Jaures Cecconi (Ed.)

CIME Summer Schools

77

# Stochastic Differential Equations

Cortona, Italy 1978

Springer

FONDAZIONE CIME ROBERTO CONTI

Jaures Cecconi (Ed.)

# Stochastic Differential Equations

Lectures given at a Summer School of the
Centro Internazionale Matematico Estivo (C.I.M.E.),
held in Cortona (Arezzo), Italy,
May 29-June 10, 1978

Springer

FONDAZIONE
CIME
ROBERTO CONTI

C.I.M.E. Foundation
c/o Dipartimento di Matematica "U. Dini"
Viale Morgagni n. 67/a
50134 Firenze
Italy
**cime@math.unifi.it**

ISBN 978-3-642-11077-1 e-ISBN: 978-3-642-11079-5
DOI:10.1007/978-3-642-11079-5
Springer Heidelberg Dordrecht London New York


Reprint of the 1st ed. C.I.M.E., Ed. Liguori, Napoli & Birkhäuser 1981
With kind permission of C.I.M.E.

Printed on acid-free paper

Springer.com

C O N T E N T S

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# STOCHASTIC PROCESSES AND STOCHASTIC DIFFERENTIAL EQUATIONS

C. DOLEANS-DADE

# STOCHASTIC PROCESSES AND STOCHASTIC DIFFERENTIAL EQUATIONS

C. Doléans-Dade

University of Illinois, Urbana

Introduction. Since Ito has defined the stochastic integral with respect to the Brownian motion, mathematicians have tried to generalize it. The first step consisted of replacing the Brownian motion by a square integrable martingale. Later H. Kunita and S. Watanabe in [10] introduced the concept of local continuous martingale and stochastic integral with respect to local continuous martingales which P. A. Meyer generalized to the non continuous case.

But in many cases one observes a certain process $X$ and there are at least two laws $P$ and $Q$ on $(\Omega,\underline{\underline{F}})$. For the law $Q$, $X$ is not a local martingale but the sum of a local martingale and a process with finite variation. We would like to talk about the stochastic integrals $\int_P \Phi_s dX_s$ and $\int_Q \Phi_s dX_s$ in the two probability spaces $(\Omega,\underline{\underline{F}},P)$ and $(\Omega,\underline{\underline{F}},Q)$. And of course we would like those two stochastic integrals to be the same.

This is why one should try to integrate with respect to semimartingales (sums of a local martingale and a process with finite variation), and this is what people have been doing for awhile (see chapters 5 and 6). Now the latest result in the theory is "one cannot integrate with respect to anything more general than semimartingales" (see chapter 3). So as it stands now the theory looks complete.

To end this introduction I wish to thank Professor J. P. Cecconi and

the C.I.M.E. for their kind invitation to this session on differential stochastic equations in Cortona; the two weeks of which I, and my family, found most enjoyable.

## STOPPING TIMES AND STOCHASTIC PROCESSES

We shall list in this chapter some definitions and properties on stopping times and stochastic processes. The proofs can be found in [1] or [2].

In all that follows $(\Omega,\underline{\underline{F}},P)$ is a given complete probability space and $(\underline{\underline{F}}_t)_{t\geq 0}$ a family of sub-$\sigma$-fields of $\underline{\underline{F}}$ verifying the "usual" following properties

a) the family $(\underline{\underline{F}}_t)_{t\geq 0}$ is non decreasing and continuous on the right

b) for each $t$, $\underline{\underline{F}}_t$ contains all the P-null sets of $\underline{\underline{F}}$ (a P-null set is a set of P-measure zero).

The $\sigma$-fields $\underline{\underline{F}}_t$ should be thought of as the $\sigma$-field of the events which occurred up to time $t$.

We will sometimes consider other probabilities $Q$ on the measurable space $(\Omega,\underline{\underline{F}})$. But we shall always assume that the probabilities $P$ and $Q$ are equivalent (i.e. they have the same null sets); and the family $(\underline{\underline{F}}_t)$ will still satisfy the "usual" conditions relatively to the probability $Q$.

### STOPPING TIMES

Suppose a gambler decides to stop playing when a certain phenomenon has occured in the game. Let $T$ be the time at which he will stop playing.

The event $\{T \le t\}$ will depend only on the observations of the gambler up to time $t$. This remark leads to the natural following definition.

1.1. Definition. A non negative random variable $T$ is a stopping time if for every $t \ge 0$ the event $\{T \le t\}$ is in $\underline{\underline{F}}_t$. (We allow the random variable $T$ to take the value $+\infty$)

1.2. Properties of stopping times:

1) if $S$ and $T$ are two stopping times so are $S \vee T$, $S \wedge T$ and $S + T$

2) if $S_n$ is a monotone sequence of stopping times, the limit $T = \lim_{n \to +\infty} S_n$ is also a stopping time.

1.3. The $\sigma$-field $\underline{\underline{F}}_T$. If $T$ is a stopping time, $\underline{\underline{F}}_T$ is the family of all the events $A \in \underline{\underline{F}}_\infty = \bigvee_t \underline{\underline{F}}_t$, such that for every $t \ge 0$ the event $A \cap \{T \le t\} \in \underline{\underline{F}}_t$.

It is easy to check that $\underline{\underline{F}}_T$ is a $\sigma$-field; it is intuitevely the $\sigma$-field of all the events that occurred up to time $T$. In particular, if $T$ is the constant stopping time $t$, $\underline{\underline{F}}_T = \underline{\underline{F}}_t$; if $S$ and $T$ are two stopping times, and if $S \le T$ a.e., then $\underline{\underline{F}}_S \subset \underline{\underline{F}}_T$.

If $T$ is a stopping time, and if $A \in \underline{\underline{F}}_T$, the r.v. $T_A$ defined by $T_A = T$ on $A$, $T_A = +\infty$ on $A^c$, is also a stopping time ($A^c$ denotes the complement of the set $A$).

Any stopping time can be approached strictly on the right by the sequence of stopping times $T_n = T + \frac{1}{n}$ (knowing everything up to the near future you know the present); the similar property on the left is false (knowing the strict past is not enough to know the present); the stopping times which can be thus announced are called predictable times.

1.4. Predictable times. A predictable time $T$ is a stopping time $T$ for which there exists a non decreasing sequence $(T_n)_{n \ge 0}$ of stopping times such that

$$\lim_{n\to+\infty} T_n = T \text{ a.e., and } \forall n \quad T_n < T \text{ a.e. on } \{T > 0\}.$$

We shall say that such a sequence $(T_n)$ announces the stopping time T.

Let T be a predictable time and $(T_n)$ a sequence announcing T; the σ-field $\underline{\underline{F}}_{T-} = \bigvee_n \underline{\underline{F}}_{T_n}$ is independent of the choice of the announcing sequence. It is the σ-field of the events occurring strictly before the time T. If $A \in \underline{\underline{F}}_{T-}$ the stopping time $T_A$ is also a predictable time. The σ-field $\underline{\underline{F}}_{T-}$ is contained in $\underline{\underline{F}}_T$, and if S is a stopping time and $S < T$ a.e., then $\underline{\underline{F}}_S \subset \underline{\underline{F}}_{T-}$.

1.5. Graph of a stopping time. If T is a stopping time, its graph $[\![T]\!]$ is the subset of $\mathbb{R}_+ \times \Omega$:

$$[\![T]\!] = \{(t,\omega);\ t = T(\omega) < +\infty\}.$$

1.6. Accessible time. An accessible time is a stopping time T, such that its graph $[\![T]\!]$ is contained in a countable union of graphs of predictable times. So there exists a partition $(A_n)$ of Ω such that on each $A_n$, the time T can be announced by a sequence $(S_{n,m})_{m\geq 0}$. But the sequence $(S_{n,m})$ depends on the set $A_n$. The time T is predictable if one can make the $(S_{n,m})$ independent of n.

1.7. Totally inaccessible time. A totally inaccessbile time is a stopping time T such that for every predictable time S, we have $P(T = S < +\infty) = 0$. In other words, one just cannot announce a totally inaccessbile time except on sets of measure zero.

1.8. Decomposition of stopping time. Let T be a stopping time; there exists a set $A \in \underline{\underline{F}}_T$ (unique in the sense that the difference of two such sets is of measure zero) such that $T_A$ is an accessible time, $T_{A^c}$ is a totally inaccessible time and $A \subset \{T < +\infty\}$.

STOCHASTIC PROCESSES

A stochastic process $X$ is a real valued function $(t,\omega) \to X_t(\omega)$ defined on $\mathbb{R}_+ \times \Omega$.

1.9. A stochastic process $Y$ is a version of a process $X$ if $\forall t \geq 0$ $P(Y_t \neq X_t) = 0$. If one looks at the values of two such processes $X$ and $Y$ at a countable number of times (which is the best one can do in reality) one can't tell them apart.

1.10. Two processes $X$ and $Y$ are indistinguishable if $P(\omega; \exists\ t$ such that $X_t(\omega) \neq Y_t(\omega)) = 0$. This is a much stronger property than the preceding one. In the following chapters we shall state theorems of the kind: "there exists a unique process such that...". It will mean, two processes having this property are indistinguishable.

1.11. A process $X$ is measurable if the application $(t,\omega) \to X_t(\omega)$ is $\underline{\underline{B}}(\mathbb{R}_+) \times \underline{\underline{F}}$ measurable ($\underline{\underline{B}}(\mathbb{R}_+)$ is the borelian $\sigma$-field on $\mathbb{R}_+$).

1.12. A process $X$ is adapted if for every $t \geq 0$ the application $\omega \to X_t(\omega)$ is $\underline{\underline{F}}_t$-measurable.

1.13. A process $X$ is progressively measurable if for each $t \geq 0$ the restriction of the application $(s,\omega) \to X_s(\omega)$ to the set $[0,t] \times \Omega$ is $\underline{\underline{B}}([0,t]) \times \underline{\underline{F}}_t$- measurable. Such a process is an adapted process.

Why is the notion of progressive measurability of any interest?

a) If $X$ is a stochastic process and $T$ is a stopping time, denotedby $X_T$ the r.v. $X_T(\omega) = X_{T(\omega)}(\omega)$; this r.v. is defined only on $\{T < +\infty\}$ (unless $X_\infty$ is defined in which case we take $X_T = X_\infty$ on $\{T = +\infty\}$). Assume that $X$ is an adapted process; is then $X_T I_{\{T<+\infty\}}$ a $\underline{\underline{F}}_T$-measurable function? No, in general; but if $X$ is progressively measurable, the r.v. $X_T I_{\{T<+\infty\}}$ is $\underline{\underline{F}}_T$-measurable.

b) Let $A$ be a progressively measurable set (i.e. $I_A$ is a progressively measurable process); then the r.v.

$$D_A(\omega) = \inf(t;\ (t,\omega) \in A)$$

is a stopping time (here we adopt the convention $\inf \phi = +\infty$). This last result is far from being trivial.

1.14. Càdlàg processes. A process $X$ is càdlàg if each of its trajectory $t \to X_t(\omega)$ is a right continuous function with finite left limits. For such a process we will denote by $X_{t-}$ the left limit at time t, and by $\Delta X_t = X_t - X_{t-}$ the jump at time $t$. The jumpsize will be $|\Delta X_t|$.

Any càdlàg adapted process is progressively measurable, and two càdlàg versions of the same process are indistinguishable.

Take a càdlàg process $X$, and define the r.v.

$$T_{1,0} = 0$$

$$T_{1,1} = \inf(t;\ |\Delta X_t| \geq 1)$$

$$T_{1,2} = \inf(t;\ t > T_{1,1},\ |\Delta X_t| \geq 1)$$

$$\cdots$$

$$T_{k,0} = 0$$

$$\cdots$$

$$T_{k,n} = \inf(t;\ t > T_{k,n-1}\ \tfrac{1}{k} \leq |\Delta X_t| < \tfrac{1}{k-1})$$

$$\cdots$$

In other words $T_{k,n}$ is the time of the $n^{th}$ jump of size $|\Delta X_t| \in [\frac{1}{k}, \frac{1}{k-1}]$. The processes $X_t$ and $X_{t-}$ are progressively measurable therefore the $T_{k,n}$ are stopping times. Each of the trajectories $t \to X_t(\omega)$ is a right continuous function with left limits; in a compact interval $[0,s]$ it has only a finite number of jumps of size bigger than a given $\varepsilon > 0$, and the set $U = \{(t,\omega);\ \Delta X_t(\omega) \neq 0\}$ is exactly the countable union of the graphs $\bigcup_{\substack{k \geq 1 \\ n \geq 1}} [T_{k,n}]$.

Each stopping time can be split into its totally inaccessible part and its accessible part. Each graph of an accessible time can be covered by a countable union of graphs of predictable times. And in the end we can find a countable number of totally inaccessible times $T_n$, and a countable number of predictable times $S_n$ such that

$$U = \{(t,\omega);\ \Delta X_t(\omega) \neq 0\} \subset \bigcup_n ([\![ S_n ]\!] \cup [\![ T_n ]\!]).$$

Moreover we can always assume that $P(T_n = T_m < +\infty) = 0$ and $P(S_n = S_m < +\infty) = 0 \quad n \neq m$. So we can cover the jump times of a càdlàg adapted process by a countable number of totally inaccessible, or predictable times. Note that at the totally inaccessible times $T_n$ we have $\Delta X_{T_n} \neq 0$ on $\{T_n < +\infty\}$, but at the predictable times $S_n$, $\Delta X_{S_n}$ can be zero on part of $\{S_n < +\infty\}$. This is what comes from using predictable times instead of accessible times.

1.15. Predictable $\sigma$-field. The predictable $\sigma$-field is the $\sigma$-field on $\mathbb{R}_+ \times \Omega$ generated by the left continuous adapted processes. This $\sigma$-field will be essential in stochastic integration (see chapters 3 and 5). A subset of $\mathbb{R}_+ \times \Omega$ is predictable if it belongs to the predictable $\sigma$-field. A process $X$ is predictable if the function $(t,\omega) \to X_t(\omega)$ is measurable with respect to the predictable $\sigma$-field. Any predictable process is progressively measurable.

It is handy to have some other systems of generators for the predictable $\sigma$-field. Here are two:

a) it is generated by the process of the form $\varphi_0^*(\omega) I_{\{0\}}(t) + \sum_{i=0}^{n-1} \varphi_i(\omega) I_{]t_i,t_{i+1}]}(t)$, where $0 \le t_0 < t_1 < \ldots < t_n < +\infty$, $\varphi_0^*$ is a bounded $\underline{\underline{F}}_0$-measurable r.v., and the r.v. $\varphi_i$ are bounded and $\underline{\underline{F}}_{t_i}$-measurable

b) it is also generated by the process of the form $\varphi_0^*(\omega) I_{\{0\}}(t) + \sum_{i=1}^{n-1} \varphi_i(\omega) I_{]\!]T_i,T_{i+1}]\!]}(t,\omega)$, where the $(T_i)$ form a nondecreasing finite sequence of stopping times, $\varphi_0^*$ is a bounded, $\underline{\underline{F}}_0$-measurable r.v., the $\varphi_i$ are $\underline{\underline{F}}_{T_i}$-measurable, bounded r.v. and $]\!]T_i,T_{i+1}]\!]$ is the stochastic interval $\{(t,\omega);\ T_i(\omega) < t \leq T_{i+1}(\omega)\}$.

If $X$ is a predictable process and $T$ a predictable time, the r.v. $X_T I_{\{T<+\infty\}}$ is $\underline{\underline{F}}_{T-}$-measurable (it is obvious for left continuous processes and extend easily to predictable processes).

1.16. Predictable times and predictable σ-fields. A r.v. $T$ is a predictable time if and only if its graph $[\![T]\!]$ is a predictable set (this is another non trivial result).

If $A$ is a predictable set, the r.v. $D_A(\omega) = \inf(t;\ (t,\omega) \in A)$ is a stopping time (1.13 and 1.15). If the graph $[\![D_A]\!]$ is included in the set $A$, $[\![D_A]\!] = A \setminus ]\!]D_A + \infty[\![$ is a predictable set and $D_A$ is a predictable time.

1.17. Càdlàg predictable processes. In particular, if $(X_t)$ is a càdlàg predictable process, the time $T_{k,n}$ of the $n^{th}$ jump of size $|\Delta X_t| \in [\frac{1}{k}, \frac{1}{k-1}]$ is a predictable time and $U = \{(t,\omega);\ \Delta X_t \neq 0\} = \bigcup_{\substack{n\geq 1\\ k\geq 1}} [\![T_{n,k}]\!]$. Furthermore the r.v. $X_{T_{n,k}}$ are $\underline{\underline{F}}_{T_{n,k}-}$-measurable (1.15).

1.13. Increasing processes and processes with finite variation. A process $A$ is an increasing process if

a) $A$ is adapted and càdlàg

b) $A_0 = 0$

c) $A_s \leq A_t$ for $s \leq t$.

A process $B$ is a process with finite variation if

a) $B$ is adapted and càdlàg

b) $B_0 = 0$

c) for each $\omega$, the trajectory $\omega \to B_t(\omega)$ has finite variation on compact intervals.

One can show that a process is a process with finite variation if and only if it is the difference of two increasing processes.

If $B$ is a process with finite variation, the Stieltjes integrals

$\int_0^t f(s)dB_s(\omega)$ exist for any bounded (or non negative) borelian function $f(s)$. The symbol $\int f(s)|dB_s|$ will denote the integral of $f(s)$ with respect to the variation of $B_s$. In particular $\int_0^t |dB_s(\omega)|$ is the variation of $B_s(\omega)$ on $[0,t]$.

An increasing process $A$ is <u>integrable</u> if $E[\int_0^\infty dA_t] < +\infty$. A process $B$ has <u>integrable variation</u> if $E[\int_0^\infty |dB_s|] < +\infty$.

If $B$ is a process with finite variation, the sums $\sum_{s\leq t} |\Delta B_s| \leq \int_0^t |dB_s|$ are finite; and the process $B$ is of the form

$$B_t = B_t^c + \sum_{s\leq t} \Delta B_s$$

where $B^c$ is a continuous process with finite variation (if $B$ is an increasing process, so is $B^c$). Using 1.14 we can write $\sum_{s\leq t} \Delta B_s$ in the form $\sum_n \Delta B_{T_n} I_{\{t\geq T_n\}}$, where $T_n$ is a sequence of stopping times.

1.19. <u>Predictable processes with finite variation</u>. Suppose now that B is a predictable process with finite variation, the stopping times $T_n$ can be taken predictable, and the $\Delta B_{T_n}$ are $\underline{\underline{F}}_{T_n-}$-measurable (see 1.17). Any <u>predictable process with finite variation</u> is therefore of the form

$$B_t = B_t^c + \sum \varphi_n I_{\{t\geq T_n\}}$$

where $B^c$ is a continuous process with finite variation, the $T_n$ are predictable times, the r.v. $\varphi_n$ are $\underline{\underline{F}}_{T_n-}$-measurable, and $\sum_n |\varphi_n| I_{\{t\geq T_n\}}$ exists for any $t$. The reader can check that conversely any process of this form is a predictable process with finite variation.

# CHAPTER II: MARTINGALES, LOCAL MARTINGALES AND SEMIMARTINGALES

We shall just give here the results necessary for Theorem 3.1 of chapter 3 which shows why semimartingales are important. The machinery on martingales and local martingales needed to construct the stochastic integrals will be seen in chapter 4.

## MARTINGALE, SUBMARTINGALE AND SUPERMARTINGALE

This section is just a summary of the classical results in martingale theory. The reader who is not familiar with the subject should consult [6] or [12].

2.1. Martingales. A martingale is an adapted process $M$ such that

a) $E[|M_t|] < +\infty \quad \forall t \geq 0$

b) $E[M_t | F_s] = M_s$ a.e. $\forall t \geq s$.

2.2. Sub and supermartingales. A super (resp. sub) martingale is an adapted process $M$ such that

a) $E[|M_t|] < +\infty \quad \forall t \geq 0$

b) $E[M_t | \underline{\underline{F}}_s] \leq M_s$ (resp. $\geq M_s$) a.e. $\forall t \geq s$.

If $M_t$ is the capital of a gambler a time $t$ the notion of martingale (resp. sub, resp. super) corresponds to the notion of fair (resp. favorable, resp. unfavorable game).

2.3. Càdlàg versions of martingales. Any martingale $M$ has a càdlàg version; therefore the term "martingale" will from now on mean "càdlàg martingale".

**2.4. If $X$ is a supermartingale (non necessarily càdlàg), for almost all $\omega$, the two limits**

$$X_{t_+} = \lim_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} \quad \text{and} \quad X_{t_-} = \lim_{\substack{s \to t \\ s < t \\ s \in \mathbb{Q}}} X_s$$

exist for each $t \in \mathbb{R}_+$ (the limits are taken over the set $\mathbb{Q}$ of the rational numbers). The process $(X_{t_+})$ is then indistinguishable from a càdlàg supermartingale. The supermartingale $(X_t)$ has a right continuous version if and only if the function $t \to E[X_t]$ is right continuous. <u>We shall always</u>, except when otherwise specified, <u>consider càdlàg supermartingales, and call them supermartingales for short</u>.

2.5. A martingale $M$ is said to be <u>uniformly integrable</u> if the family of r.v. $(M_t)_{t \geq 0}$ is uniformly integrable. For any uniformly integrable martingale $M$, the limit $M_\infty = \lim_{t \to +\infty} M_t$ exists a.e.; and for any stopping time $T$, we then have $M_T = E[M_\infty | \underline{\underline{F}}_T]$. Apply this result to a sequence $S_n$ announcing a predictable time $S$. We get $M_{S_n} = E[M_\infty | \underline{\underline{F}}_{S_n}] = E[M_S | \underline{\underline{F}}_{S_n}]$ for any $n$; and by taking limits on both sides, $E[M_S | \underline{\underline{F}}_{S-}] = M_{S-}$. So if $M$ is a uniformly integrable martingale and $S$ a predictable time, the jump at time $S$ verifies $E[\Delta M_S | \underline{\underline{F}}_{S-}] = 0$.

2.6. Let $X$ be a non negative supermartingale, and take $X_\infty = 0$, then $(X_t)_{0 \leq t \leq +\infty}$ is a supermartingale, and for any two stopping times $T$ and $S$ such that $S \leq T$ we have $X_S$ and $X_T \in L^1$, and $E[X_T | \underline{\underline{F}}_S] \leq X_S$.

2.7. Let $M$ be a non negative martingale, and let $T = \inf(t; M_t \text{ or } M_{t-} = 0)$, then $M = 0$ a.e. on $[\![T, +\infty]\!]$. In particular if $M_\infty = \lim_{t \to +\infty} M_t$ exists and if $M_\infty > 0$ a.e., we have $T = +\infty$ a.e. that is $P\{(\omega;\ \exists t \text{ such that } M_t(\omega) \text{ or } M_{t-}(\omega) = 0)\} = 0$.

2.8. <u>Jensen's inequality</u>. If $M$ is a martingale and $f(x)$ is a convex function, the process $f(M)$ is a submartingale provided $E[|f(M_t)|] < +\infty$, $\forall\, t \geq 0$.

2.9. <u>Doob's decomposition theorem</u>.

If $M$ is a uniformly integrable martingale, and $A$ is an integrable increasing process, the process $X = M - A$ is a supermartingale, satisfying the strong following integrability condition: the family of r.v. $\{X_T I_{\{T<+\infty\}}$, $T$ stopping time$\}$ is a uniformly integrable family. We shall call those supermartingales, supermartingales of class (D). Doob's decomposition theorem is just the converse statement: any supermartingale $X$ of class (D) is of the form

$$X = M - A$$

where $M$ is a uniformly integrable martingale, and $A$ is a predictable, integrable, increasing process. And this decomposition is unique. See [12], [4] and [14] for three different proofs of this theorem.

2.10. Corollaries.

1) Let $X$ be a supermartingale of class (D), and $B$ be the predictable increasing process in Doob's decomposition. The process $B$ jumps only at predictable times; at such a predictable time $T$, $\Delta B_T$ is $\underline{\underline{F}}_{T-}$-measurable (see 1.19) and we have, (2.5), if $M$ is the uniformly integrable martingale $M = X + B$

$$0 = E[\Delta M_T | \underline{\underline{F}}_{T-}] = E[\Delta X_T + \Delta B_T | \underline{\underline{F}}_{T-}] = E[\Delta X_T | \underline{\underline{F}}_{T-}] + \Delta B_T$$

And the jumps of $B$ are easy to compute.

2) If $A$ is an integrable increasing process the process $-A$ is a supermartingale of class (D); therefore there exists a unique integrable predictable increasing process $B$ such that $B - A$ is a uniformly integrable martingale. The process $B$ is called the compensator of $A$.

This generalizes to processes with integrable variation. If $A$ is such a process, there exists a unique predictable process $B$ with integrable variation such that $B - A$ is a uniformly integrable martingale.

The process $B$ is again called the compensators of $A$. From part 1 of this corollary we get:

a) if $T$ is a totally inaccessible time, and $\varphi$ an $\underline{\underline{F}}_T$-measurable function in $L^1$, the compensator of $A_t = \varphi I_{\{t \geq T\}}$ is a continuous process with integrable variation.

b) if $T$ is a predictable time, and $\varphi$ is an $\underline{\underline{F}}_T$-measurable function in $L^1$, the compensator of $A_t = \varphi I_{\{t \geq T\}}$ is the process $B_t = E[\varphi | \underline{\underline{F}}_{T-}] I_{\{t \geq T\}}$

LOCAL MARTINGALES AND PROCESSES WITH LOCALLY INTEGRABLE VARIATION

Let $X$ be a stochastic process, and $T$ a stopping time. The symbol $X^T$ will denote the process $X$ stopped at time $T$ : $X^T(\omega) = X_{t \wedge T}(\omega)$. A process $M$ is a martingale if and only if, for any constant time $n$, the process $M^n$ is a uniformly integrable martingale. And it is natural to let the constant times $n$ be stopping times $T_n$:

2.11. Definition. A localizing sequence is a nondecreasing sequence $(T_n)$ of stopping times such that $\lim_n T_n = +\infty$ a.e.

2.12. Definition. A process $M$ is a local martingale if

a) $M_0 = 0$

b) there exists a localizing sequence $(T_n)$ such that each process $M^{T_n}$ is a uniformly integrable martingale.

Such a sequence $(T_n)$ will be called a fundamental sequence for the local martingale $M$.

Remark. 1) Local martingales are necessarily càdlàg processes as we decided that here "martingale" means "càdlàg martingale".

2) The processes defined above should really be called "local martingales vanishing at time zero". We shall not use here the general

concept of local martingales. The interested reader can consult [3].

2.13. Definition. A stopping time $T$ reduces a local martingale $M$ if $M^T$ is a uniformly integrable martingale.

2.14. Theorem. Let $M$ be a local martingale then

1) a stopping time $S$ reduces $M$ if and only if the process $M^S$ is of class (D) (i.e. the family of r.v. $\{M^S_T I_{\{T<+\infty\}}$ ; $T$ stopping time$\}$ is uniformly integrable;

2) if $T$ is a stopping time reducing $M$, if $S$ is a stopping time and if $S \leq T$, then $S$ reduces $T$.

3) if $S$ and $T$ are two stopping times reducing $M$, then $S \vee T$ reduces $M$.

Proof. Parts 1 and 2 are trivial. Part 3 comes from the fact that $M^{S\vee T} = M^S + M^T - M^{S\wedge T}$.

2.15. Theorem. If a process $M$ is locally a local martingale, then it is a local martingale.

So there is no way one can get more general processes by localizing once more.

Proof. There exists a localizing sequence $(T_n)$ such that each process $M^{T_n}$ is a local martingale. Let $\underline{\underline{H}} = \{T;\ T$ stopping time, $M^T$ is a uniformly integrable martingale$\}$. And take $R = \operatorname{ess.sup}_{T\in \underline{\underline{H}}} T$. There exists a sequence $S_n$ of elements of $\underline{\underline{H}}$ which converges a.e. to $R$. Using part 3 of 2.14 we can make this sequence non decreasing. The r.v. $R$ has to be bigger than or equal to any of the $T_n$ (a.e.), so $R = +\infty$, and $S_n$ is a fundamental sequence for the process $M$.

2.16. Definition. A process $B$ has locally integrable variation if there exists a localizing sequence $(T_n)$ such that each process $B^{T_n}$ has integrable variation.

2.17. Theorem. Let $B$ be a process with locally integrable variation, there exists a unique predictable process $A$ with locally integrable variation such that $B - A$ is a local martingale. $A$ is called the compensator of $B$.

Proof. Easy consequence of the existence and uniqueness of the compensator of a process with integrable variation.

2.18. Remark. It is important to remark the following fact. If $B$ is a predictable process with finite variation, then the variation of $B$ is locally bounded: define the stopping times $T_n = \inf(t; \int_0^t |dB_s| \geq n) \wedge n$. The variation of $B$ on $[\![0,T_n[\![$ is bounded by $n$, but we know nothing on the jump of $B$ at time $T_n$. Now each time $T_n$ is predictable and can be announced by a sequence $(S_{n,m})_{m \geq 0}$; on $[\![0,S_{n,m}]\!]$ the variation of $B$ is bounded by $n$. Take the stopping times $R_k = \sup_{\substack{n \leq k \\ m \leq k}} S_{n,m}$. The sequence $(R_k)$ is a localizing sequence and on $[\![0,R_k]\!]$ the variation of $B$ is bounded by $k$.

Let $A$ be a process with finite variation. If there exists a predictable process $B$ with finite variation such that $A - B$ is a local martingale, then $B$ has locally bounded variation, and $A$ itself has locally integrable variation. We can rewrite theorem 2.17 in a stronger form

2.19. Theorem. A process $A$ with finite variation has a compensator if and only if its variation is locally integrable.

Here is now an easy way to verify that the variation is locally integrable.

2.20. Lemma. Let $A$ be a process with finite variation; we assume that there exists a localizing sequence $(T_n)$ such that for each $n$, $\sup_{s \leq T_n} |\Delta A_s| = Y_n \in L^1$. Then the variation of $A$ is locally integrable.

Proof. Take $S_n = T_n \wedge \inf(t; \int_0^t |dA_s| \geq n)$. The sequence $(S_n)$ is localizing, the variation of $A$ on $[\![0,S_n[\![$ is bounded by $n$, and

$$\int_{[\![0,S_n]\!]} |dA_s| \leq n + |\Delta A_{S_n}| \leq n + Y_n \in L^1.$$

We shall now state the fundamental lemma for local martingales.

2.21. Fundamental lemma. Let $M$ be a local martingale then

1) the increasing process $M_t^* = \sup_{s<t} |M_s|$ is locally integrable

2) the local martingale $M$ can be written in the form $M = U + V$ where $U$ is a local martingale, the jumps of $U$ are bounded by $1$ in size, and $V$ is both a local martingale and a process with finite variation. (The bound $1$ for the jump size could have been replaced by another strictly positive constant). In particular there exists a localizing sequence $(T_n)$ such that each $U^{T_n}$ is a bounded process.

Proof. 1) Let $R_n$ be a fundamental sequence for $M$. We can always assume that the $R_n$ are finite (otherwise use the fundamental sequence $R_n \wedge n$); we consider the following stopping times

$$S_n = R_n \wedge \inf\{t; |M_t| \geq n\}.$$

The martingales $M^{R_n}$ are uniformly integrable, and $S_n \leq R_n$, therefore the r.v. $M_{S_n}$ is integrable (2.5). Furthermore on $[\![0,S_n[\![$, we have $|M_t| \leq n$; and

$$M_{S_n}^* \leq n + |M_{S_n}| \in L^1.$$

As the sequence $(S_n)$ is a localizing sequence, the increasing process $M_t^*$ is locally integrable.

2) Let $A_t = \sum_{s<t} \Delta M_s I_{\{|\Delta M_s| > \frac{1}{2}\}}$. This sum is, for each $\omega$, a finite sum. Take the sequence $(S_n)$ constructed above and consider the stopping

times

$$T_n = S_n \wedge \inf\{t;\ \int_0^t |dA_s| \geq n\}.$$

The sequence $(T_n)$ is a localizing sequence, and $\int_{[\![0,T_n]\!]} |dA_t| \leq n + |\Delta A_{T_n}| \leq n + 2M^*_{T_n} \in L^1$. There exists therefore a compensator $B$ for $A$. Take $V = A - B$, $V$ is both a local martingale and a process with finite variation.

The jumps of $B$ occur only at predictable times $T$. Stop all the processes at the time $S_n$, and remember that $M^{S_n}$ is a uniformly integrable martingale. If $T$ is a predictable time, we get

$$|\Delta B_T^{S_n}| = |E[\Delta A_T^{S_n}|\underline{\underline{F}}_{T-}] = |E[\Delta M_T^{S_n} - \Delta M_T^{S_n} I_{\{|\Delta M_T|<\frac{1}{2}\}}|\underline{\underline{F}}_{T-}]| \leq 0 + \frac{1}{2}.$$

And the jumps of $U$ verify

$$|\Delta U_t| \leq |\Delta(M - A)_t| + |\Delta B_t| \leq 1$$

It is now easy to see that the sequence $(R_n)$, $R_n = \inf(t;\ |U_t| \geq n)$, is a localizing sequence and that $U^{R_n}$ is bounded by $n + 1$.

## SEMIMARTINGALES

2.22. <u>Definition</u>. <u>A process $X$ is a semimartingale if it is of the form $X = X_0 + M + A$, where $X_0$ is an $\underline{\underline{F}}_0$-measurable r.v., $M$ is a local martingale and $A$ is a process with finite variation (remember that by definition both processes $M$ and $A$ vanish at $t = 0$).</u>

This definition contains no local integrability condition. It should not if we want the semimartingales to remain semimartingales when the probability $P$ is replaced by an equivalent probability $Q$.

Here again one cannot get more general processes by localizing the notion of semimartingale

2.23. Theorem. Any process which is locally a semimartingale is a semimartingale.

Proof. We shall need the following useful lemma

2.24. Lemma. Let $X$ be a semimartingale, assume that the size of the jumps of $X$ is bounded by $a$ $(a > 0)$. Then one can write $X$ in a unique way as

$$X = X_0 + M + A$$

where $X_0$ is an $\underline{\underline{F}}_0$-measurable r.v., $M$ is a local martingale, and $A$ is a predictable process with finite variation (in fact, locally bounded variation by 2.18).

Proof. The semimartingale is of the form $X = X_0 + N + B$ where $N$ is a local martingale and $B$ a process with finite variation. The r.v. $X_0$ is uniquely determined. The jumps of the process $B$ verify

$$|\Delta B_s| \leq |\Delta M_s| + |\Delta X_s| \leq 2M_s^* + a.$$

The increasing process $Y_t = \sup_{s \leq t} |\Delta B_s|$ is locally integrable and by 2.20, the variation of $B$ is locally integrable. Let $A$ be the compensator of $B$, and let $M = N + B - A$; $M$ is a local martingale, $A$ is a predictable process with finite variation, and $X = X_0 + M + A$.

If $X = X_0 + M + A = X_0 + M' + A'$ are two such decompositions of the semimartingale $X$, $A - A'$ is a local martingale, so $A'$ is the compensator of $A$. But $A$ being predictable is its own compensator and $A = A'$.

Proof of 2.23. The process $X$ is locally a semimartingale. That is, there exists a localizing sequence of stopping times $(T_n)$, such that each $X^{T_n}$ is a semimartingale.

Take $Y_t = \sum_{s \leq t} \Delta X_s I_{\{|\Delta X_s| \geq 1\}}$. As $X$ is càdlàg the process $Y$ has finite variation; for each $n$ the process $X^{T_n} - Y^{T_n}$ is a semimartingale with jump size smaller than 1. It can therefore be written in a unique way as $X^{T_n} - Y^{T_n} = X_0 + M_n + A_n$, where $M_n$ is a local martingale and $A_n$ is a predictable process with finite variation. The uniqueness of the $A_n$ gives

$$A_n = A_{n+1}, \quad \text{and} \quad M_n = M_{n+1} \quad \text{on} \quad [\![0,T_n]\!].$$

One can patch the $A_n$ together, and the $M_n$ together to get the processes $A = \sum_n A_n I_{]\!]T_{n-1},T_n]\!]}$ $M = \sum_n M_n I_{]\!]T_{n-1},T_n]\!]}$. The process $A$ is predictable and has finite variation, the process $M$ is locally a local martingale, therefore it is a martingale (2.15) and $X = X_0 + M + A + Y$ is a semimartingale.

The following lemma is important.

2.25. Lemma. *Let $X$ be a càdlàg adapted process; we assume that there exists a sequence of stopping times $T_n$ and a sequence of semimartingales $Y_n$ such that*

1) $\lim_{n\to+\infty} T_n = +\infty$ a.e.
2) $X_n = Y_n$ on $[\![0,T_n[\![$.

*Then $X$ is a semimartingale.*

Proof. $X^{T_n} = Y_n^{T_n} - Y_{n,T_n} I_{\{t\geq T_n\}} + X_{T_n} I_{\{t\geq T_n\}}$. So for each $n$ the process $X^{T_n}$ is a semimartingale. If the sequence $T_n$ is non decreasing, 2.23 says that $X$ is a semimartingale. Otherwise, make the sequence non decreasing by remarking that if $S$ and $T$ are two stopping times, and if $X^S$ and $X^T$ are both semimartingales, then $X^{S\wedge T}$ and $X^{S\vee T} = X^S + X^T - X^{S\wedge T}$ are both semimartingales.

2.26. Examples of semimartingales.

1) any supermartingale (therefore any submartingale) is a semimartingale: let $X$ be a supermartingale, by 2.24 we just have to show that each $X^n$ is a semimartingale. But $X_t^n = E[X_n|\underline{\underline{F}}_t] + X_t^n - E[X_n|\underline{\underline{F}}_t]$,

and we just have to work with the non negative supermartingale $Z_t = X_t^n - E[X_n|\underline{\underline{F}}_t]$. For any stopping time $S$ we have $E[|Z_S|] < +\infty$ (2.6); consider the process $Z_t^* = \sup_{s<t}|Z_s|$, and the stopping times $T_n = \inf(t; |Z_s^*| \geq n\} \wedge n$. On $[\![0,T_n]\!]$ we have $|Z_t^*| \leq n + |Z_{T_n}| \in L^1$. So the supermartingale $Z$ is locally of class (D). This and the uniqueness in Doob's decomposition theorem 2.9 gives that $Z$ is of the form $Z = M - A$, where $M$ is a local martingale and $A$ is a predictable process with finite variation. Therefore $Z$ is a semimartingale, and so is $X$.

2) Let $X$ be a càdlàg process with independent increments. As $X$ is càdlàg the process $Z_t = \sum_{s\leq t} \Delta X_s I_{\{|\Delta X_s| \geq 1\}}$ has finite variation, and has independent increments; the process $Y = X - Z$ is a process with independent increments, and its jumps are all bounded in size by 1. Therefore $E[Y_t]$ exists, and $Y_t - E[Y_t]$ is a martingale. We see then that the process $X$ is a semimartingale if and only if the function $t \to E[Y_t]$ has finite variation.

We get now to the interesting result on semimartingales.

2.27. Theorem. Let $P$ and $Q$ be two equivalent probabilities on $(\Omega,\underline{\underline{F}})$. Then the semimartingales for $P$ are the semimartingales for $Q$.

We know that this result is false if we replace semimartingales by local martingales.

Proof. It is enough to show that any P-local martingale is a Q-semimartingale.

Consider the Radon Nikodym derivative $M_\infty = \frac{dQ}{dP}$ on $(\Omega,\underline{\underline{F}})$, and let $M$ be a càdlàg version of the P-martingale $E_P[M_\infty|\underline{\underline{F}}_t]$. As $P(M_\infty = 0) = 0$, the martingale is $P$ (and $Q$) a.e. strictly positive. It is easy to check that $X$ is a P-local martingale if and only if $\frac{X}{M}$ is a Q-local martingale. Write $X$ as $X = \frac{X}{M} M$. The process $\frac{X}{M}$ is a Q-local martingale, therefore a Q-semimartingale. The process $\frac{1}{M}$ is a Q-martingale (as 1 is a P-martingale). All we should know now is: 1) if $N$ is a strictly positive Q-local martingale, then $\frac{1}{N}$ is a Q-semimartingale (lemma 2.28); 2) the product of two Q-semimartingales is a Q-semimartingale (lemma 2.29).

2.28. <u>Lemma</u>. <u>If $N$ is a strictly positive Q-local martingale, then $\frac{1}{N}$ is a Q-semimartingale</u>.

<u>Proof</u>. If $N_t \geq a > 0$, everything is trivial: $\frac{1}{N_t} \leq \frac{1}{a}$, so $E[\frac{1}{N_t}] < +\infty\ \forall t$; as the function $\frac{1}{x}$ is convex, Jensen's inequality (2.8) gives that $\frac{1}{N}$ is a Q-submartingale, therefore a Q-semimartingale. Otherwise, consider the functions $f_n(x)$ defined as follows: $f_n(x)$ is the function $\frac{1}{x}$ on $[\frac{1}{n}, +\infty]$. On $[0,\frac{1}{n}]$ the graph of $f_n$ is the tangent to the curve $y = \frac{1}{x}$ at the point $\frac{1}{n}$. The functions $f_n$ are convex, and $f_n(N_t) \in L^1\ \forall t \geq 0$. So the processes $Y_n = f_n(N)$ are Q-submartingales. Consider the stopping times $T_n = \inf(t;\ N_t \leq \frac{1}{n})$. The sequence $(T_n)$ is a localizing sequence; the processes $Y_n$ and $\frac{1}{N}$ coincide on $[\![0,T_n[\![$, so $\frac{1}{N}$ itself is a Q-semimartingale (lemma 2.25).

2.29. <u>Lemma</u>. <u>The product of two semimartingales is a semimartingale</u>.

<u>Proof</u>. It is enough to show that the square of a semimartingale is a semimartingale (use $(a + b)^2 = a^2 + 2ab + b^2$). Let $X$ be a semimartingale, $X$ can be written in the form $X = M + V$ where $M$ is a locally bounded martingale, and $V$ is a process with finite variation (lemma 2.21). Let $(T_n)$ be a localizing sequence such that for each $n$, $M^{T_n}$ is a bounded martingale, and the variation of $V^{T_n}$ is bounded on $[0,T_n[$. We just have to show that each process $(M^{T_n} + V^{T_n})^2$ is a semimartingale.

The first term $(M^{T_n})^2$ is a submartingale by Jensen's inequality (2.8). The term $(V^{T_n})^2$ is a process with finite variation as, if $0 \leq t_0 < t_1 \ldots < t_n = t$ is a subdivision of $[0,t]$, we have

$$\sum_i |(V^{T_n}_{t_{i+1}})^2 - (V^{T_n}_{t_i})^2| = \sum_i |V^{T_n}_{t_{i+1}} + V^{T_n}_{t_i}|\,|V^{T_n}_{t_{i+1}} - V^{T_n}_{t_i}|$$

$$\leq 2\int_0^t |dV^{T_n}_s| \left(\sum_i |V^{T_n}_{t_{i+1}} - V^{T_n}_{t_i}|\right) \leq 2\left(\int_0^t |dV^{T_n}_s|\right)^2.$$

That leaves the term $M^{T_n}V^{T_n}$. The process $M^{T_n}$ is the difference of the two bounded non negative martingales $E[M^+_{T_n}|\underline{\underline{F}}_t]$ and $E[M^-_{T_n}|\underline{\underline{F}}_t]$; the process $V^{T_n}$ is the difference of two increasing processes. So we just have to look at the case where $M^{T_n}$ is a non negative bounded martingale, and $V^{T_n}$ is an increasing process such that $V^{T_n}_{T_{n-}}$ is bounded. The process $W^n_t = V^{T_n}_t - \Delta V_{T_n} I_{\{t \geq T_n\}}$ is a bounded increasing process, and $M^{T_n}W^n$ is a submartingale. As for each $n$ $M^{T_n}V^{T_n} = M^{T_n}W^n$ on $[\![0,T_n[\![$, the product $MV$ is a semimartingale (lemma 2.25).

Now one more definition and one more theorem and we will be finished with this long chapter.

2.30. Definitions. 1) A subdivision of $[0,+\infty]$ is any finite sequence $\tau = (t_0,t_1,\ldots,t_n)$ such that $0 \leq t_0 < t_1 <\ldots< t_n \leq +\infty$.

2) Let $X$ be a process defined on $[0,+\infty]$, we shall denote by $\mathrm{var}(X)$ the quantity

$$\mathrm{var}(X) = \sup_\tau E\left[\sum_{i=0}^{n} |E[X_{t_{i+1}} - X_{t_i}|\underline{\underline{F}}_{t_i}]|\right]$$

the sup being taken over all the subdivisions of $[0,+\infty]$

3) Let $X$ be a right continuous, adapted process defined on $[0,+\infty]$. The process $X$ is a quasimartingale on $[0,+\infty]$ if $E[|X_t|] < +\infty$, $\forall t \in [0,+\infty]$ and if $\mathrm{var}(X) < +\infty$

2.31. Theorem. Let $(X_t)_{t\geq 0}$ be an adapted, right continuous process. We extend $X$ by taking $X_\infty = 0$. The process $X$ is a quasimartingale on $[0,+\infty]$ if and only if $X$ is of the form $X = Y - Z$ where $Y$ and $Z$ are two càdlàg nonnegative supermartingales.

<u>Proof</u>. The only non trivial part is the necessary condition. For all $s \geq 0$ let $\sum(s)$ be the set of all the subdivisions of $[s,+\infty]$. For any $\tau = (t_0, t_1, \dots, t_n) \in \sum(s)$ consider

$$Y_s^\tau = E[\sum_{i=0}^{n} (E[X_{t_i} \quad X_{t_{i+1}} | \underline{\underline{F}}_{t_i}])^+ | \underline{\underline{F}}_s]$$

$$Z_s^\tau = E[\sum_{i=0}^{n} (E[X_{t_i} - X_{t_{i+1}} | \underline{\underline{F}}_{t_i})^- | \underline{\underline{F}}_s]$$

It is easy to see that $Y_s^\sigma \leq Y_s^\tau$ if the subdivision $\tau$ is finer than the subdivision $\sigma$. The quantities $E[Y_s^\tau]$ are bounded by $\mathrm{var}(X)$, so $Y_s^0 = \lim_\tau Y_s^\tau$ exists in $L^1$ (the limit is taken over the directed set $\sum(s)$ with the partial ordering $\sigma < \tau$ if $\tau$ is finer than $\sigma$.

The r.v. $Y_s^0$ can be computed by using only the subdivisions $\tau$ such that $t_0 = s$ and $t_n = +\infty$. For such a subdivision we have $Y_s^\tau - Z_s^\tau = X_s$, so $Y_s^0 - Z_s^0 = X_s$.

It is easy to see that for $s < t$ we have $Y_s^0 \geq E[Y_t^0 | \underline{\underline{F}}_s]$ and $Z_s^0 \geq E[Z_t^0 | \underline{\underline{F}}_s]$ a.e. The processes $Y^0$ and $Z^0$ are supermartingales which might not be càdlàg. But (see 2.4) for almost all $\omega$ the limits

$$Y_{t_+} = \lim_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} Y_s^0 \quad \text{and} \quad Z_{t_+} = \lim_{\substack{s \to t \\ s > t \\ s \in \mathbb{Q}}} Z_s^0 \quad \text{exist for all } t \in \mathbb{R}_+ .$$

The processes $Y_{t_+}$ and $Z_{t_+}$ are indistinguishable from càdlàg, non negative supermartingales; and as $X_t$ is right continuous the equality $X_t = Y_t^0 - Z_t^0$ becomes $X_t = Y_{t_+} - Z_{t_+}$.

2.32. <u>Corollary</u>. <u>Let $X$ be a right continuous, adapted process defined on $[0,+\infty]$. If $X$ is a quasimartingale on $[0,+\infty]$, then $X$ is of the form</u>

$$X = M + Y - Z$$

<u>where $M$ is a martingale on $[0,+\infty]$, and $Y$ and $Z$ are two non negative supermartingales on $[0,+\infty]$.</u>

Proof. Let $M_t$ be a càdlàg version of the martingale $E[X_\infty|\underline{\underline{F}}_t]$. The processes $M$ and $X - M$ are quasimartingales on $[0,+\infty]$. As $X_\infty = M_\infty$, we finish using theorem 2.31.

## CHAPTER III: WHY SEMIMARTINGALES?

Let $X$ be a right continuous process. It is natural to define $\int_{]0,t]} dX_s$ as $\int_{]0,t]} dX_s = \int_0^\infty I_{]0,t]} dX_s = X_t$. For a process $\phi = \varphi_0^* I_{\{0\}} + \varphi_0 I_{]0,t_1]} + \varphi_1 I_{]t_1,t_2]} + \ldots + \varphi_{k-1} I_{]t_{k-1},t_k]}$, where $\varphi_0^*$ and the $\varphi_i$ are r.v., the integral $\int_{]0,\infty[} \phi_s dX_s$ should be $\varphi_0(X_{t_1} - X_0) + \ldots + \varphi_k(X_{t_k} - X_{t_{k-1}})$. Nowdays the most recent trend is to extend all the processes to $\mathbb{R}$ by taking $X_t = 0$, $\phi_t = 0$ for $t < 0$; the process $X$ has then a jump at time $0$ and $\int_{[0,\infty[} \phi_s dX_s = \varphi_0^* + \int_{]0,\infty[} \phi_s dX_s$. We will not bother about this possible jump at $t = 0$, and $\int_0^t$ will always mean $\int_{]0,t]}$. Similarly $\int_s^t$ will mean $\int_{]s,t]}$.

We have a filtration $(\underline{\underline{F}}_t)$ on $(\Omega, \underline{\underline{F}})$ so we will also assume that the process $X$ is adapted; in that case the interesting processes $\phi$ are those for which $\varphi_0^*$ is $\underline{\underline{F}}_0$-measurable and the $\varphi_i$ are $\underline{\underline{F}}_{t_i}$-measurable. So that $\int_0^t \phi_s dX_s$ is itself $\underline{\underline{F}}_t$-measurable.

Let $\underline{\underline{B}}_t$ be the set of all processes $\phi$, $\phi = \varphi_0^* I_{\{0\}} + \varphi_0 I_{]0,t_1]} + \ldots + \varphi_{k-1} I_{]t_{k-1},t_k]}$, such that $\varphi_0^*$ is an $\underline{\underline{F}}_0$-measurable bounded r.v., the $\varphi_i^*$ are $\underline{\underline{F}}_{t_i}$-measurable, bounded r.v. and $t_i \le t$ $\forall i \le k$; the processes in $\underline{\underline{B}}_t$ vanish on $]t,+\infty[$; we put on $\underline{\underline{B}}_t$ the topology of the uniform convergence.

We denote by $L^0$ the vector space of the classes of r.v. with the topology of the convergence in probability. The topological vector space $L^0$ is metrisable and complete; if $\|f\|_0 = E[|f| \wedge 1]$, we get a fundamental system of neighborhoods of zero by taking the sets

$\mathcal{V}_{\varepsilon} = \{f;\ \|f\|_0 \leq \varepsilon\}$.

If $J(\phi) = \int_0^{\infty}\phi_s dX_s$ is going to have the properties of an integral, we should have at least: if the processes $\phi_n$ in $\underline{\underline{B}}_t$ converge uniformly to a process $\phi$ in $\underline{\underline{B}}_t$, then $J(\phi_n)$ converges in probability to $J(\phi)$. Now, as we are going to see, this implies that $X$ is a semimartingale.

This result is quite recent and has been proved by Dellacherie with the collaboration of Mokobodzki. Later, Letta gave a variant of the proof, which uses less analysis; this is the one we shall give here. To simplify things a little, we shall assume that the process $X$ is càdlàg, despite the fact that Meyer remarked that right continuity of $X$ is enough.

3.1. Theorem. Let $X$ be a càdlàg adapted process, if for any $t \geq 0$, the application $J$ from $\underline{\underline{B}}_t$ into $L^0$ is continuous, then $X$ is a semimartingale.

Proof. It is enough to show that, on each $[0,t]$, the process $X$ is a semimartingale. So we are going to transform $[0,t]$ into $[0,+\infty]$: the càdlàg, adapted process $X$ is defined on $[0,+\infty]$, $\underline{\underline{B}}$ is the set of the processes of the form $\phi = \varphi_0^* I_{\{0\}} + \varphi_0 I_{]0,t_1]} + \dots + \varphi_{k-1} I_{]t_{k-1},t_k]}$, where $0 < t_1 < \dots < t_k \leq +\infty$, $\varphi_0^*$ is an $\underline{\underline{F}}_0$-measurable bounded r.v., and the $\varphi_i$ are $F_{t_i}$-measurable bounded r.v. On $\underline{\underline{B}}$ we con ider the topology of the uniform convergence, and our assumption is "$J$ is a continuous function from $\underline{\underline{B}}$ into $L^0$". To show that $X$ is a semimartingale on $[0,+\infty]$, we shall show that it is a quasimartingale on $[0,+\infty]$ for an equivalent probability $Q$ and then use 2.32, 2.26 and 2.27.

As $X$ is a càdlàg process on $[0,+\infty]$, the r.v. $X^*_{\infty} = \sup_{s \leq +\infty} |X_s|$ is finite. We can find a probability $P'$ equivalent to the probability $P$ such that $\int |X^*_{\infty}| dP' < +\infty$.

Let $D_1$ be the unit ball of $\underline{\underline{B}}$ for the norm of the uniform convergence. And let $J(D_1) = A$. The set $A$ is a convex and bounded set of

$L^0$ (this as $J$ is linear and continuous). Furthermore $A$ is in $L^1(P')$. We shall show (see lemma 3.4) that there exists a probability $Q$ equivalent to $P'$ such that $L^1(P') \subset L^1(Q)$ and such that $\sup\limits_{f \in A} \int f dQ < +\infty$.

For the probability $Q$ we then have

$$\begin{aligned}
\mathrm{Var}_Q(X) &= \sup_{\tau} E_Q[\sum |E_Q[X_{t_{i+1}} - X_{t_i} | \underline{\underline{F}}_{t_i}]|] \\
&= \sup_{\phi \in D_1} E_Q[\sum \varphi_i (X_{t_{i+1}} - X_{t_i})] = \sup_{\phi \in D_1} E_Q[\int_0^\infty \phi dX_s] \\
&= \sup_{f \in A} E_Q[f] < +\infty .
\end{aligned}$$

The process $X$ is then a $Q$ quasimartingale, and therefore a $P$ semimartingale.

Let us show the existence of the probability $Q$.

3.2. Lemma. Let $A$ be a non empty convex set of $L^1$ and $K$ be a non empty convex set of $L^\infty$; we assume that $K$ is compact for the weak topology $\sigma(L^\infty, L^1)$ and that for a certain number $c$ in $\mathbb{R}$, and for any $x \in A$, the sets $H_x = \{y;\ y \in K,\ \langle y,x\rangle \leq c\}$ are all non empty. Then $\bigcap\limits_{x \in A} H_x \neq \phi$.

Proof. $H_x$ is weakly closed in $K$, so it is a weakly compact set, and we just have to show that any finite intersection of the sets $H_x$ is non empty. If $x_1, x_2, \ldots, x_n$ are in $A$, and if $\bigcap\limits_{i=1}^{n} H_{x_i} = \phi$, the two sets of $\mathbb{R}^n$, $L = \{(\langle y,x_i\rangle)_{i=1,\ldots,n};\ y \in K\}$ and $M = ]-\infty, c]^n$ are disjoint. $L$ is a convex, compact set of $\mathbb{R}^n$, and $M$ is a convex closed set of $\mathbb{R}^n$, so there exists a linear form on $\mathbb{R}^n$, $l(t_1, t_2, \ldots, t_n) = \alpha_1 t_1 + \ldots + \alpha_n t_n$ such that $\sup\limits_{t \in M} l(t) \leq \inf\limits_{u \in L} l(u)$. The coefficients $\alpha_i$ have to be non negative, and they can't be all equal to zero, so we can normalize and assume that $\sum\limits_{i=1}^{n} \alpha_i = 1$. Take the point $x = \sum\limits_{i=1}^{n} \alpha_i x_i$, the point $x$ is in $A$ and for any $y \in K$ we have

$$\langle y,x\rangle = \sum \alpha_i \langle y,x_i\rangle > \sum \alpha_i l(c,\ldots,c) = c.$$

For this $x$ we have $H_x = \phi$ and this can't be.

3.3. Lemma. Let $A$ be a convex bounded set in $L^0$, such that $A \subset L^1$. Then

1) $\forall \varepsilon > 0$, there exists $c \in \mathbb{R}$ such that $P(f \leq c) > 1 - \varepsilon$ $\forall f \in A$.

2) for any $\varepsilon > 0$ there exists $c \in \mathbb{R}$ and $g_0 \in L^\infty$ such that

$0 \leq g_0 \leq 1$, $E[g_0] \geq 1 - \varepsilon$ and $E[fg_0] \leq c$ $\forall f \in A$.

Proof. 1) Let $\varepsilon > 0$ and $V_\varepsilon = \{f;\ \|f\|_0 \leq \varepsilon\}$. As the set $A$ is bounded in $L^0$ there exists $\lambda > 0$ such that $\lambda A \subset V_\varepsilon$. For such a pair $(\varepsilon, \lambda)$ we have $P(|\lambda f| \geq 1) \leq \varepsilon$ $\forall f \in A$; and finally, taking $c = \frac{1}{\lambda}$ we get $P(f \leq c) < P(|f| \leq c) \geq 1 - \varepsilon$ $\forall f \in A$.

2) Let $\varepsilon > 0$ and let $c$ be chosen as above. Take $K = \{g;\ g \in L^\infty,\ 0 \leq g \leq 1,\ E[g] \geq 1 - \varepsilon\}$. The set $K$ is weakly closed in the unit ball of $L^\infty$; therefore $K$ is weakly compact. The set $K$ is obviously convex and for any $f \in A$, the set $H_f = \{g;\ g \in K,\ \langle g,f\rangle \leq c\}$ contains at least the function $g = I_{\{f \leq c\}}$. Lemma 3.2 implies that there exists $g_0 \in K$ such that $\sup_{f \in A} \langle g_0, f\rangle \leq c$.

3.4. Lemma. If $A \subset L^1(P')$ is a convex, bounded subset of $L^0$, there exists a probability measure $Q$ equivalent to $P'$ such that $\sup_{f \in A} \int f dQ < +\infty$.

Proof. For any integer $n$, there exists $c_n \in \mathbb{R}$ and $g_n \in L^\infty$, $0 \leq g_n \leq 1$, such that $\int g_n dP' \geq 1 - \frac{1}{n}$ and $\sup_{f \in A} \int f g_n dP' \leq c_n$. Choose a sequence of strictly positive real numbers $\alpha_n$ such that $\sum \alpha_n$ and $\sum \alpha_n |c_n|$ both converge. Take $h = \sum \alpha_n g_n$, and let $Q$ be the measure $Q = hP'$. The set $(h = 0)$ is the intersection $\bigcap_n (g_n = 0)$; so $P'(h = 0) \leq \frac{1}{n}$, $\forall n$ and the measures $P$ and $Q$ are equivalent. Furthermore $\sup_{f \in A} \int fh dP' = \sup_{f \in A} \int f dQ \leq \sum \alpha_n c_n < +\infty$. The only thing is that the finite measure $Q$ might not be a probability, but this can be easily taken care of.

We see now that there is no hope to go beyond semimartingales in stochastic integration. Can one actually integrate with respect to semimartingales? Yes and we have the following theorem.

3.5. Theorem. Let $X$ be a semimartingale and $b(\underline{\underline{B}}_t)$ be the set of all bounded predictable processes vanishing on $]t,+\infty[$. Then there exists one and only one extension $J^*$ of the function $J(\phi) = \int_0^\infty \phi_s dX_s$ to the set $\bigcup_t b(\underline{\underline{B}}_t)$ such that

1) $J^*$ is linear

2) if $(Y^n)$ is a bounded sequence of elements of $b(\underline{\underline{B}}_t)$, and if for all $(s,\omega)$, $Y(s,\omega) = \lim_n Y^n(s,\omega)$ exists, then $J^*(Y^n) \to J^*(Y)$ in $L^0$. (Note that the limit process $Y$ is automatically in $b(\underline{\underline{B}}_t)$).

Partial proof. The existence of $J^*$ will be proven in chapter 5. The unicity is trivial, using the fact that $\underline{\underline{B}}_t$ generates the $\sigma$-field of the predictable sets vanishing on $]t,+\infty[$.

Remark. Note that we are working in $L^0$, so the extension $J^*$ is the same if we replace the probability $P$ by an equivalent probability $Q$.

# CHAPTER IV: MORE ON LOCAL MARTINGALES AND SEMIMARTINGALES

## SQUARE INTEGRABLE MARTINGALES

Ito's stochastic integral theory is based on the remark that, if $B_t$ is the Brownian motion starting from zero, the process $B_t^2 - t$ is a martingale. We are going to see what takes place of the process $A_t = t$, when instead of the Brownian motion we have a martingale.

4.1. <u>Definition</u>. <u>A martingale</u> $M$ <u>is a square integrable martingale if</u> $\sup_t E[M_t^2] < +\infty$.

4.2. <u>Properties</u>. If $M$ is a square integrable martingale, $M$ has the following properties

1) $M_t^2$ is a submartingale

2) $M_\infty = \lim_{t\to+\infty} M_t$ exists a.e. and in $L^2$.

3) $E[\sup_t |M_t^2|] \leq 4 \sup_t E[|M_t|^2] = 4E[M_\infty^2]$ (Doob's inequality).

4) $E[\sum_s (\Delta M_s)^2] \leq \liminf_\tau E[\sum (M_{t_{i+1}} - M_{t_i})^2] = E[M_\infty^2] < +\infty$. (the lim inf being taken over the directed set of the partitions $\tau$ of $[0,\infty]$ with the partial ordering $\tau < \sigma$ if $\sigma$ is finer than $\tau$; the above inequality comes from $\sum_s (\Delta M_s)^2 \leq \liminf_\tau \sum (M_{t_{i+1}} - M_{t_i})^2$ and Fatou's lemma).

4.3. <u>The process</u> $\langle M,M\rangle$. One can find an increasing process $A$ such that $M^2 - A$ is a martingale as follows: the process $-M^2$ is a supermartingale of class (D) if $M$ is square integrable (use Doob's inequality in 4.2); there exists therefore a unique predictable, integrable, increasing process $A$ such that $M^2 - A$ is a uniformly integrable martingale. This increasing process is usually denoted by $A = \langle M,M \rangle$.

In the case where $M$ is the Brownian motion, which is locally a square integrable martingale, we get $\langle M,M \rangle_t = t$. Unfortunately the process $\langle M,M \rangle$ is not the good one for local martingales. For some local martingales $M$, there is no <u>predictable</u> increasing process $A$ such that $M^2 - A$ is a local martingale (see 2.19 for an intuitive explaination). This is why we have to look more closely at square integrable martingales.

4.4. <u>Decomposition of square integrable martingales</u>. Let $M$ be a square integrable martingale. It is càdlàg, therefore we can cover the jump times of $M$ by a sequence of stopping times, some of which are predictable (call them $S_n$), the others are totally inaccessible (call them $R_n$).

Look at the processes $C_t^n = \Delta M_{S_n} I_{\{t \geq S_n\}}$. The jump $\Delta M_{S_n}$ is in $L^2$. So $C_t^n$ has a compensator which is $E[\Delta M_{S_n} | \underline{\underline{F}}_{S_n -}] I_{\{t \geq S_n\}}$ (2.10), but $E[\Delta M_{S_n} | \underline{\underline{F}}_{S_n -}] = 0$ (2.5) and $C_t^n$ is a square integrable martingale.

For the process $D_t^n = \Delta M_{R_n} I_{\{t \geq R_n\}}$, things are slightly more complicated. $D_t^n$ has a compensator $\tilde{D}_t^n$ which is a continuous process with integrable variation, but $\tilde{D}_t^n$ is not the zero process. Nevertheless $\hat{D}_t^n = D_t^n - \tilde{D}_t^n$ is a uniformly integrable martingale which jumps only at time $R_n$ and $\Delta \hat{D}_{R_n}^n = \Delta M_{R_n}$.

Can we sum $\sum_n C_t^n$ and $\sum_n \hat{D}_t^n$ and get the martingale $M$ as the sum of a continuous martingale and compensated jumps? The sum $\sum_s |\Delta M_s|$ might not converge; but we know that $E[\sum_s |\Delta M_s|^2] < +\infty$, and we are going to work in $L^2$. First we need a few results.

4.5. <u>Lemma</u>. <u>If</u> $A$ <u>and</u> $B$ <u>are two process with integrable variation, if</u> $A - B$ <u>is a martingale and if</u> $\phi$ <u>is a predictable, bounded process</u> (<u>or a predictable non negative process</u>), <u>then</u> $E[\int_0^\infty \phi_s d(A - B)_s] = 0$.

<u>Proof</u>. The result is true for any process $\phi$ of the form $\phi = \varphi_0^* I_{\{0\}} + \sum_{i=1}^n \varphi_i I_{]t_i, t_{i+1}]}$, where $\varphi_0^*$ is a bounded, $\underline{\underline{F}}_0$-measurable r.v. and the $\varphi_i$

are bounded $\underline{\underline{F}}_{t_i}$ -measurable r.v. Therefore it is true for any predictable bounded (or non negative) process.

4.6. Lemma. Let $a(t)$ be a right continuous, increasing function from $[0,\infty[$ into $\mathbb{R}_+$, such that $a(0) = 0$. And let $c(s) = \inf(t;\ a(t) > s)$. Then $c(s)$ is a non decreasing, right continuous function; and for any borelian function $f$ bounded or non negative

$$\int_0^\infty f(s)\,da(s) = \int_0^\infty f(c(s))\,I_{\{c(s)<+\infty\}}\,ds\ .$$

Proof. Just check the equality for the functions of the form $f(s) = I_{]0,t]}$ (Remember that $da(s)$ and $ds$ have no mass at $s = 0$).

4.7. Lemma. If $A$ is a process with integrable variation and if $M$ is a bounded martingale

$$E[A_\infty M_\infty] = E[\int_0^t M_s\,dA_s].$$

Proof. Apply lemma 4.6 to $A_t$ and $C_t = \inf(s;\ A_s > t)$. The r.v. $C_t$ are stopping times and

$$E[\int_0^\infty M_s\,dA_s] = E[\int_0^\infty M_{C_s} I_{\{C_s<+\infty\}}\,ds] = \int_0^\infty E[M_{C_s} I_{\{C_s<+\infty\}}]\,ds$$

$$= \int_0^\infty E[M_\infty I_{\{C_s<+\infty\}}]\,ds = E[\int_0^\infty M_\infty I_{\{C_s<+\infty\}}\,ds\ ] = E[\int_0^\infty M_\infty\,dA_s] =$$

$$E[M_\infty A_\infty].$$

4.8. Lemma. If $L_t$ is a uniformly integrable process on $[0,+\infty]$, if $L_0 = 0$ and if $E[L_S] = 0$ for any stopping time $S$, then $L$ is a martingale.

Proof. Let $A \in \underline{\underline{F}}_t$, and $S = t$ on $A$, $= +\infty$ on $A^c$. We have

$$\int_A L_t\,dP + \int_{A^c} L_\infty\,dP = 0$$

But $\int_A L_\infty\,dP + \int_{A^c} L_\infty\,dP = 0$ (take $S = +\infty$) so $\int_A L_t\,dP = \int_A L_\infty\,dP$, $\forall A \in \underline{\underline{F}}_t$, and $L$ is a martingale.

4.9. Lemma. Let $M$ be a square integrable martingale, and $S$ be a predictable time. Then $C_t = \Delta M_S I_{\{t \geq S\}}$ is a square integrable martingale; and for any square integrable martingale $N$, the process $L_t = C_t N_t - \Delta C_S \Delta N_S I_{\{t \geq S\}}$ is a uniformly integrable martingale.

Proof. We have already seen that $C_t$ is a martingale. As $\Delta C_S \in L^2$, the martingale $C_t$ is square integrable. The process $L$ is uniformly integrable as $\sup_t |L_t| \leq \sup_t |C_t| \sup_t |N_t| + |\Delta C_S||\Delta N_S| \in L^1$ (see 4.2). Apply 4.7 to $C$ and $N^T$ ($N$ stopped at a stopping time $T$), we get

$$E[C_\infty N_\infty^T] = E[C_T N_T] = E[\int_0^\infty N_s^T dC_s] = E[N_S^T \Delta C_S] =$$

$$E[\Delta N_S^T \Delta C_S] = E[\Delta N_S \Delta C_S I_{\{S \geq T\}}]$$

(we used the fact that $E[N_{S-}^T \Delta C_S] = E[N_{S-}^T E[\Delta C_S | \underline{\underline{F}}_{S-}]] = 0$). So $E[L_T] = 0$ for any stopping time $T$. As $L_0 = 0$, $L$ is a martingale (4.8).

4.10. Lemma. Let $M$ be a square integrable martingale, and let $R$ be a totally inaccessible time. We consider $D_t = \Delta M_R I_{\{t \geq R\}}$, its compensator $\tilde{D}_t$ and $\hat{D}_t = D_t - \tilde{D}_t$, then

1) $\hat{D}_t$ is a square integrable martingale, and $E[\hat{D}_\infty^2] \leq 5E[(\Delta M_R)^2]$

2) for any square integrable martingale $N$, the process $L_t = D_t N_t - \Delta D_R \Delta N_R I_{\{t \geq R\}}$ is a uniformly integrable martingale.

Proof. 1) By considering $\Delta M_R^+$ and $\Delta M_R^-$ it is enough to study the case where $\phi$ is a non negative, $\underline{\underline{F}}_R$-measurable r.v. in $L^2$ and $D_t = \phi I_{\{t \geq R\}}$. If the function $\phi$ is bounded by the constant a we get

$$E[\tilde{D}_\infty^2] = 2E[\int_0^\infty \tilde{D}_s d\tilde{D}_s] = 2E[\int_0^\infty \tilde{D}_s dD_s] \leq 2E[\int_0^\infty \tilde{D}_\infty dD_s]$$

$$= 2E[\phi \tilde{D}_\infty] \leq 2aE[\tilde{D}_\infty] < +\infty .$$

(we used the formula $f(\infty)^2 = 2\int_0^\infty f(s)df(s)$, true for continuous non decreasing functions, and lemma 4.5.)

So if $\phi$ is non negative bounded, $\tilde{D}_\infty \in L^2$ and $E[\tilde{D}_\infty^2] \leq 2E[\phi\tilde{D}_\infty] \leq 2\|\phi\|_{L^2}\|\tilde{D}_\infty\|_{L^2}$ and $E[\tilde{D}_\infty^2] \leq 4E[\phi^2]$.

If the non negative function $\phi$ is not bounded, consider the functions $\phi^n = \phi \wedge n$, and define the $\tilde{D}_t^n$ corresponding to $D_t^n = \phi^n I_{\{t \geq R\}}$. The process $\tilde{D}_t^{n+1} - \tilde{D}_t^n$ is the compensator of the increasing process $D_t^{n+1} - D_t^n$, so we can construct the increasing, continuous processes $\tilde{D}_t^n$ in such a way that for each $n$, $\tilde{D}_t^{n+1} - \tilde{D}_t^n$ is itself an increasing process. For each $n$ we have

$$E[(\tilde{D}_\infty^n)^2] \leq 4E[(\phi^n)^2] \leq 4E[\phi^2] < +\infty .$$

The limit $B_\infty$ of the non decreasing sequence $\tilde{D}_\infty^n$ is therefore in $L^2$, as is each $B_t = \lim_n \tilde{D}_t^n$. For each $\omega$ the functions $t \to \tilde{D}_t^n$ converge uniformly to the function $t \to B_t$ (this as $0 \leq \tilde{D}_t^{n+1} - \tilde{D}_t^n \leq \tilde{D}_\infty^{n+1} - \tilde{D}_\infty^n$). So the process $B_t$ is continuous. By taking limits in $L^1$ in the equality $E[D_t^n - \tilde{D}_t^n | \underline{\underline{F}}_s] = D_s^n - \tilde{D}_s^n$ for $s \leq t$, we get that $D_t - B_t$ is a uniformly integrable martingale. Therefore $B_t = \tilde{D}_t$, and $E[\tilde{D}_\infty^2] = E[B_\infty^2] \leq 4E[\phi^2] < +\infty$. And the martingale $\hat{D} = D - \tilde{D}$ is square integrable.

2) As in (4.9) we see that $L$ is a uniformly integrable process. Let $T$ be a stopping time and apply 4.7 to $\hat{D}$ and the martingale $N^T$, we get

$$E[(D_T - \tilde{D}_T)N_T] = E[\int_0^\infty N_t^T d(D - \tilde{D})_t] = E[N_R^T \Delta D_R] - E[\int_0^\infty N_t^T d\tilde{D}_t]$$

$$= E[N_R^T \Delta D_R] - E[\int_0^\infty N_{t-}^T d\tilde{D}_t] = E[\Delta N_R \Delta D_R \, I_{\{T \geq R\}}].$$

(we used the fact that $N_t$ and $N_{t-}$ have the same Stieltjes integral with respect to the continuous process $\tilde{D}_t$). Again use 4.8 to show that $L$ is a martingale.

4.11. <u>Definition</u>. <u>A square integrable martingale</u> $M$ <u>is purely discontinuous if</u> $M_0 = 0$, <u>and if for any square integrable continuous martingale</u> $N$,

<u>the process</u> $MN$ <u>is a martingale</u>.

4.12. <u>Theorem</u>. <u>Let</u> $M$ <u>be a square integrable martingale, then</u> $M$ <u>can be decomposed in a unique way into the sum of a square integrable continuous martingale</u> $M^c$ <u>and a square integrable purely discontinous martingale</u> $M^d$. <u>The martingale</u> $M^c$ <u>is called the continuous part of</u> $M$, <u>and</u> $M^d$ <u>is called the compensated sum of the jumps of</u> $M$.

<u>Proof</u>. Let us come back to the $C^n$ and $\hat{D}^n$ defined in 4.4. According to 4.9 and 4.10 the products $C^n\hat{D}^n$, $C^n\hat{D}^m$, $C^nC^m$, $\hat{D}^n\hat{D}^m$ $(n \neq m)$ are all martingales. Consider the processes $Y_t^N = \sum_1^N C_t^n + \sum_1^N D_t^n$; we then have for $N' \leq N$ (using 4.9 and 4.10 again)

$$E[(Y_\infty^{N'} - Y_\infty^N)^2] = E[\sum_{N+1}^{N'} (C_\infty^n)^2 + \sum_{N+1}^{N'} (\hat{D}_\infty^n)^2] \leq 5E[\sum_{N+1}^{N'} (|\Delta X_{S_n}|^2 + |\Delta X_{R_n}|^2).$$

This and 4.2 shows that

$$E[\sup_t |Y_t^{N'} - Y_t^N|^2] \leq 5E[\sum_{N+1}^{N'} (|\Delta X_{S_n}|^2 + |\Delta X_{R_n}|^2)].$$

We can assume, by taking a subsequence (which we shall still denote $Y^N$) that $\sum_N E[\sup_t |Y_t^{N+1} - Y_t^N|^2] < +\infty$. This implies, by Borel Cantelli lemma, that for almost all $\omega$ the functions $Y_t^N(\omega)$ converge uniformly in $t$ to a function $Y_t$, and that for each $t$ the r.v. $Y_t^N$ converges to $Y_t$ in $L^2$. The process $Y$ is therefore càdlàg, it has the same jumps as $M$, and it is a square integrable martingale.

Let $X$ be a continuous square integrable martingale. Each $XY^N$ is a martingale (4.9 and 4.10). The $Y_t^N$ converge to $Y_t$ in $L^2$ so $X_tY_t^N$ converges to $X_tY_t$ in $L^1$, and $XY$ is a martingale. So $M^d = Y$ is a purely discontinuous martingale, and $M^c = M - M^d$ is a continuous martingale.

We still have to show the uniqueness of such a decomposition; if $M = M^c + M^d = N^c + N^d$ with $M^c$ and $N^c$ square integrable, continuous martingales, and $M^d$ and $N^d$ purely discontinuous square integrable martingales, $M^c - N^c = N^d - M^d$ is continuous and purely discontinuous. So $(M^d - N^d)^2$ is a martingale and $E[(M_t^d - N_t^d)^2] = E[(M_0^d - N_0^d)2] = 0\ \forall t \geq 0$.

4.13. Remark. For each $N$ the process

$$(Y_t^N)^2 - \sum_{n=1}^{N} (\Delta M_{S_n})^2 I_{\{t \leq S_n\}} - \sum_{n=1}^{N} (\Delta M_{R_n})^2 I_{\{t \geq R_n\}}$$

is a martingale (apply 4.9 and 4.10). When $N$ goes to $+\infty$, each term converges in $L^1$, and $(M_t^d)^2 - \sum_{s \leq t} (\Delta M_s)^2 = (M_t^d)^2 - \sum_{s \leq t} (\Delta M_s^d)^2$ is a uniformly integrable martingale.

4.14. The increasing process $[M,M]$. Let $M$ be a square integrable martingale such that $M_0 = 0$, and let $M^c$ and $M^d$ be its continuous and purely discontinuous part. We define

$$[M,M]_t = \langle M^c, M^c \rangle_t + \sum_{s \leq t} (\Delta M_s)^2$$

The process $\langle M^c, M^c \rangle$ is the one defined in 4.3.

The processes $(M_t^c)^2 - \langle M^c, M^c \rangle_t$, $(M_t^d)^2 - \sum_{s \leq t} (\Delta M_s)^2$ and $M^c M^d$ are all uniformly integrable martingales and so is $M^2 - [M,M]$. The increasing process $\langle M,M \rangle$ defined in 4.3 is the compensator of $[M,M]$.

If $M$ does not vanish at $t = 0$, the tendancy is to think of $M_0$ as the jump at time $0$ and to define

$[M,M] = \langle M^c - M_0, M^c - M_0 \rangle + M_0^2 + \sum_{s \leq t} (\Delta M_s)^2$. But we shall always assume $M_0 = 0$. If $M$ is the Brownian motion, we have $M = M^c$ and $[M,M]_t = t$ again.

4.15. If $M$ and $N$ are two square integrable martingales vanishing at $0$ we define:

$$\langle M,N \rangle_t = \frac{1}{2}(\langle M + N, M + N \rangle_t - \langle M,M \rangle_t - \langle N,N \rangle_t)$$

$$[M,N]_t = \frac{1}{2}([M + N, M + N]_t - [M,M]_t - [N,N]_t).$$

We have the following obvious properties

1) $<M,N>$ is the unique predictable process $B$ with integrable variation such that $MN - B$ is a uniformly integrable martingale.

2) Because of the uniqueness of $<M,N>$ we have for any stopping time $T$

$$<M^T,N> = <M,N^T> = <M^T,N^T> = <M,N>^T$$

3) $[M,N]_t = <M^c,N^c>_t + \sum_{s\leq t} \Delta M_s \Delta N_s$

4) $MN - [M,N]$ is a uniformly integrable martingale

5) for any stopping time $T$

$$[M^T,N] = [M,N^T] = [M^T,N^T] = [M,N]^T.$$

LOCAL MARTINGALES

4.16. Definition. A local martingale $M$ is purely discontinuous if $M_0 = 0$ and if for any continuous local martingale $N$ the product $MN$ is a local martingale. (As any continuous local martingale is locally a bounded martingale, it is enough to check the property for any bounded continuous martingale).

4.17. Lemma. Let $M$ be a local martingale. Suppose that $M$ is also a process with finite variation ($M_0 = 0$ by definition 2.12). Then

1) $V_t = \sum_{s\leq t} (\Delta M_s)$ is a process with locally integrable variation, and $M = V - \tilde{V}$ ($\tilde{V}$ is the compensator of $V$).

2) For any bounded continuous martingale $N$, the product $MN$ is a local martingale.

Proof. 1) the process $M_t^* = \sup_s |M_s|$ is locally integrable (2.21) so by 2.20 the variation of $V$ is locally integrable and $V$ has a compensator $\tilde{V}$. The local martingale $M - (V - \tilde{V})$ is continuous (localize so that $M$ is a uniformly integrable martingale and use 2.10). The variation of

$M - (V - \tilde{V})$ is locally integrable by 2.21 and 2.20. So $M - (V - \tilde{V})$ has a compensator which has to be $M - (V - \tilde{V})$ as $M - (V - \tilde{V})$ is continuous, but which has to be $0$ as $M - (V - \tilde{V})$ is a local martingale. So $M = V - \tilde{V}$.

2) Let $N$ be a bounded continuous martingale; by working locally we can assume that the variations of $V$ and $\tilde{V}$ are integrable. For any stopping time $T$, we get using successively 4.7 and 4.5.

$$E[M_T N_T] = E[M_\infty N_\infty^T] = E[\int_0^\infty N^T dM_s] = E[\int_0^\infty N^T dV_s]$$

$$- E[\int_0^\infty N_s^T d\tilde{V}_s] = E[\int_0^\infty N_s^T dV_s] - E[\int_0^\infty N_s^T dV_s] = E[\int_0^\infty (N_s^T - N_s^T) dV_s] = 0.$$

So the uniformly integrable process $J = MN$ is a martingale (lemma 4.8).

4.18. Theorem. *Let $M$ be a local martingale. $M$ can be written in a unique way as*

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

*where $M^c$ is a continuous local martingale, and $M^d$ is a purely discontinuous martingale. Moreover we have $(M^c)^T = (M^T)^c$ for any stopping time $T$ such that $M^T$ is a square integrable martingale.*

Proof. Use the fundamental lemma 2.21; $M = N + U$ where $N$ is a locally bounded local martingale, and $U$ is a martingale with finite variation. Let $(T_n)$ be a localizing sequence of stopping times such that each $N^{T_n}$ is a bounded (therefore square integrable) martingale. $N^{T_n}$ can be decomposed into $N^{T_n} = (N^{T_n})^c + (N^{T_n})^d$. Because of the uniqueness in 4.12, we have $(N^{T_n})^c = (N^{T_{n+1}})^c$ on $[\![0,T_n]\!]$.

Define $N^c$ as

$$N^c = \sum_{n=1}^\infty (N^{T_n})^c I_{]\!]T_{n-1},T_n]\!]} \qquad (T_0 = 0).$$

The process $N^c$ is a continuous local martingale, and $N^d = N - N^c$ is a purely discontinuous local martingale (this as each $(N^d)^{T_n} = (N^{T_n})^d$ is a purely discontinuous square integrable martingale). So $M = N^c + (N^d + U)$ where $N$ is a continuous local martingale and $N^d + U$ is a purely discontinuous local martingale (4.17).

If this decomposition unique? If $M = M^c + M^d = X^c + X^d$ are two such decompositions, the process $M^c - X^c = X^d - M^d$ is both a continuous and purely discontinuous martingale. Being continuous it is locally bounded, and we see, using 4.12, that locally $M^c - X^c = X^d - M^d = 0$.

The fact that $(M^c)^T = (M^T)^c$ for any stopping time $T$ such that $M^T$ is a square integrable martingale is again due to the uniqueness in 4.12.

4.19. Lemma. *Let $M$ be a continuous martingale, there exists a unique increasing predictable process $[M,M]$ such that $[M,M]^T = [M^T,M^T]$ for any stopping time $T$ such that $M^T$ is square integrable.*

Proof. The uniqueness and existence are both due to the existence and uniqueness of $\langle M,M\rangle$ in 4.3; and to the fact that the continuous local martingale $M$ is locally square integrable.

The process $[M,M]$.

4.20. Lemma. *Let $M$ be a local martingale, then for almost all $\omega$ the sums $\sum_{s\leq t}(\Delta M_s)^2$ converge for all $t$.*

Proof. By the fundamental lemma 2.21, $M$ is the sum of a locally bounded martingale $U$ and a process $V$ with finite variation. As $\sum_{s\leq t}|\Delta V_s|$ converges, so does $\sum_{s\leq t}|\Delta V_s|^2$. There exists a localizing sequence $(T_n)$ such that each $U^{T_n}$ is square integrable; we then have $E[\sum_{s\leq T_n}|\Delta U_s|^2] < +\infty$; and for almost all $\omega$ $\sum_{s\leq T_n}|\Delta M_s|^2 \leq 2\sum_{s\leq T_n}|\Delta U_s|^2 + 2\sum_{s\leq T_n}|\Delta V_s|^2$ converges.

4.21. Definition. *Let $M$ be a local martingale, we define* $[M,M]_t = [M^c,M^c]_t + \sum_{s\leq t}(\Delta M_s)^2$.

The process $[M,M]_t$ is an increasing process; we will see in chapter 6 that $M^2 - [M,M]$ is a local martingale.

SEMIMARTINGALES

4.22. Lemma. Let $X$ be a semimartingale; suppose $X = X_0 + M + A = X_0 + N + B$ where $M$ and $N$ are local martingales, and $A$ and $B$ are two processes with finite variation. Then $M^c = N^c$. We shall denote $X^c = M^c$, $X^c$ is called the continuous local martingale part of the semimartingale $X$.

Proof. $M - N$ is a local martingale with finite variation, so, by lemma 4.17, $M - N$ is purely discontinuous and $(M - N)^c = 0$.

4.23. Lemma. Let $X$ be a semimartingale; for almost all $\omega$ the sums $\sum_{s<t} (\Delta X_s)^2$ converge for all $t$.

Proof. Use the definition of a semimartingale and 4.20.

4.24. Definition. Let $X$ be a semimartingale, the increasing process $[X,X]$ is the process

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

As in 4.15 we shall define the processes $[X,Y]$.

4.25. Remark. One can show (see [12]) that $[X,X]_t$ is the limit in probability of the $\sum_{i=1}^{n} (X_{t_{i+1}} - X_{t_i})^2$ where the limit is taken on the directed set of the partitions of $[0,t]$. So $[X,X]_t$ is the quadratic variation of $X$ on $[0,t]$, and $[X,X]_t$ does not change if $P$ is replaced by an equivalent probability. We shall never need the result here, but it is an important one, as we have already decided that we should deal only with concepts invariant by a change of probability.

## CHAPTER V: STOCHASTIC INTEGRALS

In this chapter we will show the existence part of theorem 3.5: if $X$ is a semimartingale and $\phi$ a bounded predictable process vanishing after a certain time $t$, we can define $J^*(\phi) = \int_0^\infty \phi_s dX_s$, and $J^*$ verifies the continuity condition of 3.5. Actually we shall look at the stochastic integrals slightly differently and construct for any locally bounded predictable process $\phi$, the semimartingale $(\phi \circ X)_t = \int_0^t \phi_s dX_s$. In the case where $X$ is a square integrable martingale it will be similar to the construction of Ito's integral.

### STOCHASTIC INTEGRATION WITH RESPECT TO SQUARE INTEGRABLE MARTINGALES

5.1. Let $\underline{\underline{M}}$ be the set of all the square integrable martingales $M$. For each $M \in \underline{\underline{M}}$, the r.v. $M_\infty = \lim_{t \to +\infty} M_t$ is in $L^2(\Omega, \underline{\underline{F}}_\infty, P)$. And conversely, to each r.v. $Y \in L^2(\Omega, \underline{\underline{F}}_\infty, P)$ corresponds a square integrable martingale $M$ which is the càdlàg version of the martingale $E[Y|\underline{\underline{F}}_t]$. The vector space $L^2(\Omega, \underline{\underline{F}}_\infty, P)$ is a Hilbert space for the scalar product $(X,Y) \to E[XY]$, so $\underline{\underline{M}}$ is a Hilbert space for the scalar product $(M,N) \to E[M_\infty N_\infty]$. The corresponding norm on $\underline{\underline{M}}$ is $\|M\| = \sqrt{E[M_\infty^2]}$.

5.2. Let $M^n$ be a sequence of square integrable martingales converging in $\underline{\underline{M}}$ to a square integrable $M$. By Doob's inequality (4.2) we have $E[\sup_t |M_t^n - M_t|^2] \leq 4E[|M_\infty^n - M_\infty|^2]$. Therefore by Borel Cantelli lemma, there exists a subsequence $M^{n_k}$ such that for almost every $\omega$ the trajectory $t \to M_t^{n_k}(\omega)$ converge uniformly to the trajectory $t \to M_t(\omega)$. In particular, if each $M^n$ is continuous, so is $M$.

5.3. Definition. If $M$ is a square integrable martingale such that $M_0 = 0$, we define

$$\dot{L}^2(M) = \{H;\ H \text{ predictable process such that } E[\int_0^\infty H_s^2 d[M,M]_s < +\infty\}.$$

(The dot on the $L$ is to distinguish it from a space $L^2(M)$ not used here).

On $\dot{L}^2(M)$ we take the norm $\|H\|_{\dot{L}^2(M)} = \sqrt{E[\int_0^\infty H_s^2 d[M,M]_s]}$.

5.4. Let $\underline{\underline{B}}$ be the set of all the predictable processes $\phi$ of the form $\phi = \varphi_0^* I_{\{0\}} + \sum_{i=1}^{n-1} \varphi_i I_{]t_i, t_{i+1}]}$ where $0 = t_0 < t_1 < \ldots < t_n < +\infty$, $\varphi_0^*$ is a bounded $\underline{\underline{F}}_0$-measurable and the $\varphi_i$ are bounded, $\underline{\underline{F}}_{t_i}$-measurable r.v. For such a process $\phi$ we consider, as in chapter 3, the process $(\phi\circ M)_t = \int_0^\infty I_{]0,t]}(s)\phi_s dM_s = \sum_{1=}^{1} \varphi_i (M_{t_{i+1}\wedge t} - M_{t_i \wedge t})$. (As $M$ is a square integrable martingale vanishing at 0, we could use $I_{[0,t]}$ instead of $I_{]0,t]}$. But it is better to adopt the notation $\int_t^s = \int_{]t,s]}$.)

5.5. Lemma. 1) If $M \in \underline{\underline{M}}$, $M_0 = 0$, and $\phi \in \underline{\underline{B}}$ then the process $\phi\circ M$ is a square integrable martingale and $\|\phi\circ M\| = \|\phi\|_{\dot{L}^2(M)}$

2) If $M$ is a continuous square integrable martingale, so is $\phi\circ M$.

3) The processes $\Delta(\phi\circ M)_s$ and $\phi_s \Delta M_s$ are indistinguishable.

Proof. The proof of part 1 is trivial if one remembers that $M^2 - [M,M]$ is a uniformly integrable martingale. Parts 2 and 3 are obvious on the explicit form of $\phi\circ M$.

5.6. Theorem. Let $M \in \underline{\underline{M}}$, such that $M_0 = 0$, then

1) $\underline{\underline{B}}$ is dense in $\dot{L}^2(M)$

2) We can extend the application $\phi \to \phi\circ M$ defined on $\underline{\underline{B}}$ into an isometry $H \to H\circ M$ from $\dot{L}^2(M)$ into $\underline{\underline{M}}$.

3) if $M$ is continuous so is $H\circ M$

4) the processes $\Delta(H\circ M)_s$ and $H_s \Delta M_s$ are indistinguishable

Proof. The processes of $\underline{\underline{B}}$ generate the predictable σ-field so $\underline{\underline{B}}$ is dense in $\dot{L}^2(M)$; and the isometry $\phi \to \phi \circ M$ from $\underline{\underline{B}}$ into $\underline{\underline{M}}$ can be extended into an isometry from $\dot{L}^2(M)$ into $\underline{\underline{M}}$.

Convergence in $\underline{\underline{M}}$ implies almost everywhere uniform convergence on the trajectories for a subsequence (5.2), so part 3 and 4 follow from lemma 5.5.

5.7. Theorem. Let $M \in \underline{\underline{M}}$ such that $M_0 = 0$, and $H \in \dot{L}^2(M)$. The process $L = H \circ M$ is the unique element of $\underline{\underline{M}}$ such that

1) $L_0 = 0$

2) $[L,N] = H \circ [M,N] \ \forall N \in \underline{\underline{M}}$

(the symbol $(H \circ [M,N])_t$ denotes the Stieltjes integral $(H \circ [M,N]_t) = \int_{]0,t]} H_s d[M,N]_s$).

The proof will use Kunita and Watanabe's inequality which is just a form of Schwarz inequality.

5.8. Lemma. Kunita and Watanabe's inequality: let $M$ and $N$ be two elements of $\underline{\underline{M}}$ vanishing at time zero, and $H$ and $K$ be two bounded measurable processes then

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

Proof. Use the fact that $[M + \lambda N, M + \lambda N]_t - [M + \lambda N, M + \lambda N]_s$ is non negative for any $\lambda \in \mathbb{R}$, and $s \leq t$, to get

$$|[M,N]_t - [M,N]_s|^2 \leq ([M,M]_t - [M,M]_s)([N,N]_t - [N,N]_s).$$

This implies easily: if $H$ and $K$ are of the form $H = \sum_{i=1}^n H_i I_{]t_i,t_{i+1}]}$, $K = \sum_{j=1}^m K_j I_{]s_j,s_{j+1}]}$, where the $H_i$ and $K_j$ are bounded $\underline{\underline{F}}$-measurable r.v. then

$$(*) \qquad \left|\int_0^\infty H_s K_s d[M,N]_s\right| \leq \left(\int_0^\infty H_s^2 d[M,M]_s\right)^{\frac{1}{2}} \left(\int_0^\infty K_s^2 d[N,N]_s\right)^{\frac{1}{2}}.$$

And this extends to the case where $H$ and $K$ are bounded measurable processes. As the borelian $\sigma$-field on $\mathbb{R}_+$ is separable, there exists a measurable version of the process $J_s = \dfrac{|d[M,N]_s|}{d[M,N]_s}$ , and we can assume that $J_s$ is bounded by 1. Apply the inequality (*) to $H_s$ and $J_sK_s$ to get

$$\left| \int_0^\infty H_sK_s|d[M,N]_s|\right| \le \left(\int_0^\infty H_s^2 d[M,M]_s\right)^{\frac{1}{2}}\left(\int_0^\infty K_s^2 d[M,M]_s\right)^{\frac{1}{2}},$$

for any $H$ and $K$ bounded measurable processes. Then the inequality with $|H|$ and $|K|$ is trivial.

<u>Proof of theorem 5.7</u>. By lemma 5.8 and theorem 5.6 the application $H \to E[(H\circ M)_\infty N_\infty - \int_0^\infty H_s d[M,N]_s]$ is continuous on $\dot{L}^2(M)$. Its value, being zero on $\underline{\underline{B}}$, is zero on all $\dot{L}^2(M)$.

Let $M$ and $N$ be two elements of $\underline{\underline{M}}$ vanishing at $t = 0$, and let $H \in \dot{L}^2(M)$. The process $J_t = (H\circ M)_tN_t - \int_0^t H_s d[M,N]_s$ is uniformly integrable (use lemmas 5.8, 5.6 and Doob's inequality 4.2). Take a stopping time $T$ and $N^T$ the martingale $N$ stopped at time $T$; we get

$$0 = E[(H\circ M)_\infty N_\infty^T - \int_0^\infty H_s d[M,N^T]_s] = E[J_T]$$

So $J_t$ is a uniformly integrable martingale (lemma 4.8).

Let $M^c$ be the continuous part of $M$, and $M^d$ be its purely discontinuous part. For any continuous martingale $N^c$, the process $\int_0^t H_s d[M^c,N^c]_s$ is continuous and has finite variation. As $(M\circ M^c)_tN_t^c - \int_0^t H_s d[M^c,N^c]_s$ is a uniformly integrable martingale, we have (4.15)

$$[H\circ M^c,N^c]_t = \langle H\circ M^c,N^c\rangle_t = \int_0^t H_s d[M^c,N^c]_s.$$

For any continuous martingale $N^c$, the process $(H\circ M^d)N^c - \int H_s d[M^d,N^c]$ is a uniformly integrable martingale; but $[M^d,N^c] = 0$ by definition 4.14, so $H\circ M^d$ is purely discontinuous. As $H\circ M^c$ is continuous, and as $H\circ M = H\circ M^c + H\circ M^d$, we have $(H\circ M)^c = H\circ M^c$, and $(H\circ M)^d = H\circ M^d$.

We have now for any square integrable martingale $N$

$$[H\circ M,N]_t = [H\circ M^c,N^c]_t + \sum_{s\leq t} \Delta(H\circ M)_s \Delta N_s$$
$$= (H\circ[M^c,N^c])_t + \sum_{s\leq t} H_s \Delta M_s \Delta N_s = (H\circ[M,N])_t$$

Suppose two square integrable martingale $L$ and $L'$ verify $L_0 = L'_0 = 0$ and $[L,N] = H\circ[M,N]$ $\forall\, N \in \underline{\underline{M}}$. We then have $[L - L', N] = 0$ $\forall\, N \in \underline{\underline{M}}$, and $[L - L', L - L'] = 0$. The process $(L - L')^2$ is a martingale, $L_0 = L'_0 = 0$, so $L = L'$.

5.9. Remark. Note that in the proof of 5.7 we showed that $(H\circ M)^c = H\circ M^c$ and $(H\circ M)^d = H\circ M^d$.

5.10. Corollary. *Let $M$ and $N$ be two elements of $\underline{\underline{M}}$ vanishing at $0$, and $H$ be in $\dot{L}^2(M)$. Using the characterization 5.7 of the stochastic integral we get, for any bounded predictable process $K$*

$$K\circ[H\circ M,N] = (KH)\circ[M,N] = [KH\circ M,N]$$

*and*

$$K\circ(H\circ M) = (KH)\circ M.$$

*In particular if $T$ is a stopping time, and $K$ is the predictable process $K = I_{]\!]0,T]\!]}$, we have*

$$(H\circ M)_T = \int_0^T H_s dM_s = (K\circ(H\circ M))_\infty = ((KH)\circ M)_\infty = \int_0^\infty I_{]\!]0,T]\!]} H_s dM_s.$$

5.11. Remark. Let us come back to theorem 3.5. The set $b(\underline{\underline{B}}_t)$ was the set of all bounded predictable processes vanishing on $]t,+\infty[$. For a process $Y \in b(\underline{\underline{B}}_t)$ we shall define $J^*(Y) = \int_0^\infty I_{]0,t]} Y_s dM_s = (Y\circ M)_t$. It is obvious that $J^*$ is linear on $\bigcup_t b(\underline{\underline{B}}_t)$; let $(Y^n)$ be a bounded sequence of elements of $b(\underline{\underline{B}}_t)$, converging at each $(s,\omega)$ to a process $Y \in b(\underline{\underline{B}}_t)$. The process $Y$ is then the limit in $\dot{L}^2(M)$ of the $Y^n$, and $J^*(Y^n)$ converges in

$L^2$ to $J^*(Y)$. The stochastic integral we have constructed so far, is the one we were looking for.

STOCHASTIC INTEGRATION WITH RESPECT TO LOCAL MARTINGALES

Remember that a local martingale always vanishes at $t = 0$.

5.12. Definition. *A process* $H$ *is locally bounded, if there exists a localizing sequence* $(T_n)$ *of stopping times such that each process* $H^{T_n} I_{\{T_n>0\}}$ *is a bounded process.*

Since we have decided that we would like to be able to replace $P$ by an equivalent probability $Q$ without changing the set of processes $H$ for which $\int H_s dM_s$ is defined, we shall restrict ourselves to the processes $H$ which are predictable and locally bounded.

5.13. Let $M$ be a local martingale, and $H$ a predictable, locally bounded process. By the fundamental lemma (2.21), we have $M = U + V$ where $U$ is a locally square integrable martingale, and $V$ is a local martingale with finite variation. Let $(T_n)$ be a localizing sequence such that for each $n$, $U^{T_n}$ is a square integrable variation, $V$ is a martingale with integrable variation (we can do this by 4.17), and $H^{T_n} I_{\{T_n>0\}}$ is bounded. We want to define the stochastic integral $(H\circ M)_t$ by

$$(H\circ M)_t = \int_0^t H_s^{T_n} dU_s^{T_n} + \int_0^t H_s^{T_n} dV_s^{T_n}, \quad \text{on } [\![0,T_n]\!].$$

Lemma 5.14 will tell us that the stochastic integral *$H\circ M$ thus defined does not depend on the decomposition* $M = U + V$, *and that the process $H\circ M$ is a local martingale.*

5.14. Lemma. *Let* $V$ *be a process with integrable variation*

1) *if* $V$ *is a martingale, and if* $H$ *is a predictable process such that* $E[\int_0^\infty |H_s||dV_s|] < +\infty$, *then the Stieltjes integral* $\int_0^t H_s dV_s$ *is a*

càdlàg uniformly integrable martingale.

2) If $V$ is a square integrable martingale, and if $H$ is a bounded process the stochastic integral $(H\circ V)_t$ and the Stieljes integral $\int_0^t H_s dV_s$ are two indistinguishable processes.

Proof. Part 1 is true if $H \in \underline{\underline{B}}$. Define $\dot{L}^1(V)$ as the set of all predictable processes $H$ such that $E[\int_0^\infty |H_s||dV_s|] < +\infty$. And take on $\dot{L}^1(V)$ the norm $\|H\|_{\dot{L}^1(V)} = E[\int_0^\infty |H_s||dV_s|]$. If $H^n \to H$ in $\dot{L}^1(V)$, we have $E[\sup_t |\int_0^t H^n_s dV_s - \int_0^t H_s dV_s|] \leq E[\int_0^t |H^n - H|_s |dV_s|] \leq \|H^n - H\|_{\dot{L}^1(V)} \to 0$. We finish as in 5.6, except that the convergences are in $L^1$ instead of $L^2$.

For part 2, the stochastic integral $(H\circ V) = ((HI_{]0,t]})\circ V)_\infty$ and the Stieljes integral $\int_0^t H_s dV_s = \int_0^\infty I_{]0,t]} H_s dV_s$ coincide on $\bigcup_t b(\underline{\underline{B}}_t)$; both are linear in $H$ and verify the continuity property of theorem 3.5, so $(H\circ V)_t = \int_0^t H_s dV_s$ a.e. As the two processes are càdlàg they are indistinguishable.

5.15. Theorem. Let $M$ be a local martingale, and let $H$ be a predictable locally bounded process. Then $L = H\circ M$ is the unique local martingale $L$ such that

$$[L,N] = H\circ[M,N] \ \forall N \text{ local martingale.}$$

Proof. Let $M = U + V$ where $U$ is a locally square integrable martingale, and $V$ is a local martingale with finite variation. The process $H\circ V$ is a local martingale with finite variation, so $(H\circ V)^c = 0$ (lemma 4.17). As $U$ is locally square integrable we have (using 4.18 and 5.9) $(H\circ U)^c = H\circ U^c$ and $(H\circ U)^d = H\circ U^d$.

Now we have

$$[H\circ M,N]_t = [(H\circ M)^c, N^c]_t + \sum_{s\leq t} H_s \Delta M_s \Delta N_s$$

$$= [H\circ U^c, N^c]_t + \sum_{s\leq t} H_s \Delta M_s \Delta N_s$$

and

$$(H\circ[M,N])_t = (H\circ[U^c,N^c])_t + \sum_{s\leq t} H_s \Delta M_s \Delta N_s .$$

The continuous martingales $U^c, N^c$ and $H\circ U^c$ are locally square integrable, so by Theorem 5.7,

$$H\circ[U^c,N^c] = [H\circ U^c, N^c]$$

and $$H\circ[M,N] = [H\circ M, N]$$

Let $L$ and $L'$ be two local martingales such that $[L,N] = [L',N] = H\circ[M,N]$, $\forall N$ local martingale. We have $[L - L',N] = 0 \ \forall N$ local martingale. Let $N_\infty$ be any bounded $\underline{\underline{F}}_\infty$-measurable r.v., and let $N_t = E[N_\infty | \underline{\underline{F}}_t]$ (càdlàg version). By localizing we can suppose that $L - L' = U + V$ where $U$ is a square integrable martingale and $V$ is a martingale with integrable variation (2.21 and 4.17). We then have as $[U,N]$ and $[V,N]$ are processes with integrable variation:

$$\begin{aligned} E[(L_\infty - L'_\infty)N_\infty] &= E[U_\infty N_\infty] + E[V_\infty N_\infty] \\ &= E[\textstyle\int_0^\infty d[L - L',N]_s] = 0 \end{aligned}$$

for any bounded $\underline{\underline{F}}_\infty$-measurable r.v. $N_\infty$. So $L_\infty - L'_\infty = 0$, and $L = L'$.

5.16. <u>Remark</u>. From 5.15 we get: if $M$ is a local martingale and $H$ a predictable, locally bounded process:

1) $(H\circ M)^c = H\circ M^c$

2) $(H\circ M)^d = H\circ M^d$

3) the processes $\Delta(H\circ M)_s$ and $H_s \Delta M_s$ are indistinguishable (this fact is trivial by 5.6 and the definition of the Stieltjes integral).

4) $(H\circ M)_T = \int I_{]0,T]} H_s dM_s = \int_0^T H_s dM_s$ for any finite stopping time $T$.

STOCHASTIC INTEGRATION WITH RESPECT TO SEMIMARTINGALES

5.17. Let $X = X_0 + M + A$ be a semimartingale, and $H$ be a predictable, locally bounded process. We shall define

$$(H\circ X)_t = (H\circ M)_t + \int_0^t H_s dA_s.$$

The integral $H\circ M$ is the stochastic integral with respect to the local martingale $M$; $\int_0^t H_s dA_s$ is the Stieltjes integral with respect to the process with finite variation $A$. According to lemma 5.14 the process $H\circ X$ does not depend of the decomposition $X = X_0 + M + A$. And we have

1) $H\circ X$ is a semimartingale
2) $(H\circ X)^c = H\circ X^c$
3) the processes $\Delta(H\circ M)_s$ and $H_s \Delta M_s$ are indistinguishable
4) $(H\circ X)_T = \int_0^\infty I_{]0,T]} H_s dX_s = \int_0^T H_s dX_s$ for any finite stopping time $T$.

5.18. Remark. Let $H$ be a process which is adapted, left continuous, and has right limits everywhere, then $X_t$ is predictable and locally bounded. (Take the localizing sequence $T_n = \inf(t; |H_t| \geq n)$). Those are the only processes $H$ we will really use.

5.19. Remark. Let $B_t$ be the Brownian motion, and $\underline{\underline{F}}_t$ be $\sigma$-field $\sigma(B_s, s \leq t)$ completed with all the null sets in $\underline{\underline{F}}$. The family $\underline{\underline{F}}_t$ is then right continuous. In his lectures Friedman showed that one could integrate any bounded progressively measurable process $K$ with respect to $B$. In fact for such a $K$ there exists a bounded predictable process $H$ such that

$$P(\omega;\ K_t = H_t \text{ except for at most a countable number of } t) = 1.$$

In that case $E[\int_0^t (K_s - H_s)^2 ds] = 0\ \forall t$, and it is natural to take $\int_0^t K_s dB_s = \int_0^t H_s dB_s$, and this is the same as the integral defined in Friedman's lectures (the fact that the process $\mathfrak{H}$ is not unique is trivially no problem at all).

## CHAPTER VI: ITO'S FORMULA

If $F$ is a continuously differentiable function from $\mathbb{R}$ into $\mathbb{R}$, we have $F(t) = \int_0^t F'(s)ds + F(0)$. This formula is also valid if $V_t$ is a continuous process with finite variation, and $F(V_t) = \int_0^t F'(V_s)dV_s + F(V_0)$. How does the proof go? You write

$$F(V_t) - F(V_0) = \sum_{i=1}^{n-1} F(V_{t_{i+1}}) - F(V_{t_i})$$

where $0 = t_0 < t_1 \dots < t_n = t$ is s subdivision of $[0,t]$, then you use Taylor's formula and the fact that the quadratic variation of $V_t$ is zero to get the result. If $X$ is a continuous semimartingale, the quadratic variation of $X$ is $[X,X]$; and the sums $\sum_{i=1}^{n-1} |X_{t_{i+1}} - X_{t_i}|^3$ go to zero in probability when the subdivision $\tau = (t_0,\dots,t_n)$ gets finer and finer. So if $F$ is a function with continuous $2^{nd}$ order derivative we should get

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

If $X$ is a semimartingale, non necessarily continuous, just look at what the jumps of $F(X_t)$ are to guess the formula in Theorem 6.1.

6.1. Theorem Ito's formula. *Let $X_1, X_2, \dots, X_n$ be $n$ semimartingales; we denote by $X_t$ the $\mathbb{R}^n$-valued semimartingale $(X_{1,t},\dots,X_{n,t})$. Let $F$ be a real valued function on $\mathbb{R}^n$, which is twice continuously differentiable. Then*

$$F\circ X_t = F\circ X_0 + \sum_{i=1}^{n} \int_0^t D^i F\circ X_{s-} dX_{i,s} + \frac{1}{2} \sum_{\substack{i=1,\dots,n \\ j=1,\dots,n}} D^i D^j F\circ X_{s-} d[X_i^c, X_j^c]_s$$

$$+ \sum_{s\leq t} (F\circ X_s - F\circ X_{s-} - \sum_{i=1}^{n} D^i F\circ X_{s-} \Delta X_{i,s}).$$

Comments. $D^iF$ and $D^iD^jF$ are the derivatives of $F$. The term $X_i^c$ is the continuous martingale part of the semimartingale $X_i$. All the processes $D^iF\circ X_{s-}$ are predictable and locally bounded so the stochastic integrals exist. Let $(T_n)$ be the localizing sequence $T_n = \inf(t; |X_t| \geq n)$. on $[\![0,T_n[\![$, $|X_t|$ and $|X_{t-}|$ are bounded by $n$; let $K = \sup_{|x|\leq n} \sum_{i,j} |D^iD^jF(x)|$, we have using Taylor's $2^{nd}$ order formula

$$\left| \sum_{s<T_n} (F\circ X_s - F\circ X_{s-} - \sum_{i=1}^{n} D^i F\circ X_{s-} \Delta X_{i,s}) \right| \leq \frac{1}{2} K \sum_i \left( \sum_{s<T_n} |\Delta X_{i,s}|^2 \right)$$

So for almost all $\omega$, $\sum_{s<T_n} |F\circ X_s - F\circ X_{s-} - \sum_{i=1}^{n} D^i F\circ X_{s-} \Delta X_{i,s}|$ converges, and so does $\sum_{s\leq T_n} |F\circ X_s - F\circ X_{s-} - \sum D^i F\circ X_{s-} \Delta X_{i,s}|$. Therefore for almost all $\omega$, the sums $\sum_{s\leq t} |F\circ X_s - F\circ X_{s-} - \sum D^i F\circ X_{s-} \Delta X_{i,s}|$ converge for all $t$.

Note also that there is no condition on $F$, except the continuity of the $2^{nd}$ derivatives.

Proof. The proof will consist of simplifying the problem step by step. First note that the jump at time $s$ of the term on the right in Ito's formula is (using 5.17)

$$\sum_{i=1}^{n} D^i F\circ X_{s-} \Delta X_{i,s} + F\circ X_s - F\circ X_{s-} - \sum_{i=1}^{n} D^i F\circ X_{s-} \Delta X_{i,s} = F\circ X_s - F\circ X_{s-}.$$

This is also the jump at time $s$ of the term on the left. We will write down the proof for the case $n = 1$.

1) it is enough to proof Ito's formula, when

a) $F$, $F'$ and $F''$ are bounded

b) $X = X_0 + M + A$, where $X_0$ is a bounded random variable, $M$ is a bounded martingale and $A$ is a process with bounded variation.

By lemma 2.21, we can write

$X = X_0 + N + B$, where $N$ is a martingale with bounded jump size and $B$ a process with finite variation. Let $R_n = +\infty$ if $|X_0| \le n$, $R_n = 0$ if $|X_0| > n$, and let $T_n = \inf\{t;\ |N_t| \ge n$ or $\int_0^t |dB_s| \ge n\} \wedge R_n$. The sequence $(T_n)$ is a localizing sequence.

Take $Y_t^n = X_0 I_{\{T_n>0\}} + N_t^{T_n} + B_t^{T_n} - \Delta B_{T_n} I_{\{t \ge T_n\}}$. The r.v. $Y_0 = X_0 I_{\{T_n>0\}}$ is a bounded, $\underline{\underline{F}}_0$-measurable r.v., $M = N^{T_n}$ is bounded martingale, and $A_t = B_t^{T_n} I_{\{t \ge T_n\}}$ has bounded variation. The two processes $Y_t^n$ and $X_t$ coincide on $[\![0,T_n[\![$; and $[Y^{nc},Y^{nc}] = [X^c,X^c]$ on $[\![0,T_n]\!]$; if Ito's formula is valid for $Y_t^n$, it will be valid for $X_t$, $\forall\, t < T_n$. As the two terms in Ito's formula have the same jump at $T_n$, Ito's formula is valid for $X_t$, $\forall\, t \le T_n$.

Now $Y_t^n$ and $Y_{t-}^n$ are two bounded processes, taking their values in a compact set $D$ of $\mathbb{R}$. Let $G$ be a twice continuously differentiable function, with compact support such that $G = F$ on $D$. All we have to do is prove Ito's formula for $G$ and $Y^n$. So we can assume that $F$; $F'$ and $F''$ are bounded.

2) Furthermore we can assume that $M$ and $A$ have at most $N$ jumps.

The martingale $M$ is bounded so there exists (proof of 4.12) a sequence of martingales $M^n$ such that $M^n = M^c$ + compensated sum of a finite number of jumps of $M$, and such that $\sum_n \| M_\infty^n - M_\infty \|_{L^2} < +\infty$. This, by Doob's inequality 4.2 implies that for almost all $\omega$ the functions $t \to M_t^n$ converge uniformly to the functions $t \to M_t$; and so $M_{t-}^n \to M_{t-}^n$ a.e.

For the process $A$, things are simpler; we can have by 1.18

$$A = A^c + \sum_n \Delta A_{T_n} I_{\{t \ge T_n\}}$$

The process $A^c$ is the continuous "process with finite variation" part of $A$ and not the continuous local martingale part of $A$ which is $0$. As the variation of $A$ is bounded, there exists a subsequence $n_k$, such that $A^{n_k} = A^c + \sum_{n=1}^{n_k} \Delta A_{T_n} I_{\{t \geq T_n\}}$ verifies $\sum_k E[\int_0^\infty |d(A^{n_k} - A)_s|] < +\infty$. Let us call this sequence $A^n$. Again, for almost all $\omega$ the trajectories $t \to A_t^n(\omega)$ converge uniformly to the trajectory $t \to A_t(\omega)$. And $A_{t-}^n \to A_{t-}$ a.e.

Suppose that Ito's formula is valid for each $X^n = X_0 + M^n + A^n$, so

$$F \circ X_t^n = F \circ X_0^n + \int_0^t F' \circ X_{s-}^n dX_s^n + \frac{1}{2} \int_0^t F'' \circ X_{s-}^n d[X^{nc}, X^{nc}]_s +$$

$$\sum_{s \leq t} (F \circ X_s^n - F \circ X_{s-}^n - F' \circ X_{s-}^n \Delta X_s^n).$$

The first term $F \circ X_t^n$ converges a.e. to $F \circ X_t$.

The second term $F \circ X_0^n$ is the term $F \circ X_0$.

For the third term $\int_0^t F' \circ X_{s-}^n dX_s^n = \int_0^t F' \circ X_{s-}^n dM_s^n + \int_0^t F' \circ X_{s-}^n dA_s^n$, we have, as $M$ and $M^n$ are square integrable

$$E[|\int_0^t F' \circ X_{s-}^n dM_s^n - \int_0^t F' \circ X_{s-} dM_s|^2] \leq 2E[|\int_0^t (F' \circ X_{s-}^n - F' \circ X_{s-}) dM_s^n|^2]$$

$$+ 2E[|\int_0^t F' \circ X_{s-} d(M^n - M)_s|^2] = 2E[\int_0^t (F' \circ X_{s-}^n - F' \circ X_{s-})^2 d[M^n, M^n]_s]$$

$$+ 2\, E[\int_0^t (F' \circ X_{s-})^2 d[M^n - M, M^n - M]_s].$$

Now the process $[M,M] - [M^n,M^n]$ is an increasing process (by the definition of the $M^n$), and $(F' \circ X_{s-}^n - F' \circ X_{s-})^2$ converges to zero and remains bounded (remember that $F$, $F'$, $F''$ are bounded from now on). So $E[\int_0^t (F' \circ X_{s-}^n - F' \circ X_{s-})^2 d[M^n, M^n]_s] \to 0$.

For the other term, we have

$$E[\int_0^t (F' \circ X_{s-})^2 d[M^n - M, M^n - M]_s] \leq \sup|F'| E[[M^n - M, M^n - M]_\infty]$$

$$\leq \sup|F'| E[(M_\infty^n - M_\infty)^2] \to 0.$$

And $\int_0^t F' \circ X_{s-}^n dM_s^n$ converges in $L^2$ to $\int_0^t F' \circ X_{s-} dM_s$.

Quite similarly (in fact it is easier as we work with Stieltjes integrals) we have that $\int_0^t F'\circ X^n_{s-}dA^n_s$ converges in $L^1$ to $\int_0^t F'\circ X_{s-}dA_s$.

As $[X^{nc},X^{nc}] = [X^c,X^c]$ and as $F''\circ X^n_{s-}$ converges to $F''\circ X_{s-}$ and remains bounded, the term $\int_0^t F''\circ X^n_{s-}d[X^{nc},X^{nc}]_s$ converges in $L^1$ to $\int_0^t F''\circ X_{s-}d[X^c,X^c]$ (remember that $X^c = M^c$ is square integrable so that $E[[X^c,X^c]_\infty] < +\infty$).

That leaves $\sum_{s\leq t}(F\circ X^n_s - F\circ X^n_{s-} - F'\circ X^n_{s-}\Delta X^n_s)$. Using Taylor's formula, and the fact that $F''$ is bounded we have

$$|F\circ X^n_s - F\circ X^n_{s-} - F'\circ X^n_{s-}\Delta X^n_s| \leq C|\Delta X^n_s|^2 \leq 2C(|\Delta M^n_s|^2 + |\Delta A^n_s|^2)$$
$$\leq 2C(|\Delta M_s|^2 + |\Delta A_s|^2).$$

As $\sum_{s\leq t}|\Delta M_s|^2$ and $\sum_{s\leq t}|\Delta A_s|^2$ converge for almost all $\omega$, we see that the sums $\sum_{s\leq t}(F\circ X^n_s - F\circ X^n_{s-} - F'\circ X^n_{s-}\Delta X^n_s)$ converge uniformly in $n$; and

$$\lim_{n\to+\infty}\sum_{s\leq t}(F\circ X^n_s - F\circ X^n_{s-} - F'\circ X^n_{s-}\Delta X^n_s) = \sum_{s\leq t} F\circ X_s - F\circ X_{s-} - F'\circ X_{s-}\Delta X_s \quad \text{a.e.}$$

Each term of Ito's formula for $X^n$ converges in probability to the corresponding term for $X$ so Ito's formula is valid for $X$.

3) <u>It is enough to prove Ito's formula when</u> $X_0$ <u>is bounded,</u> $F$, $F'$ <u>and</u> $F''$ <u>are bounded and</u> $M$ <u>and</u> $A$ <u>are continuous.</u>

We are already down to the case where $M$ and $A$ have at most $N$ jumps and $M$ is square integrable. Take $B_t = \sum_{s<t}\Delta M_s$, this is a process with integrable variation (remember that in fact $B_t = \sum_{n=1}^N \Delta M_{T_n} I_{\{t\geq T_n\}}$ and that each $\Delta M_{T_n}$ is in $L^2$). Let $\tilde{B}$ be its compensator, we get $M = M^c + B - \tilde{B}$; let $C = A + B - \tilde{B}$, $C$ is a process with finite variation which has almost $2N$ jumps and $X = X_0 + M^c + C$. Let $R_1, R_2, \ldots, R_{2N}$ be the possible jump times of $C$; if Ito's formula is true for continuous $M$ and $A$ we have

$$F\circ X_{R_1-} = F\circ X_0 + \int_{]0,R_1[} F'\circ X_{s-}dX_s + \frac{1}{2}\int_{]0,R_1[} F''\circ X_{s-}d[X^c,X^c]_s$$

and

$$F\circ X_{R_k^-} = F\circ X_{R_{k-1}} + \int_{]\!]R_{k-1},R_k[\![} F'\circ X_{s-}dX_s + \frac{1}{2}\int_{]\!]R_{k-1},R_k[\![} F''\circ X_{s-}d[X^c,X^c]_s.$$

Add the jumps at times $R_k$ on both sides, and you get Ito's formula for $X$.

4) Proof in the continuous case. So we can assume that $X_0, F, F'$ and $F''$ are bounded, that $M, A$ (and $[M,M]$) are continuous. And by localizing we can also assume that $M$, $A$ and $[M,M]$ are bounded.

Taylor's formula gives

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

and there exists a non decreasing function $\varepsilon(t)$ such that $\lim_{t\to 0} \varepsilon(t) = 0$ and $|r(x,y)| \le \varepsilon(|y - x|)|y - x|^2$

Choose a constant $a > 0$, and consider the stopping times

$$T_0 = 0$$

$$\dots$$

$$T_{i+1} = t \wedge (T_i + a) \wedge \inf\{s > T_i;\ |M_s - M_{T_i}| > a,\ [M,M]_s - [M,M]_{T_i} > a,$$

$$\text{or} \quad |A_s - A_{T_i}| > a\}$$

$$\dots$$

For each $\omega$, we have $T_i(\omega) = t$ except for a finite number of $i$.

$$\begin{aligned}
F(X_t) - F(X_0) &= \sum_i F(X_{T_{i+1}}) - F(X_{T_i}) \\
&= \sum_i F'\circ X_{T_i}(X_{T_{i+1}} - X_{T_i}) \\
&+ \frac{1}{2}\sum_i F''\circ X_{T_i}(X_{T_{i+1}} - X_{T_i})^2 \\
&+ \sum_i r(X_{T_i}, X_{T_{i+1}}).
\end{aligned}$$

We have for the first sum

$$E[(\int_0^t F'\circ X_s dM_s - \sum_i F'\circ X_{T_i}(M_{T_{i+1}} - M_{T_i}))^2]$$

$$= E[\sum_i \int_{T_i}^{T_{i+1}} (F'(X_s) - F'(X_{T_i}))^2 d[M,M]_s] \leq E[\sup_i |F'(X_s) - F'(X_{T_i})|^2 [M,M]_t]$$

and the last quantity goes to zero when $a \to 0$, since $\sup_i |F'(X_s) - F'(X_{T_i})|$ remains bounded and goes to zero.

Similarly we get that $\sum_i F'\circ X_{T_i}(A_{T_{i+1}} - A_{T_i})$ converges in $L^1$ to $\int_0^t F'\circ X_s dA_s$.

The term $\sum_i F''\circ X_{T_i}(X_{T_{i+1}} - X_{T_i})^2$ splits into three terms. We have (where $C$ is a upper bound for $|F|$ $|F'|$ and $|F''|$)

$$|\sum_i F''\circ X_{T_i}(A_{T_{i+1}} - A_{T_i})^2| \leq C \sup_i |A_{T_{i+1}} - A_{T_i}| |\int_0^t |dA_s| \leq Ca\int_0^t |dA_s|$$

so this term goes to zero with $a$.

The double product is as easy to deal with, as

$$|\sum_i F''\circ X_{T_i}(A_{T_{i+1}} - A_{T_i})(M_{T_{i+1}} - M_{T_i})| \leq C \sup_i |M_{T_{i+1}} - M_{T_i}| |\int_0^t |dA_s| \leq Ca\int_0^t |dA_s|.$$

Now for the term with $M^2$, we have using the fact that $M^2 - [M,M]$ is a martingale

$$E[|\sum_i F''(X_{T_i})(M_{T_{i+1}} - M_{T_i})^2 - \sum_i F''(X_{T_i})([M,M]_{T_{i+1}} - [M,M]_{T_i})|^2]$$

$$= \sum_i E[F''(X_{T_i})^2((M_{T_{i+1}} - M_{T_i})^2 - ([M,M]_{T_{i+1}} - [M,M]_{T_i}))^2]$$

$$\leq 2C^2 \sum_i E[(M_{T_{i+1}} - M_{T_i})^4] + 2C^2 E[([M,M]_{T_{i+1}} - [M,M]_{T_i})^2]$$

$$\leq 2C^2 E[\sup_i |M_{T_{i+1}} - M_{T_i}|^2 M_t^2] + 2C^2 E[\sup_i |[M,M]_{T_{i+1}} - [M,M]_{T_i}| [M,M]_t]$$

$$\leq 2C^2 a^2 E[M_t^2] + 2C^2 a E[[M,M]_t] \to 0 \quad \text{when} \quad a \to 0.$$

That leaves the term $\sum r(X_{T_i}, X_{T_{i+1}})$:

$$E[\sum_i r(X_{T_i}, X_{T_{i+1}})] \leq E[\sum_i (X_{T_{i+1}} - X_{T_i})^2 \varepsilon(|X_{T_{i+1}} - X_{T_i}|)$$

$$\leq E[2\varepsilon(2a)\sum_i (A_{T_{i+1}} - A_{T_i})^2 + 2\varepsilon(2a)\sum_i (M_{T_{i+1}} - M_{T_i})^2]$$

$$\leq 2\varepsilon(2a)E[a\int_0^t |dA_s| + M_t^2] \to 0 \quad \text{when} \quad a \to 0.$$

And we have proven Ito's formula.

6.2. Corollary. If $M$ is a local martingale, the process $M^2 - [M,M]$ is the local martingale $2\int M_{s-}dM_s$.

Proof. Ito's formula with $F(x) = x^2$

6.3. Corollary. If $M$ is a semimartingale, and $M_0 = 0$ then the process $Z_t = \exp(M_t - \frac{1}{2}[M^c,M^c]_t) \prod_{s\leq t} (1 + \Delta M_s)e^{-\Delta M_s}$ is the only semimartingale verifying

$$Z_t = 1 + \int_0^t Z_{s-}dM_s$$

Proof: Take $N_t = M_t - \frac{1}{2}[M^c,M^c]_t$, $K_t = \prod_{s<t} (1 + \Delta M_s)e^{-\Delta M_s}$ and apply Ito's formula to the function $F(N_t,K_t) = K_t\exp(\overline{N}_t)$. Theorem 7.1 will show that it is the unique solution.

## CHAPTER VII: STOCHASTIC DIFFERENTIAL EQUATIONS

Now that we have developed a nice theory of stochastic integration, the next step is to look at stochastic differential equations. We are going to state and prove the theorem with only one semimartingale, but one can similarly study systems of stochastic differential equations.

7.1. <u>Theorem</u>. *Let $M$ be a semimartingale such that $M_0 = 0$, and let $H$ be a càdlàg adapted process. We consider a function $f(\omega,s,x)$ defined on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ such that*

($L_1$) *For $\omega,s$ fixed, $f(\omega,s,\cdot)$ is a Lipschitz function with Lipschitz constant $K$*

($L_2$) *For $s,x$ fixed, $f(\cdot,s,x)$ is $\underline{\underline{F}}_s$-measurable.*

($L_3$) *For $x,\omega$ fixed, $f(\omega,\cdot,x)$ is a left continuous function with right limits.*

*Then the stochastic integral equation*

$$X_t(\omega) = H_t(\omega) + \int_0^t f(\omega,s,X_{s-}(\omega))dM_s(\omega)$$

*has one and only one solution $(X_t)$ which is a càdlàg adapted process.*

Before proving the theorem let us show that $\int_0^t f(\omega,s,X_{s-}(\omega))dM_s(\omega)$ exists.

7.1'. <u>Lemma</u>. *If $X$ is an adapted, càdlàg process, the process $(s,\omega) \to f(\omega,s,X_{s-}(\omega))$ is adapted, left continuous and has right limits (so it is a predictable locally bounded process).*

Proof. For $t$ fixed, the function $(\omega,x) \to f(\omega,t,x)$ is $\underline{\underline{F}}_t \times \underline{\underline{B}}(\mathbb{R})$-measurable (by $(L_2)$, and the continuity in $x$, $(L_1)$). So $f(\omega,t,X_{t-}(\omega))$ is $\underline{\underline{F}}_t$-measurable.

The left continuity and the existence of right limits is easy to show.

Proof of Theorem 7.1. Let us try the classical proof for non stochastic differential equations on the stochastic integral equation

$$X_t = \int_0^t f(s,\omega,X_{s-})\,dM_s$$

in the easy following case

$(A_1)$ $M$ is of the form $M = N + A$, where $N$ is a local martingale, $A$ is a process with finite variation such that $[N,N]$ and $B = \int |dA_s|$ are both bounded by a constant $b$.

$(A_2)$ $|f(\omega,s,0)| \leq c \quad \forall(\omega,s)$

Let $\underline{H}$ = {càdlàg processes $X$ such that $X^* = \sup_t |X_t| \in L^2$ and $X_0 = 0$}

On $\underline{\underline{H}}$ we take the norm $\|X\| = \|X^*\|_{L^2}$.

We consider for each $X \in \underline{\underline{H}}$ the process

$$(WX)_t = \int_0^t f(\cdot,s,X_{s-})\,dM_s$$

7.2. Lemma. The process $WX$ is in $\underline{H}$, and if $X$ and $Y$ are in $\underline{H}$, $\|WX - WY\| \leq h \; \|X-Y\|$, where $h = K(2\sqrt{b} + b)$

Proof. $(W0)_t = \int_0^t f(\cdot,s,0)\,dN_s + \int_0^t f(\cdot,s,0)\,dA_s$. Let $L_t = \int_0^t f(\cdot,s,0)\,dN_s$, $L$ is a local martingale and

$$[L,L]_\infty = \int_0^\infty f^2(\cdot,s,0)\,d[N,N]_s \leq c^2 b \tag{5.15}$$

So by Doob's inequality 4.2

$$E[L^{*2}] \leq 4c^2 b$$

Let $V_t = \int_0^t f(\cdot,s,0)dA_s$, we have

$$V^* \leq \int_0^\infty |dV_s| \leq \int_0^\infty |f(\cdot,s,0)||dA_s| \leq cb$$

and $E[V^{*2}] \leq c^2b^2$. So the process $W0$ is in $\underline{\underline{H}}$.

Let $X$ and $Y$ be in $\underline{\underline{H}}$, and $Z = X - Y$, we have

$$(WX)_t - (WY)_t = \int_0^t [f(\cdot,s,X_{s-}) - f(\cdot,s,Y_{s-})]dN_s + \int_0^t [f(\cdot,s,X_{s-}) - f(\cdot,s,X_s)]dA_s$$

Take again $L' = \int_0^t [f(\cdot,s,X_{s-}) - f(\cdot,s,Y_{s-})]dN_s$ and $V' = \int_0^t [f(\cdot,s,X_{s-}) - f(\cdot,s,Y_{s-})]dA_s$, we get

$$[L',L']_\infty = \int_0^\infty [f(\cdot,s,X_{s-}) - f(\cdot,s,Y_{s-})]^2 d[N,N]_s \leq K^2bZ^{*2}$$

$$E[L'^{*2}] \leq 4K^2bE[Z^{*2}]$$

and

$$V^* \leq \int_0^\infty K|Z_s||dA_s| \leq KbZ^*$$

So $\| WX - WY\| \leq K\| X - Y\| (b + 2\sqrt{b})$.

As $W0 \in \underline{\underline{H}}$, this implies that for any $X \in \underline{\underline{H}}$, $WX \in \underline{\underline{H}}$

7.3. Lemma. *If $M$ and $f$ satisfy conditions $(A_1)$ and $(A_2)$, and if $h = K(b + 2\sqrt{b}) < 1$, there exists one and only one adapted càdlàg process $X_t$ which is solution of*

$$X_t = \int_0^t f(\cdot,s,X_{s-})dM_s.$$

*Proof.* There is one and only one solution $X$ in $\underline{\underline{H}}$, as $h < 1$. IF $Z$ is a càdlàg adapted process, and if $Z = \int_0^t f(\cdot,s,Z_{s-})dM_s$, consider the times $T_n = \inf(t; |Z_t| \geq n)$. The jump of $Z$ at time $T_n$ is

$$\Delta Z_{T_n} = f(\cdot,T_n,Z_{T_{n-}})|\Delta M_{T_n}| \leq (c + nK)(2b + \sqrt{2b})$$

(use $(L_1)$, $(L_2)$ and the fact that $|\Delta M_s| \leq |\Delta N_s| + |\Delta A_s| = \sqrt{\Delta[N,N]_s} + |\Delta A_s|$

$\leq (\sqrt{2b} + 2b)$. The process $Z$ is therefore locally bounded, and locally in $\underline{\underline{H}}$, and by the uniqueness of the solution in $\underline{\underline{H}}$ we have $Z = X$.

7.4. <u>Lemma</u>. <u>If</u> $M$ <u>satisfies</u> $(A_1)$ <u>and if</u> $h = K(b + 2\sqrt{b}) < 1$, <u>then there exists one and only one adapted càdlàg process</u> $X_t$ <u>which is solution of</u> $X_t = \int_0^t f(\cdot,s,X_{s-})dM_s$.

<u>Proof</u>. Let $T_n(\omega) = \inf(t;\ |f(\omega,t,0)| \geq n)$, and let $f_n(\omega,t,x) = f(\omega,t,x)I_{\{0<t\leq T_n\}}$. The functions $f_n$ satisfy $(A_2)$ with $c = n$. And each stochastic integral equation $Z_t^n = \int_0^t f_n(\cdot,s,Z_{s-}^n)dM_s$ has a unique solution $Y^n$ As $f_{n+1}(\omega,t,x) = f_n(\omega,t,x)$ on $]\!]0,T_n]\!]$, we have $Y^n = Y^{n+1}$ on $[\![0,T_n]\!]$.

An adapted càdlàg process $X$ is solution of $X_t = \int_0^t f(\cdot,s,X_{s-})dM_s$ if and only if for each $T_n$ we have $X_t^{T_n} = \int_0^t f(\cdot,s,X_{s-}^{T_n})dM_s^{T_n}$. It is now easy to see that the process $X$ defined as

$$\begin{cases} X = Y^n & \text{on } ]\!]0,T_n]\!] \\ X_0 = 0 \end{cases}$$

is the unique solution of $X_t = \int_0^t f(\cdot,s,X_{s-})dM_s$.

7.5. <u>Lemma</u>. <u>If</u> $M$ <u>satisfies</u> $A_1$, <u>if</u> $h = K(b + 2\sqrt{b}) < 1$, <u>and if</u> $H_t$ <u>is a càdlàg adapted process, then the stochastic integral equation</u>

$$(*)\qquad X_t = H_t + \int_0^t f(\cdot,s,X_{s-})dM_s$$

<u>has a unique càdlàg adapted solution.</u>

<u>Proof</u>. $X$ is solution of $(*)$ if and only if $Y = X - H$ is solution of $Y_t = \int_0^t f(\cdot,t,Y_{s-} + H_{s-})dM_s$.

Take the function $g(\omega,t,x) = f(\omega,t,x + H_{s-}(\omega))$, it satisfies properties $(L_1)$, $(L_2)$ and $(L_3)$ with the same Lipschitz constant $K$. So just apply lemma 7.4.

7.6. <u>Lemma</u>. <u>If</u> $H$ <u>is a càdlàg adapted process if</u> $M$ <u>is a semimartingale, if the jumps of</u> $M$ <u>verify</u> $|\Delta M_t| \leq \frac{b}{4}$, if $b < 1$ and if $h = K(b + 2\sqrt{b}) < 1$,

<u>then the stochastic integral equation</u>

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

<u>has one and only one càdlàg, adapted solution</u>.

<u>Proof</u>. By 2.24 we can write $M$ as $M = N + A$, where $N$ is a local martingale, and $A$ is a predictable process with locally bounded variation. The process $A$ jumps only at predictable times; for a predictable time $T$ we have

$$|\Delta A_T| = |E[\Delta M_T|\underline{\underline{F}}_{T-}] - E[\Delta N_T|\underline{\underline{F}}_{T-}]| = |E[\Delta M_T|\underline{\underline{F}}_{T-}| \le \frac{b}{4}$$

(one here should really localize to be sure that $N$ is a uniformly integrable martingale and that the variation of $A$ is integrable). The jumps $\Delta[N,N]_t = (\Delta N_t)^2$ and $|\Delta A_t|$ and therefore bounded by $\frac{b}{4}$

Let $D_t = [N,N]_t + \int_0^t |dA_s|$ and define

$T_0 = 0$

$\dots$

$T_n = \inf(t;\ t > T_{n-1},\ D_t - D_{T_{n-1}} \ge \frac{b}{2})$

$\dots$

We now have $\int_{]\!]T_n,T_{n+1}]\!]} d[N,N]_s \le b$ and $\int_{]\!]T_n,T_{n+1}]\!]} |dA_s| \le b$ for any $n$. So by lemma 7.5, we have one and only one solution on $]\!]0,T_1]\!]$ for the stochastic integral equation

$X_t = H_t + \int_0^t f(\cdot,s,X_{s-})dM_s$. Call it $X^1$.

Now look on $]\!]T_1,T_2]\!]$ at the equation

$$X_t = H_t + X^1_{T_1} + \int_{]\!]T_1,t]\!]} f(\cdot,s,X_{s-})dM_s,$$

use lemma 7.5 again to get a unique solution $X^2$, define similarly the $X^n$

on $]\!]T_{n-1},T_n]\!]$. Patch them together, it is then easy to see that the process $X$ thus defined is the unique càdlàg adapted solution of

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

<u>Proof of theorem 7.1</u>. Let $M$ be a semimartingale, let $b$ as in lemma 7.6 and let

$$T_1 \leq T_2 \leq \ldots \leq T_n \ldots$$

be the successive times at which $M$ has jumps of size $|\Delta M_t| > \frac{b}{4}$. We consider the semimartingale

$$M_t^1 = M_t I_{\{t<T_1\}} + M_{T_1-} I_{\{t\geq T_1\}}$$

and

$$H_t^1 = H_t I_{\{t<T_1\}} + H_{T_1-} I_{\{t\geq T_1\}},$$

A càdlàg adapted process $X_t$ is solution on $[\![0,T_1]\!]$ of $X_t = H_t + \int_0^t f(\cdot,s,X_{s-})dM_s$ if and only if the process $X_t^1 = X_t I_{\{t<T_1\}} + X_{T_1-} I_{\{t\geq T_1\}}$ is solution on $[\![0,T_1]\!]$

$$X_t^1 = H_t^1 + \int_0^t f(\cdot,s,X_{s-}^1)dM_s^1.$$

But on $[\![0,T_1]\!]$ the semimartingale $M^1$ satisfies the conditions of lemma 7.6, so this last stochastic integral equation has one and only one càdlàg adapted solution $X_t^1$ on $[\![0,T_1]\!]$; do the same on each $]\!]T_n,T_{n+1}]\!]$ and you get theorem 7.1.

Recently Protter and Emery have studied the stability of solutions of stochastic differential equations. The interested reader will find the results in [7], [8], [9] and [13].

Bibliography


[ 1 ] C. Dellacherie. Capacités et processus stochastiques. Springer 1972.

[ 2 ] C. Dellacherie et P. A. Meyer. Probabilités et potentiel. 1$^{er}$ volume. Hermann 1975.

[ 3 ] C. Dellacherie et P. A. Meyer. Probabilités et potentiel. 2$^{ème}$ volume (to appear).

[ 4 ] C. Doléans. Existence du processus croissant naturel associé à un potentiel de la classe (D). Z. für Wahrscheinlichkeitstheorie 9, 1967.

[ 5 ] C. Doléans-Dade et P. A. Meyer. Equations différentielles stochastiques. Séminaire de Probabilités XI, Lecture Notes 581. Springer, 1977.

[ 6 ] J. L. Doob. Stochastic processes. J. Wiley and Sons, 1953.

[ 7 ] M. Emery. Stabilité des solutions des équations différentielles stochastiques; application aux intégrales multiplicatives stochastiques. Z. für Wahrscheinlichkeitstheorie 41, 1978.

[ 8 ] M. Emery. Une topologie sur l'espace des semimartingales (to appear).

[ 9 ] M. Emery. Equations différentielles lipschitziennes: la stabilité. (to appear).

[10] H. Kunita and S. Watanabe. On square integrable martingales. Nagoya Math. J. 30, 1967.

[11] P. A. Meyer. Un cours sur les intégrales stochastiques. Séminaire de Probabilité X. Lecture Notes 511. Springer 1976.

[12] P. A. Meyer. Probabilités et Potentiel. Hermann, 1966.

[13] P. E. Protter. $H^p$ stability of solutions of stochastic differential equations (to appear).

[14] M. Rao. On decomposition theorems of Meyer. Math. Scand. 24 (1969).

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# STOCHASTIC DIFFERENTIAL EQUATIONS AND APPLICATIONS

AVNER FRIEDMAN

# Stochastic Differential Equations and Applications

by

Avner Friedman
Northwestern University
Evanston, Illinois

## §1. Brownian motion

A stochastic process $x(t)$, $t \in I$ is a family of random variables $x(t)$ defined in a measure space $(\Omega, \mathcal{F})$ or in a probability space $(\Omega, \mathcal{F}, P)$; here $x(t)$ is either real valued or n-vector valued and $I$ is an interval, usually $[0,\infty)$. Notice that $x(t)$ is a function $x(t,\omega)$, $\omega \in \Omega$.

The function $t \to x(t,\omega)$ is called a sample path $\omega$. If a.e. sample path is continuous (right continuous), we say that the process $x(t)$ is continuous (right continuous).

A process $y(t)$ is said to be a version of $x(t)$ if $P(x(t) \neq y(t)) = 0 \ \forall t$.

Theorem 1. (Kolmogorov). If

$$E|x(t)-x(s)|^{\beta} \leq C|t-s|^{1+\alpha} \qquad (t,s \in I)$$

for some positive constants $C, \beta, \alpha$ then there is a continuous version of $x(t)$.

A process $x(t)$, $t \in I$ is called separable if there exists a sequence $\{t_j\}$ dense in $I$ and a subset $N \subset \Omega$ of probability 0 such that, if $\omega \notin N$,

$$\{x(t,\omega) \in F \ \ \forall\ t \in J\} = \{x(t_j,\omega) \in F \ \ \forall\ t_j \in J\}$$

for any open set $J \subset I$ and any closed set $F \subset R^r$.

It is known (Doob) that any stochastic process $x(t)$ has a separable version.

It is not difficult to show that if $x(t)$ is separable then

$$\sup_{t\in J} x(t), \qquad \liminf_{t\to t_0} x(t), \qquad \text{etc.}$$

are measurable.

Definition. $x(t)$ is martingale (submartingale) if $E|x(t)| < \infty \ \forall\, t$, and

$$E[x(t)|x(s),\ s \le \tau] = x(\tau) \qquad (\ge x(\tau)) \qquad \forall\, t > \tau.$$

The martingale inequality: If $x(t)$ is a separable submartingale then

$$P[\sup_{s\le t} x(s) \ge \lambda] \le \frac{1}{\lambda} Ex^+(t) \qquad (\lambda > 0).$$

If $x(t)$ is a separable martingale and if $E|x(t)|^\alpha < \infty$ $(\alpha \ge 1)$ then $|x(t)|^\alpha$ is a submartingale; consequently

$$P[\sup_{s\le t}|x(s)|^\alpha \ge \lambda] \le \frac{E|x(t)|^\alpha}{\lambda}.$$

Let $\mathcal{F}_t$ be an increasing sequence of $\sigma$-fields, $t \ge 0$. A random variable $\tau$ with range in $[0,\infty]$ is called a stopping time with respect to $\mathcal{F}_t$ if

$$\{\tau \le \lambda\} \in \mathcal{F}_\lambda \quad \forall\, \lambda > 0.$$

If $\mathcal{F}_t = \sigma(x(s),\ s \le t)$, we say that $\tau$ is a stopping time with respect to $x(t)$.

Theorem 2. If $x(t)$ is right continuous martingale and $\tau$ a stopping time with respect to $x(t)$, then $y(t) \equiv x(t \wedge \tau)$ is also right continuous martingale.

Let us fix the following model of $(\Omega, \mathfrak{F})$: $\omega$ varies over the space $\Omega = M[0,\infty)$ of measurable functions from $[0,\infty)$ into $R^1$, $\mathfrak{F}_t^s$ is the $\sigma$-field generated by $\omega(u)$, $s \le u \le t$, $\mathfrak{F}_\infty^s = \bigcup_{t>s} \mathfrak{F}_t^s$, $\mathfrak{F} = \mathfrak{F}_\infty^0$.

Let $p(s,x,t,A)$ be a nonnegative function defined for $0 \le s < t < \infty$, $x \in R^1$, $A$ any Borel set in $R^1$, satisfying:

(i) $x \to p(s,x,t,A)$ is Borel measurable;

(ii) $A \to p(s,x,t,A)$ is a probability measure;

(iii) $p(s,x,t,A) = \int_{R^1} p(s,x,\lambda,dy)p(\lambda,y,t,A) \quad (s < \lambda < t)$

(the Chapman-Kolmogorov equation).

Theorem 3. Under the foregoing assumptions (i)-(iii), there exists a family of probabilities $P_{x,s}$ such that the process $x(t)$, with $x(t,\omega) = \omega(t)$, satisfies:

$$P_{x,s}[x(s,\omega) = x] = 1; \tag{1}$$

$$P_{x,s}[x(t+h,\omega) \in A \mid \mathfrak{F}_t^s] = p(t,x(t,\omega),t+h,A) \tag{2}$$

$$\text{a.e. } (t \ge s, h > 0).$$

The collection $\{\Omega, \mathfrak{F}, \mathfrak{F}_t^s, x(t), P_{x,s}\}$ is called a Markov process and $p$ is called the transition probability function.

To prove the theorem, one introduces the multi-dimensional distribution functions

$$F_{t_0 t_1 \cdots t_m}(x_0, x_1, \ldots, x_m) = \int_{-\infty}^{x_m} \cdots \int_{-\infty}^{x_1} \int_{-\infty}^{x_0} p(t_{m-1}, y_{m-1}, t_m, dy_m)$$

$$\cdots p(t_1, y_1, t_2, dy_2) p(t_0, y_0, t_1, dy_1) \pi(dy_0)$$

where $\pi(dx)$ is a distribution function on $R^1$.

Property (iii) ensures that the family $F_{t_0 \cdots t_m}$ is consistent in the sense that if $x_i \to \infty$ then

$$F_{t_0 \cdots t_{i-1} t_i t_{i+1} \cdots t_m}(x_0, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_m)$$

$$\to F_{t_0 \cdots t_{i-1} t_{i+1} \cdots t_m}(x_0, \ldots, x_{i-1} x_{i+1}, \ldots, x_m).$$

By the Kolmogorov construction there exists, then, a probability $P_{s,\pi}$ such that the finite dimensional distribution functions of $x(t)$ coincide with $F_{t_0 t_1 \cdots t_m}$, $t_0 = s$. The $P_{s,x}$ are obtained by choosing $\pi$ to be the Delta measure supported at $x$. The verification of (1) and (2) will be omitted.

Taking the expectation in (2), we obtain (with $t = s$ and $t + h$ replaced by $t$)

$$E_{x,s} f(x(t)) = \int_{R^1} p(s,x,t,dy) f(y), \tag{3}$$

for $f = I_A$, and hence for any bounded Borel measurable $f$. The property (2) is called the Markov property; it can be recast in the form

$$\text{(4)}\qquad \begin{aligned} E_{x,s}[f(x(t+h))\,|\,\mathcal{F}_t^s] &= \int_{R^1} p(t,x(t),t+h,dy)f(y) \\ &= E_{x(t),t}f(x(t+h)) \quad \forall\ f \text{ bounded, Borel measurable.} \end{aligned}$$

The more general property

$$\text{(5)}\qquad P_{x,s}[x(t+\tau)\in A\,|\,\mathcal{F}_\tau^s] = p(\tau,x(\tau),t+\tau,A) \text{ a.e.,}$$

or, equivalently,

$$\text{(6)}\qquad \begin{aligned} E_{x,s}[f(x(t+\tau))\,|\,\mathcal{F}_\tau^s] &= \int p(\tau,x(\tau),t+\tau,dy)f(y) \\ &= E_{x(\tau),\tau}f(x(t+\tau)) \quad \forall\ f \text{ bounded, Borel measurable} \end{aligned}$$

is called the strong Markov property; here $\tau$ is any stopping time with respect to $\mathcal{F}_t^s$, $t \geq s$, and $\mathcal{F}_\tau^s$ is the $\sigma$-field of all sets $A \in \mathcal{F}_\infty^s$ such that

$$A \cap (\tau \leq t) \in \mathcal{F}_t^s \quad \forall\ t \geq s.$$

The strong Markov property is a very useful tool, and we shall therefore give a general condition under which it is satisfied.

Definition. If $\forall\ \lambda > 0$ and $f$ bounded continuous,

$$(t,z) \to \int_{R^1} p(t,z,t+\lambda,dy)f(y) \text{ is continuous,}$$

then we say that $p$ satisfies the <u>Feller property</u>.

<u>Theorem 4</u>. A right continuous Markov process with $p$ satisfying the Feller property satisfies the strong Markov property.

If $p(s,x,t,A) = p(0,x,t-s,A) \equiv p(t-s,x,A)$ then we speak of <u>time-homogeneous</u> Markov process and write $P_x = P_{x,0}$. The strong Markov property becomes:

$$E_x[f(x(t+\tau)) \mid \mathcal{F}_\tau] = E_{x(\tau)} f(x(t)), \quad \mathcal{F}_\tau = \mathcal{F}_\tau^0.$$

<u>Problems</u>. 1. Let $x(t)$ be an n-dimensional continuous process. Let $\tau$ be the first time $x(t)$ hits a given closed set $A \subset R^n$. Prove that $\tau$ is a stopping time.

2. If $\tau,\sigma$ are stopping times then $\tau \wedge \sigma$ is a stopping time.

3. Let

$$p(t,x,A) = \int_A (2\pi t)^{-1/2} e^{-(x-y)^2/2t} dy. \tag{7}$$

Prove that this function satisfies the Chapman-Kolmogorov equation.

The corresponding Markov process is called a <u>Brownian motion</u> (or a <u>Wiener process</u>). One can verify directly that, in this case, if $\pi(dx)$ is the Dirac measure then

$$F_{t_0 t_1 \cdots t_m}(x_0, x_1, \ldots, x_m) = \frac{\chi_0(x_0)}{(2\pi)^{m/2}(\det \Gamma)^{1/2}} \int_{-\infty}^{x_m} \cdots \int_{-\infty}^{x_1}$$

$$\exp[-\frac{1}{2} \sum_{i,j=1}^{m} \Gamma_{ij}^{-1} y_i y_j] dy_1 \cdots dy_m$$

where $\chi_0(x_0) = 0$ if $x_0 < 0$, $=1$ if $x_0 > 0$, $(\Gamma_{ij}^{-1})$ is the inverse of $(\Gamma_{ij})$, and

$$\Gamma_{ij} = \min(t_i, t_j).$$

Random variables $X_1, \ldots, X_m$ are said to have <u>joint normal distribution</u> $N(0,\Gamma)$ if

$$f(u) \equiv E\, e^{i \sum_{j=1}^{m} u_j X_j} = e^{-\frac{1}{2} \sum_{j,k=1}^{m} \Gamma_{jk} u_j u_k}, \quad \Gamma = (\Gamma_{jk}).$$

It then follows that $\Gamma_{jk} = E\, X_j X_k$. If $\det \Gamma \neq 0$ then the joint distribution function $\rho(y)$ of $X_1, \ldots, X_m$ is given by

$$\rho(y) = \frac{1}{(2\pi)^{m/2}(\det \Gamma)^{1/2}} e^{-\frac{1}{2} \sum_{j,k=1}^{m} \Gamma_{jk}^{-1} y_j y_k}.$$

Using these facts we discover that for a Brownian motion

(8) $\qquad x(t_1)-x(t_0),\ x(t_2)-(t_1), \ldots, x(t_m)-x(t_{m-1})$

are independent if $t_0 < t_1 < t_2 < \cdots < t_m$,

(9) $x(t)-x(s)$ is normally distributed with mean 0 and variance $t-s$.

Thus, in particular,

$$E_{x_0} x(t) = 0, \quad E_{x_0} (x(t)-x(s))^2 = t - s,$$

One easily computes

$$E_{x_0} (x(t)-x(s))^4 = C(t-s)^2, \qquad C > 0.$$

Hence, by Theorem 1, $x(t)$ has a continuous version; from now on we always work with this version.

The fact that $x(t,\omega)$ is continuous in $t$, for a.a.$\omega$, means that the set of points $\omega$ for which $\omega(t) = x(t,\omega)$ is not continuous for all $t \geq 0$ is of $P_x$ measure 0. Thus, we may remove this set from our space without affecting anything. From now on we work with this slightly revised space; we continue to denote it by $\Omega$, but now the points $\omega$ of $\Omega$ are elements of $C[0,\infty)$, and the measurable sets are defined in the obvious way.

<u>Theorem 5</u>. Almost all sample paths of a Brownian motion are Hölder continuous with any exponent $\alpha$, $\alpha < 1/2$, and nowhere Hölder continuous with any exponent $\alpha \geq 1/2$.

We also mention the <u>iterated logarithm laws</u>:

$$\overline{\lim_{t \downarrow 0}} \frac{x(t)}{\sqrt{2t \log \log(1/t)}} = 1 \text{ a.s.,}$$

$$\overline{\lim_{t\uparrow\infty}} \frac{x(t)}{\sqrt{2t \log \log t}} = 1 \text{ a.s.}$$

Problems. 4. If $\tau$ is a stopping time for a Brownian motion $x(t)$, then $y(t) \equiv x(t+\tau)-x(\tau)$ is a Brownian motion independent of $\mathcal{F}_\tau$.

5. If $x(t)$ is a process satisfying (8), (9) then its continuous version is a Brownian motion with

$$P_x[x(\cdot);\ x(t_1)\in A_1,\ldots,x(t_k)\in A_k]$$

$$= P[\omega;\ x+x(t_1,\omega)\in A_1,\ldots,x+x(t_k,\omega)\in A_k]$$

and its transition probability function is given by (7).

6. If $x(t)$ is a Brownian motion then

$$E[x(t)-x(s)\,|\,\mathcal{F}_s] = 0 \text{ a.s.},$$

$$E[(x(t)-x(s))^2\,|\,\mathcal{F}_s] = t-s \text{ a.s.}$$

(The converse is also true, namely, if $x(t)$ is continuous martingale and if $x^2(t)-t$ is a martingale, then $x(t)$ satisfies (8), (9) and is thus a Brownian motion.)

An n-dimensional Brownian motion is defined analogously to a 1-dimensional Brownian motion. Thus, in (7) we take $x\in R^n$ and $A$ to be any Borel set in $R^n$. In terms of the components $(x_1(t),\ldots,x_n(t))$ of the process $x(t)$, each $x_i(t)$ corresponds to a 1-dimensional Brownian motion and the processes $x_i(t)$, $t \geq 0$ are mutually independent.

## §2. The stochastic integral

We take a 1-dimensional Brownian motion and denote it by $w(t)$; the probability and expectation corresponding to $w(0) = 0$ will be denoted by $P$ and $E$.

Let $\mathcal{F}_t$ be an increasing family of $\sigma$-fields $(t \geq 0)$ such that

$$\sigma(w(h+t)-w(t), h \geq 0) \text{ is independent of } \mathcal{F}_t.$$

For example, if $\mathcal{G}$ is a $\sigma$-field independent of the Brownian motion, we can take $\mathcal{F}_t$ to be the $\sigma$-field generated by $\sigma(w(s), s \leq t)$ and $\mathcal{G}$.

Denote by $\mathcal{B}[0,T]$ the Borel $\sigma$-field of the interval $[0,T]$.

Definition. A stochastic process $f(t)$, $0 \leq t \leq T$, is called nonanticipative with respect to (or adapted to) $\mathcal{F}_t$ if

(i) $\forall\, t \in [0,T]$, $f(t)$ is separable and $\mathcal{F}_t$ measurable;

(ii) $\forall\, T_0 \in (0,T]$ the function $(t,\omega) \to f(t,\omega)$ from $[0,T_0] \times \Omega \to R^1$ is $\mathcal{B}[0,T_0] \times \mathcal{F}_{T_0}$ measurable.

If, in addition,

$$P[\int_0^T |f(t)|^2 dt < \infty] = 1 \quad (E \int_0^T |f(t)|^2 dt < \infty)$$

then we say that $f$ belongs to $L_w^2[0,T]$ $(M_w^2[0,T])$.

A step function is a stochastic process $f(t)$ for which there exists a partition $\{t_i\}$ of the $t$-interval such that

$f(t) = f(t_i)$ if $t_i \le t < t_{i+1}$.

Lemma 1. Let $f \in L_w^2[0,T]$. Then there exist sequences of functions $g_n \in L_w^2[0,T]$, $h_n \in L_w^2[0,T]$ such that

(i) $g_n$ is continuous, $\lim_{n\to\infty} \int_0^T |f(t)-g_n(t)|^2 dt = 0$ a.s.;

(ii) $h_n$ is a step function, $\lim_{n\to\infty} \int_0^T |f(t)-h_n(t)|^2 dt = 0$ a.s.

If in addition $f \in M_w^2[0,T]$ then the above assertions are valid with $L^2$ replaced by $M^2$, and with $\int_0^T$ replaced by $E\int_0^T$.

Definition. If $f$ is a step function in $L_w^2[0,T]$, with $f(t) = f_i$ when $t_i \le t \le t_{i+1}$, the sum

$$\sum_k f(t_k)(w(t_{k+1})-w(t_k)) \tag{1}$$

is denoted by $\int_0^T f(t)dw(t)$ and is called the stochastic integral of $f$ with respect to the Brownian motion $w(t)$.

Recall that almost every sample path $t \to w(t,\omega)$ is not of bounded variation. Therefore we cannot define the integral

$$\int f(t)dw(t) \tag{2}$$

pathwise, for any (say) bounded measurable $f$. The definition which we shall soon introduce is based on approximation of $f$ by step functions and on the particular definition (1). (If for instance we change (1) just "slightly" by taking $\sum f(t_{k+1})(w(t_{k+1})-w(t_k))$, then the resulting definition for (2) will be quite different).

Lemma 2. If $f$ is a step function in $M_w^2[0,T]$,

$$E[\int_0^T f(t)dw(t)] = 0, \tag{3}$$

$$E|\int_0^T f(t)dw(t)|^2 = E\int_0^T f^2(t)dt. \tag{4}$$

The proof is direct.

Lemma 3. If $f$ is a step function in $L_w^2[0,T]$, then for any $\varepsilon > 0$, $N > 0$,

$$P[|\int_0^T f(t)dw(t)| > \varepsilon] \leq P[\int_0^T f^2(t)dt > N] + \frac{N}{\varepsilon^2}. \tag{5}$$

Proof. Define $\varphi(t) = f(t)$ if $t_k \leq t < t_{k+1}$,
$\sum_{j=0}^{k} f^2(t_j)(t_{j+1}-t_j) \leq N$ and $\varphi(t) = 0$ otherwise. Then
$E\int_0^T \varphi^2(t) \leq N$ and

$$P[|\int_0^T f(t)dw(t)| > \varepsilon] \leq P[|\int_0^T \varphi(t)dw(t)| > \varepsilon]$$
$$+ P[\int_0^T f^2(t)dt > N]$$

since $f(t) = \varphi(t)$ for all $0 < t < s$ if $\int_0^s f^2(t)dt < N$. Estimate now the first term on the right by Chebyshev's inequality.

Let $f \in L_w^2[0,T]$ and choose $f_n$ step functions in $L_w^2[0,T]$ such that

$$\int_0^T |f_n(t)-f(t)|^2 dt \xrightarrow{P} 0 \text{ if } n \to \infty$$

By Lemma 3,

$$\int_0^T f_n(t)dw(t)$$

is convergent in probability. We denote the limit of

$$\int_0^T f(t)dw(t)$$

and call it the <u>stochastic integral</u> (or the <u>Ito integral</u>) of f(t) with respect to the Brownian motion w(t).

The above definition is clearly independent of the approximation $f_n$.

<u>Theorem 4</u>. The assertions of Lemmas 1 and 2 are valid for any f in $M_w^2[0,T]$ and $L_w^2[0,T]$, respectively.

This follows by approximation.

<u>Problems</u>. 1. One can define $L_w^2[\alpha,\beta]$, $M_w^2[\alpha,\beta]$ and $\int_\alpha^\beta f(t)dw(t)$ in the obvious way. Prove that if $f \in M_w^2[\alpha,\beta]$,

$$E[\int_\alpha^\beta f(t)dw(t) \mid \mathcal{F}_\alpha] = 0$$

$$E[|\int_\alpha^\beta f(t)dw(t)|^2 \mid \mathcal{F}_\alpha] = \int_\alpha^\beta E[f^2(t) \mid \mathcal{F}_\alpha]dt.$$

2. If f,g belong to $L_w^2[\alpha,\beta]$ and if f(t) = g(t) for all

$\alpha \leq t \leq \beta$, $\omega \in \Omega_0$, then

$$\int_\alpha^\beta f(t)dw(t) = \int_\alpha^\beta g(t)dw(t) \text{ a.e. on } \Omega_0.$$

3. If $f \in L_w^2[\alpha,\beta]$, $f$ continuous, then for any sequence of partitions $(t_{n,1},\ldots,t_{n,m_n})$ of $[\alpha,\beta]$ with mesh $\to 0$

$$\Sigma f(t_{n,k})(w(t_{n,k+1})-w(t_{n,k})) \xrightarrow{P} \int_\alpha^\beta f(t)dw(t).$$

<u>Theorem 5</u>. Let $f \in L_w^2[0,T]$. Then the process

$$I(t) = \int_0^t f(s)dw(s)$$

is a martingale and, further, it has a continuous version.

In the future we work only with this version and call it the <u>indefinite integral</u>.

<u>Proof</u>. First take $f \in M_w^2[0,T]$ and approximate it by step functions $f_n$ as in Lemma 1. Let

$$I_n(t) = \int_0^t f_n(s)dw(s).$$

By Problem 1, $I_n(t)$ is a martingale. Hence, by the martingale inequality,

$$P\{\sup_{0 \leq t \leq T} |I_n(t)-I_m(t)| > \varepsilon\}$$

$$\leq \frac{1}{\epsilon^2}E\left|\int_0^T (f_n(s)-f_m(s))dw(s)\right|^2$$

$$\leq \frac{1}{\epsilon^2}E\int_0^T |f_n(s)-f_m(s)|^2 ds \to 0 \text{ if } n,m \to \infty$$

Taking $\epsilon = 1/2^k$ it follows that for some $n_k$ sufficiently large,

$$P\{\sup_{0\leq t\leq T} |I_{n_k}(t)-I_m(t)| > \frac{1}{2^k}\} < \frac{1}{k^2} \quad \text{if} \quad m \geq n_k.$$

We can choose the $n_k$ in such a way that $n_k \uparrow$ if $k \uparrow$. Hence,

$$P\{\sup_{0\leq t\leq T} |I_{n_k}(t)-I_{n_{k+1}}(t)| > \frac{1}{2^k}\} < \frac{1}{k^2} \quad (k = 1,2,\ldots).$$

Since $\Sigma k^{-2} < \infty$, the Borel-Cantelli lemma implies that

$$P\{\sup_{0\leq t\leq T} |I_{n_k}(t)-I_{n_{k+1}}(t)| > \frac{1}{2^k} \text{ i.o.}\} = 0.$$

i.e., for a.a. $\omega$

$$|I_{n_k}(t)-I_{n_{k+1}}(t)| \leq \frac{1}{2^k} \text{ for all } 0 \leq t \leq T, \text{ if } k \geq k_0(\omega).$$

But then, with probability one, $\{I_{n_k}(t)\}$ is uniformly convergent in $t\in[0,T]$. The limit $J(t)$ is therefore a continuous function in $t\in[0,T]$ for a.a. $\omega$. Since

$$\int_0^t f_n(s)dw(s) \to \int_0^t f(s)dw(s) \text{ in } L^2(\Omega),$$

it follows that

$$J(t) = \int_0^t f(s)dw(s) \text{ a.s.}$$

Thus, the indefinite integral has a continuous version.

Consider now the general case where $f \in L_w^2[0,T]$. For any $N > 0$, let

$$\chi_N(z) = \begin{cases} 1 & \text{if } z \leq N, \\ 0 & \text{if } z > N, \end{cases}$$

and introduce the function

$$f_N(t) = f(t)\chi_N(\int_0^t f^2(s)ds).$$

It is easily checked that $f_N$ belongs to $M_w^2[0,T]$. Hence, by what was already proved, a version of

$$J_N(t) = \int_0^t f_N(s)dw(s) \quad (0 \leq t \leq T)$$

is a continuous process.

Let

$$\Omega_N = \{\int_0^T f^2(t)dt < N\}.$$

If $\omega \in \Omega_N$, then $f_N(t) = f_M(t)$ for $0 \leq t \leq T$, $M > N$. By Problem 2 it follows that for a.a. $\omega \in \Omega_N$

$$J_N(t) = J_M(t) \quad \text{if} \quad 0 \leq t \leq T.$$

Therefore

$$\tilde{J}(t) = \lim_{M\to\infty} J_M(t)$$

is continuous in $t\in[0,T]$ for a.a. $\omega\in\Omega_N$.

Since $\Omega_N \uparrow$, $P(\Omega_N) \uparrow 1$ if $N \uparrow \infty$, $\tilde{J}(t)$ $(0 \le t \le T)$ is a continuous process.

But since for each $t\in(0,T]$,

$$P\{\int_0^t |f(s)-f_M(s)|^2 ds > 0\} = P\{\int_0^t f^2(s)ds > M\} \to 0$$

as $M \to \infty$, we have,

$$J_M(t) \xrightarrow{P} \int_0^t f(s)dw(s) = I(t).$$

Consequently, $I(t)$ has the continuous version $\tilde{J}(t)$.

Problems. 4. If $x(t)$ is a separable martingale then, for any $\alpha > 1$,

$$E[\sup_{0\le t\le T} |x(t)|^\alpha] \le (\frac{\alpha}{\alpha-1})^\alpha E|x(T)|^\alpha.$$

Use this fact to prove that, for $f\in M_w^2[0,T]$,

$$E[\sup_{0\le t\le T} |\int_0^t f(s)dw(s)|^2] \le 4E\int_0^T f^2(t)dt$$

5. If $f\in M_w^2[0,T]$, $\tau$ a stopping time with respect to $\mathcal{F}_t$, $0 \le \tau \le T$, then $E\int_0^{\tau\wedge t} f(s)dw(s) = 0$.

6. Define $\int_{\zeta_1}^{\zeta_1} f(t)dw(t) = \int_0^{\zeta_2} f(t)dw(t)-\int_0^{\zeta_1} f(t)dw(t)$, where

$\zeta_i$ are nonnegative random variables, $\zeta_1 \le \zeta_2$. Then

$$E \int_{\zeta_1}^{\zeta_2} f(t)dw(t) = 0$$

if $f \in M_w^2[0,T]$ provided $\zeta_1, \zeta_2$ are $\mathcal{F}_t$ stopping time, $0 \le \zeta_1 \le \zeta_2 \le T$.

§3. Ito's formula

Suppose $\xi(t)$ is a stochastic process satisfying for any $0 \le t_1 < t_2 \le T$

$$\xi(t_2)-\xi(t_1) = \int_{t_1}^{t_2} a(t)dt + \int_{t_1}^{t_2} b(t)dw(t)$$

where $a \in L_w^1[0,T]$, $b \in L_w^2[0,T]$ ($L_w^1[0,T]$ is defined similarly to $L_w^2[0,T]$, with $P[\int_0^T |f(t)|dt < \infty] = 1$). Then we say that $\xi(t)$ has stochastic differential $d\xi$ given by

$$d\xi(t) = a(t)dt + b(t)dw(t).$$

Lemma 1. Let $\{t_{n,j};\ j = 1,2,\ldots,m_n\}$ be a sequence of partitions of an interval $a < t < b$ with mesh $\delta_n \to 0$. Then

$$S_n = \sum_{j=1}^{m_n} (w(t_{n,j})-w(t_{n,j-1}))^2$$

converges in $L^2$ to the constant $b-a$.

Proof. Write $t_j = t_{n,j}$, $m = m_n$. Then

$$S_n-(b-a) = \sum_{j=1}^{n} [(w(t_j)-w(t_{j-1}))^2-(t_j-t_{j-1})].$$

Since the summands are independent and of mean 0,

$$E[S_n-(b-a)]^2 = E\sum_{j=1}^{m} [(w(t_j)-w(t_{j-1}))^2-(t_j-t_{j-1})]^2$$

$$= \sum_{j=1}^{m} E[(Y_j^2-1)(t_j-t_{j-1})]^2$$

where

$$Y_j = \frac{w(t_j)-w(t_{j-1})}{(t_j-t_{j-1})^{1/2}}.$$

The $Y_j$ are identically normally distributed; hence

$$E[S_n-(b-a)]^2 = E(Y_1^2-1)^2 \sum_{j=1}^{m} (t_j-t_{j-1})^2$$

$$\leq E(Y_1^2-1)^2\cdot(b-a)\delta_n \to 0 \text{ if } n \to \infty.$$

We shall use Lemma 1 in order to prove that

$$\text{(1)}\qquad \int_{t_1}^{t_2} w(t)dw(t) = \frac{1}{2}(w^2(t_2)-w^2(t_1)) - \frac{1}{2}(t_2-t_1).$$

Indeed,

$$\int_{t_1}^{t_2} w(t)dw(t) = \lim \Sigma\, w(t_{n,j})(w(t_{n,j+1})-w(t_{n,j}))$$

$$= \frac{1}{2} \lim \Sigma\{[w^2(t_{n,j+1})-w^2(t_{n,j})]-[w(t_{n,j+1})-w(t_{n,j})]^2\}$$

$$= \frac{1}{2}(w^2(t_2)-w^2(t_j)) - \frac{1}{2} \lim \Sigma[w(t_{n,j+1})-w(t_{n,j})]^2$$

and the last limit, in probability, is $t_2-t_1$.

We can rewrite (1) in the form

$$dw^2(t) = 2w(t)dw(t) + t. \tag{2}$$

As another example we compute

$$\int_{t_1}^{t_2} dw(t) = \lim \Sigma\, t_{n,j}[w(t_{n,j+1})-w(t_{n,j})].$$

Clearly

$$\int_{t_1}^{t_2} w(t)dt = \lim \Sigma\, w(t_{n,j+1})(t_{n,j+1}-t_{n,j}),$$

and adding both equations,

$$\int_{t_1}^{t_2} t\, dw(t) + \int_{t_1}^{t_2} w(t)dt = t_2 w(t_2)-t_1 w(t_1).$$

Thus

$$d(t\, w(t)) = w(t)dt + t\, dw(t) \tag{3}$$

<u>Theorem 2</u>. Let $f(x,t)$ be a continuous function for $(x,t)\in R^1 \times [0,\infty)$ with continuous derivatives $f_x, f_t, f_{xx}$

Let $d\xi(t) = a(t)dt + b(t)dw(t)$. Then $f(\xi(t),t)$ has a stochastic differential

(4)

$$df(\xi(t),t) = [f_t(\xi(t),t) + a(t)f_x(\xi(t),t) + \frac{1}{2}b^2(t)f_{xx}(\xi(t),t)]dt$$

$$+ f_x(\xi(t),t)b(t)dw(t).$$

This formula is called Ito's formula.

Proof. If $d\xi = adt + bdw$, define $fd\xi = fadt + fbdw$. If

$$d\xi_i(t) = a_i(t)dt + b_i(t)dw(t) \quad (i = 1,2)$$

then

(5) $$d(\xi_1(t)\xi_2(t)) = \xi_1(t)d\xi_2(t) + \xi_2(t)d\xi_1(t) + b_1(t)b_2(t)dt.$$

Indeed, for $a_i, b_i$ constants, (5) is a consequence of the rules (2), (3). When $a_i, b_i$ are step functions, (5) in its integrated form follows from its integrated form on each step of the $a_i, b_i$. For general $a_i, b_i$ it then follows by approximation.

Applying (5) one can establish by induction that

$$dw^m(t) = mw^{m-1}(t)dw(t) + \frac{1}{2}m(m-1)w^{m-2}(t)dt.$$

This can be used to show that

$$dQ(w(t)) = Q'(w(t))dw(t) + \frac{1}{2}Q''(w(t))dt$$

for a polynomial $Q$. Next one uses this relation and (5) in order to establish Ito's formula for a function $Q(x)g(t)$ where $Q$ is a polynomial and $g(t)$ is in $C^1$. By linearily and approximation one then establishes (4) in case $\xi(t) = w(t)$. Similarly one can establish (4) when $d\xi(t) = adt + bdw$ and $a,b$ are step functions and, by approximation, also for general $a,b$.

<u>Problems</u>. 1. Let $f \in M_w^{2m}[0,T]$, $m$ positive integer. Use Ito's formula with $f = x^{2m}$, $\xi(t) = \int_0^t f(s)dw$ to show that

$$E\left|\int_0^T f(t)dw(t)\right|^{2m} \le m(2m-1)^m T^{m-1} E\left[\int_0^T f^{2m}(t)dt\right] \tag{6}$$

2. Let $f \in L_w^2[0,T]$ and let $\alpha,\beta$ be positive numbers. Prove the <u>exponential martingale inequality</u>

$$P\left\{\max_{0\le t\le T}\left[\int_0^t f(\lambda)dw(\lambda) - \frac{\alpha}{2}\int_0^t f^2(\lambda)d\lambda\right] > \beta\right\} \le e^{-\alpha\beta} \tag{7}$$

[<u>Hint</u>: For $f$ bounded step function, let

$$\xi(t) = \int_0^t f(\lambda)dw(\lambda) - \frac{1}{2}\int_0^t f^2(\lambda)d\lambda$$

and apply Ito's formula with $e^x$ to deduce that

$$\zeta(t) = \zeta(s) + \int_s^t e^{\xi(\lambda)} f(\lambda)dw(\lambda)$$

where $\zeta(t) = e^{\xi(t)}$. Apply the martingale inequality. Finally approximate general $f$ by step functions].

For n-dimensional Brownian motion $w(t) = (w_1(t),\ldots,w_n(t))$ one defines

$$\int_\alpha^\beta b(t)dw(t) = \sum_{j=1}^{n} \int_\alpha^\beta b_j(t)dw_j(t)$$

where $b = (b_1,\ldots,b_n)$, $b_j \in L_w^2[\alpha,\beta]$. If $b = (b_{ij})$ then

$$E\left|\int_\alpha^\beta b(t)dw(t)\right|^2 = E\int_\alpha^\beta \sum_{i,j}(b_{ij}(t))^2 dt.$$

The concept of stochastic differential

$$d\xi(t) = a(t)dt + b(t)dw(t)$$

where $a = (a_1,\ldots,a_m)$, $b = (b_{ij})$ $(1 \le i \le m,\ 1 \le j \le n)$ is defined in the obvious way, and the corresponding Ito's formula is

$$\begin{aligned} du(\xi(t),t) = {} & [u_t(\xi(t),t) + \sum_{i=1}^{m} a_i(t)u_{x_i}(\xi(t),t) \\ & + \frac{1}{2}\sum_{\ell=1}^{n}\sum_{i,j=1}^{m} b_{i\ell}(t)b_{j\ell}(t)u_{x_ix_j}(\xi(t),t)]dt \\ & + \sum_{\ell=1}^{n}\sum_{i=1}^{m} b_{i\ell}(t)u_{x_i}(\xi(t),t)dw_\ell(t). \end{aligned} \tag{8}$$

Here $u_t, u_{x_i}, u_{x_ix_j}$ are assumed to be continuous.

The proof uses the special case $n = 1$ and the relation

$$d(w_1w_2) = w_1dw_2 + w_2dw_1 \tag{9}$$

To prove (9), simply notice that $w(t) = (w_1(t) + w_2(t))/\sqrt{2}$ is a Brownian motion, so that $dw^2 = dt + 2wdw$.

Ito's formula (8) can be written in the form

$$\text{(10)}\qquad \begin{aligned} du(\xi(t),t) &= u_t(\xi(t),t)dt + \sum_{i=1}^{n} u_{x_i}(\xi(t),t)d\xi_i \\ &+ \sum_{i,j=1}^{n} u_{x_i x_j}(\xi(t),t)d\xi_i d\xi_j \end{aligned}$$

provided we define the multiplication table:

$$dw_i dt = 0,\ dtdt = 0,\ dw_i dw_j = 0 \text{ if } i \neq j,\ dw_i dw_i = dt.$$

Introducing the matrix $(a_{ij})$ where $a_{ij} = \sum_{k=1}^{n} b_{ik} b_{jk}$ and the operator

$$\text{(11)}\qquad Lu = \frac{1}{2} \sum_{i,j=1}^{m} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{m} a_i \frac{\partial u}{\partial x_i} + \frac{\partial u}{\partial t}$$

we can also write (8) in the form

$$\text{(12)}\qquad du(\xi(t),t) = Lu(\xi(t),t)dt + u_x(\xi(t),t)\cdot b(t)dw(t).$$

It follows that if $Lu$ and $u_x \cdot b$ are in $M_w^1[0,T], M_w^2[0,T]$ respectively, then, for any stopping time $\tau$, $0 \leq \tau \leq T$,

$$\text{(13)}\qquad Eu(\xi(\tau),\tau) - Eu(\xi(0),0) = E\int_0^{\tau} Lu(\xi(s),s)ds$$

<u>Problems</u>. 3. Let $\xi(t) = \int_0^t b(t)dw(t)$ where $b$ is $n \times n$ matrix with elements $b_{ij}$ in $M_w^2[0,T]$. Suppose $d\xi_i d\xi_j = 0$ if $i \neq j$, $d\xi_i d\xi_i = dt$. Prove that $\xi(t)$ is an n-dimensional Brownian motion.

[<u>Hint</u>: Suppose $b_{ij}$ are step functions. Let $\zeta(t) = \exp[i\gamma\cdot\xi(t) + \gamma^2 t/2]$. By Ito's formula $d\zeta = i\zeta\gamma dw$.

Deduce that $E[e^{i\gamma\cdot\xi(t)}|\mathcal{F}_s] = e^{i\gamma\cdot\xi(s)-\gamma^2(t-s)/2}$ and the following fact: if $\xi(t_0),\ldots,\xi(t_n)$ have joint normal distribution $N(0,\Gamma)$ with $\Gamma = (\Gamma_{ij})$, $\Gamma_{jk} = \min(t_j,t_k)$ then $\xi(t)$ in a Brownian motion.]

## §4. Stochastic differential equations

Let $b(x,t) = (b_1(x,t),\ldots,b_n(x,t))$, $\sigma(x,t) = (\sigma_{ij}(x,t))$ where $1 \le i,j \le n$. If $\xi(t)$ is a stochastic process satisfying

$$d\xi(t) = b(\xi(t),t)dt + \sigma(\xi(t),t)dw(t), \tag{1}$$

$$\xi(0) = \xi_0 \tag{2}$$

then we say that $\xi(t)$ satisfies the system of <u>stochastic differential equations</u> (1) and the <u>initial condition</u> (2).

We shall assume that

$$\begin{aligned} &b(x,t),\sigma(x,t) \text{ are measurable in } R^n \times [0,T], \\ &|b(x,t)-b(\bar{x},t)| \le K|x-\bar{x}|, \quad |b(x,t)| \le K(1+|x|), \\ &|\sigma(x,t)-\sigma(\bar{x},t)| \le K|x-\bar{x}|, \quad |\sigma(x,t)| \le K(1+|x|) \end{aligned} \tag{3}$$

and that

$$\begin{aligned} &\xi_0 \text{ is independent of the Brownian motion } w(t), \\ &E|\xi_0|^2 < \infty. \end{aligned} \tag{4}$$

Thus, in particular, $\xi_0$ may be any constant.

We take $\mathcal{F}_t$ to be the $\sigma$-field generated by $w(s), s \leq t$ and the $\sigma$-field of $\xi_0$.

Theorem 1. Let (3), (4) hold. Then there exists a unique solution $\xi(t)$ of (1), (2) in $M_w^2[0,T]$.

Uniqueness means that if $\tilde{\xi}(t)$ is another solution in $M_w^2[0,T]$, then

$$P[\xi(t) \neq \tilde{\xi}(t)] = 0 \quad \forall\ t \in [0,T].$$

Proof. Uniqueness. If $\xi_1, \xi_2$ are two solutions, then

$$\xi_1(t)-\xi_2(t) = \int_0^t [b(\xi_1(s),s)-b_2(\xi_2(s),s)]ds$$
$$+ \int_0^t [\sigma(\xi_1(s),s)-\sigma(\xi_2(s),s)]dw(s).$$

Hence

$$E|\xi_1(t)-\xi_2(t)|^2 \leq C \int_0^t E|\xi_1(s)-\xi_2(s)|^2 ds$$

implying $E|\xi_1(t)-\xi_2(t)|^2 = 0$.

Existence. Define

$$(5) \quad \xi_{m+1}(t) = \xi_0 + \int_0^t b(\xi_m(s),s)ds + \int_0^t \sigma(\xi_m(s),s)dw(s) \qquad (m \geq 0)$$

and make the inductive assumption that $\xi_k \in M_w^2[0,T]$ and

$$(6) \qquad E|\xi_{k+1}(t)-\xi_k(t)|^2 \leq \frac{(Mt)^{k+1}}{(k+1)!} \text{ for } 0 \leq k \leq m-1,$$

where $M$ is a positive constant depending on $K,T$.

It can easily be shown that (6) holds for $k = m$ and $\xi_{m+1} \in M_w^2[0,T]$. Next, using Problem 4, §2 we find that

$$E \sup_{0\le t\le T} |\xi_{m+1}(t)-\xi_m(t)|^2 \le C \frac{(MT)^m}{m!}.$$

Hence

$$P[\sup_{0\le t\le T} |\xi_{m+1}(t)-\xi_m(t)| > \frac{1}{2^m}] \le 2^m C \frac{(MT)^m}{m!}.$$

The Borel-Cantelli lemma now implies

$$P[\sup_{0\le t\le T} |\xi_{m+1}(t)-\xi_m(t)| > \frac{1}{2^m} \text{ i.o.}] = 0.$$

Thus, for almost any $\omega$ there is an $m_0 = m_0(\omega)$ such that

$$\sup_{0\le t\le T} |\xi_{m+1}(t)-\xi_m(t)| \le \frac{1}{2^m} \text{ if } m \ge m_0(\omega).$$

It follows that the series

$$\xi_0 + \sum_{m=0}^{\infty} (\xi_{m+1}(t)-\xi_m(t))$$

is uniformly convergent in $t\in[0,T]$. Denoting the limit by $\xi(t)$ we then have $\xi_k(t) \to \xi(t)$ uniformly in $t\in[0,T]$, for almost any $\omega$. But this implies

$$b(\xi_m(t),t) \to b(\xi(t),t)$$

$$\sigma(\xi_m(t),t) \to \sigma(\xi(t),t)$$

uniformly in $t$, for a.a.$\omega$. Taking $m \to \infty$ in (5) we find that $\xi(t)$ is a solution of (1), (2).

Finally, since by (5),

$$E|\xi_{m+1}(t)|^2 \le C + C\int_0^t E|\xi_m(s)|^2 dx.$$

we get, by induction,

$$E|\xi_{m+1}(t)|^2 \le (C + C^2 t + \cdots + C^{m+2}\frac{t^{m+1}}{(m+1)!})(1 + E|\xi_0|^2)$$

so that

$$E|\xi_{m+1}(t)|^2 \le C_1.$$

The same inequality then holds for $\xi(t)$, i.e., $\xi(t) \in M_W^2[0,T]$.

Theorem 1 can be extended in two directions:

Uniqueness. If $\xi_i$ $(i = 1,2)$ is a solution of a stochastic differential system with $b^i, \sigma^i$ such that $b^1(x,t) = b^2(x,t)$, $\sigma^1(x,t) = \sigma^2(x,t)$ if $(x,t)$ is in a domain $U$, and if $\xi_1(0) = \xi_2(0) \in U$, then $\xi_1(t) = \xi_2(t)$ a.e. for all $t < \tau$ where $\tau$ is the exit time of $\xi_1(t)$ (or $\xi_2(t)$) from $U$.

Existence. The existence part remains true if the uniform Lipschitz conditions on $\sigma$, $b$ are replaced by a Lipschitz condition in every compact set.

Theorem 2. The solution $\xi(t)$ of (1), (2) satisfies

(7) $\quad P[\xi(t)\in A|\mathcal{F}_s] = P[\xi(t)\in A|\xi(s)] = p(s,\xi(s),t,A)$ a.s.

for all $t > s$ and any Borel set $A$; further, $p(s,x,t,A)$ is a transition probability function.

Proof. Let

$$\xi_k(t) = \gamma + \int_s^t b(\xi_{k-1}(\lambda),\lambda)d\lambda + \int_s^t \sigma(\xi_{k-1}(\lambda),\lambda)dw(\lambda)$$

$$\xi_0(t) = \gamma.$$

If $\gamma = \xi(s)$ then $\xi_k(t) \to \xi(t)$ (by the proof of Theorem 1). Denote by $\xi_{x,s}(t)$ the solution of (1) with $\xi(s) = x$.

By induction one can show that each $\xi_k(t)$ is measurable with respect to the $\sigma$-field generated by $\gamma$ and $\sigma(w(\lambda+s)-w(s), s \le \lambda \le t)$; more specifically, there exists a sequence of Borel measurable functions $F_n(t,x_0,x_1,\ldots,x_{\mu_m})$ such that

$$F_m(t,\gamma,w(u_{m,1}+s)-w(s),\ldots,w(u_{m,\mu_m}+s)-w(s)) \to \xi_k(t)$$

for a.a.$\omega$, uniformly in $t$, for some $0 < u_{m,i} < t - s$. Therefore, the same holds for $\xi(t)$. Thus (with another sequence $F_m$) a.s.

(8) $$\xi(t) = \lim_{m\to\infty} F_m(t,\xi(s),w(u_{m,1}+s)-w(s),\ldots,w(u_{m,\mu_m}+s)-w(s)),$$

(9) $$\xi_{x,s}(t) = \lim_{m\to\infty} F_m(t,x,w(u_{m,1}+s)-w(s),\ldots,w(u_{m,\mu_m}+s)-w(s)),$$

But

$$E[F(\xi(s),w(u_1+s)-w(s),\ldots,w(u_k+s)-w(s))|\mathcal{F}_s]$$

$$= \{E\ F(x_0,w(u_1+s)-w(s),\ldots,w(u_k+s)-w(s))\}_{x_0=\xi(s)}$$

as seen by taking first $F = F_0(x_0)F_1(x_1,x_2,\ldots,x_k)$. Taking $F = f(F_m)$, $f$ a bounded continuous function and using (8), (9), we get

$$E[f(\xi(t)|\mathcal{F}_s] = E[f(\xi(t)/\xi(s)] = \emptyset(x)|_{x=\xi(s)}$$

where $\emptyset(x) = Ef(\xi_{x,s}(t))$. By approximation, this relation holds for any bounded Borel function.

It remains to prove that $p$ is a transition probability function. The Chapman-Kolmogorov equation is the only condition that is not obvious. Notice that

(10) $$p(s,x,t,A) = P(\xi_{x,s}(t)\in A).$$

Hence

$$\int p(s,x,t,dy)\psi(y) = \int \psi(\xi_{x,s}(t))dP$$

for any bounded Borel function $\psi$. Taking $\psi(y) = p(t,y,\tau,A)$, $t < \tau$, we get

$$\int p(s,x,t,dy)p(t,y,\tau,A) = \int p(t,\xi_{x,s}(t),\tau,A)dP$$

$$= E\ p(t,\xi_{x,s}(t),\tau,A) = E\ p[\xi_{x,s}(\tau)\in A|\xi_{x,s}(t)]$$

where (7) was used in the last equality. Since the last expression is equal to

$$E\, E[I_A(\xi_{x,s}(\tau)) | \xi_{x,s}(t)] = E\, I_A(\xi_{x,s}(\tau))$$

$$= P[\xi_{x,s}(\tau) \in A] = p(s,x,\tau,A),$$

the proof is complete.

We can now identify the solution of the stochastic differential system with a Markov process. Indeed, take $\Omega$ to be the space $\mathcal{C}_T$ of all continuous n-vector functions $x(\cdot)$, $\mathcal{M}_t^s$ the $\sigma$-field generated by $x(u)$, $s \le u \le t$, and

$$P_{x,s}[x(\cdot) \in B] = P[\omega;\ \xi_{x,s}(\cdot,\omega) \in B]. \tag{11}$$

We have to verify the Markov property

$$P_{x,s}[x(t+h) \in A | \mathcal{M}_t^s] = p(t,x(t),t+h,A) \text{ a.s.} \tag{12}$$

<u>Proof of (12)</u>. Since

$$P[\xi_{x,s}(t+h) \in A | \xi_{x,s}(\lambda), \lambda \le t] = p(t,\xi_{x,s}(t),t+h,A),$$

for any $s \le t_1 < t_2 < \cdots < t_m \le t$ and Borel sets $A_1,\ldots,A_m$,

$$P[\xi_{x,s}(t+h) \in A, \xi_{x,s}(t_1) \in A_1, \ldots, \xi_{x,s}(t_m) \in A_m]$$

$$= \int_{\bigcap_{i=1}^{m} [\xi_{x,s}(t_i) = A_i]} p(t,\xi_{x,s}(t),t+h,A)\,dP$$

$$= \int I_{A_1}(\xi_{x,s}(t_1)) \cdots I_{A_m}(\xi_{x,s}(t_m))\, p(t,\xi_{x,s}(t),t+h,A)\,dP.$$

Hence, by (11),

$$P_{x,s}[x(t+h)\in A, x(t_1)\in A_1,\ldots,x(t_m)\in A_m]$$

$$= \int I_{A_1}(x(t_1)) \cdots I_{A_m}(x(t_m))p(t,x(t),t+h,A)dP_{x,s}$$

$$= \int_{\bigcap_{i=1}^{m}[x(t_i)\in A_i]} p(t,x(t),t+h,A)dP_{x,s}$$

This implies (12).

By the solution of the stochastic system (1) one means the Markov process just constructed, namely,

$$\{\mathcal{C}_T, \mathcal{M}_T^0, \mathcal{M}_t^s, x(t), P_{x,s}\}$$

Notice that $E\, f(\xi_{x,s}(t)) = E_{x,s}f(x(t))$ for any bounded Borel function.

When $b = b(x)$, $\sigma = \sigma(x)$, the Markov process is time-homogeneous.

Theorem 3. The solution of (1) satisfies the Feller property (and therefore also the strong Markov property).

Proof. It is not difficult to see that

$$(13) \qquad E \sup_{\tau\le t\le T} |\xi_{x,s}(t)-\xi_{y,\tau}(t)|^2 \le C(|x-y|^2 + |s-\tau|)$$

if $|x| \le R$, $|y| \le R$, $0 \le s \le \tau \le T$, where $C$ depends on $R,T$. By the Lebesgue bounded convergence theorem,

$$E\, f(\xi_{y,\tau}(t+\tau)) - E\, f(\xi_{x,s}(t+s)) \to 0 \text{ if } y \to x,\ \tau \to s$$

provided $f$ is a bounded continuous function. Also

$$E\, f(\xi_{x,s}(t+\tau)) \to E\, f(\xi_{x,s}(t+s)) \text{ if } \tau \to s.$$

Thus

$$(s,x) \to \int p(s,x,t+s,dy) f(y)$$

is continuous.

Problems. 1. Show that the solution of (1), (2) satisfies

$$E|\xi(t)|^{2m} \le (1 + E|\xi_0|^{2m}) e^{Ct}$$

2. Assume that $b(x,t)$, $\sigma(x,t)$ are continuous and set $a_{ij} = \sum_{k=1}^{n} \sigma_{ik}\sigma_{jk}$. Prove that

$$\frac{1}{h} E|\xi_{x,t}(t+h)-x|^4 \to 0 \text{ if } h \to 0,$$

$$\frac{1}{h} E[\xi_{x,t}(t+h)-x] \to b(x,t) \text{ if } h \to 0,$$

$$\frac{1}{h} E(\xi^i_{x,t}(t+h)-x)(\xi^j_{x,t}(t+h)-x) \to a_{ij}(x,t) \text{ if } h \to 0.$$

3. If $b(x,t)$, $\sigma(x,t)$ are continuous then for any $\varepsilon > 0$, $t \ge 0$, $x \in R^n$,

$$\frac{1}{h} \int_{|y-x|<\varepsilon} p(t,x,t+h,dy) \to 0 \text{ if } h \downarrow 0$$

$$\frac{1}{h} \int_{|y-x|>\varepsilon} (y_i - x_i) p(t,x,t,t+h,dy) \to b_i(x,t) \text{ if } h \downarrow 0.$$

A Markov process with $p$ satisfying these properties is sometimes called a <u>diffusion process.</u>

Suppose now that

$$(14)\qquad \begin{array}{l} D_x^\alpha b(x,t),\ D_x^\alpha \sigma(x,t) \text{ are continuous} \\ \text{and bounded by } C(1+|x|^\beta) \quad (0 \le |\alpha| \le 2,\ \beta > 0) \end{array}$$

for some $C > 0$. One can show that for any $f$ such that

$$(15)\qquad \begin{array}{l} D_x^\alpha f(x) \text{ is continuous and bounded} \\ \text{by } C(1+|x|^\beta) \quad (0 \le |\alpha| \le 2,\ \beta > 0) \end{array}$$

the function

$$\varphi(x) = E\, f(\xi_{x,s}(t))$$

has two continuous derivatives, and

$$|D^\alpha \varphi(x)| \le C_1(1+|x|^\gamma) \text{ if } 0 \le |\alpha| \le 2 \quad (C_1 \ge 0,\ \gamma > 0).$$

The proof is omitted. In this connection we mention also the fact that

$$u(x,t) = E\, f(\xi_{x,t}(T)$$

satisfies the <u>Kolmogorov equation</u> (or the backward parabolic equation)

$$\frac{\partial u}{\partial t} + \Sigma\, b_i \frac{\partial u}{\partial x_i} + \frac{1}{2} \Sigma\, a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} = 0 \text{ in } R^n \times (0,T).$$

Problems. 4. Consider the linear stochastic differential equation

(16) $$d\xi(t) = [\alpha(t) + \beta(t)\xi(t)]dt + [\gamma(t) + \delta(t)\xi(t)]dw(t)$$

where $\alpha, \beta, \gamma, \delta$ are bounded and measurable. Prove:

(a) if $\alpha \equiv 0$, $\gamma \equiv 0$ then the solution $\xi = \xi_0(t)$ is given by

$$\xi_0(t) = \xi_0(0)\exp\{\int_0^t [\beta(s) - \frac{1}{2}\delta^2(s)]ds + \int_0^t \delta(s)dw(s)\}.$$

(b) setting $\zeta(t) = \xi_0(t)\xi(t)$ show that $\xi(t)$ solves the equation (16) if and only if

$$\zeta(t) = \zeta(0) + \int_0^t [\xi_0(s)\alpha(s) - \gamma(s)\delta(s)]ds + \int_0^t \gamma(s)\xi_0(s)dw(s).$$

Thus the solution of (16) is $\zeta(t)/\xi_0(t)$ with $\xi(0) = \zeta(0)/\xi_0(0)$.

## §5. Probabilistic interpretation of boundary value problems

Let

$$Lu \equiv \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial u}{\partial x_i} + c(x)u$$

with coefficients defined in the closure $\bar{D}$ of a bounded domain $D \subset R^n$, and assume that

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \mu|\xi|^2 \quad (x \in D, \xi \in R^n, \mu > 0),$$

$$(1) \qquad a_{ij}, b_i \text{ uniformly Lipschitz continuous in } \bar{D},$$
$$c \le 0, \; c \text{ uniformly Hölder continuous in } \bar{D}.$$

Let $f, \varphi$ be functions defined on $D$ and $\partial D$ respectively, satisfying:

$$(2) \qquad \begin{aligned} &f \text{ is uniformly Hölder continuous in } \bar{D}, \\ &\varphi \text{ is continuous on } \partial D. \end{aligned}$$

Assume finally that $\partial D$ is in $C^2$.

Consider the Dirichlet problem

$$(3) \qquad \begin{aligned} Lu &= f \text{ in } D, \\ u &= \varphi \text{ on } \partial D. \end{aligned}$$

It is well known that there exists a unique classical solution to this problem; $u$ is continuous in $\bar{D}$ and its second derivatives are continuous (in fact, Hölder continuous) in $D$.

Since the matrix $a(x) = (a_{ij}(x))$ is positive definite and uniformly Lipschitz continuous in $\bar{D}$, there exists a square matrix $\sigma(x) = (\sigma_{ij}(x))$ which is symmetric, positive definite and uniformly Lipschitz continuous in $\bar{D}$ such that $a(x) = \sigma^2(x)$. We extend $\sigma(x)$ into $R^n$ so that it remains uniformly Lipschitz continuous; $b(x) = (b_1(x), \ldots, b_n(x))$ is extended similarly into $R^n$.

Consider the system of stochastic differential equations

$$(4) \qquad d\xi(t) = \sigma(\xi(t))dw(t) + b(\xi(t))dt.$$

<u>Theorem 1</u>. Denote by $\tau$ the exit time of $\xi(t)$ from $D$. Then $E_x\tau < \infty \ \forall\ x\in D$ and the solution of (3) can be represented in the form

$$
\begin{aligned}
u(x) &= E_x\varphi(\xi(\tau))e^{\int_0^\tau c(\xi(s))ds} \\
&\quad - E_x\int_0^\tau f(\xi(t))e^{\int_0^t c(\xi(s))ds} .
\end{aligned} \tag{5}
$$

<u>Proof</u>. Consider the function

$$h(x) = -Ae^{\lambda x_1} \text{ in } D.$$

We can choose $\lambda$ large and then $A$ large so that

$$\frac{1}{2}\Sigma a_{ij}h_{x_ix_j} + \Sigma b_ih_{x_i} \le -1.$$

By Ito's formula, for any $T < \infty$,

$$E_xh(\xi(\tau\wedge T))-h(x) \le E_x\int_0^{\tau\wedge T}(-1)dt = E_x(\tau\wedge T).$$

Since $|h(x)| \le K$ in $D$, $E_x(\tau \wedge T) \le 2K$. Taking $T \uparrow \infty$ we obtain

$$E_x\tau \le 2K < \infty \quad \forall\ x\in D. \tag{6}$$

To prove (5), denote by $V_\varepsilon$ ($\varepsilon > 0$) the closed $\varepsilon$-neighborhood of $\partial D$ and let $D_\varepsilon = D\backslash V_\varepsilon$. Let $v$ be a function in $C^2(R^n)$ which coincides with $u$ in $D_{\varepsilon/2}$. By Ito's formula and §3 (5)

$$d(v(\xi)(t))\exp[\int_0^t c(\xi(s))ds]) = Lv(\xi(t)\cdot\exp[\int_0^t c(\xi(s))ds] + A\,dw$$

for some $A \in M_w^2[0,T]$. Hence

$$E_x v(\xi(\tau_\varepsilon \wedge T))\exp[\int_0^{\tau_\varepsilon\wedge T} c(\xi(s))ds] - v(x)$$

$$= E_x \int_0^{\tau_\varepsilon \wedge T} Lv(\xi(t)\cdot\exp[\int_0^t c(\xi(s))ds]dt$$

for any $x \in D_\varepsilon$, where $\tau_\varepsilon$ is the hitting time of $V_\varepsilon$ and $T < \infty$. Noting that $v(\xi(t) = u(\xi(t))$ if $0 \le t \le \tau_\varepsilon \wedge T$ and taking $\varepsilon \to 0$, we get

$$u(x) = E_x u(\xi(\tau \wedge T))\exp[\int_0^{\tau\wedge T} c(\xi(x))ds]$$

$$-E_x \int_0^{\tau\wedge T} f(\xi(t))\exp[\int_0^t c(\xi(s))ds]dt.$$

Taking $T\uparrow\infty$ and using (6), the assertion (5) follows.

<u>Problems</u>. 1. Consider the case of one stochastic differential equation

$$d\xi(t) = \sigma(\xi(t))dw(t) + b(\xi(t))dt$$

where $\sigma(x), b(x)$ are uniformly Lipschitz and $\sigma(x) > 0$ for all $x$. The function

$$v(x) = \int_0^x \exp[-\int_0^z \frac{2b(u)}{\sigma^2(u)} du]dz$$

satisfies

$$\frac{1}{2}\sigma^2 v'' + bv' = 0.$$

Prove that if $v(-\infty) = -\infty$ then $P_x[\sup_{t>0} \xi(t) = \infty] = 1$. Similarly, if $v(\infty) = \infty$ then $P_x[\inf_{t>0} \xi(t) = -\infty] = 1$.

2. In the preceding problem, assume that $v(\infty) = \infty$, $v(-\infty) > -\infty$. Show that

$$P_x[\sup_{t>0} \xi(t) < \infty] = 1,$$

$$P_x[\lim_{t\to\infty} \xi(t) = -\infty] = 1.$$

[Hint. To prove the last part, denote by $\tau_y$ the first time $\xi(t) \le y$ and let $y < x$, $x_2 > y$. Then $P_x(\tau_y < \infty) = 1$ and by the strong Markov property

$$P_x[\sup_{t>0} \xi(t + \tau_y)] = E_x P_y[\sup_{t>0} \xi(t) \ge x_2] = \frac{v(y)-v(-\infty)}{v(x_2)-v(-\infty)}$$

Also the left hand side is $\ge P_x[\overline{\lim}\, \xi(t) \ge x_2]$. Take $y \to -\infty$ and then $x_2 \to -\infty$.]

Consider now the operator

$$Lu \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x,t) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t)\frac{\partial u}{\partial x_i} + c(x,t)u$$

and assume that for some cylinder $Q_T = D \times (0,T)$:

$\Sigma\, a_{ij}(x,t)\xi_i\xi_j \geq \mu|\xi|^2 \quad (x \in D,\ 0 \leq t \leq T,\ \xi \in R^n,\ \mu > 0)$,

$a_{ij}, b_i$ are uniformly Lipschitz continuous in $(x,t) \in \overline{Q}_T$,

$c$ is uniformly Hölder continuous in $(x,t) \in \overline{Q}_T$.

(7) $f$ is uniformly Hölder continuous in $(x,t) \in \overline{Q}_T$,

$g$ is continuous on $S_T \equiv \partial D \times [0,T]$,

$\varphi$ is continuous on $D_T = \{(x,T);\ x \in \overline{D}\}$ and $\varphi = g$ on $\partial D_T$,

$\partial D$ is in $C^2$.

Consider the initial boundary value problem

$$Lu + \frac{\partial u}{\partial t} = f \quad \text{in} \quad D \times [0,T),$$

$$u = \varphi \quad \text{on} \quad D_T, \tag{8}$$

$$u = g \quad \text{on} \quad S_T.$$

It is well known that this problem has a unique classical solution $u$.

As in the elliptic case we introduce the square root $\sigma(x,t)$ of $a(x,t) = (a_{ij}(x,t))$ and extend both $\sigma$ and $b$ as uniformly Lipschitz functions in $R^n \times [0,T]$. Introducing the system of stochastic differential equations

$$d\xi(t) = \sigma(\xi(t),t)dw(t) + b(\xi(t),t)dt \tag{9}$$

we can now state:

<u>Theorem 2</u>. The solution $u$ of (8) can be represented in the form

$$u(x,t) = E_{x,t}g(\xi(\tau),\tau)\exp[\int_t^\tau c(\xi(s),s)ds]I_{\tau<T}$$
$$+ E_{x,t}\varphi(\xi(T),T)\exp[\int_t^T c(\xi(s),s)ds]I_{\tau=T}$$
$$- E_{x,t}\int_t^\tau f(\xi(s),s)\exp[\int_t^s c(\xi(\lambda),\lambda)d\lambda]ds$$

where $\tau$ is the first time $\lambda\in[t,\tau)$ that $\xi(\lambda)$ leaves $D$; if no such $\lambda$ exists then we set $\tau = T$.

The proof is similar to the proof of Theorem 1; we apply here Ito's formula to

$$u(\xi(\lambda),\lambda)\exp[\int_t^\lambda c(\xi(s),s)ds].$$

Consider next the Cauchy problem

$$(8)\qquad \begin{aligned} Lu + \frac{\partial u}{\partial t} &= f \quad \text{in} \quad R^n \times [0,T), \\ u(x,T) &= \psi(x) \quad \text{in} \quad R^n. \end{aligned}$$

We assume that

(9) $\Sigma\, a_{ij}(x,t)\xi_i\xi_j \geq \mu|\xi|^2 \quad (x\in R^n,\ 0 \leq t \leq T,\ \xi\in R^n,\ \mu > 0),$

$a_{ij}, b_i$ are bounded and uniformly Lipschitz continuous in $R^n \times [0,T]$,

$c$ is bounded and uniformly Hölder continuous in $R^n \times [0,T]$,

$f(x,t)$ is uniformly Hölder continuous in compact subsets of $R^n \times [0,T]$ and $|f(x,t)| \leq C(1 + |x|^\alpha)$,

$\psi(x)$ is continuous in $R^n$ and $|\psi(x)| \leq C(1 + |x|^\alpha)$

where $C > 0$, $\alpha > 0$.

Under these conditions there exists a unique solution $u(x,t)$ of (8) satisfying

$$|u(x,t)| \leq C_0(1 + |x|^\beta) \quad (C_0 > 0, \beta > 0) \tag{10}$$

The first derivative $u_x$ is also bounded by the right hand side of (10) (with different constants) in every set $R^n \times [0,T-\varepsilon]$.

We can now represent $u(x,t)$ in terms of the solution $\xi(t)$ of (9):

<u>Theorem 3.</u> The solution $u(x,t)$ is given by

$$\begin{aligned} u(x,t) &= E_{x,t}\psi(\xi(T))\exp[\int_t^T c(\xi(s),s)ds] \\ &-E_{x,t}\int_t^T f(\xi(s),s)\exp[\int_t^s c(\xi(\lambda),\lambda)d\lambda]ds. \end{aligned} \tag{11}$$

The proof is left as an excercise.

Consider now the special case

$$c(x,t) \equiv 0 \tag{12}$$

and the Cauchy problem

$$
(13) \qquad \begin{aligned} Lu + \frac{\partial u}{\partial t} &= 0 \quad \text{in} \quad R^n \times [0,T], \\ u(x,T) &= \psi(x) \quad \text{in} \quad R^n. \end{aligned}
$$

The solution can be represented in terms of the fundamental solution $\Gamma(x,t;\ y,T)$ of the backward parabolic equation $L + \partial/\partial t$:

$$
(14) \qquad u(x,t) = \int_{R^n} \Gamma(x,t;\ y,T)\psi(y)dy.
$$

We recall that as a function of $(x,t)$,

$$
(L + \frac{\partial}{\partial t})\Gamma = 0.
$$

Also

$$
0 \le \Gamma(x,t;\ y,T) \le C(T-t)^{-\frac{n}{2}}\exp[-\frac{c|x-y|^2}{T-t}]
$$

for some $C > 0$, $c > 0$.

From (11) we get

$$
u(x,t) = E_{x,t}\psi(\xi(T)) = \int \psi(y)P_{x,t}(\xi(T)\in dy).
$$

Since $\psi$ is arbitrary, it follows that

$$
(15) \qquad P_{x,t}(\xi(T)\in dy) = \Gamma(x,t;\ y,T)dy,
$$

that is, <u>the transition probability function</u>, <u>considered as a measure</u> $A \to p(t,x,T,A)$, <u>has density which is the fundamental solution</u> $\Gamma(x,t;\ y,T)$ <u>of</u> $L + \partial/\partial t$.

We shall use later on the $L^p$ elliptic estimate:

$$|u|_{W^{2,p}(G)} \leq C[|Lu|_{L^p(G)} + |u|_{L^p(G)}]; \tag{16}$$

here $G$ is a bounded domain with $C^2$ boundary, the coefficients of $L$ are continuous in $\bar{G}$ and $L$ is elliptic, and $u$ is any function in $W^{2,p}(G) \cap W_0^{1,p}(G)$, $1 < p < \infty$. Recall that $W^{m,p}(G)$ is the class of functions whose first $n$ derivatives belong to $L^p$,

$$|u|_{W^{m,p}(G)} = [\sum_{|\alpha| \leq m} \int_G |D^\alpha u|^p dx]^{1/p},$$

and $W_0^{1,p}(G)$ is the completion in $W^{1,p}(G)$ of the set of $C^\infty$ functions with support in $G$. If $c(x) \leq 0$ then the term $|u|_{L^p}$ on the right hand side of (16) may be dropped out.

Let $u$ satisfy

$$Lu = f \quad \text{in} \quad G,$$
$$u = 0 \quad \text{on} \quad \partial G.$$

If $c(x) \equiv 0$ then

$$u(x) = -E_x \int_0^\tau f(\xi(t))dt, \ \tau \text{ exit time from } G,$$

so that, by the $L^p$ estimate,

$$|E_x \int_0^\tau f(\xi(t))dt|_{W^{2,p}(G)} \leq C|f|_{L^p(G)} \qquad (1 < p < \infty). \tag{17}$$

Krylov [20] has considered the much more general process

$$d\xi(t) = \sigma(t)dw(t) + b(t)dt, \ \xi(0) = 0,$$

with nonanticipative $\sigma(t), b(t)$ and proved the following estimate:

Assume that

$$|\sigma^2(t)| \leq M, \ |b(t)| \leq M|\det \sigma^2(t)|^{1/n}$$

and let $G$ be any open bounded domain with diameter $\leq D$. Then, for any $x \in G$, $f \in L^n(G)$,

$$E_x \int_0^{\tau_x} f(\xi(t) + x)| \ |\det \sigma^2(t)|^{1/n} dt \leq N|f|_{L^n(G)} \tag{18}$$

where $\tau_x$ = exit time of $\xi(t) + x$ from $G$ and $N$ is a constant depending only on $M, D$.

## 6. Stopping time problems and variational inequalities

Consider a stochastic differential system in $R^n$

$$d\xi(t) = b(\xi(t))dt + \sigma(\xi(t))dw(t) \tag{1}$$

with the usual Lipschitz condition on $b(x), \sigma(x)$, and let $G$ be a bounded domain with $C^2$ boundary. Denote by $t_G$ the exit time from $G$, and introduce the cost functional

$$\begin{aligned} J_x(\tau) = E_x[\int_0^{\tau \wedge t_G} & e^{-\alpha t} f(\xi(t))dt + e^{-\alpha\tau\wedge t_G}\varphi(\xi(\tau \wedge t_G)I_{\tau < t_G} \\ & + e^{-\alpha\tau\wedge t_G} h(\xi(\tau \wedge t_G))I_{\tau \geq t_G}] \end{aligned} \tag{2}$$

for any stopping time $\tau$ with respect to the standard $\sigma$-field $\mathcal{F}_t$ associated with the Brownian motion. Here $f(x)$, $\varphi(x)$, $h(x)$ are given functions and $\alpha$ is a given positive number (the <u>discount factor</u>).

We consider the problem of finding

$$V(x) = \inf_{\tau \in \mathfrak{A}} J_x(\tau)$$

where $\mathfrak{A}$ varies over the set of all stopping times, and finding a stopping time $\tau^*$ such that

$$V(x) = J_x(\tau^*).$$

We refer to this problem as a <u>stopping time problem</u>; $\tau^*$ will be called an <u>optimal stopping time</u>.

Let $a = \sigma\sigma^*$ and set

$$Lu = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i} - \alpha u.$$

Consider the problem: find a function $u$ satisfying:

(3)
$$\begin{aligned} &Lu + f \geq 0 \text{ a.e. in } G, \\ &u \leq \varphi \text{ in } G \\ &(Lu + f)(u - \varphi) = 0 \text{ a.e. in } G, \\ &u = h \text{ on } \partial G. \end{aligned}$$

This problem is called a <u>variational inequality</u>. If $L$ is formally selfadjoint, then $u$ is the function $v$ which minimizes

$$\frac{1}{2}(Lv,v)_{L^2(G)} + (f,v)_{L^2(G)}$$

over the functions $v$ which vary in the convex set: $v \leq \varphi$, $v = h$ on $\partial G$.

We shall now assume:

$$\varphi \text{ is in } C^2(\overline{G}), \tag{4}$$

$$f \text{ is in } C(\overline{G}), \tag{5}$$

$$h \text{ is in } C^2(\partial G) \text{ and } h \leq \varphi. \tag{6}$$

Theorem 1. Under the foregoing assumptions, there exists a unique solution $u$ of the variational inequality (3) such that

$$u \in W^{2,p}(G) \text{ for any } 1 < p < \infty. \tag{7}$$

Proof. Let $\beta_\varepsilon(t)$ be $C^\infty$ function in $t$, for any $\varepsilon > 0$, such that

$$\beta_\varepsilon(t) = 0 \text{ if } t < 0,$$

$$\beta_\varepsilon(t) \to \infty \text{ if } t > 0,\ \varepsilon \to 0,$$

$$\beta_\varepsilon'(t) \geq 0 \text{ if } t > 0,$$

and consider the Dirichlet problem

$$-Lu + \beta_\varepsilon(u-\varphi) = f \text{ in } G, \tag{8}$$

$$u = h \text{ on } \partial G.$$

By the standard theory, a unique solution exists. We now estimate the maximum of $\beta_\varepsilon(u-\varphi)$ in $\overline{G}$ by noting that if the maximum is attained at a point $x^0 \in \partial G$ then $\beta_\varepsilon(u-\varphi) = 0$, whereas if it is attained at a point $x^0 \in G$ then $u-\varphi$ also attains its maximum at $x^0$ so that $-L(u-\varphi) \geq 0$ at $x^0$. We thus find that

$$0 \leq \beta_\varepsilon(u-\varphi) \leq C.$$

We can now use the $L^p$ estimates to deduce that

$$|u|_{W^{2,p}(G)} \leq C.$$

Taking a subsequence of $u_\varepsilon$, which is weakly convergent to some $u$ in $W^{2,p}(G)$ and strongly convergent in $W^{1,p}(G)$, we easily find that $u$ solves the variational inequality.

The uniqueness follows from the following theorem which connects the variational inequality problem to the stopping time problem.

Theorem 2. Any solution $u$ of (3) which satisfies (7) is given by

$$u(x) = \inf_{\tau \in \mathfrak{A}} J_x(\tau).$$

Further,

$$u(x) = J_x(\tau^*) \tag{10}$$

where $\tau^* = \hat{\tau} \wedge t_G$, and $\hat{\tau}$ is the hitting time of the set $S = \{x \in G;\ u(x) = \varphi(x)\}$.

The set $S$ is called the <u>stopping set</u> and the set $C = \{x \in G;\ u(x) < \varphi(x)\}$ is called the <u>continuation set</u>. The relation (10) means that the optimal stopping procedure is to continue while $\xi(t)$ is in $C$ and to stop as soon as $\xi(t)$ hits $S$.

<u>Proof</u>. By Ito's formula (cf. the proof of Theorem 1, §5).

$$(11) \qquad E_x u(\xi(\hat{\tau}))e^{-\alpha\hat{\tau}} - u(x) = E_x \int_0^{\hat{\tau}} e^{-\alpha t} Lu(\xi(t))dt$$

where $\hat{\tau} = \tau \wedge \tau_\varepsilon$, $\tau_\varepsilon$ = exit time from $G_\varepsilon = G \backslash V_\varepsilon$, and $V_\varepsilon$ is an $\varepsilon$-neighborhood of $\partial G$. Actually, for Ito's formula (11) to hold one usually requires that $u \in C^2(G_\varepsilon)$. However this formula holds also if $u$ is just assumed to belong to $W^{2,p}(G_\varepsilon)$ with $p > 1 + n/2$; see [7] [15]. Using the inequalities $Lu \geq -f, u \leq \varphi$ and then taking $\varepsilon \to 0$, we obtain $u(x) \leq J_x(\tau)$. Taking in the preceding proof $\tau = \tau^*$ and noting that

$$Lu(\xi(t)) = f(\xi(t) \quad \text{if} \quad t < \tau^*,$$

$$u(\xi(\tau^*)) = \varphi(\xi(\tau^*)) \quad \text{if} \quad \tau^* < t_G,$$

$$u(\xi(\tau^*)) = h(\xi(\tau^*)) \quad \text{if} \quad \tau^* = t_G,$$

we obtain (10).

It is actually not surprising that the optimal stopping problem leads to the variational inequality. Indeed, arguing formally we have two choices at each initial position (x,0) with

$x \in G$:

(i) either stop, which implies that $V(x) \leq \varphi(x)$;

(ii) or continue for a time $\sigma$ and then proceed optimally, which implies

$$V(x) \leq E_x\Big[\int_0^\sigma e^{-\alpha t} f(\xi(t))dt + e^{-\alpha\sigma} V(\xi(\sigma))\Big];$$

the second summand on the right is obtained after applying the strong Markov property. Using Ito's formula and then dividing by $\sigma$ and taking $\sigma \to 0$, we obtain $LV + f \geq 0$.

Finally, since either (i) or (ii) is optimal, we must have $(V-\varphi)(LV + f) = 0$.

The above procedure of deriving formally differential inequalities for the optimum can be applied to a large variety of Markov optimization problems.

The system (8) is called the <u>penalized problem</u>. Consider the case where $\beta_\varepsilon(t) = t^+/\varepsilon$, so that the penalized problem becomes

$$-Lu + \frac{1}{\varepsilon}(u-\varphi)^+ = f \text{ in } G,$$

$$u = h \text{ on } \partial G.$$

Even though this $\beta_\varepsilon(t)$ is only Lipschitz in $t$, the previous proof still applies, so that $u_\varepsilon \to 0$ if $\varepsilon \to 0$. The solution $u_\varepsilon$ can be given a probabilistic interpretation, namely:

Denote by $V$ the class of all nonanticipative functions

$v(t)$ with $0 \leq v(t) \leq 1$. For any $v \in V$, define the cost functional

$$(13)\quad \bar{J}_x(v) = E_x\Big[\int_0^{t_G} [f(\xi(t)) + \frac{1}{\varepsilon}\varphi(\xi(t))v(t))] e^{-\int_0^t \frac{v(\lambda)}{\varepsilon}d\lambda} \; e^{-\alpha t}dt$$

$$+ h(\xi(t_G))e^{-\int_0^{t_G} \frac{v(\lambda)}{\varepsilon}d\lambda} e^{-\alpha t_G}\Big].$$

Then

$$(14)\qquad u_\varepsilon(x) = \inf_{v \in V} \bar{J}_x(v).$$

Problems. 1. Prove (14), by applying Ito's formula. Prove also that $u_\varepsilon(x) = \bar{J}_x(\bar{v})$ where $\bar{v}(t) = 1$ if $u_\varepsilon(\xi(t)) \geq \varphi(\xi(t))$ and $\bar{v}(t) = 0$ otherwise.

2. Prove that

$$u(x) \leq u_\varepsilon(x) \leq u(x) + C\varepsilon.$$

Consider now a functional which depends on two stopping times:

$$J_x(\sigma,\tau) = E_x\Big[\int_0^{\sigma\wedge\tau\wedge t_G} e^{-\alpha t} f(\xi(t))dt$$

$$(15)\qquad + e^{-\alpha\sigma}\varphi_1(\xi(\sigma))I_{\substack{\sigma<\tau\\ \sigma<t_G}} + e^{-\alpha\tau}\varphi_2(\xi(\tau))I_{\substack{\tau\leq\sigma\\ \tau\leq t_G}}$$

$$+ e^{-\alpha t_G}h(\xi(t_G))I_{t_G\leq\sigma\wedge\tau}\Big].$$

We call $J_x(\sigma,\tau)$ a payoff and we consider two players, the first one controls $\sigma$ and tries to minimize the payoff, and the second one controls $\tau$ and tries to maximize the payoff. This model is called a zero sum stochastic differential game.

A pair $(\sigma^*,\tau^*)$ of stopping times is called a saddle point if

$$J_x(\sigma^*,\tau) \leq J_x(\sigma^*,\tau^*) \leq J_x(\sigma,\tau^*)$$

for all $\sigma,\tau$. The number

$$V(x) = J_x(\sigma^*,\tau^*)$$

is called the value of the game.

The definition (15) is not symmetric in $\sigma,\tau$, since when $\sigma = \tau < t_G$ the function $\varphi_2$ (and not $\varphi_1$) is relevant; this however will not affect the results below (which will be symmetric in $\sigma,\tau$) provided $\varphi_2 \leq \varphi_1$. (Notice that $V(x) \geq \varphi_2(x)$ and if the results should be symmetric then $V(x) \leq \varphi_1(x)$, thus leading to the necessary condition $\varphi_2 \leq \varphi_1$).

We introduce the variational inequality with two constraints:

$$\begin{aligned} &Lu + f \geq 0 \text{ a.e. where } u > \varphi_2, \\ &Lu + f \leq 0 \text{ a.e. where } u < \varphi_1, \\ &\varphi_2 \leq u \leq \varphi_1 \text{ in } G, \\ &u = h \text{ on } \partial G. \end{aligned} \tag{16}$$

We assume the same regularity conditions on $L,f,h$ as before and, in addition,

$$\text{(17)}\qquad \begin{aligned} &\varphi_1, \varphi_2 \text{ belong to } C^2(\overline{G}), \\ &\varphi_2 \le \varphi_1 \text{ in } G,\ \varphi_2 \le h \le \varphi_1 \text{ on } \partial G. \end{aligned}$$

Theorem 3. There exists a solution $u$ of (16) which belongs to $W^{2,p}(G)$ for any $1 < p < \infty$. The solution is unique and coincides with $V(x)$. Further, the pair $(\sigma^*, \tau^*)$ where

$$\sigma^* = \text{hitting time of the set } \{u = \varphi_1\},$$
$$\tau^* = \text{hitting time of the set } \{u = \varphi_2\}$$

is a saddle point.

Problems. 3. Prove Theorem 3 by the method of proof of Theorems 1,2, introducing the penalized problem

$$-Lu + \beta_\varepsilon(u-\varphi_1) + \gamma_\varepsilon(u-\varphi_2) = f \quad \text{in} \quad G,$$

where $\gamma_\varepsilon(t) = 0$ if $t > 0$, $-\gamma_\varepsilon(t) \to \infty$ if $t < 0$, $\varepsilon \to 0$, $\gamma_\varepsilon(t) > 0$ if $t < 0$.

Theorems 1-3 can be generalized to unbounded domains $G$ and to time-dependent coefficients and data (the differential inequalities form a parabolic variational inequality). Also, instead of just controlling the stopping time, one may introduce nonanticipative controls into the stochastic differential equations [15]. There is also some work on non-zero sum stochastic differential games (see [2] [15]).

If in a variational inequality the constraint depends on the unknown solution, then we call this problem a quasi variational inequality. Such problems arise in non-zero stochastic differential games. Another model which gives rise to such a problem is when the control variable is a sequence of stopping times $\tau = (\tau_1, \tau_2, \ldots)$. We refer to [5] [6] for a model of this kind, where $\tau_1, \tau_2, \ldots$ are the time for ordering stock from the warehouse. Another model arising in quality control is studied in [1].

## §7. Stochastic switching and nonlinear elliptic equations

For any $\mu > 0$, we denote by $W^{m,p,\mu}(R^n)$ the class of functions $u$ such that

$$e^{-\mu(1+|x|^2)^{1/2}} u \in W^{2,p}(R^n)$$

Let

$$L^k u \equiv \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}^k(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i^k(x) \frac{\partial u}{\partial x_i} - \alpha u \qquad (\alpha > 0;\ k = 1,2,..)$$

be elliptic operators satisfying:

$$(1) \qquad \begin{aligned} &\sum a_{ij}^k(x)\xi_i\xi_j \geq \nu|\xi|^2 \qquad (x \in R^n,\ \xi \in R^n,\ \nu > 0), \\ &|D^\beta a_{ij}^k(x)| \leq C,\ |D^\beta b_i^k(x)| \leq C \qquad (0 \leq |\beta| \leq 2,\ x \in R^n,\ C > 0) \end{aligned}$$

and introduce the corresponding systems of stochastic differential equations

(2) $$d\xi^k(t) = \sigma^k(\xi^k(t))dw(t) + b^k(\xi^k(t))dt$$

where $\sigma$ is the positive square root of the matrix $(a^k_{ij})$.

Let $v(t)$ be any function with values in $\{1,2,3,\ldots\}$. We call $v$ a <u>control function</u> and denote by $V$ the set of all controls. To each $v \in V$ we define the trajectory $\xi^v(t)$ by

(3) $$d\xi^v(t) = \sigma^{v(t)}(\xi^v(t))dw(t) + b^{v(t)}(\xi^v(t))dt$$

with initial condition $\xi^v(0) = x$. Thus $\xi^v(t)$ coincides with $\xi^k(t)$ "as long as" $v(t) = k$. The construction of a continuous process $\xi^v(t)$ and its uniqueness can be proved by the successive approximation method of §4.

We now introduce a cost functional which depends on a sequence of given functions $f^k(x)$, for which

(4) $$|D^\beta f^k(x)| \le C \quad (|\beta| \le 2,\ x \in R^n,\ C > 0),$$

and on a discount factor $\alpha > 0$:

(5) $$J_x(v) = E_x \int_0^\infty e^{-\alpha t} f^{v(t)}(\xi^v(t))dt.$$

Consider the problem of finding

(6) $$V(x) = \inf_{v \in V} J_x(v).$$

This is a problem of optimizing the running cost $f = \{f^k\}$ when one is allowed to switch freely from one stochastic system to another.

Krylov [21] studied this problem. His main result is the following.

Theorem 1. If $\alpha$ is sufficiently large then

$$V \in W^{2,p,\mu}(R^n) \text{ for some } \mu > 0 \text{ and all } p < \infty, \tag{7}$$

$$\inf_k \{L^k V(x) + f^k(x)\} = 0 \text{ a.e. in } R^n, \tag{8}$$

and $V$ is uniquely determined by (7), (8).

Equation (8) is called the Bellman equation. Krylov's proof is probabilistic and does not extend to the corresponding problem in a domain $G$, $G \neq R^n$ (which will be defined in detail below); his proof uses, among other things, the inequality (18), §5.

Now let $G$ be a bounded domain with $C^2$ boundary $\partial G$, and define a cost functional

$$J_x(v) = E_x \int_0^T e^{-\alpha t} f^v(\xi^v(t))dt \tag{9}$$

where $T$ is the exit time from $G$; let $V(x)$ be again defined by (6). Consider the problem of characterizing $V(x)$ as the solution of the Dirichlet problem for the Bellman equation:

$$\begin{aligned} &\inf_k \{L^k u(x) + f^k(x)\} = 0 \text{ a.e. in } G, \\ &u = 0 \text{ on } \partial G. \end{aligned} \tag{10}$$

The following result is due to Evans and Friedman [12].

<u>Theorem 2.</u> Assume that the coefficients $a_{ij}^k$ are constants. Then, for any $\alpha > 0$, there exists a unique solution $u$ of (10) such that

$$u \in W_0^{1,\infty}(G) \cap W_{loc}^{2,\infty}(G),$$

and $u \equiv V$.

Before outlining the proof we introduce, as a motivation, another stochastic control problem corresponding to a finite number $m$ of the elliptic operators, say $L^1,\ldots,L^m$. The previous control variable $v(t)$ is now restricted to a countable number of switchings, and, furthermore, the switchings are cyclic, i.e., from state $i$ to state $i+1$ (where state $m+1$ is identified with state 1). Equivalently, we take the control variable to be a sequence of stopping times $\theta = (\theta_1,\theta_2,\ldots)$ with $\theta_j \uparrow \infty$.

To write down the cost $J_x^1(\theta)$, we fix positive numbers $k_1,\ldots,k_m$ and then define

$$\begin{aligned} J_x^k(\theta) = E_x\Big[&\int_0^{\theta_1} e^{-\alpha t} f^1(\xi^1(t))dt + k_1 e^{-\alpha\theta_1} \\ &+ \int_{\theta_1}^{\theta_2} e^{-\alpha t} f^2(\xi^2(t))dt + k_2 e^{-\alpha\theta_2} + \cdots \\ &+ \int_{\theta_{m-1}}^{\theta_m} e^{-\alpha t} f^m(\xi^m(t))dt + k_m e^{-\alpha\theta_m} \\ &+ \int_{\theta_m}^{\theta_{m+1}} e^{-\alpha t} f^1(\xi^1(t))dt + k_1 e^{-\alpha\theta_{m+1}} + \cdots \Big] \end{aligned}$$

Thus the switching from $\xi^i$ to $\xi^{i+1}$ incurs cost $k_i$. Set

$$V^1(x) = \inf_\theta J_x^1(\theta).$$

Similarly we can define a cost $J_x^i(\theta)$ starting with $\xi^i$ instead of $\xi^1$, i.e.,

$$J_x^i(\theta) = E_x[\int_0^{\theta_1} e^{-\alpha t} f^i(\xi^i(t))dt + k_i e^{-\alpha\theta_1}$$

$$+ \int_{\theta_1}^{\theta_2} e^{-\alpha t} f^{i+1}(\xi^{i+1}(t))dt + k_{i+1} e^{-\alpha\theta_2} + \cdots];$$

set

$$V^i(x) = \inf_\theta J_x^i(\theta).$$

Proceeding formally we arrive at the following system of variational inequalities for $u^i = V^i(x)$:

$$L^i u^i + f^i \geq 0,\ u^i \leq u^{i+1} + k_i,$$

$$(L^i u^i + f^i)(u^i - u^{i+1} - k_i) = 0 \quad \text{in} \quad G,$$

$$u^i = 0 \quad \text{on} \quad \partial G.$$

This system was studied in [3] [4] in case $m = 2$.

It is clear from the above model that if $k_i \to 0$ $(1 \leq i \leq m)$ then each $u^i(x)$ should converge to the same function $V(x)$, where $V(x)$ is defined in (6). Thus one is motivated to first solve

(11) and then take $k_i \to 0$.

In order to solve (11), we introduce a penalty term $\beta_\varepsilon(u^i-u^{i+1}-k_i)$ where $\beta_\varepsilon$ is defined as in §6 and then take $\varepsilon \to 0$. In this way one can show (even when the $a_{ij}^k$ are not constants) that there is a unique solution of (11) such that $u^i \in W^{2,p}(G)$ for any $p < \infty$. (One can also prove this result by more probabilistic methods based upon approximating the costs $J_x^i(\theta)$ by costs functionals in which a finite number of times $\theta_1 < \theta_2 < \cdots < \theta_N$ is used, and then let $N \to \infty$; see [12].)

Since we are mainly interested in solving (10), or first

$$\text{(12)}\qquad \begin{aligned} &\inf_{1\le k\le m} \{L^k u(x) + f^k(x)\} = 0 \text{ a.e. in } G,\\ &u = 0 \quad \text{on} \quad \partial G, \end{aligned}$$

it is technically simpler to work directly with the penalized problem

$$\text{(13)}\qquad \begin{aligned} &-L^i u^i + \beta_\varepsilon(u^i-u^{i+1}) - f^i = 0 \quad \text{in} \quad G,\\ &u^i = 0 \quad \text{on} \quad \partial G \quad (1 \le i \le m) \end{aligned}$$

and take $\varepsilon \to 0$, hoping to get the solution of (12) as $\lim_{\varepsilon\to 0} u^i(x)$.

The existence of a unique classical solution of (13) is rather standard. The next step, which is crucial, is to derive a priori estimates

$$\text{(14)}\qquad |Du^i|_{L^\infty(G)} \le C,$$

$$(15) \qquad |D^2u^i|_{L^\infty(G_0)} \le C \qquad (\bar{G}_0 \subset G)$$

where $C$ is independent of $\varepsilon$. (Details are omitted.) Using these estimates one proceeds to show that, as $\varepsilon \to 0$,

$$u^i(x) \to V_m(x) \qquad (1 \le i \le m,\ x \in G)$$

and $V_m(x)$ is a solution of (12). Next we take $m \to \infty$ and show that $V_m(x) \to V(x)$ where $V(x)$ is the asserted solution of (10). Uniqueness follows by the method of Krylov [21] (see also [7]).

As an immediate application of Theorem 2, consider the Dirichlet problem for a highly nonlinear elliptic equation

$$(16) \qquad \lambda u - F(u_{x_i x_j}) = f(x) \quad \text{in} \quad G,$$

$$(17) \qquad u = 0 \quad \text{on} \quad \partial G$$

where $\lambda > 0$. Assume:

$$F\colon R^{n^2} \to R \text{ is convex and } C^2,$$

$$\Sigma F_{x_{ij}} \xi_i \xi_j \ge \gamma |\xi|^2, \ \gamma > 0 \text{ (ellipticity)}$$

$$|F(x) - \Sigma F_{x_{ij}}(x) x_{ij}| \le C.$$

Then we can write (16) as a Bellman equation with

$$-L^\xi u = \lambda u - \Sigma F_{x_i x_j}(\xi) \frac{\partial^2 u}{\partial x_i \partial x_j},$$

$\xi = (\xi_{ij})$ with rational coordinates. Thus there exists a unique solution of (16), (17) in $W^{1,\infty}(G) \cap W^{2,\infty}_{loc}(G)$.

## §8. Probabilistic methods in singular perturbations

Consider the uniformly elliptic operator

$$L_\varepsilon u = \frac{\varepsilon}{2} \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial u}{\partial x_i} \qquad (\varepsilon > 0)$$

with Lipschitz continuous coefficients in a bounded domain $D$ with $C^2$ boundary, and set $Lu = L_1 u$. We shall be interested in the following problems.

Problem. 1. Denote by $u_\varepsilon$ the solution of the Dirichlet problem

$$\text{(1)} \qquad \begin{aligned} L_\varepsilon u &= 0 \quad \text{in } D, \\ u_\varepsilon &= \varphi \quad \text{on } \partial D. \end{aligned}$$

Find the behavior of $u_\varepsilon(x)$ as $\varepsilon \to 0$.

Problem 2. Denote by $\lambda_\varepsilon$, $\varphi_\varepsilon$ the principal eigenvalue and eigenfunction of

$$\text{(2)} \qquad \begin{aligned} L_\varepsilon w &= -\lambda w \quad \text{in } D, \\ w &= 0 \quad \text{on } \partial D. \end{aligned}$$

Find the behavior of $\lambda_\varepsilon$, $\varphi_\varepsilon$ as $\varepsilon \to 0$; here $\varphi_\varepsilon$ is normalized by

$$\int_D \varphi_\varepsilon^2 dx = 1, \ \varphi_\varepsilon > 0 \quad \text{in} \quad D.$$

The solution of (1) can be written in the form

(3) $$u_\varepsilon(x) = E_x \varphi(\xi^\varepsilon(\tau^\varepsilon))$$

where

$$d\xi^\varepsilon(t) = \varepsilon^{1/2}\sigma(\xi^\varepsilon(t))dw(t) + b(\xi^\varepsilon(t))dt,$$

$\sigma$ is the positive square root of $(a_{ij})$, and $\tau^\varepsilon$ is the exit time of $\xi^\varepsilon(t)$ from $D$. The behavior of $\tau^\varepsilon$ as $\varepsilon \to 0$ depends in a crucial way on the behavior of the solutions of the ordinary differential equations

(4) $$\frac{dx}{dt} = b(x), \ x(0) \in D.$$

Suppose all solutions of (4) leave $\bar{D}$ in finite time $\tau_x^0$ (depending on the initial point $x(0) = x$). It is easily shown that, for any $T < \infty$,

(5) $$\sup_{0 \le t \le T} |\xi_x^\varepsilon(t) - \xi_x^0(t)| \xrightarrow{P} 0 \quad \text{as} \quad \varepsilon \to 0$$

where $\xi_x^\varepsilon(t)$ $(\varepsilon \ge 0)$ is the solution $\xi^\varepsilon(t)$ with $\xi_x^\varepsilon(0) = x$. It follows that

$$u_\varepsilon(x) \to u_0(x) \equiv \varphi(\xi_x^0(\tau_x^0)) \quad \text{as} \quad \varepsilon \to 0$$

Consider now the extreme case where

$$b(x)\cdot\nu(x) < 0 \qquad (x \in \partial D) \tag{7}$$

where $\nu$ is the outward normal. This condition implies the solutions of (4) cannot reach $\partial D$ in any time $t > 0$. Thus $\tau_x^0 = \infty$.

We shall assume:

(A) There is a point $x^0 \in D$ such that every solution of (4) enters a given neighborhood of $x^0$ in finite time. Further, $x^0$ is a stable equilibrium point for (4) in the sense that $b(x^0) = 0$ and the Jacobian matrix of $a^{-1}b$ at $x^0$ has all negative eigenvalues.

(B) There exists a function $\psi(x)$ in $\bar{D}$ such that

$$b = \frac{1}{2} a \nabla \psi .$$

We shall consider the Dirichlet problem

$$\begin{aligned} L_\varepsilon u + \varepsilon \sum_{i=1}^{n} b_i^1 \frac{\partial u}{\partial x_i} &= 0 \quad \text{in} \quad D, \\ u &= 0 \quad \text{on} \quad \partial D \end{aligned} \tag{8}$$

where

$$b_j^1 = \frac{1}{2} \sum_{k=1}^{n} \frac{\partial a_{jk}}{\partial x_k} + \frac{1}{2} \sum_{k=1}^{n} a_{jk} \frac{\partial \psi^1}{\partial x_k} \quad \text{for some} \quad \psi^1;$$

Notice that if $a_{jk} = \text{const.}$ then we can take $\psi^1 = 0$, $b_j^1 = 0$ so that (8) reduces to (1).

Set

$$(9) \qquad C = \lim_{\varepsilon\to 0} \frac{\int_{\partial D} e^{\psi/\varepsilon} e^{\psi^1} (b\cdot\nu)\varphi \, dS}{\int_{\partial D} e^{\psi/\varepsilon} e^{\psi^1} (b\cdot\nu) dS}$$

if this limit exists.

Theorem 1. Let (7), (A),(B) hold. If $C$ exists then the solution $u_\varepsilon$ of (8) satisfies: $u_\varepsilon(x) \to C$ uniformly in bounded subsets of $D$.

For the Dirichlet problem

$$(10) \qquad \begin{aligned} &\frac{\varepsilon}{2} \sum_{j,k=1}^{n} \frac{\partial}{\partial x_j}\left(a_{jk} \frac{\partial u}{\partial x_k}\right) + \sum_{j=1}^{n} b_j \frac{\partial u}{\partial x_j} = 0 \quad \text{in} \quad D, \\ &u = 0 \quad \text{on} \quad \partial D \end{aligned}$$

we can establish a similar result (if (A),(B) hold) with

$$(11) \qquad C = \lim_{\varepsilon\to 0} \frac{\int_{\partial D} e^{\psi/\varepsilon} (b\cdot\nu)\varphi \, dS}{\int_{\partial D} e^{\psi/\varepsilon} (b\cdot\nu) dS} .$$

These formulas were discovered heuristically by Matkowsky and Schuss [23] and proved under various restrictions by Kamin [19]. The proof in the general case is due to Devinatz and Friedman [9] and exploits both probabilistic considerations and elliptic estimates.

Theorem 1 requires self-adjointness of the elliptic operator (with respect to the measure $e^{\tilde{\psi}/\varepsilon}$, $\tilde{\psi} = \psi + \varepsilon\psi^1$). Consider now the problem (1), without the condition (B), but assuming (7) and (A). Introduce the functional

$$I_T(\ ) = \int_0^T (\|\frac{d\zeta(t)}{dt} - b(\zeta(t))\|_{a^{-1}(\ (t)})^2 dt$$

if $\zeta(t)$ is absolutely continuous ($I_T(\zeta) = \infty$ if $\zeta$ is not absolutely continuous) and

$$\|X\|_{a^{-1}(x)} = (\Sigma\, a_{ij}^{-1}(x)X_iX_j)^{1/2}, \quad (a_{ij}^{-1}(x)) = \text{inverse of } (a_{ij}(x)).$$

Let $\zeta(t)$ $(0 \leq t \leq T)$ vary over all continuous curves such that $\zeta(0) = x^0$, $\zeta(T) = y$, $\zeta(t) \in D$ if $0 < t < T$, where $y$ is a fixed point on $\partial D$. Let

$$V(y) = \inf I_T(\zeta)$$

over all such $\zeta$, for all $T < \infty$.

$V(y)$ is called a quasi potential. It measures in some sense the amount of work required to move a particle from $x^0$ to $y$ against the dynamical system (4). It is easy to verify that $V(y)$ is Lipschitz continuous. It can also be shown that, under the conditions of Theorem 1, $V(y) = 4\psi(y)$ at the points where $V$ takes its maximum. Denote this set of points by $\Sigma$.

Theorem 2. If $\varphi \equiv \text{const.} = C$ on $\Sigma$, then (under the assumptions (7),(A)) $u_\varepsilon(x) \to C$ uniformly on compact subsets of $D$.

This result is due to Ventcel and Freidlin [25]. Their proof is based on the following asymptotic estimates for any open set $G$ and any closed set $H$ is the space $\mathcal{C}_T$:

$$\text{(12)} \qquad \underline{\lim}_{\epsilon\to 0}[2\epsilon \log P_x^\epsilon(G)] \geq -\inf_{\omega\in G_x} I_T(\omega),$$

$$\text{(13)} \qquad \overline{\lim}_{\epsilon\to 0}[2\epsilon \log P_x^\epsilon(H)] \leq -\inf_{\omega\in H_x} I_T(\omega),$$

where $P_x^\epsilon$ is $P_x$ induced by $\xi^\epsilon(t)$ and

$$G_x = \{\omega\in G;\ \omega(0) = x\},\ H_x = \{\omega\in H;\ \omega(0) = x\}.$$

For proofs see also [15].

It remains an open problem to determine whether $\lim u_\epsilon$ exists when the only assumptions are (7) and (A).

We now consider Problem 2.

Lemma 3. Define

$$\Lambda = \sup\{\lambda \geq 0,\ \sup_{x\in D} E_x e^{\lambda\tau} < \infty\}$$

where $\tau$ = exit time from $D$. Then $\Lambda = \lambda_0$ where $\lambda_0$ is the principal eigenvalue of $L$.

For proof, see [15].

Theorem 4. Let (7) and (A) hold. Then the principal eigenvalue $\lambda_\epsilon$ satisfies

$$\text{(14)} \qquad -2\epsilon \log \lambda_\epsilon \to V^* \qquad (V^* = \min_{y\in\partial D} V(y))$$

Theorem 5. If in addition to (7), (A), $a_{ij} \equiv \delta_{ij}$ and $b = \nabla\psi$, then the principal eigenfunction $\varphi_\varepsilon$ satisfies

$$\varphi_\varepsilon(x) \to \text{const.} = C \qquad \left(\int_D C^2 dx = 1\right) \tag{15}$$

uniformly in compact subsets of $D$ and boundedly in $D$.

The proof of Theorem 4 (which was originally asserted by Ventcel [24]) is proved in Friedman [15]. The proof of Theorem 5 is due to Devinatz and Friedman [8]. The proofs use the Ventcel-Freidlin estimates (12), (13) and (in the case of Theorem 5) some elliptic estimates.

Theorem 5 holds for general $a_{ij}$ provided $a^{-1}b = \nabla\psi$ in a neighborhood of $x^0$. It remains an open question to prove the theorem without this restriction.

Other results are known on the behavior of $\lambda_\varepsilon, \varphi_\varepsilon$ under different type of conditions on $b(x)$. For instance, if $b(x)$ has zero of order $\nu$ at $x^0$ and all solutions of (4) with $x(0) \neq x^0$ leave $\bar{D}$ in finite time then

$$\lambda_\varepsilon = O(\varepsilon^{(\nu-1)/(\nu+1)}),$$

$$u_\varepsilon^2 \to \text{Dirac delta function supported at } x^0.$$

For proof see [8] and the references given there.

We finally mention that the Ventcel-Freidlin estimates have been used to obtain the precise asymptotic behavior of other quantities; for instance, the Green function $q_\varepsilon(t,x,y)$ of

$L_\varepsilon - \partial/\partial t$; see [15].

Problems. 1. Let $D \subset D^*$ and denote by $\lambda_0, \lambda_0^*$ the principal eigenvalues corresponding to $L$ in $D$ and $D^*$ respectively. Prove that $\lambda_0 \geq \lambda_0^*$.

2. If $\tau_x^0 = \infty$ for some $x \in D$ then $\lambda_\varepsilon \to 0$ if $\varepsilon \to 0$.

3. Use the Ventcel-Freidlin estimates to show that for any $\delta > 0$

$$P_x^\varepsilon[\sup_{0<t<T} |\xi^\varepsilon(t) - \xi^0(t)| > \delta] \leq e^{-\frac{c}{\varepsilon}}$$

for all $\varepsilon$ small, where $c$, is a positive constant.

4. Let $u_\varepsilon$ satisfy

$$Lu_\varepsilon - \frac{\partial u_\varepsilon}{\partial t} = 0 \quad \text{in} \quad D \times (0,T),$$

$$u_\varepsilon(x,0) = 0 \quad \text{if} \quad x \in D,$$

$$u_\varepsilon(x,t) = 1 \quad \text{if} \quad x \in \partial D,\ 0 < t < T.$$

Use the Ventcel-Freidlin estimates to prove that

$$\lim_{\varepsilon \to 0} [2\varepsilon \log u_\varepsilon(x,t)] = -I(t,x,\partial D)$$

where $I(t,x,\partial D) = \inf I_t(\varphi)$, $\varphi$ varying over all functions in $T$ satisfying: $\varphi(0) = x$, $\min_{0 \leq s \leq t} \text{dist.}(\varphi(s), \partial D) = 0$.

## References


[1] R. F. Anderson and A. Friedman, Multi-dimensional quality control problem, Trans. Amer. Math. Soc., to appear.

[2] A. Bensoussan and A. Friedman, Non-zero sum stochastic differential games with stopping times and free boundary problems, Trans. Amer. Math. Soc., 231 (1977), 275-327.

[3] A. Bensoussan and A. Friedman, On the support of the solution of a system of quasi variational inequalities, J. Math. Anal. and Appl., to appear.

[4] A. Bensoussan and J. L. Lions, Contrôle impulsionnel et systèmes d'inequations quasi variationnelles, C. R. Acad. Sc. Paris, 278 (1974), 747-751.

[5] A. Bensoussan and J. L. Lions, Nouvelles methodes en contrôle impulsionnel, Appl. Math. and Optimization, 1 (1975), 289-312.

[6] A. Bensoussan and J. L. Lions, Temps d'arrêt et contrôle impulsionnel: Inéquations variationnelles et quasi variationnelles d'evolution, Univ. Paris IX, Cahier de Math. de la Décision, 1975, no. 7523.

[7] A. Bensoussan and J. L. Lions, Temps d'arrêt optimal, Dunod, 1978.

[8] A. Devinatz and A. Friedman, Asymptotic behavior of the principal eigenfunction for singularly perturbed Dirichlet problem, Indiana Univ. Math. J., 27 (1978), 143-157.

[9] A. Devinatz and A. Friedman, The asymptotic behavior of a singularly perturbed Dirichlet problem, Indiana Univ. Math. J., to appear.

[10] J. L. Doob, Stochastic Processes, Wiley, New York, 1967.

[11] E. B. Dynkin, Markov Processes, Vols. 1,2, Springer-Verlag, Berlin, 1965.

[12] C. L. Evans and A. Friedman, Optimal stochastic switching and the Dirichlet problem for the Bellman equation, to appear.

[13] A. Friedman, Partial Differential Equations of Parabolic Type, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

[14] A. Friedman, Stochastic Differential Equations and Applications, Vol. 1, Academic Press, New York, 1975.

[15] A. Friedman, Stochastic Differential Equations and Applications, Vol. 2, Academic Press, New York, 1976.

[16] I. I. Gikhman and A. V. Skorohod, The Theory of Stochastic Processes I, Springer-Verlag, Berlin, 1974.

[17] I. I. Gikhman and A. V. Skorohod, The Theory of Stochastic Processes II, Springer-Verlag, Berlin, 1975.

[18] I. I. Gikhman and A. V. Skorohod, The Theory of Stochastic Processes III, Springer-Verlag, Berlin, to appear.

[19] S. Kamin, On elliptic singular perturbation problems with turning points, SIAM J. Appl. Math., to appear.

[20] N. V. Krylov, An inequality in the theory of stochastic integrals, Th. Prob. Appl., 16 (1971), 438-448.

[21] N. V. Krylov, Control of a solution of a stochastic integral equation, Th. Prob. Appl., 17 (1972), 114-130.

[22] H. P. McKean, Stochastic Integrals, Academic Press, New York, 1971.

[23] B. J. Matkowsky and Z. Schuss, On the exit problem for randomly perturbed dynamical systems, SIAM J. Appl. Math., 33 (1977), 365-382.

[24] A. D. Ventcel, On the asymptotic behavior of the greatest eigenvalue of a second order elliptic differential operator with a small parameter in the highest derivatives, Soviet Math. Sokl., 13 (1972), 13-17.

[25] A. D. Ventcel and M. I. Freidlin, On small random perturbations of dynamical systems, Russian Math. Surveys, 25 (1970), 1-56 [Uspehi Math. Nauk, 25 (1970), 3-55].

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# THEORY OF DIFFUSION PROCESSES

D. STROOCK - S.R.S. VARADHAN

Theory of Diffusion Processes

D. Stroock and S.R.S. Varadhan
University of Colorado and C.I.M.S., N.Y.U.

## Section I

Let $x(t)$ be a Markov process and assume that

(1.1) $\quad E[\varphi(t+h) - \varphi(x(t)) | x(s) \quad \text{for} \quad s \leq t] = hL_t\varphi(x(t)) + o(h)$

for $\varphi \in C_0^\infty(R^d)$. It is not difficult to check that $L_t$ must be a linear operator which is quasi-local (i.e., for each $x \in R^d$ and $\varepsilon > 0$ there is a constant $C_\varepsilon < \infty$ such that $|L_t\varphi(x)| \leq C_\varepsilon \|\varphi\|$ for all $\varphi \in C_0^\infty(R^d \setminus \overline{B(x,\varepsilon)})$. Here and throughout $\|\cdot\|$ denotes the uniform norm.) Moreover, $L_t$ must satisfy the weak maximum principle in that if $\varphi(x) = \max_{y \in R^d} \varphi(y)$ then certainly $L_t\varphi(x) \leq 0$. From these observations one can conclude that $L_t$ ought to be of the form

$$L_t\varphi(x) = 1/2 \sum_{i,j=1}^{d} a^{ij}(t,x) \frac{\partial^2\varphi}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b^i(t,x) \frac{\partial\varphi}{\partial x_i}$$

$$+ \int_{R^d \setminus \{0\}} \left(\varphi(x+y) - \varphi(x) - \frac{\langle y, \nabla\varphi(x)\rangle}{1+|y|^2}\right) M(t,x;dy)$$

where $((a^{ij}(t,x)))$ is an element of the class $\mathcal{S}_d$ of symmetric non-negative definite matrices and $M(t,x;\cdot)$ is a $\sigma$-finite non-negative measure on $R^d \setminus \{0\}$ such that $\int_{R^d \setminus \{0\}} \frac{|y|^2}{1+|y|^2} M(t,x;dy) < \infty$. The assertion that

$L_t$ must have this form is the analytic statement of the renowned Lévy - Khinchine decomposition theorem. In particular, it says that the process $x(t+h) - x(t)$, for small $h$, behaves like the independent increment process whose Gaussian part has covariance $a(t,x(t))$ and drift $b(t,x(t))$ and whose Poisson jump part has Lévy measure $M(t,x(t);\cdot)$. Since our attention in these lectures will be devoted to processes which are continuous with respect to $t$, we can and will assume from now on that the jump part of $L_t$ is absent so that

$$(1.2) \qquad L_t = 1/2 \sum_{i,j=1}^{d} a^{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(t,x) \frac{\partial}{\partial x_i}$$

The central theme of these lectures will be the investigation of what can be said when one tries to pursue the preceding line of reasoning in the opposite direction. That is, suppose that an $L_t$ of the form in (1.2) is given. Then there are two key questions which we wish to answer: i) is there a continuous process $x(\cdot)$ for which (1.1) obtains, and ii) is there at most one such process if one also specifies the initial data? Before these questions can be studied it is essential to give a precise formulation of what we mean by a stochastic process satisfying (1.1).

Since we are going to be restricting our attention to continuous processes, our basic sample space will be $\Omega = C([0,\infty),R^d)$ endowed with the topology of uniform convergence on compact intervals. As such $\Omega$ is a complete separable metric space and we will denote by $\mathcal{M}$ the Borel field over $\Omega$. Given $\omega \in \Omega$ and $t \geq 0$, we use $x(t,\omega)$ to denote the position of $\omega$ at time $t$. In this way $x(t)$ becomes an $R^d$ - valued random variable on $\Omega$ for each $t \geq 0$. Next define $\mathcal{M}_t = \sigma(x(s): 0 \leq s \leq t)$ for $t \geq 0$. Clearly $\mathcal{M}_t$ is a sub $\sigma$ - algebra of $\mathcal{M}$ for each $t \geq 0$. Moreover, one can easily check that $\mathcal{M} = \sigma(\bigcup_{t\geq 0} \mathcal{M}_t)$. From now on <u>a stochastic process</u>

satisfying (1.1) will be for us a probability measure $P$ on $(\Omega, \mathcal{M})$ such that

$$(1.1') \qquad \lim_{h \downarrow 0} \frac{1}{h} E^P[\varphi(x(t+h))) - \varphi(x(t))|\mathcal{M}_t] = L_t\varphi(x(t))$$

for all $t \geq 0$ and $\varphi \in C_0^\infty(R^d)$. Observe that in this formulation the paramount role is played by the measure $P$ whereas the paths $x(\cdot)$ are relegated to the position of artifacts. We next want to manipulate (1.1′) into a more convenient form. Let $0 \leq t_1 \leq t$ be given. Then

$$\begin{aligned}
\lim_{h \downarrow 0} \frac{1}{h} E^P[\varphi(x(t+h)) - \varphi(x(t))|\mathcal{M}_{t_1}] \\
&= \lim_{h \downarrow 0} E^P[\frac{1}{h} E^P[\varphi(x(t+h)) - \varphi(x(t))|\mathcal{M}_t] \,|\, \mathcal{M}_{t_1}] \\
&= E^P[L_t f(x(t))|\mathcal{M}_{t_1}] \,.
\end{aligned}$$

Hence for $0 \leq t_1 \leq t_2$ :

$$E^P[\varphi(x(t_2)) - \varphi(x(t_1))|\mathcal{M}_{t_1}] = E^P[\int_{t_1}^{t_2} L_t\varphi(x(t))dt|\mathcal{M}_{t_1}]$$

or equivalently:

$$(1.4) \qquad E^P[\varphi(x(t_2)) - \int_0^{t_2} L_t\varphi(x(t))dt|\mathcal{M}_{t_1}] = \varphi(x(t_1)) - \int_0^{t_1} L_t\varphi(x(t))dt \,.$$

That is, the quantity

$$(1.5) \qquad X_\varphi(t) \equiv \varphi(x(t)) - \int_0^t L_s\varphi(x(s))ds$$

is conditionally constant under $P$. This version of (1.1′) is quite pleasing on both intuitive as well as technical grounds. Indeed, if the second order part of $L_t$ is absent and $L_t = \sum_{i=1}^d b^i(t,x) \frac{\partial}{\partial x_i}$, then we should expect the process associated with $L_t$ to be concentrated on integral curve of $\sum_{i=1}^{d} b^i(t,x) \frac{\partial}{\partial x_i}$ , in which case $X_\varphi(t)$ would be actually (not just conditionally) constant.

Processes which are conditionally constant play such an important role in probability theory that they have been given a special name: they are called <u>martingales</u>. With this terminology we can now formulate our problem in its final form. Given $L_t$ as in (1.2), we will say the probability measure $P$ on $(\Omega, \mathcal{M})$ <u>solves the martingale problem for</u> $L_t$ <u>starting from</u> $(s,x)$ <u>if</u>:

$$(1.6) \qquad \begin{cases} \text{a)} \ P(x(t) = x \ , \ 0 \leq t \leq s) = 1 \\ \\ \text{b)} \ X_\varphi(t \vee s) \ \text{is a } P\text{-martingale for all } \varphi \in C_0^\infty(R^d) \end{cases}$$

where $X_\varphi$, is defined by (1.5). We propose to study the following questions:

i) Existence: for each $(s,x)$ is there a solution $P$ to the martingale problem for $L_t$ starting from $(s,x)$ ?

ii) Uniqueness: for each $(s,x)$ is there at most one such $P$ ?

In addition, we will be interested in finding out what conclusions can be drawn from affirmative answers to i) and ii).

In order to carry out this program, we are going to require various preliminaries of a more or less standard nature. These fall quite naturally into two categories: the general theory of probability measures on $(\Omega, \mathcal{M})$, and the theory of martingales. The rest of this lecture will be devoted to the first of these topics.

Let $M(\Omega)$ stand for the set of all probability measures on $(\Omega, \mathcal{M})$ . The topology on $M(\Omega)$ with which we will be concerned is the so called <u>weak topology</u>: the smallest topology with respect to which $P \to E^P[F]$ is continuous for all $F \in C_b(\Omega)$. It is possible to find a metric on $M(\Omega)$ so that $M(\Omega)$ with the weak topology becomes a complete separable metric space. More important for our purposes is that we can characterize compact subsets of $M(\Omega)$. In fact, by Prokharov's theorem, $\Gamma \subseteq M(\Omega)$ is pre-compact if and only if for each $\varepsilon > 0$ there is a compact $K_\varepsilon \subseteq \Omega$ such that $\inf_{P \in \Gamma} P(K_\varepsilon) \geq 1 - \varepsilon$. Since the compact subsets of $\Omega$ are characterized by the Azela-Ascoli theorem, we now can say the $\Gamma \subseteq M(\Omega)$ is pre-compact if and only if

$$(1.7)\quad \begin{cases} \lim\limits_{A \uparrow \infty} \inf\limits_{P \in \Gamma} P(|x(0)| \leq A) = 1 \\ \lim\limits_{\delta \downarrow 0} \inf\limits_{P \in \Gamma} P(\sup\limits_{\substack{0 \leq s < t \leq T \\ t - s < \delta}} |x(t) - x(s)| \leq \rho) = 1 \quad , \quad T \geq 0 \text{ and } \rho > 0 \end{cases}$$

Since the second part of the condition (1.7) is in practice difficult to check, it is well to have more easily verified sufficient conditions. One such criterion is that of Kolmogorov: let $\Gamma \subseteq M(\Omega)$ satisfy

$$(1.8)\qquad \sup_{P \in \Gamma} E^P[|X(0)|] < \infty$$

and suppose that for each $T > 0$ there is a $C_T < \infty$ such that

$$(1.9)\qquad \sup_{P \in \Gamma} E^P[|x(t_2) - x(t_1)|^4 \leq C_T(t_2 - t_1)^2 \quad , \quad 0 \leq t_1 < t_2 \leq T \ .$$

Then $\Gamma$ satisfies (1.7) and is therefore pre-compact. As an example of how Kolmogorov's criterion can be applied, let $X_1, \ldots, X_n, \ldots$ be independent $R^d$-valued standard normal random variables (i.e., $P(X_n \in A) = \frac{1}{(2\pi)^{d/2}} \int_A e^{-|x|^2/2} dx)$ and define

$$S_n(t) = \frac{1}{n^{1/2}} \sum_1^{[nt]} X_k + n^{1/2}(t - \frac{[nt]}{n}) X_{[nt]+1} \ , \ t \geq 0 \ .$$

where $[x]$ is the largest integer in $x$. Check that

$$E[|S_n(t_2) - S_n(t_1)|^4] \le C(t_2 - t_1)^2 \ , \ 0 \le t_1 < t_2 \ ,$$

where $C < \infty$ is independent of $n \ge 1$. Now let $P_n$ on $(\Omega, \mathcal{M})$ denote the distribution of $S_n(\cdot)$ (clearly $t \to S_n(t)$ is continuous). Then $P_n(x(0) = 0) = 1$ and $\sup_{n \ge 1} E^{P_n}[|x(t_2) - x(t_1)|^4] \le C(t_2 - t_1)^2$, $0 \le t_1 < t_2$. Hence $\{P_n: n \ge 1\}$ is pre-compact. Let $P$ denote any limit point of $\{P_n: n \ge 1\}$ and check for an $N \ge 1$, $0 < t_1 < \ldots < t_N$, and $\varphi_1, \ldots, \varphi_N \in C_b(R^d)$:

$$E^P[\varphi(x(t_1)) \ldots \varphi(x(t_N))]$$

$$= \int \ldots \int \varphi_1(y_1) g(t_1, y_1) \varphi_2(y_2) g(t_2 - t_1, y_2 - y_1) \ldots \varphi_N(y_N) g(t_N - t_{N-1}, y_N - y_{N-1}) \cdot dy_1 \ldots dy_N$$

where $g(t,y) = \dfrac{1}{(2\pi t)^{d/2}} e^{-|y|^2/2t}$. This shows that all limit points of $\{P_n: n \ge 1\}$ coincide and therefore that $\lim_{n \to \infty} P_n$ exists. It also proves the existence of a $P$ on $(\Omega, \mathcal{M})$ satisfying (1.10). The measure satisfying (1.10) is the famous one constructed by Wiener and will be denoted by $\mathcal{W}$. We will later see that $\mathcal{W}$ is the one and only solution to the martingale problem for $1/2\Delta$ starting from $(0, \mathcal{O})$.

Another important fact about probability measures on $\Omega$ is the following. Let $\mathcal{M}'$ be a countably generated sub $\sigma$-algebra of $\mathcal{M}$ and let $P \in M(\Omega)$. Then conditional expectations with respect to $P$ given $\mathcal{M}'$ are given by integration. To be precise, there is a family $\{P_\omega\} \subseteq M(\Omega)$ such that $\omega \to P_\omega(A)$ is $\mathcal{M}'$-measurable for all $A \in \mathcal{M}$, $P_\omega(A) = \chi_A(\omega)$ for all $\omega \in \Omega$ and $A \in \mathcal{M}'$, and $E^P[P_\cdot(A), B] = P(A \cap B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{M}'$. In particular, $E^P[F|\mathcal{M}'] = E^{P_\cdot}[F]$ (a.s.,P) for all bounded $\mathcal{M}$-measure $F: \Omega \to C$. In the future we will call such a family $\{P_\omega\}$ a <u>regular conditional probability distribution of</u> $P$ <u>given</u> $\mathcal{M}'$ and we will abbreviate this statement by <u>r.c.p.d. of</u> $P|\mathcal{M}'$. Of particular importance to us will be the case when $\mathcal{M}'$ has a special form. Namely, let $\tau: \Omega \to [0,\infty) \cup \{\infty\}$ be a stopping time (i.e., $\{\tau \le t\} \in \mathcal{M}_t$ for all $t \ge 0$) and set $\mathcal{M}_\tau = \{A \in \mathcal{M}: A \cap \{\tau \le t\} \in \mathcal{M}_t$ for all $t \ge 0\}$. One can show that $\mathcal{M}_\tau = \sigma(x(t \wedge \tau): t \ge 0)$ and therefore that $\mathcal{M}_\tau$ is countably generated and that the atom in $\mathcal{M}_\tau$ containing $\omega_0$ is the $\{\omega \in \Omega: x(t,\omega) = x(t,\omega_0)$ for $0 \le t \le \tau(\omega_0)\}$. Hence if $\{P_\omega\}$ is a r.c.p.d. of $P|\mathcal{M}_\tau$, then $P_\omega(x(t) = x(t,\omega)\ ,\ 0 \le t \le \tau(\omega)) = 1$.

One final property of probability measures on $\Omega$ is that they can be "glued together" at stopping times. That is, let $P$ be a probability measure on $(\Omega, \mathcal{M}_\tau)$ and let $\{Q_\omega\} \subseteq M(P)$ be a family such that for all $s \ge 0$ and $A \in \mathcal{M}^s \equiv \sigma(x(t): t \ge s)$ the map $\omega \to Q_\omega(A)$ is $\mathcal{M}_s$-measurable on $\{\tau \ge s\}$ and $Q_\omega(x(\tau(\omega)) = x(\tau(\omega),\omega)) = 1$. Define $\delta_\omega \otimes_{\tau(\omega)} Q_\omega$ on $(\Omega, \mathcal{M})$ to be the measure such that $\delta_\omega \otimes_{\tau(\omega)} Q_\omega(A \cap B) = \chi_A(\omega) Q_\omega(B)$ for $A \in \mathcal{M}_{\tau(\omega)}$ and $B \in \mathcal{M}^{\tau(\omega)}$. Finally, define $P \otimes_{\tau(\cdot)} Q.$ by $P \otimes_{\tau(\cdot)} Q.(A) = E^P[\delta. \otimes_{\tau(\cdot)} Q.(A)]$ for $A \in \mathcal{M}$. Then $P \otimes_{\tau(\cdot)} Q.$ is the unique probability measure $R$ on $(\Omega, \mathcal{M})$ such that $R$ coincides with $P$ on $\mathcal{M}_\tau$ and $\{\delta_\omega \otimes_{\tau(\omega)} Q_\omega\}$ is a r.c.p.d. of $R|\mathcal{M}_\tau$.

## <u>Section II</u>

Let $a: [0,\infty) \times R^d \to S_d$ and $b: [0,\infty) \times R^d \to R^d$ be bounded measurable functions and define $L_t$ accordingly by (1.2). In this lecture we are going to derive various equivalent formulations of this martingale problem. Each of these formulations has its own special virtues and the ability to go from one to another will facilitate our understanding of the questions raised in Lecture I. The basic tool used in proving their equivalence is the following elementary lemma.

Lemma (2.1): Let $X: [0,\infty) \times \Omega \to C$ be a continuous $P$-martingale and let $Y: [0,\infty) \times \Omega \to C$ be a continuous function which is adapted to $\{\mathcal{M}_t: t \geq 0\}$ (i.e., $Y(t)$ is $\mathcal{M}_t$-measurable) and has the additional property that for all $T > 0$ and $\omega \in \Omega$ the total variation of $Y(\cdot\,,\omega)$ on $[0,T]$ is bounded by some constant $C_T < \infty$ which doesn't depend on $\omega$. Then $X(t)Y(t) - \int_0^t X(s)Y(ds)$ is again a $P$-martingale.

Proof: Let $0 \leq t_1 < t_2$ and $A \in \mathcal{M}_{t_1}$ be given. Then

$$E^P[X(t_2)Y(t_2) - X(t_1)Y(t_1),A]$$
$$= E^P[(X(t_2) - X(t_1))Y(t_1),A] + E^P[X(t_2)(Y(t_2) - Y(t_1)),A]$$
$$= E^P[X(t_2)(Y(t_2) - Y(t_1)),A] .$$

Next, let $\{t_1 = s_0 < \dots < s_n = t_2\}$ be partition points of the interval $[t_1,t_2]$. Then

$$E^P[X(t_2)(Y(t_2) - Y(t_1)),A] = \sum_{k=0}^{n-1} E^P[X(t_2)(Y(s_{k+1}) - Y(s_k)),A]$$
$$= \sum_{k=0}^{n-1} E^P[X(s_{k+1})(Y(s_{k+1}) - Y(s_k)),A]$$
$$\to E^P[\int_{t_1}^{t_2} X(s)Y(ds),A]$$

as $\max_{0\leq k\leq n} |s_{k+1} - s_k| \to 0$. Q.E.D.

Theorem (2.2): Let $P \in M(\Omega)$ satisfy $P(x(t) = x\ ,\ 0 \leq t \leq s) = 1$. Then the following statements are equivalent:

a) $P$ solves the martingale problem for $L_t$ starting from $(s,x)$,

b) for all $\theta \in R^d$, $X_\theta(t \vee s)$ is a P-martingale where

$$(2.3)\qquad X_\theta(t) = \exp[\langle\theta, x(t)-x(s) - \int_s^t b(s,x(s))ds\rangle - \frac{1}{2}\int_0^t \langle\theta, a(s,x(s))\theta\rangle ds] ,$$

c) for all $\theta \in R^d$, $X_{i\theta}(t \vee s)$ is a P-martingale.

Moreover, if P satisfies a), then for all $R > 0$

$$(2.4)\qquad P(\sup_{0\le t\le T} |x(t) - x(s) - \int_0^t b(s,x(s))ds| \ge R) \le 2de^{-R^2/2TAd^{1/2}}$$

where $A = \sup_{(s,x)} \sup_{|\theta|=1} \langle\theta, a(s,x)\theta\rangle$. Finally, if P satisfies a) and $f \in C_b^{1,2}([0,\infty) \times R^d)$ has the property that for each $T > 0$, $\sup_{0\le t\le T} (|f(t,x)| + |(\frac{\partial f}{\partial t} + L_t f)(t,x)|)$ is dominated by $C_T e^{\lambda_T|x|}$ for some $C_T < \infty$ and $\lambda_T > 0$, then $X_f(t \vee s)$ is a P-martingale where

$$(2.5)\qquad X_f(t) \equiv f(t,x(t)) - \int_s^t (\frac{\partial}{\partial u} + Lu)f(u,x(u))du .$$

<u>Proof</u>: Assume that $(s,x) = (0,\mathcal{O})$. Let P satisfy a). By an easy limiting procedure, we can conclude that $X_\varphi(t)$ is a P-martingale for all $\varphi \in C_b^2(R^d)$. Now let $\varphi \in C_b^2(R^d)$ be uniformly positive and take $X(t)$ and $Y(t)$ in Lemma (2.1) equal to $X_\varphi(t)$ and $\exp[-\int_0^t \frac{L_u\varphi(x(u))}{\varphi(x(u))}du]$, respectively. Then, by Lemma (2.1):

$$\varphi(x(t))\exp[-\int_0^t \frac{L_u\varphi}{\varphi}(x(u))du]$$

is a P-martingale. In particular, if $R > 0$ and $\varphi(x) = e^{\langle\theta,x\rangle}$ on $B(\mathcal{O},R)$ then, by Doob's stopping time theorem, $X_\theta(t \wedge \tau_R)$ is a P-martingale, where $\tau_R = \inf\{t \ge 0: |x(t)| \ge R\}$. Note that

$$X_\theta^2(t) = X_{2\theta}(t)\exp[\int_0^t \langle\theta, a(u,x(u))\theta\rangle du]$$

$$\le X_{2\theta}(t)e^{A|\theta|^2 t} ,$$

and so $\sup_{R>0} E^P[X_\theta^2(t\wedge\tau_R)] \le e^{A|\theta|^2 t} \sup_{R>0} E^P[X_{2\theta}(t \wedge \tau_R)] = e^{A|\theta|^2 t}$.

This shows that $\{X_\theta(t \wedge \tau_R): R > 0\}$ is uniformly P-integrable. Since $X_\theta(t \wedge \tau_R) \to X_\theta(t)$ as $R \uparrow \infty$, we have now proved that a) $\Rightarrow$ b).

Next assume that P satisfies b). Then, by an easy analytic continuation arguement, P must satisfy c). Finally, if P satisfies c), take $X(t)$ and $Y(t)$ in Lemma (2.1) equal to $X_{i\theta}(t)$ and $\exp[-i < \theta , \int_0^t b(u,x(u))du\rangle - \frac{1}{2}\int_0^t \langle\theta,a(u,x(u))\theta\rangle du]$ , respectively. Then, by Lemma (2.1), $X_\varphi(t)$ is a P-martingale with $\varphi(x) = e^{i\langle\theta,x\rangle}$. Since every $\varphi \in C_0^\infty(R^d)$ can be expressed as $\int e^{i\langle\theta,x\rangle}\psi(\theta)d\theta$ for some $\psi \in \mathscr{S}(R^d)$, it is easy to see that $X_\varphi(t)$ is a P-martingale for all $\varphi \in C_0^\infty(R^d)$.

The proof of (2.4) runs as follows. Let $|\theta| = 1$ be fixed. For $\lambda > 0$ we have by b) and Doob's inequality:

$$
\begin{aligned}
P(\sup_{0\le t\le T} \langle\theta,x(t) - \int_0^t b(u,x(u))du\rangle \geq R)&\\
&\leq P(\sup_{0\le t\le T} X_{\lambda\theta}(t) \geq \exp(\lambda R - \frac{\lambda^2}{2}\int_0^t \langle\theta,a(u,x(u))\theta\rangle du))\\
&\leq e^{-\lambda R + \lambda^2/2\, AT}.
\end{aligned}
$$

Setting $\lambda = \frac{R}{AT}$, we arrive at

$$
P(\sup_{0\le t\le T} \langle\theta,x(t) - \int_0^t b(u,x(u))du\rangle \geq R) \leq e^{-R^2/2AT};
$$

and (2.4) follows quickly from this.

The final part of the theorem can now be proved in two easy steps. First, one shows that $X_f(t)$ is a P-martingale for all $f \in C_b^{1,2}([0,\infty) \times R^d)$. This is easily done by initially assuming that $f \in C_b^\infty([0,\infty) \times R^d)$ and applying a) to f as a function of x and the fundamental theorem of calculus to f as a function of t. After this has been done, it is easy to pass to all $f \in C_b^{1,2}([0,\infty) \times R^d)$. The second step is to use an approximation procedure and apply (2.4) to justify the passage to the limit. Q.E.D.

As a preliminary application of Theorem (2.2), we point out that uniqueness of solutions to the martingale for $L_t$ can be easily proved whenever one has a strong enough existence theorem for the P.D.E.:

$$\frac{\partial u}{\partial t} + L_t u = -\varphi\,, \quad 0 \le t < T \tag{2.6}$$
$$\lim_{t \uparrow T} u(t,\cdot) = 0$$

Indeed, suppose that (2.6) admits a smooth solution $u$ for every $T > 0$ and $\varphi \in C_0^\infty(R^d)$. Then, for any solution $P$ to the martingale for $L_t$ starting from $(s,x)$:

$$u(t \vee s, x(t \vee s)) + \int_s^{t \vee s} \varphi(x(u))du$$

is a P-martingale for $t \le T$. Thus

$$u(s,x) = E^P[\int_s^T \varphi(x(u))du] \tag{2.7}$$

for all $T > s$ and $\varphi \in C_0^\infty(R^d)$. This means that the one dimensional time marginals of $P$ are unique determined. To complete the proof that $P$ itself is unique, we require the next theorem.

Theorem (2.8): Let $P$ solve the martingale problem for $L_t$ starting from $(s,x)$ and let $\tau \ge s$ be a stopping time. If $\{P_\omega\}$ is a r.c.p.d. of $P|\mathcal{M}_\tau$, then there is a P-null set $N \in \mathcal{M}_\tau$ such that whenever $\omega \notin N$ the measure $\delta_{x(\tau(\omega),\omega)} \otimes_{\tau(\omega)} P_\omega$ solves the martingale problem for $L_t$ starting from $(\tau(\omega), x(\tau(\omega),\omega))$ (here $\delta_{x(\tau(\omega),\omega)}$ stands for the point mass on the path which is constantly equal to $x(\tau(\omega),\omega)$).

The proof of Theorm (2.8) is not difficult, but it is somewhat tedious. The idea is to show that for each $\varphi \in C_0^\infty(R^d)$, $X_\varphi(t \vee \tau(\omega))$ is a $P_\omega$-martingale for P-almost all $\omega$. Since $C_0^\infty(R^d)$ is separable, one can isolate one P-null set $N \in \mathcal{M}_\tau$ such that $X_\varphi(t \vee \tau(\omega))$ is a $P_\omega$-martingale for all $\omega \notin N$ and $\varphi \in C_0^\infty(R^d)$.

Given Theorem (2.8), we can now easily complete argument begun above. Indeed, by Theorem (2.8) plus (2.7), we have that

$$E^P[\int_t^T \varphi(x(u))du|\mathcal{M}_t] = u(t,x(t)) \quad (\text{a.s.},P) \tag{2.9}$$

for $s \le t < T$, since $E^P[\int_t^T \varphi(x(u))du|\mathcal{M}_t] = E^{P_\cdot}[\int_t^T \varphi(x(u))du]$ (a.s.,P) where $P_\cdot$ is the r.c.p.d. of $P|\mathcal{M}_t$. But from (2.9), it is immediate that all finite dimensional time marginals of $P$ are uniquely determined, and therefore $P$ itself is unique.

The preceding line of reasoning applies to many choices of $L_t$. For instance, if $((a^{ij}(f,x)))$ and $(b^i(t,x))$ are bounded and Hölder continuous and if $((a^{ij}(t,x)))$ is uniformly positive definite or if $((a^{ij}(t,x)))$ and $(b^i(t,x))$ are sufficiently smooth, then (2.6) admits good solutions. In particular, if $L_t = 1/2\Delta$, then the corresponding martingale problem has at most one solution for each $(s,x)$ and in fact

$$P(x(t_2) \in \Gamma|\mathcal{M}_{t_1}) = \int_\Gamma g(t_2 - t_1, y - x(t_1))dy \quad (\text{a.s.},P) \tag{2.10}$$

for all $s \le t_1 < t_2$.

Conversely, if $P$ on $(\Omega,\mathcal{M})$ satisfies $P(x(t) = x,\ 0 \le t \le s) = 1$ and (2.10), then $P$ solves the martingale problem for $1/2\Delta$ starting from $(s,x)$. To see this, note that from (2.10) one gets

$$E^P[e^{i\langle\theta,x(t_2)\rangle}|\mathcal{M}_{t_1}] = e^{i\langle\theta,x(t_1)\rangle - \frac{|\theta|^2}{2}(t_2 - t_1)}$$

and therefore that $\exp[i\langle\theta,x(t)\rangle + \frac{|\theta|^2}{2} t$ is a P-martingale for all $\theta \in R^d$. By Theorem (2.2), this means that $P$ solves the martingale problem for $1/2\Delta$. We are now in a position to identify Wiener measure with solutions to the martingale problem for $1/2\Delta$. Starting from (1.10), it is an easy matter to see that $\mathcal{W}(x(0) = 0) = 1$ and that (2.10) is satisfied with $P$ replaced by $\mathcal{W}$. Thus $\mathcal{W}$ is the unique solution to the martingale problem for $1/2\Delta$ starting from $(0, \mathcal{O})$. To get the solution starting from a general $(s,x)$, we take advantage of the translation invariance of $1/2\Delta$. That is, define $\Phi_{s,x}: \Omega \to \Omega$ so that $x(t,\Phi_{s,x}(\omega)) = x + x((t-s) \vee 0,\omega)$, $t \geq 0$. Clearly $\Phi_{s,x}(\omega)$ is jointly continuous in $(s,x)$ and $\omega$. Let $\mathcal{W}_{s,x} = \mathcal{W} \circ \Phi_{s,x}^{-1}$. An elementary computation identifies $\mathcal{W}_{s,x}$ as the unique solution to the martingale problem for $1/2\Delta$ starting from $(s,x)$.

## Section III

We begin in this lecture to prepare the machinery for our attack on the question of uniqueness. Crucial to this enterprise is the relationship between the martingale problem and Itô stochastic integral equations. Throughout this lecture we will be assuming that $((a^{ij}))$ is uniformly positive definite. Under this assumption we will show that $P$ solves the martingale problem for $L_t$ starting from $(s,x)$ if and only if

$$(3.1) \qquad x(t) = x + \int_s^{t \wedge s} a^{1/2}(u,x(u))d\beta(u) + \int_s^{t\wedge s} b(u,x(u))du \ , \ t \geq 0 \ , \qquad (\text{a.s.},P)$$

where $\beta(\cdot)$ is a <u>P-Brownian motion after time $s$</u> (i.e., $\beta(t)$ is $\mathcal{M}_t$-measurable for all $t \geq s$, $\beta(\cdot,\omega)$ is continuous on $[s,\infty)$ for P-almost all $\omega$, $\beta(s) = 0$ (a.s.,P), and

$$(3.2) \qquad P(\beta(t_2) \in \Gamma | \mathcal{M}_{t_1}) = \int_\Gamma g(t_2 - t_1, y - \beta(t_1))dy \quad (\text{a.s.},P)$$

for $s \leq t_1 < t_2$). The first integral on the right hand side of (3.1) is to be interpreted in the sense of Itô and is well-defined since $\beta(\cdot)$ is a Brownian motion $\beta(\cdot)$ relative to the family $\{\mathcal{M}_t: t \geq s\}$. It must be emphasized that just because (3.1) holds, there is <u>no</u> implication here that $x(\cdot)$ is $\beta(\cdot)$-measurable. That is, $\sigma(\beta(u): s \leq u \leq t)$ may be strictly smaller than $\sigma(x(u)\colon s \le u \le t)$.

To see that (3.1) implies that $P$ solves the martingale problem for $L_t$ starting from $(s,x)$, we need only invoke Itô's formula and thereby conclude that if (3.1) holds:

$$\varphi(x(t)) - \int_s^t L_u\varphi(x(u))du = \varphi(x) + \int_s^t \langle \nabla\varphi(x(u)) , d\beta(u)\rangle \quad , \quad t \ge s \quad .$$

Since the right hand side of the preceding is a P-martingale, there can be no doubt that $P$ solves the martingale problem for $L_t$ starting from $(s,x)$.

The converse statement is not quite so easily proved. To facilitate the presentation, we will assume that $s = 0$, $x = \mathcal{O}$, and that $b \equiv 0$. Given a solution $P$ to the martingale problem for $L_t$ starting from $(0,\mathcal{O})$, we proceed to develop the theory of Itô-type stochastic integrals for the process $x(\cdot)$. This is done, word for word, in the same way as in the case of Brownian motion. That is, we start with bounded measurable functions $\theta\colon [0,\infty) \times \Omega \to R^d$ such that $\theta(t)$ is $\mathcal{M}_t$-measurable for all $t \ge 0$ and $\theta(t) = \theta(\frac{[nt]}{n})$, $t \ge 0$, for some $n \ge 1$. The definition of $\int_0^t \langle\theta(u) , dx(u)\rangle$ is then given by the Riemann sum:

$$\sum_{k=1}^{[nt]} \langle\theta(\tfrac{k-1}{n}) , x(\tfrac{k}{n}) - x(\tfrac{k-1}{n})\rangle + \langle\theta(\tfrac{[nt]}{n}), x(t) - x(\tfrac{[nt]}{n})\rangle .$$

We then observe that for such $\theta$'s:

$$(3.2) \qquad X_\theta(t) = \exp[\int_0^t \langle\theta(u),dx(u)\rangle - \frac{1}{2}\int_0^t \langle\theta(u),a(u,x(u))\theta(u)\rangle\, du]$$

is a P-martingale. To see this, assume that $0 \le t_1 \le t_2 \le \frac{[nt_1]+1}{n}$ and let $\{P_\omega\}$ be a r.c.p.d. of $P|\mathcal{M}_{t_1}$. Then, by b) of Theorem (2.2):

$$E^P[X_\theta(t_2)|\mathcal{M}_{t_1}]$$

$$= X_\theta(t_1,\cdot)E^{P\cdot}[\exp[\langle\theta\left(\frac{[nt_1]}{n},\cdot\right), x(t_2) - x(t_1)\rangle - \frac{1}{2}\int_{t_1}^{t_2}\langle\theta\left(\frac{[nt_1]}{n},\cdot\right), a(u,x(u))\theta\left(\frac{[nt_1]}{n},\cdot\right)\rangle du]]$$

$$= X_\theta(t_1) \quad .$$

Clearly this proves that $X_\theta(t)$ is a P-martingale. Now replace $\theta$ by $\lambda\theta$ where $\lambda \in R^1$. Differentiating once and then twice with respect to $\lambda$, after setting $\lambda = 0$ we see that:

$$\int_0^t \langle\theta(u), dx(u)\rangle \tag{3.3}$$

and

$$\left(\int_0^t \langle\theta(u), dx(u)\rangle\right)^2 - \int_0^t \langle\theta(u), a(u,x(u))\theta(u)\rangle du \tag{3.4}$$

are P-martingale. From here it is an easy task to complete the definition of $\int_0^t \langle\theta(u), dx(u)\rangle$ for general bounded measurable $\theta\colon [0,\infty) \times \Omega \to \mathbb{C}^d$ which are adapted to $\{\mathcal{M}_t\colon t \geq 0\}$. The procedure is identical to the one used in the Brownian case. In this way, we arrive at a definition of $\int_0^t \langle\theta(u), dx(u)\rangle$ for such $\theta$'s; and the integral $\int_0^t \langle\theta(u), dx(u)\rangle$ enjoys the following properties:

a) $\int_0^t \langle\theta(u), dx(u)\rangle$ is an a.s. continuous P-martingale,

b) $\left(\int_0^t \langle\theta(u), dx(u)\rangle\right)^2 - \int_0^t \langle\theta(u), a(u,x(u))\theta(u)\rangle du$ is a P-martingale,

c) $X_\theta(t)$ is a P-martingale, where $X_\theta$ is given by (3.2).

Next suppose that $\sigma\colon [0,\infty) \times \Omega \to R^d \otimes R^d$ is a bounded measurable function which is adapted to $\{\mathcal{M}_t\colon t \geq 0\}$. We then define $\int_0^t \sigma(u)dx(u)$ so that for all $\theta \in R^d$: $\langle \theta , \int_0^t \sigma(u)dx(u)\rangle = \int_0^t \langle \sigma^*(u)\theta, dx(u)\rangle$ . In particular, if $\sigma(u) = a^{1/2}(u,x(u))$ and $\beta(t) = \int_0^t \sigma(u)dx(u)$ , then for each $\theta \in R^d$:

$$\exp[i\langle \theta, \beta(t)\rangle + \frac{|\theta|^2}{2} t]$$

$$= \exp[i\int_0^t \langle \sigma(u)\theta, dx(u)\rangle + \frac{1}{2}\int_0^t \langle \sigma(u)\theta, a(u,x(u))\sigma(u)\theta\rangle du]$$

is a P-martingale. But this means that $\beta(\cdot)$ is a P-Brownian motion. Moreover, it is not hard to show that for any bounded measurable $\theta\colon [0,\infty) \times \Omega \to \mathbf{C}^d$ which is adapted to $\{\mathcal{M}_t\colon t \geq 0\}$ ,

$$\int_0^t \langle \sigma(u)\theta(u), dx(u)\rangle = \int_0^t \langle \theta(u), d\beta(u)\rangle \ .$$

In particular, for $\theta \in R^d$ :

$$\langle \theta, x(t)\rangle = \int_0^t \langle \theta, dx(u)\rangle$$

$$= \int_0^t \langle \sigma(u)a^{1/2}(u,x(u))\theta, dx(u)\rangle$$

$$= \int_0^t \langle a^{1/2}(u,x(u))\theta, d\beta(u)\rangle = \langle \theta , \int_0^t a^{1/2}(u,x(u))d\beta(u)\rangle$$

We have therefore demonstrated how to go from a solution to the martingale problem to the equation

$$x(t) = \int_0^t a^{1/2}(u,x(u))d\beta(u) \quad .$$

The equivalence between solutions to the martingale problem and stochastic integral equations opens up the possibility of studying the martingale problem with Itô's methods. First consider the question of existence. If there is somewhere a probability space $(E,\mathcal{F},Q)$ on which there is a Brownian motion $\beta(\cdot)$ relative to an increasing family

$\{\mathcal{F}_t : t \ge 0\}$ of sub $\sigma$-algebras and if there is a measurable function $\xi : [0,\infty) \times E \to R^d$ adapted to $\{\mathcal{F}_t : t \ge 0\}$ such that

$$\xi(t) = x + \int_s^{s \vee t} a^{1/2}(u,\xi(u))d\beta(u) + \int_s^{s \vee t} b(u,\xi(u))du \quad ,$$

then the distribution on $(\Omega, \mathcal{M})$ of $\xi(\cdot)$ under $Q$ solves the martingale problem for $L_t$ starting from $(s,x)$. Indeed, this assertion is proved in exactly the same way as we proved (3.1) is a sufficient condition for $P$ to solve the martingale problem. Thus we now see that existence for the martingale problem can be proved whenever on some space there is existence for the stochastic integral equation associated with the coefficients of $L_t$.

Next we investigate whether uniqueness for the martingale problem can be deduced from uniqueness for the associated stochastic differential equation. To be specific, let $\sigma = a^{1/2}$ and assume that

$$\sup_{0 \le t \le T} \left(\|\sigma(t,x) - \sigma(t,y)\| + |b(t,x) - b(t,y)|\right) \le C|x - y| \tag{3.5}$$

for all $T > 0$. If $P$ solves the martingale problem for $L_t$ starting for $(s,x)$ and $\beta(\cdot)$ is a P-Brownian motion for which (3.1) obtains, define $\xi_0(\cdot) \equiv x$ and

$$\xi_n(t) = x + \int_s^{s \vee t} \sigma(u,\xi_{n-1}(u))d\beta(u) + \int_s^{s \vee t} b(u,\xi_{n-1}(u))du \text{ , } n \ge 1.$$

Clearly each $\xi_n(\cdot)$ is a functional of $\beta(\cdot)$ and therefore its distribution under $P$ is the same as that of the analogous quantity under any other solution to the same martingale problem. Furthermore, $\xi_n(\cdot) \to \xi(\cdot)$ where

$$\xi(t) = x + \int_s^{t \vee s} \sigma(u,\xi(u))d\beta(u) + \int_s^{t \vee s} b(u,\xi(u))du \text{ .}$$

Clearly $\xi(\cdot)$'s distribution again is the same for all solutions to the same martingale problem. Finally, by pathwise uniqueness, $\xi(\cdot) = x(\cdot)$ (a.s.,P). Thus the condition in (3.5) is enough to guarantee uniqueness for the martingale problem via uniqueness for the corresponding stochastic differential equation. Actually a more refined technique shows that after the notion of uniqueness for stochastic differential equations has been properly formulated, then uniqueness for the martingale problem is always a consequence of uniqueness for the corresponding stochastic differential equation. This more refined technique is intimately connected with the determination of the circumstances under which $x(\cdot)$ is a functional of the $\beta(\cdot)$ in (3.1). We will take this subject up again in Section V.

## Section IV

We open this section with a quite general existence theorem for solutions to the martingale problem.

Given $A \in S_d$, $B \in R^d$, and $(s,x) \in [0,\infty) \times R^d$, define $\Phi^{(A,B)}_{\{s,x\}}$: $\Omega \to \Omega$ by $x(t,\Phi^{(A,B)}_{\{s,x\}}\omega) = x + A^{1/2}x((t-s) \vee 0,\omega) + ((t-s) \vee 0)B$ and let $W^{(A,B)}_{s,x} = W \circ (\Phi^{(A,B)}_{s,x})^{-1}$. It is clear that $((A,B),(s,x)) \to W^{(A,B)}_{s,x}$ is a continuous map. Moreover, a simple computation suffices in order to check that $W^{(A,B)}_{s,x}$ is the unique solution to the martingale problem for $1/2 \sum_{i,j=1}^{d} A^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} B^i \frac{\partial}{\partial x_i}$ starting from $(s,x)$. Now suppose that

$a: [0,\infty) \times R^d \to S_d$ and $b: [0,\infty) \times R^d \to R^d$ are bounded continuous functions. Given $n \geq 1$, define $P^{(m)}_n$ inductively on $m$ by

$$P^{(0)}_n = W^{(a(0,\sigma),b(0,\sigma))}_{0,\sigma}$$

and

$$P_n^{(m)} = P_n^{(m-1)} \otimes_{m/n} \mathbb{W}_{m/n\,,\,x(m/n,\cdot)}^{(a(m/n,x(m/n\,,\,\cdot)),b(m/n,x(m/n,\cdot))}\quad , \quad m \geq 1 \quad .$$

Since it is clear that $P_n^{(m+1)}$ restricted to $\mathcal{M}_{m/n}$ coincides with $P_n^{(m)}$ restricted to $\mathcal{M}_{m/n}$ , standard extention theorems tell us that there is a unique $P_n$ on $(\Omega\,,\mathcal{M})$ such that $P_n$ on $\mathcal{M}_{m/n}$ coincides with $P_n^{(m)}$ on $\mathcal{M}_{m/n}$ . Furthermore, it is not hard to check by induction that for any $\varphi \in C^2(R^d)$ which, together with its first and second or derivatives, grows no faster than an exponential:

$$X_\varphi^{(n)}(t) \equiv \varphi(x(t)) - \int_0^t \sum_{i,j=1}^d a^{ij}\left(\frac{[nu]}{n},x\left(\frac{[nu]}{n}\right)\right)\frac{\partial^2\varphi}{\partial x_i\partial x_j}(x(u))du$$
$$- \int_0^t \sum_{i=1}^d b^i\left(\frac{[nu]}{n}\,,\,x\left(\frac{[nu]}{n}\right)\right)\frac{\partial\varphi}{\partial x_i}(x(u))du$$

is a $P_n$-martingale. In particular, one can see from this that for each $T > 0$ there is a constant $C_T$ which is independent of $n \geq 1$ such that

$$E^{P_n}[|x(t_2) - x(t_1)|^4] \leq C_T(t_2 - t_1)^2 \;, \; 0 \leq t_1 < t_2 \leq T \quad .$$

Since $P_n(x(0) = \mathcal{O}) = 1$ for all $n \geq 1$ , we now see that $\{P_n\colon n \geq 1\}$ is pre-compact in $M(\Omega)$ . Let $\{P_{n'}\}$ be a convergent subsequence with limit $P$ . Then $P(x(0) = \mathcal{O}) = 1$ . Moreover, if $\varphi \in C_0^\infty(R^d)$ , then $a\left(\frac{[nt]}{n}\,,x\left(\frac{[nt]}{n}\,,\omega\right)\right) \to a(t,x(t,\omega))$ and $b\left(\frac{[nt]}{n}\,,x\left(\frac{[nt]}{n}\,,\omega\right)\right) \to b(t,x(t,\omega))$ uniformly as $(t,\omega)$ ranges over compact subsets of $[0,\infty) \times \Omega$ . Hence if $0 \leq t_1 < t_2$ and $F$ is a bounded $\mathcal{M}_{t_1}$-measurable function, then

$$\begin{aligned} E^P[X_\varphi(t_2)F] &= \lim_{n\to\infty} E^{P_n}[X_\varphi^{(n)}(t_2)F] \\ &= \lim_{n\to\infty} E^{P_n}[X_\varphi^{(n)}(t_1)F] = E^P[X_\varphi(t_1)F] \quad , \end{aligned}$$

where $X_\varphi(t) = \varphi(x(t)) - \int_0^t L_u\varphi(x(u))du$ . From here it is an easy step to conclude that $P$ solves the martingale problem for $L_t$ starting from $(0, \mathcal{O})$. By a trivial change in notation, we could have carried out the same line of reasoning to produce a solution starting from any $(s,x)$. Thus we have proved the next theorem.

Theorem (4.1): Let $a: [0,\infty) \times R^d \to S_d$ and $b: [0,\infty) \times R^d \to R^d$ be bounded continuous functions and define $L_t$ accordingly. Then for each $(s,x)$ there is a solution to the martingale problem for $L_t$ starting from $(s,x)$.

The rest of this section is devoted to the development of the Cameron-Martin-Girsonov formula. This formula will enable us to reduce both the question of existence as well as uniqueness when $a$ is positive definite to the case in which $b \equiv 0$.

Let $a$ be uniformly positive definite. Given a solution $P$ to the martingale problem for $L_t^0 = \frac{1}{2} \sum_{i,j=1}^{d} a^{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j}$ starting from $(s,x)$, define

$$(4.2) \quad R(t) = \exp\Big[\int_s^{t \vee s} \langle a^{-1} b(u,x(u)), dx(u)\rangle - \frac{1}{2}\int_s^{t\vee s} \langle b(u,x(u)), a^{-1}(u,x(u)) b(u,x(u))\rangle du\Big] .$$

Then $R(t)$ is a $P$-martingale. Thus there is a unique $Q$ on $(\Omega, \mathcal{M})$ such that $Q(A) = E^P[R(t),A]$ for all $t \geq 0$ and $A \in \mathcal{M}_t$. We claim that $Q$ solves the martingale problem for $L_t = L_t^0 + \sum_{i=1}^{d} b^i(t,x) \frac{\partial}{\partial x_i}$ starting from $(s,x)$. To see this, let $\theta_0 \in R^d$ be given and define $\theta(t) = \theta_0 + a^{-1} b(t,x(t))$. Then for $s \leq t_1 < t_2$ and $A \in \mathcal{M}_{t_1}$:

$$E^Q[X_{\theta_0}(t_2),A] = E^P[X_{\theta_0}(t_2)R(t_2),A]$$
$$= E^P[X_\theta(t_2),A] = E^P[X_\theta(t_1),A]$$
$$= E^P[X_{\theta_0}(t_1)R(t_1),A] = E^Q[X_{\theta_0}(t_1),A] \quad ,$$

where $X_{\theta_0}(t) = \exp[\langle\theta_0, x(t\vee s) - x\rangle - \frac{1}{2}\int_s^{t\vee s}\langle\theta_0, a(u,x(u))\theta_0\rangle du]$ and $X_\theta(t) = \exp[\int_s^{t\vee s}\langle\theta(u), dx(u)\rangle - \frac{1}{2}\int_s^{t\vee s}\langle\theta(u), a(u,x(u))\theta(u) > du]$. This justifies our claim. Conversely, suppose that $Q$ is a solution for $L_t$ starting from $(s,x)$ and define $S(t) = 1/R(t)$. Then, by extending the considerations of Section 3 to cover $b \neq 0$, one can show that $S(t)$ is $Q$-martingale and therefore that there is a $P$ on $(\Omega, \mathcal{M})$ such that $P(A) = E^Q[S(t),A]$, $t \geq 0$ and $A \in \mathcal{M}_t$. Reasoning as we did above, one can now check that $P$ solves for $L_t^0$ starting from $(s,x)$. With these remarks, we have the following important theorem.

<u>Theorem (4.3):</u> Let $a: [0,\infty) \times R^d \to S_d$ and $b: [0,\infty) \times R^d \to R^d$ be bounded measurable functions and assume that $a$ is uniformly positive definite. Then $Q$ solves the martingale problem for $L_t =$
$\frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(t,x)\frac{\partial}{\partial x_i}$ starting from $(s,x)$ if and only if there is a solution $P$ to the martingale problem for
$L_t^0 = \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}$ such that for all $t \geq 0$ and $A \in \mathcal{M}_t$
$Q(A) = E^P[R(t),A]$ where $R(t)$ is defined in (4.2). In particular, existence (uniqueness) for the martingale corresponding to $L_t$ follows from existence (uniqueness) for the martingale problem associated with $L_t^0$.

## Section 5

We saw in the preceding section that the problem of proving uniqueness for solutions to the martingale problem for the case of general coefficients $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$ can be reduced to the case $b(\cdot,\cdot) \equiv 0$, when $a(\cdot,\cdot)$ is uniformly elliptic. There are other procedures which will be useful in proving the uniqueness of solutions to the martingale problem.

<u>Localization</u>. Suppose $\{G_\alpha\}$ is an open covering of $[0,\infty) \times R^d$ and for each $\alpha$ we have coefficients $\{a_\alpha(\cdot,\cdot),b_\alpha(\cdot,\cdot)\}$ such that

(i) $\{a_\alpha(\cdot,\cdot),b_\alpha(\cdot,\cdot)\} \equiv \{a(\cdot,\cdot),b(\cdot,\cdot)\}$ on $G_\alpha$ and

(ii) For each $\alpha$ we have a unique measurable family $\{P^\alpha_{s,x}\}(s,x) \in [0,\infty) \times R^d$ of solutions to the martingale problem corresponding to $\{a_\alpha(\cdot,\cdot),b_\alpha(\cdot,\cdot)\}$ indexed by the starting points $(s,x) \in [0,\infty) \times R^d$.

Then the solution to the martingale problem corresponding to $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$, if it exists, is unique for every starting point $(s,x) \in [0,\infty) \times R^d$.

<u>Outline of the proof</u>: Let $P_1$ and $P_2$ be two solutions of the martingale problem corresponding to $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$ starting from the same point $(s_0,x_0)$. Let $B_R$ be the open ball of radius R in $[0,\infty) \times R^d$ around $(s_0,x_0)$. Let $\hat{\tau}_R = \inf\{t: (t,x(t)) \notin B_R\}$ be the exit time from the ball. Clearly $\hat{\tau}_R$ is a stopping time and $\hat{\tau}_R(\omega) \to \infty$ as $R \to \infty$ for each $\omega \in \Omega$. It is therefore sufficient to show that $P_1$ and $P_2$ agree on $\mathcal{M}_{\hat{\tau}_R}$ for each $R < \infty$. Let us fix a value of $R < \infty$.

It follows from standard compactness arguments that we can find a number $\delta = \delta_R > 0$ such that for any $(s,x)$ in the closed ball $\bar{B}_R$ around $(s_0,x_0)$ we can find a value of $\alpha$ such that

$S_\delta(s_0,x_0) \subset G_\alpha$ where $S_\delta(s_0,x_0)$ is the sphere around $(s_0,x_0)$ of radius $\delta$. We now define stopping times $\tau_1,\tau_2,\ldots$ by $\tau_0 = s_0$ and $\tau_n = \inf\{t: t \geq \tau_{n-1}, |(t,x(t)) - (\tau_{n-1},x(\tau_{n-1})| \geq \delta\}$ for $n \geq 1$. By the continuity of paths $\tau_n \to \infty$ as $n \to \infty$. If we define $\sigma_n = \tau_n \wedge \hat{\tau}_R$ it is clearly sufficient to prove that $P_1$ and $P_2$ agree on $M_{\sigma_n}$ for every n. We will prove it by induction on n. Let us take the case $n = 1$. Since $(s_0,x_0) \in B_R$ we can find $\alpha_0$, such that $\{a_{\alpha_0}(\cdot,\cdot),b_{\alpha_0}(\cdot,\cdot)\} \equiv \{a(\cdot,\cdot),b(\cdot,\cdot)\}$ on $S_\delta(s_0,x_0)$ and we have unique solutions $\{P^{\alpha_0}_{s,x}\}$ corresponding to $\{a_\alpha(\cdot,\cdot),b_\alpha(\cdot,\cdot)\}$. Let us construct solutions $\tilde{P}_1$ and $\tilde{P}_2$ by the relations

$$\tilde{P}_i = P_i \text{ on } M_{\tau_1} \text{ for } i = 1,2$$

$$\text{r.c.p.d. of } \tilde{P}_i|_{M_{\tau_1}} = P^{\alpha_0}_{\tau_1(\omega),x(\tau_1(\omega),\omega)} \quad \text{on } M^{\tau_1(\omega)} \quad \text{for } i = 1,2.$$

Then by the properties of martingales $\tilde{P}_i$ are solutions to the martingale problem corresponding to $\{a_{\alpha_0}(\cdot,\cdot),b_{\alpha_0}(\cdot,\cdot)\}$ and by uniqueness $\tilde{P}_1 \equiv \tilde{P}_2$. Therefore $P_1 \equiv P_2$ on $M_{\tau_1}$ and a fortiori on $M_{\sigma_1}$.

Assume that we have the result for $n = \ell$. Then $P_1 \equiv P_2$ on $M_{\sigma_\ell}$ and the r.c.p.d.'s $P_1^\omega$ and $P_2^\omega$ of $P_1$ and $P_2$ given $M_{\sigma_\ell}$ are again solutions to the martingale problem corresponding to $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$ starting from $(\sigma_\ell,x(\sigma_\ell))$, at least for almost all $\omega$ with respect to $P_1$ or $P_2$. Now the argument for $n = 1$ applies to the conditional distributions $P_1^\omega$ and $P_2^\omega$. The role of $(s_0,x_0)$ is played by $(\sigma_\ell,x(\sigma_\ell))$ and the new $\sigma_1$ is the same as $\sigma_{\ell+1}$. One therefore obtains that $P_1^\omega$ and $P_2^\omega$ agree on $M^{\sigma_\ell(\omega)}_{\sigma_{\ell+1}}$ for almost all $\omega$. Consequently $P_1 = P_2$ on $M_{\sigma_{\ell+1}}$ and the induction is complete.

The impact of the above argument is that uniqueness is a local property of the coefficients.

One Dimensional Marginals

Suppose we have a measurable family $p(s,x,t,\cdot)$ of probability measures indexed by $s < t$ and $x \in R^d$ such that for any solution $P$ to the martingale problem starting from any point $(s_0,x_0)$ corresponding to $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$

$$P[x(t) \in A] = p(s_0,x_0,t,A) \text{ for } t > s_0 \text{ and } A \in \mathcal{B}(R^d).$$

Then the solution to the martingale problem corresponding to $\{a(\cdot,\cdot),b(\cdot,\cdot)\}$ is unique for any starting point $(s_0,x_0)$, and is the Markov process with transition probabilities $p(s,x,t,\cdot)$. In particular $p(s,x,t,\cdot)$ satisfies the Chapman-Kolmogorov equations.

Proof: Let us consider the r.c.p.d. $Q_\omega$ of any solution $P$ starting from $(s_0,x_0)$ given the $\sigma$-field $M_{t_0}$ for some $t_0 > s_0$. The solution $Q_\omega$ is again a solution to the martingale problem starting from $(t_0,x(t_0))$. By our assumption we have

$$Q_\omega[x(t) \in A] = p(t_0,x(t_0),t,A) \quad \text{a.e. } P$$

for $t > t_0$ and $A \in \mathcal{B}(R^d)$. $P$ is therefore the Markov process with transition probabilities $p(s,x,t,\cdot)$ starting from $(s_0,x_0)$. $P$ is therefore unique and moreover $p(\cdot,\cdot,\cdot,\cdot)$ must satisfy the Chapman-Kolmogorov equations.

Remark. It now follows by conditioning with respect to any $M_\tau$ where $\tau$ is a stopping time that the r.c.p.d. of $P$ given $M_\tau$ is the solution starting from $(\tau, x(\tau))$ for almost all $\omega$. In other words the family of unique solutions $\{P_{s,x}\}$ has the strong Markov property.

## Section 6

We will continue our discussion of various circumstances under which either a reduction or a complete solution of the problem of uniqueness is possible.

### Random Time Change

Let $\phi(x)$ be a measurable function of $x$ in $R^d$ and satisfy the the bounds $0 < c_1 < \phi(x) \leq C_1 < \infty$ for all $x \in R^d$. We introduce a map $T_\phi$ of $\Omega \to \Omega$ as follows:

$$(T_\phi \omega)(t) = \omega(\tau_\phi(t))$$

where $\tau_\phi(t)$ is a solution of

$$\int_0^{\tau_\phi(t)} \phi\big(x(s,\omega)\big)\, ds = t .$$

Since $\phi$ is bounded above and below, we have a unique solution $\tau_\phi(t)$ of the above equation, which is a stopping time (as a function of $\omega$) for each $t \geq 0$. Moreover $\tau_\phi(t)$ is nondecreasing in $t$ and tends to $\infty$ as $t \to \infty$ for each $\omega$. In fact

$$\frac{t}{C_1} \leq \tau_\phi(t) \leq \frac{t}{c_1}$$

for all $\omega$ and $t$.

If $\phi,\psi$ are two functions of the above type then the maps $T_\phi$ and $T_\psi$ have the property

$$T_\phi \circ T_\psi = T_\psi \circ T_\phi = T_{\phi\psi}$$

The above property is easily verified by computing the derivative

$$\frac{d\tau_\phi(t)}{dt} = \frac{1}{\phi\big(x(\tau_\phi(t),\omega)\big)} = \frac{1}{\phi\big(x(t,T_\phi\omega)\big)}$$

In particular $T_\phi$ and $T_{1/\phi}$ are inverses of each other. One can also verify that

$$T_\phi \mathcal{M}_t \subset \mathcal{M}_{\tau_\phi(t)}$$

Suppose now that we have coefficients which are independent of time and we denote by L the operator

$$L = \frac{1}{2} \sum a^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum b^j(x) \frac{\partial}{\partial x_j}$$

and P is a solution corresponding to L , starting from the point $x_0$ at time 0. Then

$$f\big(x(t)\big) - \int_0^t (Lf)\big(x(s)\big)\, ds$$

is a $(\Omega, \mathcal{M}_t, P)$ martingale for all $f \in C_0^\infty(R^d)$. By Doob's stopping theorem,

$$f\big(x(\tau_\phi(t))\big) - \int_0^{\tau_\phi(t)} (Lf)\big(x(s)\big)\, ds$$

is a martingale relative to $(\Omega, \mathcal{M}_{\tau_\phi(t)}, P)$. We can rewrite this as

$$f\big(y(t)\big) - \int_0^t \frac{1}{\phi(y(s))} (Lf)\big(y(s)\big)\, ds$$

is a martingale where $y(t) = x(\tau_\phi(t)) = (T_\phi\omega)(t)$. Since the field generated by $y(s)$ for $0 \le s \le t$ is contained in $\mathcal{M}_{\tau_\phi(t)}$ we can say that

$$f\big(x(t)\big) - \int_0^t \frac{1}{\phi\big(x(s)\big)} (Lf)\big(x(s)\big)\, ds$$

is a martingale relative to $(\Omega, \mathcal{M}_t, Q)$ where $Q = PT_\phi^{-1}$. In other words, the transformation $T_\phi$ maps solutions of L into solutions of $\frac{1}{\phi}$ L. Since the mappint $T_\phi$ has the inverse $T_{1/\phi}$ we conclude that the solutions corresponding to L and the solutions corresponding to $\frac{1}{\phi}$ L for the same starting point (note that $(T_\phi\omega)(0) = \omega(0)$) are in one to one correspondence. In particular existence or uniqueness for L ensures the existence or uniqueness for $\frac{1}{\phi}$ L provided $\phi$ is bounded above and below.

Remark. Let us consider the case of a diffusion in $R^1$ corresponding to

$$L = \frac{1}{2} a(x) \frac{\partial^2}{\partial x^2} + b(x) \frac{\partial}{\partial x}$$

where a(x) and b(x) are bounded and measurable and a(x) in addition has the lower bound $a(x) \geq c > 0$. Then by the Cameron-Martin-Girsanov formula, the existence and uniqueness for L is the same as that for $L_0 = \frac{1}{2} a(x) \frac{\partial^2}{\partial x^2}$ and by the random time change discussed above it is the same as that for $\Delta_0 = \frac{1}{2} \frac{\partial^2}{\partial x^2}$. Since the only solution for $\Delta_0$ is the Brownian motion we conclude that we have existence and uniqueness for any starting point for the given operator L.

Connection with Itô's Theory

Let a(t,x) be such that $a(t,x) = \sigma(t,x)\sigma^*(t,x)$ for all t,x. Suppose we try to solve Itô's equation (in the more general sense) i.e. we look for a solution x(t), on some $(E,F_t,P)$, where there is also a Brownian motion $\beta(\cdot)$ which is given, of the equation

$$x(t) = x + \int_0^t \sigma(s,x(s))d\beta(s) + \int_0^t b(s,x(s))\, ds \quad .$$

More precisely we are looking for a measure $\tilde{P}$ on $C[[0,\infty);R^{2d}]$ starting from $(x,0)$ at time $0$, which solves the martingale problem corresponding to

$$a(t,x,y) = \left\{\begin{array}{c|c} a(t,x) & \sigma(t,x) \\ \hline \sigma^*(t,x) & I \end{array}\right\}$$

and

$$\tilde{b}(t,x,y) = \left\{b(t,x),0\right\}$$

This means that the first component is a solution to the martingale problem corresponding to $\{a(t,x),b(t,x)\}$, the second component is Brownian motion and the two are related by Itô's equations. One knows that any solution to the martingale problem corresponding to $\{a(t,x),b(t,x)\}$ can be exhibited as the first component of a solution $\tilde{P}$ corresponding to $\{\tilde{a},\tilde{b}\}$ with any choice of $\sigma$ such that $\sigma\,\sigma^* = a$.

Pathwise uniqueness can be phrased in terms of a solution to the martingale problem for an even bigger system. Consider for instance

$$\hat{a}(t,x,x',y) = \left\{\begin{array}{c|c|c} a(t,x) & \sigma(t,x)\sigma^*(t,x') & \sigma(t,x) \\ \hline \sigma(t,x')\sigma^*(t,x) & a(t,x') & \sigma(t,x') \\ \hline \sigma^*(t,x) & \sigma^*(t,x') & I \end{array}\right\}$$

and

$$\hat{b}(t,x,x',y) = \left\{b(t,x),\ b(t,x'),\ 0\right\}$$

A solution $\hat{P}$ to the martingale problem corresponding to $\{\hat{a},\hat{b}\}$

starting from $(x,x,0)$ on $C[[0,\infty); R^{3d}]$ is just the distribution of $x(t)$, $x'(t)$ and $\beta(t)$ where $\beta$ is a Brownian motion and $x(\cdot)$, $x'(\cdot)$ are two solutions of Itô's equation in terms of the Brownian motion starting from the same point $x$. Pathwise uniqueness is therefore the same as every such $\hat{P}$ living on the diagonal $x(t) \equiv x'(t)$ for all $t$.

To see that pathwise uniqueness implies that the solution to the martingale problem is unique, we need a construction which starts with two solutions $P_1$, $P_2$ to the martingale problem starting from the same point $x$ and ends up with a solution $\hat{P}$ starting from $(x,x,0)$ corresponding to $\{\hat{a},\hat{b}\}$ which has $P_1,P_2$ for the marginals for the first and second components respectively. Since $\hat{P}$ lives on the diagonal we will conclude that $P_1 = P_2$.

This construction carried out by Yamada and Watanabe is as follows: We can start from $P_1$ and $P_2$ and construct two solutions $\tilde{P}_1$ and $\tilde{P}_2$ starting from $(x,0)$ corresponding to $\{\tilde{a},\tilde{b}\}$. The second component is Brownian motion under both $\tilde{P}_1$ and $\tilde{P}_2$. Let us denote by $R_1$ and $R_2$ the r.c.p.d. of the first component given the second component. Let us denote by $W$ the Wiener measure. We denote points in $C[[0,\infty); R^{3d}]$ by three components $\omega_1$, $\omega_2$, $\omega_3$ and write

$$\hat{P}(d\omega_1,d\omega_2,d\omega_3) = W(d\omega_3)\; R_1(\omega_3;\, d\omega_1)\; R_2(\omega_3;\, d\omega_2)$$

In other words we make the first two components independent under $\hat{P}$ given the third component which is Brownian motion. This clearly works.

We also deduce from this that $R_1(\omega_3, d\omega_1)$ and $R_2(\omega_3, d\omega_2)$ must be degenerate distributions for almost all $\omega_3$. In other words the solution $x(\cdot)$ to Itô's equation is really a measurable functional of the Brownian path even though we did not know it to begin with.

## Section 7

We saw in a preceding section that if the equation

$$(*) \qquad \frac{\partial u}{\partial t} + L_t u = f$$

with $u(T,x) = 0$ can be solved for $0 \le t \le T$ and for every $T < \infty$ for sufficiently many functions $f$ then we can conclude uniqueness. Let us pursue the point a little further. Suppose $P_1$ and $P_2$ are two solutions starting from $(s_0, x_0)$. Let us define

$$\Lambda_i(f) = E^{P_i}\left[\int_{s_0}^{T} f(s,x(s))\, ds\right] \qquad \text{for } i = 1, 2.$$

Our aim is to conclude that $\Lambda_1 \equiv \Lambda_2$. For each $f$ for which we can solve the equation (*) with a smooth solution we can get $\Lambda_1(f) = \Lambda_2(f)$. Sometimes one can prove that the set of functions f for which (*) is solvable is a dense class in some $L_p$ space. Then one has to go back and show that $\Lambda_1$ and $\Lambda_2$ belong to the appropriate dual space so that one can still conclude that $\Lambda_1 \equiv \Lambda_2$.

We are now going to prove the uniqueness for coefficients $\{a,b\}$ such that $a$ is uniformly elliptic, a is bounded and continuous in t and x, and b is bounded and measurable. Because of the Cameron-Martin-Ginsanov formula we can assume that b is identically zero. By the localization procedure we can assume that $a$ is uniformly close to a constant matrix $a_0$, which after a linear transformation of $R^d$ can be assumed to be I. The problem of uniqueness therefore can be reduced to the following two propositions.

<u>Proposition 1</u>. Let us consider $a^{ij}(t,x) = \delta_{ij} + \varepsilon^{ij}(t,x)$ where $|\varepsilon^{ij}(t,x)| \leq \varepsilon_0$ for some $\varepsilon_0 > 0$. For a suitable choice of $\varepsilon_0$ and a suitable choice of $p_0$ for every T, the equation (*) has a smooth solution for a class of functions f which is dense in $L_{p_0}[[0,T] \times R^d]$.

<u>Proposition 2</u>. For any solution P to the martingale problem corresponding to $\{a^{ij}(t,x),0\}$ starting from any point $(s_0,x_0)$ with $0 < s_0 \leq T$, where $T < \infty$ is arbitrary but given,

$$\Lambda(f) = E^P[\int_{s_0}^{T} f(s,x(s))\, ds]$$

is a bounded linear functional on $L_{p_0}\{[0,T] \times R^d\}$. [Here $p_0$ is the same as in Proposition 1 and the coefficients also satisfy the same bounds $|a^{ij}(t,x) - \delta_{ij}| \leq \varepsilon_0$ as in Proposition 1.]

<u>Outline of Proof</u>: Let us fix an interval [0,T] and consider the

function

$$p(s,x,t,y) = \frac{1}{[2\pi(t-s)]^{d/2}} e^{-\frac{\|y-x\|^2}{2(t-s)}} \quad \text{for} \quad 0 \le s \le t \le T$$

We define an operator G acting on functions f defined on $[0,T]\times R^d$ by

$$(Gf)(s,x) = \int_s^T \int_{R^d} f(t,y)\, p(s,x,t,y)\, dt\, dy$$

Since G is a convolution and $\int_0^T \int_{R^d} p(s,x,t,y)\, dt\, dy < \infty$ we conclude that G is a bounded operator from $L_p[[0,T]\times R^d]$ into itself. Moreover $\int_0^T \int_{R^d} |p(s,x,t,y)|^\alpha\, dy\, dt$ is finite for $\alpha < (d+2)/d$. This in turn implies that

$$\|Gf\|_\infty \le C_{\alpha'} \|f\|_{\alpha'}$$

if $\alpha' > d/2 + 1$. If $Gf = u$ for some nice f then

(i) $u(s,x) \to 0$ as $s \to T$ for all x

(ii) u is smooth and

$$\frac{\partial u}{\partial s} + \frac{1}{2} \Delta u = -f$$

We have the following important bound from the theory of singular integrals.

For any $\alpha$ in the range $1 < \alpha < \infty$ there is a constant $M_\alpha$ such that

$$\left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_{L_\alpha[[0,T]\times R^d]} \le M_\alpha \|f\|_{L_\alpha[[0,T]\times R^d]}$$

We shall choose and fix a value of $\alpha$ larger than $d/2 + 1$. Let us call it $\alpha_0$. The corresponding bound $M_{\alpha_0}$ will be denoted by $M_0$.

Now consider the coefficients $a^{ij}(t,x)$ such that

$$a^{ij}(t,x) = \delta_{ij} + \varepsilon^{ij}(t,x)$$

where $|\varepsilon^{ij}(t,x)| \leq \varepsilon_0$ and $\varepsilon_0$ is to be chosen presently. We want to solve

$$\frac{\partial u}{\partial s} + \frac{1}{2} \sum a^{ij}(s,x) \frac{\partial^2 u}{\partial x_i \partial x_j} = - f$$

i.e.

$$\frac{\partial u}{\partial s} + \frac{1}{2} \Delta u + \frac{1}{2} \sum \varepsilon^{ij}(s,x) \frac{\partial^2 u}{\partial x_i \partial x_j} = - f$$

if we try for a $u$ of the form $u = Gg$ then we have to solve

$$- g + \frac{1}{2} \sum \varepsilon^{ij}(s,x) \frac{\partial^2 Gg}{\partial x_i \partial x_j} = - f$$

Denoting

$$Eg = \frac{1}{2} \sum \varepsilon^{ij}(s,x) \frac{\partial^2 Gg}{\partial x_i \partial x_j}$$

we want to solve

$$- g + Eg = - f$$

or

$$(I - E)g = f$$

The problem formally is to invert $(I - E)$. However

$$\|Eg\|_{\alpha_0} \leq \frac{d^2}{2} \varepsilon_0 M_0 \|g\|_{\alpha_0} .$$

If $\varepsilon_0$ is chosen so that $d^2 M_0 \varepsilon_0 \leq 1$, we shall have

$$\|Eg\|_{\alpha_0} \leq \frac{1}{2} \|g\|_{\alpha_0} .$$

This in turn implies that $(I - E)^{-1}$ exists as a bounded operator in $L_{\alpha_0}[[0,T] \times R^d]$ and

$$\| (I - E)^{-1} f \|_{\alpha_0} \leq 2 \| f \|_{\alpha_0}$$

Formally

$$u = G(I - E)^{-1} f$$

The set of f's for which u is nice are those for which $(I-E)^{-1}f$ is nice. We do not know what they are except that they are dense in $L_{\alpha_0}$. This proves Proposition 1.

We now turn to Proposition 2. Let P be some solution to the martingale problem corresponding to $\{a(\cdot,\cdot),0\}$ of the sort discussed above.

We can find a Brownian motion on our space such that

$$x(t) = x_0 + \int_0^t \sigma(s,x(s))\, d\beta(s) .$$

We can assume without loss of generality that the starting time is 0 and that the starting point is some $x_0 \in R^d$. Let us define

$$\sigma_n(s,\omega) = \sigma\left(\frac{j}{n}, x\left(\frac{j}{n}\right)\right) \quad \text{if } \frac{j}{n} \leq s < \frac{j+1}{n}$$

and

$$x_n(t) = x_0 + \int_0^t \sigma_n(s,\omega)\, d\beta(s) .$$

We define the linear functionals

$$\Lambda_n(f) = E^P\left[\int_0^T f(s,x_n(s))\, ds\right]$$

Since, by the properties of stochastic integrals $x_n(t) \to x(t)$

uniformly in probability as $n \to \infty$. We conclude that for each smooth f

$$\lim_{n\to\infty} \Lambda_n(f) = \Lambda(f) \ .$$

Moreover each $x_n(\cdot)$ as a stochastic process is piecewise Brownian motion. Therefore there is clearly a constant $C_n$ such that

$$|\Lambda_n(f)| \leq C_n \| f \|_{\alpha_0}$$

Our problem in completing the proof of Proposition 2 is to obtain a uniform bound on $C_n$ independent of n. Let us pick a function $u = Gg$ for some g. Then

$$u(t,x_n(t)) - \int_0^t \left(\frac{\partial u}{\partial s} + \frac{1}{2} \sum a_n^{ij}(s,\omega) \frac{\partial^2 u}{\partial x_i \partial x_j}\right)(s,x_n(s)) \ ds$$

is a martingale, where $a_n^{ij}(s,\omega) = \delta_{ij} + \varepsilon_n^{ij}(s,\omega)$ for some $\varepsilon^{ij}(\cdot,\cdot)$ with $|\varepsilon_n^{ij}(\cdot,\cdot)| \leq \varepsilon_0$ uniformly. Therefore

$$u(t,x_n(t)) + \int_0^t g(s,x_n(s)) \ ds - \frac{1}{2} \int_0^t \sum \varepsilon_n^{ij}(s,\omega) \frac{\partial^2 Gg}{\partial x_i \partial x_j} (s,x_n(s)) \ ds$$

is a $(\Omega,M_t,P)$ martingale. Equating expectations at $t = 0$ and $t = T$ and using the bounds on $\varepsilon_n^{ij}(\cdot,\cdot)$

$$|u(0,x_0)| \geq |E \int_0^T g(s,x_n(s))ds]| - \frac{1}{2} \varepsilon_0 E[\int_0^T \sum \left|\frac{\partial^2 Gg}{\partial x_i \partial x_j}\right| (s,x_n(s))ds]$$

In other words

$$|\Lambda_n(g)| \leq |u(0,x_0)| + \frac{\varepsilon_0}{2} \sum_{i,j} \Lambda_n\left(\left|\frac{\partial^2 Gg}{\partial x_i \partial x_j}\right|\right)$$

Let us denote $\partial^2 Gg/\partial x_i \partial x_j$ by $H_{ij}(s,x)$. Then

$$|\Lambda_n(g)| \leq |u(0,x)| + \frac{\varepsilon_0}{2} \sum_{i,j} \Lambda_n(|H_{ij}|)$$

Taking the supremum over all g with $\|g\|_{\alpha_0} \leq 1$ we obtain for the norm $\|\Lambda_n\|^{\alpha_0}$ in the dual space of $L_{\alpha_0}$

$$\|\Lambda_n\|^{\alpha_0} \leq C'_{\alpha_0} + \|\Lambda_n\|^{\alpha_0} \frac{1}{2} \varepsilon_0 d^2 M_{\alpha_0}.$$

Since $\|\Lambda_n\|^{\alpha_0} \leq C_n$ is finite we have

$$\|\Lambda_n\|^{\alpha_0} \leq 2C'_{\alpha_0} \qquad \text{(using } \varepsilon_0 d^2 M_{\alpha_0} \leq 1)$$

This completes the proof of Proposition 2.

## Section 8

Here we are interested in proving the convergence of the diffusion process corresponding to $\{a_n,b_n\}$ to the one corresponding to $\{a,b\}$. We shall illustrate the methods in the simplest situation. Let us suppose that $\{a_n\}$ and $\{b_n\}$ are bounded uniformly in s,x and n. We shall further suppose that $a_n$ and $b_n$ are converging to a,b uniformly on compact subsets of $[0,\infty)\times R^d$ and that $a_n$, $b_n$, a, b are all continuous functions of s and x. Finally each $a_n$ and a is assumed to be uniformly elliptic. Let $(s_n,x_n)$ be starting points converging to $(s,x)$. We denote by $P_n$ the solution for $\{a_n,b_n\}$ starting from $(s_n,x_n)$ and by P the solution for $\{a,b\}$ starting from $(s,x)$.

<u>Proposition</u>: $P_n$ converges weakly to P as $n \to \infty$.

Outline of Proof: From the fact that the coefficients $\{a_n, b_n\}$ are uniformly bounded and that $(s_n, x_n)$ varies over a bounded set we conclude that $P_n$ is weakly conditionally compact. Let $Q$ be any limit point. Without loss of generality we can assume that $P_n \Rightarrow Q$ as $n \to \infty$. If we prove that $Q$ is a solution to the martingale problem for $\{a,b\}$ starting from $(s,x)$ then by the uniqueness $Q$ must equal $P$ and we have proved the proposition. We note that for $f \in C_0^\infty(R^d)$

$$f(x(t)) - \int_{s_n}^{t} \psi_n(s,x(s))\, ds$$

is a $(\Omega, M_t, P_n)$ martingale. Here

$$\psi_n(s,x) = \frac{1}{2} \sum a_n^{ij}(s,x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum b_n^j(s,x) \frac{\partial f}{\partial x_j}(x)$$

Clearly $\psi_n(s,x) \to \psi(s,x)$ uniformly on compact sets where

$$\psi(s,x) = \frac{1}{2} \sum a^{ij}(s,x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum b^j(s,x) \frac{\partial f}{\partial x_j}(x)$$

Just as in the proof of existence we condlue that

$$f(x(t)) - \int_{s}^{t} \psi(s,x(s))\, ds$$

is a martingale relative to $(\Omega, M_t, Q)$. This completes the proof. We note of course that

$$Q[\omega: x(s) = x] = 1$$

because

$$P_n[\omega: x(s_n) = x_n] = 1 \qquad \text{for each } n.$$

One can prove by these techniques that under very general conditions certain Markov chains converge to diffusions. Let us for simplicity treat the time homogeneous case. Suppose for each $h > 0$ we have a transition probability $\pi_h(x,dy)$ on $R^d$ satisfying for each $f \in C_0^\infty(R^d)$

$$\lim_{h\to 0} \frac{1}{h} \int [f(y)-f(x)]\, \Pi_h(x,dy) = (Lf)(x)$$

uniformly on compact subsets of $R^d$, where

$$Lf = \frac{1}{2} \sum a^{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum b^j(x) \frac{\partial f}{\partial x_j}(x) \; .$$

We assume here that the coefficients are continuous and bounded on $R^d$ and that $\{a(x)\}$ is uniformly elliptic. Let us fix a starting point $x_0$ at time 0 and construct a Markov chain which at times $jh$, $j = 1,2,\ldots$ jump according to $\Pi_h(x,dy)$. We can interpolate linearly in between so that we have a measure $P_h$ on the space $\Omega$ for each $h > 0$. We want to show that

$$\lim_{h\to 0} P_h = P$$

where $P$ solves the martingale problem for $L$ starting from $x_0$ at time 0.

Sketch of proof: Let us pick a function $\phi(x)$ with $\phi(x) = 1$ for $|x| \leq 1$ and $\phi(x) = 0$ for $|x| \geq 2$ which is smooth and satisfies $0 \leq \phi(x) \leq 1$. We define $\phi_R(x) = \phi(x/R)$ and consider

$$\Pi_h^R(x,A) = \phi_R(x)\, \Pi_h(x,A) + [1-\phi_R(x)]\, \chi_A(x)$$

We define $P_h^R$ just as $P_h$ was defined relative to $\Pi_h$ and $L_R$ is the operator

$$(L_R f)(x) = \phi_R(x)(Lf)(x) \; .$$

Clearly

$$\lim_{h\to 0} \frac{\Pi_h^R f - f}{h} = L_R f \quad \text{uniformly}$$

for each $f \in C_0^\infty(R^d)$. One can now show that $\{P_h^R\}$ is relatively compact as $h \to 0$ on $\Omega$ and any limit point is a solution corresponding to $L_R$.

The compactness is established by the techniques that we have already seen. To identify the limit we note that

$$f(x(nh)) - \sum_{j=0}^{n-1} (\Pi_h^R f - f)(x(jh))$$

is a martingale relative to $(\Omega, M_{nh}, P_h)$. If we let $h \to 0$ along a subsequence so that $P_h \Rightarrow Q$ we conclude that

$$f\big(x(t)\big) - \int_0^t \big(L_R f\big)\big(x(s)\big)\, ds$$

is a martingale relative to $(\Omega, M_t, Q)$. Q must therefore necessarily agree with P on $M_{\tau_R}$ where $\tau_R$ is the exit time from the sphere of radius R. In particular

$$\limsup_{h\to 0} P_h^R[\sup_{0\le s\le T} |x(s)| \ge \ell] \le Q[\sup_{0\le s\le T} |x(s)| \ge \ell] = P[\sup_{0\le s\le T} |x(s)| \ge \ell]$$

Therefore

$$\limsup_{\ell\to\infty}\ \limsup_{R\to\infty}\ \limsup_{h\to 0} P_h^R[\sup_{0\le s\le T} |x(s)| \ge \ell] = 0$$

The last inequality implies that for large R the difference between $P_h^R$ and $P_h$ is uniformly small as $h \to 0$ on any finite time interval. Consequently one can interchange the limits as $R \to \infty$ and $h \to 0$ and conclude that $P_h$ itself converges to P as $h \to 0$.

## Section 9

We will be concerned here with the situation in which we cannot assert uniqueness. To fix ideas let us suppose that $\{a,b\}$ is bounded and continuous. We have existence but no uniqueness. For each x let $C_x$ be the set of solutions starting from x at time 0. $C_x$ perhaps consists of more than one element. We want to pick a $P_x$ from each $C_x$ such that $\{P_x\}$ is a strong Markov family. The crucial property is that for any stopping time $\tau$ the r.c.p.d. of P given $M_\tau$ is in the class $C_{x(\tau)}$ provided we reset the time $\tau$ as 0. The idea is to pick a bounded continuous function $f_1(x)$ on $R^d$, $\lambda > 0$, and consider for each x the set $C_x^1$ defined by

$$C_x^1 = \left\{P:\ P\in C_x,\ E^P[\int_0^\infty e^{-\lambda_1 t} f_1(x(t))\,dt] = \sup_{P\in C_x} [\int e^{-\lambda_1 t} f_1(x(t))dt]\right.$$

By ideas very similar to that of dynamic programming one can show that $C_x^1$ inherits from $C_x$ the property of being closed under conditioning.

We now pick $\lambda_2$ and $f_2$ and define

$$C_x^2 = \left\{P:\ P\in C_x^1;\ E^P[\int_0^\infty e^{-\lambda_2 t} f_2(x(t))dt] = \sup_{P\in C_x^1} E^P[\int_0^\infty e^{-\lambda_2 t} f_2(x(t))dt]\right\}$$

and so on. Such $C_x^n$ will have the property of being closed under conditioning. If we go through $(\lambda_j, f_j)$ which is dense among all pairs $(\lambda, f)$ then denoting by $D_x$ the intersection $\bigcap_n C_x^n$ we see that $D_x$ is closed under conditioning and furthermore if $P_1$ and $P_2$ are in $D_x$,

$$E^{P_1}[\int_0^\infty e^{-\lambda t} f(x(t))\,dt] = E^{P_2}[\int_0^\infty e^{-\lambda t} f(x(t))\,dt\,],$$

for all  and f. This means that

$$P_1[x(t) \in A] \equiv P_2[x(t) \in A] \quad \text{for all } t \geq 0 \text{ and } A \in \mathcal{B}(R^d).$$

By the method through which we proved uniqueness this in turn implies that each $D_x$ consists only of a single element $P_x$ and they of course automatically form a strong Markov family.

There is also a natural converse in the sense that starting from all strong Markov families $\{P_x\}$ and mixing them up one can recover the collection $C_x$.

In other words any nonuniqueness of solutions to the martingale problem arises from nonuniqueness of the Markov semigroups whose infinitesimal generators are extensions of L from smooth functions.

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# WAVE PROPAGATION AND HEAT CONDUCTION IN A RANDOM MEDIUM

G. C. PAPANICOLAOU

# Wave Propagation and Heat Conduction in a Random Medium

G. C. Papanicolaou
Courant Institute, New York University

## INTRODUCTION

We shall give a fairly selfcontained account of some results on waves in random media and related problems that we have considered in the past few years [1]-[6]. These results rely upon properties of solutions of differential equations with random coefficients, i.e., stochastic equations. We restrict attention to one-dimensional problems so that we are dealing with stochastic ordinary differential equations. There are a few results, at present, dealing with multidimensional problems at [cf. 12] but we shall not discuss these here.

## 1. FORMULATION OF THE PROBLEMS

We consider a one-dimensional medium occupying the interval $[0,L]$ with a wave of unit amplitude incident from $x<0$. Let $u(x)\exp\{-i\omega t\}$ denote the complex-valued wave field at location $x$ at time $t$. The time factor will be omitted as is customary. The field $u(x)$ satisfies the one-dimensional reduced wave equation

$$\frac{d^2u(x)}{dx^2} + k^2n^2(x)\,u(x) = 0, \qquad 0 < x < L. \tag{1.1}$$

Here $n(x)$ is the index of refraction, $k = \omega/c$ is the wave number and $c$ is the free space propagation speed. The index of refraction $n(x)$ is a random process with known statistical properties to be described below.

A wave of unit amplitude is incident from the left which is free space. Therefore,

$$u(x) = e^{ikx} + R\, e^{-ikx}, \qquad x < 0, \tag{1.2}$$

where $R = R(L,k)$ is the reflection coefficient. It is a complex-valued random variable with $|R| \le 1$. The region to the right of $[0,L]$ is also free space so that the transmitted wave is

$$u(x) = T\, e^{ikx} \tag{1.3}$$

where $T = T(L,k)$ is the transmission coefficient.

Equation (1.1) for $u(x)$ in $0 < x < L$ is supplemented by requiring that $u(x)$ and $du(x)/dx$ be continuous at $x = 0$ and $x = L$. This yields the two point boundary conditions

$$\frac{1}{2}\left[u(x) + \frac{1}{ik}\frac{du(x)}{dx}\right]_{x=0} = 1 \tag{1.4}$$

$$\frac{1}{2}\left[u(x) - \frac{1}{ik}\frac{du(x)}{dx}\right]_{x=L} = 0\,. \tag{1.5}$$

Equations (1.1), (1.4) and (1.5) determine $u(x)$ completely. Then $R$ and $T$ are given by

$$R(L,k) = \frac{1}{2}\left[u(x) - \frac{1}{ik}\frac{du(x)}{dx}\right]_{x=0} \tag{1.6}$$

$$T(L,k) = \frac{1}{2}\, e^{-ikx}\left[u(x) + \frac{1}{ik}\frac{du(x)}{dx}\right]_{x=L}, \tag{1.7}$$

where $u(x) = u(x;L,k)$ but we suppress dependence on $L$ and $k$. Note that we have the conservation relation

$$|T|^2 + |R|^2 = 1 \tag{1.8}$$

which says that the wave energy per unit time transmitted through $[0,L]$ plus the wave energy per unit time reflected equals the incident energy per unit time which is normalized to one.

We shall now describe the class of random indices of refraction $n(x)$ which we will consider. We assume that

$$n^2(x) = 1 + g\big(y(x)\big) \tag{1.9}$$

where $y(x)$, $x \geq 0$, is a Markov process on a state space $S$ which is a compact metric space and $g(y)$ is a continuous function from $S$ to $[-1/2,1/2]$, i.e. $|g(y(x))| \leq 1/2$ for all $x \geq 0$. The Markov process $y(x)$ is assumed to be ergodic (more technical assumptions are introduced later) with the initial value $y(0)$ distributed according to the invariant distribution so that $\{y(x), x \geq 0\}$ is also a stationary process. We shall denote by $E\{\cdot\}$ expectation relative to this stationary Markov process and we shall assume that

$$E\Big\{g\big(y(x)\big)\Big\} = 0 \ , \quad E\Big\{g^2\big(y(x)\big)\Big\} = \alpha^2 > 0 \tag{1.10}$$

and

$$\frac{1}{\alpha^2}\int_0^\infty E\Big\{g\big(y(x)\big)\ g\big(y(0)\big)\Big\}\, dx = \ell > 0 \ . \tag{1.11}$$

The dimensionless variable $\alpha$ is a measure of the size of the fluctuations in the refractive index while the parameter $\ell$ with the dimensions of length is a measure of the amount of correlation in the fluctuations at different points.

The problem at hand has three natural dimensionless parameters: $\alpha$, $kL$ and $k\ell$. The quantity of principal interest to us, the transmission coefficient $T$, is thus a functional of process $y(\cdot)$ and a function of the parameters $L$, $k$, $\ell$ and $\alpha$, i.e.,

$$T = T(L,k,\ell,\alpha) \ , \tag{1.12}$$

with $\alpha$, $kL$ and $k\ell$ dimensionless parameters.

One can study the statistical properties of the transmission coefficient $T$

in at least four interesting asymptotic limits which we now enumerate.

(i) Rapid fluctuation limit (law of large numbers). This means that $\ell$ is small while all other parameters are fixed. Thus, we replace $\ell$ by $\varepsilon^2\ell$ with $\varepsilon \to 0$. It is easily seen that $T(L,k,\varepsilon^2\ell,\alpha) \to 1$ with probability one and this is not difficult to prove as will be explained in the next section. If the index of refraction has the form $n^2(x) = \zeta^2(x)\bigl(1 + g(y(x))\bigr)$, instead of (1.9), with $\zeta(x)$ a deterministic refractive index then $T(L,k,\varepsilon^2\ell,\alpha) \to$ Deterministic transmission coefficient corresponding to $\zeta(x)$. Returning to the case $T \to 1$ one may now analyze the limiting distribution of $\varepsilon^{-1}[T(L,k,\varepsilon^2\ell,\alpha) - 1]$ as $\varepsilon \to 0$ and show that it is a complex-valued Gaussian distribution. This is also easy to show and it is discussed further in the next section.

(ii) White noise limit. This means that we replace $\alpha$ by $\alpha/\varepsilon$ and $\ell$ by $\varepsilon^2\ell$ and let $\varepsilon \to 0$. We shall see later that $T(L,k,\varepsilon^2\ell,\alpha/\varepsilon)$ converges weakly as $\varepsilon \to 0$ to a random variable whose distribution can be obtained explicitly, in principle. However, it is very difficult to extract physically useful information from the limit since the necessary calculations are prohibitively complex. Physically, the white noise limit is just what the words mean: if $g^\varepsilon(x)$ denotes the scaled random process $g\bigl(y(x)\bigr)$ then its correlation $E\{g^\varepsilon(x + x')\, g^\varepsilon(x')\}$ tends to $2\,\alpha^2\ell\;\delta(x)$ as $\varepsilon \to 0$ where $\delta(x)$ is the Dirac delta function.

(iii) Weak fluctuations -- large scattering region. Here we replace $\alpha$ by $\alpha\varepsilon$ and $L$ by $L/\varepsilon^2$. It is easy to see that $T(L/\varepsilon^2,k,\ell,\varepsilon\alpha) = T(L,k/\varepsilon^2,\varepsilon^2\ell,\varepsilon\alpha)$ $= T(L/\varepsilon,k/\varepsilon,\varepsilon\ell,\varepsilon\alpha)$ and hence the limit could have been called the weak fluctuations -- high frequency -- short correlation limit. The important thing to notice here is that the dimensionless parameter $k\ell$ is independent of $\varepsilon$. This means that the wave length of the incident wave $\lambda = 2\pi/k$ is comparable to the correlation length and both are small compared to the size of the scattering region (and the fluctuations are also small). This is the asymptotic limit of basic physical interest for waves in random media. It turns out, fortunately that practically everything one wants about $T$ can be computed explicitly in this asymptotic limit [1]-[2].

(iv) Large scattering region - Large wave lengths. Here we replace L by $L/\varepsilon^2$ and k by $\varepsilon k$ and note that $T(L/\varepsilon^2,\varepsilon k,\ell,\alpha) = T(L,k/\varepsilon,\varepsilon^2\ell,\alpha) = T(L/\varepsilon,k,\varepsilon\ell,\alpha)$ which means that this limit could have been called the small wave length -- small correlation length limit with the correlation length much smaller than the wave length. Note that we do not assume weak fluctuations here (just as we did not assume weak fluctuations in case (i)). The present scaling is of interest in the heat conduction problem [4], [7] which is introduced below.

Having enumerated several intersting asymptotic limits we now pose the following problems:

Problem I: Show that $T(L/\varepsilon^2,k,\ell,\varepsilon\alpha)$ has a limiting distribution as $\varepsilon \to 0$ and compute this distribution explicitly as a function of L, k, $\ell$ and $\alpha$.

Problem II: Same as Problem I, for $T(L/\varepsilon^2,\varepsilon k,\ell,\alpha)$.

It turns out, not so surprisingly, that the limiting distributions in I and II are in fact the same but their parametric dependence on k and $\ell$ is different. We return to Problems I and II in Sections 3 and 4 respectively.

We shall introduce the heat conduction problem next. One can also give a less phenomenological derivation than the one below for the quantity of interest [14].

Suppose that the incident wave from the left is not monochromatic but a general time dependent pulse so that for $x < 0$

$$u(t,x) = \int_{-\infty}^{\infty} e^{-i\omega t}[e^{ikx} + R\, e^{-ikx}]\, G(d\omega) \; . \tag{1.13}$$

Here $G(d\omega)$ is a complex-valued measure on $R^1$ such that $G^*(-d\omega) = G(d\omega)$, star being complex conjugation. We assume that G is statistically independent of the fluctuation process $y(x)$, that $\langle G(d\omega)\rangle = 0$, and that

$$\langle G(d\omega)\, G^*(-d\omega')\rangle = \theta(\omega)\ \delta(\omega+\omega')\ d\omega\ d\omega' . \tag{1.14}$$

We use angular brackets $\langle\ \rangle$ to denote expectation involving G which is distinct from $E\{\cdot\}$. Thus we assume that the wave incident from the left is a stationary random function of time, statistically independent of the scattering medium and with power spectral density $\theta(\omega)$.

By (1.3) and linearity, the transmitted wave is

$$u(t,x) = \int_{-\infty}^{\infty} e^{-i\omega t}\, e^{ikx}\, T(L,k)\ G(d\omega)\ , \qquad x \geq L . \tag{1.15}$$

Since $k = \omega/c$ this is the same as

$$u(t,x) = \int_{-\infty}^{\infty} e^{-i\omega(t-x/c)}\ T\left(L, \frac{\omega}{c}\right) G(d\omega)\ , \qquad x \geq L \tag{1.16}$$

and this is a real-valued process. From (1.14), (1.16) and the identity $T^*(L,k) = T(L,-k)$, it follows that

$$\langle u(t+s,x)\ u(t,x)\rangle = \int_{-\infty}^{\infty} e^{-i\omega s}\left|T\left(L, \frac{\omega}{c}\right)\right|^2 \theta(\omega)\ d\omega, \qquad s \geq 0 . \tag{1.17}$$

The quantity on the left is the time correlation function of the transmitted wave with time lag s. We see that it is independent of $t \geq 0$ and $x \geq L$. Of particular interest is the variance of average transmitted wave defined by

$$J_\theta(L) = \int_{-\infty}^{\infty} E\left\{\left|T\left(L, \frac{\omega}{c}\right)\right|^2\right\} \theta(\omega)\ d\omega . \tag{1.18}$$

Since both $|T(L,\omega/c)|^2$ and $\theta(\omega)$ are even functions of $\omega$, we have

$$J_\theta(L) = 2c\int_0^{\infty} E\left\{|T(L,k)|^2\right\} \theta(kc)\ dk . \tag{1.19}$$

To simulate a heat bath at temperature equal to $\theta_0$ on the left $(x \leq 0)$ we shall take $\theta(kc) \equiv \theta_0 > 0$ in $[0,1]$ and zero for $kc > 1$. Only the behavior of $\theta$ near $k = 0$ matters as we shall see. Since no waves impinge from the right, the medium on the right is at temperature zero. Thus we define the average rate of heat conduction by

$$J(L) = 2c\theta_0 \int_0^1 E\left\{|T(L,k)|^2\right\} dk \ . \tag{1.20}$$

Problem III: Determine the asymptotic behavior of $J(L)$ as $L \to \infty$.

We return to this problem in Section 4 and show [4] that $J(L) \sim L^{-1/2}$ as $L \to \infty$.

There are many other interesting problems one can pose about the behavior of $T(L,k,\ell,\alpha)$ or even the wave amplitude $u(x;L,k,\ell,\alpha)$ at interior points $0 < x < L$. Some of them are discussed in the references (and the many more references to be found therein).

One can show [8], [9] by continuous-time analogs of Furssenberg's theorems [10] that $|T(L,k,\ell,\alpha)|^2 \to 0$ as $L \to \infty$ with probability one. In fact, $E\{|T(L,k,\ell,\alpha)|^2\}$ decays exponentially to zero as $L \to \infty$, as is shown in [8]. The exact constant of exponential decay is not known, however. We return to this question in Section 6.

## 2. LIMIT THEOREMS FOR STOCHASTIC EQUATIONS

The approach to Problems I, II, III of the previous section that we shall follow is this. Consider the wave field $u(x;L,k)$ satisfying (1.1) subject to (1.4) and (1.5). Define the forward and backward traveling wave amplitudes A and B by (cf. [1])

$$A(x,k) = \frac{1}{2} e^{-ikx} [u(x) - \frac{i}{k} \frac{du(x)}{dx}] \tag{2.1}$$

$$B(x,k) = \frac{1}{2} e^{ikx} [u(x) + \frac{i}{k} \frac{du(x)}{dx}] \ .$$

Then (1.1), (1.4) and (1.5) yield the following equations for A and B.

$$\frac{d}{dx} \begin{pmatrix} A(x,k) \\ B(x,k) \end{pmatrix} = \frac{ikg(y(x))}{2} \begin{pmatrix} 1 & e^{-2ikx} \\ -e^{2ikx} & -1 \end{pmatrix} \begin{pmatrix} A(x,k) \\ B(x,k) \end{pmatrix} , \quad 0 < x < L, \tag{2.2}$$

with

$$(2.3) \qquad A(0,k) = 1 \ , \qquad B(L,k) = 0 \ .$$

Furthermore, (1.6) and (1.7) become

$$(2.4) \qquad B(0,k) = R(L,k), \ A(L,k) = T(L,k) \ .$$

Now let

$$(2.5) \qquad m(x;k) = \frac{ikg(y(x))}{2} \begin{pmatrix} 1 & e^{-2ikx} \\ -e^{2ikx} & -1 \end{pmatrix}$$

and let $Y(x;k)$ be the $2 \times 2$ fundamental solution matrix of (2.2) i.e.

$$(2.6) \qquad \frac{d}{dx} Y(x;k) = m(x;k) \ Y(x;k) \ , \qquad Y(0;k) = I.$$

From the structure of m in (2.5) we see that Y has the form

$$(2.7) \qquad Y = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} , \qquad |a|^2 - |b|^2 = 1,$$

i.e., $Y(x;k)$, $x \geq 0$ is a random process with values in the group $SU(1,1)$ and k is a parameter. It also follows from (2.3), (2.4) and (2.7) that $T(L;k)$ $(= T(L,k,\ell,\alpha))$ is a function of $a(L;k)$, $b(L;k)$ and in particular

$$(2.8) \qquad |T(L;k)|^2 = \frac{1}{|a(L;k)|^2} \ .$$

Therefore the problems of the previous section reduce to the determination of the asymptotic distribution of $Y(L;k)$ $(= Y(L;k,\ell,\alpha))$, under various scalings. But Y is the solution of the stochastic <u>initial</u> value problem (2.6). So we need some theorems regarding the asymptotic behavior of stochastic ODE's depending on a small parameter.

Following [5] we shall describe next some results concerning asymptotics for stochastic ODE's. In Section 6 we shall prove in some detail one particular result; the others follow in much the same way.

First we consider initial value problems scaled as in Case (i) in Section 1. We shall denote by $t \geq 0$ the independent variable, by $y^{\varepsilon}(t)$ the scaled fluctuating coefficients and by $x^{\varepsilon}(t)$ the scaled random process which is the solution of

$$\frac{dx^{\varepsilon}(t)}{dt} = F\left(x^{\varepsilon}(t), y^{\varepsilon}(t), \varepsilon\right), \qquad x^{\varepsilon}(0) = x. \tag{2.9}$$

Here $F(x,y,\varepsilon)$ is a given function from $R^n \times S \times (0,1]$ to $R^n$ so that $x^{\varepsilon}(t)$ is $R^n$-valued. Recall that the fluctuation process $y^{\varepsilon}(t)$ takes values in some compact metric space S (the space of coefficients). The discussion here will be informal to avoid detailed listing of regularity and other hypotheses necessary for a complete treatment.

Suppose that $\{y(t),\ t \geq 0\}$ is a given ergodic Markov process on S (more details in Section 6) and that $y^{\varepsilon}(t) = y(t/\varepsilon^2)$. Suppose also that $F = F(x,y)$ and with x fixed put

$$\bar{F}(x) = E\{F(x, y(t))\}. \tag{2.10}$$

Here and in the sequel $E\{\cdot\}$ refers to expectation relative to the probability measure of the ergodic Markov process $\{y(t),\ t \geq 0\}$ regarded as a stationary process, i.e. with $y(0)$ distributed according to the invariant measure of the process.

Let $\bar{x}(t)$ be the solution of

$$\frac{d\bar{x}(t)}{dt} = \bar{F}(\bar{x}(t)), \qquad \bar{x}(0) = x. \tag{2.11}$$

It is reasonable to expect that as $\varepsilon \to 0$, $x^{\varepsilon}(t) \to \bar{x}(t)$ in probability. In fact it is not hard to show that [2,13,15] for each $\delta > 0$ and $T < \infty$,

$$P\left\{\sup_{0 \leq t \leq T} \left|x^{\varepsilon}(t) - \bar{x}(t)\right| > \delta\right\} \to 0 \quad \text{as} \quad \varepsilon \to 0. \tag{2.12}$$

The reader can verify easily that this result covers Case (i) of the previous section.

One can now consider the behavior of the fluctuation process $\varepsilon^{-1}\left(x^{\varepsilon}(t)-\bar{x}(t)\right)$. Again one can show [2,13,15] that this process converges weakly as $\varepsilon \to 0$ to a Gaussian Markov process whose infinitesimal parameters are easily identified. Since we shall not use this result, we shall not discuss it further (cf. also [17]).

Next we consider scaling corresponding to Case (ii). Now $y^{\varepsilon}(t) = y(t/\varepsilon^2)$ as above, but

$$F(x,y,\varepsilon) = \varepsilon^{-1}F^{(1)}(x,y) + F^{(2)}(x,y) + O(\varepsilon) \tag{2.13}$$

in (2.9) and

$$E\{F^{(1)}(x,y(t))\} \equiv 0 \ . \tag{2.14}$$

The asymptotic behavior of $x^{\varepsilon}(t)$ in this case was treated first by Khasminskii [16]. It is also treated in [5] and in several other places referred to in [1] and [2]. The result is that $x^{\varepsilon}(t)$ converges weakly to the diffusion Markov process $x(t)$ as $\varepsilon \to 0$ and the infinitesimal generator of this process is given by

$$\begin{aligned} L\, f(x) = \int_0^{\infty} E\left\{F^{(1)}(x,y(0))\cdot \frac{\partial}{\partial x}\left(F^{(1)}(x,y(s))\cdot \frac{\partial f(x)}{\partial x}\right)\right\} ds \\ + E\left\{F^{(2)}(x,y(0))\cdot \frac{\partial f(x)}{\partial x}\right\}. \end{aligned} \tag{2.15}$$

Case (iii) corresponds to $y^{\varepsilon}(t) = y(t/\varepsilon^2)$ as above but now $F(x,y,\varepsilon)$ in (2.9) depends also explicitly on $t/\varepsilon^2$ i.e.

$$F = \varepsilon^{-1} F^{(1)}(x,y,t/\varepsilon^2) + F^{(2)}(x,y,t/\varepsilon^2) + O(\varepsilon) \ , \tag{2.16}$$

and again

$$E\{F^{(1)}(x,y,\tau)\} \equiv 0 \ . \tag{2.17}$$

The asymptotic behavior of $x^{\varepsilon}(t)$ in this case was also obtained in [16] and the result is that $x^{\varepsilon}(t)$ converges weakly to the diffusion process $x(t)$ as $\varepsilon \to 0$ whose infinitesimal generator is given by

$$(2.18)\quad L\, f(x) = \lim_{T\uparrow\infty} \frac{1}{T}\int_0^T dt \int_0^\infty ds\; E\left\{F^{(1)}(x,y(0),t)\cdot \frac{\partial}{\partial x}\left(F^{(1)}(x,y(s),t+s)\frac{\partial f(x)}{\partial x}\right)\right\}$$
$$+ \lim_{T\uparrow\infty} \frac{1}{T}\int_0^T dt\; E\left\{F^{(2)}(x,y(t),t)\cdot \frac{\partial f(x)}{\partial x}\right\}.$$

This result is also proven in [5] and formula (2.18) for the limit generator is the basic tool for the computations carried out in [1]-[3].

Case (iv) differs from Case (iii) only insofar as instead of (2.16) we have

$$(2.19)\qquad F = \varepsilon^{-1}F^{(1)}(x,y,t/\varepsilon) + F^{(2)}(x,y,t/\varepsilon) + O(\varepsilon).$$

The generator of the limiting diffusion process is given now by

$$(2.20)\quad Lf(x) = \lim_{T\uparrow\infty} \frac{1}{T}\int_0^T dt \int_0^\infty ds\; E\left\{F^{(1)}(x,y(0),t)\cdot \frac{\partial}{\partial x}\left(F^{(1)}(x,y(s),t)\cdot \frac{\partial f(x)}{\partial x}\right)\right\}$$

This is the formula used in the heat conduction problem [4]. In Section 5 we shall give a brief proof of this result following [5].

In addition to the above limit theorems in which $\varepsilon \to 0$ and $t$ stays fixed in some finite interval $0 \leq t \leq T < \infty$, one may also consider what happens to $x^{\varepsilon}(t)$ as $t \to \infty$ while $\varepsilon > 0$ is fixed but small. Some results regarding this question are given in [6].

Another type of question that can be asked is this: In case (i) where (2.12) holds, suppose $U$ is a set of trajectories that does not contain the trajectory $x(t)$, $0 \leq t \leq T$. What is the probability that $x^{\varepsilon}(\cdot) \in U$, for $\varepsilon$ small? This probability is exponentially small and the exponential constant can be computed explicitly [18].

## 3. APPLICATION TO THE TRANSMISSION COEFFICIENT

Formula (2.18) can be applied to the matrix-valued process $Y^{\varepsilon}(x;k)$ which is the scaled version (Case (iii)) of (2.6)

$$(3.1) \quad \frac{dY^{\varepsilon}}{dx} = \frac{1}{\varepsilon}\,\frac{ikg(y(x/\varepsilon^2))}{2}\begin{pmatrix} 1 & e^{-2ikx/\varepsilon^2} \\ -e^{2ikx/\varepsilon^2} & -1 \end{pmatrix} Y^{\varepsilon}, \qquad Y^{\varepsilon}(0,k) = I.$$

We find that $Y^{\varepsilon}(x;k)$ converges weakly as $\varepsilon \to 0$ to a diffusion Markov process with values in SU(1,1) and it remains to find its generator explicitly [1].

The calculations for this are straightforward and in addition (cf. [1]) one can compute the limiting distribution function of the transmission coefficient explicitly. One also finds that

$$(3.2) \quad \lim_{\varepsilon\downarrow 0} E\left\{\left|T(L/\varepsilon^2,k,\ell,\varepsilon)\right|^2\right\} = 2\pi\int_0^{\infty} e^{-\gamma L(t^2+1/4)}\,\frac{t\sinh \pi t}{\cosh^2 \pi t}\,dt$$

where $\gamma = \gamma(k)$ is given by

$$(3.3) \quad \gamma(k) = \frac{k^2}{2}\int_0^{\infty} E\left\{g\big(y(s)\big)\; g\big(y(0)\big)\right\}\cos 2ks\; ds$$

and should be thought of as the proper form of $\ell$ here since $\ell$ does not appear in the result (3.2). In [1] and [2] we also give references to other papers where formula (3.2) is derived.

Formula (2.19) can be applied to the matrix valued process $Y^{\varepsilon}(x;k)$ which is now taken to be the scaled version of (2.6) under case (iv):

$$(3.4) \quad \frac{d}{dx}Y^{\varepsilon} = \frac{1}{\varepsilon}\,\frac{ikg\big(y(x/\varepsilon^2)\big)}{2}\begin{pmatrix} 1 & e^{-2ikx/\varepsilon} \\ -e^{-2ikx/\varepsilon} & 1 \end{pmatrix} Y^{\varepsilon}, \quad Y^{\varepsilon}(0,k) = I.$$

Now the generator of the limiting process has the same form as that of (3.1) (cf. (4.11) in [1], Part I) but the coefficients change (as explained in [4] also). Formula (3.2) is valid again except that now $\gamma = \gamma(k)$ is given by

$$(3.5) \qquad \gamma(k) = \frac{k^2}{2} \int_0^\infty E\left\{g\big(y(s)\big)\ g\big(y(0)\big)\right\} ds$$

which is in fact $k^2\ell/2$.

## 4. APPLICATION TO THE HEAT CONDUCTION PROBLEM

We rewrite the average rate of heat conduction (1.20) in the form

$$(4.1) \qquad \sqrt{L}\, J(L) = 2c\theta_0 \int_0^{\sqrt{L}} E\left\{\left|T(L, \frac{k}{\sqrt{L}})\right|^2\right\} dk$$

If we set $L = \varepsilon^{-2}$ and use the results of the previous section we see that

$$(4.2) \qquad \lim_{\varepsilon \downarrow 0} E\left\{\left|T(\varepsilon^{-2}, \varepsilon k, \ell, 1)\right|^2\right\} = 2\pi \int_0^\infty e^{-\gamma(t^2+1/4)} \frac{t \sinh \pi t}{\cosh^2 \pi t}\, dt$$

where $\gamma(k)$ is given by (3.5) i.e. $\gamma = k^2\ell/2$. Now we take the limit $L \to \infty$ in (4.1) and use (4.2) assuming that we can take the limit inside the integral. We obtain

$$(4.3) \qquad \lim_{L \uparrow \infty} \sqrt{L}\, J(L) = 2c\theta_0 (2\pi) \int_0^\infty \int_0^\infty e^{-k^2\ell(t^2+1/4)/2} \frac{t \sinh \pi t}{\cosh^2 \pi t}\, dt\, dk$$

$$= \frac{c\theta_0 (2\pi)^{3/2}}{\sqrt{\ell}} \int_0^\infty (t^2 + \tfrac{1}{4})^{-1/2} \frac{t \sinh \pi t}{\cosh^2 \pi t}\, dt\ .$$

This settles problem III of Section 1.

It remains to verify that one can take the limit $L \to \infty$ inside the integral sign in (4.1). This requires special considerations that are due largely to Pastur and Feldman [8] and are considered in Section 6.

## 5. PROOF OF A LIMIT THEOREM

We shall now give a brief argument that shows how the result leading to (2.19) is obtained.

The process $x^\varepsilon(t)$ under consideration is assumed to satisfy

$$\frac{dx^\varepsilon(t)}{dt} = \frac{1}{\varepsilon} F\left(x^\varepsilon(t), y^\varepsilon(t), t/\varepsilon\right), \qquad x^\varepsilon(0) = x \tag{5.1}$$

where $F(x,y,\tau)$ is a function on $R^n \times S \times [0,\infty)$ with values in $R^n$ and $y^\varepsilon(t) \equiv y(t/\varepsilon^2)$ where $\{y(t),\ t \geq 0\}$ is an ergodic Markov process with values in S which is a compact metric space. We shall assume that F is a differentiable function of x and $\tau$ and continuous in y. Regarding $y(t)$, $t \geq 0$, we assume that it is a right continuous Markov process whose infinitesimal generator

$$Qf(y) = \lim_{h\downarrow 0} \frac{1}{h} \left[ E_y\left\{f\left(y(h)\right)\right\} - f(y) \right] \tag{5.2}$$

is a bounded operator on $C(S)$. Here $E_y\{\cdot\}$ is expectation with respect to the probability $P_y$ on trajectories in S that start from the point y.

We have assumed that $\{y(t),\ t \geq 0\}$ is ergodic. Actually we need that this be so with an exponential approach rate so that the operator $-Q^{-1}$ is well defined and bounded on functions which average to zero with respect to the invariant measure. In other words we assume that the null space of Q is one-dimensional, spanned by the onstant functions, and that the Fredholm alternative is valid.

The process $\left(x^\varepsilon(t), y^\varepsilon(t)\right)$ is a Markov process with values in $R^n \times S$. Let $L_t^\varepsilon$ denote the infinitesimal generator of this process. It is easily seen that

$$L_t^\varepsilon f(x,y) = \frac{1}{\varepsilon^2} Qf(x,y) + \frac{1}{\varepsilon} F(x,y,t/\varepsilon) \cdot \frac{\partial f(x,y)}{\partial x}. \tag{5.3}$$

Note that the infinitesimal generator depends on t so that the process is not time homogeneous. The pobbability measures $P_{x,y}^\varepsilon$ on trajectories in $R^n \times S$ starting from $(x,y)$ can be characterized by the martingale approach of Stroock and Varadhan: it is that probability measure $D([0,\infty), R^n \times S)$ (Skorohod space) for which

$$f\left(x(t), y(t), t\right) - \int_0^t \left(L_s^\varepsilon + \frac{\partial}{\partial s}\right) f\left(x(s), y(s), s\right) ds \tag{5.4}$$

is a martingale for each $f(x,y,t)$ which is bounded and has bounded x and s derivatives.

For the asymptotic analysis we require that

$$\int_S F(x,y,\tau)\ \bar{P}(dy) \equiv 0 \tag{5.5}$$

where $\bar{P}(dy)$ is the invariant measure of $\{y(t),\ t \geq 0\}$. This is hypothesis (2.17). Let $\psi(y,dz)$ be the kernel of $-Q^{-1}$ so that

$$-\ Q^{-1}h(y) = \int_S \psi(y,dz)\ h(z) \tag{5.6}$$

where h is bounded and

$$\int h(y)\ \bar{P}(dy) = 0\ . \tag{5.7}$$

Let $f(x)$ be a fixed function with bounded and continuous second derivatives. Define the operator

$$\bar{L}_\tau f(x) = \int_S \int_S \bar{P}(dy)\ \psi(y,dz)\ F(x,y,\tau) \cdot \frac{\partial}{\partial x} \left(F(x,z,\tau) \cdot \frac{\partial f(x)}{\partial x}\right) \tag{5.8}$$

We shall assume that for each f fixed

$$\mathcal{L}\ f(x) = \lim_{T\uparrow\infty} \int_0^T ds\ \bar{L}_{\tau+s}\ f(x) \tag{5.9}$$

exists uniformly in x and $\tau$ and is independent of $\tau$ as shown. It can be verified easily that (5.9) and (2.19) are identical.

Now for the limit theorem we proceed as follows. Fix $f(x)$ smooth (say $C^\infty$ with compact support) and let for $\varepsilon > 0$, $\lambda > 0$,

$$f^{(\varepsilon,\lambda)}(x,y,\tau) = f(x) + \varepsilon\ \psi_f^{(1)}(x,y,\tau) + \varepsilon\tilde{\psi}_f^{(1,\lambda)}(x,\tau) + \varepsilon^2\psi_f^{(2,\lambda)}(x,y,\tau) \tag{5.10}$$

where

$$\psi_f^{(1)}(x,y,\tau) = \int \psi(y,dz)\ F(x,z,\tau) \cdot \frac{\partial f(x)}{\partial x} \tag{5.11}$$

$$(5.12) \qquad \tilde{\psi}_f^{(1,\lambda)}(x,\tau) = \int_0^\infty e^{-\lambda s}[\bar{L}_{\tau+s} f(x) - Lf(x)]\, ds$$

and

$$(5.13) \quad \psi^{(2,\lambda)}(x,y,\tau) = \int \psi(y,dz) \left[F(x,z,\tau)\cdot \frac{\partial \psi_f^{(1)}(x,z,\tau)}{\partial x} - \bar{L}_\tau f(x) + \frac{\partial \psi_f^{(1)}}{\partial \tau} + F(x,z,\tau)\cdot \frac{\partial \tilde{\psi}^{(1,\lambda)}(x,\tau)}{\partial x}\right]$$

Then,

$$(5.14) \qquad \left(L_t^\varepsilon + \frac{\partial}{\partial t}\right) f^{(\varepsilon,\lambda)}\left(x,y,\frac{t}{\varepsilon}\right) = Lf(x) + \lambda\tilde{\psi}_f^{(1,\lambda)}\left(x,\frac{t}{\varepsilon}\right) + O(\varepsilon) \; .$$

In fact, the point of the above constructions is precisely to obtain (5.14) which is a formalized perturbation theory.

We return to (5.4) and put for f the function $f^{(\varepsilon,\lambda)}(x,y,t/\varepsilon)$. We see then from (5.10) and (5.14) that

$$(5.15) \quad f(x^\varepsilon(t)) - \int_0^t Lf(x^\varepsilon(s))\, ds = -\lambda\, \tilde{\psi}^{(1,\lambda)}\left(x^\varepsilon(t), \frac{t}{\varepsilon}\right) + O(\varepsilon) + \text{Martingale}.$$

Assuming that we have shown weak compactness for the process $x^\varepsilon(\cdot)$ (which is not difficult to show [5]) then we can pass to the limit in (5.15) along a convergent subsequence. Because of (5.9),

$$\limsup_{\lambda \downarrow 0} \sup_{x,\tau} \lambda\, \tilde{\psi}^{(1,\lambda)}(x,\tau) = 0$$

and hence we conclude that for any limit of the process $x^\varepsilon(\cdot)$ the expression

$$f\big(x(t)\big) - \int_0^t Lf\big(x(s)\big)\, ds$$

is a martingale. Since $L$ is a diffusion operator with smooth coefficients this martingale problem has a unique solution. Then $x^\varepsilon(\cdot)$ converges weakly to the diffusion process generated by $L$.

## 6. THE INSTABILITY OF THE HARMONIC OSCILLATOR

For the result (4.3) to hold we need the following estimate [8] which allows interchange of limit and integration.

There is a constant C independent of k and a positive function z(k) for $k > 0$, such that

(6.1) $$E\left\{\left|T\left(L, \frac{k}{\sqrt{L}}\right)\right|^2\right\} \le C\, e^{-z(k)L} , \qquad L \ge 0.$$

Moreover, $z(k) \to 0$ as $k \to 0$ but there is a constant $z_0 > 0$ such that

(6.2) $$\lim_{k \downarrow 0} \frac{1}{k^2} z(k) = z_0 .$$

As is easily seen from [8] and elsewhere, the estimate (6.1) quickly reduces to the following problem. Consider the <u>initial</u> value problem for (1.1) and introduce polar coordinates

(6.3) $$u = e^r \cos\theta , \qquad \frac{du}{dx} = -k\, e^r \sin\theta .$$

Then $(r(x), \theta(x))$ are solutions of the system

(6.4) $$\frac{dr}{dx} = \frac{k}{2}\, g\big(y(x)\big) \sin 2\theta(x)$$

$$\frac{d\theta}{dx} = k\left(1 + g\big(y(x)\big)\cos^2\theta(x)\right)$$

and we shall take $r(0) = 0$ in the sequel. It is easily seen that

(6.5) $$\left|T\left(L, \frac{k}{\sqrt{L}}\right)\right|^2 \le \frac{4}{2 + e^{2r(L)}}$$

where of course $r(L) = r(L;k)$. Thus, (6.1) is implied by

(6.6) $$E\left\{\frac{4}{2 + e^{2r(L)}}\right\} \le C\, e^{-z(k)L} , \qquad L \ge 0 ,$$

which is what we shall prove.

For the proof that follows we must strengthen our hypotheses on $\{y(t),\ t \ge 0\}$ as follows. Recall that we have assumed it is a right continuous Markov process on S with bounded infinitesimal generator Q which satisfies the Fredholm alternative. We shall now assume that

$$Qh(y) = q(y) \int_S \pi(y,dy')\ h(y') - q(y)h(y) \tag{6.7}$$

where q is continuous on S and strictly positive and the probability measures $\pi(y,A)$ have a continuous density relative to a reference measure $\phi$ on S and this density is strictly positive. This hypothesis implies, as is well known [19], the Fredholm alternative for Q. We also assume that

$$\ell = \int_0^\infty E\left\{g\big(y(s)\big)\ g\big(y(0)\big)\right\} ds \tag{6.8}$$

$$= \int \bar{P}(dy) \int \psi(y,dy')\ g(y')\ g(y) \qquad \text{(cf. Section 5)}$$

is a positive number (it is always nonnegative since it is equal to 1/2 the power spectral density of the stationary process $g\big(y(s)\big)$ at zero frequency).

The proof of (6.6) is in two parts. One deals with the case $k > 0$ fixed and the other with the case $k \to 0$.

<u>Part I</u> $k > 0$ (see [8]).

The process $\big(y(x),\theta(x)\big)$ (cf. (6.4)) is a Markov process on $S \times T$ ($T$ = the unit circle) with infinitesimal generator

$$L = Q + k\big(1 + g(y)\ \cos^2\theta\big)\ \frac{\partial}{\partial\theta}\ . \tag{6.9}$$

Let $V(y,\theta;k)$ be defined by

$$V = \frac{k}{2}\ g(y)\ \sin\ 2\theta \tag{6.10}$$

and note that

$$r(L) = \int_0^L V\big(y(s),\theta(s)\big)\ ds\ . \tag{6.11}$$

LEMMA 1. For each real $\beta$ the operator $L+\beta V$ generates a positivity preserving semigroup $R^V(t)$ on the bounded measurable functions on $S \times T$. This semigroup has an isolated maximal eigenvalue $\lambda = \lambda(\beta,k)$ and strictly positive corresponding right and left eigenvectors i.e.

$$(6.12) \qquad R^V(t)v = e^{\lambda t} v , \qquad (v = v(y,\theta;\beta,k))$$

$$\tilde{v}\, R^V(t) = e^{\lambda t} \tilde{v} , \qquad t \geq 0 .$$

Moreover $\lambda$, $v$ and $\tilde{v}$ are differentiable functions of $\beta$.

This lemma is proved by noting that by the Feynman-Kac formula we have explicitly

$$(6.13) \quad R^V(t)\, f(y,\theta) = E_{y,\theta}\left\{ e^{\beta\int_0^t V(y(s),\theta(s))ds} f(y(t),\theta(t)) \right\}$$

where $E_{y,\theta}\{\cdot\}$ is expectation relative to the measure of the process $\{(y(t),\theta(t)),\ t \geq 0\}$. The positivity preserving property is seen from (6.13).

For the existence of an isolated maximal eigenvalue with positive right and left null vectors it suffices to show [11] that there is a $t_0 < \infty$ and a constant $\gamma > 0$ such that for all $A \subset T \times S$

$$(6.14) \qquad R^V(t_0)\ \chi_A(y,\theta) \geq \gamma\ \phi(A)$$

Since $V$ is bounded and continuous, $|g| \leq 1$ and $Q$ has the form (6.7) this is easily obtained. Finally the differentiability in $\beta$ is a consequence of the isolated nature of $\lambda$ and is not hard to show.

LEMMA 2. We have that $\lambda(0,k) \equiv 0$,

$$(6.15) \qquad \frac{d}{d\beta}\lambda(\beta,k)\Big|_{\beta=0} = I(k) > 0 , \qquad k > 0$$

and $\lambda(\beta,k)$ is a convex function of $\beta$.

Proof: The fact that $I(k)$ is positive is a continuous-time analog of

Fursterberg's result [10] and it is proved in [8] and [9]. It is a nontrivial fact which we shall not, however, consider in detail. The convexity follows from the formula

$$\lambda(\beta,k) = \lim_{t\uparrow\infty} \frac{1}{t} \log \| R^V(t)\| , \tag{6.16}$$

the Feynman-Kac formula and Hölder's inequality.

Following [8] we now have

$$\begin{aligned} E\left\{\frac{4}{2+e^{2r(L)}}\right\} &= E\left\{\frac{4}{2+e^{2r(L)}}, \frac{r(L)}{L} < I - \delta\right\} \\ &\quad + E\left\{\frac{4}{2+e^{2r(L)}}, \frac{r(L)}{L} \geq I - \delta\right\} \\ &\leq 2P\left\{\frac{r(L)}{L} < I - \delta\right\} + 4e^{-2(I-\delta)L} , \end{aligned} \tag{6.17}$$

where $\delta > 0$ is chosen below. To estimate the probability on the right, we note that for $\beta > 0$

$$\begin{aligned} P\left\{r(L) < (I-\delta)L\right\} &= P\left\{-\beta r(L) > -\beta(I-\delta)L\right\} \\ &\leq e^{\beta(I-\delta)L} E\left\{e^{-\beta r(L)}\right\} \end{aligned} \tag{6.18}$$

But from (6.11), (6.16) and the subadditive nature of $\log \| R^V(t)\|$ we conclude that there is a constant C independent of k and L such that

$$\begin{aligned} E\left\{r(L) < (I-\delta)L\right\} &\leq C e^{\beta(I-\delta)L+\lambda(-\beta)L} \\ &= C e^{-L[-\lambda(-\beta)-\beta(I-\delta)]} \end{aligned} \tag{6.19}$$

If $\delta > 0$ is small enough so that $I-\delta > 0$, then we have $\beta = \beta^* > 0$ that maximizes the exponent on the right in (6.19) and $\beta^* = \beta^*(k)$. Thus, from (6.17) and (6.19) we get

$$z(k) = \min\left\{2(I-\delta), -\lambda(-\beta^*) - \beta^*(I-\delta)\right\} > 0 \tag{6.20}$$

which completes part I.

Part II. $k \downarrow 0$.

LEMMA 3. For $k > 0$, small we have

$$I(k) = \frac{k^2 \ell}{4} + o(k^2) \tag{6.21}$$

$$\lambda(\beta,k) = \frac{k^2 \ell}{4} \left(\frac{1}{2}\beta^2 + \beta\right) + o(k^2) \tag{6.22}$$

where $\ell$ is given by (6.8).

Proof: Let $\lambda_0(\beta) = \frac{\ell}{4}\left(\frac{1}{2}\beta^2 + \beta\right)$. We shall show that there are constants $\hat{C}_1$ and $\hat{C}_2$ such that

$$\frac{\hat{C}_2}{\hat{C}_1} e^{k^2[\lambda_0(\beta)-o(1)]t} \le \| R^V(t) \| \le \frac{\hat{C}_1}{\hat{C}_2} e^{k^2[\lambda_0(\beta)+o(1)]t} \tag{6.23}$$

which along with (6.16) gives the result (6.22). The result (6.21) is obtained by a similar argument.

Let $\psi(y,dz)$ be the kernel of $-Q^{-1}$ (cf. Section 5) and define

$$f_1(y,\theta) = \frac{\beta}{2} \sin 2\theta \int_S \psi(y,dy')\, g(y') \tag{6.24}$$

$$h_1(\theta) = -\int_0^\theta \left[\frac{\beta^2 \ell}{4} \sin^2\theta + \beta\ell \cos^2\theta \cos 2\theta - \frac{\ell}{4}\left(\frac{1}{2}\beta^2 + \beta\right)\right] d\theta \tag{6.25}$$

$$\begin{aligned} f_2(y,\theta) = \int_S \psi(y,dy') \Bigg[ & \frac{\beta}{2} g(y') \sin 2\theta\, f_1(y',\theta) + \frac{\beta}{2} g(y') h_1(\theta) \sin 2\theta \\ & + \frac{\partial f_1(y',\theta)}{\partial\theta} + g(y') \frac{\partial f_1(y',\theta)}{\partial\theta} \cos^2\theta + g(y') \frac{\partial h_1(\theta)}{\partial\theta} \cos^2\theta \Bigg] . \end{aligned} \tag{6.26}$$

By direct calculation as in Section 5 we find that

$$(L + \beta V)\left(1 + k\tilde{f}_1 + k^2 f_2\right) = k^2 \lambda_0(\beta) + f_3 \tag{6.27}$$

where $\tilde{f}_1 = f_1 + h_1$ and $f_3 = O(k^3)$ uniformly in $(y,\theta) \in S \times T$.

Let

$$f^{(k)} = 1 + k\tilde{f}_1 + k^2 f_2 . \tag{6.28}$$

For k sufficiently small there exist constants $\hat{c}_1$ and $\hat{c}_2$ such that

$$0 < \hat{c}_2 \le f^{(k)}(y,\theta) \le \hat{c}_1 < \infty \,. \tag{6.29}$$

If 1 denotes the function identically equal to one then the Feynman-Kac formula gives

$$\begin{aligned}(R^V(t)1)(y,\theta) &\le \frac{1}{\hat{c}_2} \sup_{y,\theta} E_{y,\theta}\left\{e^{\beta r(t)} f^{(k)}\big(y(t),\theta(t)\big)\right\} \\ (R^V(t)1)(y,\theta) &\ge \frac{1}{\hat{c}_1} \inf_{y,\theta} E_{y,\theta}\left\{e^{\beta r(t)} f^{(k)}\big(y(t),\theta(t)\big)\right\}\end{aligned} \tag{6.30}$$

Now choose $\delta > 0$ small so that

$$\begin{aligned}[L+\beta V-k^2(\lambda_0(\beta) \pm \delta\sqrt{k})]\, f^{(k)} &= k^2[\mp \delta\sqrt{k} + o(k)] \\ &\gtreqless 0\end{aligned} \tag{6.31}$$

for all k small. From Dynkin's identity (integrated semigroup identity) we have

$$\begin{aligned} &e^{-k^2(\lambda_0(\beta)\pm\delta\sqrt{k})t}\, E_{y,\theta}\left\{e^{\beta r(t)} f^{(k)}\big(y(t),\theta(t)\big)\right\} \\ &= f^{(k)}(y,\theta) + E_{y,\theta}\left\{\int_0^t e^{-k^2(\lambda_0(\beta)\pm\sqrt{k}\delta)s} \right. \\ &\qquad \left. [L+\beta V - k^2(\lambda_0 \pm \delta\sqrt{k}\,)]\, f^{(k)}\big(y(s),\theta(s)\big)\, ds\right\}\end{aligned} \tag{6.32}$$

From (6.31) and (6.32) we obtain the inequalities

$$\begin{aligned}E_{\theta,y}\left\{e^{\beta r(t)} f^{(k)}\big(y(t),\theta(t)\big)\right\} &\le \hat{c}_1\, e^{k^2(\lambda_0(\beta)+o(1))t} \\ E_{\theta,y}\left\{e^{\beta r(t)} f^{(k)}\big(y(t),\theta(t)\big)\right\} &\ge \hat{c}_2\, e^{k^2(\lambda_0(\beta)-o(1))t}\end{aligned} \tag{6.33}$$

Combining this with (6.30) yields (6.23) and the lemma is proved.

Now to find $z_0$ in (6.2) we repeat the argument (6.17)-(6.19) and use Lemma 3. By picking the $\delta > 0$ appropriately we actually obtain $z_0 = \frac{\ell}{2}(3 - \sqrt{8}) > 0$.

This completes the proof of (6.1) and (6.2).

REFERENCES

[1] W. Kohler and G. C. Papanicolaou, Power statistics for waves in one dimension and comparison with radiative transport theory, J. Math. Phys. 14 (1973) pp. 1733-1745 and 15 (1974), pp. 2186-2197.

[2] W. Kohler, and G. C. Papanicolaou, Wave propagation in a randomly inhomogeneous ocean, Springer Lecture Notes in Physics # 70 (1977), edited by J. B. Keller and J. Papadakis.

[3] W. Kohler and G. C. Papanicolaou, Fluctuation phenomena in under water sound propagation, I and II, Proceedings of Conference on Stochastic Equations and Applications, Academic Press (1977), edited by J. D. Mason.

[4] J. B. Keller, G. C. Papanicolaou and J. Weilenmann, Heat conduction in a one-dimensional random medium, Comm. Pure Appl. Math. 32 (1978) to appear.

[5] G. C. Papanicolaou, D. Stroock and S. R. S. Varadhan, Martingale approach to some limit theorems, in Statistic Mechanics, Dynamical Systems and the Duke Turbulence Conference, by D. Ruelle, Duke Univ., Math. Series Vol. III, Durham, North Carolina, 1977.

[6] G. Blankenship and G. C. Papanicolaou, Stability and control of stochastic systems with wide-boud noise disturbances, I, SIAM J. Appl. Math. 34 (1978), pp. 437-476.

[7] O'Connor, A. J., A central limit theorem for the disordered harmonic chain, Comm. Math. Phys. 45 (1975) pp. 67-77.

[8] L. A. Pastur and E. P. Feldman, Wave transmittance for a thick layer of a randomly inhomogeneous medium, Soviet Phys. JETP 40 (1975), pp. 241-243.

[9] L. A. Pastur, Spectra of random Jacobi matrices and Schrödinger equations with random potentials on the whole axis, Preprint, FTINT Akad. Nauk Ukr. SSR, Khar'kov, 1974.

[10] H. Furstenberg, Noncommuting random pooducts, Trans. Am. Math. Soc. 108

(1963), pp. 337-428.

[11] T. Harris, Branching Processes, Springer, Berlin, 1963.

[12] V. I. Klyatskin, Statistical properties of dynamical systems with randomly fluctuating parameters, Nauka, Moskow, 1975.

[13] I. I. Gihman and A. V. Skorohod, Stochastic Differential Equations, Springer, Berlin, 1972.

[14] A. Casher and J. L. Lebowitz, Heat flow in regular and disturbed harmonic chains, J. Math. Phys., 12 (1971) pp. 1701-1711.

[15] R. Z. Khasminskii, On stochastic processes defined by differential equations with a small parameter, Theor. Prob. Appl. 11 (1966), pp. 211-228.

[16] R. Z. Khasminskii, A limit theorem for solutions of differential equations with a random right-hand side, Theor. Prob. Appl. 11, pp pp. 390-406.

[17] B. White and J. Franklin, A limit theorem for stochastic two-point boundary value problems of ordinary differential equations, Comm. Pure Appl. Math., to appear.

[18] M. I. Freidlin, Fluctuations in dynamical systems with averaging, Dokl. Akad. Nauk SSSR, 17 (1976) pp. 104-108.

[19] J. L. Doob, Stochastic Processes, J. Wiley, New York, 1953.

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# A STOCHASTIC PROBLEM IN PHYSICS

CECILE DEWITT-MORETTE

A STOCHASTIC PROBLEM IN PHYSICS

Cecile DeWitt-Morette
Department of Astronomy and Center for Relativity
University of Texas, Austin, TX 78712

## Introduction

The world is global and stochastic and physical laws are local and deterministic. Thus the problems discussed at this Summer School are the very fabric of physics. But physics asks some questions which go beyond the territory which has been explored here. I shall present one of them, show how far physicists have gone toward its solution and mention an important problem of current interest.

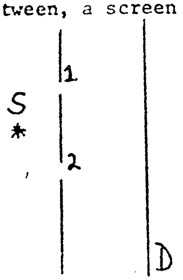
Probability theory begins with a probability space $(\Omega,\mathcal{F},P)$. The careful definition of the $\sigma$-field $\mathcal{F}$ of subsets of $\Omega$ and of the probability measure $P$ has given us many powerful theorems. It is also possible, and often preferable in physics, to define P as a promeasure,[1] namely as a projective family of bounded measures defined on the system of finite dimensional spaces $\mathcal{G}$ known as the projective system of $\Omega$. We thus start from $(\Omega,\mathcal{G},P)$ rather than $(\Omega,\mathcal{F},P)$. This is excellent for statistical mechanics. Unfortunately, in quantum mechanics, we have to deal with families of unbounded measures on the projective system $\mathcal{G}$ of $\Omega$. And this is the key issue in the study of Feynman path

---

[1] Also called cylindrical measure. See for instance [Bourbaki].

integrals.

When Feynman was a graduate student in the early forties, he felt uncomfortable with quantum mechanics and, rather than let familiarity become a substitute for understanding, he analyzed[1] a basic quantum phenomenon in his own terms: He considered a source of electrons S; a plane detector D and, in between, a screen with 2 slits which could be closed or open.




The detector measures the number of electrons arriving on the plane D as a function of position. It is turned on for an amount of time which can be considered as infinite. If one does these experiments:

i. slit 1 open, slit 2 closed

ii. slit 1 closed, slit 2 open

iii. both slits open,

one finds that the probability $P_3$ measured by the detector in the third experiment is not the sum $P_1 + P_2$ of the probabilities measured in the first two experiments. But one finds that there is an additive quantity, called probability amplitude, whose absolute value squared is the probability $P(b,t_b;a,t_a)$ that the electron known to be at $\underline{a}$ at time $t_a$ be found at $\underline{b}$ at time $t_b$:

$$P_i(b,t_b;a,t_a) = |K_i(b,t_b;a,t_a)|^2 \qquad i = 1,2,3$$

$$K_3 = K_1 + K_2$$

This result can be generalized to an infinite number of slits and the probability amplitude for a transition from $(a,t_a)$ to $(b,t_b)$ is the sum over all possible paths

$$x;\ T \to R^3 \quad \text{such that} \quad x(t_a) = a \quad \text{and} \quad x(t_b) = b.$$

The requirement that, in the limit $\hbar = 0$, quantum physics goes over to classical physics implies that

$$K(b,t_b;a,t_a) = \text{"}\sum_{\Omega}\text{"} \exp(iS(x)/\hbar)$$

[1] Read the first chapter of [Feynman and Hibbs] for a beautiful account of this analysis.

where S is the action defined by the Lagrangian L.

$$S(x) = \int_T L(x(t),\dot{x}(t))dt$$

The system need not be a particle in $R^3$. For instance consider a system whose configuration space is M. One can write the probability amplitude for a transition from $(a \in M, t_a)$ to $(b \in M, t_b)$ by a similar path integral. The space of paths $x \in \Omega$ is then the space of continuous paths $x: T \to M$, such that $x(t_a) = a$ and $x(t_b) = b$.

I have now set up all the necessary physical concepts to show why (1) physicists need "P" to be more general than a probability measure. (2) They need "Ω" to be endowed with a variety of structures. Indeed:

1. The fact that we have to work with unbounded measures comes from the fact that we sum probability amplitudes rather than probabilities. Where a probabilist has

$$d\gamma(u) = (2\pi i)^{-d/2} (\mathrm{Det}\ \Gamma^{-1})^{1/2} \exp(-\Gamma_{kj}^{-1} u^k u^j/2)\ du^1 \ldots du^n \tag{1,a}$$

we have

$$d\gamma(u) = (2\pi i)^{-d/2} (\mathrm{Det}\ \Gamma^{-1})^{1/2} \exp(i\Gamma_{kj}^{-1} u^k u^j/2\ du^1 \ldots du^n \tag{2,a}$$

It is clear that we cannot use the probabilists' estimates and we have been forced to investigate different approaches.

i. Feynman did not know the Wiener integral and invented his own calculus. He replaced a path x by n of its values $x(t_1)\ldots x(t_n)$, and computed the limit $n = \infty$ of the discretized problem. He discovered "experimentally" that what is now known as the Stratonovitch integral gives the "right" result if the problem is simple enough--for instance, if the configuration space $M = R^d$. With his admittedly crude tool, Feynman was able to construct a fantastically good computational procedure known as the Feynman diagrams. The Feynman diagrams are used widely in nearly all branches of physics. The diagram rules can be applied and even refined without knowing anything about path integration. They have been justified by several methods and have often eclipsed path integration.

ii. Another approach pioneered in particular by Montroll and Nelson is based on analytic continuation; either the time or the mass is complexified. The main activity in this domain is euclidean field theory.

iii. I shall speak today of a method which proceeds neither by discretization nor by analytical continuation. It has given versatile tools which can a fortiori be used in probability theory. It defines an object on $\Omega$ called prodistribution. Because prodistributions are defined directly on $\Omega$, one can investigate what happens when $\Omega$ is endowed with a variety of other structures.

2. The space $\Omega$, in physics, is often the space of paths mapping the time $T \subset R$ into the configuration space M, or into the phase space $T^*M$ of a system. Too often the global properties of the configuration space of a system are ignored, and one thinks of the configuration space of a system with d degrees of freedom as $R^d$. But even the simplest systems, a pendulum, a system of indistinguishable particles, a rigid body rotator, etc., have configuration spaces which are multiply connected riemannian spaces. A path integral formulation of quantum physics is an integral over $\Omega$. It reflects the global properties of $\Omega$ and the various structures put on $\Omega$.

I shall now introduce prodistributions and explain briefly[1] how they can be used to compute path integrals explicitly. Let us go back to P considered as a promeasure. We could have defined a promeasure by its Fourier transform, i.e. by a family of functions on the dual of the projective system. For instance, instead of defining a gaussian promeasure by a projective family of gaussians on finite dimensional spaces of the type (1,a) we can define it by their Fourier transforms $\mathcal{F}\gamma$ on the dual spaces

$$\mathcal{F}\gamma(\xi) = \exp(-\Gamma^{ij}\xi_i\xi_j/2) \tag{1,b}$$

At this point we can remove the condition that the measures $\gamma$ be bounded. Indeed whereas $\gamma$ is a set function, $\gamma(U) = \int_U d\gamma(u)$, its Fourier transform $\mathcal{F}\gamma$ is defined pointwise. Whereas "i" plays havoc in equation (2,a) it is quite manageable in its Fourier transform

$$\mathcal{F}\gamma(\xi) = \exp(-i\Gamma^{ij}\xi_i\xi_j/2) \tag{2,b}$$

---

[1] A detailed account will appear in [DeWitt-Morette, Maheshwari, B. Nelson, 1979].

In other words, instead of considering a projective family of bounded measures, we can consider a projective family of tempered distributions. Dieudonné has proposed to call a projective family of tempered distributions a "prodistribution."

Since time is limited I shall work with an example. The Feynman-Kac formula suggests itself since you are working with the Kac formula and I work with the Feynman formula. Given

$$ih \frac{\partial\psi}{\partial t} = (-\frac{h^2}{2m}\Delta + V)\psi \quad , \qquad \Delta = |g|^{-1/2}\partial_\alpha g^{\alpha\beta}|g|^{1/2}\partial_\beta$$

$$\psi(t_a, a) = \phi(a)$$

Write down the path integral representation of the solution and compute it. The problem is sufficiently complicated to display the power of prodistributions.

Answer:

$$\psi(t_b, b) = \int_{\Omega_+} dw_+^W(y)\, \exp\left(-\frac{i}{\mu^2 m}\int_{t_a}^{t_b} V(\mathrm{Dev}_b(\mu y, t)dt\right) \phi(\mathrm{Dev}_b(\mu y, t_a)) \tag{3}$$

In particular the propagator $K(t_b, b; t_a, a)$ is obtained by choosing the initial wave function to be

$$\phi(\mathrm{Dev}_b(\mu y, t_a)) = \delta(\mathrm{Dev}_b(\mu y, t_a) - a) \tag{4}$$

The following notation has been used.

i. $\mu = (h/m)^{1/2}$

ii. $\mathrm{Dev}_b$ is the development mapping from the space of $L^{2,1}$ paths[1] on the tangent space $T_b M$ (tangent space to M at b) to the space of $L^{2,1}$ paths of M. If $M = R^d$, then

$$\mathrm{Dev}_b(\mu x, t) = b + \mu x(t)$$

---

[1] Space of square integrable functions whose first weak derivatives are square integrable.

In general, $\mathrm{Dev}_b(\mu x,\cdot)$ is a path X on M such that $\dot{X}(t)$ is equal to the parallel transport of $\dot{x}(t)$ from b to $X(t)$ along X. Thus equation (3) is defined for paths X on M, but the variable of integration x is a path on $T_bM$. Elworthy[1] has shown that the development mapping defines a measureable mapping from the space of continuous paths on $T_bM$ into the space of continuous paths on M.

iii. $\Omega_+$ is the space of continuous paths on $T_bM$ such that $x(t_b) = 0$.

iv. $w_+^W$ is the prodistribution on $\Omega_+$ defined by its Fourier transform on the dual $\Omega_+'$ of $\Omega_+$.

$$\text{Let } x \in \Omega_+ \text{ and } \mu \in \Omega_+' , \qquad \langle\mu,x\rangle = \int d\mu_\alpha(t)x^\alpha(t) \qquad a = 1,\ldots d$$

A gaussian prodistribution w on a space of continuous paths defined on T is a prodistribution whose Fourier transform is of the form

$$\mathcal{F}w(\mu) = \exp(-\tfrac{i}{2}\,W(\mu,\mu))$$

$$W(\mu,\nu) = \int_T d\mu_\alpha(t) \int_T d\nu_\beta(s)G^{\alpha\beta}(t,s)$$

The Wiener prodistribution $w_+^W$ on $\Omega_+$ is the gaussian prodistribution whose covariance is

$$G^{\alpha\beta}(t,s) = \inf\,(t_b - t,\ t_b - s)$$

Equation (3) in the flat case is the Feynman-Kac formula. Note that we integrate over $\Omega_+$ (paths vanishing at $t_b$) and not on $\Omega_-$ (paths vanishing at $t_a$, where the covariance $G^{\alpha\beta}(t,s) = \inf(t-t_a, s-t_a)$). This is conceptually simpler (sum over all paths ending at b) and computationally easier. This is the form one obtains readily by working with product integrals.

Equation (3) has been derived by Elworthy for the probalistic case (solution of the heat diffusion equation). The theory of prodistributions makes it possible to use Elworthy's construction for the Schrödinger equation.

Computation of equation (3).

Consider a linear continuous mapping P from $\Omega_+$ eitner into itself or into

[1] [Elworthy 1978]

another space. Say

$P: \Omega_+ \to U$ by $x \to u$ ; let P be the transposed mapping between the respective duals $U'$ and $\Omega'_+$

$\tilde{P}: U' \to \Omega'_+$ by $\xi \mapsto \mu$ . If $F: \Omega_+ \to R$ is such that $F = f \circ U$, then

$$\int_{\Omega_+} F(x)dw(x) = \int_U f(u)dw_P(u)$$

where $\mathfrak{F}w_P = \mathfrak{F}w \circ \tilde{P}$ .

This simple relation is the clue for many explicit calculations and we carry out one calculation in the appendix.

The explicit calculation of (3) proceeds via several linear mappings. I shall mention only a couple of them:

1. Map $y \mapsto x$ such that $b + \mu y(t) = q(t) + \mu x(t)$ where q ia the path whose development is a solution of the Euler-Lagrange equation of the problem such that $q(t_b) = b$. The boundary value $q(t_a)$ is related to the initial wave function. For instance if one computes the propagator (eq. 4) one chooses $q(t_a) = a$.

Set $\mathrm{Dev}(q + \mu x, t) = Y(t,x,\mu)$ and expand the integrand in equation (3) in powers of $\mu$.

Set $\delta Y(\cdot,x) = \frac{\partial}{\partial\mu} \mathrm{Dev}(q + \mu x, \cdot)\Big|_{\mu=0} = \mathrm{Dev}'(q)x$

$\mathrm{Dev}'(q)$ is a linear mapping from the set of vector fields along q into the set of vector fields along $\mathrm{Dev}(q)$.

2. It is easy to construct the linear mapping which "absorbs" terms of the form $(\nabla_\alpha \nabla_\beta V)\delta Y^\alpha \delta Y^\beta$ . The image under this mapping of the Wiener gaussian $w_+^W$ is a gaussian whose covariance is an elementary kernel of the Jacobi equation of the system (alias the small disturbance equation, alias the variational equation of the action S). The problem of solving a partial differential equation (Schrödinger equation) is then reduced to solving an ordinary differential equation (Jacobi equation) and much is known about this second order linear homogeneous ordinary equation.[1]

[1] cf Jacobi, Poincaré, Sturm Liouville, etc., etc. ...,

Finally one obtains

$$\psi = \psi_{WKB}(1 + \sum_{k=1}^{\infty} \hbar^k A_k)$$

where

$$\psi_{WKB}(b,t_b) = \exp(i\bar{S}(b,t_b;a,t_a)/\hbar)(2\pi i\hbar)^{-n/2} \left|\det\partial^2 \bar{S}/\partial b^\alpha \partial a^\beta\right|^{1/2}$$

and where $\bar{S}$ is the action along the classical path from $(a,t_a)$ to $(b,t_b)$ and where the terms $A_k$ are given by integrals over finite dimensional spaces. $A_k$ is a "moment integral" very easy to compute in the flat case, in principle for any k. It is very difficult to compute in the Riemannian case.

Prodistributions have been used in a variety of problems: scattering states, bound states, quantum properties of systems whose classical solutions have caustics, etc. A versatile technology has been developed to obtain explicit answers.

Problems on curved spaces have been solved. The next problem we plan to investigate is path integration on curved spacetimes. This is not a simple generalization of path integration on curved spaces: if one replaces the Laplacian by a d'Alembertian, one loses ellipticity. On the other hand we do not want to touch field theory until we understand what happens on curved spacetimes.

## Appendix

Example. Let w be the Wiener measure on the space $\Omega$ of continuous paths $x: T \to R$ such that $x(t_a) = 0$.

$$\mathcal{F}w(\mu) = \exp(-W(\mu,\mu)/2) \qquad \text{for} \quad \mu \in \Omega'$$

$$W(\mu,\mu) = \int_T d\mu(t) \int_T d\mu(s) \quad G(t,s)$$

$$G(t,s) = \inf(t-t_a, s-t_a)$$

Compute $I = \int_\Omega F(x)dw(x)$ where

$$F(x) = f(x(t_1),\ x(t_2) - x(t_1),\ \ x(t_n) - x(t_{n-1}))$$

$$t_a = t_0 < t_1 \ldots < t_n = t_b.$$

Answer,

$$F = f \circ P \quad \text{where} \quad P: x \mapsto u = \{u^1 \ldots u^{n+1}\}$$

$$u^k = x(t_k) - x(t_{k-1}) = <\delta_{t_k} - \delta_{t_{k-1}}, x>$$

It follows that $I = \int_{R^n} f(u)dw_P(u)$ where $\mathcal{F}w_P = \mathcal{F}w \circ \tilde{P}$ .

The transposed $\tilde{P}$ of P is defined by

$$\tilde{P}: R^n \to \Omega' \quad \text{by} \quad \xi \to \mu \quad \text{such that}$$

$$<\tilde{P}\xi,x>_\Omega = <\xi,Px>_{R^n}$$

where $<\ ,>_\Omega$ is the duality in $\Omega$, $<\mu,x>_\Omega = \int_T d\mu(t)x(t)$

and $<\ ,>_{R^n}$ is the duality in $R^n$, $<\xi,u>_{R^n} = \Sigma\xi_i u^i$ ,

One can read off immediately

$$\tilde{P}\xi = \sum_i \xi_i(\delta_{t_i} - \delta_{t_{i-1}})$$

Hence

$$\mathcal{F}w_P = \exp(-W(\tilde{P}\xi,\tilde{P}\xi)/2)$$

A quick calculation gives

$$W(\tilde{P}\xi,\tilde{P}\xi) = \Sigma \quad \xi_i^2(t_i - t_{i-1}) \quad .$$

It follows that

$$I = \int_{R^n} f(u^1,\ldots u^n)d\gamma_1(u^1)\ldots d\gamma_n(u^n)$$

with $\quad d\gamma_k(u^k) = (2\pi)^{-1/2} \exp(-(u^k)^2/2(t_k-t_{k-1}))du^k \quad .$

This example, possibly the best known result of probability theory, was chosen to display on familiar grounds, methods used in computing explicitly the WKB approximation of the wave function on curved spaces (eq. 3).

References

Bourbaki, N. (1969) Elements de mathematique, Chapter IX, Volume VI, also referred to as Fascicule 35 or No. 1343 of the Actualites Scientifiques et Industrielle, Hermann, Paris.

DeWitt-Morette, C., A. Maheshwari and B. Nelson (1979) Path Integration in Non-Relativistic Quantum Mechanics, Physics Reports.

Elworthy, K. D. (1978) "Stochastic dynamical systems and their flows," to appear in Proceedings of the Conference on Stochastic Analysis, Northwestern University.

Feynman, R. P. and A. R. Hibbs (1965) Quantum Mechanics and Path Integrals, McGraw-Hill, New York.

CENTRO INTERNAZIONALE MATEMATICO ESTIVO

(C.I.M.E.)

# THE EMBEDDING PROBLEM FOR STOCHASTIC MATRICES

G. S. GOODMAN

## §1. Statement of the embedding problem.

An $n \times n$ matrix $P = [p_{ij}]$, with non-negative entries, is said to be stochastic if the entries along any row sum to one. In 1938, G. Elfving [2] formulated the embedding problem for stochastic matrices, essentially as follows.

For what matrices P can there be found a value $t_0 > 0$ and a continuous family of stochastic matrices $P(s,t)$ on $0 \leq s \leq t \leq t_0$ that satisfies the functional equation

(1.1.) $P(s,t) = P(s,u)P(u,t)$ whenever $(s \leq u \leq t)$

has

(1.2.) $P(s,s) = I$ for all $0 \leq s \leq t_0$

and is such that

(1.3.) $P(0,t_0) = P$?

Here I denotes, as usual, the identity matrix.

In probability theory, the entries $p_{ij}(s,t)$, $i,j = 1,\ldots,n$, of $P(s,t)$ are regarded as transition probabilities of an n-state, non-homogeneous Markov process in continuous time, i.e., $p_{ij}(s,t)$ is the conditional probability that the process will be found in state j at time t, given that it was in state i at time s, and equation (1.1.) is known as the Chapman - Kolmogorov equation. The continuity of $P(s,t)$ reflects certain hypotheses concerning the nature of the sample paths. Thus the embedding problem is concerned with determining exactly which stochastic matrices P can serve as transition matrices of an n-state Markov process.

For $2 \times 2$ matrices, the problem had been solved in 1932 by Fréchet [3], and the solution was rediscovered by Elfving. The necessary and sufficient condition for embeddability in this case is that

$$\operatorname{tr} P - 1 = \det P > 0.$$

When $n > 2$, the problem is still open. It is known that $\det P > 0$ is necessary, but is not sufficient, as a result cited in §5 below abundantly shows.

## §2. The Kolmogorov Equations

In [5], I showed that if such an embedding family $P(s,t)$ exists, then the simple change of time scale, replacing s and t by

(2.1) $\quad -\log\det P(0,s)$ and $-\log\det P(0,t)$

converts the family into one which is lipschitzian in each variable and can be identified with the general solution of the Kolmogorov forward and backward differential equations

K1 $$\frac{dP}{dt} = PQ(t) \quad \text{a.e.} \qquad (0 \leq t \leq t_0),$$

K2 $$\frac{dP}{ds} = -Q(s)P \quad \text{a.e.} \qquad (0 \leq s \leq t_0),$$

resp. (understood in the Carathéodory sense). Here, for fixed values of s and t, the Q's denote intensity matrices i.e., $Q(t)$ is given by

(2.2) $$Q(t) = \lim_{u,v \to t} \frac{P(u,v) - I}{v - u} \qquad (u \leq t \leq v,\ v-u \neq 0),$$

and a similar formula holds for $Q(s)$. The main work in [5] was to show that these limits exist a.e. when the above time scale is used.

It follows from (2.3) that the intensity matrices $Q = [q_{ij}]$ $i,j = 1,\dots, n$, satisfy the conditions

(2.3) $$\begin{cases} -1 \leq q_{ii} \leq 0 \text{ for all } i,\ 1 \geq q_{ij} \geq 0 \quad \text{whenever } i \neq j, \\ \sum_{j=1}^{n} q_{ij} = 0 \text{ all } i, \\ \text{while (2.1) implies that} \\ \sum_{i=1}^{n} q_{ii} = -1 \end{cases}$$

They thus form an $n(n-q)$ dimensional simplex. Moroever, the conditions (2.3) characterise the intensity matrices, for, given any one-parameter of such matrices with measurable entries, we can integrate the Kolmogorov differential equations, subject to the initial values (1.2), and generate a unique family $P(s,t)$ of stochastic matrices that satisfies (1.1) and yields (2.2). It is only the proof that the elements of $P(s,t)$ are non-negative that is not routine: a simple way out is to use product integration, cf. [8]

It follows that <u>either one of the Kolmogorov equations (K1) or (K2), together with the constraints (2.3), can be regarded as a control system that generates stochastic matrices, with the intensity matrices, varying measurably, playing the rôle of controls.</u>

3. <u>Control-theoretic formulation of the embedding problem.</u>

In [5], I pointed out that by replacing the functional equation (1.1) by the control eqation (K1), subject to the constraints (2.3), the embedding problem is converted into an equivalent <u>reachability problem</u>, viz.,

> <u>What matrices P can be reached at</u> $t_0$ <u>from the identity matrix</u> I <u>at</u> $t = 0$ <u>by solutions</u> $P(t) = P(0,t)$ of (K1)?

Of course, if we want, we can use (K2) and ask

> <u>What matrices</u> P <u>at</u> $s = 0$ <u>can be steered to the identity matrix</u> I <u>at</u> $s = t_0$ <u>along solutions</u> $P(s) = P(s,t_0)$ of (K2)?

In both cases, $t_0$ plays the role of a parameter. The two problems are equivalent, and the second can be put into the same form as the first by replacing s by $t_0-s$, thereby changing the $-$ sign in (K2) to +.

The investigation of the embedding problem by control-theoretic means became one of the main tasks of a research project, sponsored by the Scientific Affairs Division of Nato, in which the principal investigators were Søren Johansen from Copenhagen and myself.

4. Some properties of the reachable set.

From general considerations concerning semigroups of positive matrices [1], it follows that the reachable set is contractable to certain of its boundary points (which correspond to values $t_0 = \infty$). In [9], Johansen proved from the differential quations that the contractions can be done along rays, so that the reachable set is actually starlike with respect to these points.

The basic existence theorem of Filippov, in the form given by Lee and Markus [12], Roxin [15] and myself [4], implies that the set of matrices that can be reached in time $t_0$ is compact, and the set of all reachable matrices is compact relative to GL (n).

The fact that the sections $t_0$=const. of the reachable set is arc-wise connected is almost immediate. For if $P_1$ and $P_2$ are two embeddable matrices, associated with the embedding families $P_1(s,t)$ and $P_2(s,t)$, resp., each reachable in time $t_0$, then

$$P(u) = P_1(0,u)\ P_2(u,t_0) \qquad (0 \leq u \leq t_0)$$

represents an absolutely continuous curve wich joins $P_1$ and $P_2$. For each fixed u, P(u) is reachable in time $t_0$, its control law being that of $P_1$ from 0 to u and that of $P_2$ from u to $t_0$. The same argument shows that the sections $t_0$=const. of the set reachable by bang-bang controls is also arcwise connected.

The reachable set has certain symmetries, which go back to the fact that the order in which the states are labeled in a Markov process is irrelevant. Thus, the reachable set is carried onto itself by orthogonal transformations induced by permutation matrices. One could try to normalize the reachable matrices by requiring that the elements along the main diagonal be arranged in an increasing, or decreasing, order, but this is not always convenient.

## §5. The bang-bang conjecture

The theory of sliding regimes or chattering controls [13] asserts that any embeddable matrix can be approximated (along its whole trajectory) by finite products of elementary matrices, i.e., matrices that are generated when the controls are fixed at the extreme points of the control region. These elementary matrices turn out to be precisely those stochastic matrices which differ from the identity by the presence of precisely one non-zero, off-diagonal element. Johansen [9] has observed that their trajectories are rectilinear.

Some years ago, I conjectured that the bang-bang principle holds and that every embeddable matrix is a finite product of elementary matrices. The conjecture was suggested by results of Loewner [14] on totally positive and on doubly-stochastic matrices. It is easy to see that it holds, trivially, when $n=2$, for then any stochastic matrix $P$ can be written as the product of two elementary matrices, so long as it satisfies the embedding condition $\mathrm{tr}P - 1 > 0$. In general, one might expect that the number of terms would depend on $\det P$ as well as upon $n$, but I suspect that it depends upon $n$ alone and equals $n(n-1)$

When $n \geq 2$, Johansen proved [9] that every matrix in the interior of the reachable set can be reached by bang-bang controls with a finite number of switches. Owing to the work of Krener [11], this is now seen to be a general property of certain control systems. Since there is no bound on the number of switches, it is not possible to conclude that the bang-bang principle holds for matrices on the boundary of the reachable set.

A considerable amount of effort has gone into the study of the bang-bang conjecture. Recent results in the case $n=3$ are reported below in §11.

## §6 The determinantal inequality.

While the result of section 4 give a certain amount of qualitative information about the set of all embeddable matrices, they fail to yield any criterion for deciding whether a given stochastic matrix P is embeddable or not.

To remedy this, we may appeal to a result proved in [5]. There it was noted that the differential equation (K1), together with the constraints (2.3), yield a differential inequality for the product of the diagonal elements in $P(t)$, just by omitting the terms in the equation which are non-negative. Integrating this inequality and using the Jacobi-Liouville formula for the determinant (or(2.1)directly) gives then the following inequality which must be satisfield by the elements of any embeddable matrix P:

(6.1) $$\prod_{i=1}^{n} p_{ii} \geq \det P > 0.$$

The same inequality occurs in the theories of positive-definite and totally positive matrices.

The inequality (5.1) is a strong necessary condition for embeddability, and it can be used to show that there are stochastic matrices arbitrarily close to the identity which are not embeddable (cf. §7 below).

The set of stochastic matrices which satisfies (6.1) is a semigroup, and in [5] I conjectured that it is precisely the semigroup of embeddable matrices, i.e., that the condition (6.1) is not only necessary, but also sufficient for embeddability. Shortly thereafter, David Williams pointed out to me that equality can hold in (6.1) for an embeddable matrix P only if some off-diagonal element vanishes. His proof was based upon the functional equation (1.1), but it is equally apparent when one checks the differential inequality described above.

Williams' remark shows, for example, that the $3\times3$ matrix whose entries on the main diagonal are each 1/4, while the remaining elements are each 3/8, is not embeddable, even though (6.1) is satisfied. In §8, we shall see how his remark can be used to establish that the set of embeddable matrices is <u>not</u> convex when $n>2$. (In fact, its convex hull is not known).

## §7 <u>Geometrical representation of stochastic matrices.</u>

One of the most captivating features of the embedding problem is that it is completely equivalent to a problem of geometry, or, at least, of kinematics. In this and the next few sections, I shall explain how this comes about. A more complete account will appear in [7].

For simplicity, let us restrict ourselves to $3\times3$ matrices. Considering the rows of a stochastic matrix P as vectors in threespace relative to a fixed coordinate system, we see that they specify the vertices of an oriented triangle $\langle P\rangle$ lying in the plane through the three unit vectors. The identity matrix I corresponds to the unit triangle $\langle I\rangle$ and fixes the orientation. All the other stochastic matrices describe subtriangles of I, and the inclusion is strict except for permutation matrices.

then to conclude from (1.1) --(1.3) that every embeddable matrix P belongs to this semigroup. (The same can be said for (6.1), which, of course, implies (7.1), but (7.1) has been derived without use of the differential equations.)

8 <u>A pre-order for stochastic matrices.</u>

Now let us return to our main theme. Having associated to each $3 \times 3$ stochastic matrix P a triangle $\langle P \rangle$, we can introduce a pre-order in the class of stochastic matrices by defining

(8.1) $\quad P \prec R$ <u>if and only if</u> $\quad \mathrm{co}\langle P\rangle \subset \mathrm{co}\langle R\rangle$

Thus, the inclusion refers to the points of the simplex spanned by the vertices of P . The pre-order fails to be a partial order because it is not anti-symmetric: the ordering of the vertices has got lost in the set inclusion. Indeed, $P \prec R$ and $R \prec P$ mean that P and R are congruent, so that P and R are congruent under the action of a permutation matrix (cf. §4).

The pre-order just introduced can be put into an analytical form that shows that it agrees with the pre-order natural to any transformation semigroup (cf. [4], p. 14) viz.,

(8.2) $P \prec R$ <u>if and only if</u> $SR = P$ <u>for some stochastic matrix</u> $S$.

A proof will be given in [7] . Since P, considered as an operator on contravariant vectors, represents the unique affine transformation that carries $\langle I \rangle$ onto $\langle P \rangle$ , whereupon $\mathrm{co}\langle I\rangle$ goes onto $\mathrm{co}\langle P\rangle$, and the some holds for R in place of P, we recognize in (8.2) the analog of a well-known property of projection operators Specialization to elementary matrices (§4 ) reveals further analogies still.

It should be noted that (8.2) continues to hold if P, R and S are restricted to lie in the subsemigroup of stochastic matrices with positive determinant.

Proposition (8.2) allows us to characterize the $3 \times 3$ stochastic matrices which are indecomposable, i.e., they cannot be factored, except trivially, in the semigroup of stochastic matrices. Indeed, if $\langle P \rangle$ is such that each of its vertices lies on a different side of $\langle I \rangle$, then $P \prec R$ implies that $R = I$ or $P$, unless the vertices of $\langle P \rangle$ coincide with the vertices of $\langle I \rangle$, in which case $R$ is a permutation matrix, and the factorization might again be considered trivial. Conversely, if $P$ does not have the configuration stated, it is easy to see that a non-trivial $R$ can be found such that $co\langle P \rangle \subset co\langle R \rangle$, whence $P$ has $R$ as a factor. Note that, in the first case, it is possible for (7.1) to be satisfied, but not (6.1), and such P's can lie arbitrarily close to the identity I, without being embeddable. They are, in fact, convex combinations of I and permutation matrices; consequently, there are rays emanating from I in the semigroup of stochastic matrices that do not meet the set of embeddable matrices except at I.

The embeddable matrices are far from being indecomposable, for they are precisely the stochastic matrices which are infinitely factorizable, in a sense that can be made precise [6]. This characterization of the embeddable matrices dates from my early work in the field, in 1969, and is linked to cognate ideas in the theory of schlicht functions. It embraces the intuitive idea that embeddable matrices can be decomposed into infinitesimal matrices and then reassembled from them, which underlies the results of §2. The same ideas were taken up later by Johansen, who improved them and gave an elegant derivation of Kolmogorov differential equations based upon them, in [8].

## §9. A geometric formulation of the embedding problem.

Looking back at the preceding development, we see that the embedding problem of §1 can now be put into a simple geometrical form.

For, suppose we fix $t_0$ and set

$$P(s) = P(s,t_0) \qquad (0 \leq s \leq t_0).$$

Then (1.1) becomes

$$(9.1) \qquad P(s) = P(s,u)\, P(u) \qquad (0 \leq s \leq u \leq t_0),$$

and (1.2) and (1.3) become

$$(9.2) \qquad P(0) = P \text{ and } P(t_0) = I,$$

resp. By (8.2), (9.1) tells us that

$$P(s) \prec P(u) \quad \text{whenever} \quad s \leq u \quad (0 \leq s,\ u \leq t_0),$$

i.e., P (s) is a monotone increasing function of s. Consequently, by (8.1),

$$\mathrm{co} \langle P(s) \rangle \subset \mathrm{co} \langle P(u) \rangle \quad \text{whenever} \quad s \leq u \quad (0 \leq s,\ u \leq t_0),$$

so that, by (9.2), the (ordered!) triangle P (s) expands continuously from P until it reaches I and the corresponding vertices coincide.

This, then, is the geometry that was lurking in the formulas of §1. The Chapman-Kolmogorov equation (1.1), which represents the Markov property in the probabilistic interpretation, turns out to be just a kinematic constraint. Moreover, the passage from the analytical formulation of the embedding problem to its geometrical interpretation can be reversed. We can start with the geometrical problem, defining continuity in an obvious way, and use the results of §7 and §8 to arrive at the analytical formulation given in §1.

The embedding problem for $3\times 3$ stochastic matrices is thus completely equivalent to the following

Locus Problem. <u>What is the locus of the ordered triple of points ($\underline{p}_1, \underline{p}_2, \underline{p}_3$) such that the ordered triangle $\langle P\rangle$ with these points as its vertices can be espanded continuously until it coincides with a given ordered triangle $\langle I\rangle$?</u>

Obviously, the problem can be reformulated to ask for the ultimate position of contracting triangles that start start off at $\langle I\rangle$.

When we trace trough the developments of §2 and §3, we see that the locus problem is equivalent to the reachability problem for the control system (K2), when the "natural parameter"

$$-\log\left[\text{area of } \langle \hat{P}(s)\rangle\right]$$

is used as a time scale. Remarkably, this time scale also has a probabilistic interpretation in terms of the expectation of certain ancillary random variables connected with the Markov process associated with the embedding family $\hat{p}(s,t)$.

The geometrical formulation of the embedding problem, and its connection with probability and control, I presented in a seminar in the Dept. of Computing and Control, at Imperial College, London, in April, 1972. The formulation is so simple that Søren Johansen was later able to present it on Danish television in a program intended to illustrate what type of problems come up in pure mathematics.

§ 10. <u>Bang-bang controls in the restricted embedding problem</u>.

The formulation of the embedding problem just given bypasses tha differential equations and attaches itself directly to the formulation in terms of the functional equation (1.1) given in §1 . Nevertheless, it is interesting to see what the notion of bang-bang control means in the geometrical problem.

As we have already observed, bang-bang controls give rise to responses which are finite products of elementary matrices, i.e., stochastic matrices which have only a single non-zero off-diagonal element. The effect of an elementary matrix on a figure $\langle P \rangle$ is easily seen to be that it moves just one vertex of P along a side joining it to one or the other of remaining vertices. A bang-bang control therefore consist of a succession of such simple moves, one at a time.

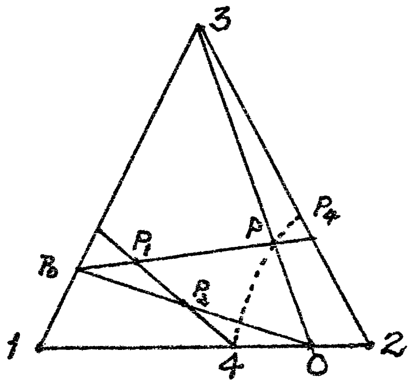
We now formulate a restricted embedding problem.

Restricted Problem. Given two points $p_1$ and $p_2$ in $\langle I \rangle$, what is the locus of points $p_3$ such that the ordered triangle $\langle P \rangle$ with these points as its vertices can be expanded continuously until it coincides with $\langle I \rangle$?

It is immediate from the foregoing that the locus of admissible points $p_3$ is starlike about $p_1$ and $p_2$, and it may be empty.

Clearly, if we knew the solution of the restricted embedding problem for all different positions of $p_1$ and $p_2$, we would know the solution of the original embedding problem.

The use of bang-bang controls in the restricted problem leads to an interesting construction. Suppose that $p_1$, $p_2$ and $\langle I \rangle =$ 1,2,3 have the configuration given in the diagram. The line through $p_1$ and $p_2$ meets the side 12 at a point labled 4. Let 0 be any point on the segment 24 and draw a ray from 0 through $p_2$; it will meet the side 13 at a point $p_0$. Now draw a ray from $p_0$ through $p_1$; it will intersect the line through 3 and 0 at a point p.




I claim that p is an admissible value of $p_3$ and the resulting triangle <P> can be expanded to <I> by a bang-bang control in five moves. The proof is simple. Start with the triangle <I> and move 2 to 0, 1 to $p_0$, 3 to p, 0 to $p_2$ and $p_0$ to $p_1$, in that order, always along line segments. In this way, the triangle <I> ends up coincident with <P>. Reversing the steps then expands <P> to <I>, as required.

As 0 varies from 4 to 2, p sweeps out a continuous curve that reaches from 4 to a certain point $p_4$ on 23. The curve is well-known: it is an arc of a hyperbola, and the construction is due to Braiken-ridge and Maclaurin, around 1733. The construction places the pencil of lines through 3 in projective relationship with the pencil through $p_1$; the points of intersection of corresponding lines realize Steiner's definition of a conic, 1832.

It can be verified that the segment $4p_4$ is precisely the locus of points in which

$$p_{11}p_{22}p_{33} = \det P,$$

in an obvious notation. (Since $p_1$ and $p_2$, the first two rows of P, are fixed, this is a linear equation for $p_3$.) The discussion of the determinal inequality in §6 tells us that equality can hold for an embeddable matrix only at the endpoints. If $p_3 = 4$, the triangle P is degenerate, but there are points

$p_3$ arbitrarily near to 4 which are admissible. It follows that the set of embeddable matrices is not convex, for the section in which $p_1$ and $p_2$ are constant is not convex (if it were, its closure would be convex).

Every point $p_3$ in $\langle I \rangle$ that lies inside the connected domain bounded by the line through $p_1$ and $p_2$ and the arc through 4 and $p_4$ gives rise to a triangle $\langle P \rangle$ that can be expanded to $\langle I \rangle$ in six moves. To see this, just draw the point $p_3$ away from $p_1$ on a rectilinear path until it reaches the boundary : the resulting triangle can then be expanded to $\langle I \rangle$ in five moves, as the reader can check. Actually, the domain indicated is starlike with respect to $p_1$. The crux of the proof is to realize that no line through $p_1$ can meet the arc from 4 to $p_4$ in more than one point. But that is clear, since $p_1$ already belongs to the other branch of the conic.
It can likewise be shown that bang-bang controls will work when $p_3$ is located on the line segment joining $p_2$ to $p_4$ or on the segment that joins 4 to the intersection of the line through $1p_1$ with 23.

The results of the section date from the Spring of 1972. Apart from the discussion of the determinal equality and the identification of the conic, they were rediscovered by Johansen and Ramsey [10] who employed them in an attack on the bang-bang conjecture of §4.

## §11. A characterization of the reachable set.

In §4 we pointed out that set of all reachable matrices is bounded and closed relative to GL(n) and in§5 we remarked that finite products of elementary matrices are dense in the reachable set.

These properties can be used to characterize the reachable set, as follows.

Suppose that we can find a set R with these properties:

1) I belongs to R

2) $KR \in R$ for any non-singular elementary matrix K

3) R is closed relative to GL(n)

4) every matrix in R is reachable from I;

then R is the reachable set from I.

To see this, it is enough to observe that 1) and 2) imply that finite products of elementary matrices belong to R , hence R is dense in the reachable set, while by 3), R is closed, so that it contains the reachable set. But the latter set also contains R, because of 4) hence they coincide.

The advantage of this scheme is that it allows us to test whether an explicitly given set R is the reachable set or not.

## 12. Recent work an the bang-bang conjecture

Recently, Johansen has exploited a variant of the above procedure to characterize the set of $3\times 3$ matrices reachable in $t_o$ for $t_o$ in the interval $0, \log_e 2$ . For R he takes a closed set which he can describe explicitly and which has the property that every matrix in it can be expressed as the product of at most six elementary matrices. Then, using results on the restricted imbedding problem, he establishes that the remaining properties 1) and 2 cited above are valid, provided that det K is sufficiently large. This allows him to conclude that the described set coincides with the set reachable in $t_o \leq \log_e 2$ , and at the same time it establishes the bang-bang conjecture, first for matrices corresponding to $t_o$ in this interval, but then for the whole reachable set just by iteration, where now the number of factors is proportional to $t_o$.

It is of some interest to observe that in adopting this approach, Johansen did not have to establish a priori that his set R contained all matrices reachable in $t_o \leq \log_e 2$ that are the product of six elementary matrices. That is a consequence of his final result.

It may very well be that a modification of this procedure woned lead to a proof of the "strong bang-bang conjecture" that any reachable matrix at all can be expressed as the product of at most six elementary matrices*.

Then an algebraic decomposition theorem will have been established by geometric means.

*Note added August 10, 1978

I believe that I can now establish geometrically this conjecture by going back to the criterion of § 11.

References

1. Brown, D.R., on clans of non negative matrices, Proc. Amer Math. Soc. 15 (1964), 671-674.
2. Elfving, G., Über die Interpolation von Markoffschen Ketten, Soc. Sci. Fennica Comment. Phys. Math 10, No 3 (1938), 1-8
3. Fréchet, M., Solution continue la plus général d'une équation fonctionelle de la théorie des probabilités "en chaine",Bull. Soc. Math. France 60 (1932), 242-280.
4. Goodman, G.S., On a theorem of Scorza-Dragoni and its application to optimal control, in "Math. Theory of Control", Balakrishnan & Neustadt eds., Academic Press, N.Y. (1967), 165-180.
5. Goodman, G.S., An intrinsic time for non-stationary finite Markov Chains, Zeit f. Warsch. 16 (1970), 164-170.
6. Goodman, Control theory in transformation semigroups, in "Geometric Methods in System Theory", Mayne & Brockett eds., Reidel, Dordrecht. (1973), 215-226.
7. Goodman, G.S., A geometrical formulation of the embedding problem for stochastic matrices, to appear.
8. Johansen, S., A central limit theorem for finite semigroups and its application to the embedding problem for finite state Markov chains, Zeit. F. Warsch. 26 (1973), 191-195.
9. Johansen, S., The bang-bang problem for stochastic matrices,Zeit, f. Warsch. 26 (1973), 191-195.
10. Johansen, S. & Ramsey, F., A representation theorem for imbeddable $3\times3$ stochastic matrices, Preprint n. 5 , Aug 1973, Inst. of Math. Stat., Univ. of Copenhagen.
11. Krener, A., A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems,SIAM J. Control(1971)43-52
12. Lee, E. B. & Markus, L., Optimal control for nonlinear processes Arch. Rat. Mech. Anal. 8 (1961), 36-58.
13. Lee, E. B. & Markus, L., "Foundations of Optimal Control Theory", Wiley, New York (1967).
14. Loewner, C., On semigroups in analysis and geometry, Bull Amer. Math. Soc. (1964), 1-15.
15. Roxin E., The existence of optimal controls, Mich. Math. J. 2 (1962), 109-119.