\mathcal{H}^2$ and $H^n \in b\mathbb{L}$ such that H^n is Cauchy under d_X, then $H^n \cdot X$ is Cauchy in $\mathcal{H}^2$.

Proof. We have $||H^n \cdot X - H^m \cdot X||_{\mathcal{H}^2} = d_X(H^n, H^m)$. □

(2.9) *Theorem.* *Let X be a semimartingale in* $\mathcal{H}^2$ *and let* $H \in b\mathcal{P}$. *If* $H^n \in b\mathbb{L}$ *and* $J^n \in b\mathbb{L}$ *such that* $\lim_n d_X(H^n, H) = \lim_n d_X(J^n, H) = 0$, *then* $H^n \cdot X$ *and* $J^n \cdot H$ *tend to the same limit in* $\mathcal{H}^2$.

Proof. Let $Y = \lim_n H^n \cdot X$ and let $Z = \lim_n J^n \cdot X$, in $\mathcal{H}^2$.
Then

$$\begin{aligned} ||Y - Z||_{\mathcal{H}^2} &\leq ||Y - H^n \cdot X||_{\mathcal{H}^2} + ||H^n \cdot X - J^n \cdot X||_{\mathcal{H}^2} + ||J^n \cdot X - Z||_{\mathcal{H}^2} \\ &\leq 2\varepsilon + ||H^n \cdot X - J^n \cdot X||_{\mathcal{H}^2} \quad (n \geq N_\varepsilon) \\ &\leq 2\varepsilon + d_X(H^n, J^n) \\ &\leq 2\varepsilon + d_X(H^n, H) + d_X(H, J^n) \\ &\leq 4\varepsilon, \text{ and the result follows.} \quad \square \end{aligned}$$

We can now make the:

(2.10) *Definition.* Let X be a semimartingale in $\mathcal{H}^2$ and let $H \in b\mathcal{P}$. Let $H^n \in b\mathbb{L}$ be such that $\lim d_X(H^n, H) = 0$. *The stochastic integral* $H \cdot X$ is the (unique) semimartingale Y in $\mathcal{H}^2$ given by $\lim H^n \cdot X = Y = H \cdot X$, with convergence in $\mathcal{H}^2$.

(2.11) *Theorem.* *Let X be a semimartingale in* $\mathcal{H}^2$. *Then* $E\{(\sup_t |X_t|)^2\} \leq c||X||^2_{\mathcal{H}^2}$.

Proof. Let $X^*_\infty = \sup_t |X_t|$. Then $X^*_\infty \leq \bar{N}^*_\infty + \int_0^\infty |d\bar{B}_s|$, and by Doob's maximal quadratic inequality,

$$E((\bar{N}^*_\infty)^2) \leq 4E(\bar{N}^2_\infty) = 4E([\bar{N},\bar{N}]_\infty) .$$

Thus $E\{(X^*_\infty)^2\} \leq 2E\{(\bar{N}^*_\infty)^2\} + 2E\{(\int_0^\infty |d\bar{B}_s|)^2\}$

$$\leq 8||[\bar{N},\bar{N}]_\infty^{\frac{1}{2}}||_{L^2}^2 + 2||\int_0^\infty |d\bar{B}_s|\,||_{L^2}^2$$

and the result follows. □

(2.12) Corollary. If X^n is a sequence of semimartingales converging to X in $\mathcal{H}^2$ then there exists a subsequence n_k such that $\lim_{n_k \to \infty} (X^{n_k} - X)_\infty^* = 0$ a.s..

Proof. Since $(X^n - X)^*$ converges to 0 in L^2 by (2.11), there exists a subsequence converging a.s.. □

We are now in a position to investigate some of the properties of this more general stochastic integral. The bilinearity is evident, and we state it without proof.

(2.13) Theorem. Let X, Y be $\mathcal{H}^2$ semimartingales and $H,K \in b\mathcal{P}$. Then $(H + K) \cdot X = H \cdot X + K \cdot X$, and $H \cdot (X + Y) = H \cdot X + H \cdot Y$.

(2.14) Theorem. Let X be a square-integrable martingale and let $H \in b\mathcal{P}$. Then $H \cdot X$ is a square integrable martingale.

Proof. Clearly X is a semimartingale in $\mathcal{H}^2$. Let $H^n \in b\mathbb{L}$ such that $\lim_{n\to\infty} d_X(H^n,H) = 0$. Then $H^n \cdot X$ is a square integrable martingale by III.5.11 for each n. The theorem follows by L^2-convergence. □

(2.15) Theorem. Let X be an $\mathcal{H}^2$-semimartingale with paths of finite variation on compacts. Let $H \in b\mathcal{P}$. Then $H \cdot X$ agrees with a path by path Lebesgue-Stieltjes integral.

Proof. Let $H^n \in b\mathbb{L}$ such that $\lim d_X(H^n,H) = 0$. Then $H^n \cdot X$ is a Lebesgue-Stieltjes integral for each n, and the result follows by passing to the limit. □

(2.16) Theorem. Let X be an $\mathcal{H}^2$-semimartingale and $H \in b\mathcal{P}$. Then $\Delta(H \cdot X) = H(\Delta X)$.

Proof. Let $H^n \in b\mathbb{L}$ such that $\lim_{n\to\infty} d_X(H^n,H) = 0$. Then there exists a subsequence n_k such that $\lim_{n_k\to\infty} (H^{n_k} \cdot X - H \cdot X)^*_\infty = 0$ a.s., by (2.12). This implies $\Delta(H^{n_k} \cdot X) \to \Delta(H \cdot X)$ outside of an evanescent set. However since $H^{n_k} \in \mathbb{L}$, we know $\Delta(H^{n_k} \cdot X) = H^{n_k}(\Delta X)$. Therefore $\lim_{n\to\infty} H^{n_k}_t(\omega) = \frac{\Delta(H\cdot X)_t(\omega)}{\Delta X_t(\omega)}$, on $\{\Delta X_t \neq 0\}$, hence the limit exists. If

$$\Lambda = \{\omega: \text{there exists } t > 0 \text{ such that} \lim H^{n_k}(\omega) \neq H_t(\omega) \text{ and } \Delta X_t \neq 0\},$$

and if $P(\Lambda) > 0$, then we would contradict that $\lim_{n\to\infty} d_X(H^n,H) = 0$, since

$$\lim_n d_X(H^n,H) \geq \lim_{n_k} \| 1_\Lambda (\int_0^\infty (H^{n_k}_s - H_s)^2 d(\Delta \bar{N}_s)^2)^{\frac{1}{2}} + 1_\Lambda \int_0^\infty |H^{n_k}_s - H_s| d|\Delta \bar{B}_s| \|_{L^2},$$

and if $\Delta X_s \neq 0$, then $|\Delta \bar{N}_s| + |\Delta \bar{B}_s| > 0$. Thus $P(\Lambda) = 0$, and we have

$$\Delta(H\cdot X)_t = \lim_{n_k} H^{n_k}_t \Delta X_t = H_t \Delta X_t. \quad \square$$

(2.17) Theorem. Let X be an $\mathcal{H}^2$-semimartingale, and let $H,K \in b\mathcal{P}$. Then $H \cdot (K \cdot X) = (HK) \cdot X$.

Proof. This follows from the result for $H,K \in b\mathbb{L}$ (III.5.10), and then by taking limits. $\square$

(2.18) Theorem. Let X, Y be $\mathcal{H}^2$-semimartingales and let $H,K \in b\mathcal{P}$. Then

$$[H\cdot X, K\cdot Y]_t = \int_0^t H_s K_s d[X,Y]_s$$

and, in particular,

$$[H\cdot X, H\cdot X]_t = \int_0^t H_s^2 d[X,X]_s.$$

Proof. As in the proof of Theorem III.6.16, it suffices to show

$[H\cdot X,Y]_t = \int_0^t H_s d[X,Y]_s$. Let $H^n \in b\mathbb{L}$ such that $d_X(H^n,H) \to 0$. By stopping we can also assume $Y_- \in b\mathbb{L}$, and it is then easy to check that $d_X(H^n Y_-, HY_-)$ also tends to 0. We have, by III.6.16,

(2.19) $[H^n\cdot X,Y] = \int H^n d[X,Y] \to \int H d[X,Y]$;

Let $Z^n = H^n \cdot X$, and by (2.12) we know there is a subsequence n_k such that $\lim_{n_k\to\infty} (Z^{n_k} - Z)^*_\infty = 0$ a.s., where $Z = H \cdot X$. Then

$$[Z^{n_k},Y] = Z^{n_k}Y - (Y_-) \cdot Z^{n_k} - (Z^{n_k}_-) \cdot Y = Z^{n_k}Y - (Y_- H^{n_k}) \cdot X - (Z^{n_k}_-) \cdot Y,$$

by integration by parts and by (2.17). Taking limits we have

$$\lim_{n_k\to\infty} [Z^{n_k},Y] = ZY - Y_- \cdot (H\cdot X) - Z_- \cdot Y$$

$$= ZY - Y_- \cdot (Z) - Z_- \cdot Y$$

$$= [Z,Y] = [H\cdot X,Y].$$

Combining this with (2.19) yields the result. □

Now let X be <u>any</u> semimartingale with $X_0 = 0$ for simplicity. Let $X = M + A$ be a decomposition where the local martingale has bounded jumps, which we know exists by IV.4.6. Define:

$$T_n = \inf\{t > 0: \ |M_t| > n \ \text{ or } \int_0^t |dA_s| > n\}.$$

Then

$$X^{T_n-} = \begin{cases} X_t & \text{if } T_n(\omega) > t \\ X_{T_n-} & \text{if } T_n(\omega) \leq t, \end{cases}$$

and $X^{T_n-} = M^{T_n} + A^{T_n-} - (\Delta M_{T_n}) 1_{[T_n,\infty[}$; if c is a bound for the jumps of M, then $|X^{T_n-}| \leq 2n + 2c$; that is, it is bounded. Note that X^{T_n-} is a semimartingale and is in $\mathcal{H}^2$. This allows the extension of the stochastic integral to arbitrary semimartingales.

(2.20) <u>Definition</u>. Let X be a semimartingale and $H \in b\mathcal{P}$. Let T^n be stopping times increasing to ∞ such that X^{T_n-} is in $\mathcal{H}^2$. Define $H \cdot X$ to be $H \cdot (X^{T_n-})$ on $[0,T_n[$ for each n, and call $H \cdot X$ <u>the stochastic integral</u>.

Note that if $T_m > T_n$ in Definition (2.20), then if $H^n \in b\mathbb{L}$ converge to $H \in b\mathcal{P}$ in $d_{(X^{T_m-})}(\cdot,\cdot)$, then they converge as well in $d_{(X^{T_n-})}(\cdot,\cdot)$, so the integral is well defined. We can also further extend the class of integrands. A process $H \in \mathcal{P}$ is said to be <u>locally bounded</u> if there exist stopping times T^k increasing to ∞ a.s. such that $(H-H_0)^{T^k}$ is in $b\mathcal{P}$ for each k.

(2.21) <u>Definition</u>. Let X be a semimartingale and let $H \in \mathcal{P}$ be locally bounded. The <u>stochastic integral</u> $H \cdot X$ is defined to be $H_0 X_0 + (H - H_0)^{T^k} \cdot X$ on $[0,T^k]$.

It is now a simple matter to check that all the properties (2.13) through (2.18) still hold for this mild extension.

These techniques can be carried further, but we do not do so here. By developing the semimartingale topology, which is closely tied to the $\mathcal{H}^2$ norm, one can extend the stochastic integral to the space of predictable, <u>integrable</u> processes. We refer the interested reader to [4] and [44].

3. Stochastic Integration Depending on a Parameter

In order to present the theory of local time for semimartingales and its concomitant formulas, it is first necessary to establish some measurability results, which is the purpose of this paragraph. We begin with a clever theorem due to C. Stricker and M. Yor [45], which builds on an idea of C. Doléans-Dade [12]. We let $(A,\mathcal{A})$ be a measurable space.

(3.1) Theorem. Let X_n be a sequence of functions on $A \times \Omega$ which are $\mathcal{A}\otimes\mathcal{F}$-measurable. Suppose $X_n(a,\cdot)$ converges in probability on Ω for each $a \in A$. Then there exists an $\mathcal{A}\otimes\mathcal{F}$-measurable function X such that $X(a,\cdot) = \lim_n X_n(a,\cdot)$, in probability, for every $a \in A$.

Proof. Set $n_0(a) = 1$ and define inductively

$$n_k(a) = \inf\{m > n_{k-1}(a) : \sup_{p,q\geq m} P[|X_p(a,\cdot) - X_q(a,\cdot)| > 2^{-k}] \leq 2^{-k}\}.$$

Set $Y_k(a,\omega) = X_{n_{k(a)}}(a,\omega)$. (Note that $a \to n_k(a)$ is measurable). For each $k \geq 1$ and every a,

$$P\{|Y_{k+1}(a,\cdot) - Y_k(a,\cdot)| > 2^{-k}\} \leq 2^{-k},$$

and the Borel-Cantelli lemma implies that $Y_k(a,\cdot)$ converges a.s. Next let Λ^a be the set where Y_k converges and define

$$X(a,\omega) = \begin{cases} \lim_{k\to\infty} Y_k(a,\omega), & \omega \in \Lambda^a \\ 0 & \omega \notin \Lambda^a \end{cases}$$

Then $\Lambda^a \in \mathcal{A}\otimes\mathcal{F}$, and hence X is jointly measurable. □

We will not have use of it, but we record here nonetheless a useful corollary that follows trivially.

(3.2) Corollary. Let X_n be as in (3.1), except assume that X_n converges in $L^p(p \geq 1)$ for every $a \in A$. Then there exists an X on $A \times \Omega$ that is $\mathcal{A} \otimes \mathcal{F}$-measurable and such that $X_n(a,\cdot)$ converges in L^p to $X(a,\cdot)$ for every a.

Proof. Convergence in L^p implies convergence in probability, thus the corollary follows from (3.1) □

At this point let us recall two definitions.

(3.3) Definition. The predictable σ-algebra $\mathcal{P}$ on $\mathbb{R}_+ \times \Omega$ is the smallest σ-algebra making all left continuous, adapted processes measurable; the optional σ-algebra $\mathcal{O}$ on $\mathbb{R}_+ \times \Omega$ is the smallest σ-algebra making all right continuous, adapted processes measurable.

(3.4) Theorem. Let X be a semimartingale and let $H(a,t,\omega) = H^a_t(\omega)$ be $\mathcal{A} \otimes \mathcal{P}$- measurable and bounded. Then there is a function $Z(a,t,\omega) \in \mathcal{A} \otimes \mathcal{O}$ such that $Z(a,\cdot,\cdot)$ is a cadlag, adapted version of the stochastic integral $H^a \cdot X$, each $a \in A$.

Proof. Let $\mathcal{H} = \{H \in b\mathcal{A} \otimes \mathcal{P}:$ the conclusion holds$\}$. If $K \in b\,\mathcal{P}$ and $f \in b\mathcal{A}$, and if $H(a,t,\omega) = f(a)K(t,\omega)$, then

$\int H^a_s dX_s = \int f(a)K_s dX_s = f(a) \int K_s dX_s \in \mathcal{A} \otimes \mathcal{O}$, and so $H \in \mathcal{H}$. Note that H of this form generate $b\mathcal{A} \otimes \mathcal{P}$.

Suppose next that $H^n \in \mathcal{H}$ and H^n converges boundedly to a process $H \in b\mathcal{A} \otimes \mathcal{P}$. By the monotone class theorem it will suffice to show that the limit process $H \in \mathcal{H}$. Let $\hat{\Omega} = \mathbb{R}_+ \times \Omega$, and define $\hat{P} = \lambda \times P$, where $\lambda(dt) = e^{-t}dt$ on $\mathbb{R}_+$. It then suffices to show, thanks to Theorem (3.1), that $H^{n,a} \cdot X$ converges in $\hat{P}$ probability for each a.

Suppose first $X \in \mathcal{H}^2$. Since

$$E\{\sup_t(\int_0^t H_s^{n,a}dX_s - \int_0^t H_s^a dX_s)^2\} \le c|| \int H_s^{n,a}dX_s - \int H_s^a dX_s||^2_{\mathcal{H}^2}$$

by Theorem (2.11), it suffices to show that $\lim_{n\to\infty} ||H^{n,a} \cdot X - H^a \cdot X||_{\mathcal{H}^2} = 0$, each $a \in \mathcal{A}$, because if $\sup_t Y^n(t,\cdot)$ tends to 0 in P-probability, then $Y^n(t,\omega)$ tends to 0 in $\hat{P}$-probability as well. Let $X = \bar{N} + \bar{B}$ be the canonical decomposition of X. Observe that

$$||H^{n,a} \cdot X - H^a \cdot X||^2_{\mathcal{H}^2} \le 2E(\int_0^\infty (H_s^{n,a} - H_s^a)^2 d[\bar{N},\bar{N}]_s)$$
$$+ 2E(\int_0^\infty |H_s^{n,a} - H_s^a||d\bar{B}_s|) \quad ,$$

which tends to zero by the dominated convergence theorem, since $H^{n,a}$ and H^a are bounded, and we have the result.

For a general semimartingale X not necessarily in $\mathcal{H}^2$, let T^n be stopping times increasing to ∞ a.s. with $T^0 = 0$ such that $X^{T^n -} \in \mathcal{H}^2$, for each n. Then $H^a \cdot X^{T^n -} \in \mathcal{A} \otimes \mathcal{O}$; but $H^a \cdot X^{T^n -} = H^a \cdot X 1_{[0,T^n[} + H^a \cdot X_{T^n -} 1_{[T^n,\infty[}$, with each term in $A \otimes \mathcal{O}$. Since

$$H^a \cdot X = \sum_{n=0}^\infty H^a \cdot X 1_{[T^n,T^{n+1}[} \quad ,$$

we have $H_a \cdot X \in \mathcal{A} \otimes \mathcal{O}$.

(3.5) Theorem. Let $X(a,t, \ ) = X_t^a(\omega)$ be $\mathcal{A} \otimes \mathcal{O}$ - measurable and a semimartingale for each $a \in \mathcal{A}$. Then there exists a version of $[X^a,X^a]$ which is $\mathcal{A} \otimes \mathcal{O}$- measurable.

Proof. Let σ_n be a sequence of fixed times wich form a sequence of partitions of [0,s] tending to the identity (cf. III.6.1). Then

$$(X_0^a)^2 + \sum_i (X^{a,t_{i+1}^n} - X^{a,t_i^n})^2$$

converges to $[X^a,X^a]_s$ in ucp for each $a \in A$; and this holds for each $s > 0$. Let $\hat{P}(dt,d\omega) = \lambda(dt) \times P(d\omega)$, where $\lambda(dt) = e^{-t}dt$ on $\mathbb{R}_+$. We define $\hat{\Omega} = \mathbb{R}_+ \times \Omega$, so that $\hat{P}$ is a probability on Ω. We set $Y_s^{n,a} = (X_0^a)^2 + \sum_i (X^{a,t_{i+1}^n} - X^{a,t_i^n})^2$, and then $Y^{n,a} \in \mathcal{A} \otimes \mathcal{O}$. It suffices to show that $Y^{n,a}$ converges to $[X^a,X^a]$ in $\hat{P}$-probability for each a by Theorem (3.1). By Theorem III.6.2 we know that $Y^{n,a}$ converges to $[X^a,X^a]$ in ucp for P, for each a. This implies $\lim_n P\{\sup_{s \le t} |Y_s^{n,a} - [X^a,X^a]_s| > \delta\} = 0$, each $\delta > 0$, and each $t > 0$. Let $\varepsilon > 0$ be given, and choose m so large that $\hat{P}([m,\infty[\times \Omega) < \varepsilon$. Then

$$\hat{P}\{|Y^{n,a} - [X^a,X^a]| > \delta\} \le \hat{P}\{1_{[0,m[\times\Omega}(|Y^{n,a} - [X^a,X^a]| > \delta)\} + \hat{P}([m,\infty) \times \Omega)$$

$$\le P\{\sup_{s \le m} |Y_s^{n,a} - [X^a,X^a]_s| > \delta\} + \varepsilon$$

which yields the result. □

(3.6) Theorem. Let X be a semimartingale, and let $H(a,t,\omega) = H_t^a$ be $\mathcal{A} \otimes \mathcal{P}$-measurable such that $H_t^a(\omega)$ is locally bounded for each fixed $a \in A$. Let $Z^a = H^a \cdot X$ be the $\mathcal{A} \otimes \mathcal{O}$-measurable version of the stochastic integral. Then $\int_A Z_t^a \mu(da)$ is a version of $H \cdot X$, where $H = \int H^a \mu(da)$, for μ a finite measure on $(A, \mathcal{A})$.

Proof. First suppose $H(a,t,\omega) = f(a)K(t,\omega)$ for $f \in b\mathcal{A}$ and $K \in b\mathcal{P}$. Set $H \cdot X = f(a) K \cdot X$, the result then being clear. Next set $\mathcal{H} = \{H \in b\mathcal{A} \otimes \mathcal{P}$ such that the conclusion holds$\}$. Then $\mathcal{H}$ is a vector space containing the generating collection $f(a)K(t,\omega)$. The result for bounded functions follows by the monotone class theorem, and the extension to locally bounded integrands is simple. □

4. The Local Time of a Semimartingale

We have seen with "Itô's lemma" (III.7.1) that if $f: \mathbb{R} \to \mathbb{R}$ is $\mathcal{C}^2$ and X is a semimartingale, then $f(X)$ is again a semimartingale. That is, semimartingales are preserved under $\mathcal{C}^2$ transformations. This property extends slightly: semimartingales are preserved under convex transformations, as (4.1) shows. (Indeed this is the best one can do in general: if $B = (B_t)_{t \geq 0}$ is standard Brownian motion and $Y_t = f(B_t)$ is a semimartingale, then f must be the difference of convex functions [6]). Local time appears in the extension of Itô's lemma to convex functions (4.15).

(4.1) Theorem. Let $f: \mathbb{R} \to \mathbb{R}$ be convex and let X be a semimartingale. Then $f(X)$ is a semimartingale and one has

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + A_t$$

where f' is the left derivative of f and A is an adapted, right continuous, increasing process. Moreover $\Delta A_t = f(X_t) - f(X_{t-}) - f'(X_{t-})\Delta X_t$.

Proof. First suppose $|X|$ is bounded by n. Let g be a positive $\mathcal{C}^\infty$ function with compact support in $]-\infty,0]$ such that $\int_{-\infty}^{\infty} g(s)ds = 1$. Let $f_n(t) = n\int_{-\infty}^{\infty} f(t+s)g(ns)ds$. Then f_n is convex and $\mathcal{C}^2$ and moreover f_n' increases to f' as n tends to ∞. By Itô's lemma (III.7.1)

$$(4.2)\quad f_n(X_t) = f_n(X_0) + \int_0^t f_n'(X_{s-})dX_s + A_t^n$$

where $A_t^n = \sum_{0<s\leq t} \{f_n(X_s) - f_n(X_{s-}) - f'_n(X_{s-})\Delta X_s\} + \frac{1}{2}\int_0^t f''_n(X_{s-})d[X,X]_s^c$.

The convexity of f implies that A^n is an increasing process. Letting n tend to ∞, (4.2) becomes:

$$(4.3)\quad f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + A_t$$

where $\lim_n A_t^n = A_t$ in L^2, and where the convergence of the stochastic integral terms is in $\mathcal{H}^2$ on $[0,t]$.

We now compare the jumps on both sides of the equation (4.3). Since $\int_0^0 f'(X_{s-})dX_s = 0$ we have that $A_0 = 0$. When $t > 0$, the jump of the left side of (4.3) is $f(X_t) - f(X_{t-})$, while the jump of the right side equals $f'(X_{t-})\Delta X_t + \Delta A_t$; therefore $\Delta A_t = f(X_t) - f(X_{t-}) - f'(X_t)\Delta X_t$, and (4.1) is established for $|X|$ bounded by n.

For general X, let $T^n = \inf\{t: |X_t| \geq n\}$. Then $Y = X1_{]0,T^n[}$ is a bounded semi-martingale, hence $f(Y_t) = f(Y_0) + \int_0^t f'(Y_{s-})dY_s + A_t^n$. This is true for each n, hence one has $f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + A_t^n$ on $[0,T^n[$. Therefore $(A^{n+1})^{T^n-} = (A^n)^{T^n-}$ and we can define $A = A^n$ on $[0,T^n[$, each n. The general result now follows. □

For x a real variable let x^+, x^- be the functions $x^+ = \max(x,0)$ and $x^- = -\min(x,0)$.

(4.4) Corollary. Let X be a semimartingale. Then $|X|$, X^+, X^- are all semimartingales.

Proof. The functions $f(x) = |x|$, $g(x) = x^+$ and $h(x) = x^-$ are all convex; the result then follows by (4.1). □

For x,y real variables, let $x \vee y \equiv$ maximum (x,y) and $x \wedge y \equiv$ minimum (x,y).

(4.5) Corollary. Let X,Y be semimartingales. Then $X \vee Y$ and $X \wedge Y$ are semimartingales.

Proof. Since semimartingales form a vector space and $x \vee y = \frac{1}{2}(|x-y| + x + y)$ and $x \wedge y = \frac{1}{2}(y + x - |x-y|)$, the result is an immediate consequence of (4.4). □

We can summarize the surprisingly broad stability properties of semimartingales.

(4.6) Theorem. The space of semimartingales is a vector space, an algebra, a lattice, and is stable under $\mathcal{C}^2$, and more generally under convex transformations.

Proof. This is a combination of III.2.1, III.6.5, 4.5, III.7.1, and 4.1. □

We next define a function that is the left derivative of the convex function $h_0(x) = |x|$: define

$$(4.7) \quad \begin{cases} \text{sign}(x) = \begin{cases} 1 & x > 0 \\ -1 & x \leq 0 \end{cases} \\ h_a(x) = |x-a|. \end{cases}$$

Then sign (x-a) is the left derivative of $h_a(x)$. By (4.1) we have

(4.8) $|X_t-a| = |X_0-a| + \int_0^t \text{sign}\,(X_{s-}-a)dX_s + A_t^a$,

where A_t^a is increasing.

(4.9) Definition. Let X be a semimartingale. The local time at a of X, denoted $L_t^a = L^a(X)_t$, is defined to be the process given by

$L_t^a = A_t^a - \sum_{s\le t}\{h_a(X_s) - h_a(X_{s-}) - h_a'(X_{s-})\Delta X_s\}$, using the notation of (4.1), (4.7) and (4.8).

Note that the local time L^a as defined in (4.9) is continuous in the variable t by (4.1); moreover, if $\mathcal{A}$ are the Borel sets in $\mathbb{R}$, we know by (3.4) that we can find a version of L^a that is $\mathcal{A}\otimes\mathcal{O}$-measurable. We always choose this measurable version of the local time.

The next theorem is quite simple yet crucial to proving the properties of L^a that justify its name.

(4.10) Theorem. Let X be a semimartingale and let L^a be its local time at a. Then

(4.11) $(X_t-a)^+ = (X_0-a)^+ + \int_0^t 1_{(X_{s-} > a)}dX_s + \sum_{0<s\le t} 1_{(X_{s-} > a)}(X_s-a)^-$
$\qquad + \sum_{0<s\le t} 1_{(X_{s-} \le a)}(X_s-a)^+ + \frac{1}{2} L_t^a$

(4.12) $(X_t-a)^- = (X_0-a)^- - \int_0^t 1_{(X_{s-} \le a)}X_s + \sum_{0<s\le t} 1_{(X_{s-}>a)}(X_s-a)^-$
$\qquad + \sum_{0<s\le t} 1_{(X_{s-}\le a)}(X_s-a)^+ + \frac{1}{2} L_t^a$

Proof. Applying (4.1) to the convex functions $f(x) = (x-a)^+$ and $g(x) = (x-a)^-$ we get

$$g(X_t) = g(X_0) + \int_0^t g'(X_{s-})dX_s + B_t^-$$

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + B_t^+ .$$

Next let

$$C_t^- = B_t^- - \sum_{0<s\le t}\{g(X_s) - g(X_{s-}) - g'(X_{s-})\Delta X_s\}$$

$$C_t^+ = B_t^+ - \sum_{0<s\le t}\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\},$$

and subtracting the formulas we get $B_t^+ - B_t^- = 0$ and hence $C_t^+ = C_t^-$; also $C_t^+ + C_t^- = L_t^a$; thus $C_t^+ = C_t^- = \frac{1}{2} L_t^a$, and the proof is complete. □

The next theorem, together with the "occupation time density" formula, are the traditional justifications for the terminology "local time".

(4.13) Theorem. For a.a. ω, the measure in t, $dL_t^a(\omega)$, is carried by the set $\{s: X_{s-}(\omega) = X_s(\omega) = a\}$.

Proof. Since L_t^a has continuous paths, the measure $dL_t^a(\omega)$ is diffuse, and since $\{s: X_{s-}(\omega) = a\}$ and $\{s: X_{s-}(\omega) = X_s(\omega) = a\}$ differ by at most a countable set, it will suffice to show that $dL_t^a(\omega)$ is carried by the set $\{s: X_{s-}(\omega) = a\}$.

Suppose S, T are stopping times and that $0 < S \le T$ such that $[S,T[\subset \{X_- < a\}$. Then $X \le a$ on $[S,T[$ as well, hence by (4.11) we have

$$(X-a)_T^+ - (X-a)_S^+ = \int_S^T 1_{(X_{s-} > a)} dX_s + \sum_{S<s\le T} 1_{(X_{s-} > a)}(X_s-a)^- + \sum_{S<s\le T} 1_{(X_{s-} < a)}(X_s-a)^+ + \frac{1}{2}(L_T^a - L_S^a).$$

Eliminating the terms that must be zero yields:

$$(X-a)_T^+ = (X_T-a)^+ + \frac{1}{2}(L_T^a - L_S^a)$$

whence $\frac{1}{2}(L_T^a - L_S^a) = 0$ and $L_T^a = L_S^a$.

Next for $r \in \mathbb{Q}$, the rationals, define the stopping times $S_r(\omega)$, $r > 0$, by

$$S_r(\omega) = \begin{cases} r & \text{if } X_{r-}(\omega) < a \\ \infty & \text{if } X_{r-}(\omega) \geq a \end{cases} .$$

Then define

$$T_r(\omega) = \inf\{t > S_r(\omega) : X_{t-}(\omega) \geq a\}.$$

Then $[S_r, T_r[\subset \{X_- < a\}$, and moreover the interior of the set $\{X_- < a\}$ equals $\bigcup_{\substack{r\in\mathbb{Q}\\ r>0}}]S_r, T_r[$. As we have now seen, dL^a does not charge the interior of the set $\{X_- < a\}$. This set is open on the left, hence it differs from its interior by an at most countable set, and since dL^a is diffuse it doesn't charge countable sets. Thus dL^a does not charge the set $\{X_- < a\}$ itself.

Analogously one can show dL^a does not charge $\{X_- > a\}$. Hence its support is contained in the set $\{X_- = a\}$, and we are done. □

The next theorem gives a very satisfying generalization of Itô's lemma (III.7.1)

(4.14) Theorem. Let f be the difference of two convex functions, let f' be its left derivative, and let μ be the signed measure (when restricted to compacts) which is the second derivative of f in the generalized function sense. Then the following equation holds:

$$(4.15)\quad f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + \sum_{0<s\leq t}\{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\} + \frac{1}{2}\int_{-\infty}^{\infty}\mu(da)L_t^a$$

Proof. First note that for each fixed (t,ω) the function $a \to L_t^a(\omega)$ has compact support as a consequence of (4.13) and that semimartingales have cadlag paths which are hence bounded on $[0,t]$ for each ω. Therefore the last integral on the right side of (4.15) has meaning.

For simplicity let us assume that μ is a signed measure with finite total mass. Define a function g by

$$(4.16)\quad g(x) = \int |x-y|\mu(dy).$$

It is well known that f and g differ by at most an affine function $h(x) = ax + b$. Equation (4.15) is clearly true for affine functions (e.g. III.7.1), so it suffices to establish (4.15) for functions of the form (4.16), and we assume f is of the form $f(x) = \int |x-y|\mu(dy)$. Then $f'(x) = \int \text{sign}\,(x-y)\mu(dy)$ and hence if $J_t^y = \sum_{0<s\leq t} |X_s - y| - |X_{s-} - y| - \text{sign}\,(X_{s-} - y)\Delta X_s$, then

(4.17) $\int J_t^y \mu(dy) = \sum_{0<s\leq t} \{f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s\}$.

Also, letting $H_t^y = |X_t - y| - |X_0 - y|$, one has

(4.18) $\int H_t^y \mu(dy) = f(X_t) - f(X_0)$.

Next consider $Z_t^y = \int_0^t \text{sign}\,(X_{s-} - y)dX_s$, and let Z be the $\mathcal{A}\otimes\mathcal{O}$ measurable version, which we know exists (3.4). By the Fubini lemma (3.6) we have

(4.19) $\int Z_t^y \mu(dy) = \int_0^t f(X_{s-})dX_s$.

Since $\frac{1}{2} L_t^a = H_t^a - J_t^a - Z_t^a$, we have

(4.20) $\frac{1}{2} \int L_t^a \mu(da) = \int H_t^a \mu(da) - \int J_t^a \mu(da) - \int Z_t^a \mu(da)$

and combining (4.20) with (4.17),(4.18) and (4.19) yields the formula (4.15) and the proof is complete. □

The next formula gives an interpretation of semimartingale local time as an occupation density relative to the random "clock" $d[X,X]_s^c$.

(4.21) Corollary. Let X be a semimartingale with local time $(L^a)_{a\in\mathbb{R}}$. Let g be a bounded Borel measurable function. Then a.s.

$$\int_{-\infty}^{\infty} L_t^a g(a)da = \int_0^t g(X_{s-})d[X,X]_s^c \quad .$$

Proof. Let f be convex and $\mathcal{C}^2$. Comparing (4.15) with Itô's lemma III.7.1 shows that:

(4.22) $$\int_{-\infty}^{\infty} L_t^a f''(a)da = \int_0^t f''(X_{s-})d[X,X]_s^c$$

where $\mu(da)$ is of course f"(a)da. Since (4.22) holds for any continuous and positive function f", a monotone class argument shows that (4.22) must hold, up to a P-null set, for any bounded, Borel measurable function g. □

We record here an important special case of (4.21).

(4.23) Corollary. Let X be a semimartingale with local time $(L^a)_{a\in\mathbb{R}}$. Then

$$[X,X]_t^c = \int_{-\infty}^{\infty} L_t^a \, da.$$

We conclude with the useful Meyer-Tanaka formula.

(4.24) Corollary. Let X be a semimartingale with continuous paths. Then

$$|X_t| = |X_0| + \int_0^t \text{sign}\,(X_s)dX_s + L_t^0.$$

Proof. This is merely (4.14) with f(x) = |x|, which implies $\mu(da) = \varepsilon_0(da)$, point mass at 0. □

Observe that if f(x) = |x| , then f" = $\delta(x)$, the "delta function at 0", which of course is a generalized function, or "distribution". Thus (4.24) gives the intuitive interpretation of local times as $L_t^0 = \int_0^t \delta(X_s)dX_s$, and $L_t^a = \int_0^t \delta(X_s - a)dX_s$, for continuous semimartingales.

VI. STOCHASTIC DIFFERENTIAL EQUATIONS

1. Introduction

When K. Itô first defined the stochastic integral relative to the Wiener process, his chief purpose was to study stochastic differential equations. By considering a system of equations of the form ($1 \leq i \leq n$):

$$X_t^i = X_0^i + \sum_{j=1}^{d} \int_0^t f_j^i(s,X_s)dW_s^j + \int_0^t g^i(s,X_s)ds, \text{ where } W = (W^1,\ldots,W^d) \text{ is a}$$

d-dimensional Wiener process (or "Brownian motion"), Itô proved under Lipschitz hypotheses on f and g that a continuous, unique solution existed, and that it was a continuous strong Markov process, time homogeneous if f and g were autonomous. This provided a probabilistic method of studying multidimensional diffusions. The previous methods had relied on an analysis of the infinitesimal generators of their transition semigroups. These are partial differential operators, and hence this lead to a study of elliptic partial differential equations and a concomitant obfuscation of the probabilistic content.

The differential "dW" however also has an interpretation as "white noise" in statistical communication theory. Here the Markov nature of the solutions is not the issue and one considers more general coefficients that have a non-anticipating dependence on the history as well as current state of the solution. Furthermore, with the advent of the semimartingale integral the more general "semimartingale noise" can be taken as a driving term. It is these more general equations that will be considered in this chapter.

2. Norms for Semimartingales

Recall that $\mathbb{D}$ represents the space of all adapted processes with cadlag paths. For $H \in \mathbb{D}$, let $H^* = \sup_t |H_t|$, and define

(2.1) $\qquad ||H||_{\underline{\underline{S}}^p} = ||H^*||_{L^p} \qquad\qquad (1 \leq p \leq \infty).$

For $H \in \mathbb{D}$, the left continuous version H_- of H is in $\mathbb{L}$, and we set $||H_-||_{\underline{\underline{S}}^p} = ||H||_{\underline{\underline{S}}^p}$. If A is a semimartingale with paths of finite variation, a natural definition of a norm would be $||A|| = ||\int_0^\infty |dA_s|\ ||_{L^p}$, where $|dA_s(\omega)|$ denotes the total variation measure on $\mathbb{R}_+$ induced by $s \to A_s(\omega)$. Since semimartingales do not in general have such nice paths, however, such a norm is not appropriate. Instead we use the norm given in V.2.3, which was used to extend the space of integrands of the semimartingale integral. Our definition here is slightly different, but it can be shown to be equivalent (cf [36]). Let Z be a semimartingale. By IV.5.6 we know Z is decomposable. Let $Z = Z_0 + N + A$ be any decomposition where $N_0 = A_0 = 0$. Define for $1 \leq p \leq \infty$:

(2.2) $\qquad j_p(N,A) = ||[N,N]_\infty^{\frac{1}{2}} + |Z_0| + \int_0^\infty |d\,A_s|\ ||_{L^p}.$

(2.3) Definition. Let Z be a semimartingale. For $1 \leq p \leq \infty$ define

$$||Z||_{\underline{\underline{H}}^p} = \inf_{Z=Z_0+N+A} j_p(N,A)$$

where the infimum is taken over all possible decompositions $Z = Z_0 + N + A$ where N is a local martingale, $A \in \mathbb{D}$ with paths of finite variation on compacts, and $A_0 = N_0 = 0$.

Corollary (2.5) below shows that this norm generalizes the $\underline{\underline{H}}^p$ norm for local martingales, which has given rise to a martingale theory analogous to the theory of Hardy spaces in complex analysis. We do not pursue this topic (cf, e.g., [33]).

(2.4) Theorem. Let Z be a semimartingale. Then $\|[Z,Z]_\infty^{\frac{1}{2}}\|_{L^p} \leq \|Z\|_{\underline{\underline{H}}^p}$ $(1 \leq p \leq \infty)$.

Proof. We may assume without loss that $Z_0 = 0$ and let $Z = M + A$, $M_0 = A_0 = 0$, be a decomposition of Z. Then

$$\begin{aligned}
[Z,Z]_\infty^{\frac{1}{2}} &\leq [M,M]_\infty^{\frac{1}{2}} + [A,A]_\infty^{\frac{1}{2}} \\
&= [M,M]_\infty^{\frac{1}{2}} + (\sum(\Delta A_s)^2)^{\frac{1}{2}} \\
&\leq [M,M]_\infty^{\frac{1}{2}} + \sum_s |\Delta A_s| \\
&\leq [M,M]_\infty^{\frac{1}{2}} + \int_0^\infty |dA_s| ,
\end{aligned}$$

with the equality above a consequence of III.6.11. Taking L^p norms yields $\|[Z,Z]_\infty^{\frac{1}{2}}\|_{L^p} \leq j_p(M,A)$ and the result follows. □

(2.5) Corollary. If Z is a local martingale then $\|Z\|_{\underline{\underline{H}}^p} = \|[Z,Z]_\infty^{\frac{1}{2}}\|_{L^p}$.

Proof. Without loss assume $Z_0 = 0$. Then $Z = 0 + Z + 0$ is a decomposition of Z, hence $\|Z\|_{\underline{\underline{H}}^p} \leq j_p(Z,0) = \|[Z,Z]_\infty^{\frac{1}{2}}\|_{L^p}$. By (2.4) we have $\|[Z,Z]_\infty^{\frac{1}{2}}\|_{L^p} \leq \|Z\|_{\underline{\underline{H}}^p}$, whence equality. □

We do not have need of the $\underline{\underline{H}}^p$ norm for all $p(1 \leq p \leq \infty)$; we consider next the important case $p = 2$.

(2.6) Theorem. Let Z be a semimartingale. Then $\|Z\|_{\underline{\underline{S}}^2} \leq \sqrt{8}\, \|Z\|_{\underline{\underline{H}}^2}$.

Proof. A semimartingale $Z \in \mathbb{D}$, so $\|Z\|_{\underline{\underline{S}}^2}$ makes sense. Without loss assume $Z_0 = 0$, and let $Z = M + A$ be a decomposition with $M_0 = A_0 = 0$. Then

(2.7) $\quad ||Z||^2_{\underline{\underline{S}}^2} = E\{(Z^*_\infty)^2\} \leq E\{(M^*_\infty + \int|dA_s|)^2\} \leq 2E\{(M^*_\infty)^2 + (\int|dA_s|)^2\}$,

using $(a + b)^2 \leq 2a^2 + 2b^2$. Doob's inequality says that $E\{(M^*_\infty)^2\} \leq 4E\{(M^2_\infty)\}$ $= 4E\{[M,M]_\infty\}$, the last being a simple consequence of III.6.2. Continuing (2.7) we have

$$\begin{aligned} ||Z||^2_{\underline{\underline{S}}^2} &\leq 2E\{(M^*_\infty)^2 + (\int|dA_s|)^2\} \\ &\leq 8E\{[M,M]_\infty + (\int|dA_s|)^2\} \\ &\leq 8E\{([M,M]^{\frac{1}{2}}_\infty + \int|dA_s|)^2\} \\ &\leq 8[j_2(M,A)]^2 , \end{aligned}$$

and taking square roots yields the result. □

(2.8) Corollary. On the space of semimartingales the $\underline{\underline{H}}^2$ norm is stronger than the $\underline{\underline{S}}^2$ norm.

(2.9) Theorem. Let Z be a semimartingale, $D \in \mathbb{D}$, and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ $(1 \leq p \leq \infty;\ 1 \leq q \leq \infty)$. Then

$$||\int D_{s-}\, dZ_s||_{\underline{\underline{H}}^r} \leq ||D||_{\underline{\underline{S}}^p}\ ||Z||_{\underline{\underline{H}}^q}$$

Proof. Without loss assume $Z_0 = 0$, and let $D_- \cdot Z$ denote $\int D_{s-} dZ_s$. If $Z = M + A$, $M_0 = A_0 = 0$ is a decomposition of Z, then $D_- \cdot M + D_- \cdot A$ is a decomposition of $D_- \cdot Z$, hence $||D_- \cdot Z||_{\underline{\underline{H}}^r} \leq j_r(D_- \cdot M, D_- \cdot A)$. Also $[D_- \cdot M, D_- \cdot M]$ $= \int D^2_{s-} d[M,M]_s$ (III.6.16), hence

$$\begin{aligned} j_r(D_- \cdot M, D_- \cdot A) &= ||(\int D^2_{s-} d[M,M]_s)^{\frac{1}{2}} + \int|D_{s-}||dA_s|\,||_{L^r} \\ &\leq ||D^*([M,M]^{\frac{1}{2}}_\infty + \int_0^\infty |dA_s|)||_{L^r} \end{aligned}$$

$$\leq ||D^*_\infty||_{L^p} || [M,M]^{\frac{1}{2}}_\infty + \int_0^\infty |dA_s| \; ||_{L^q}$$

$$= ||D||_{\underline{\underline{S}}^p} \; j_q(M,A),$$

with the last inequality following from Hölder's inequality. The above implies that

$$||D_- \cdot Z||_{\underline{\underline{H}}^r} \leq ||D||_{\underline{\underline{S}}^p} \; j_q(M,A)$$

for any such decomposition. $Z = M + A$. Taking infinums over all such decompositions yields the result. □

For a process $X \in \mathbb{D}$ and a stopping time T, recall that:

$$X^T = X_t 1_{[0,T[} + X_T 1_{[T,\infty[}$$

$$X^{T-} = X_t 1_{[0,T[} + X_{T-} 1_{[T,\infty[} \quad .$$

(2.10) Definition. A process X is locally (respectively prelocally) in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$) if there exists a sequence of stopping times $(T_k)_{k\geq 1}$ increasing to ∞ a.s. such that X^{T_k} (resp. X^{T_k-}) is in $\underline{\underline{S}}^p$ (resp. $\underline{\underline{H}}^p$), $1 \leq p \leq \infty$.

While there are many semimartingales which are not locally in $\underline{\underline{H}}^p$, all semimartingales are prelocally in $\underline{\underline{H}}^p$.

(2.11) Theorem. Let Z be a semimartingale. Then $Z - Z_0$ is prelocally in $\underline{\underline{H}}^p$, all p, $1 \leq p \leq \infty$.

Proof. Without loss assume $Z_0 = 0$ and let $Z = M + A$ be any decomposition, such that M has jumps bounded by a constant ε(IV.4.7). Define inductively:

$$T_0 = 0;$$

$$T_{k+1} = \inf\{t \geq T_k : [M,M]^{\frac{1}{2}}_t + \int_0^t |dA_s| \geq k+1\} \; .$$

The sequence $(T_k)_{k\geq 1}$ are stopping times increasing to ∞ a.s. Moreover

(2.12) $$Z^{T_k-} = (M^{T_k}) + (A^{T_k-} - \Delta M_{T_k} 1_{[T_k,\infty[}) = N + B$$

is a decomposition of Z^{T_k-}. Also, since $[M,M]_{T_k} = [M,M]_{T_k-} + (\Delta M_{T_k})^2$, equation (2.12) yields:

$$\begin{aligned} j_\infty(N,B) &= \| [N,N]_\infty^{\frac{1}{2}} + \int |dB_s| \|_{L^\infty} \\ &= \| ([M,M]_{T_k-} + (\Delta M_{T_k})^2)^{\frac{1}{2}} + \int |dB_s| \|_{L^\infty} \\ &\leq \| (k^2 + \varepsilon^2)^{\frac{1}{2}} + (k + \varepsilon) \|_{L^\infty} < \infty . \end{aligned}$$

Therefore $Z^{T_k-} \in \underline{\underline{H}}^\infty$ and hence it is in $\underline{\underline{H}}^p$ as well, $1 \leq p \leq \infty$. □

(2.15) Definition. Let Z be a semimartingale in $\underline{\underline{H}}^\infty$ and let $\alpha > 0$. A finite sequence of stopping times $0 \leq T_0 \leq T_1 \leq \ldots \leq T_k$ is said to α-slice Z if $Z = Z^{T_k-}$ and $\|(Z - Z^{T_i})^{T_{i+1}-}\|_{\underline{\underline{H}}^\infty} \leq \alpha$, $0 \leq i \leq k-1$. If such a sequence of stopping times exists, we say Z is α-sliceable, and we write $Z \in \mathcal{S}(\alpha)$.

(2.14) Theorem. Let Z be a semimartingale.

(i) For $\alpha > 0$, if $Z \in \mathcal{S}(\alpha)$ then for every stopping time T, $Z^T \in \mathcal{S}(\alpha)$ and $Z^{T-} \in \mathcal{S}(2\alpha)$.

(ii) For every $\alpha > 0$, there exists an arbitrarily large stopping time T such that $Z^{T-} \in \mathcal{S}(\alpha)$.

Proof. Always $\|Z^T\|_{\underline{\underline{H}}^\infty} \leq \|Z\|_{\underline{\underline{H}}^\infty}$, and since $Z^{T-} = M^T + (A^{T-} - \Delta M_T 1_{[T,\infty[})$ one concludes $||Z^{T-}||_{\underline{\underline{H}}^\infty} \leq 2||Z||_{\underline{\underline{H}}^\infty}$, and part (i) follows.

Next consider (ii). If semimartingales Z and Y are α-sliceable, then by (i) Z + Y is 2α-sliceable. Without loss assume $Z_0 = 0$, and let $Z = M + A$, $M_0 = A_0$, where M has bounded jumps. We consider M and A separately.

For A, let $T_0 = 0$, $T_{k+1} = \inf\{t \geq T_k\colon \int_{T_k}^{t} |dA_s| \geq \alpha/2$ or $\int_0^{t-} |dA_s| \geq k\}$. Then $A^{T_k-} \in \mathcal{S}(\alpha/2)$ for each k, and the stopping times (T_k) increase to ∞ a.s.

For M, let $R_0 = 0$, $R_{k+1} = \inf\{t \geq R_k\colon [M,M]_t - [M,M]_{R_k} \geq \beta^2$ or $[M,M]_t \geq k\}$. Then $M^{R_k-} \in \underline{\underline{H}}^\infty$, each k, and moreover:

$$(M - M^{R_k})^{R_k-} = M^{R_{k+1}} - M^{R_k} - (\Delta M_{R_{k+1}} 1_{\{R_{k+1} > R_k\}}) 1_{[R_k,\infty[}$$

hence

$$||(M - M^{R_k})^{R_{k+1}-}||_{\underline{\underline{H}}^\infty}$$

$$\leq ||([M,M]_{R_{k+1}} - [M,M]_{R_k})^{\frac{1}{2}} + |\Delta M_{R_{k+1}}|\,||_{L^\infty}$$

$$= ||((\Delta M_{R_{k+1}})^2 + [M,M]_{R_{k+1}-} - [M,M]_{R_k})^{\frac{1}{2}} + |\Delta M_{R_{k+1}}|\,||_{L^\infty}$$

$$\leq ||(\beta^2 + \beta^2)^{\frac{1}{2}} + \beta||_{L^\infty} = (1 + \sqrt{2})\beta\ .$$

Thus for each k $M^{R_k-} \in \mathcal{S}((1 + \sqrt{2})\beta)$, and we take $\beta = \dfrac{\alpha}{2+2\sqrt{2}}$ and the result follows. □

3. Existence and Uniqueness of Solutions

We begin by stating a theorem whose main virtue is its simplicity. It is, of course, a trivial corollary of Theorem (3.10).

(3.1) Theorem. Let Z be a semimartingale and let $f: \mathbb{R}_+ \times \Omega \times \mathbb{R} \to \mathbb{R}$ such that:

(i) for fixed x, $(t,\omega) \to f(t,\omega,x)$ is in $\mathbb{L}$;

(ii) for each (t,ω), $|f(t,\omega,x) - f(t,\omega,y)| \leq K(\omega)\,|x-y|$ for some finite random variable K.

Let X_0 be finite and $\mathcal{F}_0$-measurable. Then the equation

$$X_t = X_0 + \int_0^t f(s, \cdot, X_{s-})dZ_s$$

admits a solution. The solution is unique and it is a semimartingale.

Of course one could state such a theorem for a finite number of differentials dZ^j $(1 \leq j \leq d)$ and for a finite system of equations. The next definition describes the coefficients we will consider. $\mathbb{D}^n$ consists of processes $\underset{\sim}{X} = (X^1,\dots,X^n)$ where each $X^i \in \mathbb{D}$ $(1 \leq i \leq n)$.

(3.2) Definition. Let $n \geq 1$ and $F: \mathbb{D}^n \to \mathbb{D}$. F is said to be $\mathbb{D}^n$-Lipschitz if

(i) for any stopping time T and $\underset{\sim}{X}, \underset{\sim}{Y} \in \mathbb{D}^n$, $\underset{\sim}{X}^{T-} = \underset{\sim}{Y}^{T-}$ implies $F(\underset{\sim}{X})^{T-} = F(\underset{\sim}{Y})^{T-}$;

(ii) $(F(\underset{\sim}{X}) - F(\underset{\sim}{Y}))^*_t \leq K(\omega)||\underset{\sim}{X} - \underset{\sim}{Y}||^*_t$

The above definition is a slight modification of one first proposed by M. Emery [18].

(3.3) Lemma. Let $J \in \underline{\underline{S}}^2$, let F be $\mathbb{D}$-Lipschitz and suppose $F(0) = 0$, and that $K(\omega) \leq c$ a.s.. Let Z be a semimartingale in $\underline{\underline{H}}^\infty$ such that $||Z||_{\underline{\underline{H}}^\infty} \leq \frac{1}{2\sqrt{8c}}$.

Then the equation

$$X_t = J_t + \int_0^t F(X)_{s-}\, dZ_s$$

has a solution in $\underline{\underline{S}}^2$, it is unique, and moreover

$$||X||_{\underline{\underline{S}}^2} \leq 2||J||_{\underline{\underline{S}}^2}$$

Proof. Define $\Lambda: \underline{\underline{S}}^2 \to \underline{\underline{S}}^2$ by $\Lambda(X)_t = J_t + \int_0^t F(X)_{s-}dZ_s$. Then by (2.9) and (2.6) the operator is $\frac{1}{2}$ - Lipschitz, and the fixed point theorem gives existence and uniqueness. Indeed

$$\begin{aligned} ||X||_{\underline{\underline{S}}^2} &\leq ||J||_{\underline{\underline{S}}^2} + ||\int F(X)_{s-}dZ_s||_{\underline{\underline{S}}^2} \\ &\leq ||J||_{\underline{\underline{S}}^2} + \sqrt{8}\, ||F(X)||_{\underline{\underline{S}}^2} \, ||Z||_{\underline{\underline{H}}^\infty} ; \\ &\leq ||J||_{\underline{\underline{S}}^2} + \frac{1}{2c} ||F(X)||_{\underline{\underline{S}}^2} . \end{aligned}$$

Since $||F(X)||_{\underline{\underline{S}}^2} = ||F(X) - F(0)||_{\underline{\underline{S}}^2} \leq c||X||_{\underline{\underline{S}}^2}$, we have

$||X||_{\underline{\underline{S}}^2} \leq ||J||_{\underline{\underline{S}}^2} + \frac{1}{2} ||X||_{\underline{\underline{S}}^2}$, which yields the estimate. □

(3.4) Lemma. Let $J \in \underline{\underline{S}}^2$, F $\mathbb{D}$-Lipschitz with $F(0) = 0$ and $K(\omega) \leq c < \infty$ a.s.. Let Z be a semimartingale such that $Z \in \mathcal{S}(\frac{1}{2\sqrt{8}c})$. Then the equation

$$X_t = J_t + \int_0^t F(X)_{s-}dZ_s$$

has a solution in $\underline{\underline{S}}^2$, it is unique, and moreover $||X||_{\underline{\underline{S}}^2} \leq C(c,Z)||J||_{\underline{\underline{S}}^2}$, where $C(c,Z)$ is a constant depending only on c and Z.

Proof. Let $z = ||Z||_{\underline{\underline{H}}^\infty}$ and $j = ||J||_{\underline{\underline{S}}^2}$. Let $T_0, T_1, \ldots, T_k$ be the slicing times for Z, and consider the equations

(3.5(i)): $X = J^{T_i-} + \int F(X)_{s-}dZ_s^{T_i-}$. Equation (3.5(0)) has the trivial solution $X \equiv 0$, and its $\underline{\underline{S}}^2$ norm is 0. Assume that equation (3.5(i)) has a unique solution X^i, and let $x^i = ||X^i||_{\underline{\underline{S}}^2}$.

Stopping next at T_i instead of T_i- , let Y^i denote the unique solution of

(3.6(i)): $Y^i = J^{T_i} + \int F(Y^i)_{s-} dZ_s^{T_i}$, and set $y^i = ||Y^i||_{\underline{\underline{S}}^2}$. Since

$Y^i = X^i + \{\Delta J_{T_i} + F(X^i)_{T_i-} \Delta Z_{T_i}\} 1_{[T_i,\infty[}$, we conclude that

$$||Y^i||_{\underline{\underline{S}}^2} \leq ||X^i||_{\underline{\underline{S}}^2} + 2||J||_{\underline{\underline{S}}^2} + \sqrt{8}||F(X^i)||_{\underline{\underline{S}}^2}||Z||_{\underline{\underline{H}}^\infty}$$
$$\leq x^i + 2j + \sqrt{8}\, c\, x^i z$$
$$= x^i(1 + \sqrt{8}\, c\, z) + 2j ;$$

hence

(3.7) $y^i \leq 2j + x^i(1 + \sqrt{8}\, c.\, z)$.

We set for $U \in \mathbb{D}$, $D_i U = (U - U^{T_i})^{T_{i+1}-}$. Since each solution X of (3.5(i+1)) satisfies $X^{T_i} = Y^i$ on $[0, T_{i+1}[$, we can change the unknown by $U = X - (Y^i)^{T_{i+1}-}$, to get the equations:

(3.8(i)) $U = D_i J + \int F(Y^i + U)_{s-} d\, D_i Z_s$; however $F(0) \neq 0$, and so we define $G^i(\cdot) = F(Y^i + \cdot\,) - F(Y^i)$, and thus (3.8(i)) can be equivalently expressed as:

(3.9)i)) $U = (D_i J + \int F(Y^i)_{s-}\, d\, D_i Z_s) + \int G_i(U)_{s-} d\, D_i Z_s$.

We can now apply Lemma (3.3) to (3.9(i)), and thus it has a unique solution in $\underline{\underline{S}}^2$, and its norm u^i is majorized by $u^i \leq 2||D_i J + \int F(Y^i)_{s-} d\, D_i Z_s||_{\underline{\underline{S}}^2}$

$$\leq 2(2j + \sqrt{8}\, c y^i \frac{1}{2\sqrt{8}c}) \leq 4j + y^i .$$

Thus we conclude equation (3.5(i+1)) has a unique solution in $\underline{\underline{S}}^2$ with norm x^{i+1} dominated by (using (3.7)):

$$x^{i+1} \le y^i + y^i \le 4j + 2y^i \le 8j + 2(1+\sqrt{8}cz)x^i .$$

We next iterate from i = 0 to k-1 and conclude that $x^k \le 8\{\frac{(2+2\sqrt{8}c\,z)^k - 1}{1+\sqrt{8}c\,z}\}j$.

Finally, since $Z = Z^{T_k-}$, we have seen that the equation $X = J + \int F(X)_{s-}dZ_s$ has a unique solution in $\underline{\underline{S}}^2$, and moreover $X = X^k + J - J^{T_k-}$; thus

$||X||_{\underline{\underline{S}}^2} \le x^k + 2j$, hence $C(c,Z) \le 2 + 8\{\frac{(2+2\sqrt{8}cz)^k - 1}{1+\sqrt{8}cz}\}$. □

(3.10) Theorem. *Given a vector of semimartingales* $\underset{\sim}{Z} = (Z^1, \dots, Z^d)$, *processes* $J^i \in \mathbb{D}$ $(1 \le i \le n)$, *and operators* F^i_j *which are* $\mathbb{D}^n$*-Lipschitz* $(1 \le i \le n;\ 1 \le j \le d)$, *then the system of equations*

$$X^i_t = J^i_t + \sum_{j=1}^{d} \int_0^t F^i_j(\underset{\sim}{X})_{s-}\, dZ^j_s$$

$(1 \le i \le n)$ *has a solution in* $\mathbb{D}^n$, *and it is unique. Moreover if* $(J^i)_{i \le n}$ *is a vector of semimartingales, then so is* $(X^i)_{i \le n}$.

Proof. The proof for systems is the same as the proof for one equation but with more cumbersome notation; hence we give here the proof for n = d = 1. Thus we will consider the equation

$$(3.11) \qquad X_t = J_t + \int_0^t F(X)_{s-}dZ_s .$$

Assume that $\max_{i,j} K^i_j(\omega) \le c < \infty$ a.s.. Also, by considering the equation:

$$(3.12) \qquad X_t = \{J_t + \int_0^t F(0)_{s-}dZ_s\} + \int_0^t G(X)_{s-}dZ_s ,$$

where G(X) = F(X) - F(0), it suffices to consider the case where F(0) = 0.

Let T be an arbitrarily large stopping time such that $J^{T-} \in \underline{\underline{S}}^2$ and such that $Z^{T-} \in \mathcal{S}(\frac{1}{4\sqrt{8}c})$. Then by Lemma (3.4) there exists a unique solution in $\underline{\underline{S}}^2$ of:

$$X(T)_t = J_t^{T-} + \int_0^t F(X(T))_{s-} dZ_s^{T-} .$$

By the uniqueness in $\underline{\underline{S}}^2$ one has, for $R > T$, that $X(R)^{T-} = X(T)^{T-}$, and thus one can define a process X on $\Omega \times [0,\infty[$ by $X = X(T)$ on $[0,T[$. Thus we have existence.

Suppose next Y is another solution. Let S be arbitrarily large such that $(X - Y)^{S-}$ is bounded, and let $R = \min(S,T)$, which can also be taken arbitrarily large. Then X^{R-} and Y^{R-} are both solutions of

$$U = J^{R-} + \int_0^t F(U)_{s-} dZ_s^{R-} ,$$

and since $Z^{R-} \in \mathcal{S}(\frac{1}{2\sqrt{8}c})$, we know that $X^{R-} = Y^{R-}$ by the uniqueness established in Lemma (3.4). Thus $X = Y$, and we have uniqueness.

We have assumed that $K(\omega) \leq c < \infty$ a.s.. Suppose instead that $K(\omega) < \infty$ a.s.. Since $K(\omega) < \infty$ a.s., choose a constant c such that $P\{K \leq c\} > 0$. Let $\Omega_n = \{K \leq c + n\}$, each $n = 1,2,3,\dots$. Define a new probability P_n by $P_n(A) = P(A \cap \Omega_n)/P(\Omega_n)$, on the space Ω_n equipped with the filtration $\mathcal{F}_t^n = \mathcal{F}_t|_{\Omega_n}$, the trace of $\mathcal{F}_t$ on Ω_n. Then it is a simple consequence of the definition of a semimartingale (III.1.1) that the restriction of Z in equation (3.11) to $\Omega_n \times [0,\infty[$ is an $(\mathcal{F}_t^n)_{0 \leq t < \infty}$-semimartingale. Let Y^n be the unique solution on Ω_n that we have seen exists. For $m > n$ we have $\Omega_m \supseteq \Omega_n$ and $P_n \ll P_m$. Hence by (III.5.3) we know P_m-stochastic integrals are indistinguishable from P_n-stochastic integrals, whence Y^m restricted to $\Omega_n \times [0,\infty[$ is a solution on $(\Omega_n, \mathcal{F}_t^n)$. Thus Y^m is P_n-indistinguishable from Y^n, and we can thus define a solution Y on $\mathbb{R}_+ \times \Omega$ by setting $Y = Y^n$ on $\Omega_n = \{\omega: K(\omega) \leq c + n\}$. □

(3.13) Definition. Let $n \geq 1$ and $F: \mathbb{D}^n \to \mathbb{D}$. F is said to be locally $\mathbb{D}^n$-Lipschitz if

(i) for any stopping time T and X, Y in $\mathbb{D}^n$, $X^{T-} = Y^{T-}$ implies $F(X)^{T-} = F(Y)^{T-}$;

(ii) there exists a process $K(t,\omega)$ such that $\sup_{t \in \Lambda} K(t,\omega) < \infty$ a.s. for compact sets Λ and such that $(F(X) - F(Y))^*_t \leq K(t,\cdot)\|X - Y\|^*_t$.

(3.14) Definition. A stopping time T is called an explosion time for a process X if

(i) for any stopping time $S < T$ a.s., $X1_{[0,S[} \in \mathbb{D}^n$;

(ii) $\limsup_{\substack{t \to T \\ t < T}} \|X_t\| = \infty$ a.s.

(3.15) Theorem. Given a vector of semimartingales $Z = (Z^1, \ldots, Z^d)$, processes $J^i \in \mathbb{D}(1 \leq i \leq n)$ and operators F^i_j which are locally $\mathbb{D}^n$ Lipshitz $(1 \leq i \leq n;\ 1 \leq j \leq d)$, then there exists a stopping time T and a process X such that if $S < T$ then $X1_{]0,S[} \in \mathbb{D}^n$ and such that X is a solution of

(3.16) $$X^i_t = J^i_t + \sum_{j=1}^{d} \int_0^t F^i_j(X)_{s-}\,dZ^j_s$$ on $[0,S]$. Moreover there is a maximal time T such that $T = \infty$ or T is an explosion time for X.

Proof. Let $f_k: \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^\infty$ functions with compact support and which take values between 0 and 1, and such that the interiors of the compact sets $\Lambda_k = \{x: f_k(x) = 1\}$ increase to $\mathbb{R}^n$; that is, $\bigcup_{k=1}^{\infty} \Lambda_k = \mathbb{R}^n$. For each (i,j) define a new coefficient $G^i_{j,k}$ by $G^i_{j,k}(X) = f_k(X)F^i_j(X)$. Then $G^i_{j,k}$ is $\mathbb{D}^n$-Lipschitz. Let X^k denote the solution of (3.16) with $G^i_{j,k}$ replacing F^i_j, and define stopping times S_k by:

$$S_k(\omega) = \inf\{t: X_t^k(\omega) \notin \Lambda_k\} .$$

By the uniqueness established in (3.10) $X^k = X^{k+1}$ on $[0,S_k[$, hence $S_k \leq S_{k+1}$ a.s., and we set $T = \lim_{k\to\infty} S_k$. Using uniqueness define $X(t,\omega)$ by:

$$X(t,\omega) = X^k(t,\omega) \text{ on } [0,S_k[, \text{ each } k.$$

Then X satisfies (3.16) on $[0,T[$.

We conclude by showing that either $X_{S_k}(\omega)$ or $X_{S_k-}(\omega)$ is in the complement of Λ_k and thus $\limsup_{\substack{t\to T\\ t<T}} ||X_t|| = \infty$ a.s. First suppose that $X_{S_k-}(\omega) = \lambda \in \Lambda_k$. Then $X_{S_k(\omega)-}(\omega) = X^k_{S_k(\omega)-}(\omega)$, since $f_k(\lambda) = 1$. The processes X and X^k then have the same jump at $S_k(\omega)$ and hence they have the same value at time $S_k(\omega)$; that is, $X_{S_k(\omega)}(\omega) = X^k_{S_k(\omega)}(\omega)$. However by right continuity and the definition of $S_k(\omega)$, we then must have $X_{S_k}(\omega) = X^k_{S_k}(\omega) \notin \Lambda_k$. □

One can easily find conditions which guarantee that the explosion time T in (3.15) is a.s. infinite, thus improving a bit on Theorem (3.10); we do not pursue these issues here.

4. The Semimartingale Topology

We begin with a key lemma, followed by an example which illustrates that a straightforward attempt to extend the lemma to a more general situation is doomed. This motivates the semimartingale topology, which gives the correct description of the stability of the solutions of stochastic differential equations. In this paragraph we will consider only the one dimensional case, the extensions to systems being clear and simple.

(4.1) Lemma. Let K,c be constants and suppose given the equations:

(4.2) $$X_t = J_t + \int_0^t F(X)_{s-} dZ_s$$

(4.3n) $$X_t^n = J_t^n + \int_0^t F^n(X^n)_{s-} dZ_s^n$$

where

(i) J and $(J^n)_{n\geq 1}$ are in $\underline{\underline{S}}^2$ (respectively $\underline{\underline{H}}^2$) and J^n tends to J in $\underline{\underline{S}}^2$ (resp. $\underline{\underline{H}}^2$);

(ii) F and $(F^n)_{n\geq 1}$ are all $\mathbb{D}$-Lipshitz with the same Lipschitz constant k and F^n all verify $\|F^n(U)\|_{\underline{\underline{S}}^\infty} \leq c$ for all $U \in \mathbb{D}$, and $F^n(X)$ converges to $F(X)$ in $\underline{\underline{S}}^2$, where X is the solution of (4.2);

(iii) Z is a semimartingale in $\mathcal{S}\left(\frac{1}{2\sqrt{8}\,k}\right)$, $(Z^n)_{n\geq 1}$ are in $\underline{\underline{H}}^2$, and Z^n converges to Z in $\underline{\underline{H}}^2$.

Then the solutions X^n of (4.3n) converge to X in $\underline{\underline{S}}^2$ (respectively $\underline{\underline{H}}^2$).

Proof. Suppose $J, (J^n)_{n\geq 1}$ are in $\underline{\underline{S}}^2$ and J^n converges to J in $\underline{\underline{S}}^2$. Then

$$X - X^n = J - J^n + (F(X) - F^n(X))_- \cdot Z + (F^n(X) - F^n(X^n))_- \cdot Z$$
$$+ F^n(X^n) \cdot (Z - Z^n).$$

Let $Y_n = (F(X) - F^n(X))_- \cdot Z + F^n(X^n) \cdot (Z - Z^n)$.

Then

(4.4) $X - X^n = J - J^n + Y^n + (F^n(X) - F^n(X^n))_- \cdot Z.$

For $U \in \mathbb{D}$ define G^n by:

$G^n(U) = F^n(X) - F^n(X - U)$. Then $G^n(U)$ is $\mathbb{D}$-Lipschitz with constant k and $G^n(0) = 0$. Take $U = X - X^n$, and (4.4) becomes

$$U = (J - J^n + Y^n) + G^n(U)_- \cdot Z$$

By (3.4) we have

$$||X - X^n||_{\underline{\underline{S}}^2} \leq C(k,Z)||J - J^n + Y^n||_{\underline{\underline{S}}^2} \leq C(k,Z)\{||J - J^n||_{\underline{\underline{S}}^2} + ||Y^n||_{\underline{\underline{S}}^2}\}.$$

Since $C(k,Z)$ is independent of n and $\lim_{n\to\infty} ||J - J^n||_{\underline{\underline{S}}^2} = 0$ by hypothesis, it suffices to show $\lim_{n\to\infty} ||Y^n||_{\underline{\underline{S}}^2} = 0$.

But

(4.5) $$||Y^n||_{\underline{\underline{S}}^2} \leq ||(F(X) - F^n(X))_- \cdot Z||_{\underline{\underline{S}}^2} + ||F^n(X^n)_- \cdot (Z - Z^n)||_{\underline{\underline{S}}^2}$$
$$\leq \sqrt{8}||F(X) - F^n(X)||_{\underline{\underline{S}}^2} ||Z||_{\underline{\underline{H}}^\infty} + \sqrt{8}\, ||F^n(X^n)_-||_{\underline{\underline{S}}^\infty}||Z-Z^n||_{\underline{\underline{H}}^2}$$

by (2.6) and (2.9). Since $||Z||_{\underline{\underline{H}}^\infty} < \infty$ by hypothesis and since $\lim_{n\to\infty} ||F(X) - F^n(X)||_{\underline{\underline{S}}^2} = \lim_{n\to\infty} ||Z - Z^n||_{\underline{\underline{H}}^2} = 0$, again by hypothesis, we are done.

Note that if we knew J^n, $J \in \underline{\underline{H}}^2$ and that J^n converged to J in $\underline{\underline{H}}^2$, then

$$||X - X^n||_{\underline{\underline{H}}^2} \leq ||J - J^n||_{\underline{\underline{H}}^2} + ||Y^n||_{\underline{\underline{H}}^2} + k||X - X^n||_{\underline{\underline{S}}^2} ||Z||_{\underline{\underline{H}}^\infty};$$

we have seen already that $\lim_{n\to\infty} ||X - X^n||_{\underline{\underline{S}}^2} = 0$, hence it suffices to show $\lim_{n\to\infty} ||Y^n||_{\underline{\underline{H}}^2} = 0$. Proceeding as in (4.5) we obtain the result. □

(4.6) Example. This example illustrates first the importance in (4.1) of the strong restriction that $||F^n(U)||_{\underline{\underline{S}}^\infty} \leq c$, for all $U \in \mathbb{D}$. We take $\Omega = [0,1]$, P to be Lebesgue measure on $[0,1]$, and $(\mathcal{F}_t)_{t\geq 0}$ equal to $\mathcal{F}$, the Lebesgue sets on $[0,1]$. Let $\phi(t) = \min(t,1)$, $t \geq 0$. Let $f_n(\omega) \geq 0$ and set $Z^n_t = \phi(t)f_n(\omega)$, $\omega \in [0,1]$, and finally take $J^n_t = J_t \equiv 1$, all t. Thus the equations (4.2) and (4.3n) are respectively:

$$X_t = 1 + \int_0^t X_{s-} dZ_s$$
$$X_t = 1 + \int_0^t X^n_{s-} dZ^n_s$$

which are elementary continuous exponential equations and have solutions:

$$X_t^n = \exp(Z_t^n) = \exp(f_n(\omega)\phi(t))$$

$$X_t = \exp(Z_t) = \exp(f(\omega)\,\phi(t)).$$

Suppose that $\lim_{n\to\infty} E\{f_n^2\} = 0$ but that $\lim_{n\to\infty} E\{f_n^p\} \neq 0$ for $p > 2$ (Cf Lemma (4.7) below which establishes (more than) the existence of such functions.) Then the Z^n converge to 0 in $\underline{\underline{H}}^2$ but X^n does not converge to $X = 1$ (since $f = 0$) in $\underline{\underline{S}}^p$, for any $p \geq 1$. Indeed, $\lim_{n\to\infty} E\{f_n^p\} \neq 0$ for $p > 2$ implies $\lim_{n\to\infty} E\{e^{tf_n}\} \neq 1$ for any $t > 0$. □

Note that the above example shows that when the coefficients are not uniformly bounded, it is the perturbation of the semimartingale differentials that create problems. However it is not a priori clear that a weaker claim is also false: could one perhaps have that Z^n converges to Z in $\underline{\underline{H}}^2$ implies that X_n converges to X prelocally in $\underline{\underline{S}}^1$? Before we continue with our example, we state a real variables lemma. We refer the reader to [41] for a proof.

(4.7) Lemma. There exist nonnegative measurable functions $(g_n)_{n\geq 1}$ on [0,1] such that

(i) $\lim_{n\to\infty} \int_0^1 g_n(X)dx = 0$

(ii) $\limsup_{n\to\infty} \int_\Lambda [g_n(x)]^p dx = +\infty$

for all $p > 1$ and all Lebesgue sets Λ such that $P(\Lambda) > 0$.

Example (4.6) continued. We continue with the same notation as in example (4.6). We have seen that $\lim_{n\to\infty} ||Z^n||_{\underline{\underline{H}}^2} = 0$, but $\overline{\lim}_{n\to\infty} ||X - X^n||_{\underline{\underline{S}}^p} \neq 0$, all $p \geq 1$. We now want to show that X^n does not tend to X even prelocally in $\underline{\underline{S}}^p$. (Since all processes are continuous here, "locally" and "prelocally" are the same.) Let T be a stopping time with $P(T > 0) > 0$. Since $\mathcal{F}_t = \mathcal{F}$, all $t > 0$, a "stopping time" is simply a nonnegative random variable. Choose $\lambda \le 1$ such that $\Lambda = \{T > \lambda\}$ and $P(\Lambda) > 0$. Let our functions (f_n) be as the functions (g_n) in (4.7). Then

$$\|(X^n)^T\|_{\underline{\underline{S}}^1} = \|(e^{Z^n})^T\|_{\underline{\underline{S}}^1} = E\{e^{Z^n_T}\}$$
$$\ge E\{(Z^n_T)^p\}$$
$$\ge E\{1_\Lambda \phi(T)^p (f_n)^p\}$$
$$\ge \lambda^p E\{1_\Lambda (f_n)^p\}$$

and so $\limsup_{n\to\infty} \|(X^n)^T\|_{\underline{\underline{S}}^1} = \infty$, which makes it impossible that X^n tends locally in $\underline{\underline{S}}^1$ to $X = 1$. □

The preceding example is what motivates the semimartingale topology, since it is not possible to obtain general results of the form: if Z^n tends to Z in $\underline{\underline{H}}^p$, then X_n tends to X locally (or even prelocally) in $\underline{\underline{S}}^q$, for some q.

We begin our treatment of the semimartingale topology with an alternative presentation of the topology "u.c.p." that we used in Chapter III.

(4.8) Definition. The subspace of $\mathbb{D}$ of processes having a finite limit at ∞ is denoted $\mathbb{D}_\infty$. The subspace of $\mathbb{D}_\infty$ consisting of total semimartingales is denoted $\underline{\underline{T}}_\infty$.

For $X \in \mathbb{D}_\infty$ we define

$$\|X\|_{\mathbb{D}_\infty} = E\{\min(X^*_\infty, 1)\} .$$

Note that $\|\cdot\|_{\mathbb{D}_\infty}$ satisfies the triangle inequality, and that $\|X\|_{\mathbb{D}_\infty} = 0$ implies $X = 0$. But in general $\|aX\|_{\mathbb{D}_\infty} \ne |a|\, \|X\|_{\mathbb{D}_\infty}$, so $\|\cdot\|_{\mathbb{D}_\infty}$ is not a norm. It does induce a distance, however.

(4.9) Definition. For $X, Y \in \mathbb{D}_\infty$, set $d_\infty(X,Y) = \|X - Y\|_{\mathbb{D}_\infty}$.

For $X \in \mathbb{D}$, $X_{t \wedge n} \in \mathbb{D}_\infty$, thus we can set $||X||_{\mathbb{D}} = \sum_{n=0}^{\infty} 2^{-n} E\{\min (X_n^*, 1)\}$. (Again, $||\cdot||_{\mathbb{D}}$ does not define a norm). Also, we set

$$d(X,Y) = ||X - Y||_{\mathbb{D}} \quad \text{for } X,Y \in \mathbb{D}.$$

The proof of the next theorem is left to the reader.

(4.10) Theorem. $(\mathbb{D}_\infty, d)$ and $(\mathbb{D}, d)$ are both complete metric spaces.

(4.11) Theorem. The topology on $\mathbb{D}$ induced by d is the same as the one induced by u.c.p.

Proof. If X^n converges to X in u.c.p. then $\lim_n (X^n - X)_k^* = 0$ in probability, each k. Therefore by Lebesgue's dominated convergence theorem

$$\lim_{n\to\infty} \sum_k 2^{-k} E\{\min ((X^n - X)_k^*, 1))\}$$

$$= \sum_k 2^{-k} \lim_{n\to\infty} E\{\min ((X^n - X)_k^*, 1))\}$$

$$= 0.$$

Conversely let X^n converge to X in $(\mathbb{D},d)$ and suppose there exists $\varepsilon, \delta > 0$ such that for some $t > 0$, $\limsup_{n\to\infty} P\{(X^n - X)_t^* > \varepsilon\} \geq \delta$. Let $k_\delta \geq t$ and let $X^{n'}$ be a subsequence such that $P\{(X^{n'} - X)_t^* > \varepsilon\} \geq \delta$, all n'. Then

$\lim_{n'\to\infty} \sum_k 2^{-k} E\{\min (X^{n'} - X)_k^*, 1) \geq 2^{-k_\delta} \min (\varepsilon\delta, 1) > 0$ which is a contradiction. □

For the rest of this paragraph we will use the

(4.12) Notational Convention. For a process $X \in \mathbb{D}$ and a stopping time T, set

$$X^T = X_{t \wedge T} 1_{\{T>0\}}$$

$$X^{T-} = \{X_t 1_{[0,T[} + X_{T-} 1_{[T,\infty[}\} 1_{\{T>0\}} .$$

Note that $t \wedge T$ = minimum (t,T); the difference with previous notation is that we multiply by $1_{\{T>0\}}$.

(4.13) Definition. Let $1 \leq p \leq \infty$ and let $X,(X^n)_{n\geq 1}$ be semimartingales (respectively $X,(X^n) \in \underline{\underline{T}}_\infty$). Then (X^n) is said to converge locally in $\underline{\underline{H}}^p$ (resp. $\underline{\underline{H}}^p_\infty$) if there exists a sequence of stopping times $(T^k)_{k\geq 1}$ increasing to ∞ a.s. (respectively: such that there exists $K(\omega) < \infty$ with $k \geq K(\omega)$ implying $T^k(\omega) = \infty$) and such that for all k $(X^n)^{T^k}$ and X^{T^k} are in $\underline{\underline{H}}^p$ and $\lim_{n\to\infty} ||(X - X^n)^{T^k}||_{\underline{\underline{H}}^p} = 0$.

(4.14) Definition. Let $1 \leq p \leq \infty$ and let $X,(X^n)_{n\geq 1}$ be semimartingales (respectively $X,(X^n)_{n\geq 1} \in \underline{\underline{T}}_\infty$). Then X^n is said to converge prelocally in $\underline{\underline{H}}^p$ (resp. $\underline{\underline{H}}^p_\infty$) if there exist stopping times $(T^k)_{k\geq 1}$ as in (4.13) such that $\lim_{n\to\infty} ||(X^n - X)^{T^k-}||_{\underline{\underline{H}}^p} = 0$, each $k \geq 1$.

Let X be in $\underline{\underline{T}}_\infty$ and let $X = M + A$ be a decomposition of X with $A_0 = 0$.

$$j_0(M,A) = E\{\min(1,([M,M]_\infty^{\frac{1}{2}} + \int_0^\infty |dA_s|))\} + \sup_T E\{|\Delta M_T\} \tag{4.15}$$

where the supremum is taken over all stopping times T. (Note that j_0 is not exactly analogous to the j_p defined in (2.2)). Next for a semimartingale X in $\underline{\underline{T}}_\infty$ set

$$||X||_{\underline{\underline{H}}^0_\infty} = \inf_{X = M + A} j_0(M,A) \tag{4.16}$$

where the infinum is taken over all possible decompositions $X = M + A$ with $A_0 = 0$. Note that $||\cdot||_{\underline{\underline{H}}^0_\infty}$ lacks homogeneity and is not a norm. It does, however, verify that $||X||_{\underline{\underline{H}}^0_\infty} = 0$ implies $X = 0$ (cf (4.19) below).

(4.17) Theorem. Let $X \in \underline{\underline{T}}_\infty$. Then $||X||_{\underline{\underline{H}}^0_\infty} \leq 1$.

Proof. Let $\varepsilon > 0$. By IV.4.7 we know there exists a decomposition of X, $X = M + A$, where the jumps of the local martingale term M are bounded by ε. Therefore $j_0(M,A) \leq 1 + \sup_T E\{|\Delta M_T|\} \leq 1 + \varepsilon$ for this decomposition. Since ε was arbitrary and since $||\cdot||_{\underline{\underline{H}}^0_\infty}$ is the infimum over all decompositions, the result follows. □

(4.18) Definition. The topology on $\underline{\underline{T}}_\infty$ induced by convergence in $||\cdot||_{\underline{\underline{H}}^0_\infty}$ is called the semimartingale topology.

The next theorem is an important characterization of the semimartingale topology. It is due to M. Emery [19, 20].

(4.19) Theorem. *Let $1 \leq p < \infty$ and let $(X^n)_{n\geq 1}$, X be semimartingales in $\underline{\underline{T}}_\infty$.*

(i) *If (X^n) is Cauchy in the semimartingale topology then there exists a subsequence which converges prelocally in $\underline{\underline{H}}^1$ to a semimartingale Y;*

(ii) *If (X^n) converges to X prelocally in $\underline{\underline{H}}^1$ then X^n converges to X in the semimartingale topology.*

Proof. We first prove (ii). Replacing X^n with $X^n - X$ it suffices to consider the case where X^n converges to 0 prelocally in $\underline{\underline{H}}^1$. Suppose X^n does not converge to 0 in the semimartingale topology. Then there exists an $\alpha > 0$ and a subsequence n' such that $||X^{n'}||_{\underline{\underline{H}}^0_\infty} \geq 2\alpha$. Let T be a stopping time such that $P(T< \infty) < \alpha/4$ and such that $(X^{n'})^{T-}$ tends to 0 in $\underline{\underline{H}}^1$. For large enough n' there exists a decomposition of $(X^{n'})^{T-}$: $(X^{n'})^{T-} = M^{n'} + A^{n'}$ where as well

$$j_1(M^{n'},A^{n'}) = E\{[M^{n'},M^{n'}]^{\frac{1}{2}}_\infty + \int_{0-}^{\infty} |dA^{n'}_s|\} \leq \alpha/4.$$

Let $X^{n'} = N^{n'} + B^{n'}$ be another decomposition of $X^{n'}$ such that $N^{n'}_0 = 0$ and $N^{n'}$ has all jumps bounded by 1(IV.4.6). Then

$$X^{n'} = (X^{n'})^{T-} + \{X^{n'} - (X^{n'})^{T}\} + \{(X^{n'})^{T} - (X^{n'})^{T-}\}$$

$$= M^{n'} + A^{n'} + \{N^{n'} - B^{n'} - (N^{n'})^{T} - (B^{n'})^{T}\} + \{(X^{n'})^{T} - (X^{n'})^{T-}\}$$

$$= \{M^{n'} + N^{n'} - (N^{n'})^{T}\} + \{A^{n'} + B^{n'} - (B^{n'})^{T} + (X^{n'})^{T} - (X^{n'})^{T-}\}$$

$$= L^{n} + C^{n} \quad .$$

Note that $L^n = M^n$ and $C^n = A^n$ on $\{T = \infty\}$. For semimartingales U,V and a stopping time T it is an immediate corollary of III.6.17 that $[U^T,V] = [U,V]^T$, and thus:

$$[L^{n'},L^{n'}]_{\infty}^{\frac{1}{2}} + \int_{0-}^{\infty} |dC^{n'}_s| = [M^{n'},M^{n'}]_{\infty}^{\frac{1}{2}} + \int_{0-}^{\infty} |dA^{n'}_s| \text{ on } \{T = \infty\}. \text{ Therefore}$$

(4.20) $$E\{\min (1,([L^{n'},L^{n'}]_{\infty}^{\frac{1}{2}} + \int_{0-}^{\infty} |dC^{n'}_s|))\}$$

$$\leq E\{\min (1,([M^{n'},M^{n'}]_{\infty}^{\frac{1}{2}} + \int_{0-}^{\infty} |dA^{n'}_s|))\} + P\{T < \infty\}$$

$$\leq \alpha/4 + \alpha/4 = \alpha/2.$$

Note further that for a stopping time S

$$|\Delta L^{n'}_S| \leq |\Delta M^{n'}_S| + |\Delta N^{n'}_S| 1_{\{T < \infty\}}$$

$$\leq [M^{n'},M^{n'}]_{\infty}^{\frac{1}{2}} + 1_{\{T < \infty\}}$$

since the jumps of $N^{n'}$ are bounded by one, and since $|\Delta M_S| \leq (\sum_s (\Delta M_s)^2)^{\frac{1}{2}} \leq [M,M]_{\infty}^{\frac{1}{2}}$ (III.6.14). Then

(4.21) $$E\{|\Delta L^{n'}_S|\} \leq E\{[M^{n'},M^{n'}]_{\infty}^{\frac{1}{2}}\} + P\{T < \infty\} \leq \alpha/4 + \alpha/4 = \alpha/2.$$

Thus $J_0(L^{n'},C^{n'}) \leq \alpha$ by combining (4.20) and (4.21), whence $||X^{n'}||_{\underline{\underline{H}}^0_\infty} \leq j_0(L^{n'},C^{n'}) \leq \alpha$, which contradicts $||X^{n'}||_{\underline{\underline{H}}^0_\infty} \geq 2\alpha$, since $\alpha > 0$.

<u>Proof of (i)</u>: Suppose (X^n) is Cauchy in the semimartingale topology. Without loss of generality assume $X_0 = 0$ and $||X^{n+1} - X^n||_{\underline{\underline{H}}^0_\infty} \leq 2^{-n}$, since we are

dealing with subsequences. Let $X^{n+1} - X^n = M^n + A^n$ be a decomposition with $j_0(M^n,A^n) \leq 2^{-n+1}$, each n. Let

$$C_t = \sum_n ([M^n,M^n]_t^{\frac{1}{2}} + \int_{0-}^{t} |dA_s^n|)$$

which is a.s. finite valued on $[0,\infty]$ since X and X^n are all in $\underline{\underline{T}}_\infty$. Define

$$T_k(\omega) = \inf\{t: C_t(\omega) > k\}.$$

Thus there exists a r.v. K such that if $k \geq K(\omega)$ then $T_k(\omega) = \infty$, almost all ω.

(4.22) Set $A_t^n = (\sum_{k=0}^{n} A_t^k)^{T-}$, where $A^0 = 0$, and $A = \lim A^n$ which exists since $C_\infty < \infty$ a.s. Then $||(A - A^n)^{T-}||_{\underline{\underline{H}}^1} \leq ||\int_0^{T-} |d(A - A^n)_s| ||_{L^1} \leq || \sum_{\ell=n+1}^{\infty} \int_0^{T-} |dA_s^\ell| ||_{L^1}$

which tends to 0 as n tends to ∞ by dominated convergence.

Since
$$[M^n,M^n]_T^{\frac{1}{2}} \leq [M^n,M^n]_{T-}^{\frac{1}{2}} + \Delta[M^n,M^n]_T^{\frac{1}{2}}$$
$$= [M^n,M^n]_{T-}^{\frac{1}{2}} + |\Delta M_T^n| ,$$
$$\leq k + |\Delta M_T^n| ,$$

and $\sup_S E\{|\Delta M_S^n|\} \leq 2^{-n+1}$, we have $||(M^n)^T||_{\underline{\underline{H}}^1} \leq k + 2^{-n+1}$ and each $M^n \in \underline{\underline{H}}^1$.

Moreover

$$|| \sum_{\ell=n+1}^{\infty} (M^\ell)^T ||_{\underline{\underline{H}}^1} \leq \sum_{\ell=n+1}^{\infty} ||(M^\ell)^T||_{\underline{\underline{H}}^1}$$
$$\leq \sum_{\ell=n+1}^{\infty} E\{[M^\ell,M^\ell]_T^{\frac{1}{2}}\}$$
$$\leq \sum_{\ell=n+1}^{\infty} E\{[M^\ell,M^\ell]_{T-}^{\frac{1}{2}} + |\Delta M_T|\}$$
$$\leq \sum_{\ell=n+1}^{\infty} E\{[M^\ell,M^\ell]_{T-}^{\frac{1}{2}}\} + \sum_{\ell=n+1}^{\infty} E\{|\Delta M_T^\ell|\}$$
$$\leq E\{\sum_{\ell=n+1}^{\infty} [M^\ell,M^\ell]_{T-}^{\frac{1}{2}}\} + \sum_{\ell=n+1}^{\infty} 2^{-\ell+1}$$

which tends to 0 as n tends to ∞, again by the dominated convergence theorem.

We have thus seen that $(M^n)^T = \sum_{\ell=0}^{n} (M^\ell)^T$ is a Cauchy sequence of local martingales in $\underline{\underline{H}}^1$. However the space of $\underline{\underline{H}}^1$ local martingales is a Banach space; that is, it is complete (cf [11]). Let

$$(4.23) \qquad M = \lim_{n\to\infty} (M^n)^T$$

where the limit is in $\underline{\underline{H}}^1$, and where M is a local martingale. Combining A from (4.22) and M from (4.23), define

$$X^{(k)} = M + A.$$

By the uniqueness of limits $X^{(\ell)} = X^{(k)}$ on $[0,T^k[$ if $\ell > k$, and we define

$$X = X^{(k)} \text{ on } [0,T^k[\ .$$

Let $X_\infty = \lim_{t\to\infty} X_t$, a limit we know exists since the sequence $T^k(\omega)$ is eventually equal to ∞, a.s. Moreover since X^{T_k-} is a semimartingale for each k, we deduce $X \in \underline{\underline{T}}_\infty$. The theorem now follows. □

(4.24) Comment. In the preceding proof in order to show M as defined in (4.23) existed, we needed the result that the space $\underline{\underline{H}}^1$ of local martingales is a Banach space. This fact is a simple consequence of the Burgess Davis inequality of martingale theory, a fairly deep result. One can avoid the necessity of invoking this theorem by replacing (i) in Theorem (4.19) with the following:

(i') *If (X^n) converges to X in the semimartingale topology then there exists a subsequence which converges to X prelocally in $\underline{\underline{H}}^1$.*

(4.25) Corollary. *The space $\underline{\underline{T}}_\infty$ equipped with the semimartingale topology is a topological vector space, metrizable into a complete metric space.*

Proof. Let $d(X,Y) = ||X - Y||_{\underline{\underline{H}}^0_\infty}$ for $X,Y \in \underline{\underline{T}}_\infty$.

By Theorem (4.19) (i) we have semimartingale convergence related to $\underline{\underline{H}}^1$

convergence, and thus X^n converges to X for d if and only if for each subsequence $X^{n'}$ we can extract a sub-subsequence $X^{n''}$ with converges to X prelocally in $\underline{\underline{H}}^1$. It follows that $\underline{\underline{T}}_\infty$ is a topological vector space that is metrizable. The completeness also follows from (4.19). □

(4.26) Comment. The choice $p = 1$ made here is arbitrary. Émery [19] has established analogous results for $1 \leq p < \infty$.

Let X be a semimartingale. For a positive integer n let $X_t^n = X_{t \wedge n}$. Then $X^n \in \underline{\underline{T}}_\infty$. Define

(4.27) $$||X||_{\underline{\underline{H}}^0} = \sum_n 2^{-n} ||X^n||_{\underline{\underline{H}}_\infty^0} .$$

(4.28) Definition. The *semimartingale topology on the space of semimartingales is the topology induced by* $|| \cdot ||_{\underline{\underline{H}}_\infty^0}$.

For a semimartingale X with $X_t^n = X_{t \wedge n}$, then $X^{(k)}$ converges to X in the semimartingale topology if and only if $X^{(k)n}$ converge to X^n is the semimartingale topology for $\underline{\underline{T}}_\infty$ for every $n = 1,2.3,\ldots$. Thus one verifies easily that (4.19), (4.24) and (4.25) remain valid for the semimartingale topology on the space of semimartingales.

(4.29) Theorem. *Let X^n be a sequence of semimartingales that is prelocally Cauchy in the semimartingale topology. Then there exists a semimartingale X such that X^n converges to X in the semimartingale topology.*

Proof. By stopping at a finite fixed time we need only establish the theorem for $\underline{\underline{T}}_\infty$ and its semimartingale topology. Thus we assume $X^n \in \underline{\underline{T}}_\infty$, each n, and let T^k be stopping times increasing to ∞ a.s. such that $\lim_{k\to\infty} P\{T_k < \infty\} = 0$ and let $Y^k \in \underline{\underline{T}}_\infty$ such that $\lim_{n\to\infty} (X^n)^{T^k -} = Y^k$. Since $(Y^{k+1})^{T^k -} = (Y^k)^{T^k -}$, we define

$$Y = Y^k \text{ on } [0,T^k]$$

Replace X^n with $X^n - Y$, and suppose then that X^n does not converge to 0 in $\underline{\underline{H}}^0_\infty$. Let n' be a subsequence such that $||X^{n'}||_{\underline{\underline{H}}^0_\infty} \geq \alpha > 0$. For each k there exists a further subsequence that converges prelocally in $\underline{\underline{H}}^1$ (by (4.19)). For k = 1, let $n_{i,1}$ be the sequence; for k = 2, $n_{i,1,2}$ is a further subsequence. Using a diagonalization procedure and taking

$$Z^n = X^{n_{i,1,2,\ldots,n}}$$

we have

(4.30) $$||Z^n||_{\underline{\underline{H}}^0_\infty} \geq \alpha > 0$$

(4.31) $(Z^n)^{T^k-}$ converges prelocally in $\underline{\underline{H}}^1$ to 0 as n tends to ∞ for each k.

Property (4.31) implies that Z^n converges prelocally in $\underline{\underline{H}}^1$ to 0 as n tends to ∞. By (4.19) we conclude Z^n converges to 0 in the semimartingale topology, which contradicts (4.30) above. □

5. Stability of Solutions of Stochastic Differential Equations

We consider here equations and sequences of equations of the form:

(5.1) $$X_t = J_t + \int_0^t F(X)_{s-} dZ_s$$

(5.2n) $$X^n_t = J^n_t + \int_0^t F^n(X^n)_{s-} dZ^n_s$$

and we wish to find conditions on J^n, F^n, and Z^n converging to H, F, and Z such that X^n converges to X. Lemma (4.1) is a tantalizing preliminary result, but as we saw in Example (4.6) this lemma does not extend in a straightforward manner. The semimartingale topology of paragraph four gives the correct notion of convergence.

(5.3) Theorem. Let $(F^n)_{n\geq 1}$, F be $\mathbb{D}$-Lipschitz with Lipschitz constants all bounded by $k > 0$. Let $(J^n)_{n\geq 1}$, J, $(Z^n)_{n\geq 1}$, Z be semimartingales and suppose J^n tends to J and Z^n tends to Z in the semimartingale topology. Let X^n be the solution of (5.2n), X the solution of (5.1) and suppose $F^n(X)$ tends to $F(X)$ in u.c.p. Then X^n tends to X in the semimartingale topology.

Proof. By prelocal stopping at an arbitrarily large time T we can assume that J, J^n, Z, and Z^n, all $n \geq 1$, are in $\underline{\underline{H}}^2$ (see Lemma (5.7) following this proof). Also by prelocal stopping we can further assume that $Z \in \mathcal{S}\left(\frac{1}{4\sqrt{8}\,k}\right)$ and that $|F(X)|$ is uniformly bounded by a constant $C > \infty$. Moreover by (4.19), trivially extended to the case $p = 2$ (4.26), we can further assume - by passing to a subsequence if necessary - that Z^n converges to Z and J^n converges to J in $\underline{\underline{H}}^2$.

We introduce truncation operators B^x defined by (for $x \geq 0$):

(5.4) $B^x(X) = \min\{x, \sup(-x,X)\}$

Then B^x is $\mathbb{D}$-Lipschitz with Lipschitz constant 1, each $x \geq 0$. Consider the equations

(5.5n) $$Y_t^n = J_t^n + \int_0^k (B^{k+c+1}F^n)(Y^n)_{s-}\, dZ_s^n .$$

By Lemma (4.1) we have Y^n converges to X in $\underline{\underline{H}}^2$. Note that $(F^n(X) - F(X))^*$ tends to 0 in $L^2(dP)$ by hypothesis and that $(Y^n - X)^*$ also tends to 0 in $L^2(dP)$ by (2.6). Passing to a further subsequence we assume they both converge to 0 a.s.

Next define:

$$T_\ell = \inf\{t \geq 0: \exists\, m \geq \ell: |Y_t^m - X_t| + |F^m(X)_t - F(X)_t| \geq 1\}$$

The times T_ℓ increase and $\lim_{\ell\to\infty} P(T_\ell = \infty) = 1$. By prelocal stopping at $T_{\ell-}$ for $n \geq \ell$ we have $(Y^n - X)^*$ and $(F^n(X) - F(X))^*$ bounded by one. Note that stopping at $T_{\ell-}$ changes Z being in $\underline{\underline{\mathcal{S}}}(\frac{1}{4\sqrt{8}\,k})$ to $Z \in \underline{\underline{\mathcal{S}}}(\frac{1}{2\sqrt{8}\,k})$, by (2.14). Then

$$\begin{aligned}|F^n(Y^n)| &\leq |F^n(Y^n) - F^n(X)| + |F^n(X) - F(X)| + |F(X)| \\ &\leq k(Y^n - X)^* + (F^n(X) - F(X))^* + F(X)^* \\ &\leq k + 1 + c\,,\end{aligned}$$

which implies $B^{k+c+1}F^n(Y^n) = F^n(Y^n)$, and hence for J^n, Z^n stopped at an arbitrarily large time Q, Y^n is a solution of

$$Y_t^n = J_t^n + \int_0^T F^n(Y^n)_{s-}\, dZ_s^n$$

and thus by uniqueness of solutions we have $Y^n = X^n$ on $[0,Q[$. Thus we have seen that we can find a sequence of stopping times Q^h increasing to ∞ a.s. such that stopping all processes at Q^h for fixed h implies that there exists a subsequence n' such that $X^{n'}$ converges prelocally in $\underline{\underline{H}}^2$ to X, which by (4.19) implies that X^n converges to X in the semimartingale topology. □

Let Lip (k) be the space of $\mathbb{D}$-Lipschitz operators having a Lipschitz constant smaller than or equal to k, and give Lip (k) the topology of u.c.p. in $\mathbb{D}$. Let $\underline{\underline{\mathcal{S}}}$ be the space of semimartingales equipped with the semimartingale topology. Then a restatement of Theorem (5.3) would be:

(5.6) Theorem. The operator Λ: $\underline{\underline{\mathcal{S}}}$ x Lip (k) x $\underline{\underline{\mathcal{S}}}$ into $\underline{\underline{\mathcal{S}}}$ given by $\Lambda(J, F, Z) = X$, where

$$X_t = J_t + \int_0^t F(X)_{s-}\, dZ_s,$$

is continuous.

(5.7) Lemma. Let $(W^n)_{n\geq 1}$ be a sequence of semimartingales, $W_0^n = 0$. There exists a sequence of stopping times $(T^m)_{m\geq 1}$ increasing to ∞ a.s. such that $(W^n)^{T^m-} \in \underline{\underline{H}}^2$, each n,m.

Proof. For each n, let $W^n = N^n + A^n$ be a decomposition with the jumps of N^n bounded by one (IV.4.7). Let $C_t^n = [N^n,N^n]_t + \int_0^t |dA_s^n|$, and $S_k^n = \inf\{t: C_t^n \geq k\}$. Then for each n $(S_k^n)_{k\geq 0}$ is a sequence of stopping times increasing to ∞ a.s. Hence there exists an integer c(n,p) such that $P\{S_{c(n,p)}^n \leq p\} < 2^{-n-p}$. Set $R_m = \inf_{n\geq 1, p\geq m} S_{c(n,p)}^n$. Then R_m is a stopping time and $R_m \leq S_{c(n,m)}^n$, each n. Therefore $(C^n)^{R_m-}$ is bounded. Moreover R_m is increasing and

$$P(R_m < m) \leq \sum_{n\geq 1,\ p\geq m} P(R_{c(n,p)}^n < m)$$

$$\leq \sum_{n\geq 1,\ p\geq m} P(R_{c(n,p)}^n < p) \leq 2^{-m+1} .$$

Therefore $\lim_{n\to\infty} R_m = \infty$ a.s. We have found a sequence of stopping times R_m increasing to ∞ a.s. such that $(C^n)^{R_m-} \leq \delta(m)$ is bounded. However since

$$(M^n)^{R_m-} = (N^n)^{R_m} + \{(A^n)^{R_m} - \Delta N_{R_m}^n 1_{[R_m,\infty[}\}$$

$$= L^{n,m} + B^{n,m} ,$$

and $[L^{n,m}, L^{n,m}]_\infty^{\frac{1}{2}} + \int_0^\infty |dB_s^{n,m}| \leq 2(\delta(m) + 1)$, we have that

$$||(M^n)^{R_m-}||_{\underline{\underline{H}}^2} \leq j_2(L^{n,m}, B^{n,m}) \leq 2(\delta(m) + 1) < \infty. \qquad \square$$

VII. STOCHASTIC DIFFERENTIAL EQUATIONS AND MARKOV PROCESSES

1. Introduction

As mentioned in the introduction to Chapter VI, K. Itô used his integral with respect to the Wiener process to study multidimensional diffusions; that is, continuous strong Markov processes. Now that we have a theory of stochastic differential equations with general semimartingales as differentials, one might well ask how such an equation can be related to an underlying Markov process. Some answers to this general question are given in Theorems (3.5) and (3.12).

2. The Markov Framework

The subject of Markov processes is both highly developed and highly technical. We use the notation of R. Blumenthal and R. Getoor [2]; the results of this chapter are taken from [6] and [40]. Our treatment here will be reasonably self-contained. We use the notational convention $X \in \underline{\underline{F}}/\underline{\underline{G}}$ to mean that if G is a set in the σ-algebra $\underline{\underline{G}}$, then $X^{-1}(G) \in \underline{\underline{F}}$.

Throughout this chapter Ω will denote an underlying space on which are defined:

(2.1) (i) a semigroup $(\theta_t)_{t \geq 0}$ of operators;

(ii) a right continuous process $X = (X_t)_{t \geq 0}$ taking values in a topological space E with Borel σ-field $\underline{\underline{E}}$;

(iii) X satisfies $X_{t+s} = X_t \circ \theta_s$ $(t,s \geq 0)$;

(iv) a filtration $(\underline{\underline{M}}^o_t)_{t \geq 0}$ of separable σ-fields on Ω such that $X_t \in \underline{\underline{M}}^o_t/\underline{\underline{E}}$; $\underline{\underline{M}}^o = \bigvee_t \underline{\underline{M}}^o_t$; and $\theta_t \in \underline{\underline{M}}^o/\underline{\underline{M}}^o$;

(v) a probability kernel $P^x(d\omega)$ from $(E,\underline{\underline{E}}^*)$ into $(\Omega,\underline{\underline{M}}^o)$, where $\underline{\underline{E}}^* = \bigcap_{\mu \text{ finite measure}} \underline{\underline{E}}^\mu$ is the σ-field of universally measurable sets on E, and where $\underline{\underline{E}}^\mu$ denotes the completion of $\underline{\underline{E}}$ with respect to the measure μ.

For each probability law μ on $(E,\underline{\underline{E}})$ we set, for $\Lambda \in \underline{\underline{E}}$, $P^{\mu}(\Lambda) = \int_E P^x(\Lambda)\mu(dx)$.
Then $\underline{\underline{M}}^{\mu}$ denotes the completion of $\underline{\underline{M}}^{o}$ under P^{μ}, and $\underline{\underline{M}}^{\mu}_t$ denotes $\underline{\underline{M}}^{o}_t$ with the P^{μ}-null sets of $\underline{\underline{M}}^{\mu}$ adjoined. Finally we set $\underline{\underline{M}} = \cap\, \underline{\underline{M}}^{\mu}$ and $\underline{\underline{M}}_t = \cap\, \underline{\underline{M}}^{\mu}_t$, with the intersections over all μ finite. One should think of the probability P^{μ} on Ω as the law governing X when it has initial distribution μ.

Let $\underline{\underline{F}}^{o}_t = \sigma(X_s;\ s \le t)$. This is the natural filtration of X; we let $\underline{\underline{F}}^{\mu}_t$, $\underline{\underline{F}}^{\mu}$, $\underline{\underline{F}}$, $\underline{\underline{F}}_t$ be defined analogously to $\underline{\underline{M}}$.

(2.2) Definition. The collection $\underset{\sim}{X} = (\Omega, \underline{\underline{M}}, \underline{\underline{M}}_{t+}, \theta_t, X_t, P^x)$ is a normal Markov (respectively strong Markov) process if for every bounded $Z \in \underline{\underline{F}}$, every $t \ge 0$ (respectively: every stopping time T), and all probabilities μ,

$$E^{\mu}\{Z\cdot\theta_t \,|\, \underline{\underline{M}}_{t+}\} = E^{X_t}[Z]$$

(respectively: T replacing t), and if $P^x(X_0 = x) = 1$, all $x \in E$.

The following is a stronger form of the strong Markov property.

(2.3) Definition. A filtration $(\underline{\underline{M}}_t)_{t\ge 0}$ is a strong Markov filtration if for every finite $\underline{\underline{M}}_{t+}$ - stopping time T we have

(i) $\underline{\underline{M}}^{o}_{(T+s)+} = \underline{\underline{M}}^{o}_{T+} \vee \theta_T^{-1}(\underline{\underline{M}}^{o}_{s+})$, all $s,t \ge 0$;

(ii) for all bounded $Z \in \underline{\underline{M}}$ we have $Z\circ\theta_T \in \underline{\underline{M}}$ and $E^{\mu}\{Z\circ\theta_T \,|\, \underline{\underline{M}}_{T+}\} = E^{X_T}[Z]$, for all μ.

To see how a strong Markov filtration naturally arises, consider two different realizations of reflecting Brownian motion as follows: let $\underset{\sim}{W} = (\Omega, \underline{\underline{F}}, \underline{\underline{F}}_t, \theta_t, W_t, P^x)$ be a standard Wiener process with natural filtration $(\underline{\underline{F}}_t)_{t\ge 0}$. Set $X_t = |W_t|$. Then $\underset{\sim}{X} = (\Omega, \underline{\underline{F}}, \underline{\underline{F}}_t, \theta_t, X_t, P^x)$ satisfies (2.2), but $(\underline{\underline{F}}_t)_{t\ge 0}$ is not a strong Markov filtration for X, since the future of W after time t depends both on $X_s = |W_s|$ and on sign (W_s), and not simply on X_s

alone. Next let $A_t = \int_0^t 1_{[0,\infty[}(W_s)ds$, and let $\tau_t = \inf\{s: A_s > t\}$ and set $Y_t = W_{\tau_t}$. Then it is well known (e.g. [2, p. 212]) that $\underset{\sim}{Y} = (\Omega, \underline{\underline{F}}, \underline{\underline{F}}_{\tau_t}, \theta_{\tau_t}, Y_t, P^x)$ is a strong Markov process and moreover $\underline{\underline{M}}_t = \underline{\underline{F}}_{\tau_t}$ is a strong Markov filtration for Y.

(2.4) Definition. A process Y is said to be additive (respectively strongly additive) if

(i) $Y_0 = 0$ a.s.

(ii) for every $s,t \geq 0$ we have $Y_{s+t} = Y_s + Y_t \circ \theta_s$ a.s. (resp. for all $t \geq 0$ and all $\underline{\underline{M}}_t$ - stopping times S we have $Y_{S+t} = Y_S + Y_t \circ \theta_S$ a.s.).

The "shift" operators $(\theta_t)_{t \geq 0}$ operate on random variables. We can also "shift" processes with the help of the following operator, invented by M. Sharpe.

(2.5) Definition. A "big shift" Θ_s is defined by setting for a given process Y:

$$(\Theta_s Y)_t = Y_{t-s} \circ \theta_s 1_{[s,\infty[}(t) \quad , \; s,t \geq 0.$$

In the definition of additivity (2.4), the equalities were given a.s.. That is, we had $Y_{s+t}(\omega) = Y_s(\omega) + Y_t \circ \theta_s(\omega)$ for all $\omega \notin N_{s,t}$, where $P^\mu(N_{s,t}) = 0$, all μ. A priori the null set $N_{s,t}$ depends on s and t. A process Y is called perfectly additive if

$$Y_{s+t}(\omega) = Y_s(\omega) + Y_t \circ \theta_s(\omega) \quad , \text{ all } \omega \notin N, \text{ all } s,t \geq 0,$$

where $P^\mu(N) = 0$ for all probabilities μ on $(E, \underline{\underline{E}})$, and where N does not depend on s and t. The proof of the next result is quite technical and we do not give it. (See [6, p. 171] and [46] for a proof.)

(2.6) Theorem. If $(\underline{\underline{M}}_t)_{t>0}$ is a strong Markov filtration and Y is an additive, $(\underline{\underline{M}}_{t+})$ - adapted right continuous real-valued process, then Y is indistinguishable from an adapted perfectly additive process $\tilde{Y}$ which is moreover strongly additive as well.

By indistinguishable we mean $P^{\mu}\{\omega$: there exists $t \geq 0$ with $Y_t(\omega) \neq \tilde{Y}_t(\omega)\} = 0$, for all probabilities μ on $(E,\underline{\underline{E}})$.

In view of the Theorem (2.6) we henceforth assume that all additive processes are both perfectly additive and strongly additive.

(2.7) Lemma. A process Y is additive if and only if $(\Theta_s Y)_t = Y_t - Y_{t\wedge s}$.

Proof. Using $(\Theta_s Y)_t = Y_{t-s} \circ \theta_s 1_{(t > s)}$ and the additivity of Y we have

$$(\Theta_s Y)_t = (Y_{(t-s)+s} - Y_s) 1_{(t>s)} = (Y_t - Y_s) 1_{(t>s)} = Y_t - Y_{s\wedge t} .$$

On the other hand if $Y_t - Y_{t\wedge s} = (\Theta_s Y)_t$, then $Y_t - Y_{t\wedge s} = (Y_t - Y_s) 1_{(t>s)}$ $= (Y_{(t-s)+s} - Y) 1_{(t>s)} = Y_{t-s} \circ \theta_s 1_{(t>s)}$,

which implies $Y_u \circ \theta_s = Y_{u+s} - Y_s$, where $u = t - s > 0$; hence Y is additive. □

(2.8) Corollary. Let Y be an additive semimartingale. Then $(\Theta_s Y)$ is again a semimartingale, for each $s \geq 0$.

Proof. By Lemma (2.7), $(\Theta_s Y)_t = Y_t - Y_{t\wedge s}$, and the result is obvious. □

We next give some technical results that we will need in the next paragraph.

(2.9) Proposition. Let H^n be a sequence of $\underline{\underline{M}}_{t+}$ - measurable random variables such that $P^x - \lim_{n\to\infty} H^n$ exists for each $x \in E$. Then there exists a r.v. $H \in \underline{\underline{M}}_{t+}$ such that $P^x - \lim_{n\to\infty} H^n = H$ for every $x \in E$.

Proof. P^x - lim denotes convergence in P^x probability. Set $n_0(x) = 0$ and $n_k(x) = \inf\{m > n_{k-1}(x): \sup_{p,q\geq m} P^x(|H^p - H^q| > 2^{-k}) \leq 2^{-k}\}$. Then each n_k is $\underline{\underline{E}}^*$ - measurable, whence $(x,\omega) \longrightarrow Z_k^x(\omega) \equiv H^{n_k(x)}(\omega)$ is $\underline{\underline{E}}^* \otimes \underline{\underline{M}}_{t+}$ - measurable. Therefore $\liminf_{k\to\infty} Z_k^x(\omega)$ is $\underline{\underline{E}}^* \otimes \underline{\underline{M}}_{t+}$ - measurable. Since $P^x\{|Z_k^x - Z_{k+1}^x| > 2^{-k}\} \leq 2^{-k}$, the Borel-Cantelli lemma implies that Z_k^x tends to Z^x, P^x a.s. The hypothesis that $P^x - \lim_{n\to\infty} H^n$ exists implies that $n_k(x)$ tends to ∞ for all $x \in E$. Thus $P^x - \lim_{n\to\infty} H^n = Z^x$. Next set $H(\omega) = Z^{X_0(\omega)}(\omega)$, and H is $\underline{\underline{M}}_{t+}$ - measurable. Since $P^x(X_0 = x) = 1$, we are done. □

(2.10) Corollary. Let Y be a semimartingale for every P^x, and let $H \in \mathbb{L}$. Then there exists a process $H \cdot Y$ which is a version of the semimartingale integral process for every P^x.

Proof. First suppose $H \in \underline{\underline{S}}$. Then $H = H_0 1_{\{0\}} + \sum_{i=1}^{n} H_i 1_{]T_i, T_{i+1}]}$, and the semimartingale integral is given by

$$J_Y(H) = \int H_s dY_s = H_0 Y_0 + \sum_{i=1}^{n} H_i (Y^{T_{i+1}} - Y^{T_i}),$$

which is defined independently of x, but nevertheless is the stochastic integral for every P^x, $x \in E$. For a given $H \in \mathbb{L}$, choose $\varepsilon > 0$, and define inductively stopping times (T_k^ε) as follows: $T_1^\varepsilon(\omega) = 0$

$$T_{k+1}^\varepsilon(\omega) = \inf\{t > T_k^\varepsilon(\omega): |H_t(\omega) - H_{T_k^\varepsilon}(\omega)| \geq \varepsilon\}$$

where $T_{k+1}^\varepsilon(\omega) = +\infty$ if the set is empty or if $T_k^\varepsilon(\omega) = +\infty$. Since $H \in \mathbb{L}$, it is easy to see that the (T_k^ε) are a.s. equal to stopping times. Taking $\varepsilon = 1/n$, we set

$$H_t^n = \sum_k H_{T_k^{1/n}} 1_{]T_k^{1/n} \wedge n,\ T_{k+1}^{1/n} \wedge n]}$$

Then $H^n \in \underline{\underline{S}}$, each n, and the H^n converge to H uniformly on compacts. Recall that the semimartingale integral was defined as the extension by continuity of J_Y from $\underline{\underline{S}}$ to $\underline{\underline{W}}$ in ucp (III.4.6). Then $P^x - \lim_{n\to\infty} (H^n \cdot Y)_t = Z_t^x$ is a version of the P^x-semimartingale integral process of H with respect to Y. Since $H^n \cdot Y$ is defined independently of x and is $\underline{\underline{M}}_{t+}$ - measurable, by (2.9) we know there exists a r.v. Z_t, each t, such that $Z_t^x = \tilde{Z}_t$ P^x - a.s., all x, and that $\tilde{Z}_t \in \underline{\underline{M}}_{t+}$.

We need to show that there is a process Z in $\mathbb{D}$ that can serve as our semimartingale integral for each P^x, all $x \in E$. We know $Z^x \in \mathbb{D}$ and $Z_t = Z_t^x$, P^x- a.s., each t. Let $\Lambda_t = \{\omega: \tilde{Z}_r(\omega) = f(r)$, where f is a right continuous function with left limits, and $r \in \mathbb{Q} \cap [0,t]\}$. Then it is standard that $\Lambda_t \in \underline{\underline{M}}_{t+}$ (e.g., IV-T-18 of [11]), and moreover Λ_t decreases when t increases. We set:

$$Z_t(\omega) = \begin{cases} \lim_{r\to t;\ r>t;\ r\in\mathbb{Q}} \tilde{Z}_r(\omega); & \omega \in \bigcup_{s>t} \Lambda_s \\ 0 & \text{otherwise} \end{cases}$$

Then $Z \in \mathbb{D}$ and is $\underline{\underline{M}}_{t+}$ - adapted. Moreover since each $Z^x \in \mathbb{D}$ and also $\tilde{Z}^x = \tilde{Z}$ P^x- a.s., we have $P^x(\Lambda_t) = 1$ and hence $Z_t = Z_t^x$ P^x-a.s.

(2.11) Lemma. Let H^n be a sequence of random variables such that $P^x\text{-}\lim_{n\to\infty} H^n = H$, all $x \in E$. Then $P^x\text{-}\lim_{n\to\infty} H^n \circ \theta_s = H \circ \theta_s$ all $s \geq 0$, all $x \in E$.

Proof. By (2.9) we may assume that H is $\underline{\underline{M}}$-measurable. Let $g(x) = \min(1,|x|)$. Then $P^x\text{-} \lim_{n\to\infty} H^n = H$ is equivalent to $\lim_{n\to\infty} E^x\{g(H^n - H)\} = 0$. But:

$$E^x\{g(H^n \circ \theta_s - H \circ \theta_s)\} = E^x\{g(H^n - H)\circ\theta_s\}$$

$$= E^x\{E^x\{g(H^n - H)\circ\theta_s | \underline{\underline{M}}_{s+}\}\}$$

$$= E^x\{E^{X_s}\{g(H^n - H)\}\} .$$

But $E^y\{g(H^n - H)\}$ is a bounded function of y tending to 0 as n tends to ∞ for all $y \in E$; thus $\lim_{n\to\infty} E^{X_s}\{g(H^n - H)\} = 0$ as well, and the result follows from the dominated convergence theorem. □

(2.12) Theorem. Let $H \in \mathbb{L}$ and let Y be an additive semimartingale. Then $\theta_s(H \cdot Y) = (\theta_s H)\cdot(\theta_s Y)$.

Proof. Recall that $\theta_s Y$ is again a semimartingale (2.8), and that there exists a version of the stochastic integral valid for all P^x simultaneously (2.10), so that the statement of the theorem makes sense.

First let $H \in \underline{\underline{S}}$ so that

$$H = H_0 1_{\{0\}} + \sum_{i=0}^{k} H_i 1_{]T_i, T_{i+1}]} .$$

Then $H \cdot Y = H_0 Y_0 + \sum_{i>0}^{k} H_i (Y^{T_{i+1}} - Y^{T_i})$, and a simple calculation shows that the result holds in this case. Replacing (2.9) with that of Lemma (2.11), the proof is now analogous to that of Corollary (2.10) and we leave it to the reader. □

3. Markov Solutions of Stochastic Differential Equations

Throughout this paragraph $X = (\Omega, \underline{\underline{M}}, \underline{\underline{M}}_{t+}, \theta_t, X_t, P^x)$ will denote an underlying normal strong Markov process as defined in (2.2). The assumptions and notations of paragraph two will be used freely.

We consider here stochastic differential equations of the form:

$$Y_t^i = H_t^i + \sum_{j=1}^{d} \int_0^t F_j^i(\underset{\sim}{Y})_{s-} dZ_s^j ,$$

$1 \leq i \leq n$. However techniques for systems are the same as for one equation, so we take $n = 1$. Also since no extra problems arise when $d > 1$, for simplicity we consider only $d = 1$. Thus we will study

$$(3.1) \qquad Y_t = H_t + \int_0^t F(Y)_{s-} dZ_s$$

and we will try to relate Y to the underlying process X. Our basic result (Theorem (3.5)) is that under appropriate hypotheses on H, F, and Z, the process (Y,X) is strong Markov.

To allow arbitrary initial distributions for the solution process Y we enlarge our probability space Ω as follows: set $\overline{\Omega} = \Omega \times \mathbb{R}$; $\overline{\underline{M}}_t^0 \otimes \underline{\mathbb{R}}$ and $\overline{\underline{M}}^0 = \underline{M}^0 \otimes \underline{\mathbb{R}}$, where $\underline{\mathbb{R}}$ denotes the Borel sets of $\mathbb{R}$. Let $\overline{P}^{x,y} = P^x \otimes \varepsilon_y$, product measure, where ε_y denotes the point mass Dirac measure at $\{y\}$.

Next set $\overline{E} = E \times \mathbb{R}$ and we let $(\overline{\underline{H}}_t)_{t \geq 0}$ denote the smallest right continuous filtration satisfying $\underline{M}_t \otimes \mathbb{R} \subseteq \overline{\underline{H}}_t$ for all $t \geq 0$.

There is a one-to-one correspondence between a process $\overline{H}$ on $\overline{\Omega}$ and a family of processes $(H^y)_{y \in \mathbb{R}}$ on Ω, and it is given by $H_t^y(\omega) = \overline{H}_t(y,\omega)$, where $(y,\omega) \in \overline{\Omega}$. We let $\overline{\mathbb{D}}$ denote the space of $\overline{\underline{H}}_t$- adapted processes with paths in $\mathbb{D}$, $\overline{P}^{x,y}$ a.s. for each $(x,y) \in \overline{E}$.

We now turn to the coefficients in equation (3.1).

(3.2) Definition. Given a semimartingale Z we say F is $\mathbb{D}$-extendably Lipschitz if F is $\mathbb{D}$-Lipschitz and can be extended to be $\overline{\mathbb{D}}$-Lipschitz on $\overline{\Omega} \times \mathbb{R}_+$.

Recall that $\mathbb{D}$-Lipshitz was defined in Chapter VI (VI.3.2). We need one more technical result.

(3.3) Theorem. *Let $\overline{H} = (H^y)_{y \in \mathbb{R}} \in \mathbb{D}$, and $\overline{Z} = (Z^y)_{y \in \mathbb{R}} \in \overline{\mathbb{D}}$ such that Z^y is a*

<u>semimartingale for</u> P^x , <u>all</u> $x \in E$ <u>and all</u> $y \in \mathbb{R}$; <u>let</u> F <u>be extendably</u> $\mathbb{D}$-<u>Lipschitz</u>. <u>Then there exists a process</u> $\bar{Y} = (Y^y)_{y\in\mathbb{R}}$ <u>in</u> $\bar{\mathbb{D}}$ <u>such that</u> Y^y <u>is the</u> P^x <u>solution of</u>

$$Y_t^y = H_t^y + \int_0^t F(Y^y)_{s-}\, dZ_s^y$$

<u>for all</u> $y \in \mathbb{R}$ <u>and all</u> $x \in E$.

<u>Proof</u>. For $(\omega,y) \in \bar{\Omega}$ define $\bar{\theta}_t(\omega,y) = (\theta_t(\omega),y)$, and define $\bar{X}_t(\omega,y) = (X_t(\omega),y)$. Then $\underset{\sim}{\bar{X}} = (\bar{\Omega}, \bar{\underline{\underline{H}}}^0, \bar{\underline{\underline{H}}}_{t+}^0, \bar{\theta}_t, \bar{X}_t, \bar{P}^{x,y})$ is a normal strong Markov process. Since $\bar{P}^{x,y} = P^x \otimes \varepsilon_y$, it is trivial that $\bar{Z}$ is a semimartingale $(\bar{P}^{x,y})$ when Z^y is one P^x. Thus for any $\bar{H} \in \bar{\mathbb{L}}$ we can find a version of the semimartingale integral $\bar{H}\cdot\bar{Z}$ such that $\bar{H}\cdot\bar{Z} = \int H_s^y dZ_s^y$, $P^{x,y}$ - a.s. That is, $(\bar{H}\cdot\bar{Z})^y = H^y \cdot Z^y$ P^x- a.s., all $x \in E$.

Next define $\bar{Y}(1) = \bar{H}$, and inductively define $\bar{Y}(n+1) = \bar{H} + \bar{F}(\bar{Y}(n))_- \cdot \bar{Z}$, where $\bar{F}(\bar{U})^y = F(U^y)$. Then $\bar{Y}(n) \in \bar{\mathbb{D}}$ for each n. Using the Lipschitz hypothesis on F one can verify $\bar{P}^{x,y}$- $\lim_{n\to\infty} \bar{Y}(n)_t = \bar{Y}_t$, where $\bar{Y}$ is a solution of the equation such that $\bar{Y} = (Y^y)$, Y^y being the solution for all P^x, $x \in E$. An argument analogous to the one used in the proof of (2.10) yields the result. □

In Chapter VI we allowed our coefficients to depend on the entire history of the solution up to the present state. Such latitude in the coefficients is too great if we are to hope to have a Markov property for the solution.

(3.4) <u>Definition</u>. A coefficient F is said to be <u>strongly homogeneous</u> if

(i) for all $Y \in \mathbb{D}$ and stopping times T $\theta_T(F(Y))$ is indistinguishable from $F(\theta_T Y)$ on $[T,\infty[$;

(ii) for all $\omega \in \Omega$, $s \geq 0$, $t > s$ and H, $K \in \mathbb{D}$ such that $H_r(\omega) = K_r(\omega)$ when $s < r < t$, then $F(H)_t(\omega) = F(K)_t(\omega)$.

Examples of strongly homogeneous coefficients are $F(Y)_{t-}(\omega) = f(Y_{t-}(\omega))$ with $f \in C^1$; $G(Y)_{t-}(\omega) = g(X_{t-}(\omega), Y_{t-}(\omega))1_{(0,\infty)}(t)$, assuming X has left limits and g is smooth.

(3.5) Theorem. Let $H, Z \in \mathbb{D}$ be strongly additive and suppose Z is a P^x-semimartingale for all x. Let F be extendably $\mathbb{D}$-Lipschitz and strongly homogeneous and let $\overline{Y} = (Y^y)_{y \in \mathbb{R}} \in \mathbb{D}$ such that Y^y is a solution for each $y \in \mathbb{R}$ of

(3.6) $$Y_t^y = y + H_t + \int_0^t F(Y^y)_{s-}\, dZ_s.$$

Then

(3.7) $$\overline{E}^{x,y}\{f(X_{S+t}, \overline{Y}_{S+t}) \mid \underline{\underline{H}}_S\} = \overline{E}^{(X_S, \overline{Y}_S)}\{f(X_t, \overline{Y}_t)\}$$

for all bounded $f \in \underline{\underline{E}} \otimes \underline{\underline{R}}$, for all $(x,y) \in \overline{E}$, and for all finite $\underline{\underline{H}}_t$- stopping times S.

Proof. Fix an a.s. finite stopping time S. Let $H_t' = H_t - H_{t\wedge S}$, $Z_t' = Z_t - Z_{t\wedge S}$. For $G \in \underline{\underline{H}}_S$ and finite, let $\tilde{Y}^G$ denote a solution of the equation

$$\tilde{Y}_t^G = G 1_{[S,\infty[}(t) + H_t' + \int_0^t F(\tilde{Y}^G)_{s-}\, dZ'_s.$$

Step 1. We first show that if $G = Y_S^y$, then $\tilde{Y}_t^G = Y_t^y$ on $\{t \geq S\}$. We let

$\tilde{Y}_u = Y_u^y 1_{\{u<S\}} + \tilde{Y}_u^G 1_{\{u\geq S\}}$, which is in $\mathbb{D}$. On $\{t \leq S\}$ we have

$$\tilde{Y}_t = \tilde{Y}_t^G = Y_S^y + H_t' + \int_0^t F(\tilde{Y}^G)_{s-} dZ'_s$$

which can be reexpressed on $\{t \geq S\}$ as

$$= y + H_t + \int_0^S F(Y^y)_{s-}\, dZ_s + \int_0^t F(\tilde{Y}^G)_{s-}\, dZ'_s .$$

By the homogeneity of F we have $F(Y^y)_u = F(\tilde{Y})_u$ if $u \leq S$ and also $F(\tilde{Y}^G)_u = F(\tilde{Y})_u$ if $u > S$. However $Z'_u = 0$ on $\{u \leq S\}$, hence

$$(3.8) \qquad \tilde{Y}_t = y + H_t + \int_0^t F(\tilde{Y})_{s-}\, dZ_s \text{ on } \{t \geq S\}.$$

But (3.8) is clearly valid for $t < S$ as well, and thus $\tilde{Y}$ is a solution of equation (3.6), and step 1 follows from the uniqueness of solutions.

Step 2: We next show $\Theta_S Y^S$ and $\tilde{Y}^y$ are indistinguishable, where $\tilde{Y}^y \equiv \tilde{Y}^G$ if $G \equiv y$. We have:

$$\Theta_S Y^y = y1_{[S,\infty[} + \Theta_S H + \Theta_S(F(Y^y)\cdot Z)$$
$$= y1_{[S,\infty[} + \Theta_S H + (\Theta_S F(Y^y) \cdot \Theta_S Z)$$

by Theorem (2.12). Using the additivity of H and Z as well as Lemma (2.7), we get:

$$\Theta_S Y^y = y1_{[S,\infty[} + H' + F(\Theta_S(Y^y))\cdot Z' \quad \text{a.s.}$$

Thus $\Theta_S Y^y$ is a solution of the same equation that $\tilde{Y}^y$ is, and step 2 follows from the uniqueness of such solutions.

Step 3. In this step we will see that

$$\overline{E}^{x,y}\{\overline{V}f(X_{S+t}, \overline{Y}_{S+t})\} = \overline{E}^{x,y}\{\overline{V}\overline{E}^{(X_S,\overline{Y}_S)}[f(X_t, \overline{Y}_t)]\}$$

for all bounded $\overline{V}$ in $\underline{\underline{H}}_S$. Since $\underline{\underline{H}}_t$ is contained in the $\overline{P}^{x,y}$ completion of $\underline{\underline{H}}_t \otimes \{\phi, \mathbb{R}\}$, we need establish only that

$$\overline{E}^{x,y}\{Vf(X_{S+t},\overline{Y}_{S+t})\} = \overline{E}^{x,y}\{VE^{X_S,\overline{Y}_S}[f(X_t, \overline{Y}_t)]\}$$

for arbitrary stopping time S, $t \geq 0$, bounded $V \in \underline{\underline{H}}_S$ and bounded $f \in \underline{\underline{E}} \otimes \underline{\underline{R}}$.

Set $G = Y_S^y$. Then

$$(3.9) \qquad \overline{E}^{x,y}\{Vf(X_{S+t},\overline{Y}_{S+t})\} = E^x\{Vf(X_{S+t},Y^y{}_{S+t})\} = E^x\{Vf(X_{S+t},\tilde{Y}^G_{S+t})$$

by step 1. However by (3.3) we know $(\omega,y) \to \tilde{Y}^y_{S+t}(\omega)$ is $\underline{\underline{H}}_{S+t}\otimes\underline{\underline{R}}$-measurable. Since $G\in \underline{\underline{H}}_S$, equation (3.4) yields

$$\overline{E}^{x,y}\{Vf(X_{S+t},\overline{Y}_{S+t})\} = \int P^x(d\omega)V(\omega)E^x\{(f(X_{S+t},\tilde{Y}^{G(\omega)}_{S+t})|\underline{\underline{H}}_S\}(\omega)$$

$$= \int P^x(d\omega)V(\omega)E^x\{f(X_t,Y_t^{G(\omega)}\circ\theta_S(\omega))|\underline{\underline{H}}_S\}(\omega)$$

by step 2. Using the strong Markov property of X and continuing the equation:

$$= \int P^x(d\omega)V(\omega)E^{X_S(\omega)}\{f(X_t,Y_t^{G(\omega)})\}$$

$$= \int P^x(d\omega)V(\omega)\overline{E}^{(X_S(\omega),G(\omega))}\{f(X_t,\overline{Y}_t)\}$$

$$= E^x\{V\overline{E}^{(X_S,Y_S^y)}[f(X_t,\overline{Y}_t)]\}$$

$$= \overline{E}^{x,y}\{V\overline{E}^{(X_S,Y_S)}[f(X_t,\overline{Y}_t)]\}.$$

This completes the proof. □

(3.10) Comment. Observe that equation (3.7) implies that $(\Omega, \underline{\underline{H}}, \underline{\underline{H}}_t, (X_t,\overline{Y}_t), \overline{P}^{x,y})$ is a right continuous strong Markov process as defined in [2], but unfortunately it is not one according to our definition (2.2). The problem is that the shift operators $(\theta_t)_{t\geq 0}$ for $(X,\overline{Y})$ are not present. We could define a shift by $\overline{\theta}_t(\omega,y) = (\theta_t(\omega), \overline{Y}_t(\omega,y))$ in which case we would have $X_{t+s} = X_t \circ \overline{\theta}_s$ on $\overline{\Omega}$, but the relation $\overline{\theta}_{t+s} = \overline{\theta}_t \circ \overline{\theta}_s$, for example, only holds a.s. and not identically. This is, however, a false problem since being a Markov process is really a statement about only the transition semigroup, and one can always realize a process on a new space with the same transition semigroup that would be a strong Markov process in the sense of definition (2.2). (cf. [6, p. 271])

(3.11) Comment. The hypotheses that H and Z be additive in Theorem (3.5) may at first seem strange. However examples abound: if the Markov process X is itself a semimartingale one can take $Z_t \equiv X_t - X_0$. If X is not a semimartingale then there always exist functions $f: E \to \mathbb{R}$ such that $Z_t = f(X_t) - f(X_0)$ is an additive semimartingale for all P^x (cf. [6]).

In the classical Itô theory the underlying Markov process X is just a Wiener process, and one has two semimartingale differentials: $dZ_t^1 = dX_t$, the Wiener process, and $dZ_t^2 = dt$, Lebesgue measure. Due to the independence of the increments, however, the conclusion is stronger than that of Theorem (3.5): the solution itself is strong Markov. This result extends in our framework as follows:

(3.12) Theorem. *Let X have stationary and independent increments and let $Z_t = X_t - X_0$. Let F be extendably $\mathbb{D}$-Lipschitz and strongly homogeneous and let $\overline{Y} = (Y^y)_{y \in \mathbb{R}} \in \overline{\mathbb{D}}$ such that Y^y is a solution, for each $y \in \mathbb{R}$, of:*

$$Y_t^y = y + \int_0^t F(Y^y)_{s-} \, dZ_s \ .$$

Then $\overline{E}^{x,y}\{f(\overline{Y}_{S+t}) | \underline{\underline{H}}_S\} = \overline{E}^{(x,\overline{Y}_S)}\{f(\overline{Y}_t)\}$ for all bounded $f \in \underline{\underline{R}}$, $(x,y) \in \overline{E}$, and finite stopping times S.

Proof. The proof proceeds exactly as steps one and two of the proof of Theorem (3.5), taking H = 0 of course. However using the elementary fact that the independence of the increments of X is also valid for stopping times [that is, $\sigma(X_{S+t} - X_S;\ t \geq 0)$ is independent of $\underline{\underline{H}}_S$], (cf. [1], for example), we note that $\tilde{Y}^G$ as defined just before step one is independent of $\underline{\underline{H}}_S$.

Step 3. Let $V \in \underline{\underline{H}}_S$ and $f \in \underline{\underline{R}}$ both be bounded. Then as in (3.9):

$$(3.13) \qquad \overline{E}^{x,y}\{Vf(\overline{Y}_{S+t})\} = E^x\{Vf(Y^y_{S+t})\}$$
$$= E^x\{Vf(\tilde{Y}^G_{S+t})\}$$
$$= \int P^x(d\omega)V(\omega)E^x\{f(\tilde{Y}^{G(\omega)}_{S+t})|\underline{\underline{H}}_S\}(\omega).$$

However $\tilde{Y}^y_{S+t}$ is independent of $\underline{\underline{H}}_S$, hence $E^x\{f(\tilde{Y}^{G(\omega)}_{S+t})|\underline{\underline{H}}_S\}(\omega) = h(x,G(\omega))$, where $h(x,y) = E^x\{f(\tilde{Y}^y_t)\}$, using the stationarity of the increments. Thus (3.13) becomes:

$$\overline{E}^{x,y}\{Vf(\overline{Y}_{S+t})\} = \int P^x(d\omega)V(\omega)h(x,G(\omega))$$
$$= E^x\{V\overline{E}^{(x,Y^y_S)}[f(\overline{Y}_t)]\}$$
$$= E^{x,y}\{V\overline{E}^{(x,\overline{Y}_S)}[f(\overline{Y}_t)]\},$$

which completes the proof. □

VIII. REFERENCES

1. Bichteler, K.: Stochastic integration and L^p theory of semimartingales. Ann. Prob. 9, 49-89 (1981).

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

3. Bretagnolle, J. L.: Processus à accroissements indépendants. Ecole d'Eté de Probabilités. Springer Lect. Notes in Math. 307, 1-26 (1973).

4. Chou, C. S., Meyer, P. A., Stricker, C.: Sur les intégrales stochastiques de processus prévisibles non bornés. Springer Lect. Notes in Math. 784, 128-139 (1980).

5. Chung, K. L., Williams, R. J.: Introduction to Stochastic Integration. Boston: Birkhauser (1983).

6. Çinlar, C., Jacod, J., Protter, P., Sharpe, M. J.: Semimartingales and Markov processes. Z. Wahrscheinlichkeits. Geb. 54, 161-219 (1980).

7. Courrège, P.: Intégrale stochastique par rapport à une martingale de carré intégrable. Seminaire Brélot-Choquet-Dény, 7 année (1962-63). Institut Henri Poincaré, Paris.

8. Dellacherie, C.: Capacités et Processus Stochastiques. Springer-Verlag: Berlin (1972).

9. ___________: Mesurabilité des débuts et théorème de section: le lot à la portée de toutes les bourses. Springer Lect. Notes in Math. 850, 351-360 (1981).

10. ___________: Un survol de la théorie de l'intégrale stochastique. Stochastic Proc. Appl. 10, 115-144 (1980).

11. Dellacherie, C., Meyer, P. A.: Probabilities and Potential B. North-Holland: Amsterdam (1982).

12. Doléans-Dade, C.: Intégrales stochastique dépendant d'un paramètre. Publ. Inst. Stat. Univ. Paris 16, 23-34 (1967).

13. ___________: On the existence and unicity of solutions of stochastic differential equations. Z. Wahrschein. Geb. 36, 93-101 (1976).

14. Doléans-Dade, C., Meyer, P. A.: Intégrales stochastiques par rapport aux martingales locales. Lect. Notes in Math. 124, 77-107 (1970).

15. ____________________: Equations différentielles stochastiques. Lect. Notes in Math. 581, 376-382 (1977).

16. Doob, J. L.: Stochastic Processes. Wiley: New York (1953).

17. Elliott, R. J.: Stochastic Calculus and Applications. Springer-Verlag: New York (1982).

18. Emery, M.: Stabilité des solutions des équations différentielles stochastiques. Application aux intégrales multiplicatives stochastiques. Z. Wahrschein. Geb. 41, 241-262 (1978).

19. __________: Une topologie sur l'espace des semimartingales. Springer Lect. Notes in Math. 721, 260-280 (1979).

20. __________: Equations differentielles stochastiques Lipschitziennes, étude de la stabilité. Springer Lect. Notes in Math. 721, 281-293 (1979).

21. Itô, K.: Stochastic integral. Proc. Japan Acad. 20, 519-524 (1944).

22. Itô, K., Watanabe, S.: Transformation of Markov processes by multiplicative functionals. Ann. Inst. Fourier, Grenoble 15, 13-20 (1965).

23. Jacod, J.: Calcul Stochastique et Problèmes de Martingales. Springer Lect. Notes in Math. 714 (1979).

24. Kunita, H., Watanabe, S.: On square integrable martingales. Nagoya J. Math. 30, 209-245 (1967).

25. Lenglart, E.: Transformation des martingales locales par changement absolument continu de probabilités. Z. Wahrscheinlich. Geb. 39, 65-70 (1977).

26. __________: Semimartingales et intégrales stochastiques en temps continu. Revue du CETHEDEC-Ondes et Signal 75, 91-160 (1983).

27. Letta, G.: Martingales et Intégration Stochastique. Scuola Normale Superiore: Pisa (1984).

28. Métivier, M.: Semimartingales: A Course on Stochastic Processes. Walter de Gruyter: Berlin (1982).

29. Métivier, M., Pellaumail, J.: On a stopped Doob's inequality and general stochastic equations. Ann. Probability 8, 96-114 (1980).

30. __________: Stochastic Integration. Academic Press: New York (1980).

31. Meyer, P. A.: A decomposition theorem for supermartingales. Illinois J. Math. 6, 193-205 (1962).

32. __________: Intégrales stochastiques I, II, III, IV. Springer Lect. Notes in Math. 39, 72-165 (1967).

33. __________: Le dual de H^1 est BMO (cas continu). Springer Lect. Notes in Math 321, 136-145 (1973).

34. __________: Un cours sur les intégrales stochastiques. Springer Lect. Notes in Math. 511, 246-400 (1976).

35. __________: Le théorème fondamental sur les martingales locales. Springer Lect. Notes in Math. 581, 463-464.

36. __________: Inégalités de normes pour les intégrales stochastiques. Springer Lect. Notes in Math. 649, 757-762 (1978).

37. Pellaumail, J.: Sur l'intégrale stochastique et la décomposition de Doob-Meyer. Astérisque 9, 1-124 (1974).

38. Protter P.: On the existence, uniqueness, convergence and explosions of solutions of stochastic differential equations. Ann. Probability 5, 243-261 (1977).

39. __________: Right-continuous solutions of systems of stochastic integral equations. J. Multivariate Analysis 7, 204-214 (1977).

40. __________: Markov solutions of stochastic differential equations. Z. Wahrscheinlich. Geb. 41, 39-58 (1977).

41. __________: H^p stability of solutions of stochastic differential equations. Z. Wahrscheinlich. Geb. 44, 337-352.

42. __________: Stochastic integration without tears (with apology to P. A. Meyer). To appear in Stochastics.

43. Schwartz, L.: Semimartingales and their Stochastic Calculus on Manifolds. Les Presses de l'Université de Montréal: Montréal (1984).

44. Stricker, C.: Quelques remarques sur la topologie des semimartingales. Applications aux intégrales stochastiques. Springer Lect. Notes in Math. 850, 499-522 (1981).

45. Stricker, C., Yor, M.: Calcul stochastique dépendant d'un paramètre. Z. Wahrscheinlich. Geb. 45, 109-133 (1978).

46. Walsh, J. B.: The perfection of multiplicative functionals. Springer Lect. Notes in Math. 258, 233-242 (1972).

47. Yan, Jia-An: Caractérisation d'une classe d'ensembles convexes de L^1 ou H^1. Springer Lect. Notes in Math. 784, 220-222 (1980).