Lectures on Diffusion Problems and Partial Differential Equations

By

S. R. S. Varadhan

Notes by

Pl. Muthuramalingam
Tara R. Nanda

Tata Institute of Fundamental Research, Bombay
1989

Author

S. R. S. Varadhan

Courant Institute of Mathematical Sciences
251, Mercer Street
New York. N. Y. 10012.
U.S.A.

ISBN 3-540-08773-7. Springer-Verlag, Berlin, Heidelberg. New York
ISBN 0-387-08773-7. Springer-Verlag, New York. Heidelberg. Berlin

Printed by N. S. Ray at the Book Centre Limited
Sion East, Bombay 400 022 and published by H. Goetze
Springer-Vertal, Heidelberg, West Germany

PRINTED IN INDIA

Contents

Preface

THESE ARE NOTES based on the lectures given at the T.I.F.R. Centre, Indian Institute of Science, Bangalore, during July and August of 1977. Starting from Brownian Motion, the lectures quickly got into the areas of Stochastic Differential Equations and Diffusion Theory. An attempt was made to introduce to the students diverse aspects of the theory. The last section on Martingales is based on some additional lectures given by K. Ramamurthy of the Indian Institute of Science. The author would like to express his appreciation of the efforts by Tara R. Nanda and PL. Muthuramalingam whose dedication and perseverance has made these notes possible.

S.R.S. Varadhan

1. The Heat Equation

LET US CONSIDER the equation 1

$$u_t - \frac{1}{2}\Delta u = 0 \tag{1}$$

which describes (in a suitable system of units) the temperature distribution of a certain homogeneous, isotropic body in the absence of any heat sources within the body. Here

$$u \equiv u(x_1, \ldots, x_d, t); \quad u_t \equiv \frac{\partial u}{\partial t}; \quad \Delta u = \sum_{i=1}^{d} \frac{\partial^2 u}{\partial x_i^2},$$

t represents the time ranging over $[0, \infty)$ or $[0, T]$ and $x \equiv (x_1 \ldots x_d)$ belongs to $\mathbb{R}^d$.

We first consider the initial value problem. It consists in integrating equation (1) subject to the initial condition

$$u(0, x) = f(x). \tag{2}$$

The relation (2) is to be understood in the sense that

$$\operatorname*{Lt}_{t \to 0} u(t, x) = f(x).$$

Physically (2) means that the distribution of temperature throughout the body is known at the initial moment of time.

We assume that the solution u has continuous derivatives, in the space coordinates upto second order inclusive and first order derivative in time.

It is easily verified that

$$u(t,x) = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{2t}\right); \quad |x|^2 = \sum_{i=1}^{d} x_i^2, \tag{3}$$

2 satisfies (1) and

$$u(0,x) = \underset{t\to 0}{\mathrm{Lt}}\, u(t,x) = \delta(x) \tag{4}$$

Equation (4) gives us a very nice physical interpretation. The solution (3) can be interpreted as the temperature distribution within the body due to a unit sourse of head specified at $t = 0$ at the space point $x = 0$. The linearity of the equation (1) now tells us that (by superposition) the solution of the initial value problem may be expected in the form

$$u(t,x) = \int\limits_{\mathbb{R}^d} f(y)p(t,x-y)dy, \tag{5}$$

where

$$p(t,x) = \frac{1}{(2\pi t)^{d/2}} \exp -\frac{|x|^2}{2t}.$$

Exercise 1. Let $f(x)$ be any bounded continuous function. Verify that $p(t,x)$ satisfies (1) and show that

(a) $\int p(t,x)dx = 1, \ \forall t > 0$;

(b) $\underset{t\to 0+}{\mathrm{Lt}} \int p(t,x)f(x)dx = f(0)$;

(c) using (b) justify (4). Also show that (5) solves the initial value problem.

(Hints: For (a) use $\int\limits_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$. For part (b) make the substitution $y = \dfrac{x}{\sqrt{\pi}}$ and apply Lebesgue dominated convergence theorem).

Since equation (1) is linear with constant coefficients it is invariant under time as well as space translations. This means that translates of
3 solutions are also solutions. Further, for $s \geq 0$, $t > 0$ and $y \in \mathbb{R}^d$,

$$u(t,x) = \frac{1}{[2\pi(t+s)]^{d/2}} \exp -\frac{|x-y|^2}{2(t+s)} \tag{6}$$

and for $t > s$, $y \in \mathbb{R}^d$,

$$u(t,x) = \frac{1}{[2\pi(t-s)]^{d/2}} \exp -\frac{|x-y|^2}{2(t-s)} \tag{7}$$

are also solutions of the heat equation (1).

The above method of solving the initial value problem is a sort of trial method, viz. we pick out a solution and verify that it satisfies (1). But one may ask, how does one obtain the solution? A partial clue to this is provided by the method of Fourier transforms. We pretend as if our solution $u(t,x)$ is going to be very well behaved and allow all operations performed on u to be legitimate.

Put $v(t,x) = \hat{u}(t,x)$ where $\widehat{\ }$ stands for the Fourier transform in the space variables only (in this case), i.e.

$$v(t,x) = \int\limits_{\mathbb{R}^d} u(t,y) e^{i\ x-y} dy.$$

Using equation (1), one easily verifies that

$$v_t(t,x) = \frac{1}{2}|x|^2 v(t,x) \tag{8}$$

with

$$v(0,x) = \hat{f}(x). \tag{9}$$

The solution of equation (8) is given by

$$v(t,x) = \hat{f}(x) e^{-t|x|^2/2}. \tag{10}$$

We have used (9) in obtaining (10).

Exercise 2. Verify that 4

$$\hat{p}(t,x) = \exp -\left(\frac{t|x|^2}{2}\right).$$

Using Exercise 2, (10) can be written as

$$v(t,x) = \hat{u}(t,x) = \hat{f}(x)\hat{p}(t,x). \tag{11}$$

The right hand side above is the product of two Fourier transforms and we know that the Fourier transform of the convolution of two funtions is given by the product of the Fourier transforms. Hence $u(t,x)$ is expected to be of the form (5).

Observe that if f is non-negative, then u is nonnegative and if f is bounded by M then u is also bounded by M in view of part (a) of Exercise 1.

The Inhomogeneous Equation. Consider the equation

$$v_t - \frac{\Delta v}{2} = g, \quad \text{with} \quad v(0,x) = 0,$$

which describes the temperature within a homogeneous isotropic body in the presence of heat sources, specified as a function of time and space by $g(t,x)$. For $t > s$,

$$u(t,x) = \frac{1}{[2\pi(t-s)]^{d/2}} \exp -\frac{|x-y|^2}{2(t-s)}$$

is a solution of $u_t(t,x) - \dfrac{1}{2}\Delta u(t,x) = 0$ corresponding to a unit source at $t = s$, $x = y$. Consequently, a solution of the inhomogeneous problem is obtained by superposition.

Let

$$v(t,x) = \int_{\mathbb{R}^d} \int_0^t g(s,y) \frac{1}{[2\pi(t-s)]^{d/2}} \exp\left(-\frac{|x-y|^2}{2(t-s)}\right) dy\, ds$$

5 i.e.

$$v(t,x) = \int_0^t w(t,x,s)ds$$

where

$$w(t,x,s) = \int_{\mathbb{R}^d} g(s,y) \frac{1}{[2\pi(t,s)]^{d/2}} \exp\left(-\frac{|x-y|^2}{2(t-s)}\right) dy.$$

Exercise 3. Show that $v(t,x)$ defined above solves the inhomogeneous heat equation and satisfies $v(0,x) = 0$. Assume that g is sufficiently smooth and has compact support. $v_t - \frac{1}{2}\Delta v = \underset{s\to t}{\mathrm{Lt}}\, w(t,x,s)$ and now use part (b) of Exercise (1).

Remark 1. We can assume g has compact support because in evaluating $v_t - \frac{1}{2}\Delta v$ the contribution to the integral is mainly from a small neighbourhood of the point (t,x). Outside this neighbourhood

$$\frac{1}{[2\pi(t-s)]^{d/2}} \exp\left(-\frac{|x-y|^2}{2(t-s)}\right)$$

satisfies

$$u_t - \frac{1}{2}\Delta u = 0.$$

2. If we put $g(s,y) = 0$ for $s < 0$, we recognize that $v(t,x) = g * p$. Taking spatial Fourier transforms this can be written as

$$v(t,\xi) = \int_0^t g(s,\xi) \exp -\frac{1}{2}(t-s)|\xi|^2 d\xi,$$

or

$$\frac{\partial \hat{v}}{\partial t} = \frac{\partial v}{\partial t} = g(t,\xi) + \frac{1}{2}\Delta v = \left(g(t,\xi) + \frac{1}{2}\Delta v\right).$$

Therefore

$$\frac{\partial v}{\partial t} - \frac{1}{2}\Delta v = g.$$

Exercise 4. Solve $w_t - \frac{1}{2}\Delta w = g$ on $[0,\infty)\times\mathbb{R}^d$ with $w = f$ on $\{0\}\times\mathbb{R}^d$ 6
(Cauchy problem for the heat equation).

Uniqueness. The solution of the Cauchy problem is unique provided the class of solutions is suitably restricted. The uniqueness of the solution is a consequence of the Maximum Principle.

Maximum Principle. *Let u be smooth and bounded on* $[0,T]\times\mathbb{R}^d$ *satisfying*

$$u_t - \frac{\Delta u}{2} \geq 0 \quad \text{in} \quad (0,T]\times\mathbb{R}^d \quad \text{and} \quad u(0,x)\geq 0,\ \forall x\in\mathbb{R}^d.$$

Then

$$u(t,x)\geq 0\ \ \forall, \quad t\in[0,T] \quad \text{and} \quad \forall x\in\mathbb{R}^d.$$

Proof. The idea is to find minima for u or for an auxillary function.

Step 1. Let v be *any* function satisfying

$$v_t - \frac{\Delta v}{2} > 0 \quad \text{in} \quad (0,T]\times\mathbb{R}^d.$$

Claim. v cannot attain a minimum for $t_0\in(0,T]$. Assume (to get a contradiction) that $v(t_0,x_0)\leq v(t,x)$ for some $t_0>0$ and for all $t\in[0,T]$, $\forall x\in\mathbb{R}^d$. At a minimum $v_t(t_0,x_0)\leq 0$, (since $t_0\neq 0$) $\Delta v(t_0,x_0)\geq 0$. Therefore

$$\left(v_t - \frac{\Delta v}{2}\right)(t_0,x_0)\leq 0.$$

Thus, if v has any minimum it should occur at $t_0 = 0$.

Step 2. Let $\epsilon > 0$ be arbitrary. Choose α such that

$$h(t,x) = |x|^2 + \alpha t$$

7 satisfies

$$h_t - \frac{\Delta h}{2} = \alpha - d > 0 \quad (\text{say } \alpha = 2d).$$

Put $v_\epsilon = u + \epsilon h$. Then

$$\frac{\partial v_\epsilon}{\partial t} - \frac{1}{2}\Delta v_\epsilon > 0.$$

As u is bounded, $v_\epsilon \to +\infty$ as $|x| \to +\infty$, v_ϵ must attain a minimum. This minimum occurs at $t = 0$ by Step 1. Therefore,

$$v_\epsilon(t,x) \geq v_\epsilon(0,x_0) \quad \text{for some} \quad x_0 \in \mathbb{R}^d,$$

i.e.

$$v_\epsilon(t,x) \geq u(0,x_0) + \epsilon|x_0|^2 > 0,$$

i.e.

$$u(t,x) + \epsilon h(t,x) > 0, \ \forall \epsilon.$$

This gives

$$u(t,x) \geq 0.$$

This completes the proof. □

Exercise 5. (a) Let L be a linear differential operator satisfying $Lu = g$ on Ω (open in $\mathbb{R}^d$) and $u = f$ on $\partial\Omega$. Show that u is uniquely determined by f and g if and only if $Lu = 0$ on Ω and $u = 0$ on $\partial\Omega$ imply $u = 0$ on Ω.

(b) Let u be a bounded solution of the heat equation $u_t - \frac{1}{2}\Delta u = g$ with $u(0,x) = f(x)$. Use the maximum principle and part (a) to show that u is unique in the class of all bounded functions.

(c) Let

$$g(t) = \begin{cases} e^{-1/t^2}, & \text{if } t > 0, \\ 0, & \text{if } t \leq 0, \end{cases}$$

$$u(t,x) = \sum_{k=0}^{\infty} \frac{g^{(k)}(t/2)x^{2^k}}{(2k)!}, \quad \text{on} \quad R \times R.$$

Then

$$u(0,x) = 0, \quad u_t = \frac{\Delta u}{2}, \quad u \not\equiv 0,$$

i.e. u satisfies

$$u_t - \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0, \quad \text{with} \quad u(0,x) = 0.$$

This example shows that the solution is not unique because, u is not bounded. (This example is due to Tychonoff).

Lemma 1. *Let* $p(t,x) = \frac{1}{(2\pi t)^{d/2}} \exp -\frac{|x|^2}{2t}$ *for* $t > 0$. *Then*

$$p(t,\cdot) * p(s,\cdot) = p(t+s,\cdot).$$

Proof. Let f be any bounded continuous function and put

$$u(t,x) = \int_{\mathbb{R}^d} f(y)p(t,x-y)dy.$$

Then u satisfies

$$u_t - \frac{1}{2}\Delta u = 0, \quad u(0,x) = f(x).$$

Let

$$v(t,x) = u(t+s,x).$$

Then

$$v_t - \frac{1}{2}\Delta v = 0, \quad v(0,x) = u(s,x).$$

This has the unique solution

$$v(t,x) = \int u(s,y)p(t,x-y)dy.$$

Thus

$$\int_{\mathbb{R}^d} f(y)p(t+s,x-y)dy = \iint f(z)p(s,y-z)p(t,x-y)dz\,dy.$$

9 This is true for all f bounded and continuous. We conclude, therefore, that

$$p(t, \cdot) * p(s, \cdot) = p(t + s, \cdot).$$

□

Exercise 6. Prove Lemma 1 directly using Fourier transforms.

It will be convenient to make a small change in notation which will be useful later on. We shall write $p(s, x, t, y) = p(t-s, y-x)$ for every x, y and $t > s$. $p(s, x, t, y)$ is called the *transition probability*, in dealing with Brownian motion. It represents the probability density that a "Brownian particle" located at space point x at time s moves to the space point y at a later time t.

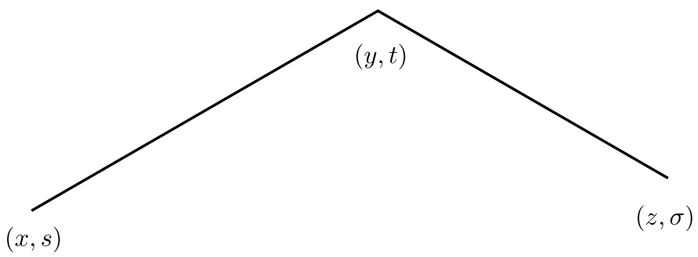
Note . We use the same symbol p for the transition probability; it is function of four variables and there will not be any ambiguity *in* using the same symbol p.

Exercise 7. Verify that

$$\int\limits_{\mathbb{R}^d} p(s, x, t, y) p(t, y, \sigma, z) dy = p(s, x, \sigma, z), \ \ s < t < \sigma.$$

(Use Exercise 6).

Remark . The significance of this result is obvious. The probability that the particle goes from x at time s to z at time σ is the sum total of the probabilities, that the particle moves from x at s to y at some intermediate time t and then to z at time σ.



10 In this section we have introduced Brownian Motion corresponding to the operator $\frac{1}{2}\Delta$. Later on we shall introduce a more general diffusion process which corresponds to the operator $\frac{1}{2}\sum a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum b_j \frac{\partial}{\partial x_j}$.

2. Kolmogorov's Theorem

Definition. LET $(\Omega, \mathscr{B}, P)$ BE A probability space. A stochastic process in $\mathbb{R}^d$ is a collection $\{X_t : t \in I\}$ of $\mathbb{R}^d$-valued random variables defined on $(\Omega, \mathscr{B})$. **11**

Note 1. I will always denote a subset of $\mathbb{R}^+ = [0, \infty)$.

2. X_t is also denoted by $X(t)$.

Let $\{X_t : t \in I\}$ be a stochastic process. For any collection t_1, $t_2, \ldots, t_k$ such that $t_i \in I$ and $0 \leq t_1 < t_2 < \ldots < t_k$ and any Borel set Λ, in $\mathbb{R}^d \times \mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (k times),. define

$$F_{t_1} \ldots t_k(\Lambda) = P(w \in \Omega : (X_{t_1}(w), \ldots, X_{t_k}(w)) \in \Lambda).$$

If

$$\{t_1, \ldots, t_k\} \subset \{s_1, \ldots, s_\ell\} \subset I, \quad \text{with} \quad l \geq k$$

such that

$$s_1^{(0)} < \ldots < s_{n_0}^{(0)} < t_1 < s_1^{(1)} \ldots < s_n^{(1)} < t_2 \ldots < t_k < s_1^{(k)} \ldots < s_{n_k}^{(k)},$$

let then

$$\pi : \mathbb{R}^d \times \cdots \times \mathbb{R}^d (1 \text{ times}) \to \mathbb{R}^d \times \cdots \times \mathbb{R}^d (k \text{ times})$$

be the canonical projection. If $E_{t_i} \subset \mathbb{R}^d$ is any Borel set in $\mathbb{R}^d$, $i = 1, 2, \ldots, k$, then

$$\pi^{-1}(E_{t_1} \times \cdots \times E_{t_k}) = \mathbb{R}^d \times \cdots \times E_{t_1} \times \mathbb{R}^d \times \cdots \times E_{t_2} \times \cdots \times \mathbb{R}^d$$

(l times). The following condition always holds.

(*) $\quad E_{t_1} \ldots t_k(E_{t_1} \times \cdots \times E_{t_k}) = F_{s_1} \ldots s_1(\Pi^{-1}(E_{t_1} \times \cdots \times E_{t_k}))$.

If $(*)$ holds for an arbitrary collection $\{F_{t_1} \ldots t_k : 0 \leq t_1 < t_2 \ldots < t_k\}$ 12
$(k = 1, 2, 3 \ldots)$ of distributions then it is said to satisfy the *consistency condition.*

Exercise 1. (a) Verify that $F_{t_1} \ldots t_k$ is a probability measure on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (k times).

(b) Verify $(*)$. (If B_m denotes the Borel σ field of $\mathbb{R}^m$, $B_{m+n} = B_m \times B_n$).

The following theorem is a converse of Exercise 1 and is often used to identify a stochastic process with a family of distributions satisfying the consistency condition.

Kolmogorov's Theorem.

Let $\{F_{t_1,t_2,\ldots t_k} 0 \leq t_1 < t_2 < \ldots < t_k < \infty\}$ be a family of probability distributions (on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$, k times, $k = 1, 2, \ldots$) satisfying the consistency condition. Then there exists a measurable space $(\Omega_k, \mathbb{B})$, a unique probability measure P an $(\Omega_k, \mathscr{B})$ and a stochastic process $\{X_t : 0 \leq t < \infty\}$ such that the family of probability measures associated with it is precisely

$$\{F_{t_1,t_2,\ldots t_k} : 0 \leq t_1 < t_2 < \ldots < t_k < \infty\}, \quad k = 1, 2, \ldots.$$

A proof can be found in the APPENDIX. We mention a few points about the proof which prove to be very useful and should be observed carefully.

13 1. The space Ω_k is the set of all $\mathbb{R}^d$-valued functions defined on $[0, \infty)$:
$$\Omega_K = \prod_{t \in [0,\infty)} \mathbb{R}^d$$

2. The random variable X_t is the t^{th}-projection of Ω_K onto $\mathbb{R}^d_t$.

3. $\mathscr{B}$ is the smallest σ-algebra with respect to which all the projections are measurable.

4. P given by

$$P(w : X_{t_1}(w) \in A_1, \ldots X_{t_k}(w) \in A_k) = F_{t_1 \ldots t_k}(A_1 \times \cdots \times A_k)$$

   where A_i is a Borel set in $\mathbb{R}^d$, is a measure on the algebra generated by $\{X_{t_1}, \ldots X_{t_k}\}(k = 1, 2, 3 \ldots)$ and extends uniquely to $\mathscr{B}$.

Remark . Although the proof of Kolmogorov's theorem is very constructive the space Ω_K is too "large" and the σ-algebra $\mathscr{B}$ too "small" for practical purposes. In applications one needs a "nice" collection of $\mathbb{R}^d$-valued functions (for example continuous, or differentiable functions), a "large" σ-algebra on this collection and a probability measure concentrated on this family.

3. The One Dimensional Random Walk

BEFORE WE TAKE up Brownian motion, we describe a one dimensional random walk which in a certain limiting case possesses the properties of Brownian motion. 14

Imagine a person at the position $x = 0$ at time $t = 0$. Assume that at equal intervals of time $t = \tau$ he takes a step h either along the positive x axis or the negative x axis and reaches the point $x(t) = x(t - \tau) + h$ or $x(t) = x(t - \tau) - h$ respectively. The probability that he takes a step in either direction is assumed to be $1/2$. Denote by $f(x, t)$ the probability that after the time $t = n\tau$ (n intervals of time τ) he reaches the position x. If he takes m steps to the right (positive x-axis) in reaching x then there are ${}^nC_m$ possible ways in which he can achieve these m steps. Therefore, the probability $f(x, t)$ is ${}^nC_m(\frac{1}{2})^n$.

$f(x, t)$ satisfies the difference equation

$$f(x, t + \tau) = \frac{1}{2} f(x - h, t) + \frac{1}{2} f(x + h, t) \tag{1}$$

and

$$x = h(m - (n - m)) = (2m - n)h. \tag{2}$$

To see this one need only observe that to reach $(x, t + \tau)$ there are two ways possible, viz. $(x - h, t) \to (x, t + \tau)$ or $(x + h, t) \to (x, t + \tau)$ and the probability for each one of these is $1/2$. Also note that by definition of f,

$$f(h,\tau) = \frac{1}{2} = f(-h,\tau), \tag{3}$$

15 so that

$$f(x, t+\tau) = f(h,\tau) f(x-h,t) + f(-h,\tau) f(x+h,t). \tag{4}$$

The reader can identify (4) as a "discrete version" of convolution. By our assumption,

$$f(0,0) = 1, \quad f(x,0) = 0 \quad \text{if} \quad x \neq 0. \tag{5}$$

We examine equation (1) in the limit $h \to 0$, $\tau \to 0$. To obtain reasonable results we cannot let h and τ tend to zero arbitratily. Instead we assume that

$$\frac{h}{\tau} \to 1 \quad \text{as} \quad h \to 0 \quad \text{and} \quad \tau \to 0. \tag{6}$$

The physical nature of the problem suggests that (6) should hold. To see this we argue as follows. Since the person is equally likely to go in either direction the average value of x will be 0. Therefore a reasonable measure of the "progress" made by the person is either $|x|$ or x^2. Indeed, since x is a random variable (since m is one) one gets, using (2),

$$E(x) = 2E(m) - n = 0, \quad E(x^2) = h^2 E((2m-n)^2) = h^2 n.$$

(Use $\sum\limits_{m=0}^{n} m\,{}^nC_m \left(\frac{1}{2}\right)^n = \frac{n}{2}$, $\sum\limits_{m=0}^{n} {}^nC_m \left(\frac{1}{2}\right)^n = \frac{n(n+1)}{4}$)

Thus

$$E\left\{\frac{x^2}{t}\right\} = \frac{1}{t} E(x^2) = \frac{h^2 n}{n\tau} = \frac{h^2}{\tau},$$

and as t becomes large we expect that the average distance covered per
16 unit time remains constant. (This constant is chosen to be 1 for reasons that will become apparent later). This justifies (6). In fact, a simple argument shows that if $\frac{h^2}{\tau} \to 0$ or $+\infty$, x may approach $+\infty$ in a finite time which is physically untenable.

(1) now gives

$$f(x,t+\tau) - f(x,t) = \frac{1}{2}\{f(x-h,t) - f(x,t) + f(x,h,t) - f(x,t)\}.$$

Assuming sufficient smoothness on f, we get in the limit as $h, \tau \to 0$ and in view of (6),

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2} \tag{7}$$

(to get the factor $1/2$ we choose $\frac{h^2}{\tau} \to 1$). This is the equation satisfied by the probability density f. The particle in this limit performs what is known as *Brownian motion* to which we now turn our attention.

References.

[1] GNEDENKO: *The theory of probability*, Ch. 10.

[2] *The Feynman Lectures on physics*, Vol. 1, Ch. 6.

4. Construction of Wiener Measure

ONE EXAMPLE WHERE the Kolmogorov construction yields a probability measure concentrated on a "nice" class Ω is the Brownian motion. 17

Definition. A Brownian motion with starting point x is an $\mathbb{R}^d$-valued stochastic process $\{X(t) : 0 \leq t < \infty\}$ where

(i) $X(0) = x = \text{constant}$;

(ii) the family of distribution is specified by

$$F_{t_1} \ldots t_k(A) = \int\limits_A p(0, x, t_1, x_1) p(t_1, x_1, t_2, x_2) \ldots$$
$$p(t_{k-1}, x_{k-1}, t_k, x_k) dx_1 \ldots dx_k$$

for every Borel set A in $\mathbb{R}^d \times \cdots \mathbb{R}^d$ (k times).

N.B. The stochastic process appearing in the definition above is the one given by the Kolmogorov construction.

It may be useful to have the following picture of a Brownian motion. The space Ω_k may be thought of as representing particles performing Brownian movement; $\{X_t : 0 \leq t < \infty\}$ then represents the trajectories of these particles in the space $\mathbb{R}^d$ as functions of time and $\mathscr{B}$ can be considered as a representation of the observations made on these particles.

Exercise 2. (a) Show that $F_{t_1 \ldots t_k}$ defined above is a probability measure on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (k times).

(b) $\{F_{t_1 \ldots t_k} : 0 \leq t_1 < t_2 < \ldots t_k < \infty\}$ satisfies the consistency condition. (Use Fubini's theorem).

18 (c) $X_{t_1} - x, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}$ are independent random variables and if $t > s$, then $X_t - X_s$ is a random variable whose distribution density is given by

$$p(t-s, y) = \frac{1}{[2\pi(t-s)]^{d/2}} \exp\left(-\frac{1}{2}(t-s)^{-1}|y|^2\right).$$

(Hint: Try to show that $X_{t_1} - x$, $X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}$ have a joint distribution given by a product measure. For this let ϕ be any bounded real measurable function on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (k times). Then

$$\underset{X_{t_1}-x, X_{t_2}-X_{t_1}, \ldots, X_{t_k}-X_{t_{k-1}}}{E(\phi(Z_1, \ldots, Z_k))} = \underset{X_{t_1}, \ldots, X_{t_k}}{E} (\phi(Z_1 - x, \ldots, Z_k - Z_{k-1}))$$

where $\underset{X_{t_1} \ldots X_{t_k}}{E(\phi)}$ is the expectation of ϕ with respect to the joint distribution of $(X_{t_1} \ldots, X_{t_k})$. You may also require the change of variable formula).

Problem. Given a Brownian motion with starting point x our aim is to find a probability P_x on the space $\Omega = C([0, \infty); \mathbb{R}^d)$ of all continuous funcitons from $[0, \infty) \to \mathbb{R}^d$ which induces the Brownian motion. We will thus have a *continuous realisation* to Brownian motion. To achieve this goal we will work with the collection $\{F_{t_1, \ldots, t_k} : 0 \leq t_1 < t_2 < \ldots < t_k\}$ where $t_i \in D$, a countable dense subset of $[0, \infty)$.

19 **Step 1.** The first step is to find a probability measure on a "smaller" space and lift it to $C([0, \infty); \mathbb{R}^d)$. Let

$$\Omega = C([0, \infty); \mathbb{R}^d),$$

D a countable dense subset of $[0, \infty)$; $\Omega(D) = \{F : D \to \mathbb{R}^d\}$ where f is uniformly continuous on $[0, N] \cap D$ for $N = 1, 2, \ldots$. We equip Ω with the topology of uniform convergence on compact sets and $\Omega(D)$ with the topology of uniform convergence on sets of the form $D \cap K$

where $K \subset [0, \infty)$ is compact; Ω and $\Omega(D)$ are separable metric spaces isometric to each other.

Exercise 3. Let

$$p_n(f, g) = \sup_{0 \le t \le n} |f(t) - g(t)| \quad \text{for} \quad f, g \in \Omega$$

and

$$p_{n,D}(f, g) = \sup_{\substack{0 \le t \le n \\ t \in D}} |f(t) - g(t)| \quad \text{for} \quad f, g \in \Omega(D).$$

Define

$$\rho(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(f, g)}{1 + p_n(f, g)}, \ \forall f, g \in \Omega,$$
$$\rho_D(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_{n,D}(f, g)}{1 + p_{n,D}(f, g)}, \ \forall f, g \in \Omega(D).$$

Show that

(i) $\{f_n\} \subset \Omega$ converges to f if and only if $f_n \to f$ uniformly on compact subsets of $[0, \infty)$;

(ii) $\{f_n\} \subset \Omega(D)$ converges to f if and only if $f_{n|D \cap K} \to f_{|D \cap K|}$ uniformly for every compact subset K of $[0, \infty)$;

(iii) $\{(P_1, \ldots, P_d)\}$ where P_i is a polynomial with rational coefficients 20
is a countable dense subset of Ω;

(iv) $\{(P_{1D}, \ldots, P_{dD})\}$ is a countable dense subset of $\Omega(D)$;

(v) $\tau : \Omega \to \Omega(D)$ where $\tau(f) = f_{|D}$ is a (ρ, ρ_D)-isometry of Ω onto $\Omega(D)$;

(vi) if $V(f, \epsilon, n) = \{g \in \Omega : p_n(f, g) < \epsilon\}$ for $f \in \Omega$, $\epsilon > 0$ and

$$V_D(f, \epsilon, n) = \{g \in \Omega(D) : p_{n,D}(f, g) < \epsilon\} \quad \text{for} \quad f \in \Omega(D), \epsilon > 0,$$

then

$$\{V(f, \epsilon, n) : f \in \Omega, \epsilon > 0, n = 1, 2 \ldots\}$$

is a base for the topology of Ω and

$$\{V_D(f,\epsilon,n) : f \in \Omega(D), \epsilon > 0, n = 1, 2, \ldots\}$$

is a base for the topology of $\Omega(D)$.

Remark. By Exercise 3(v) any Borel probability measure on $\Omega(D)$ can be lifted to a Borel probability measure on Ω.

2nd Step. Define the modulus of continuity $\Delta_D^{T,\delta}(f)$ of a function f on D in the interval $[0,T]$ by

$$\Delta_D^{T,\delta}(f) = \sup\{|f(t) - f(s)| : |t - s| < \delta t, s \in D \cap [0,T]\}$$

As D is countable one has

Exercise 4. (a) Show that $f : \Delta_D^{N,\frac{1}{j}}(f) \leq \frac{1}{k}\}$ is measurable in the σ-algebra generated by the projections

$$\pi_t : \pi\{\mathbb{R}_t^d : t \in D\} \to \mathbb{R}_t^d$$

21 *Proof.* The lemma is equivalent to showing that $\mathscr{B} = \sigma(\mathscr{E})$. As each of the projection $\pi_{t_1 \ldots t_k}$ is continuous, $\sigma(\mathscr{E}) \subset \mathscr{B}$. To show that $\mathscr{B} \subset \sigma(\mathscr{E})$, it is enough to show that $V_D(f,\epsilon,n) \in \mathscr{E}$ because $\Omega(D)$ is separable. (Cf. Exercise 3(iv) and 3(vi)). By definition

$$\begin{aligned} V_D(f,\epsilon,n) &= \{g \in \Omega(D) : P_{n,D}(f,g) < \epsilon\} \\ &= \bigcup_{m=1}^{\infty} \left\{g \in \Omega(D) : p_{n,D}(f,g) \leq \epsilon - \frac{1}{m}\right\} \\ &= \bigcup_{m=1}^{\infty} \{g \in \Omega(D) : |g(t_i) - f(t_i)| \leq \epsilon - \frac{1}{m}, \ \forall t_i \in D \cap [0,n]\}. \end{aligned}$$

The result follows if one observes that each π_{t_i} is continuous. □

Remark 1. The lemma together with Exercise 4(b) signifies that the Kolmogorov probability P_x on $\pi\{\mathbb{R}_t^d : t \in D\}$ is defined on the topological Borel σ-field of $\Omega(D)$.

2. The proof of the lemma goes through if $\Omega(D)$ is replaced by Ω.

Step 3. We want to show that $P_x(\Omega(D)) = 1$. By Exercise 4(b) this is equivalent to showing that $\underset{j\to\infty}{\text{Lt}} P(\Delta_D^{N,1/j}(f) \leq \frac{1}{k}) = 1$ for all N and k. The lemmas which follow will give the desired result.

Lemma (Lévy). *Let $X_1, \ldots X_n$ be independent random variables, $\epsilon > 0$ and $\delta > 0$ arbitrary. If*

$$P(|X_r + X_{r+1} + \cdots + X_\ell| \geq \delta) \leq \epsilon$$

$\forall$ r, ℓ such that $1 \leq r \leq \ell \leq n$, then

$$P(\sup_{1\leq j\leq n} |X_1 + \cdots + X_j| \geq 2\delta) \leq 2\epsilon.$$

(see Kolmogorov's theorem) for every $j = 1, 2, \ldots,$ for every $N = 1, 2, \ldots$ and for every $k = 1, 2, \ldots$. (Hint: Use the fact that the projections are continuous). **22**

(b) *Show that $\Omega(D) = \bigcap_{N=1}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{j=1}^{\infty} \{\Delta_D^{N,\frac{1}{j}}(f) \leq \frac{1}{k}\}$ and hence $\Omega(D)$ is measurable in $\pi\{\mathbb{R}_t^d : t \in D\}$.*

Let $\pi_{t_1 \ldots t_k} : \Omega(D) \to \mathbb{R}^d \times \mathbb{R}^d \times \cdots \mathbb{R}^d$ (k times) be the projections and let

$$\mathscr{E}_{t_1\ldots t_k} = \pi^{-1}_{t_1\ldots t_k}(\underbrace{\mathscr{B}(\mathbb{R}^d)\times \cdots \times\mathscr{B}(\mathbb{R}^d)}_{k \text{ times}}).$$

Put

$$\mathscr{E} = \cup\{\mathscr{E}_{t_1\ldots t_k} : 0 \leqq t_1 < t_2 < \ldots < t_k < \infty; t_i \in D\}.$$

Then, as

$$\mathscr{E}_{t_1\ldots t_k} \cup \mathscr{E}_{s_1\ldots s_1} \subset \mathscr{E}_{\tau_1\ldots\tau_m},$$

where

$$\{t_1 \ldots t_k, s_1 \ldots s_1\} \subset \{\tau_1 \ldots \tau_m\},$$

$\mathscr{E}$ is an algebra. Let $\sigma(\mathscr{E})$ be the σ-algebra generated by $\mathscr{E}$.

Lemma . *Let $\mathscr{B}$ be the (topological) Borel σ-field of $\Omega(D)$. Then $\mathscr{B}$ is the σ-algebra generated by all the projections*

$$\{\pi_{t_i\ldots t_k} : 0 \leq t_1 < t_2 < \ldots < t_k, t_i \in D\}.$$

Remark. By subadditivity it is clear that 23

$$P\left(\sup_{1\leq j\leq n}|X_1+\cdots+X_j|\geq 2\delta\right)\leq n\epsilon.$$

Ultimately, we shall let $n\to\infty$ and this estimate is of no utility. The importance of the lemma is that it gives an estimate independent of n.

Proof. Let $S_j = X_1+\cdots+X_j$, $E=\{\sup_{1\leq j\leq n}|S_j|\geq 2\delta\}$. Put

$$\begin{aligned}
E_1 &= \{|S_1|\geq 2\delta\},\\
E_2 &= \{|S_1|<2\delta,\ |S_2|\geq 2\delta\},\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\\
E_n &= \{|S_j|<2\delta, 1\leq j\leq n-1, |S_n|\geq 2\delta\}.
\end{aligned}$$

Then

$$\begin{aligned}
E &= \bigcup_{j=1}^{n} E_j,\ E_j\cap E_i=\phi \quad \text{if} \quad j\neq i;\\
P\{E\cap(|S_n|\leq\delta) &= P\left(\bigcup_{j=1}^{n}(E_j\cap(|S_n|\leq\delta))\right)\\
&\leq P\left\{\bigcup(E_i\cap(|S_n-S_j|\geq\delta))\right\}\\
&\leq \sum_{j=1}^{n}P(E_j)P(|S_n-S_j|\geq\delta) \quad \text{(by independence)}\\
&\leq \epsilon P(E) \quad \text{(by data)}.\\
&= P\{E\cap(|S_n|>\delta)\}\leq P(|S_n|>\delta)\leq\epsilon \quad \text{(by data)}.
\end{aligned}$$

Combining the two estimates above, we get

$$P(E)\leq\epsilon+\epsilon P(E).$$

24 If $\epsilon>\dfrac{1}{2}$, $2\epsilon>1$. If $\epsilon<\dfrac{1}{2}$, $\dfrac{\epsilon}{1-\epsilon}\leq 2\epsilon$. In either case $P(E)\leq 2\epsilon$. □

Lemma . *Let $\{X(t)\}$ be a Brownian motion, $I \subset [0, \infty)$ be a finite interval, $F \subset I \cap D$ be finite. Then*

$$P_x\left(\operatorname{Sup}_{t,\sigma\in F} |X(t) - X(\sigma)| \geq 4\delta\right) \leq C(d)\frac{|I|^2}{\delta^4},$$

where $|I|$ is the length of the interval and $C(d)$ a constant depending only on d.

Remark. Observe that the estimate is independent of the finite set F.

Proof. Let $F = \{t_i : 0 \leq t_1 < t_2 < \ldots < t_k < \infty\}$.

Put

$$X_1 = X(t_2) - X(t_1), \ldots, X_{k-1} = X(t_k) - X(t_{k-1}).$$

Then $X_1 \ldots X_{k-1}$ are independent (Cf. Exercise 2(c)). Let

$$\epsilon = \sup_{1\leq r\leq 1\leq k-1} P_x(|X_r + X_{r+1} + \cdots + X_1| \geq \delta).$$

Note that

$$\begin{aligned}
&P_x(|X_r + \cdots + X_1| \geq \delta) = P(|X(t') - X(t'')| \geq \delta) \text{ for some } t', t'' \text{ in } F\\
&\quad \leq \frac{E(|X(t') - X(t'')|^4)}{\delta^4} \quad \text{(see Tchebyshey's inequality in Appendix)}\\
(*)&\quad \leq \frac{C'(t'' - t')}{\delta^4} \quad (C'' = \text{constant})\\
&\quad \leq \frac{C'|I|^2}{\delta^4}.
\end{aligned}$$

Therefore $\epsilon \leq \dfrac{C'|I|^2}{\delta^4}$. Now

$$\begin{aligned}
&P_x(\sup_{t,\sigma\in P} |X(t) - X(\sigma)| \geq 4\delta)\\
&\quad P_x(\sup_{1\leq i\leq k} |X(t_i) - X(t_1)| \geq 2\delta)\\
&= P_x(\sup_{i\leq j\leq k-1} |X_1 + \cdots + X_j| \geq 2) \leq 2\epsilon \quad \text{(by previous lemma)}\\
&\quad \frac{2C'|I|^2}{\delta^4} = \frac{C|I|^2}{\delta^4}.
\end{aligned}$$

□

Exercise 5. Verify (∗).

(Hint: Use the density function obtained in Exercise 2(c) to evaluate the expectation and go over to "popular" coordinates. (The value of C' is $d(2d+1)$)).

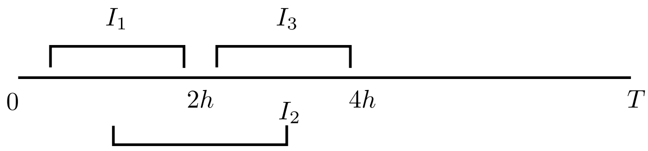
Lemma .

$$P_x\left\{\sup_{\substack{|t-s|\leq h\\ t,s\in[0,t]\cap D}} |X(t)-X(s)|>\rho\right\} = P_x(\Delta_D^{T,h}>\rho)$$

$$\leq \phi(T,\rho,h) = C\frac{h}{\rho^4}\left(\left[\frac{T}{h}\right]+1\right).$$

Note that $\phi(T,\rho,h)\to 0$ as $h\to 0$ for every fixed T and ρ.

Proof. Define the intervals $I_1, I_2, \ldots$ by



$$I_k = [(k-1)h, (k+1)h]\cap(0,T],\ k=1,2,\ldots.$$

Let $I_1, I_2, \ldots I_r$ be those intervals for which

$$I_j\cap[0,T]\neq\phi\,(j=1,2,\ldots,r).$$

26 Clearly there are $[\frac{T}{h}]+1$ of them. If $|t-s|\leq h$ then t, $s\in I_j$ for some j, $1\leq j\leq r$. Write $D=\bigcup\limits_{n=1}^{\infty} F_n$ where $F_n\subset F_{n+1}$ and F_n is finite. Then

$$P_x\left\{\sup_{\substack{|t-s|\leq h\\ t,s\in[0,T]\cap D}} |X(t)-X(s)|>\rho\right\} = P_x\left\{\bigcup_{n=1}^{\infty}\left(\sup_{\substack{|t-s|\leq h\\ t,s\in D\cap F_n}} |X(t)-X(s)|>\rho\right)\right\}$$

$$= \sup_n P_x\left\{\sup_j \sup_{t,s\in F_n} (|X_{I_j}(t)-X(s)|>\rho)\right\}$$

$$
\begin{aligned}
&\le \sup_n \sum_{j=1}^{r} P_x\left(\sup_{t,s\in F_n}(|X_{I_j}(t)-X(s)|>\rho)\right)\\
&\sup_n\left(\left[\frac{T}{h}\right]+1\right)\frac{C(2h)^2}{(\rho/4)^4} \quad \text{by the last lemma}\\
&\le \phi(T,\rho,h).
\end{aligned}
$$

□

Theorem . $P_x(\Omega(D)) = 1$.

Proof. It is enough to show that

$$\operatorname*{Lt}_{j\to\infty} P_x\left(\Delta_D^{N,\frac{1}{j}}(f)\le\frac{1}{k}\right)=1 \quad \text{(See Exercise 4(b)).}$$

But this is guaranteed by the previous lemma. □

Remark. 1. It can be shown that the outer measure of Ω is 1.

2. Ω is not measurable in $\prod\limits_{t\ge 0}\mathbb{R}^d_t$.

Let $\tilde{P}_x$ be the measure on Ω induced by P_x on $\Omega(D)$. We have already remarked that P_x is defined on the (topological Borel σ field of $\Omega(D)$. As P_x is a probability measure, $\tilde{P}_x$ is also a probability measure. It should now be verified that $\tilde{P}_x$ is really the probability measure consistent with the given distribution. **27**

Theorem . *$\tilde{P}_x$ is a required probability measure for a continuous realization of the Brownian motion.*

Proof. We have to show that

$$F_{t_1\ldots t_k} = \tilde{P}_x\pi^{-1}_{t_1\ldots t_k} \quad \text{for all} \quad t_1,t_2\ldots t_k \quad \text{in} \quad [0,\infty).$$

Step 1. Let $t_1,\ldots,t_k\in D$. Then

$$P_x(\pi^{-1}_{t_1\ldots t_k}(A_1\times\cdots\times A_k)) = P_x(\tau\pi_{t_1\ldots t_k}(A_1\times\cdots\times A_k))$$

for every A_i Borel in $\mathbb{R}^d$. The right side above is

$$P_x(\pi^{-1}_{t_1\ldots t_k}(A_1\times\cdots\times A_k)) = F_{t_1\ldots t_k}(A_1\times\cdots\times A_k)$$

(by definition of P_x).

Step 2. We know that $T_{t_1,t_2\ldots t_k} = \tilde{P}_x \pi_{t_1,t_2,\ldots,t_k}$ provided that $t_1, t_2, \ldots, t_k \in D$. Let us pick $t_1^{(n)}, \ldots, t_k^{(n)}$, such that each $t_i^{(n)} \in D$ and $t_k^{(n)} \to t_k$ as $n \to \infty$. For each n and for each fixed $f : \mathbb{R}^d \to \mathbb{R}$ which is bounded and continuous,

$$E^{F_{1}^{(n)},\ldots t_k^{(n)}}[f(x_1,\ldots,x_k)] = E^{\tilde{P}_x}[f(x(t_1^{(n)},\ldots,x(t_k^{(n)})))].$$

28 Letting $n \to \infty$ we see that

$$E^{F_{t_1,\ldots,t_k}}[f(x_1,\ldots,x_k)] = E^{P_x}[f(x(t_1),\ldots,x(t_k))]$$

for all $t_1, \ldots, t_k$. This completes the proof. □

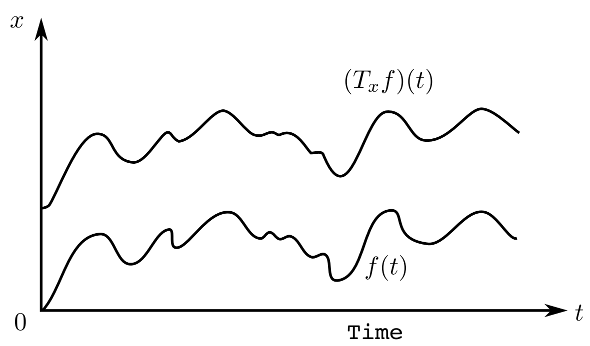
The definition of the Brownian motion given earlier has a built-in constraint that all "trajectories" start from $X(0) = x$. This result is given by

Theorem . $\tilde{P}_0\{f : f(0) = 0\} = 1$.

Proof. Obvious; because $E^{\tilde{P}_x}[\phi(x(0))] = \phi(x)$. □

Note. In future $\tilde{P}_x$ will be replaced by P_x and $\tilde{P}_0 = P_0$ will be denoted by P.

Let $T_x : \Omega \to \Omega$ be the map given by $(T_x f)(t) = x + f(t)$. T_x translates every 'trajectory' through the vector x.



Let us conceive a real Brownian motion of a system of particles. The operation T_x means that the system is translated in space (along with

everything else that affects it) through a vector x. The symmetry of the
29 physicl laws governing this motion tells us that any property exhibited by the first process should be exhibited by the second process and vice versa. Mathematically this is expressed by

Theorem . $P_x = PT_x^{-1}$.

Proof. It is enough to show that

$$P_x(T_x\pi^{-1}_{t_1\ldots t_k}(A_1\times\cdots\times A_k)) = P(\pi^{-1}_{t_1\ldots t_k}(A_1\times\cdots\times A_k))$$

for every A_i Borel in $\mathbb{R}^d$. Clearly,

$$T_x\pi^{-1}_{t_1\ldots t_k}(A_1\times\cdots\times A_k) = \pi^{-1}_{t_1\ldots t_k}(A_1 - x\times\cdots\times A_k - x).$$

Thus we have only to show that

$$\begin{aligned}\int_{A_1-x}\int\cdots\int_{A_k-x} & p(0,x,t_1,x_1)\ldots p(t_{k-1},x_{k-1},t_k,x_k)dx_1\ldots dx_k\\ = \int_{A_1}\cdots\int_{A_k} & p(0,0,t_1,x_1)\ldots p(t_{k-1},x_{k-1},t_k,x_k)dx_1\ldots dx_k,\end{aligned}$$

which is obvious. $\square$

Exercise. (a) If $\beta(t,\cdot)$ is a Brownian motion s tarting at $(0,0)$ then $\dfrac{1}{\sqrt{\epsilon}}\beta(\epsilon t)$ is a Brownian motion starting at $(0,0)$ for every $\epsilon > 0$.

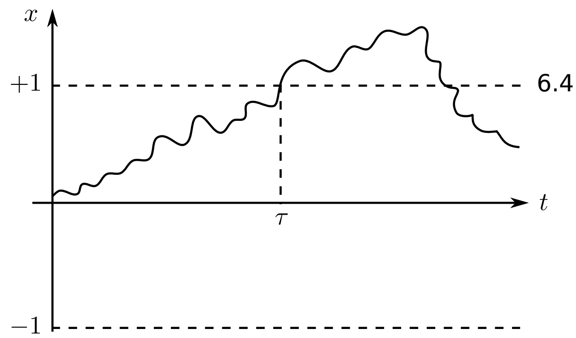
(b) If X is a d-dimensional Brownian motion and Y is a d'-dimensional Brownian motion then (X,Y) is a $d+d'$ dimensional Brownian motion provided that X and Y are independent.

(c) If $X_t = (X_t^1,\ldots,X_t^d)$ is a d-dimensional Brownian motion, then X_t^j is a one-dimensional Brownian motion. $(j = 1,2,\ldots d)$.

$$\begin{aligned}\tau(w) &= \inf\{t : |X_t(w)| \geq +1\}\\ &= \inf\{t : |w(t)| \geq 1\}\end{aligned}$$

$\tau(w)$ is the first time the particle hits either of the horizontal lines 30

+1 or −1.



2. Let $\{X_t\}$ be a d-dimensional Brownian motion, G any closed set in $\mathbb{R}^d$. Define
$$\tau(w) = \inf\{t : w(t) \in G\}.$$
This is a generalization of Example 1. To see that τ is a stopping time use
$$\{\tau \leq s\} = \bigcap_{n=1}^{\infty} \varlimsup_{\substack{\theta\in[0,s] \\ \theta \text{ rational}}} \{w : w(\theta) \in G_n\},$$
where
$$G_n = \left\{x \in \mathbb{R}^d : d(x, G) \leq \frac{1}{n}\right\}.$$

3. Let (X_t) be a d-dimensional Brownian motion, C and D disjoint closed sets in $\mathbb{R}^d$. Define
$$\tau(w) = \inf\{t; w(t) \in C \text{ and for some } s \leq t, w(s) \in D\}.$$
$\tau(w)$ is the first time that w hits C after visiting D.

5. Generalised Brownian Motion

LET Ω BE ANY space, $\mathscr{F}$ a σ-field and $(\mathscr{F}_t)$ an increasing family of sub σ-fields such that $\sigma(\cup\mathscr{F}_t) = \mathscr{F}$. Let P be a measure on $(\Omega, \mathscr{F})$. 31

$$X(t,w) : [0,w) \times \Omega \to \mathbb{R}^d$$

is called a *Brownian motion relative to* $(\Omega, \mathscr{F}_t, P)$ if

(i) $X(t,w)$ is progressively measurable with respect to $\mathscr{F}_t$;

(ii) $X(t,w)$ is a.e. continuous in t;

(iii) $X(t,w) - X(s,w)$ for $t > s$ is independent of $\mathscr{F}_s$ and is distributed normally with mean 0 and variance $t - s$, i.e.

$$P(X(t,\cdot) - X(s,\cdot) \in A|\mathscr{F}_s) = \int_A \frac{1}{[2\pi(t-s)]^{d/s}} \exp -\frac{|y|^2}{2(t-s)} dy.$$

Note. 1. The Brownian motion constructed previously was concentrated on $\Omega = C([0,\infty);\mathbb{R}^d)$, $\mathscr{F}$ was the Borel field of Ω, $X(t,w) = w(t)$ and $\mathscr{F}_t = \sigma\{X(s) : 0 \leq s \leq t\}$. The measure P so obtained is often called the *Wiener measure*.

2. The above definition is more general because

$$\sigma\{X(s) : 0 \leq s \leq t\} \subset \mathscr{F}_t.$$

Exercise. (Brownian motion starting at time s). Let $\Omega = C([s, \infty); \mathbb{R}^d)$, $\mathscr{B}$ = Borel field of Ω. Show that for each $x \in \mathbb{R}^d$ $\exists$ a probability measure P^s_x on Ω such that

32 (i) $P^s_x\{w : w(s) = x\} = 1$;

(ii) $P^s_x(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k)$

$$= \int_{A_1}\int_{A_2}\ldots\int_{A_k} p(s, x, t_1, x_1)p(t_1, x_1, t_2, x_2)\ldots$$
$$\ldots p(t_{k-1}x_{k-1}, t_k, x_k)dx_1 \ldots dx_k,$$
$$\forall\ s < t_1 < \ldots < t_k.$$

For reasons which will become clear later, we would like to shift the measure P^s_x to a measure on $C([0, \infty); \mathbb{R}^d)$. To do this we define

$$T : C([s, \infty); \mathbb{R}^d) \to C([0, \infty); \mathbb{R}^d)$$

by

$$(Tw)(t) = \begin{cases} w(t), & \text{if } t \geq s, \\ w(s), & \text{if } t \leq s. \end{cases}$$

Clearly, T is continuous. Put

$$P_{s,x} = P^s_x T^{-1}.$$

Then

(i) $P_{s,x}$ is a probability measure on the Borel field of $C([0, \infty); \mathbb{R}^d)$;

(ii) $P_{s,x}\{w : w(s) = x\} = 1$.

6. Markov Properties of Brownian Motion

Notation. 1. A random variable of a stochastic process $\{X(t)\}_{t \in I}$ shall be denoted by X_t or $X(t)$. $0 \leq t < \infty$. **33**

2. $\mathscr{F}_s$ will denote the σ-algebra generated by $\{X_t : 0 \leq t \leq s\}$; $\mathscr{F}_{s+} = \{\mathscr{F}_a : a > s\}$; $\mathscr{F}_{s-}$ will be the σ-algebra generated by $\cup\{\mathscr{F}_a : a < s\} s > 0$. It is clear that $\{\mathscr{F}_t\}$ is an increasing family.

3. For the Brownian motion, $\mathscr{B}$ = the σ-algebra generated by $\cup\{\mathscr{F}_t : t < \infty\}$ will be denoted by $\mathscr{F}$.

Theorem . *Let $\{X_t : 0 \leq t < \infty\}$ be a Brownian motion. Then $X_t - X_s$ is independent of $\mathscr{F}_s$.*

Proof. Let

$$0 \leq t_1 < t_2 < t_3 < \ldots < t_k \leq s.$$

Then the σ-algebra generated by $X_{t_1}, \ldots, X_{t_k}$ is the same as the σ-algebra generated by

$$X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_k} - X_{t_{k-1}}.$$

Since $X_t - X_s$ is independent of these increments, it is independent of $\sigma\{X_{t_1}, \ldots, X_{t_k}\}$. This is true for every finite set $t_1, \ldots, t_k$ and therefore $X_t - X_s$ is independent of $\mathscr{F}_s$.

Let us carry out the following calculation very *formally*.

$$P[X_t \in A \mid \mathscr{F}_s](w) = P[X_t - X_s \in B \mid \mathscr{F}_s](w), \ B = A - X_s(w),$$

$$= P[X_t - X_s \in B], \quad \text{by independence,}$$

34 i.e.

$$P[X_t \in A \mid \mathscr{F}_s](w) = \int_A \frac{1}{(2\pi t)^{d/2}} \exp - \frac{|y - X_s(w)|^2}{2(t-s)} dy.$$

□

This formal calculation leads us to

Theorem .

$$P[X_t \in A \mid \mathscr{F}_s](w) = \int_A \frac{1}{(2\pi t)^{d/2}} \exp - \frac{|y - X_s(w)|^2}{2(t-s)} dy.$$

where A is Borel in $\mathbb{R}^d$, $t > s$.

Remark. It may be useful to note that $p(s, X_s(w), t, y)$ can be thought of as a conditional probability density.

Proof. (i) We show that

$$f_A(w) = \int_A \frac{1}{(2\pi t)^{d/2}} \exp - \frac{|y - X_s(w)|^2}{2(t-s)} dy$$

is $\mathscr{F}_S$-measurable. Assume first that A is bounded and Borel in $\mathbb{R}^d$. If $\omega_n \to \omega$, then $f_A(\omega_n) \to f_A(\omega)$, i.e. f_A is continuous and hence $\mathscr{F}_s$-measurable. The general case follows if we note that any Borel set can be got as an increasing union of a countable number of bounded Borel sets.

(ii) For any $C \in \mathscr{F}_s$ we show that

$$(*) \qquad \int_C^{X} X_t^{-1}(A) dP(\omega) = \int_C \int_A \frac{\exp - |y - X_s(\omega)|^2}{(2\pi(t-s))^{d/2}} dy\, dP(\omega).$$

It is enough to verify $(*)$ for C of the form

$$C = \{\omega : (X_{t_1}(\omega), \ldots, X_{t_k}(\omega)) \in A_1 \times \cdots \times A_k;\ 0 \leq t_1 < \ldots < t_k \leq s\},$$

where A_i is Borel in $\mathbb{R}^d$ for $i = 1, 2 \ldots k$. The left side of $(*)$ is then

$$\int\limits_{A_i\times\cdots\times A_k\times A} p(0,0,t_1,x_{t_1})p(t_1,x_{t_1},t_2,x_{t_2})\ldots p(t_k,x_{t_k},t,x_t)dx_{t_1}\ldots dx_t.$$

35

To compute the right side define

$$f:\mathbb{R}^{(k+1)d}\to\mathbb{B}$$

by

$$f(u_1,\ldots,u_k,u) = X_{A_1}(u_1)\ldots X_{A_k}(u_k)p(s,u,t,y).$$

Clearly f is Borel measurable. An application of Fubini's theorem to the right side of $(*)$ yields

$$\begin{aligned}
&\int\limits_A dy \int\limits_\Omega X_{A_1}(X_{t_1}(\omega))\ldots X_{A_k}(X_{t_k}(\omega))p(s,X_s(\omega),t,y)dP(\omega)\\
&= \int\limits_A dy \int\limits_{\substack{\mathbb{R}^d\times\cdots\times\mathbb{R}^d\\ (k+1)\quad \text{times}}} f(x_1\ldots x_k,x_s)dF_{t_1\ldots t_k},s\\
&= \int\limits_A dy \int\limits_{A_1\times\cdots\times A_k\times\mathbb{R}^d} p(0,0,t_1,x_1)\ldots p(t_{k-1},t_k,x_k)\\
&\qquad\qquad p(t_k,x_k,s,x_s)p(s,x_s,t,y)dx_1\ldots dx_k dx_s\\
&= \int\limits_{A_1\times\cdots\times A_k\times A} p(0,0,t_1,x_1)\ldots p(t_{k-1},x_{k-1},t_k,x_k)p(t_k,x_k,t,y)\\
&\qquad\qquad dx_1\ldots,dx_k dy\\
&\qquad\qquad \text{(by the convolution rule)}\\
&= \text{left side.}
\end{aligned}$$

□

Examples of Stopping Times.

1. Let (X_t) be a one-dimensional Brownian motion. Define τ by

$$\{\tau \leq s\} = \bigcap_{n=1}^{\infty} \overline{\lim_{\substack{\theta_1,\theta_2 \\ \theta_1,\theta_2 \text{ rational in } [0,s]}}} \{w : w(\theta_1) \in D_n, w(\theta_2) \in C_n\},$$

36 where

$$D_n = \left\{x \in \mathbb{R}^d : d(x, D) \leq \frac{1}{n}\right\}, C_n = \left\{x \in \mathbb{R}^d : d(x, C) \leq \frac{1}{n}\right\}$$

Exercise 1. Let τ be as in Example 1.

(a) If $A = \{w : X_1(w) \leq \tau\}$ show that $A \notin \mathscr{F}_\tau$.
(Hint: $A \cap \{\tau \leq 0\} \notin \mathscr{F}_0$). This shows that $\mathscr{F}_\tau \subsetneqq \mathscr{F}_0$.

(b) $P_0\{w : \tau(w) = \infty\} = 0$.
(Hint: $P_0\{w : |w(t)| < 1\} \leq \int\limits_{|y| \leq t^{-1/2}} e^{-1/2|y|^2} dy \ \forall t$).

Theorem . *(Strong Markov Property of Brownian Motion). Let τ be any finite stopping time, i.e. $\tau < \infty$ a.e. Let $Y_t = X_{\tau+t} - X_\tau$. Then*

1. *$P[(Y_{t_1} \in A_1, \ldots, Y_{t_k} \in A_k) \cap A] = P(X_{t_1} \in A_1, \ldots X_{t_k} \in A_k) \cdot P(A)$, $\forall\, A \in \mathscr{F}_\tau$ and for every A_i Borel in $\mathbb{R}^d$. Consequently,*

2. *(Y_t) is a Brownian motion.*

3. *(Y_t) is independent of $\mathscr{F}_\tau$.*

The assertion is that a Brownian motion starts afresh at every stopping time.

Proof.

Step 1. Let τ take only countably many values, say s_1, s_2, $s_3 \ldots$. Put $E_j = \tau^{-1}\{s_j\}$. Then each E_j is $\mathscr{F}_\tau$-measure and

$$\Omega = \bigcup_{j=1}^{\infty} E_j,\ E_j \cap E_i = \emptyset\ j \neq i.$$

37 Fix $A \in \mathscr{F}_\tau$.

$$
\begin{aligned}
P[(Y_{t_1} &\in A_1, \ldots, Y_{t_k} \in A_k) \cap A] \\
&= \sum_{j=1}^{\infty} P[(Y_{t_1} \in A_1, \ldots, Y_{t_k} \in A_k) \cap A \cap E_j] \\
&= \sum_{j=1}^{\infty} P[(X_{t_1+s_j} - X_{s_j}) \in A_1, \ldots, X_{t_k+s_j} - X_{s_j} \in A_k) \cap A \cap E_j] \\
&= \sum_{j=1}^{\infty} P[(X_{t_1} \in A_1), \ldots, (X_{t_k} \in A_k)] P(A \cap E_j) \\
&\qquad \text{(by the Markov property)} \\
&= P(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k) \cdot P(A)
\end{aligned}
$$

Step 2. Let τ be any stopping time; put $\tau_n = \dfrac{[n\tau]+1}{n}$. A simple calculation shows that τ_n is a stopping time taking only countably many values. As $\tau_n \downarrow \tau$, $\mathscr{F}_\tau \subset \mathscr{F}_{\tau_n} \forall n$. Let $Y_t^{(n)} = X_{\tau_n+t} - X_{\tau_n}$.

By Step 1,

$$
\begin{aligned}
&P[(Y_{t_1}^{(n)} < x_1, \ldots, Y_{t_k}^{(n)} < x_k) \cap A] \\
&= P(X_{t_1} < x_1, \ldots, X_{t_k} < x_k) \cdot P(A)
\end{aligned}
$$

(where $x < y$ means $x_i < y_i$ $i = 1, 2, \ldots, d$) for every $A \in \mathscr{F}_\tau$. As all the Brownian paths are continuous, $Y_t^{(n)} \to Y_t$ a.e. Thus, if $x_1, \ldots, x_k$ is a point of continuity of the joint distribution of $X_{t_1}, \ldots, X_{t_k}$, we have

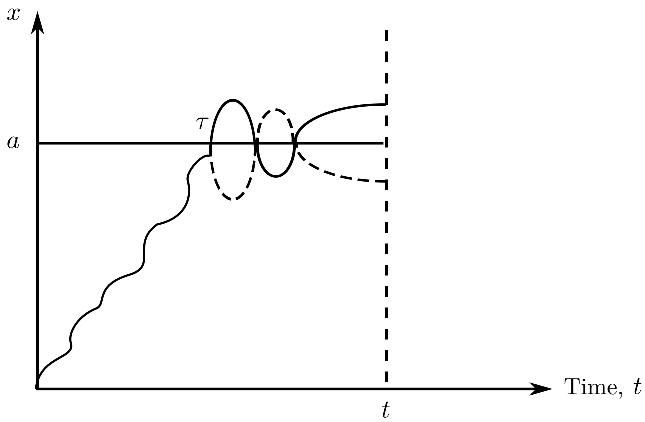
$$
P[(Y_{t_1} < x_1, \ldots, Y_{t_k} < x_k) \cap A] = P(X_{t_1} < x_1, \ldots, X_{t_k} < x_k) P(A)
$$

$\forall A \in \mathscr{F}_\tau$. Now assertion (1) follows easily. 38

For (2), put $A = \Omega$ in (1), and (3) is a consequence of (1) and (2). □

7. Reflection Principle

LET (X_t) BE A one-dimensional Brownian motion. Then $P(\sup\limits_{0\leq s\leq t} X_s \geq a) = 2P(X_t \geq a)$ with $a > 0$. 39
This gives us the probability of a Brownian particle hitting the line $x = a$ some time less than or equal to t. The intuitive idea of the proof is as follows.



Among all the paths that hit a before time t exactly half of them end up below a at time t. This is due to the reflection symmetry. If $X_s = a$ for some $s < t$, reflection about the horizontal line at a gives a one - one correspondence between paths with $X_t > a$ and paths with $X_t < a$. Therefore

$$P\left\{\max_{0\leq s\leq t} X_s \geq a, X_t > a\right\} = \left\{\max_{0\leq s\leq t} X_s \geq a, X_t < a\right\}$$

Since $P\{X_t = a\} = 0$, we obtain

$$\begin{aligned} P\left\{\sup_{0\le s\le t} X_s \ge a\right\} &= P\left\{\sup_{0\le s\le t} X_s \ge a, X_t > a\right\} + P\left\{\sup_{0\le s\le t} X_s \ge a, X_t > a\right\} \\ &= 2P\{X_t \ge a\} \end{aligned}$$

40 We shall now give a precise argument. We need a few elementary results.

Lemma 1. *Let $X_n = \sum\limits_{k=1}^{n} Y_k$ where the Y_k are independent random variables such that $P\{Y_k \in B\} = P\{-Y_k \in B\} \forall$ Borel set $B \subset R$ (i.e. Y_k are symmetric). Then for any real number a,*

$$P\left\{\max_{1\le i\le n} X_i > a\right\} \le 2P\{X_n > a\}$$

Proof. It is easy to verify that a random variable is symmetric if and only if its characteristic function is real. Define

$$\begin{gathered} A_i = \{X_1 \le a, \ldots X_{i-1} \le a, X_i > a\}, i = 1, 2, \ldots, n; \\ B = \{X_n > a\} \end{gathered}$$

Then $A_i \cap A_j = \emptyset$ if $i \ne j$. Now,

$$\begin{aligned} P(A_i \cap B) &\ge P(A_i \cap \{X_n \ge X_i\}) \\ &= P(A_i)P(X_n \ge X_i), \quad \text{by independence.} \\ &= P(A_i)P(Y_{i+1} + \cdots + Y_n \ge 0). \end{aligned}$$

As $Y_{i+1}, \ldots, Y_n$ are independent, the characteristic function of $Y_{i+1} + \cdots + Y_n$ is the product of the characteristic functions of $Y_{i+1} + \cdots + Y_n$, so that $Y_{i+1} + \cdots + Y_n$ is symmetric. Therefore

$$P(Y_{i+1} + \cdots + Y_n \ge 0) \ge \frac{1}{2}.$$

Thus $P(A_i \cap B) \ge \dfrac{1}{2} P(A_i)$ and

$$P(B) \ge \sum_{i=1}^{n} P(A_i \cap B) \ge \frac{1}{2} \sum P(A_i) \ge \frac{1}{2} P\left(\bigcup_{i=1}^{n} A_i\right),$$

41 i.e.

$$2P(B) \geq P\left(\bigcup_{i=1}^{n} A_i\right),$$

or

$$P\left\{\max_{1\leq i\leq n} X_i > a\right\} \leq 2P\{X_n > a\}$$

$\square$

Lemma 2. *Let $Y_i, \ldots, Y_n$ be independent random variables. Put $X_n = \sum_{k=1}^{n} Y_k$ and let $\tau = \min\{i : X_i > a\}$, $a > 0$ and $\tau = \infty$ if there is no such i. Then for each $\epsilon > 0$,*

(a) $P\{\tau \leq n-1, X_n - X_\tau \leq -\epsilon\} \leq P\{\tau \leq n-1, X_n \leq a\} + \sum_{j=1}^{n-1} P(Y_j > \epsilon)$.

(b) $P\{\tau \leq n-1, X_n > a+2\epsilon\} \leq P\{\tau \leq n-1, X_n - X_\tau > \epsilon\} + \sum_{j=1}^{n-1} P\{Y_j > \epsilon\}$

(c) $P\{X_n > a + 2\epsilon\} \leq P\{\tau \leq n-1, X_n > a + 2\epsilon\} + P\{Y_n > 2\epsilon\}$.

If, further, $Y_1, \ldots, Y_n$ are symmetric, then

(d) $P\{\max_{1\leq i\leq n} X_i > a, X_n \leq a\} \geq P\{X_n > a + 2\epsilon\} - P\{Y_n \geq 2\epsilon\} - 2\sum_{j=1}^{n-1} P\{Y_j > \epsilon\}$

(e) $P\{\max_{1\leq i\leq n} X_i > a\} \geq 2P\{X_n > a + 2\epsilon\} - 2\sum_{j=1}^{n} P\{Y_j > \epsilon\}$

Proof. (a) Suppose $w \in \{\tau \leq n-1, X_n - X_\tau \leq -\epsilon\}$ and $w \in \{\tau \leq n-1, X_n \leq a\}$. Then $X_n(w) > a$ and $X_n(w) + \epsilon \leq X_{\tau(w)}(w)$ or, $X_{\tau(w)}(w) > a + \epsilon$.

By definition of $\tau(w)$, $X_{\tau(w)-1}(w) \leq a$ and therefore,

$$Y_{\tau(w)}(w) = X_{\tau(w)}(w) - X_{\tau(w)-1}(w) > a + \epsilon - a = \epsilon$$

if $\tau(w) > 1$; if $\tau(w) = 1$, $Y_{\tau(w)}(w) = X_{\tau(w)}(w) > a + \epsilon > \epsilon$.

Thus $Y_j(w) > \epsilon$ for some $j \leq n-1$, i.e. 42

$$w \in \bigcup_{j=1}^{n-1} \{Y_j > \epsilon\}.$$

Therefore

$$\{\tau \le n-1, X_n - X_\tau \le -\epsilon\} \subset \{\tau \le n-1, X_n \le a\} \bigcup_{j=1}^{n-1} \{Y_j > \epsilon\}$$

and (q) follows.

(b) Suppose $w \in \{\tau \le n-1, X_n > a+2\epsilon\}$ but $w \in \{\tau \le n-1, X_n - X_\tau > \epsilon\}$. Then

$$X_n(w) - X_{\tau(w)}(w) \le \epsilon,$$

or, $X_{\tau(w)}(w) > a+\epsilon$ so that $Y_{\tau(w)}(w) > \epsilon$ as in (a); hence $Y_j(w) > \epsilon$ for some $j \le n-1$. This proves (b).

(c) If $w \in \{X_n > a+2\epsilon\}$, then $\tau(w) \le n$; if $w \notin \{\tau \le n-1, X_n > a+2\epsilon\}$, then $\tau(w) = n$ so that $X_{n-1}(w) \le a$; therefore $Y_n(w) = X_n(w) - X_{n-1}(w) > 2\epsilon$. i.e. $w \in \{Y_n > 2\epsilon\}$. Therefore

$$\{X_n > a+2\epsilon\} \subset \{\tau \le n-1, X_n > a+2\epsilon\} \cup \{Y_n > 2\epsilon\}.$$

This establishes (c).

(d) $P\{\max\limits_{1\le i\le n} X_i > a, X_n \le a\} = P\{\tau \le n-1, X_n \le a\}$

$$\ge P\{\tau \le n-1, X_n - X_\tau \le -\epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon), \quad \text{by (a)},$$

$$P\left[\bigcup_{k=1}^{n-1} \{\tau = k, X_n - X_k \le -\epsilon\}\right] - \sum_{j=1}^{n-1} P(Y_j > \epsilon)$$

$$= \sum_{k=1}^{n-1} P\{\tau = k, X_n - X_k \le -\epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon)$$

$$= \sum_{k=1}^{n-1} P\{\tau = k\} P\{X_n - X_k \le -\epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon)$$

(by independence)

$$= \sum_{k=1}^{n} P\{\tau = k\}P\{X_n - X_k > \epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon) \quad \text{(by symmetry)}$$

$$= P\{\tau \le n-1, X_n - X_\tau \ge \epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon)$$

$$\ge P\{\tau \le n-1, X_n - X_\tau > \epsilon\} - \sum_{j=1}^{n-1} P(Y_j > \epsilon)$$

$$\ge P\{\tau \le n-1, X_n > a + 2\epsilon\} - 2\sum_{j=1}^{n-1} P\{Y_j > \epsilon\} \quad \text{(by (b))}$$

$$\ge P\{X_n > a + 2\epsilon\} - P\{Y_n > 2\epsilon\} - 2\sum_{j=1}^{n-1} P\{Y_j > \epsilon\} \quad \text{(by (c))}$$ **43**

This proves (d).

(e) $$\begin{aligned} P\{\max_{1\le i\le n} X_i > a\} &= P\{\max_{1\le i\le n} X_i > a, X_n \le a\} + P\{\max_{1\le i\le n} X_i > a, X_n > a\} \\ &= P\{\max_{1\le i\le n} X_i > a, X_n \le a\} + P\{X_n > a\} \\ &= P\{X_n > a + 2\epsilon\} - P\{Y_n > 2\epsilon\} + P\{X_n > a\} \\ &\qquad - 2\sum_{j=1}^{n-1} P\{Y_j > \epsilon\} \quad \text{(by (d))} \end{aligned}$$

Since $P\{X_n > a + 2\epsilon\} \le P\{X_n > a\}$ and

$$P\{Y_n > 2\epsilon\} \le P\{Y_n > \epsilon\} \le 2P\{Y_n > \epsilon\},$$

we get

$$P\{\max_{1\le i\le n} X_i > a\} \ge 2P\{X_n > a + 2\epsilon\} - 2\sum_{j=1}^{n} P(Y_j > \epsilon)$$

This completes the proof. **44**

□

Proof of the reflection principle.

By Lemma 1

$$p = P\left\{\max_{1\le j\le n} X\left(\frac{jt}{n}\right) >\right\} \le 2P(X(t) > a).$$

By Lemma 2(e),

$$p \geq 2P(X(t) > a + 2\epsilon) - 2\sum_{j=1}^{n} P\left\{\left(X\left(\frac{jt}{n}\right) - X\left(\frac{(j-1)t}{n}\right)\right) > \epsilon\right\}.$$

Since $X\left(\frac{jt}{n}\right) - X\left(\frac{(j-1)t}{n}\right)$ are independent and identically distributed normal random variables with mean zero and variance $\frac{t}{n}$ (in particular they are symmetric),

$$\begin{aligned} P\left(\left(X\left(\frac{jt}{n}\right) - X\left(\frac{(j-1)t}{n}\right)\right) > \epsilon\right) &= P\left(\left(X\left(\frac{t}{n}\right) - X(0)\right) > \epsilon\right) \\ &= P\left(X\left(\frac{t}{n}\right) > \epsilon\right). \end{aligned}$$

Therefore

$$p \geq 2P(X(t) > a + 2\epsilon) - 2n\, P\left(X\left(\frac{t}{n}\right) > \epsilon\right).$$

$$\begin{aligned} P(X(t/n) > \epsilon) &= \int_{\epsilon}^{\infty} \frac{1}{\sqrt{(2t/n)}} e^{-x^2/\frac{2t}{n}} dx \\ &= \int_{\epsilon\sqrt{n}/\sqrt{(2t)}}^{\infty} \frac{e^{-x^2}}{\sqrt{x}} dx \leq \frac{\epsilon\sqrt{n}/\sqrt{(2t)}}{\sqrt{\pi}} \int_{\epsilon\sqrt{n}/\sqrt{(2t)}}^{-1\infty} xe^{-x^2} dx \end{aligned}$$

or

$$P(X(t/n) > \epsilon) \leq \frac{1}{2\sqrt{\pi}} e^{-\epsilon^2 n/2t} \cdot \frac{\sqrt{(2t)}}{\epsilon\sqrt{n}}.$$

Therefore

$$nP(X(t/n) > \epsilon) \leq \frac{n}{2\epsilon}\sqrt{(2t)}/\sqrt{(\pi n)}e^{-\epsilon^2 n/2t} \to 0 \quad \text{as} \quad n \to +\infty.$$

45

By continuity,

$$P\left\{\max_{1\leq\leq n} X(jt/n) > a\right\} \to P\left\{\max_{0\leq s\leq t} X(s) > a\right\}.$$

We let n tend to ∞ through values $2, 2^2, 2^3, \ldots$ so that we get

$$\begin{aligned} &2P\{X(t) > a + 2\epsilon\} - 2n\, P\{X(t/n) > \epsilon\} \\ \leq &P\left\{\max_{1\leq j\leq n} X(t/n) > a\right\} \leq 2P\{X(t) > a\}, \end{aligned}$$

or

$$\begin{aligned} 2P\{X(t) > a\} \leq 2P\{X(t) \geq a\} \leq &P\left\{\max_{0\leq s\leq t} X(t) > a\right\} \\ &2P\{X(t) > a\}, \end{aligned}$$

on letting $n \to +\infty$ first and then letting $\epsilon \to 0$. Therefore,

$$\begin{aligned} P\left\{\max_{0\leq s\leq t} X(s) > a\right\} &= 2P\{X(t) > a\} \\ &= 2\int_a^\infty 1/\surd(2\pi t) e^{-x^2/2t} dx. \end{aligned}$$

AN APPLICATION. Consider a one-dimensional Brownian motion. A particle starts at 0. What can we say about the behaviour of the particle in a small interval of time $[0, \epsilon)$? The answer is given by the following result.

$$\begin{aligned} P(A) \equiv P\{w : \forall\epsilon > 0, \exists\, t, s \text{ in } [0, \epsilon) \text{ such that } X_t(w) > 0 \text{ and} \\ X_s(w) < 0\} = 1. \end{aligned}$$

INTERPRETATION. Near zero all the particles oscillate about their 46
starting point. Let

$$\begin{aligned} A^+ &= \{w : \forall \in > 0 \exists\, t \in [0, \epsilon) \text{ such that } X_t(w) > 0\}, \\ A^- &= \{w : \forall\epsilon > 0 \exists\, s \in [0, \epsilon) \text{ such that } X_s(w) < 0\}. \end{aligned}$$

We show that $P(A^+) = P(A^-) = 1$ and therefore $P(A) = P(A^+ \cap A^-) = 1$.

$$A^+ \supset \bigcap_{n=1}^{\infty} \left\{\sup_{0\leq t\leq 1/n} w > 0\right\} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \left(\sup_{0\leq t\leq 1/n} w(t) \geq 1/m\right)$$

Therefore

$$\begin{aligned}P(A^+) &\geq \operatorname*{Lt}_{n\to\infty} \sup_{m\to\infty} P\left(\sup_{0\leq t\leq 1/n} w(t) \geq 1/m\right)\\ &\geq 2 \operatorname*{Lt}_{n\to\infty m\to\infty} \sup P(w(1/n) \geq 1/m) \quad \text{(by the reflection principle)}\\ &\geq 1.\end{aligned}$$

Similarly $P(A^-) = 1$.

Theorem . *Let $\{X_t\}$ be a one-dimensional Brownian motion, $A \subset (-\infty, a)$ $(a > 0)$ and Borel subset of $\mathbb{R}$. Then*

$$\begin{aligned}&P_0\{X_t \in A, X_s < a\ \forall s \text{ such that } 0 \leq s \leq t\}\\ &= \int\limits_A 1/\sqrt{(2\pi t)}e^{-y^2/2t}dy - \int\limits_A 1/\sqrt{(2\pi t)}e^{-(2a-y)^2/2t}dy\end{aligned}$$

Proof. Let $\tau(w) = \inf\{t : w(t) \geq a\}$. By the strong Markov property of Brownian motion,

$$P_0\{B(X(\tau + s) - X(\tau) \in A)\} = P_0(B)P_0(X(s) \in A)$$

for every set B in $\mathscr{F}_t$. This can be written as

$$E(X_{(X(\tau+s)-X(\tau)\in A)}|\mathscr{F}_\tau) = P_0(X(s) \in A)$$

47 Therefore

$$E(X_{(X(\tau+\ell(w))-X(\tau)\in A)}|\mathfrak{F}_\tau) = P_0(X(\ell(w)) \in A)$$

for every function $\ell(w)$ which is $\mathscr{F}_\tau$-measurable. Therefore,

$$P_0((\tau \leq t) \cap ((X(\tau + \ell(w)) - X(\tau)) \in A) = \int\limits_{\{\tau\leq t\}} P_0(X(\ell(w)) \in A)dP(w)$$

In particular, take $\ell(w) = t - \tau(w)$, clearly $\ell(w)$ is $\mathscr{F}_\tau$-measurable. Therefore,

$$P_0((\tau \leq t)((X(t) - X(\tau)) \in A)) = \int\limits_{\{\tau\leq t\}} P_0(X(\ell(w) \in A)dP(w)).$$

Now $X(\tau(w)) = a$. Replace A by $A - a$ to get

$$(*) \qquad P_0((\tau \leq t) \cap (X(t) \in A)) = \int\limits_{\{\tau \leq t\}} P_0(X(\ell(w) \in A - a)dP(w))$$

Consider now

$$\begin{aligned} P_{2a}(X(t) \in A) &= P_0(X(t) \in A - 2a) \\ &= P_0(X(t) \in 2a - A) \quad \text{(by symmetry of } x) \\ &= P_0((\tau \leq t) \cap (X(t) \in 2a - A)). \end{aligned}$$

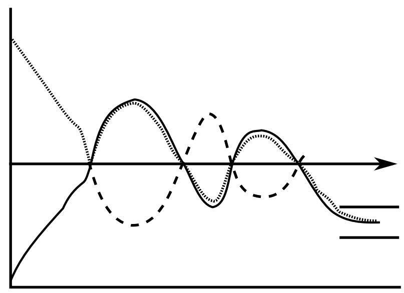
The last step follows from the face that $A \subset (-\infty, a)$ and the continuity of the Brownina paths. Therefore

$$\begin{aligned} P_{2a}(X(t) \in A) &= \int\limits_{\{\tau \leq t\}} P_0(X(\ell(w)) \in a - A)dP(w), \quad \text{(using } *) \\ &= P_0((\tau \leq t) \cap (X(t) \in A)). \end{aligned}$$

Now the required probability

$$\begin{aligned} P_0\{X_t \in A, X_s < a \,\forall s \in 0 \leq s \leq t\} &= P_0\{X_t \in A\} - P_0\{(\tau \leq t) \cap (X_t \in A)\} \\ &= \int\limits_A 1/\sqrt{(2\pi t)} e^{-y^2/2t} dy - \int\limits_A 1/\sqrt{(2\pi t)} e^{-(2a-y)^2/2t} dy. \end{aligned}$$

The intuitive idea of the previous theorem is quite clear. To obtain **48**
the paths that reach A at time t without hitting the horizontal line $x = a$, we consider all paths that reach A at time t and subtract those paths that hit the horizontal line $x = a$ before time t and then reach A at time t. To see exactly which paths reach A at time t after hitting $x = a$ we consider a typical path $X(w)$.



□

The reflection principle (or the strong Markov property) allows us to replace this path by the dotted path (see Fig.). The symmetry of the Brownian motion can then be used to reflect this path about the line $x = a$ and obtain the path shown in dark. Thus we have the following result:

the probability that a Brownian particle starts from $x = 0$ at $t = 0$ and reaches A at time t after it has hit $x = a$ at some time $\tau \leq t$ is the same as if the particle started at time $t = 0$ at $x = 2a$ and reached A at time t. (The continuity of the path ensures that at some time $\tau \leq t$, this particle has to hit $x = a$).

49 We shall use the intuitive approach in what follows, the mathematical analysis being clear, thorugh lengthy.

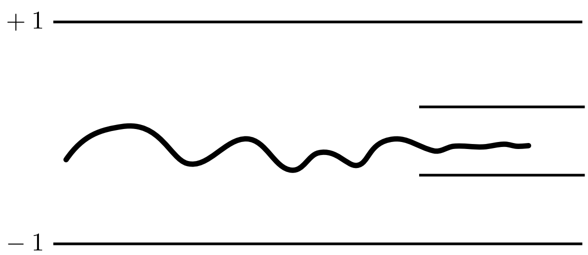
Theorem . *Let $X(t)$ be a one-dimensional Brownian motion, $A \subset (-1, 1)$ any Borel subset of $\mathbb{R}$. Then*

$$P_0\left[\sup_{0\leq s\leq t} |X(s)| < 1, X(t) \in A\right] = \int_A \phi(t, y)dy,$$

where

$$\phi(t, y) = \sum_{n=-\infty}^{\infty} (-1)^n/\sqrt{(2\pi t)}e^{-(y-2n)^2/2t}.$$

Proof.



Let E_n be the set of those trajections which (i) start at $x = 0$ at time $t = 0$ (ii) hit $x = +1$ at some time $\tau_1 < t$ (iii) hit $x = -1$ at some later time $\tau_2 < t$ (iv) hit $x = 1$ again at a later time $\tau_3 < t \ldots$ and finally reach A at time t. The number of τ's should be equal to n at least, i.e.

$$E_n = \{w : \text{ there exists a sequence } \tau_1, \ldots \tau_n \text{ of}$$

stopping times such that $0 < \tau_1 < \tau_2 < \ldots < t_{\tau_n} < t$, $X(\tau_j) = (-1)^{j-1}$, $X(t) \in A\}$. Similarly, let

$$\begin{aligned} F_n = \{w : \text{ there exists a sequence } \tau_1, \ldots, \tau_n \text{ of stopping times} \\ 0 < \tau_1 < \tau_2 < \ldots < \tau_n < t, X(\tau_j) = (-1)^j, X(t) \in A\} \end{aligned}$$

Note that **50**

$$\begin{aligned} &E_1 \supset E_2 \supset E_3 \supset \ldots, \\ &F_1 \supset F_2 \supset F_3 \supset \ldots, \\ &F_n \supset E_{n+1};\ E_n \supset F_{n+1}, \\ &E_n \cap F_n = E_{n+1} \cup F_{n+1}. \end{aligned}$$

Let

$$\phi(t, A) = P\left[\sup_{0 \leq s \leq t} |X(s)| < 1, X(t) \in A\right].$$

Therefore

$$\phi(t, A) = P[X(t) \in A] - P\left[\sup_{0 \leq s \leq t} |X(s)| \geq 1, X(t) \in A\right]$$

$$= \int_A 1/\sqrt{(2\pi t)} e^{-y^2/2t} dy - P[(E_1 \cup F_1) \cap A_0],$$

where

$$A_0 = \{X(t) \in A\} = \int_A 1/\sqrt{(2\pi t)} e^{-y^2/2t} dy - P[(E_1 \cap A_0) \cup (F_1 \cap A_0)].$$

Use the fact that $P[A \cup B] = P(A) + P(B) - P(A \cap B)$ to get

$$\phi(t, A) = \int_A 1/\sqrt{(2\pi t)} e^{-y^2/2t} dy - P[E_1 \cap A_0] - P[F_1 \cap A_0] + P[E_1 \cap F_1 \cap A_0],$$

as $E_1 \cap F_1 = E_2 \cup F_2$. Proceeding successively we finally get

$$\phi(t, A) = \int_A 1/\sqrt{(2\pi t)} e^{-y^2/2t} dy + \sum_{n=1}^{\infty} (-1)^n P[E_n \cap A_0] + \sum_{n=1}^{\infty} (-1)^n P[F_n \cap A_0]$$

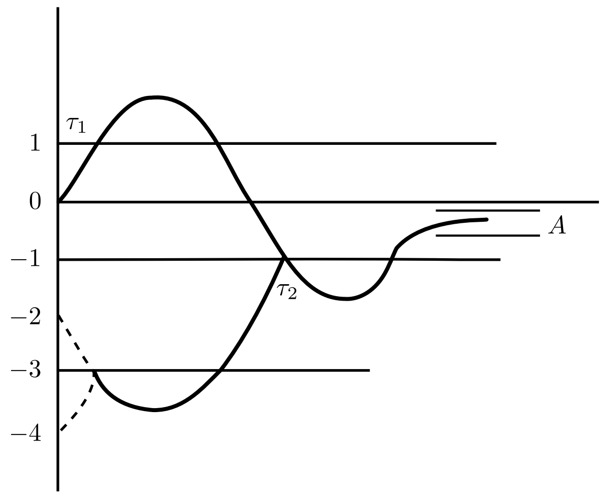
We shall obtain the expression for $P(E_1 \cap A_0)$ and $P[E_2 \cap A_0]$, the other terms can be obtained similarly.

$E_1 \cap A_0$ consists of those trajectries that hit $x = \pm 1$ at some time $\tau \leq t$ and then reach A at time t. Thus $P[E_1 \cap A_0]$ is given by the previous theorem by

$$\int_A 1/\sqrt{(2\pi t)} e^{-(y-2)^2/2t} dy.$$

51 $E_2 \cap A_0$ consists of those trajectories that hit $x = \pm 1$ at time τ_1, hit $x = -1$ at time τ_2 and reach A at time $t(\tau_1 < \tau_2 < t)$.

According to the previous theorem we can reflect the trajectory upto τ_2 about $x = -1$ so that $P(E_2 \cap A_0)$ is the same as if the particle starts at $x = -2$ at time $t = 0$, hits $x = -3$ at time τ_1 and ends up in A at time t. We can now reflect the trajectory



upto time τ_1 (the dotted curve should be reflected) about $x = -3$ to obtain the required probability as if the trajectory started at $x = -4$. Thus,

$$P(E_2 \cap A_0) = \int_A e^{-(y+4)^2/2t/\sqrt{(2\pi t)}} dy.$$

Thus

$$\phi(t, A) = \sum_{n=-\infty}^{\infty} (-1)^n \int_A 1/\sqrt{2\pi t} e^{-(y-2n)^2/2t} dy$$
$$= \int_A \phi(t, y) dy.$$

The previous theorem leads to an interesting result: **52**

$$P\left[\sup_{0 \le s \le t} |X(s)| < 1\right] = \int_{-1}^{1} \phi(t, y) dy$$

Therefore

$$P\left[\sup_{0 \le s \le t} |X(s)| \ge 1\right] = 1 - P\left[\sup_{0 \le s \le t} |X(s)| < 1\right]$$

$$= -1 - \int_{-1}^{1} \phi(t,y)dy,$$

$$\phi(t,y) = \sum_{n=-\infty}^{\infty} (-1)^n / \sqrt{(2\pi t)} e^{-(y-2n)^2/2t}$$

Case (i). t is very small.

In this case it is enough to consider the terms corresponding to $n = 0$, ± 1 (the higher order terms are very small). As y varies from -1 to 1,

$$\phi(t,y) \simeq 1/\sqrt{(2\pi t)} \left[e^{-y^2/2t} - e^{-(y-2)^2/2t} - e^{-(y+2)^2/2t}\right].$$

Therefore

$$\int_{-1}^{1} \phi(t,y)dy \simeq 4/\sqrt{(2\pi t)} e^{-1/2t}.$$

Case (ii). t is large. In this case we use Poisson's summation formula for $\phi(t,y)$:

$$\phi(t,y) = \sum_{k=0}^{\infty} e^{-(2k+1)^2\pi^2 t/8} Cos\{(k+1)/2\pi y\},$$

to get

$$\int_{-1}^{1} \phi(t,y)dy \simeq 4/\pi e^{-\pi^2 t/8}$$

53 for large t. Thus, $P(\tau > t) = 4/\pi e^{-\pi^2 t/8}$.

This result says that for large values of t the probability of paths which stay between -1 and $+1$ is very very small and the decay rate is governed by the factor $e^{-\pi^2 t/8}$. This is connected with the solution of a certain differential equation as shall be seen later on. □

8. Blumenthal's Zero-One Law

LET X_t BE A d-dimensional Brownian motion. If $A \in \mathscr{F}_{0+} = \bigcap\limits_{t>0} \mathscr{F}_t$, then $P(A) = 0$ or $P(A) = 1$. 54

Interpretation. If an event is observable in every interval $[0,t]$ of time then either it always happens or it never happens.

We shall need the following two lemmas.

Lemma 1. *Let $(\Omega, \mathscr{B}, P)$ be any probability space, $\mathscr{C}_a$ sub-algebra of $\mathscr{B}$. Then*

(a) *$L^2(\Omega, \mathscr{C}, P)$ is a closed subspace of $L^2(\Omega, \mathscr{B}, P)$.*

(b) *If $\pi : L^2(\Omega, \mathscr{B}, P) \to L^2(\Omega, \mathscr{C}, P)$ is the projection map then $\pi f = E(f|\mathscr{C})$.*

Proof. Refer appendix. □

Lemma 2. *Let $\Omega = C([0,\infty); \mathbb{R}^d)$, P_0 the probability corresponding to the Brownian motion. Then the set $\{\phi(\pi_{t_1}, \ldots, t_k) \in \phi$ is continuous, bounded on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$ (k times), $\pi_{t_1}, \ldots, t_k$ the canonical projection) is dense in $L^2(\Omega, \mathscr{B}, P)$.*

Proof. Functions of the form $\phi(x(t_1), \ldots, x(t_k)$ where ϕ runs over continuous functions is clearly dense in $L_2(\Omega, \mathscr{F}_{t_1, t_2, \ldots, t_k}, P)$ and

$$\bigcup_k \bigcup_{t_1, \ldots, t_k} L_2(\Omega, \mathscr{F}_{t_1, \ldots, t_k}, P)$$

is clearly dense in $L_2(\Omega, \mathscr{B}, P)$. $\square$

55 **Proof of zero-one law.** Let

$$H_t = L^2(\Omega, \mathscr{F}_t, P), H = L^2(\Omega, \mathscr{B}, P), H_{0+} = \bigcap_{t>0} H_t.$$

Clearly $H_{0+} = L^2(\Omega, \mathscr{F}_{0+}, P)$.

Let $\pi_t : H \to H_t$ be the projection. Then $\pi_t f \to \pi_{0+} f \,\forall_f$ in H. To prove the law it is enough to show that H_{0+} contains only constants, which is equivalent to $\pi_{0+} f =$ constant $\forall f$ in H. As π_{0+} is continuous and linear it is enough to show that $\pi_{0+}\phi =$ const $\forall\phi$ of the Lemma 2:

$$\begin{aligned} \pi_{0+}\phi = \operatorname*{Lt}_{t\to 0} \pi_t\phi &= \operatorname*{Lt}_{t\to 0} E(\phi|_t) \quad \text{by Lemma 1,} \\ &= \operatorname*{Lt}_{t\to 0} E(\phi(t_1, \ldots, t_k)|\mathscr{F}_t). \end{aligned}$$

We can assume without loss of generality that $t < t_1 < t_2 < \ldots < t_k$.

$$\begin{aligned} E(\phi(t_1, \ldots, t_k)|\mathscr{F}_t) = \int \phi(y_1, \ldots, y_k) 1/\sqrt{(2\pi(t_1 - t))} e^{-|y_1 - X_t(w)|^2/2(t_1 - t)} \ldots \\ \ldots 1/\sqrt{(2\pi(t_k - t_{k-1}))} e^{\frac{-|y_k - y_{k-1}|^2}{2(t_k - t_{k-1})}} dy_1 \ldots dy_k. \end{aligned}$$

Since $X_0(w) = 0$ we get, as $t \to 0$,

$$\pi_{0+}\phi = \text{ constant}.$$

This completes the proof.

APPLICATION. Let $\alpha \geq 1$ $A = \{w : \int\limits_0^1 |w(t)|/t^\alpha < \infty\}$. Then $A \in \mathscr{F}_{0+}$.

For, if $0 < s < 1$, then $\int\limits_s^1 |w(t)|/t^\alpha < \infty$. Therefore $w \in A$ or not according
56 as $\int\limits_0^s |w(t)|/t^\alpha dt$ converges or not. But this convergence can be asserted

by knowing the history of w upto time s. Hence $A \in \mathscr{F}_s$. Blumenthal's law implies that

$$\int\limits_0^1 |w(t)|/t^\alpha dt < \infty \text{ a.e.w., or, } \int\limits_0^1 |w(t)|/t^\alpha dt = \infty \text{ a.e.w.}$$

A precise argument can be given along the following lines. If $0 < s < 1$,

$$\begin{aligned} A &= \{w : \int\limits_0^s |w(t)|/t^\alpha < \infty\} \\ &= \{w : \sup I_{n,s}(w) < \infty\} \end{aligned}$$

where $I_{n,s}(w)$ is the lower Riemannian sum of $|w(t)^n|/t^\alpha$ corresponding to the partition $\{0, s/n, \ldots, s\}$ and each $I_{n,s} \in \mathscr{F}_s$.

9. Properties of Brownian Motion in One Dimension

WE NOW PROVE the following. **57**

Lemma . *Let* (X_t) *be a one-dimensional Brownian motion. Then*

(a) $P(\overline{\lim}\, X_t = \infty) = 1$*; consequently* $P(\overline{\lim}\, X_t < \infty) = 0$.

(b) $P(\underline{\lim}\, X_t = -\infty) = 1$*; consequently* $P(\underline{\lim}\, X_t > -\infty) = 0$.

(c) $P(\underline{\lim}\, X_t = -\infty);\ \overline{\lim}\, X_t = \infty) = 1$.

SIGNIFICANCE. By (c) almost every Brownian path assumes each value infinitely often.

Proof.

$$\begin{aligned}\{\overline{\lim}\, X_t = \infty\} &= \bigcap_{n=1}^{\infty}(\overline{\lim}\, X_t > n)\\ &= \bigcap_{n=1}^{\infty}(\overline{\lim_{\theta \text{ rational}}}\, X_\theta > n) \quad \text{(by continuity of Brownian paths)}\end{aligned}$$

First, note that

$$P_0\left[\sup_{0\le s\le t} X(s) \le n\right] = 1 - P_0\left[\sup_{0\le s\le t} X(s) > n\right]$$

$$= 1 - 21/\sqrt{(2\pi t)} \int\limits_{n}^{\infty} e^{-y^2/2t} dy$$

$$= \sqrt{(2/\pi t)} \int\limits_{0}^{n} e^{-y^2/2t} dy.$$

Therefore, for any x_0 and t,

$$P\left[\sup_{t_0 \le s \le t} X(s) \ge n | X(t_0) = x_0\right] = P_0\left[\sup_{0 \le s \le t - t_0} X(s) \ge n - x_0\right]$$

(independent increments) which tends to 1 as $t \to \infty$. Consequently,

$$P_0\left[\sup_{t \ge t_0} X(t) \ge n\right] = EP\left[\sup_{t \ge t_0} X(t) \ge n | X(t_0)\right]$$
$$= E1 = 1.$$

58

In other words,

$$P_0\left[\limsup_{t \to \infty} X(t) \ge n\right] = 1$$

for every n. Thus

$$P(\overline{\lim}\, X_t = \infty) = 1.$$

(b) is clear if one notes that $w \to -w$ leaves the probability invariant.

(c) $P(\overline{\lim}\, X_t = \infty, \underline{\lim}\, X_t = -\infty)$

$$= P(\overline{\lim}\, X_t = \infty) - P(\underline{\lim}\, X_t > -\infty, \overline{\lim}\, X_t = \infty).$$
$$\ge 1 - P(\underline{\lim}\, X_t > -\infty)$$
$$= 1.$$

□

Corollary . *Let* (X_t) *be a d-dimensional Brownian motion. Then*

$$P(\overline{\lim}\, |X_t| = \infty) = 1.$$

Remark. If $d \geq 3$ we shall see later that $P(\underset{t\to\infty}{\text{Lt}} |X_t| = \infty) = 1$. i.e. almost every Brownian path "wanders" off to ∞.

Theorem . *Almost all Brownian paths are of unbounded variation in any interval.*

Proof. Let I be any interval $[a, b]$ with $a < b$. For $n = 1, 2, \ldots$ define

$$V_n(wQ_n) = \sum_{i=1}^{n} |w(t_i) - w(t_{i-1})| \, (t_i = a + (b-a)i/n, i = 0, 1, 2, \ldots n),$$

The variation corresponding to the partioin Q_n dividing $[a, b]$ into n equal parts. Let 59

$$U_n(w, Q_n) = \sum_{i=1}^{n} |(w(t_i) - w(t_{i-1})|^2.$$

If

$$A_n(w, Q_n) \sup_{1\leq i\leq n} |w(t_i) - w(t_{i-1})|,$$

then

$$A_n(w, Q_n) V_n(w, Q_n) \geq U_n(w, Q_n).$$

By continuity $\underset{n\to\infty}{\text{Lt}} A_n(w, Q_n) = 0.$ □

Claim. $\underset{n\to\infty}{\text{Lt}} E[(U_n(w, Q_n) - (b-a))^2] = 0.$

Proof.

$$\begin{aligned}
&E[(U_n - (b-a))^2] \\
&= E\left\{\sum_{j=1}^{n} [(X_{t_j} - X_{t_{j-1}})^2 - (b - a/n)]\right\}^2 \\
&E[(\sum (Z_j^2 - b - a/n))^2], \ Z_j = X_{t_j} - X_{t_{j-1}}, \\
&= nE[(Z_1^2 - b - a/n)^2]
\end{aligned}$$

(because Z_j are independent and identically distributed).

$$= n[E(Z_1^4) - (b - a/n)^2] = 2(b - a/n)^2 \to 0.$$

Thus a subsequence $U_{n_i} \to b - a$ almost everywhere. Since $A_{n_i} \to 0$ it follows that $V_{n_i}(w, Q_n) \to \infty$ almost everywhere. This completes the proof. □

60 **Note.** $\{w : w$ is of bounded variation on $[a, b]\}$ can be shown to be measurable if one proves

Exercise. Let f be continuous on $[a, b]$ and define $V_n(f, Q_n)$ as above. Show that f is of bounded variation on $[a, b]$ iff $\sup\limits_{n=1,2,\ldots} V_n(f, Q_n) < \infty$.

Theorem . *Let t be any fixed real number in $[0, \infty)$, $D_t = \{w : w$ is differentiable at $t\}$. Then $P(D_t) = 0$.*

Proof. The measurability of D_t follows from the following observation: if f is continuous then f is differentiable at t if and only if

$$\operatorname*{Lt}_{\substack{r \to 0 \\ r \text{ rational}}} \frac{f(t + r) - f(t)}{r}.$$

exists. Now

$$D_t = \bigcup_{m=1}^{\infty} w : |\frac{w(t + h) - w(t)}{h}| \leq M, \text{ for all } h \neq 0, \text{ rational}\}$$

and

$$P\left\{w : |\frac{X_{t+h} - X_t}{h}| \leq M \ \forall h \in Q, h \neq 0\right\} \leq 2 \inf_h \int_0^{M\sqrt{h}} \frac{1}{\sqrt{(2\pi)}} e^{-|y|^2/2} dy = 0$$

□

Remark. A stronger result holds:

$$P\left(\bigcup_{t \geq 0} D_t\right) = 0.$$

Hint: $\bigcup_{0\le t\le 1} D_t \bigcup_{=1} \bigcup_{m=1} {}_{n-m} {}_{i=1} \bigcup_{k=i+1,i+2,i+3}^{n+2} \left\{ w : w\left(\frac{k}{n}\right) - w\left(\frac{k-1}{n}\right) | \le \frac{71}{n} \right\}$

and

$$P\left({}_{i=1} \bigcup_{k=i+1,\ldots,i+3}^{n+2} \left\{ w : w\left(\frac{k}{n}\right) - w\left(\frac{k-1}{n}\right) | \le \frac{71}{n} \right\} \quad \text{const}/\sqrt{n} \right)$$

This construction is due to A. Dvoretski, P. Erdos & S. Kakutani.

10. Dirichlet Problem and Brownian Motion

LET G BE ANY bounded open set in $\mathbb{R}^d$. Define the exit time $\tau_G(w)$ as follows: 61

$$\tau_G(w) = \{\inf t : w(t) \notin G\}.$$

If $w(0) \in G$, $\tau_G(w) = 0$; if $w(0) \in G$, $\tau_G(w)$ is the first time w escapes G or, equivalently, it is the first time that w hits the boundary ∂G of G. Clearly $\tau_G(w)$ is a stopping time. By definition $X_{\tau_G}(w) \in \partial G$, $\forall w$ and X_{τ_G} is a random variable. We can define a Borel probability measure on ∂G by

$$\begin{aligned} \pi_G(x, \Gamma) &= P_x(X_{\tau_G} \in \Gamma) \\ &= \text{probability that } w \text{ hits } \Gamma. \end{aligned}$$

If f is a bounded, real-valued measurable funciton defined on ∂G, we define

$$u(x) = E_x(f(X_{\tau_G})) = \int\limits_{\partial G} f(y)\pi_G(x, dy)$$

where

$$E_x = E^{P_x}.$$

In case G is a sphere centred around x, the exact form of $\pi_G(x, \Gamma)$ is computable.

Theorem . *Let $S = S(0; r) = \{y \in \mathbb{R}^d : |y| < r\}$. Then*

$$\pi_S(0, r) = \frac{\textit{surface area of } \Gamma}{\textit{surface area of } S}.$$

62 *Proof.* The distributions $\{F_{t_1,\ldots,t_k}\}$ defining Brownian motion are invariant under rotations. Thus $\pi_S(0,\cdot)$ is a rotationally invariant probability measure. The result follows from the fact that the only probability measure (on the surface of a sphere) that is invariant under rotations is the normalised surface area. □

Theorem . *Let G be any bounded region, f a bounded measurable real valued function defined on ∂G. Define $u(x) = E_x(f(X_{\tau_G}))$. Then*

(i) *u is measurable and bounded;*

(ii) *u has the mean value property; consequently,*

(iii) *u is harmonic in G.*

Proof. (i) To prove this, it is enough to show that the mapping $x \to P_x(A)$ is measurable for every Borel set A.

Let $\mathscr{C} = \{A \in \mathscr{B} : x \to P_x(A)$ is measurable$\}$

It is clear that $\pi^{-1}_{t_1,\ldots,t_k}(B) \in \mathscr{C}$, $\forall$ Borel set B in $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$. As $\mathscr{C}$ is a monotone class $\mathscr{C} = \mathscr{B}$.

(ii) Let S be any sphere with centre at x, and $S \subset G$. Let $\tau = \tau_S$ denote the exit time through S. Clearly $\tau \leq \tau_G$. By the strong Markov property,

$$u(X_\tau) = E(f(X_{\tau_G})|\mathscr{F}_\tau).$$

Now

$$\begin{aligned}
u(x) &= E_x(f(X_{\tau_G})) = E_x(E(f(X_{\tau_G}))|\mathscr{F}_\tau) \\
&= E_x(u(X_\tau)) = \int\limits_{\partial S} u(y)\pi_S(x, dy) \\
&= \frac{1}{|S|}\int\limits_{\partial S} u(y)dS;\ |S| = \text{ surface area of } S.
\end{aligned}$$

63

(iii) is a consequence of (i) and (ii). (See exercise below).

□

Exercise ′. Let u be a bounded measurable function in a region G satisfying the mean value property, i.e.

$$u(x) = \frac{1}{|S|}\int\limits_{\partial S} u(y)dS$$

for every sphere $S\ G$. Then

(i) $u(x) = \dfrac{1}{v \in lS}\int\limits_{S} u(y)dy.$

(ii) Using (i) show that u is continuous.

(iii) Using (i) and (ii) show that u is harmonic.

We shall now solve the boundary value problem under suitable conditions on the region G.

Theorem . *Let G, f, u be as in the previous theorem. Further suppose that*

(i) *f is continuous;*

(ii) *G satisfies the exterior cone condition at every point of ∂G, i.e. for each $y \in \partial G$ there exists a cone C_h with vertex at the point y of height h and such that $C_h - \{y\} \subset$ exterior of G. Then*

$$\lim_{x\to y, x\in G} u(x) = f(y),\ \forall y \in \partial G.$$

64

Proof.

Step 1. $P_y\{w : w(0) = y, w$ remains in $\overline{G}$ for some positive time$\} = 0$.

Let $A_n = \{w : w(0) = y, w(s) \in \overline{G}$ for $0 \leq s \leq 1/n\}$,

$B_n = \Omega - A_n$, $A = \bigcup\limits_{n=1}^{\infty} A_n$, $B = \bigcap\limits_{n=1}^{\infty} B_n$.

As A_n's are increasing, B_n's are decreasing and $B_n \in \mathscr{F}_{1/n}$; so that $B \in \mathscr{F}_{0+}$. We show that $P(B) > 0$, so that by Bluementhal's zero-one law, $P(B) = 1$, i.e. $P(A) = 0$.

$$P_y(B) = \lim_{n\to\infty} P_y(B_n) \geq \overline{\lim_{n\to\infty}}\, P_y\{w : w(0) = y, w(\frac{1}{2n}) \in C_h - \{y\}\}$$

Thus

$$P_y(b) \geq \overline{\lim} \int\limits_{C_h-\{y\}} 1/\sqrt{(2\pi/2n)^d} \exp(-|z-y|^2/2/2n)dz$$
$$= \int\limits_{C_\infty} 1/\sqrt{(2\pi)} e^{-|y|^2/2} dy,$$

where C_∞ is the cone of infinite height obtained from C_h. Thus $P_y(B) > 0$.

Step 2. If C is closed then the mapping $x \to P_x(C)$ is upper semi-continuous.

For, denote by X_C the indicator function of C. As C is closed (in a metric space) $\exists$ a sequence of continuous functions f_n decreasing to X_C such that $0 \leq f_n \leq 1$. Thus $E_x(f_n)$ decreases to $E_x(X_C) = P_x(C)$. Clearly $x \to E_x(F_n)$ is continuous. The result follows from the fact that the infimum of any collection of continuous functions is upper semi-continuous.

65 **Step 3.** Let $\delta > 0$,

$$N(y; \delta) = \{z \in \partial G : |z - y| < \delta\},$$
$$B_\delta = \{w : w(0) \in G, X_{\tau_G}(w) \in \partial G - N(y; \delta)\},$$

i.e. B_δ consists of trajectories which start at a point of G and escape for the first time through ∂G at a point not in $N(y; \delta)$. If $C_\delta = \overline{B}_\delta$, then

$$C_\delta \cap \{w : w(0) = y\} \subset A \cap \{w : w(0) = y\}$$

where A is as in Step 1.

For, suppose $w \in C_\delta \cap \{w : w(0) = y\}$. Then there exists $w_n \in B_\delta$ such that $w_n \to w$ uniformaly on compact sets. If $w \notin A \cap \{w : w(0) = y\}$ there exists $\epsilon > 0$ such that $w(t) \in \overline{G}\, \forall t$ in $(0, \epsilon]$. Let $\delta^* = \inf\limits_{0 \leq t \leq \epsilon} d(w(t), G - N(y, \delta))$. Then $\delta^* > 0$. If $t_n = \tau_G(w_n)$ and t_n does not converge to 0, then there exists a subsequence, again denoted by t_n, such that $t_n \geq k\epsilon > 0$ for some $k \in (0, 1)$. Since $w_n(k\epsilon) \in \overline{G}$ and $w_n(k\epsilon)$, $w(k\epsilon) \in \overline{G}$, a contradiction. Thus we can assume that t_n converges to 0 and also that $\epsilon \geq t_n \forall n$, But then

$$(*) \qquad |w_n(t_n) - w(t_n)| \geq \delta^*.$$

However, as w_n converges to w uniformly on $[0, \epsilon]$,

$$w_n(t_n) - w(t_n) \to w(0) - w(0) = 0$$

contradicting (*). Thus $w \in A\{w : w(0) = y\}$.

Step 4. $\lim\limits_{x \to y, x \in G} P_x(B_\delta) = 0$.

For, 66

$$\begin{aligned}
\overline{\lim_{x\to y}}\, P_x(B_\delta) &\leq \overline{\lim_{x\to y}}\, P_x(C_\delta) \leq P_y(C_\delta) \quad \text{(by Step 2)} \\
&= P_y(C_\delta \cap \{w : w(0) = y\}) \\
&\leq P_y(A) \quad \text{(by Step 3)} \\
&= 0.
\end{aligned}$$

Step 5.

$$\begin{aligned}
|u(x) - f(y)| &= |\int_\Omega f(X_{\tau_G}(w)) dP_x(w) - \int_\Omega f(y) dP_x(w)| \\
&\leq \int_{\Omega - B_\delta} |f(X_{\tau_G}(w)) - f(y)| dP_x(w) + |\int_{B_\delta} (f(X_{\tau_G}(w)) - f(y)) dP_x(w)| \\
&\leq \int_{\Omega - B_\delta} |f(X_{\tau_G}(w)) - f(y)| dP_x(w) + 2\|f\|_\infty P_x(B_\delta)
\end{aligned}$$

and the right hand side converges to 0 as $x \to y$ (by Step 4 and the fact that f is continuous). This proves the theorem.

□

Remark. The theorem is local.

Theorem . *Let $G = \{y \in \mathbb{R}^d : \delta < |y| < R\}$, f any continuous function on $\partial G = \{|y| = \delta\} \cap \{|y| = R\}$. If u is any harmonic function in G with boundary values f, then $u(x) = E_x(f(X_{\tau_G}))$.*

Proof. Clearly G has the exterior cone property. Thus, if

$$v(x) = E_x(f(X_{\tau_G})),$$

then v is harmonic in G and has boundary values f (by the previous theorem). The result follows from the uniqueness of the solution of the Dirichlet problem for the Laplacian operator.

The function $f = 0$ on $|y| = R$ and $f = 1$ on $|y| = \delta$ is of special interest. Denote by $\cup^{R,0}_{\delta,1}$ the corresponding solution of the Dirichlet problem. □

67 **Exercise.** (i) If $d = 2$ then

$$U^{R,0}_{\delta,1}(x) = \frac{\log R - \log |x|}{\log R - \log \delta}, \ \forall x \in G.$$

(ii) If $d \geq 3$ then

$$U^{R,0}_{\delta,1}(x) = \frac{|x|^{-n+2} - R^{-n+2}}{\delta^{-n+2} - R^{-n+2}}.$$

Case (i): $d = 2$. Then

$$\frac{\log R - \log |x|}{\log R - \log \delta} = U^{R,0}_{\delta,1}(x).$$

Now,

$$E_x(f(X_{\tau_G})) = \int\limits_{|y|=\delta} \pi_G(x, dy) = P_x(|X_{\tau_G}| = \delta),$$

i.e.

$$\frac{\log R - \log |x|}{\log R - \log \delta} = P_x(|X_{\tau_G}| = \delta)$$

$$P_x \quad \text{(the particle hits } |y| = \delta \text{ before it hits } |y| = R).$$

Fix R and let $\delta \to 0$; then $0 = P_x$ (the particle hits 0 before hitting $|y| = R$).

Let R take values $1, 2, 3, \ldots$, then $0 = P_x$ (the particle hits 0 before hitting any of the circles $|y| = N$). Recalling that

$$P_x(\overline{\lim}\, |X_t| = \infty) = 1,$$

we get

Proposition . *A two-dimensional Brownian motion does not visit a point.*

Next, keep δ fixed and let $R \to \infty$, then,

$$1 = P_x(|w(t)| = \delta \quad \textit{for some time} \quad t > 0).$$

Since any time t can be taken as the starting time for the Brownian motion, we have 68

Proposition . *Two-dimensional Brownian motion has the recurrence property.*

Case (ii): $d \geq 3$. In this case

$$\begin{aligned} &P_x(w : w \text{ hits } |y| = \delta \text{ before it hits } |y| = R) \\ &= (1/|x|^{n-2} - 1/R^{n-2})/(1/\delta^{n-2} - 1/R^{n-2}). \end{aligned}$$

Letting $R \to \infty$ we get

$$P_x(w : w \text{ hits } |y| = \delta) = (\delta/|x|)^{n-2}$$

which lies strictly between 0 and 1. Fixing δ and letting $|x| \to \infty$, we have

Proposition . *If the particle start at a point for away from* 0 *then it has very little chance of hitting the circle $|y| = \delta$.*

If $|x| \leq \delta$, then

$$P(w \textit{ hits } S_\delta) = 1 \textit{ where } S_\delta = \{y \in R^d : |y| = \delta\}.$$

Let

$$V_\delta(x) = (\delta/|x|)^{n-2} \text{ for } |x| \geq \delta.$$

In view of the above result it is natural to extend V_δ to all space by putting $V_\delta(x) = 1$ for $|x| \leq \delta$. As Brownian motion has the Markov property

$$P_x\{w : w \text{ hits } S_\delta \text{ after time } t\}$$
$$= \int V_\delta(y) 1/\sqrt{(2\pi t)^d} \exp -|y|^2/2t \, dy \to 0 \text{ as } t \to +\infty.$$

69 Thus $P(w$ hits S_δ for arbitrarily large $t) = 0$. In other words, $P(w : \varliminf_{t\to\infty} |w(t)| \geq \delta) = 1$. As this is true $\forall \delta > 0$, we get the following important result.

Proposition . $P(\varliminf_{t\to\infty} |w(t)| = \infty) = 1$,

i.e. for $d \geq 3$, the Brownian particle wander away to $+\infty$.

11. Stochastic Integration

LET $\{X_t : t \geq 0\}$ BE A one-dimensional Brownian motion. We want **70**
first to define integrals of the type $\int\limits_0^\infty f(s)dX(s)$ for real functions $f \in L^1[0,\infty)$. If $X(s,w)$ is of bounded variation almost everywhere then we can give a meaning to $\int\limits_0^\infty f(s)dX(s,w) = g(w)$. However, since $X(s,w)$ is not bounded variation almost everywhere, $g(w)$ is not defined in the usual sense.

In order to define $g(w) = \int\limits_0^\infty f(s)dX(s,w)$ proceed as follows.

Let f be a step function of the following type:

$$f = \sum_{i=1}^n a_i X_{[t_i,t_{i+1})}, 0 \leq t_1 < t_2 < \ldots < t_{n+1}.$$

We naturally define

$$\begin{aligned} g(w) = \int\limits_0^\infty f(s)dX(s,w) &= \sum_{i=1}^n a_i(X_{t_{i+1}}(w) - X_{t_i}(w)) \\ &= \sum_{i=1}^n a_i(w(t_{i+1}) - w(t_i)). \end{aligned}$$

g satisfies the following properties:

(i) g is a random variable;

(ii) $E(g) = 0$; $E(g^2) = \sum a_i^2 (t_{i+1} - t_i) = \|f\|_2$.

This follows from the facts that (a) $X_{t_{i+1}} - X_{t_i}$ is a normal random variable with mean 0 and variance $(t_{i+1} - t_i)$ and (b) $X_{t_{i+1}} - X_{t_i}$ are independent increments, i.e. we have

$$E\left(\int_0^\infty f dX\right) = 0,\ E\left(|\int_0^\infty f dX|^2\right) = \|f\|_2^2.$$

71

Exercise 1. If

$$f = \sum_{i=1}^{n} a_i X_{[t_i, t_{i+1})}, 0 \le t_1 < \ldots < t_{n+1},$$
$$g = \sum_{i=1}^{m} b_i X_{[s_i, s_{i+1})}, 0 \le s_1 < \ldots < s_{m+1},$$

Show that

$$\int_0^\infty (f + g) dX(s, w) = \int_0^\infty f dX(s, w) + \int_0^\infty g dX(s, w)$$

and

$$\int_0^\infty (\alpha f) dX(s, w) = \alpha \int_0^\infty f dX(s, w),\ \forall \alpha \in \mathbb{R}.$$

Remark. The mapping $f \to \int\limits_0^\infty f dX$ is therefore a linear $L^2_{\mathbb{R}}$-isometry of the space S of all simple functions of the type

$$\sum_{i=1}^{n} a_i X_{[t_i, t_{i+1})}, (0 \le t_1 < \ldots < t_{n+1})$$

into $L^2(\Omega, \mathscr{B}, P)$.

Exercise 2. Show that S is a dense subspace of $L^2[0,\infty)$.
Hint: $C_c[0,\infty)$, i.e. the set of all continuous functions with compact support, is dense in $L^2[0,\infty)$. Show that S contains the closure of $C_c[0,\infty)$.

Remark. The mapping $f \to \int\limits_0^\infty f dX$ can now be uniquely extended as an isometry of $L^2[0,\infty)$ into $L^2(\Omega,\mathscr{B},P)$.

Next we define integrals fo the type **72**

$$g(w) = \int\limits_0^t X(s,w)dX(s,w)$$

Put $t = 1$ (the general case can be dealt with similarly). It seems natural to define

$$(*) \qquad \int\limits_0^1 X(s,w)dX(s) = \underset{\sup|t_j-t_{j-1}|\to 0}{\mathrm{Lt}} \sum_{j=1}^n X(\xi j)(X(t_j)-X(t_{j-1}))$$

where $0 = t_0 < t_1 < \ldots < t_n = 1$ is a partion of $[0,1]$ with $t_{j-1} \leq \xi_j \leq t_j$. In general the limit on the right hand side may not exist. Even if it exists it may happen that depending on the choice of ξ_j, we may obtain different limits. To consider an example we choose $\xi_j = t_j$ and then $\xi_j = t_{j-1}$ and compute the right hand side of $(*)$. If $\xi_j = t_{j-1}$,

$$\begin{aligned}\sum_{j=1}^n X_{\xi_j}(X_{t_j}-X_{t_{j-1}}) &= \sum_{j=1}^n X_{t_{j-1}} - (X_{t_j}-X_{t_{j-1}})\\ &= \frac{1}{2}\sum_{j=1}^n (X_{t_j}) - (X_{t_{j-1}}) - \frac{1}{2}\sum_{j=1}^n (X_{t_j}-X_{t_{j-1}})\end{aligned}$$

$$\frac{1}{2}[X^2(1)-X^2(0)] - \frac{1}{2} \text{ as } n\to\infty, \text{ and } \sup|t_j-t_{j-1}|\to 0,$$

arguing as in the proof of the result that Brownian motion is not of bounded variation. If $\xi_j = t_j$,

$$\underset{\substack{n\to\infty\\ \mathrm{Sup}|t_j-t_{j-1}|\to 0}}{\mathrm{Lt}} \sum_{j=1}^n X_{t_j}(X_{t_j}-X_{t_{j-1}}) = 1/2X(1) - 1/2X(0) + 1/2.$$

73 Thus we get different answers depending on the choice of ξ_j and hence one has to be very careful in defining the integral. It turns out that the choice of $\xi_j = t_{j-1}$ is more appropriate in the definition of the integral and gives better results.

Remark. The limit in $(*)$ should be understood in the sense of convergence probability.

Exercise 3. Let $0 \leq a < b$. Show that the "left integral" ($\xi_j = t_{j-1}$) is given by

$$L\int\limits_a^b X(s)dX(s) = \frac{X^2(b) - X^2(a) - (b-a)}{2}$$

and the "right integral" ($\xi_j = t_j$) is given by

$$R\int\limits_s^b X(s)dX(s) = \frac{X^2(b) - X^2(a) + (b-a)}{2}.$$

We now take up the general theory of stochastic integration. To motivate the definitions which follow let us consider a d-dimensional Brownian motion $\{\beta(t) : t \geq 0\}$. We have

$$E[\beta(t+s) - \beta(t) \in A|\mathscr{F}_t] = \int\limits_A 1/\sqrt{(2\pi s)} e^{-|y|^2/2s} dy.$$

Thus

$$E(f(\beta(t+s) - \beta(t))|\mathscr{F}_t] = \int f(y) 1/\sqrt{(2\pi s)} e^{-|y|^{2/2s}} dy.$$

In particular, if $f(x) = e^{ix.u}$,

$$\begin{aligned} E[e^{iu}(\beta(t+s) - \beta(t))|\mathscr{F}_t] &= \int e^{iu.y} 1/\sqrt{(2\pi s)} e^{-|y|^2/2s} dy \\ &= e^{\frac{-s|u|^2}{2}}. \end{aligned}$$

74 Thus

$$E[e^{iu.\beta(t+s)}|\mathscr{F}_t] = e^{iu.\beta(t)}e^{-s|u|^2/2},$$

or,

$$E[e^{iu.\beta(t+s)+(t+s)|u|^2/2}|\mathscr{F}_t] = e^{iu.\beta(t)+t|u|^2/2}.$$

Replacing iu by θ we get

$$E[e^{\theta.\beta(s)-|s\theta|^2/2} \mid \mathscr{F}_t] = e^{\theta.\beta(t)-t|\theta|^{2/2}},\ s > t, \forall\theta.$$

It is clear that $e^{\theta.\beta(t)-t|\theta|^2/2}$ is $\mathscr{F}_t$-measurable and a simple calculation gives

$$E(e^{\theta.\beta(t)-|\theta|^2 t/2|}) < \infty\ \forall\theta.$$

We thus have

Theorem . *If $\{\beta(t) : t \geq 0\}$ is a d-dimensional Brownian motion then* $\exp[\theta.\beta(t) - |\theta|^2 t/2]$ *is a Martingale relative to $\mathscr{F}_t$, the σ-algebra generated by $(\beta(s) : s \leq t)$.*

Definition. Let $(\Omega, \mathscr{B}, P)$ be a probability space $(\mathscr{F}_t)_{t\geq 0}$ and increasing family of sub-σ-algebras of $\mathscr{F}$ with $\mathscr{F} = \sigma(\bigcup\limits_{t\geq 0} \mathscr{F}_t)$.

Let

(i) $a : [0,\infty)\times\Omega \to [0,\infty)$ be bounded and progressively measurable;

(ii) $b : [0,\infty)\times\Omega \to \mathbb{R}$ be bounded and progressively measurable;

(iii) $X : [0,\infty)\times\Omega \to \mathbb{R}$ be progressively measurable, right continuous on $[0,\infty)$, $\forall\ w \in \Omega$, and continous on $[0,\infty)$ almost everywhere on Ω;

(iv) $Z_t(w) = e^{\theta X(t,w)-\theta\int\limits_0^t b(s,w)ds-\frac{\theta^2}{2}\int\limits_0^t a(s,w)ds}$ 75

be a Martingale relative to $(\mathscr{F}_t)_{t\geq 0}$.

Then $X(t,w)$ is called an Ito process corresponding to the parameters b and a and we write $X_t \in I[b,a]$.

N.B. The progressive measurability of $X \Rightarrow X_t$ is $\mathscr{F}_t$-measurable.

Example. If $\{\beta(t) : t \geq 0\}$ is a Brownian motion, then $X(t, w) = \beta_t(w)$ is an Ito process corresponding to parameters 0 and 1. (i) and (ii) are obvious. (iii) follows by right continuity of β_t and measurability of β_t relative to $\mathscr{F}_t$ and (iv) is proved in the previous theorem.

Exercise 4. Show that $Z_t(w)$ defined in (iv) is $\mathscr{F}_t$-measurable and progressively measurable.

[Hint:

(i) Z_t is right continuous.

(ii) Use Fubini's theorem to prove measurability].

Remark. If we put $Y(t, w) = X(t, w) - \int\limits_0^t b(s, w)ds$ then $Y(t, w)$ is progressively measurable and $Y(t, w)$ is an Ito process corresponding to the parameters $0, a$. Thus we need only consider integrals of the type $\int\limits_0^t f(s, w)dY(s, w)$ and *define*

$$\int\limits_0^t f(s, w)dX(s, w) = \int\limits_0^t f(s, w)dY(s, w) + \int\limits_0^t f(s, w)b(s, w)ds.$$

(Note that *formally* we have $dY = dX - dbt$).

76 **Lemma .** *If $Y(t, w) \in I[0, a]$, then*

$$Y(t, w) \quad \textit{and} \quad Y^2(t, w) - \int\limits_0^t a(s, w)ds$$

are Martingales relative to $(\mathscr{F}_t)$.

Proof. To motivate the arguments which follow, we first give a formal proof. Let

$$Y_\theta(t) = e^{\theta Y(t,w) - \frac{\theta^2}{2}\int\limits_0^t a(s,w)ds}.$$

Then $Y_\theta(t)$ is a martingale, $\forall\theta$. Therefore $\dfrac{Y_\theta - 1}{\theta}$ is a Martingale, $\forall\theta$. Hence (formally),

$$\lim_{\theta\to 0}\frac{Y_\theta - 1}{\theta} = Y'_\theta|_{\theta=0}$$

is a Martingale.

Step 1. $Y(t,\cdot) \in L^k(\Omega,\mathscr{F},P)$, $k = 0,1,2,\ldots$ and $\forall t$. In fact, for every real θ, $Y_\theta(t)$ is a Martingale and hence $E(Y_\theta) < \infty$. Since a is bounded this means that

$$E(e^{\theta Y(t,\cdot)}) < \infty,\ \forall\theta.$$

Taking $\theta = 1$ and -1 we conclude that $E(e^{|Y|}) < \infty$ and hence $E(|Y|^k) < \infty$, $\forall k = 0,1,2,\ldots$. Since Y is an Ito process we also get

$$\sup_{|\theta|\leq\alpha} E\left(\left[e^{Y(t,\cdot)-\frac{\theta^2}{2}\int\limits_0^t a\,ds}\right]^k\right) < \infty$$

$\forall k$ and for every $\alpha > 0$.

Step 2. Let $X_\theta(t) = [Y(t,\cdot) - \theta\int\limits_0^t a\,ds]Y(t) = \dfrac{d}{d\theta}Y_\theta(t,\cdot)$. **77**

Define

$$\phi_A(\theta) = \int\limits_A (X_\theta(t,\cdot) - X_\theta(s,\cdot))dP(w)$$

where $t > s$, $A \in \mathscr{F}_s$. Then

$$\int\limits_{\theta_1}^{\theta_2}\phi_A(\theta)d\theta = \int\limits_{\theta_1}^{\theta_2}\int\limits_A [X_\theta(t,\cdot) - X_\theta(S,\cdot)]dP(w)d\theta.$$

Since a is bounded, $\sup\limits_{|\theta|\leq\alpha} E([Y_\theta(t,\cdot)]^k) < \infty$, and $E(|Y|^k) < \infty$, $\forall k$; we can use Fubini's theorem to get

$$\int\limits_{\theta_1}^{\theta_2}\phi_A(\theta)d\theta = \int\limits_A\int\limits_{\theta_1}^{\theta_2}[X_\theta(t,\cdot) - X_\theta(s,\cdot)]d\theta\, dP(w).$$

or

$$\int_{\theta_1}^{\theta_2} \phi_A(\theta)d\theta = \int_A Y_{\theta_2}(t,\cdot) - Y_{\theta_1}(t,\cdot)dP(w) - \int_A Y_{\theta_1}(s,\cdot) - Y_{\theta_1}(s,\cdot)dP(w).$$

Let $A \in \mathscr{F}_s$ and $t > s$; then, since Y is a Martingale,

$$\int_{\theta_1}^{\theta_2} \phi_A(\theta)d\theta = 0.$$

This is true $\forall \theta_1 < \theta_2$ and since $\phi_A(\theta)$ is a continuous function of θ, we conclude that

$$\phi_A(\theta) = 0, \ \forall \theta.$$

In particular, $\phi_A(\theta) = 0$ which means that

$$\int_A Y(t,\cdot)dP(w) = \int_A Y(s,\cdot)dP(w), \ \forall A \in \mathscr{F}_s, \ t > s,$$

78 i.e., $Y(t)$ is a Martingale relative to $(\Omega, \mathscr{F}_t, P)$.

To prove the second part we put

$$Z_\theta(t,\cdot) = \frac{d^2}{d\theta^2} Y_\theta(t)$$

and

$$\psi_A(\theta) = \int_A \{Z_\theta(t,\cdot) - Z_\theta(s,\cdot)\}dP(w).$$

Then, by Fubini,

$$\int_{\theta_1}^{\theta_2} \psi_A(\theta)d\theta = \int_A \int_{\theta_1}^{\theta_2} Z_\theta(t,\cdot) - Z_\theta(s,\cdot)d\theta \, dP(w).$$

or,

$$\int_{\theta_1}^{\theta_2} \psi_A(\theta) d\theta = \phi_A(\theta_2) - \phi_A(\theta_1)$$

$$= 0 \text{ if } A \in \mathscr{F}_s, \ t > s.$$

Therefore

$$\psi_A(\theta) = 0, \ \forall \theta.$$

In particular, $\psi_A(\theta) = 0$ implies that

$$Y^2(t,w) - \int_0^t a(s,w) ds$$

is an $(\Omega, \mathscr{F}_t, P)$ Martingale. This completes the proof of lemma 1. $\square$

Definition. A function $\theta : [0, \infty) \times \Omega \to \mathbb{R}$ is called simple if there exist reals $s_0, s_1, \ldots, s_n, \ldots$

$$0 \leq s_0 < s_1 < \ldots < s_n \ldots < \infty,$$

s_n increasing to $+\infty$ and

$$\theta(s,w) = \theta_j(w)$$

if $s \in [s_j, s_{j+1})$, where $\theta_j(w)$ is $\mathscr{F}_{s_j}$-measurable and bounded. **79**

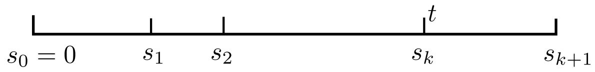
Definition. Let $\theta : [0, \infty) \times \Omega \to \mathbb{R}$ be a simple function and $Y(t,w) \in I[0,a]$. We define the stochastic integral of θ with respect to Y, denoted

$$\int_0^t \theta(s,w) dY(s,w)),$$

by

$$\xi(t,w) = \int_0^t \theta(s,w) dY(s,w)$$

$$= \sum_{j=1}^{k} \theta_{j-1}(w)[Y(s_j, w) - Y(s_{j-1}, w)] + \theta_k(w)[Y(t, w) - Y(s_k, w)].$$



Lemma 2. *Let $\sigma : [0, \infty) \times \Omega \to \mathbb{R}$ be a simple function and $Y(t, w) \in I[0, a]$. Then*

$$\xi(t, w) = \int_0^t \sigma(s, w) dY(s, w) \in I[0, a\sigma^2].$$

Proof. (i) By definition, σ is right continuous and $\sigma(t, w)$ is $\mathscr{F}_t$-measurable; hence it is progressively measurable. Since a is progressively measurable and bounded

$$a\sigma^2 : [0, \infty) \times \Omega \to [0, \infty)$$

is progressively measurable and bounded.

80 (ii) From the definition of ξ it is clear that $\xi(t, \cdot)$ is right continuous, continous almost everywhere and $\mathscr{F}_t$-measurable therefore ξ is progressively measurable.

(iii) $Z_t(w) = e^{[\theta\xi(t,w) - \frac{\theta^2}{2}\int_0^t a\sigma^2 ds]}$

is clearly $\mathscr{F}_t$-measurable $\forall\theta$. We show that

$$E(Z_t) < \infty, \ \forall t \text{ and } E(Z_{t_2} | \mathscr{F}_{t_1}) = Z_{t_1} \text{ if } t_1 < t_2.$$

We can assume without loss of generality that $\theta = 1$ (if $\theta \neq 1$ we replace σ by $\theta\sigma$). Therefore

$$Z_t(w) = e^{[\xi(t,w) - \int_0^t a\sigma^2 ds]}.$$

Since a and σ are bounded uniformly on $[0, t]$, it is enough to show that $E(e^{\xi(t,w)}) < \infty$. By definition,

$$\xi(t, w) = \sum_{j=1}^{k} \theta_{j-1}(w)[Y(s_j, w) - Y(s_{j-1}, w)] + \theta_k(w)(Y(t, w) - Y(s_k, w)).$$

The result $E(e^{\xi(t,w)}) < \infty$ will follow from the generalised Holder's inequality provided we show that

$$E(e^{\theta(w)[Y(t,w)-Y(s,w)]}) < \infty$$

for every bounded function θ which is $\mathscr{F}_s$-measurable. Now

$$E(e^{\theta[Y(t,\cdot)-Y(s,\cdot)]}|\mathscr{F}_s) =$$

constant for every constant θ, since $Y \in I[0, a]$. Therefore

$$E(e^{\theta(w)[Y(t,\cdot)-Y(s,\cdot)]}|\mathscr{F}_s) = \text{ constant}$$

for every θ which is bounded and $\mathscr{F}_s$-measurable. Thus **81**

$$E(e^{\theta(w)[Y(t,\cdot)-Y(s,\cdot)]}) < \infty.$$

This proves that $E(Z_t(w)) \in \infty$.

Finally we show that

$$E(Z_{t_2}|\mathscr{F}_{t_1}) = Z_{t_1}(w), \quad \text{if} \quad t_1 < t_2.$$

Consider first the case when t_1 and t_2 are in the same interval

$$[s_k, s_{k+1}).$$

Then

$$\xi(t_2, w) = \xi(t_1, w) + \sigma_k(w)[Y(t_2, w) - Y(t_1, w)] \quad \text{(see definition)},$$

$$\int_0^{t_2} a\sigma^2(s, w)ds = \int_0^{t_1} a\sigma^2(s, w)ds + \int_{t_1}^{t_2} a\sigma^2(s, w)ds.$$

Therefore

$$E(Z_{t_2}(w)|\mathscr{F}_{t_1}) = Z_{t_1}(w)E(\exp[\theta\sigma_k(w)[Y(t_2,w) - Y(t_1,w)] - \frac{\theta^2}{2}\int_{t_1}^{t_2} a\sigma^2 ds)|\mathscr{F}_{t_1})$$

as $Y \in I[0,a]$.

$$(*) \qquad E(\exp[\theta(Y(t_2,w) - T(t_1,w)) - \frac{\theta^2}{2}\int_{t_1}^{t_2} a(s,w)ds]|\mathscr{F}_{t_1}) = 1$$

and since $\sigma_k(w)$ is $\mathscr{F}_{t_1}$-measurable $(*)$ remains valid if θ is replaced by $\theta\sigma_k$. Thus

$$E(Z_{t_2}|\mathscr{F}_{t_1}) = Z_{t_1}(w).$$

The general case follows if we use the identity

$$E(E(X|\mathscr{C}_1)|\mathscr{C}_2) = E(X|\mathscr{C}_2) \quad \text{for} \quad \mathscr{C}_2 \subset \mathscr{C}_1.$$

Thus Z_t is a Martingale and $\xi(t,w) \in I[0, a\sigma^2]$. □

82 **Corollary .** *(i) $\xi(t,w)$ is a martingale; $E(\xi(t,w)) = 0$;*

(ii) $\xi^2(t,w) - \int_0^t a\sigma^2 ds$

is a Martingale with

$$E(\xi^2(t,w)) = E(\int_0^t a\sigma^2(s,w)ds.$$

Proof. Follows from Lemma 1. □

Lemma 3. *Let $\sigma(s,w)$ be progressively measurable such that for each t,*

$$E(\int_0^t \sigma^2(s,w)ds) < \infty.$$

Then there exists a sequence $\sigma_n(s, w)$ of simple functions such that

$$\lim_{n\to\infty} E\left(\int_0^t |\sigma_n(s, w) - \sigma(s, w)|^2 ds\right) = 0.$$

Proof. We may assume that σ is bounded, for if $\sigma_N = \sigma$ for $|\sigma| \leq N$ and 0 if $|\sigma| > N$, then $\sigma_n \to \sigma$, $\forall (s, w) \in [0, t] \times \Omega$. σ_N is progressively measurable and $|\sigma_n - \sigma|^2 \leq 4|\sigma|^2$. By hypothesis $\sigma \in L([0, t] : \Omega)$.

Therefore $E(\int\limits_0^t |\sigma_n - \sigma| ds) \to 0$, by dominated convergence. Further, we can also assume that σ is continuous. For, if σ is bounded, define

$$\sigma_h(t, w) = 1/h \int_{(t-h)v0}^{t} \sigma(s, w)ds.$$

σ_n is continuous in t and $\mathscr{F}_t$-measurable and hence progressively measurable. Also by Lebesgue's theorem

$$\sigma_h(t, w) \to \sigma(t, w), \quad \text{as} \quad h \to 0, \forall t, w.$$

Since σ is bounded by C, σ_h is also bounded by C. Thus **83**

$$E(\int_0^t |\sigma_h(s, w) - \sigma(s, w)|^2 ds) \to 0.$$

(by dominated convergence). If σ is continuous, bounded and progressively measurable, then

$$\sigma_n(s, w) = \sigma\left(\frac{[ns]}{n}, w\right)$$

is progressively measurable, bounded and simple. But

$$\operatorname*{Lt}_{n\to\infty} \sigma_n(s, w) = \sigma(s, w).$$

Thus by dominated convergence

$$E\left(\int_0^t |\sigma_n - \sigma|^2 ds\right) \to 0 \quad \text{as} \quad n \to \infty.$$

□

Theorem . *Let $\sigma(s,w)$ be progressively measurable, such that*

$$E(\int_0^t \sigma^2(s,w)ds) < \infty$$

for each $t > 0$. Let (σ_n) be simple approximations to σ as in Lemma 3. Put

$$\xi_n(t,w) = \int_0^t \sigma_n(s,w)dY(s,w)$$

where $Y \in I[0,a]$. Then

(i) $\operatorname*{Lt}_{n\to\infty} \xi_n(t,w)$ *exists uniformly in probability, i.e. there exists an almost surely continuous $\xi(t,w)$ such that*

$$\operatorname*{Lt}_{n\to\infty} P\left(\sup_{0\le t\le T} |\xi_n(t,w) - \xi(t,w)| \ge \epsilon\right) = 0$$

for each $\epsilon > 0$ and for each T. Moreover, ξ is independent of the sequence (σ_0).

84 (ii) *The map $\sigma \to \xi$ is linear.*

(iii) *$\xi(t,w)$ and $\xi^2(t,w) - \int_0^t a\sigma^2 ds$ are Martingales.*

(iv) *If σ is bounded, $\xi \in I[0, a\sigma^2]$.*

Proof. (i) It is easily seen that for simple functions the stochastic integral is linear. Therefore

$$(\xi_n - \xi_m)(t, w) = \int_0^t (\sigma_n - \sigma_m)(s, w) dY(s, w).$$

Since $\xi_n - \xi_m$ is an almost surely continuous martingale

$$P\left(\sup_{0\le t\le T} |\xi_n(t, w) - \xi_m(t, w)| \ge \epsilon\right) \le \frac{1}{\epsilon^2} E[(\xi_n - \xi_m)^2(T, w)].$$

This is a consequence of Kolmogorov inequality (See Appendix). Since

$$(\xi_n - \xi_m)^2 - \int_0^t a(\sigma_n - \sigma_m)^2 ds$$

is a Martingale, and a is bounded,

$$(*) \qquad E[(\xi_n - \xi_m)^2(T, w)] = E\left(\int_0^T (\sigma_n - \sigma_m)^2 a\, ds\right).$$

$$\le \text{const}\, \frac{1}{\epsilon^2} E\left(\int_0^T (\sigma_n - \sigma_m)^2 ds\right).$$

Therefore

$$\operatorname*{Lt}_{n,m\to\infty} E[(\xi_n - \xi_m)^2(T, w)] = 0.$$

Thus $(\xi_n - \xi_m)$ is uniformly Cauchy in probability. Therefore there exists a progressively measurable ξ such that

$$\operatorname*{Lt}_{n\to\infty} P\left(\sup_{0\le t\le T} |\xi_n(t, w) - \xi(t, w)| \ge \epsilon\right) = 0, \ \forall \epsilon > 0, \ \forall T.$$

It can be shown that ξ is almost surely continuous.

If (σ_n) and (σ'_n) are two sequences of simple functions approximating σ, then $(*)$ shows that 85

$$E[(\xi_n - \xi'_n)^2(T, w)] \to 0.$$

Thus

$$\underset{n}{\mathrm{Lt}}\, \xi_n = \underset{n}{\mathrm{Lt}}\, \xi'_n,$$

i.e. ξ is independent of (σ_n).

(ii) is obvious.

(iii) (*) shows that $\xi_n \to \xi$ in L and therefore $\xi_n(t, \cdot) \to \xi(t, \cdot)$ in L^1 for each fixed t. Since $\xi_n(t, w)$ is a martingale for each n, $\xi(t, w)$ is a martingale.

(iv) $\xi_n^2(t, w) - \int\limits_0^t a\sigma_n^2$ is a martingale for each n.

Since $\xi_n(t, w) \to \xi(t, w)$ in L^2 for each fixed t and

$$\xi_n^2(t, w) \to \xi^2(t, w) \quad \text{in} \quad L^1 \quad \text{for each fixed } t.$$

For $\xi_n^2(t, w) - \xi^2(t, w) = (\xi_n - \xi)(\xi_n + \xi)$ and using Hölder's inequality, we get the result.

Similarly, since

$$\sigma_n \to \sigma \text{ in } L^2([0, t] \times \Omega),$$
$$\sigma_n^2 \to \sigma^2 \text{ in } L^1([0, t] \times \Omega),$$

and because a is bounded $a\sigma_n^2 \to a\sigma^2$ in $L^1([0, t] \times \Omega)$. This shows that $\xi_n^2(t, w) - \int\limits_0^t a\sigma_n^2 ds$ converges to

$$\xi^2(t, w) - \int\limits_0^t a\sigma^2 ds$$

for each t in L^1. Therefore

$$\xi^2(t,w) - \int\limits_0^t a\sigma^2 ds$$

is a martingale. **86**

(v) Let σ be bounded. To show that $\xi \in I[0,\sigma^2]$ it is enough to show that

$$e^{\theta\xi(t,w) - \frac{\theta^2}{2}\int\limits_0^t a\sigma^2 ds}$$

is a martingale for each θ, the other conditions being trivially satisfied. Let

$$Z_n(t,w) = e^{\theta\xi_n(t,w) - \frac{\theta^2}{2}\int\limits_0^t a\sigma_n^2 ds}$$

We can assume that $|\sigma_n| \leq C$ if $|\sigma| \leq C$ (see the proof of Lemma 3).

$$Z_n = \exp\left[2\theta\xi_n(t,w) - \frac{(2\theta)^2}{2}\int\limits_0^t a\sigma_n^2 ds + \theta^2\int\limits_0^t a\sigma_n^2 ds\right].$$

Thus

$$(**) \qquad E(Z_n) \leq \text{const}\, E\left(e^{2\theta\xi_n(t,w) - \frac{(2\theta)^2}{2}\int\limits_0^t a\sigma_n^2 ds}\right) = \text{const}$$

since Z_n is a martingale for each θ. A subsequence Z_{n_i} converges to

$$e^{\theta\xi(t,w) - \frac{\theta^2}{2}\int\limits_0^t a\sigma^2 ds}$$

almost everywhere (P). This together with (**) ensures uniform integrability of (Z_n) and therefore

$$e^{\theta\xi(t,w) - \frac{\theta^2}{2}\int\limits_0^t a\sigma^2 ds}$$

is a martingale. Thus ξ is an Ito process, $\xi \in I[0, a\sigma^2]$.

□

Definition. With the hypothesis as in the above theorem we define the stochastic integral

$$\xi(t,w) = \int_0^t \sigma(s,w)dY(s,w).$$

87

Exercise. Show that $d(X+Y) = dX + dY$.

Remark. If σ is bounded, then ξ satisfies the hypothesis of the previous theorem and so one can define the integral of ξ with respect to Y. Further, since ξ itself is Itô, we can also define stochastic integrals with respect to.

Examples. 1. Let $\{\beta(t) : t \geq 0\}$ be a Brownian motion; then $\beta(t,w)$ is progressively measurable (being continuous and $\mathscr{F}_t$-measurable). Also,

$$\int_\Omega \int_0^t \beta^2(s)ds\, dP = \int_0^t \int_\Omega \beta^2(s)dP\, ds = \int_0^t sds = \frac{t}{2}$$

Hence

$$\int_0^t \beta(s,w)d\beta(s,w)$$

is well defined.

2. Similarly $\int_0^t \beta(s/2)d\beta(s)$ is well defined.

3. However

$$\int_0^t \beta(2s)d\beta(s)$$

is not well defined, the reason being that $\beta(2s)$ is not progressively measurable.

Exercise 5. Show that $\beta(2s)$ is not progressively measurable. 88
(Hint: Try to show that $\beta(2s)$ is not $\mathscr{F}_s$-measurable for every s. To show this prove that $\mathscr{F}_s \neq \mathscr{F}_{2s}$).

Exercise 6. Show that for a Brownian motion $\beta(t)$, the stochastic integral

$$\int_0^1 \beta(s,\cdot)d\beta(s,\cdot)$$

is the same as the left integral

$$L\int_0^1 \beta(s,\cdot)d\beta(s,\cdot)$$

defined earlier.

12. Change of Variable Formula

WE SHALL PROVE the 89

Theorem . *Let σ be any bounded progressively measurable function and Y be an Ito process. If λ is any progressively measurable function such that*

$$E\left(\int_0^t \lambda^2 ds\right) < \infty, \ \forall t,$$

then

$$(*) \qquad \int_0^t \lambda d\xi(s,w) = \int_0^t \lambda(s,w)\sigma(s,w)dY(s,w),$$

where

$$\xi(t,w) = \int_0^t \sigma(s,w)dY(s,w).$$

Proof.

Step 1. Let λ and σ be both simple, with λ bounded. By a refinement of the partition, if necessary, we may assume that there exist reals $0 = s_0, s_1, \ldots, s_n, \ldots$ increasing to $+\infty$ such that λ and σ are constant on $[s_j, s_{j+1})$, say $\lambda = \lambda_j(w)$, $\sigma = \sigma_j(w)$, where $\lambda_j(w)$ and $\sigma_j(w)$ are $\mathscr{F}_{s_j}$-measurable. In this case (*) is a direct consequence of the definition.

Step 2. Let λ be simple and bounded. Let (σ_n) be a sequence of simple bounded functions as in Lemma 3. Put

$$\xi_n(t,w) = \int_0^t \sigma_n(s,w)dY(s,w)$$

By Step 1,

$$(**) \qquad \int_0^t \lambda d\xi_n = \int_0^t \lambda\sigma_n dY(s,w).$$

90 Since λ is bounded, $\lambda\sigma_n$ converges to $\lambda\sigma$ in $L^2([0,t]\times\Omega)$. Hence, by definition, $\int_0^t \lambda\sigma_n dY(s,w)$ converges to $\int_0^t \lambda\sigma dY$ in probability.

Further,

$$\int_0^t \lambda d\xi_n(s,w) = \lambda(s_0,w)[\xi_n(s_1,w)-\xi_n(s_0,w)]+\cdots$$
$$+\cdots+\lambda(s_k,w)[\xi_n(t,w)-\xi_n(s_{k-1},w)],$$

where $s_0 < s_1 < \ldots\ldots$ is a partition for λ, and $\xi_n(t,w)$ converges to $\xi(t,w)$ in probability for every t. Therefore

$$\int_0^t \lambda d\xi_n(s,w)$$

converges in probability to

$$\int_0^t \lambda d\xi(s,w).$$

Taking limit as $n\to\infty$ in (**) we get

$$\int_0^t \lambda d\xi(s,w) = \int_0^t \lambda\sigma dY(s,w).$$

Step 3. Let λ be any progressively measurable function with

$$E(\int\limits_0^t \lambda^2 ds) < \infty, \ \forall t.$$

Let λ_n be a simple approximation to λ as in Lemma 3. Then, by Step 2,

$$(***) \qquad \int\limits_0^t \lambda_n(s,w)d\xi(s,w) = \int\limits_0^t \lambda_n(s,w)\sigma(s,w)dY(s,w).$$

By definition, the left side above converges to

$$\int\limits_0^t \lambda(s,w)d\xi(s,w)$$

in probability. As σ is bounded $\lambda_n\sigma$ converges to $\lambda\sigma$ in $L^2([0,t]\times\Omega)$. **91**
Therefore

$$P\left(\sup_{0\leq t\leq T} |\int\limits_0^t \lambda_n\sigma dY(s,w) - \int\limits_0^t \lambda\sigma \ dy(s,w)| \geq \epsilon\right)$$
$$||a||_\infty 1/\epsilon^2 E\left(\int\limits_0^t (\lambda_n\sigma - \lambda\sigma)^2 ds\right)$$

(see proof of the main theorem leading to the definition of the stochastic integral). Thus

$$\int\limits_0^t \lambda_n\sigma dY(s,w)$$

converges to

$$\int\limits_0^t \lambda\sigma \ dY(s,w)$$

in probability. Let n tend to + in (***) to conclude the proof.

□

13. Extension to Vector-Valued Itô Processes

Definition. Let $(\Omega, \mathscr{F}, P)$ be a probability space and $(\mathscr{F}_t)$ an increasing **92**
family of sub σ-algebras of $\mathscr{F}$. Suppose further that

(i) $$a : [0, \infty) \times \Omega \to S^d_+$$

is a probability measurable, bounded function taking values in the class of all symmetric positive semi-definite $d \times d$ matrices, with real entries;

(ii) $$b : [0, \infty) \times \Omega \to \mathbb{R}^d$$

is a progressively measurable, bounded function;

(iii) $$X : [0, \infty) \times \Omega \to \mathbb{R}^d$$

is progressively measurable, right continuous for every w and continuous almost everywhere (P);

(iv) $$Z(t, \cdot) = \exp[\langle \theta, X(t, \cdot)\rangle - \int\limits_0^t \langle \theta, b(s, \cdot)\rangle ds - \frac{1}{2} \int\limits_0^t \langle \theta,\ a(s, \cdot)\theta\rangle ds]$$

is a martingale for each $\theta \in \mathbb{R}^d$, where

$$\langle x, y\rangle = x_1 y_1 + \cdots + x_d y_d,\ x,\ y \in R^d.$$

Then X is called an Itô process corresponding to the parameters b and a, and we write $X \in I[b,a]$

Note. 1. $Z(t,w)$ is a real valued function.

2. b is progressively measurable if and only if each b_i is progressively measurable.

93 3. a is progressively measurable if and only if each a_{ij} is so.

Exercise 1. If $X \in I[b,a]$, then show that

(i) $$X_i \in I[b_i, a_{ii}],$$

(ii) $$Y = \sum_{i=1}^{d} \theta_i X_i \in I[\langle\theta, b\rangle, \langle\theta, a\theta\rangle],$$

where
$$\theta = (\theta_1, \ldots, \theta_d).$$

(Hint: (ii) (i). To prove (ii) appeal to the definition).

Remark. If X has a multivariate normal distribution with mean μ and covariance (ρ_{ij}), then $Y = \langle\theta, X\rangle$ has also a normal distribution with mean $\langle\theta, \mu\rangle$ and variance $\langle\theta, \rho\theta\rangle$. Note the analogy with the above exercise. This analogy explains why at times b is referred to as the "mean" and a as the "covariance".

Exercise 2. If $\{\beta(t) : t \geq 0\}$ is a d-dimensional Brownian motion, then $\beta(t,w) \in I[0,I]$ where $I = d \times d$ identity matrix.

As before one can show that $Y(t,\cdot) = X(t,\cdot) - \int_0^t b(s,w)ds$ is an Itô process with parameters 0 and a.

Definition. Let X be a d-dimensional Ito process. $\sigma = (\sigma_1, \ldots, \sigma_d)$ a d-dimensional progressively measurable function such that

$$E\left(\int_0^t \langle\sigma(s,\cdot), \sigma(s,\cdot)\rangle > ds\right)$$

is finite or, equivalently,

$$E\left(\int\limits_0^t \sigma_i^2(s,\cdot)ds\right) < \infty, \quad (i = 1, 2, \ldots d).$$

94

Then by definition

$$\int\limits_0^t \langle\sigma(s,\cdot), dX(s,\cdot)\rangle = \sum_{i=1}^d \int\limits_0^t \sigma_i(s,\cdot)dX_i(s,\cdot).$$

Proposition . *Let X be a d-dimensional Itô process $X \in I[b,a]$ and let σ be progressively measurable and bounded. If*

$$\xi_i(t,\cdot) = \int\limits_0^t \sigma_i dX_i(s,\cdot),$$

then

$$\xi = (\xi_1, \ldots, \xi_d) \in I[B, A],$$

where

$$B = (\sigma_1 b_1, \ldots, \sigma_d b_d) \quad \text{and} \quad A_{ij} = \sigma_i \sigma_j a_{ij}.$$

Proof. (i) Clearly A_{ij} is progressively measurable and bounded. Since $a \in S^d_+$, $A \in S^d_+$.

(ii) Again B is progressively measurable and bounded.

(iii) Since σ is bounded, each $\xi_i(t,\cdot)$ is an Itô process; hence ξ is progressively measurable, right continuous, continuous almost everywhere (P). It only remains to verify the martingale condition.

Step 1. Let $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$. By hypothesis,

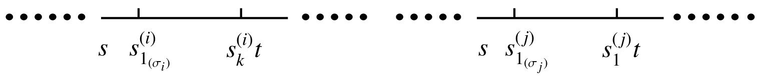
(*) $$E(\exp[(\theta_1 X_1 + \cdots + \theta_d X_d)|_s^t \cdots \int\limits_s^t (\theta_1 b_1 + \cdots + \theta_d b_d)du$$

$$- \frac{1}{2} \int_0^t \sum \theta_i \theta_j a_{ij} ds] | \mathscr{F}_s) = 1.$$

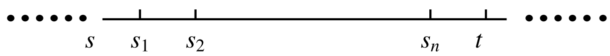
95 Assume that each σ_i is constant on $[s, t]$, $\sigma_i = \sigma_i(w)$ and $\mathscr{F}_s$-measurable. Then (*) remains true if θ_i are replaced by $\theta_i \theta_i(w)$ and since σ_i's are constant over $[s, t]$, we get

$$\begin{aligned} E(\exp[\int_0^t \sum_{i=1}^d \theta_i \sigma_i(s, \cdot) dX_i(s, \cdot) - \int_0^t \theta_i b_i \sigma_i(s, \cdot) ds \\ - \frac{1}{2} \int_0^t \sum \theta_i \theta_j \sigma_i(s, \cdot) \sigma_j(s, \cdot) a_{ij} ds] |_s) \\ \exp\left[\int_0^s \sum_{i=1}^d \theta_i \sigma_i(s, \cdot) dX_i(s, \cdot) - \int_0^s \langle \theta, B \rangle du - 1 \int_0^s \langle \theta, A\theta \rangle du \right]. \end{aligned}$$

Step 2. Let each σ_i be a simple function.



By considering finer partitions we may assume that each σ_i is a step function,



finest partition

i.e. there exist points $s_0, s_1, s_2, \ldots, s_n, s = s_0 < s_1 < \ldots < s_{n+1} = t$, such that on $[s_j, s_{j+1})$ each σ_i is a constant and ${}_{s_j}$-measurable. Then (**) holds if we use the fact that if $\mathscr{C}_1 \supset \mathscr{C}_2$.

$$E(E(f | \mathscr{C}_1) | \mathscr{C}_2) = E(f | \mathscr{C}_2).$$

Step 3. Let σ be bounded, $|\sigma| \leq C$. Let $(\sigma^{(n)})$ be a sequence of sim-
96 ple functions approximating σ as in Lemma 3. (**) is true if σ_i is replaced by $\sigma_i^{(n)}$ for each n. A simple verification shows that the expression $Z_n(t, \cdot)$, in the parenthes is on the left side of (**) with σ_i replaced by $\sigma_i^{(n)}$, converges to

$$Z(t, \cdot) =$$
$$= \operatorname{Exp}\Big(\int_0^t \sum_i \theta_i \sigma_i(s, \cdot) ds - \int_0^t \sum_i \theta_i b_i \sigma_i(s, \cdot) ds -$$
$$-\frac{1}{2}\int_0^t \sum_{i,j} \theta_i \theta_j \sigma_i \sigma_j a_{ij} ds\Big)$$

as $n \to \infty$ in probability. Since $Z_n(t, \cdot)$ is a martingale and the functions σ_i, σ_j, a are all bounded,

$$\sup_n E(Z_n(t, \cdot)) < \infty.$$

This proves that $Z(t, \cdot)$ is a martingale. □

Corollary . *With the hypothesis as in the above proposition define*

$$Z(t) = \int_0^t \langle \sigma(s, \cdot), dX(s, \cdot) \rangle.$$

Then

$$Z(t, \cdot) \in I[\langle \sigma, b \rangle, \sigma a \sigma^*]$$

where σ^* is the transpose of σ.

Proof. $Z(t, \cdot) = \xi_1(t, \cdot) + \cdots + \xi_d(t, \cdot)$. □

Definition. Let $\sigma(s, w) = (\sigma_{ij}(s, w))$ be a $n \times d$ matrix of progressively measurable functions with

$$E\left(\int_0^t \sigma_{ij}^2(s, \cdot) ds\right) < \infty.$$

If X is a d-dimensional Itô process, we define 97

$$\left(\int_0^t \sigma(s,\cdot)dX(s,\cdot)\right)_i = \sum_{j=1}^{d}\int_0^t \sigma_{ij}(s,\cdot)dX_j(s,\cdot).$$

Exercise 3. Let

$$Z(t,w) = \int_0^t \sigma(s,\cdot)dY(s,\cdot),$$

where $Y \in I[0,a]$ is a d-dimensional Itô process and σ is as in the above definition. Show that

$$Z(t,\cdot) \in I[0,\sigma a\sigma^*]$$

is an n-dimensional Ito process, (assume that σ is bounded).

Exercise 4. Verify that

$$E(|Z(t)|^2) = E\left(\int_0^t \operatorname{tr}(\sigma a\sigma^*)ds\right).$$

Exercise 5. Do exercise 3 with the assumption that $\sigma a\sigma^*$ is bounded.

Exercise 6. State and prove a change of variable formula for stochastic integrals in the case of several dimensions.
(Hint: For the proof, use the change of variable formula in the one dimensional case and $d(X+Y) = dX + dY$).

14. Brownian Motion as a Gaussian Process

SO FAR WE have been considering Brownian motion as a Markov process. We shall now show that Brownian motion can be considered as a Gaussian process. 98

Definition. Let $X \equiv (X_1, \ldots, X_N)$ be an N-dimensional random variable. It is called an N-variate normal (or Gaussian) distribution with mean $\mu \equiv (\mu_1, \ldots, \mu_N)$ and covariance A if the density function is

$$\frac{1}{(2\pi)^{N/2}} \frac{1}{(\det A)^{1/2}} \exp\left(-\frac{1}{2}[(X-\mu)A^{-1}(X-\mu)^*]\right)$$

where A is an $N \times N$ positive definite symmetric matrix.

Note. 1. $E(X_i) = \mu_i$.

2. $\mathrm{Cov}(X_i, X_j) = (A)_{ij}$.

Theorem . $X \equiv (X_1, \ldots, X_N)$ *is a multivariate normal distribution if and only if for every* $\theta \in \mathbb{R}^N$, $\langle \theta, X \rangle$ *is a one-dimensional Gaussian random variable.*

We omit the proof.

Definition. A stochastic process $\{X_t : t \in I\}$ is called a Gaussian process if $\forall t_1, t_2, \ldots, t_N \in I$, $(X_{t_1}, \ldots, X_{t_N})$ is an N-variate normal distribution.

Exercise 1. Let $\{X_t : t \geq 0\}$ be a one dimensional Brownian motion. Then show that

(a) X_t is a Gaussian process. 99

(Hint: Use the previous theorem and the fact that increments are independent)

(b) $E(X_t) = 0$, $\forall t$, $E(X(t)X(s)) = s \wedge t$.

Let $\rho : [0, 1] = [0, 1] \to \mathbb{R}$ be defined by

$$\rho(s, t) = s \wedge t.$$

Define $K : L^2_{\mathbb{R}}[0, 1] \to L^2_{\mathbb{R}}[0, 1]$ by

$$Kf(s) = \int\limits_0^1 \rho(s, t) f(t) dt.$$

Theorem . *K is a symmetric, compact operator. It has only a countable number of eigenvalues and has a complete set of eigenvectors.*

We omit the proof.

Exercise 2. Let λ be any eigenvalue of K and f an eigenvector belonging to λ. Show that

(a) $\lambda f'' + f = 0$ with $\lambda f(0) = 0 = \lambda f'(1)$.

(b) Using (a) deduce that the eigenvalues are given by $\lambda_n = 4/(2n + 1)^2\pi^2$ and the corresponding eigenvectors are given by

$$f_n = \sqrt{2}\,\mathrm{Sin}\, 1/2[(2n + 1)\pi t] n = 0, 1, 2, \ldots.$$

Let $Z_0, Z_1, \ldots, Z_n \ldots$ be identically distributed, independent, normal random variables with mean 0 and variance 1. Then we have

Proposition . $Y(t, w) = \sum\limits_{n=0}^{\infty} Z_n(w) f_n(t) \sqrt{\lambda_n}$
converges in mean for every real t.

100 *Proof.* Let $Y_m(t,w) = \sum\limits_{i=0}^{m} Z_i(w) f_i(t)\sqrt{\lambda_i}$. Therefore

$$E\{(Y_{n+m}(t,\cdot) - Y_n(t,\cdot))^2\} = \sum_{n+1}^{n+m} f_i^2(t)\lambda_i,$$

$$E(\|Y_{n+m}(\cdot) - Y_n(\cdot)\|^2 \leq \sum_{n+1}^{n+m} \lambda_i \to 0.$$

□

Remark. As each $Y_n(t,\cdot)$ is a normal random variable with mean 0 and variance $\sum\limits_{i=0}^{n} \lambda_i f_i^2(t)$, $Y(t,\cdot)$ is also a normal random variable with mean zero and variance $\sum\limits_{i=0}^{\infty} \lambda_i f_i^2$. To see this one need only observe that the limit of a sequence of normal random variables is a normal random variable.

Theorem (Mercer).

$$\rho(s,t) = \sum_{i=0}^{\infty} \lambda_i f_i(t) f_i(s), \ (s,t) \in [0,1] \times [0,1].$$

The convergence is uniform.

We omit the proof.

Exercise 3. Using Mercer's theorem show that $\{X_t : 0 \leq t \leq 1\}$ is a Brownian motion, where

$$X(t,w) = \sum_{n=0}^{\infty} Z_n(w) f_n(t)\sqrt{\lambda_n}.$$

This exercise now implies that

$$\int_0^1 X^2(s,w)ds = (L^2 - \text{norm of } X)^2$$

$$= \sum \lambda_n Z_n^2(w),$$

since $f_n(t)$ are orthonormal. Therefore **101**

$$E(e^{-\lambda \int\limits_0^1 X^2(s,\cdot)ds}) = E(e^{-\lambda \sum\limits_{n=0}^{\infty} \lambda_n Z_n^2(w)}) = \prod_{n=0}^{\infty} E(e^{-\lambda \lambda_n Z_n^2})$$

(by independence of Z_n)

$$= \prod_{n=0}^{\infty} E(e^{-\lambda \lambda_n Z_0^2})$$

as $Z_0, Z_n \ldots$ are identically distributed. Therefore

$$E(e^{-\lambda \int\limits_0^1 X^2(s,\cdot)ds}) = \prod_{n=0}^{\infty} 1/\surd(1 + 2\lambda\lambda_n)$$

$$= \prod_{n=0}^{\infty} 1/\surd\left(1 + \frac{8\ 8\lambda}{(2n+1)^2\Pi^2}\right)$$

$$= 1/\surd(\cosh)\surd(2\lambda).$$

APPLICATION. If $F(a) = P(\int\limits_0^1 X^2(s)ds < a)$, then

$$\int\limits_0^{\infty} e^{-\lambda a} dF(a) = \int\limits_{-\infty}^{\infty} e^{-\lambda a} dF(a)$$

$$= E(e^{-\lambda \int_0^1 X^2(s)ds}) = 1/\surd(\cosh)\surd(2\lambda).$$

15. Equivalent For of Itô Process

LET $(\Omega, \mathscr{F}, P)$ BE A probability space with $(\mathscr{F}_t)_{t\geq 0}$ and increasing family of sub σ-algebras of $\mathscr{F}$ such that $\sigma(\underset{t\geq 0}{U}\mathscr{F}_t) = \mathscr{F}$. Let 102

(i) $a : [0, \infty)\times\Omega \to S_d^+$ be a progressively measurable, bounded function taking values in S_d^+, the class of all $d\times d$ positive semidefinite matrices with real entries;

(ii) $b : [0, \infty) \times \Omega \to \mathbb{R}^d$ be a bounded, progressively measurable function;

(iii) $X : [0, \infty) \times \Omega \to \mathbb{R}^d$ be progressively measurable, right continuous and continuous a.s. $\forall (s, w) \in [0, \infty) \times \Omega$.

For $(s, w) \in [0, \infty) \times \Omega$ define the operator

$$L_{s,w} = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(s, w)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{d} b_j(s, w)\frac{\partial}{\partial x_j}.$$

For f, u, h belonging to $C_0^\infty(\mathbb{R}^d)$, $C_0^\infty([0, \infty) \times \mathbb{R}^d)$ and $C_b^{1,2}([0, \infty) \times \mathbb{R}^d)$ respectively we define $Y_f(t, w)$, $Z_u(t, w)$, $P_h(t, w)$ as follows:

$$Y_f(t, w) = f(X(t, w)) - \int_0^t (L_{s,w}(f)(X(s, w))ds,$$

$$Z_u(t,w) = u(t, X(t,w)) - \int_0^t \left(\frac{\partial u}{\partial s} + L_{s,w}u\right)(s, X(s,w))ds,$$

$$P_h(t,w) = \exp[h(t, X(t,w)) - \int_0^t \left(\frac{\partial h}{\partial s} + L_{s,w}h\right)(s, X(s,w)ds-$$

$$-\frac{1}{2}\int_0^t \langle a(s,w)\nabla_x h(s, X(s,w)), \nabla_x h(s, X(s,w))\rangle ds].$$

103 **Theorem .** *The following conditions are equivalent.*

(i) $X_\theta(t,w) = \exp[\langle\theta, X(t,w)\rangle - \int_0^t \langle\theta, b(s,w)\rangle ds - \int_0^t \langle\theta, a(s,w)\theta\rangle ds]$

is a martingale relative to $(\Omega, \mathscr{F}_t, P)$, $\forall\theta \in \mathbb{R}^d$.

(ii) $X_\lambda(t,w)$ *is a martingale* $\forall_\lambda$ *in* $\mathbb{R}^d$. *In particular* $X_{i\theta}(t,w)$ *is a martingale* $\forall\theta \in \mathbb{R}^d$.

(iii) $Y_f(t,w)$ *is a martingale for every* $f \in C_0^\infty(\mathbb{R}^d)$

(iv) $Z_u(t,w)$ *is a martingale for every* $u \in C_0^\infty([0,\infty) \times \mathbb{R}^d)$.

(v) $P_h(t,w)$ *is a martingale for every* $h \in C_b^{1,2}[(0,\infty) \times \mathbb{R}^d)$.

(vi) *The result (v) is true for functions* $h \in C^{1,2}([0,\infty) \times \mathbb{R}^d)$ *with linear growth, i.e. there exist constants A and B such that* $|h(x)| \leq A|x| + B$.

The functions $\dfrac{\partial h}{\partial t}$, $\dfrac{\partial h}{\partial x_i}$, and $-\dfrac{\partial^2 h}{\partial x_i \partial x_j}$ which occur under the integral sign in the exponent also grow linearly.

Remark. The above theorem enables one to replace the martingale condition in the definition of an Itô process by any of the six equivalent conditions given above.

Proof. (i) (ii). $X_\lambda(t,\cdot)$ is $\mathscr{F}_t$-measurable because it is progressively measurable. That $E(|X_\lambda(t,w)|) < \infty$ is a consequence of (i) and the fact that a is bounded.

The function $\lambda \xrightarrow{\phi} \dfrac{X_\lambda(t,w)}{X_\lambda(s,w)}$ is continuous for fixed t, s, w, $(t > s)$. Morera's theorem shows that ϕ is analytic. Let $A \in \mathscr{F}_s$. Then

$$\int\limits_A \frac{X_\lambda(t,w)}{X_\lambda(s,w)} dP(w)$$

is analytic. By hypothesis, 104

$$\int\limits_A \frac{X_\lambda(t,w)}{X_\lambda(s,w)} dP(w) = 1, \ \forall \lambda \in \mathbb{R}^d.$$

Thus $\int\limits_A \dfrac{X_\lambda(t,w)}{X_\lambda(s,w)} dP(w) = 1$, $\forall$ complex λ. Therefore

$$E(X_\lambda(t,w)|\mathscr{F}_s) = X_\lambda(s,w),$$

proving (ii). (ii) $\Rightarrow$ (iii). Let

$$A(t,w) = \exp\left[-i\int\limits_0^t \langle\theta, b(s,w)\rangle ds + \frac{1}{2}\int\limits_0^t \langle\theta, a(s,w)\theta\rangle ds\right], \theta \in \mathbb{R}^d.$$

By definition, A is progressively measurable and continuous. Also $|\dfrac{dA}{dt}(t,w)|$ is bounded on every compact set in $\mathbb{R}$ and the bound is independent of w. Therefore $A(t,w)$ is of bounded variation on every interval $[0,T]$ with the variation $\|A\|_{[0,T]}$ bounded uniformly in w. Let $M(t,w) = X_{i\theta}(t,w)$. Therefore

$$\sup_{0\le t\le T} |M(t,w)| \le e^{1/2\ T} \sup_{0\le t\le T} |\langle\theta, a\theta\rangle|.$$

By (ii) $M(t,\cdot)$ is a martingale and since

$$E\left(\sup_{0\le t\le T} |M(t,w)|\ \|A\|_{[0,T]}(w)\right) < \infty, \forall T,$$

$$M(t,\cdot)A(t,\cdot) - \frac{1}{2}\int\limits_0^t M(s,\cdot) dA(s,\cdot)$$

is a martingale (for a proof see Appendix), i.e. $Y_f(t, w)$ is a martingale when $f(x) = e^{i\langle\theta,x\rangle}$.

Let $f \in C_0^\infty(\mathbb{R}^d)$. Then $f \in \mathscr{F}(\mathbb{R}^d)$ the Schwartz-space. Therefore by the Fourier inversion theorem

$$f(x) = \int\limits_{\mathbb{R}^d} \hat{f}(\theta) e^{i\langle\theta,x\rangle} d\theta.$$

105 On simplification we get

$$Y_f(t, w) = \int\limits_{\mathbb{R}^d} \hat{f}(\theta) Y_\theta(t, w) d\theta$$

where $Y_\theta \equiv Y_e i\langle\theta, x\rangle$. Clearly $Y_f(t, \cdot)$ is progressively measurable and hence $\mathscr{F}_t$-measurable.

Using the fact that

$$E(|Y_\theta(t, w)|) \leq 1 + t\, d|\theta|\ ||b||_\infty + \frac{d^2}{2}|\theta|^2\ ||a||_\infty,$$

the fact that $\mathscr{F}(\mathbb{R}^d) \subset L^1(\mathbb{R}^d)$ and that $\mathscr{F}(\mathbb{R}^d)$ is closed under multiplication by polynomials, we get $E(|Y_f(t, w)|) < \infty$. An application of Fubini's theorem gives $E(Y_f(t, w)|\mathscr{F}_s) = Y_f(s, w)$, if $t > s$. This proves (iii).

(iii) $\Rightarrow$ (iv). Let $u \in C_0([0, \infty) \times \mathbb{R}^d)$.

Clearly $Z_u(t, \cdot)$ is progressively measurable. Since $Z_u(t, w)$ is bounded for every w, $E(|Z_u(t, w)|) < \infty$. Let $t > s$. Then

$$\begin{aligned}
&E(Z_u(t, w) - Z_u(s, w)|\mathscr{F}_s) = \\
&= E(u(t, X(t, w) - u(s, X(s, w)|\mathscr{F}_s) - E(\int\limits_s^t (\frac{\partial u}{\partial \sigma} + L_{\sigma,w} u)(\sigma, X(\sigma, w) d\sigma|\mathscr{F}_s) \\
&\quad = E(u(t, X(t, w) - u(t, X(s, w))|\mathscr{F}_s) + E(u(t, X(s, w) - u(s, X(s, w))|\mathscr{F}_s) - \\
&\quad - E(\int\limits_s^t (\frac{\partial u}{\partial \sigma} + L_\sigma u_w)(\sigma, X(\sigma, w)) d\sigma|\mathscr{F}_s)
\end{aligned}$$

$$
\begin{aligned}
&= E(\int\limits_s^t (L_{\sigma,w}u)(t, X(\sigma,w))d\sigma|\mathscr{F}_s) +\\
&+ E(\int\limits_s^t (\frac{\partial u}{\partial\sigma}(\sigma, X(s,w))d\sigma|\mathscr{F}_s) -\\
&- E(\int\limits_s^t (\frac{\partial u}{\partial\sigma} + L\sigma^u, w)(\sigma, X(\sigma,w))d\sigma|\mathscr{F}_s), \quad \text{by (iii)}\\
&= E(\int\limits_s^t [L_{\sigma,w}u(t, X(\sigma,w)) - L_{\sigma,w}u(\sigma, X(\sigma,w))]d\sigma|\mathscr{F}_s)\\
&+ E(\int\limits_s^t [\frac{\partial u}{\partial\sigma}(\sigma, X(s,w)) - \frac{\partial u}{\partial\sigma}(\sigma, X(\sigma,w))]d\sigma|\mathscr{F}_s)\\
&= E(\int\limits_s^t (L_{\sigma,w}u(t, X(\sigma,w)) - L_{\sigma,w}u(\sigma, X(\sigma,w))]d\sigma|\mathscr{F}_s)\\
&- E(\int\limits_s^t d\sigma \int\limits_s^\sigma L_{\rho,w}\frac{\partial u}{\partial\sigma}(\sigma, X(\rho,w))d\rho|\mathscr{F}_s)
\end{aligned}
$$

106

The last step follows from (iii) (the fact that $\sigma > s$ gives a minus sign).

$$
\begin{aligned}
&= E(\int\limits_0^t d\sigma \int\limits_\sigma^t \frac{\partial}{\partial\rho} L_{\sigma,w}u(\rho, X(\sigma,w))d\rho|\mathscr{F}_s)\\
&- E(\int\limits_s^t d\sigma \int\limits_s^\sigma L_{\rho,w}\frac{\partial u}{\partial\sigma}(\sigma, X(\rho,w))d\rho|\mathscr{F}_s)\\
&= 0
\end{aligned}
$$

(by Fubini). Therefore $Z_u(t,w)$ is a martingale.

Before proving (iv) $\Rightarrow$ (v) we show that (iv) is true if $u \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d$. Let $u \in C_b^{1,2}$.

(*) Assume that there exists a sequence $(u_n) \in C_0^\infty[[0,\infty)\times\mathbb{R}^d]$ such that
$$u_n \to u, \frac{\partial u_n}{\partial t} \to \frac{\partial u}{\partial t}, \frac{\partial u_n}{\partial x_i} \to \frac{\partial u}{\partial x_i}, \frac{\partial u_n}{\partial x_i \partial x_j} \to \frac{\partial^2 u}{\partial x_i \partial x_j}$$

107 uniformly on compact sets.

Then $Z_{u_n} \to Z_u$ pointwise and $\sup\limits_n(|Z_{u_n}(t,w)|) < \infty$.

Therefore Z_u is a martingale. Hence it is enough to justify (*).

For every $u \in C_b^{1,2}([0,\infty)\times\mathbb{R}^d)$ we construct a $u \in C_b^{1,2}((-\infty,\infty)\times\mathbb{R}^d) \equiv C_b^{1,2}(\mathbb{R}\times\mathbb{R}^d)$ as follows. Put

$$u(t,x) = \begin{cases} u(t,x), \text{ if } t \geq 0, \\ C_1 u(-t,x) + C_2 u(-\frac{t}{2},x), \text{ if } t < 0; \end{cases}$$

matching $\dfrac{\partial \tilde{u}}{\partial t}, \dfrac{\partial u}{\partial t}$ at $t = 0$ and $\hat{u}(t,x)$ and $u(t,x)$ at $t = 0$ and $\tilde{u}(t,x)$ and $u(t,x)$ at $t = 0$ yields the desired constants C_1 and C_2. In fact $C_1 = -3$, $C_2 = 4$. (*) will be proved if we obtain an approximating sequence for $\tilde{u}$. Let $S : \mathbb{R}$ be any C function such that if $|x| \leq 1$,

$$S(x) = \begin{cases} 1, & \text{if } |x| \leq 1, \\ 0, & \text{if } |x| \geq 2. \end{cases}$$

Let $S_n(x) = S\left(\dfrac{|x|^2}{n}\right)$ where $|x|^2 = x_1^2 + \cdots + x_{d+1}^2$. Pur $u_n = S_n\tilde{u}$. This satisfies (*).

(iv) $\Rightarrow$ (v). Let
$$h \in C_b^{1,2}([0,\infty)\times\mathbb{R}^d).$$

Put $u = \exp(h(t,x))$ in (iv) to conclude that

$$M(t,w) = e^{h(t,X(t,w))} - \int\limits_0^t e^{h(s,X(s,w))}\left[\frac{\partial h}{\partial s} + L_{s,w}h + \frac{1}{2}\langle\nabla_x h, a\nabla_x h\rangle ds\right]$$

is a martingale.

108 Put

$$A(t,w) = \exp -\left[\int_0^t \frac{\partial h}{\partial s}(s,w) + L_{s,w} - (s,w) + \frac{1}{2}\langle a(s,w)\nabla_x h, \nabla_x h\rangle ds\right].$$

$A((t,w))$ is progressively measurable, continuous everywhere and

$$\|A\|_{[0,T]}(w) \leq C_1 \in C_2 T$$

where C_1 and C_2 are constants. This follows from the fact that $|\frac{dA}{dt}|$ is uniformly bounded in w. Also $\sup\limits_{0\leq t\leq T} |M(t,w)|$ is uniformly bounded in w. Therefore

$$E(\sup_{0\leq t\leq T} |M(t,w)|\ \|A\|_{[0,T]}(w)) < \infty.$$

Hence $M(t,\cdot)A - \int\limits_0^t M(s,\cdot)dA(s,\cdot)$ is a martingale. Now

$$\frac{dA(s,w)}{A(s,w)} = -\left[\frac{\partial h}{\partial s}(s,w) + L_{s,w}h(s,w) + \frac{1}{2}\langle a\nabla_x h, \nabla_x h\rangle\right]$$

Therefore

$$M(t,w) = e^{h(t,X(t,w))} + \int_0^t e^{h(s,X(s,w))}\frac{dA(s,w)}{A(s,w)}.$$

$$M(t,w)A(t,w) = P_h(t,w) + A(t,w)\int_0^t e^{h(s,X(s,w))}\frac{dA(s,w)}{A(s,w)}$$

$$\int_0^t M(s,\cdot)dA(s,\cdot) = \int_0^t e^{h(s,X(s,w))}dA(s,w)$$

$$+\int_0^t dA(s,w)\int_0^s e^{h(\sigma,X(\sigma,w))}\frac{dA(\sigma,w)}{A(\sigma,w)}$$

Use Fubini's theorem to evaluate the second integral on the right above and conclude that $P_h(t,w)$ is a martingale.

(vi) $\Rightarrow$ (i) is clear if we take $h(t,x) = \langle\theta, x\rangle$. It only remains to prove that (v) $\Rightarrow$ (vi).

(v) $\Rightarrow$ (vi). The technique used to prove this is an important one and we shall have occasion to use it again.

109 **Step 1.** 0 Let $h(t,x) = \theta_1 x_1 + \theta_2 x_2 + \cdots \theta_d x_d = \langle\theta, x\rangle$ for every $(t,x) \in [0,\infty) \times \mathbb{R}^d$, θ is some fixed element of $\mathbb{R}^d$. Let

$$Z(t) = \exp\left[\langle\theta, X_t\rangle - \int_0^t \langle\theta, b\rangle ds - \frac{1}{2}\int_0^t \langle\theta, a\theta\rangle ds\right]$$

We claim that $Z(t,\cdot)$ is a supermartingale.

Let $f : \mathbb{R} \to \mathbb{R}$ be a C^∞ function with compact support such that $f(x) = x$ in $|x| \leq 1/2$ and $|f(x)| \leq 1$, $\forall x$. Put $f_n(x) = nf(x/n)$. Therefore $|f_n(x)| \leq C|x|$ for some C independent of n and x and $f_n(x)$ converges to x.

Let $h_n(x) = \sum\limits_{i=1}^{d} \theta_i f_n(x_i)$. Then $h_n(x)$ converges to $\langle\theta, x\rangle$ and $|h_n(x)| \leq C'|x|$ where C is also independent of n and x. By (v),

$$Z_n(t) = \exp\left[h_n(t, X_t) - \int_0^t \left(\frac{\partial h_n}{\partial s} + L_{s,w} h\right) ds - \frac{1}{2}\int_0^t \langle a\nabla_x h_n, \nabla_x h_n\rangle ds\right]$$

is a martingale. As $h_n(x)$ converges to $\langle\theta, x\rangle$, $Z_n(t,\cdot)$ converges to $Z(t,\cdot)$ pointwise. Consequently

$$E(Z(t)) = E(\underline{\lim}\, Z_n(t)) \leq \underline{\lim}\, E(Z_n(t)) = 1$$

and $Z(t)$ is a supermartingale.

Step 2. $E(\exp B \sup\limits_{0\leq s\leq t} |X(s,w)|) < \infty$ for each t and B. For, let $Y(w) = \sup\limits_{0\leq s\leq t} |X(s,w)|$, $Y_i(w) = \sup\limits_{0\leq s\leq t} |X_i(s,w)|$ where $X = (X_1, \ldots, X_d)$. Clearly $Y \leq Y_1 + \cdots + Y_d$. Therefore

$$E(e^{BY}) \leq E(e^{BY}1 e^{BY}2 \ldots e^{BY}d).$$

110 The right hand side above is finite provided $E(e^{BY}i) < \infty$ for each i as can be seen by the generalised Holder's inequality. Thus to prove the assertion it is enough to show $E(e^{BY}i) < \infty$ for each $i = 1, 2, \ldots d$ with aB' different from B; more specifically for B' bounded.

Put $\theta_2 = 0 = \theta_3 = \ldots = \theta_d$ in Step 1 to get

$$u(t) = \exp[\theta_1 X_1(t) - \int\limits_0^t \theta_1 b_1(s, \cdot) ds - \frac{1}{2}\theta_1^2 \int\limits_0^t a_{11}(s, \cdot) ds]$$

is a supermartingale. Therefore

$$P\left(\sup_{0 \le s \le t} u(s, \cdot) \ge \lambda\right) \le \frac{1}{\lambda} E(u(t)) = \frac{1}{\lambda}, \ \forall \lambda > 0.$$

(Refer section on Martingales). Let c be a common bound for both b_1 and a_{11} and let $\theta_1 > 0$. Then $(*)$ reads

$$P\left(\sup_{0 \le s \le t} \exp \theta_1 X_1(s) \ge \lambda \exp(\theta_1 ct + \frac{1}{2}\theta_1^2 ct)\right) \le \frac{1}{\lambda}.$$

Replacing λ by

$$e^{\lambda\theta_1} e^{-ct\theta_1 - 1/2ct\theta_1^2}$$

we get

$$P\left(\sup_{0 \le s \le t} \exp \theta_1 X_1(s) \ge \exp \lambda\theta_1\right) \le e^{-\lambda\theta_1 + \theta_1 ct + 1/2\theta_1^2 ct},$$

i.e.

$$P\left(\sup_{0 \le s \le t} X_1(s) \ge \lambda\right) \le e^{-\lambda\theta_1 + \theta_1 ct + 1/2\theta_1^2 ct}, \ \forall \theta_1 > 0.$$

Similarly

$$P\left(\sup_{0 \le s \le t} -X_1(s) \ge \lambda\right) \le e^{-\lambda\theta_1 + \theta_1 ct + 1/2\theta_1^2 tc}, \ \forall \theta_1 > 0.$$

As

$$\{Y_1(w) \ge \lambda\} \left\{\sup_{0 \le s \le t} X_1(s) \ge \lambda\right\} \cup \left\{\sup_{0 \le s \le t} -X_1(s) \ge \lambda\right\},$$

we get 111

$$P\{Y_1 \geq \lambda\} \leq 2e^{-\lambda\theta_1+\theta_1 ct+1/2\theta_1^2 ct}, \ \forall\theta_1 > 0.$$

Now we get

$$\begin{aligned} E(\exp BY_1) &= \frac{1}{B}\int\limits_0^\infty \exp(Bx)P(Y_1 \geq x)dx \quad (\text{since } Y_1 \geq 0) \\ &\leq \frac{2}{B}\int\limits_0^\infty \exp(Bx - x\theta_1 + \theta_1 ct + \frac{1}{2}\theta_1^2 ct)dx \\ &< \infty, \quad \text{if} \quad B < \theta_1 \end{aligned}$$

This completes the proof of step 2.

Step 3. $Z(t,w)$ is a martingale. For

$$|Z_n(t,w)| = Z_n(t,w)$$

$$= \exp\left[h_n(X_t) - \int\limits_0^t \left(\frac{\partial h_n}{\partial s} + L_{s,w}h_n\right)dx - \frac{1}{2}\int\limits_0^t \langle a\nabla_x h_n,\ \nabla_x h_n\rangle ds\right]$$

$$\leq \exp\left[h_n(X_t) - \int\limits_0^t L_{s,w}h_n\right]$$

(since a is positive semidefinite and $\partial h_n/\partial s = 0$).

Therefore $|Z_n(t,w)| \leq A\exp(B\sup\limits_{0\ s\ t}|X(s,w)|)$ (use the fact that $|h_n(s)| \leq C|x|$ and $\dfrac{\partial h_n}{\partial x_i}, \dfrac{\partial^2 h_n}{\partial x_i \partial x_j}$ are bounded by the same constant). The result now follows from the dominated convergence theorem and Step 2. ◻

Remark. In Steps 1, 2 and 3 we have proved that (v) ⇒ (i). The idea of
112 the proof was to express $Z(t,\cdot)$ as a limit of a sequence of martingales proving first that $Z(t,\cdot)$ is a supermartingale. Using the supermartingale inequality it was then shown that (Z_n) is a uniformly integrable family proving thereby that $Z(t,\cdot)$ is a martingale.

Step 4. Let $h(t,x) \in C^{1,2}([0,\infty)\times\mathbb{R}^d)$ such that $h(t,x)$, $\dfrac{\partial h}{\partial s}(t,x)$, $\dfrac{\partial h}{\partial x_i}(t,x)$, $\dfrac{\partial^2 h}{\partial x_i \partial x_j}(t,x)$ are all dominated by $\alpha|x|+\beta$ for some suitable scalars α and β. Let ϕ_n be a sequence of real valued C^∞ functions defined on $\mathbb{R}^d$ such that

$$\phi_n = \begin{cases} 1 \text{ on } |x| \leq n \\ 0 \text{ on } |x| \geq 2n \end{cases}$$

and suppose there exists a common bound C for

$$\phi_n, \frac{\partial \phi_n}{\partial x_i}, \frac{\partial^2 \phi_n}{\partial x_i \partial x_j} (\forall n).$$

Let $h_n(t,x) = h(t,x)\phi_n(x)$. By (v) $Z_{h_n}(t,w)$ is a martingale. The conditions on the function h and ϕ_n's show that

$$|Z_{h_n}(t,w)| \leq A \exp\left(B \sup_{0\leq s\leq t} |X(s,w)|\right)$$

where A and B are constants. By Step 2, (Z_{h_n}) are uniformly integrable. Also $Z_{h_n}(t,\cdot)$ converges pointwise to $P_h(t,\cdot)$ (since $h_n \to h$ pointwise). By the dominated convergence theorem $P_h(t,\cdot)$ is a martingale, proving (vi).

16. Itô's Formula

Motivation. Let $\beta(t)$ be a one-dimensional Brownian motion. We have 113
seen that the left integral

$$(*) \qquad L\left[2\int_0^t \beta(s,\cdot)d\beta\right] = [\beta^2(t,\cdot) - \beta^2(0,\cdot) - t]$$

Formally (*) can be written as

$$d\beta^2(t) = 2\beta(t)d\beta(t) + dt.$$

For, on integrating we recover (*).

Newtonian calculus gives the result:

$$d\,f(\beta(t)) = f'(\beta(t))d\beta(t) + \frac{1}{2}f''(\beta(t))d\beta^2(t) + \cdots$$

for reasonably smooth functions f and β. If β is of bounded variation, only the first term contributes something if we integrate the above equation. This is because $\sum d\beta^2 = 0$ for a function of bounded variation. For the Brownian motion we have seen that $\sum d\beta^2 \to$ a non zero value, but one can prove that $\sum d\beta^3, \ldots$ converge to 0. We therefore expect the following result to hold:

$$d\,f(\beta(t)) \approx f'(\beta(t))d\beta(t) + \frac{1}{2}f''(\beta(t))d^2\beta(t).$$

We show that for a one-dimensional Brownian motion

$$\sum(d\beta)^3, \quad \sup(d\beta)^4, \ldots$$

all vanish.

$a = t_0$ t_1 t_2 t_3 $b = t_{n+1}$

$$E(\sum(d\beta)^3) = E\left(\sum_{i=0}^{n}[\beta(t_{i+1}) - \beta(t_i)]^3\right) = \sum_{i=0} E[(\beta(t_{i+1}) - \beta(t_i)]^3$$
$$= \sum_{i=1}^{n} 0 = 0,$$

114 because $\beta(t_{i+1}) - \beta(t_i)$ is a normal random variable with mean zero and variance $t_{i+1} - t_i$. Similarly the higher odd moments vanish. Even moments of a normal random variable with mean 0 and variance σ^2 are connected by the formula

$$\mu_{2k+2} = \sigma^2(2k+1)\mu_{2k}, k > 1.$$

So

$$\sum(d\beta)^4 = \sum_{i=0}^{n}(\beta(t_{i+1}) - \beta(t_i))^4.$$

Therefore

$$E(\sum(d\beta)^4) = \sum_{i=0}^{n} E([\beta(t_{i+1}) - \beta(t_i))^4]$$
$$= 3\sum_{i=0}^{n}(t_{i+1} - t_i)^2;$$

the right hand side converges to 0 as the mesh of the partition goes to 0. Similarly the higher order even moments vanish.

More generally, if $\beta(t, \cdot)$ is a d-dimensional Brownian motion then we expect

$$df(\beta(t)) \approx \nabla f(\beta(t)) \cdot d\beta(t) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j} d\beta_i d\beta_j.$$

However $\sum d\beta_i d\beta_j = 0$ if $i \neq j$ (see exercise below). Therefore

$$df(\beta(t)) \approx \nabla f(\beta(t)) \cdot d\beta(t) + \frac{1}{2}\Delta f(\beta(t)) d\beta^2(t).$$

The appearance of Δ on the right side above is related to the heat equation.

Exercise 4. Check that $d\beta_i d\beta_j = \delta_{ij} dt$. **115**
(Hint: For $i = j$, the result was proved earlier. If $i \neq j$, consider a partition $0 = t_0 < t_1 < \ldots < t_n = t$. Let $\Delta_k \beta_i = \beta_i(t_k) - \beta_i(t_{k-1})$. Then

$$E\left(\sum_{k=1}^{n} \Delta_k \beta_i \Delta_k \beta_j\right)^2 = \sum_k E(\Delta_k^2 \beta_i \Delta_k^2 \beta_j) + 2 \sum_{k \neq l} E[(\Delta_k \beta_i)(\Delta_1 \beta_j)];$$

the right side converges to 0 as $n \to \infty$ because $\Delta_k \beta_i$ and $\Delta_\ell \beta_j$ are independent for $k \neq \ell$).

Before stating Itô's formula, we prove a few preliminary results.

Lemma 1. *Let $X(t, \cdot) \in I[b, 0]$ be a one-dimensional Itô process. Then*

$$X(t, \cdot) - X(0, \cdot) = \int_0^t b(s, \cdot) ds \quad a.e.$$

Proof. $\exp[\theta X(t, \cdot) - \theta X(0, \cdot) - \theta \int_0^t b(s, \cdot) ds]$ is a martingale for each θ. Therefore

$$E(\exp[\theta(X(t, \cdot) - X(0, \cdot)) - \theta \int_0^t b(s, \cdot) ds]) = \text{ constant } = 1, \ \forall t.$$

Let

$$W(t, \cdot) = X(t, \cdot) - X(0, \cdot) - \int_0^t b(s, \cdot) ds.$$

Then

$$E(\exp \theta W(t, \cdot)) = \text{ Moment generating function of } w = 1, \ \forall t.$$

Therefore $\theta(t, \cdot) = 0$ a.e. □

116 **Remark.** If $X(t,\cdot) \in I[0,0]$ then $X(t,\cdot) = X(0,\cdot)$ a.e.; i.e. $X(t,\cdot)$ is a trivial process.

We now state a theorem, which is a particular case of the theorem on page 103.

Theorem. *If $h \in C^{1,2}([0,\infty) \times \mathbb{R}^d)$ such that* (i) $|h(x)| \leq A|x| + B$, $\forall x \in [0,\infty) \times \mathbb{R}^d$, *for constants A and B* (ii) $\dfrac{\partial h}{\partial t}, \dfrac{\partial h}{\partial x_i}, \dfrac{\partial^2 h}{\partial x_i \partial x_j}$ *also grow linearly, then*

$$\exp[h(t,\beta(t,\cdot) - \int\limits_0^t \left(\frac{\partial h}{\partial s} + \frac{1}{2}\Delta h\right)(s,\beta(s,\cdot) - \frac{1}{2}\int\limits_0^t |\nabla h|^2(s,\beta(s,\cdot))ds]$$

is a martingale.

Ito's Formula. Let $f \in C_0^{1,2}([0,\infty)\times\mathbb{R}^d)$ and let $\beta(t,\cdot)$ be a d-dimensional Brownian motion. Then

$$\begin{aligned} f(t,\beta(t)) - f(0,\beta(0)) &= \int\limits_0^t \frac{\partial f}{\partial s}(s,\beta(s,\cdot))ds + \\ &+ \int\limits_0^t \langle \nabla f(s,\beta(s,\cdot)), d\beta(s,\cdot)\rangle + \frac{1}{2}\int\limits_0^t \Delta f(s,\beta(s,\cdot))ds. \end{aligned}$$

where

$$\Delta \equiv \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_d^2}.$$

Proof.

Step 1. Consider a $(d+1)$-dimensional process defined by

$$\begin{aligned} X_0(t,\cdot) &= f(t,\beta(t,\cdot)), \\ X_j(t,\cdot) &= \beta_j(t,\cdot). \end{aligned}$$

We claim that $X(t,\cdot) \equiv (X_0, X_1, \ldots, X_d)$ is a $(d+1)$-dimensional Itô-process with parameters

$$b = \left[\left(\frac{\partial f}{\partial s} + \frac{1}{2}\Delta f\right)(s, \beta(s,\cdot)), 0, 0, \ldots 0\right] \quad d \text{ terms}$$

and **117**

$$a = \begin{bmatrix} a_{00} & a_{01} & \ldots & a_{0d} \\ a_{10} & & & \\ \ddots & & & I_{d\times d} \\ a_{d0} & & & \end{bmatrix}$$

where

$$\begin{aligned} a_{00} &= |\nabla_x f|^2(s, \beta(s,\cdot)), \\ a_{0j} &= \left(\frac{\partial}{\partial x_j} f\right)(s, \beta(s,\cdot)), \\ a_{j0} &= a_{0j}. \end{aligned}$$

For, put $h = \lambda f(t,x) + \langle \theta, x\rangle$, $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, in the previous theorem. Then

$$\frac{\partial h}{\partial s} = \lambda \frac{\partial f}{\partial s}, \Delta h = \lambda \Delta f, \frac{\partial h}{\partial x_j} = \lambda \frac{\partial f}{\partial x_j} + \theta_j.$$

Therefore we seen that

$$\exp[\lambda f(t, \beta(t,\cdot)) + \langle \theta, \beta(t,\cdot)\rangle - \lambda \int_0^t \left(\frac{\partial f}{\partial s} + \frac{1}{2}\Delta_x f\right)(s, \beta(s,\cdot) ds$$
$$- \frac{1}{2}\lambda^2 \int_0^t |\nabla f|^2(s, \beta(s,\cdot)) ds - \frac{1}{2}|\theta|^2 t - \lambda \langle \theta, \int_0^t \nabla(f(s, \beta(s,\cdot))) ds\rangle]$$

is a martingale.

Consider $(\lambda, \theta) a \binom{\lambda}{\theta}$. We have

$$a \begin{bmatrix} \lambda \\ \theta \end{bmatrix} = \begin{bmatrix} a_{00}\lambda + \sum_{j=1}^d a_{0j}\theta_j \\ \rho + \theta \end{bmatrix}, \quad \rho = \lambda \begin{bmatrix} a_{10} \\ \ddots \\ a_{d0} \end{bmatrix}.$$

Therefore

$$(\lambda,\theta)a\begin{bmatrix}\lambda\\ \theta\end{bmatrix} = a_{00}\lambda^2 + \lambda\sum_{j=1}^{d} a_{0j}\theta_j + \sum_{j=1}^{d} a_{0j}\theta_j +$$
$$= \lambda^2|\nabla f|^2 + 2\lambda\frac{\partial f}{\partial x_j}\theta_j + |\theta|^2.$$

118

Thus (*) reads

$$\exp[\lambda f(t,\beta(t,\cdot) + \langle\theta,\beta(t,\cdot)\rangle - \lambda\int_0^t b_0(s,\cdot)ds - \frac{1}{2}\int_0^t \langle\alpha, a\alpha\rangle ds]$$

is a martingale where $\alpha = (\lambda,\theta) \in \mathbb{R}^{d+1}$. This proves the claim made above.

Step 2. Derine $\sigma(s,\cdot) = (1, -\nabla_x f(s,\beta(s,\cdot)))$ and let

$$Z(t,\cdot) = \int_0^t \langle\sigma(s,\cdot), dX(s,\cdot)\rangle \text{ where } X \approx (X_0, X_1, \ldots, X_d)$$

is the $(d+1)$-dimensional Itô process obtained in Step 1. Since $f \in C_b^{1,2}$, $Z(t,\cdot)$ is an Itô process with parameters $\langle\sigma, b\rangle$ and $\sigma a\sigma^*$:

$$\langle\sigma, b\rangle = \frac{\partial f}{\partial s} + \frac{1}{2}\Delta f,$$
$$a\sigma^* = \begin{bmatrix} a_{00} & \rho \\ \rho^* & I\end{bmatrix}\begin{bmatrix}1\\ -\nabla_x f\end{bmatrix}$$
$$= \begin{bmatrix} a_{00} - \langle\rho, \nabla_x f\rangle \\ \rho^* - \nabla_x f\end{bmatrix} = \begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}$$

Therefore $\sigma a\sigma^* = 0$. Hence by Lemma 1,

$$Z(t,\cdot) - Z(0,\cdot) - \int_0^t \langle\sigma, b\rangle ds = 0 \text{(a.e.)},$$

$$Z(t,\cdot) = \int_0^t dX_0(s) - \int_0^t \langle \nabla f(s,\beta(s,\cdot)), d\beta(s,\cdot)\rangle$$

$$= f(t,\beta(t)) - f(0,\beta(0)) - \int_0^t \langle \nabla f(f,\beta(s,\cdot)), d\beta(s,\cdot)\rangle$$

119

Hence $Z(0) = 0$. Thus

$$f(t,\beta(t)) - f(0,\beta(0)) - \int_0^t \langle \nabla f(s,\beta(s,\cdot)) d\beta(s,\cdot)\rangle -$$

$$- \int_0^t \left(\frac{\partial f}{\partial s} + \frac{1}{2}\Delta_x f\right)(s,\beta(s,\cdot))ds = 0 \text{ a.e.}$$

This estabilished Itô's formula.

□

Exercise. (Itô's formula for the general case). Let

$$\phi(t,x) \in C_b^{1,2}([0,\infty) \times \mathbb{R}^d).$$

If $X(t,\cdot)$ is a d-dimensional Itô process corresponding to the parameters b and a, then the following formula holds:

$$\phi(t, X(t,w)) - \phi(0, X(0,w))$$

$$= \int_0^t \frac{\partial \phi}{\partial s}(s, X(s,x))ds + \int_0^t \langle \nabla_x \phi, dX\rangle + \frac{1}{2}\int_0^t \sum a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} ds.$$

This is also written as

$$d\phi(s, X(s,w)) = \phi_s ds + \langle \nabla_x \phi, dX\rangle + \frac{1}{2}\sum a_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j} dx.$$

To prove this formula proceed as follows.

(i) Take $h(t,x) = \lambda\phi(t,x) + \langle\theta, x\rangle$ in (vi) of the theorem on the equivalence of Itô process to conclude that

$$Y(t,\cdot) = (\phi(t, X(t,\cdot)), X(t,\cdot))$$

is a $(d+1)$-dimensional Ito process with parameters

$$b' = \left(\frac{\partial\phi}{\partial t} + L_{s,w}\phi, b\right)$$

120 and

$$A = \left|\begin{array}{cc} \langle a\nabla_x\phi, \nabla_x\phi\rangle, & a\nabla_x\phi \\ 1\times 1 & 1\times d \\ a\nabla_x\phi & a \\ d\times 1 & d\times d \end{array}\right|$$

(ii) Let $\sigma(t,x) = (1, -\nabla_x\phi(t,x))$ and

$$Z(t,\cdot) = \int_0^t \langle\sigma(s, X(s,\cdot)),\ dY(s,\cdot)\rangle.$$

The assumptions on ϕ imply that Z is an Itô process corresponding to

$$(\langle\sigma, b'\rangle, \sigma A\sigma^*) \equiv \left(\frac{\partial\phi}{\partial t} + \frac{1}{2}\sum a_{ij}\frac{\partial^2\phi}{\partial x_i\partial x_j}, 0\right).$$

(iii) Use (ii) to conclude that

$$Z(t,\cdot) = \int_0^t \langle\sigma, b'\rangle ds \quad \text{a.e.}$$

This is Itô's formula.

(iv) (Exercise) Verify that Itô's formula agrees with the formula obtained for the case of Brownian motion.

Note. Observe that Itô's formula does not depend on b.

Examples. 1. Let $\beta(t)$ be a one-dimensional Brownian motion. Then

$$\begin{aligned}
d(e^t\phi(\beta(t)) &= e^t d\phi(\beta(t)) + \phi(\beta(t))d(e^t)\\
&= e^t\phi'(\beta(t))d\beta(t) + \phi(\beta(t))e^t dt +\\
&\quad + \frac{1}{2}\phi''(\beta(t))e^t dt.
\end{aligned}$$

2. To evaluate $d(e^{\int_e^t V(\beta(s,\cdot))ds}u(t,\beta(t)))$ where V is a smooth function, **121**
put

$$X_2(t,\cdot) = \int_0^t V(\beta(s,\cdot))ds,\ b = \begin{bmatrix} 0 \\ V((t,\cdot)) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$a = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad \text{Let } X_1(t,\cdot) = \beta(t,\cdot), X = (X_1, X_2).$$

Then

$$\begin{aligned}
&\exp[\theta_1 X_1(t,\cdot) - \theta_2 X_2(t,\cdot) - \theta_1 \int_0^t b_1(X(s,\cdot))ds\\
&- \theta_2 \int_0^t b_2(X(s,\cdot))ds - \frac{1}{2}\int_0^t \langle a\theta,\theta\rangle ds]\\
&\exp[\theta_1 X_1(t,\cdot) - \frac{\theta_1^2}{2}t];
\end{aligned}$$

the right side is a martingale. Therefore (X_1, X_2) is a 2-dimensional Itô process with parameters b and a and one can use Itô's formula to write

$$\begin{aligned}
d\left(e^{\int_0^t V(\beta(s,\cdot))ds}u(t,\beta(t))\right) &= d\left(e^{X_2(t)}u(t,\beta(t))\right)\\
&= e^{\int_0^t V(\beta(s,\cdot))}\frac{\partial}{\partial t}u(t,\beta(t))dt +\\
+e^{\int_0^t V(\beta(s,\cdot))dt}\frac{\partial}{\partial x}u(t,\beta(t))d(t) &+ e^{\int_0^t V(\beta(s,\cdot))ds}u(t,\beta(t))dX_2
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}e^{\int_0^t V(\beta(s,\cdot))ds}\frac{\partial}{\partial t}u(t,\beta(t))dt\\
&= e^{\int_0^t V(\beta(s,\cdot))dt}\frac{\partial}{\partial t}u(t,\beta(t))dt + \frac{\partial}{\partial t}u(t,\beta(s))d(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}u(t,\beta(t))dt\\
&+u(t,\beta(t))V(\beta(t,\cdot)dt].
\end{aligned}
$$

122 3. Let σ_i, f_i, $i = 1, 2, \ldots k$ be bounded and progressively measurable relative to a one-dimensional Brownian motion $(\Omega, \mathscr{F}_t, P)$. Write

$$X_i(t,\cdot) = \int_0^t \sigma_i(s,\cdot)d\beta(s,\cdot) + \int_0^t f_i(s,\cdot)ds.$$

Then $X_i(t,\cdot)$ is an Itô process with parameters $(\int_0^t f_i(s,\cdot)ds, \sigma_i^2)$ and $(X_1, \ldots, X_k)$ is an Itô process with parameters

$$B = \left(\int_0^t f_1(s,\cdot)ds, \ldots, \int_0^t f_k(s,\cdot)ds\right),$$

$$A = (A_{ij}) \text{ where } A_{ij} = \sigma_i\sigma_j\delta_{ij}.$$

If $\phi \equiv \phi(t, X_1(t)\ldots, X_k(t))$, then by Itô's formula

$$
\begin{aligned}
d\phi &= \frac{\partial\phi}{\partial s}ds + \frac{\partial\phi}{\partial x_1}dX_1 + \cdots + \frac{\partial\phi}{\partial x_k}dX_k\\
&+ \frac{1}{2}\sum \sigma_i\sigma_j\delta_{ij}\frac{\partial^2\phi}{\partial x_i\partial x_j}ds\\
&= \frac{\partial\phi}{\partial s}ds + \frac{\partial\phi}{\partial x_1}dX_1 + \cdots + \frac{\partial\phi}{\partial x_k}dX_k\\
&+ \frac{1}{2}\sum_{i=1}^k \sigma_i^2\frac{\partial\phi}{\partial x_i}ds
\end{aligned}
$$

Exercise. Take $\sigma_1 = 1$, $\sigma_2 = 0$, $f_1 = f_2 = 0$ above and verify that if

$$\phi = e^{X_2(t,\cdot)}u(t,\beta(t)),$$

then one gets the result obtained in Example 2 above.

We give below a set of rules which can be used to calculate $d\phi$ in practice, where ϕ is as in Example 3 above.

1. With each $d\beta$ associate a term $\sqrt{(dt)}$
2. If $\phi = \phi(t, X_1, \ldots, X_k)$, formally differentiate ϕ using ordinary calculus retaining terms upto the second order to get **123**

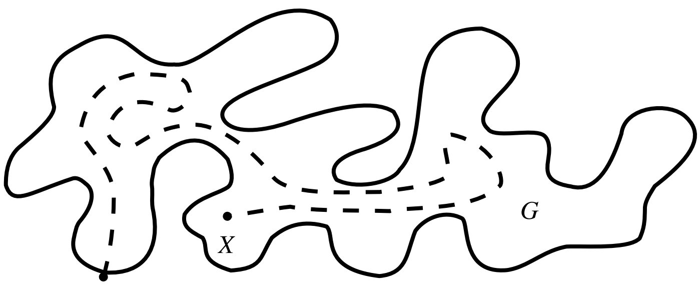
   $$(*) \quad d\phi = \frac{\partial\phi}{\partial t}dt + \frac{\partial\phi}{\partial x_1}dX_1 + \cdots + \frac{\partial\phi}{\partial x_k}X_k + \frac{1}{2}\sum \frac{\partial^2\phi}{\partial x_i \partial x_j}dX_i dX_j$$

3. Formally write $dX_i = f_i dt + \alpha_i d\beta_i$, $dX_j = \sigma_j d\beta_j + f_j dt$.
4. Multiply $dX_i dX_j$ and retain only the first order term in dt., For $d\beta_i d\beta_j$ substitute $\delta_{ij}dt$. Substitute in (*) to get the desired formula.

Illustration of the use of Itô Calculus. We refer the reader to the section on Dirichlet problem. There it was shown that

$$u(x) = \int\limits_G u(y)\pi(x, dy) = E(u(X(\tau)))$$

satisfies $\Delta u = 0$ in a region G with $u = u(X(\tau))$ on the boundry of G (here τ is the first hitting time).



The form of the solution is given directly by Itô's formula (without having recourse to the mean value property). If $u = u(X(t))$ satisfies $\Delta u = 0$ then by Itô's formula

$$du(X(t)) = \langle \nabla u, dX \rangle.$$

Therefore

$$u(X(t)) = u(X(0)) + \int_0^t \langle \nabla u(X(s)), dX(s) \rangle$$

124

Assuming ∇u to be bounded, we see that $u(X(t))$ is a martingale. By the optional stopping theorem

$$E(u(X(\tau)) = u(X(0)) = u(x).$$

Thus Itô's formula connects solutions of certain differential equations with the hitting probabilities.

17. Solution of Poisson's Equations

LET $X(t,\cdot)$ BE A d-dimensional Brownian motion with $(\Omega, \mathscr{F}_t, P)$ as usual. Let $u(x) : \mathbb{R}^d \to \mathbb{R}$ be such that $\frac{1}{2}\Delta u = f$. Assume $u \in C_b^2(\mathbb{R}^d)$. Then $L_{s,w} u \equiv \frac{1}{2}\Delta u = f$ and we know that $u(X(t,\cdot)) - \int\limits_0^t f(X(s,\cdot))ds$ is a $(\Omega, \mathscr{F}_t, P)$-martingale. Suppose now that $u(x)$ is defined only on an open subset $G \subset \mathbb{R}^d$ and $\frac{1}{2}\Delta u = f$ on G. We would like to consider 125

$$Z(t,\cdot) = u(X(t,\cdot)) - \int\limits_0^t f(X(s,\cdot))ds$$

and ask whether $Z(t,\cdot)$ is still a martingale relative to $(\Omega, \mathscr{F}_t, P)$. Let $\tau(w) = \inf\{t, X(t,w) \in \partial G\}$. Put this way, the question is not well-posed because $Z(t,\cdot)$ is defined only upto time $\tau(w)$ for u is not defined outside G. Even if at a time $t > \tau(w) X(t,w) \in G$, one needs to know the values of f for $t > \tau(w)$ to compute the integral.

To answer the question we therefore proceed as follows. Let $A_t = [w : \tau(w) > t]$. As t increases, A_t have decreasing measures. We shall give a meaning to the statement '$Z(t,\cdot)$ is a martingale on A_t'. Define

$$\overline{Z}(t,\cdot) = u(X(\tau \wedge t,\cdot)) - \int\limits_0^{\tau \wedge t} f(X(s,\cdot))ds.$$

Therefore

$$\overline{Z}(t,\cdot)=\begin{cases} Z(t), & \text{on } A_t \\ Z(\tau,\cdot), & \text{on } (A_t)^c. \end{cases}$$

126 Since $Z(t,\cdot)$ is progressively measurable upto time τ, $\overline{Z}(t,\cdot)$ is $\mathscr{F}_t$-measurable.

Theorem . *$\overline{Z}(t,\cdot)$ is a martingale.*

Proof. Let G_n be a sequence of compact sets increasing to G such that $G_n \subset G^0_{n+1}$. Choose a C^∞ function ϕ_n such that $\phi_n = 1$ on G_n and support $\phi_n \subset G$. Put $u_n = \phi_n u$ and $f_n = \dfrac{1}{2}\Delta u_n$. Then

$$Z_n(t,\cdot)=u_n(X(t,\cdot))-\int\limits_0^t f_n(X(s,\cdot))ds$$

is a martingale for each n. Put

$$\tau_n=\inf\{t: X(t,\cdot)\notin G_n\}$$

Then $Z_n(\tau_n\wedge t,\cdot)$ is also a martingale (See exercise below). But

$$Z_n(\tau_n\wedge t)=Z(\tau_n\wedge t).$$

Therefore $M_n(t,\cdot)=Z(\tau_n\wedge t,\cdot)$ is a martingale. Observe that $\tau_n\leq\tau_{n+1}$ and since $G_n\uparrow G$ we have $\tau_n\uparrow\tau$. Therefore $Z(\tau_n\wedge t)\to Z(\tau\wedge t)$ (by continuity); also $|M_n(t,\cdot)|\leq\|u\|_\infty+\|f\|_\infty t$. Therefore $Z(\tau\wedge t)=\overline{Z}(t,\cdot)$ is a martingale. ◻

Exercise. If $M(t,\cdot)$ is a $(\Omega,\mathscr{F}_t,P)$-martingale, show that for my stopping time τ, $M(\tau\wedge t,\cdot)$ is also a martingale relative to $(\mathscr{F}_t)$.
[Hint: One has to show that if $t_2>t_1$,

$$\int\limits_A M(\tau\wedge t_2,w)dP(w)=\int\limits_A M(\tau\wedge t_1,w)dP(w), \forall A\in\mathscr{F}_{t_1}.$$

The right side =

$$\int\limits_{A\cap(\tau>t_1)} M(t_1, w)dP(w) + \int\limits_{A\cap(\tau<t_1)} M(\tau, w)dP(w).$$

127 The left side

$$= \int\limits_{A\cap(\tau>t_1)} M(t_2, w)dP(w) + \int\limits_{A\cap(\tau<t_1)} M(\tau, w)dP(w).$$

Now use optional stopping theorem].

Lemma . *Let G be a bounded region and τ be as above. Then $E_x(\tau) < \infty$, $\forall x \in G$, where $E_x = E^{P_x}$.*

Proof. Without loss of generality we assume that G is a sphere of radius R. The function $u(x) = \dfrac{R^2 - |x|^2}{d} \geq 0$ and satisfies $\dfrac{1}{2}\Delta u = -1$ in G. By the previous theorem

$$u(X(\tau \wedge t, \cdot) + \int\limits_0^{\tau\wedge t} ds$$

is a martingale. Therefore

$$E_x(u(X(\tau \wedge t, \cdot))) + E_x(\tau \wedge t) = u(X(0)) = u(x).$$

Therefore $E_x(\tau \wedge t) \leq u(x)$ (since $u \geq 0$). By Fatou's lemma, on letting $t \to \infty$, we obtain $E_x(\tau) \leq u(x) < \infty$. Thus the mere existence of a u satisfying $\dfrac{1}{2}\Delta u = 1$ helps in concluding that $E_x(\tau) < \infty$. □

Theorem . *Let $u \in C_b^2(G)$ and suppose that u satisfies*

$$(*) \qquad \frac{1}{2}\Delta u = f \text{ in } G,$$
$$u = g \text{ on } \partial G.$$

Then $u(x) = E_x[g] - E_x[\int\limits_0^\tau f(X(s, \cdot))ds]$ solves (*).

Remark. The first part of the solutin $u(x)$ is the solution of the homogeneous equation, and the second part accounts for the inhomogeneous term.

128 *Proof.* Define $\overline{Z}(t, \cdot) = u(X(\tau \wedge t)) - \int\limits_0^{\tau\wedge t} f(X(s, \cdot))ds$. Then $\overline{Z}$ is a martingale. Also $|\overline{Z}| \leq ||u||_\infty + \tau||f||_\infty$. Therefore, by the previous Lemma, $\overline{Z}(t, \cdot)$ is a uniformly integrable martingale. Therefore we can equate the expectations at time $t = 0$ and at time $t = \infty$ to get

$$u(x) = E_x(g) - E_x[\int\limits_0^\tau f(X(s, \cdot))ds].$$

□

18. The Feynman-Kac Formula

WE NOW CONSIDER the modified heat equation 129

$$(*) \qquad \frac{\partial u}{\partial t} + \frac{1}{2}\Delta u + V(x)u(t,x) = 0, \quad 0 \leq t \leq T,$$

where $u(T,x) = f(x)$. The Feynman-Kac formula says that the solution for $s \leq T$ is given by

$$(**) \qquad u(s,x) = E_{s,x}(e^{\int_s^T V(X(s))dS} f(X(T))).$$

Observe that the solution at time s depends on the expectation with respect to the process starting at time s.

Note. (**) is to be understood in the following sense. If (*) admits a smooth solution then it must be given by (**). We shall not go into the conditions under which the solution exists. Let

$$Z(t,\cdot) = u(t, X(t,\cdot))e^{\int_s^t V(X(\sigma,\cdot)d\sigma)}, \quad t \geq s.$$

By Ito's formula (see Example 2 of section 16), we get

$$Z(t,\cdot) = Z(s,\cdot) + \int_s^t e^{\int_s^t V(X(\sigma,\cdot)d\sigma)} \langle \nabla u(\lambda, X(\lambda)), dX(\lambda) \rangle,$$

provided that u satisfies (*). Assume tentatively that V and ∇u are bounded and progressively measurable. Then $Z(t,\cdot)$ is a martingale. Therefore

$$E_{s,x}(Z(T,\cdot)) = E_{z,x}(Z(s,\cdot)),$$

or
$$E_{s,x}(u(T, X(T)))e^{\int_s^T V(X(\sigma,\cdot))d\sigma} = u(s, x).$$

This proves the result.

130 We shall now remove the condition ∇_u is bounded and prove the uniqueness of the solution corresponding to (*) under the assumption that V is bounded above and

$$|u(t, x)| \leq e^{A+|x|^\alpha}, \quad \alpha < 2, \quad \text{on} \quad [s, T).$$

In particular, the Feynman-Kac formula extends the uniqueness theorem for the heat equation to the class of unbounded functions satisfying a growth condition of the form given above.

Let ϕ be a C^∞ function such that $\phi = 1$ on $|X| \leq R$, and $\phi = 0$ outside $|x| > R + 1$. Put $u_R(t, x) = u(t, x)\phi$,

$$Z_R(t, x) = u_R(t, x)e^{\int_s^t V(X(\sigma)d\sigma)}.$$

By what we have already proved, $Z_R(t, \cdot)$ is a martingale. Let

$$\tau_R(\omega) = \inf\{t : t \geq s\omega(t) \in S(0; R) = \{|x| \leq R\}\}.$$

Then $Z_R(t \wedge \tau_R, \cdot))$ is also a martingale, i.e.

$$u_R(t \wedge \tau_R, X(t \wedge \tau_R, \cdot))e^{\int_s^{t\wedge\tau_R} V(X(\sigma))d\sigma}$$

is a martingale. Equating the expectations at time $t = s$ and time $t = T$ and using the fact that

$$u_R(t \wedge \tau_R, X(t \wedge \tau_R, \cdot)) = u(t \wedge \tau_R, X(t \wedge \tau_R, \cdot))),$$

we conclude that

$$\begin{aligned} u(s, x) &= E_{s,x}[u(\tau_R \wedge T, X(\tau_R \wedge T, \cdot))e^{\int_s^{\tau_R\wedge T} V(X(s))ds}] \\ &= E_{s,x}[X_{(\tau_R \wedge T)} f(X(T))e^{\int_s^T V(X(s))ds}] + \\ &\quad + E_{s,x}[X_{(\tau_R \leq T)} u(\tau_R, X_{(\tau_R)} e^{\int_s^{\tau_R} V(X(s))ds}] \end{aligned}$$

131 Consider the second term on the right:

$$|E_{s,x}[X_{(\tau_R \le T)} u(\tau_R, X(\tau_R)) e^{\int_s^{\tau_R} V(X(s))ds}]|$$
$$\le A' e^{R^\alpha} P[_R \le T]$$

(where A' is a constant given in terms of the constants A and T and the bound of V)

$$= A' e^{R^\alpha} P[\sup_{s \le \sigma \le T} |X(\sigma)| \ge R].$$

$P[\sup\limits_{s \le \sigma \le T} |X(\sigma)| \ge R]$ is of the order of $e^{-c(T)R^2}$ and since $\alpha < 2$, the second term on the right side above tends to 0 as $R \to \infty$. Hence, on letting $R \to +\infty$, we get, by the bounded convergence theorem,

$$u(s, x) = E_{s,x}[f(X(T)) e^{\int_s^T V(X(s))ds}]$$

Application. Let $\beta(t, \cdot)$ be a one-dimensional Brownian motion. Recall (Cf. Reflection principle) that $P\{\sup\limits_{0 \le s \le t} |\beta(s)| \le 1\}$ is of the order of $\dfrac{4}{\pi} e -$ $\dfrac{\pi^2 t}{8}$. The Feynman-Kac formula will be used to explain the occurance of the factor $\dfrac{\pi^2}{8}$ in the exponent. First observe that $\dfrac{\pi^2}{8} = \dfrac{\lambda^2}{2}$ where λ is the first positive root of Cos $\lambda = 0$. Let

$$\tau(w) = \inf\{t : |\beta(t)| \ge 1\}.$$

Then

$$P\{\sup_{0 \le s \le t} |\beta(s, \cdot)| \le 1\} = P\{\tau \ge t\}.$$

Let $\phi(x) = E_x[e^{\lambda\tau}]$, $\lambda < 0$. We claim that ϕ satisfies 132

$$(*) \qquad \begin{aligned} &\frac{1}{2}\phi'' + \lambda\phi = 0, \quad |x| < 1, \\ &\phi = 1, \quad |x| = 1. \end{aligned}$$

Assume ϕ to be sufficiently smooth. Using Itô's formula we get

$$d(e^{\lambda t}\phi(\beta(t)) = e^{\lambda t}\phi'(\beta(t))d\beta(t) + [\lambda\phi(\beta(t)) + \frac{1}{2}\phi''(\beta(t))]e^{\lambda t}dt.$$

Therefore

$$e^{\lambda t}\phi(\beta(t)) - \phi(\beta(0)) = \int_0^t e^{\lambda s}\phi'(\beta(s))d\beta(s) +$$

$$+ \int_0^t [\lambda\phi(\beta(s)) + \frac{1}{2}\phi''(\beta(s))]e^{\lambda s}ds,$$

i.e.

$$e^{\lambda t}\phi(\beta(t)) - \phi(\beta(0)) - \int_0^t [\lambda\phi(\beta(s)) + \frac{1}{2}\phi''(\beta(s))]e^{\lambda s}ds$$

is a martingale. By Doob's optional sampling theorem we can stop this martingale at time τ, i.e.

$$e^{\lambda(t\wedge\tau)}\phi(\beta(t\wedge\tau)) - \phi(\beta(0)) - \int_0^{t\wedge\tau} [\lambda\phi(\beta(s)) + \frac{1}{2}\phi''(\beta(s))]e^{\lambda s}ds$$

is also a martingale. But for $s \leq t \wedge \tau$,

$$\lambda\phi + \frac{1}{2}\phi'' = 0.$$

Thus we conclude that $\phi(\beta(t\wedge\tau))e^{\lambda(\tau\wedge t)}$ is a martingale. Since $\lambda < 0$ and $\phi(\beta(t\wedge\tau))$ is bounded, this martingale is uniformly integrable. Therefore equating the expectation at $t = 0$ and $t = \infty$ we get (since $\phi(\beta(\tau)) = 1$)

$$\phi(x) = E_x[e^{\lambda\tau}].$$

133 By uniqueness property this must be the solution. However (*) has a solution given by

$$(x) = \frac{\mathrm{Cos}(\sqrt{(2\lambda x)})}{\mathrm{Cos}(\sqrt{(2\lambda)})}.$$

Therefore

$$E_0[e^{\lambda\tau}] = \frac{1}{\mathrm{Cos}(\sqrt{(2\lambda)})}(\lambda < 0), \tag{1}$$

If $F(t) = P(\tau \geq t)$, then

$$\int_0^\infty e^{\lambda t} dF(t) = E_0(e^{\lambda\tau}).$$

A theorem on Laplace transforms now tells us that (1) is valid till we cross the first singularity of $\dfrac{1}{\mathrm{Cos}(\sqrt{(2\lambda)})}$. This occurs at $\lambda = \dfrac{\pi^2}{8}$. By the monotone convergence theorem

$$E_0[e^{\tau\pi^2/8}] = +\infty$$

Hence $\int_0^\infty e^{\lambda t} dF(t)$ converges for $\lambda < \dfrac{\pi^2}{8}$ and diverges for $\lambda \geq \dfrac{\pi^2}{8}$. Thus $\dfrac{\pi^2}{8}$ is the supremum of λ for which $\int_0^\infty e^{\lambda t} dF(t)$ converges, i.e. sup $[\lambda : E_0(e^{\lambda\tau})]$ exists, i.e. the decay rate is connected to the existence or the non existence of the solution of the system (*). This is a general feature and prevails even in higher dimensions.

19. An Application of the Feynman-Kac Formula. The Arc Sine Law.

LET $\beta(t,\cdot)$ BE THE one-dimensional Brownian motion with $\beta(0,\cdot)=0$. 134
Define

$$\xi_t(w)=\frac{1}{t}\int\limits_0^t X_{[0,\infty)}(\beta(s,w))ds.$$

$\xi_t(w)$ is a random variable and denotes the fraction of the time that a Brownian particle stays above the x-axis during the time interval $[0,t]$. We shall calculate

$$P[w:\xi_t(w)\leq a]=F_t(a)$$

Brownian Scaling. Let $X_t(s)=\dfrac{1}{\sqrt{t}}\beta(ts)$. Then X_t is also a Brownian motion with same distribution as that of $\beta(s)$. We can write

$$\xi_t(w)=\int\limits_0^1 X_{[0,\infty)}(X_t(s,w))ds.$$

The $\xi_t(w)$ = time spent above the x-axis by the Brownian motion $X_t(s)$ in $[0,1]$. Hence $F_t(a)$ is independent of t and is therefore denoted by $F(a)$. Suppose we succeed in solving for $F(a)$; if, now,

$$\xi_t^*(w) = \int_0^t X_{[0,\infty)}(\beta(s))ds = t\xi_t,$$

then the amount of time $\beta(s,\cdot) > 0$ in $[0,t]$ is t (amount of time $X_t(s) > 0$ in $[0,1]$). Hence we can solve for $P[\xi_t^*(w) \leq a] = G_t(a)$. Clearly the solution of G_t is given by

$$G_t(a) = F(a/t).$$

135 It is clear that

$$F(a) = \begin{cases} 0 & \text{if} \quad a \leq 0, \\ 1 & \text{if} \quad a \geq 1. \end{cases}$$

Hence it is enough to solve for $F(a)$ in $0 \leq a \leq 1$. Let

$$u_\lambda(t,x) = E_x[e^{-(\lambda \int_0^t X_{[0,\infty)}(\beta(s,w))ds)}]$$

Then

$$u_1(t,0) = E[e^{-\xi(w))}] = \int_0^1 e^{-tx} dF(x).$$

Also note that $u_\lambda(t,x)$ is bounded by 1, if $\lambda \geq 0$. By the Feynman-Kac formula (appropriately modified in case $\frac{1}{2}\Delta$ is replaced by $-\frac{1}{2}\Delta$) $u_1(t,x)$ satisfies

$$(*) \qquad \begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2} - u, \quad && x > 0, \\ &= \frac{1}{2}\frac{\partial^2 u}{\partial x^2}, && x < 0, \end{aligned}$$

and $u(0,x) = 1$. Let

$$\phi_\alpha(x) = \alpha \int_0^\infty u(t,x)e^{-\alpha t}dt, \quad \alpha > 0, \quad \text{where} \quad u = u_1,$$

$$= \int_0^\infty u(t,x)(-de^{-\alpha t}).$$

A simple integration by parts together with (*) gives the following system of ordinary differential equations for ϕ_α:

$$-\frac{1}{2}\phi_\alpha'' + (\alpha+1)\phi_\alpha = \alpha, \quad x > 0,$$
$$-\frac{1}{2}\phi_\alpha'' + \alpha\phi_\alpha = \alpha, \quad x < 0.$$

These have a solution

$$\phi_\alpha(x) = \frac{\alpha}{\alpha+1} + Ae^{x\sqrt{(2(\alpha+1))}} + Be^{-x\sqrt{(2(\alpha+1))}}, \quad x > 0,$$
$$= 1 + Ce^{x\sqrt{(2\alpha)}} + De^{-x\sqrt{(2\alpha)}}, \quad x < 0.$$

136

However u is bounded by 1 (see definition of $u_\lambda(t,x)$). Therefore ϕ_α is also bounded by 1. This forces $A = D = 0$. We demand that ϕ_α and $\frac{d\phi_\alpha}{dx}$ should match at $x = 0$. (Some justification for this will be given later on).

$$\operatorname*{Lt}_{x\to 0+} \phi_\alpha(x) = \operatorname*{Lt}_{x\to 0-} \phi_\alpha(x)$$

gives

$$\frac{\alpha}{1+\alpha} + B = C + 1.$$

Similarly we get $-B\sqrt{(2(\alpha+1))} = C\sqrt{(2\alpha)}$ by matching $\frac{d\phi_\alpha}{dx}$. Solving for B and C we get

$$B = \frac{\sqrt{\alpha}}{(1+\alpha)(\sqrt{\alpha} + \sqrt{(\alpha+1)})},$$
$$C = \frac{1}{\sqrt{(1+\alpha)}(\sqrt{\alpha} + \sqrt{(\alpha+1)})}.$$

Therefore

$$\phi_\alpha(0) = \frac{\alpha}{\alpha+1} + B = \frac{\sqrt{\alpha}}{\sqrt{(\alpha+1)}},$$

i.e.

$$\int_0^\infty E[e^{-t\xi_t}\alpha e^{-\alpha t}]dt = \frac{\sqrt{\alpha}}{\sqrt{(\alpha+1)}}.$$

Using Fubini's theorem this gives

$$E[\frac{\alpha}{\alpha+\xi}] = \sqrt{(\frac{\alpha}{\alpha+1})},$$

or

$$E[\frac{1}{1+\xi/a}] = \sqrt{(\frac{1}{1+(1/\alpha)})}, \quad \text{i.e.} \quad \int_0^1 \frac{dF(x)}{1+\gamma x} = \frac{1}{\sqrt{(1+\gamma)}}.$$

137

This can be inverted to get

$$dF(x) = \frac{2}{\pi}\frac{dx}{\sqrt{(x(1-x))}}.$$

(Refer tables on transforms or check directly that

$$\frac{2}{\pi}\int_0^1 \frac{1}{1+\beta x}\frac{dx}{\sqrt{(x(1-x))}} = \frac{1}{\sqrt{(1+\beta)}}$$

by expanding the left side in powers of (β). Therefore

$$F(a) = \frac{2}{\pi} \quad \arcsin \quad (\sqrt{a}), \quad 0 \leq a \leq 1.$$

Hence $G_t(a) = \frac{2}{\pi}$ arcsin $(\sqrt{(\frac{a}{t})})$, $0 \leq a \leq t$, i.e.

$$P[\xi_t \leq a] = \frac{2}{\pi} \quad \arcsin \quad (\sqrt{(\frac{a}{t})}), \quad 0 \leq a \leq t.$$

This result goes by the name of *arc sine law* for obvious reasons.

We now give some justification regarding the matching conditions used above.

The equation we solved was

$$\alpha\phi - \frac{1}{2}\phi'' + V\phi = f$$

where ϕ was bounded $V \geq 0$. Suppose we formally use Itô's formula to calculate

$$\begin{aligned} & d\Big(\phi(\beta(t)e^{-\int_0^t(\alpha+V)(\beta(s)ds)}\big) \\ &= e^{-\int_0^t(\alpha+V)(\beta(s,\cdot))ds}[-f(\beta(s,\cdot))dt + \frac{d\phi}{dX}(\beta(t))d\beta(t)] \end{aligned}$$

(see Example 2 of Itô's formula). Therefore

$$(Z(t,\cdot) = \phi(\beta(t,\cdot))e^{-\int_0^t(\alpha+V)(\beta(s,\cdot))ds} + \int_0^t f\beta(s,\cdot)\exp(-\int_0^s(\alpha+V)d\sigma)ds$$

is a martingale. Since ϕ, f are bounded and $V \geq 0$, **138**

$$|Z(t,\cdot)| \leq \|\phi\|_\infty + \|f\|_\infty \int_0^\infty e^{-\alpha s}ds \leq \|\phi\|_\infty + C\|f\|_\infty.$$

Therefore $Z(t,\cdot)$ is uniformly integrable. Equating the expectations at time 0 and ∞ gives

$$(*) \qquad \phi(0) = E_0 \int_0^\infty [f(\beta(s,\cdot))e^{-\alpha s - \int_0^s V(\beta(\sigma)d\sigma)}]ds.$$

This is exactly the form obtained by solving the differential equations. In order to use Itô's formula one has to justify it. If we show that Itô's formula is valid for functions having a discontinuity in the second derivative, (*) will be a legitimate solution and in general there is no reason why the second derivatives (or higher derivatives) should be matched. This partially explains the need for matching ϕ and $\frac{d\phi}{dx}$ only.

Proposition . *Let $\beta(t,\cdot)$ denote a one-dimensional Brownian motion. Suppose $\phi \in C_b^1$ and satisfied*

$$\alpha\phi - \frac{1}{2}\phi'' + V\phi = f,$$

Then

$$\phi(\beta(t)) - \int_0^t f(\beta(s))ds$$

is a martingale.

Proof. Let $(\phi_\epsilon) \in C_b^2$ such that $\alpha\phi_\epsilon - \frac{1}{2}\phi_\epsilon'' + V\phi_\epsilon + V\phi_\epsilon = f_\epsilon$ and such that (i) ϕ_ϵ converges to ϕ uniformly on compact sets, (ii) ϕ_ϵ' converges to ϕ' uniformly on compact sets, (iii) ϕ_ϵ'' converges pointwise to ϕ'' except at 0. We may suppose that the convergence is bounded. □

139 **Claim.** $\int_0^t f_\epsilon(\beta(s))ds$ *converges to* $\int_0^t f(\beta(s))ds$ *a.e.* As $f_\epsilon(\beta(s))$ converges to $f(\beta(s))$ except when $\beta(s) = 0$, it is enough to prove that

(*) $P[w$: Lebesgue measure $(s : \beta(s) = 0) > 0] = 0$. Let $X_{\{0\}}$ denote the indicator function of $\{0\}$. Then

$$E\int_0^t X_{\{0\}}(\beta(s))ds = \int_0^t EX_{\{0\}}(\beta(s))ds = 0.$$

Thus (*) holds and establishes the claim. Now

$$\phi_\epsilon(\beta(t)) - \int_0^t f_\epsilon(\beta(s))ds$$

is a uniformly bounded martingale converging to

$$\phi(\beta(t)) - \int_0^t f(\beta(s))ds.$$

Therefore

$$\phi(\beta(t)) - \int_0^t f(\beta(s))ds$$

is a martingale.

20. Brownian Motion with Drift

LET $\Omega = C[0,\infty;\mathbb{R}^d]$, $\mathscr{F}$ = BOREL σ-FIELD of Ω, $\{X(t,\cdot)\} \equiv$ Brownian motion, $\mathscr{F}_t = \sigma[X(s,\cdot) : 0 \leq s \leq t]$, $P_x \equiv$ probability measure on Ω corresponding to the Brownian motion starting at time 0 at x. $\mathscr{F} = \sigma(\underset{t\geq 0}{U}\mathscr{F}_t)$. Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be any bounded measurable funciton. Then the map $(s,w)| \to b(w(s))$ is progressively measurable and 140

$$Z(t,\cdot) = \exp[\int\limits_0^t \langle b(X(s,\cdot)), dX(s,\cdot)\rangle - \frac{1}{2}\int\limits_0^t |b(X(s,\cdot))|^2 ds]$$

is a martingale relative to $(\Omega, \mathscr{F}_t, P_x)$. Define Q_x^t on $\mathscr{F}_t$ by

$$Q_x^t(A) = \int\limits_A Z(t,\cdot) dP_x,$$

i.e. $Z(t,\cdot)$ is the Radon-Nikodym derivative of Q_x^t with respect to P_x on $\mathscr{F}_t$.

Proposition . (i) *Q_x^t is a probability measure.*

(ii) *$\{Q_x^t : t \geq 0\}$ is a consistent family on $\underset{t\geq 0}{\cup}\mathscr{F}_t$, i.e. if $A \in \mathscr{F}_{t_1}$ and $t_2 \geq t_1$ then $Q_x^t 1(A) = Q_x^t 2(A)$.*

Proof. Q_x^t being an indefinite integral, is a measure. Since $Z(t,\cdot) \geq 0$, Q_x^t is a positive measure. $Q_x^t(\Omega) = E_x(Z(t,\cdot)) = E_x(Z(0,\cdot)) = 1$. This proves (i). (ii) follows from the fact that $Z(t,\cdot)$ is a martingale.

If $A \in \mathscr{F}_t$, we define

$$Q_x(A) = Q_x^{t(A)}.$$

141 The above proposition shows that Q_x is well defined and since $(\mathscr{F}_t)$ is an increasing family, Q_x is finitely additive on $\bigcup\limits_{t \geq 0} \mathscr{F}_t$. ☐

Exercise. Show that Q_x is countably additive on the algebra $\bigcup\limits_{t \geq 0} \mathscr{F}_t$.

Then Q_x extends as a measure to $\mathscr{F} = \sigma(\bigcup\limits_{t \geq 0} \mathscr{F}_t)$. Thus we get a family of measures $\{Q_x : x \in \mathbb{R}^d\}$ defined on $(\Omega, \mathscr{F})$.

Proposition . *If $s < t$ then*

$$Q_x(X_t \in A|\mathscr{F}_s) = Q_{X(s)}(X(t-s) \in A) \quad a.e.$$

Definition. If a family of measures $\{Q_x\}$ satisfies the above property it is called a homogeneous Markov family.

Proof. Let $B \in \mathscr{F}_s$. Therefore $B \cap X_t^{-1}(A) \in \mathscr{F}_t$ and by definition,

$$\begin{aligned}
Q_x((X(t) \in A) \cap B)) &= \int\limits_{B \cap X_t^{-1}(A)} Z(t, \cdot) dP_x \\
E^{P_x}(Z(t, \cdot)\chi_B^{(w)}\chi_A(X(t, \cdot))) & \\
&= E^{P_x}(E^{P_x}(\frac{Z(t,\cdot)}{Z(s,\cdot)} Z(s,\cdot)\chi_B^{(w)}\chi_A(X(t,\cdot))|\mathscr{F}_s) \\
&= E^{P_x}([\chi_B Z(s,\cdot))E^{P_x}(\frac{Z(t,\cdot)}{Z(s,\cdot)}\chi_A(\chi(t,\cdot))]|\mathscr{F}_s)
\end{aligned}$$

(since $B \in \mathscr{F}_s$ and $Z(s, \cdot)$ is $\mathscr{F}_s$-measurable)

$$\begin{aligned}
(1) \qquad &= E^{Q_x}[\chi_B E^{P_x}(\frac{Z(t,\cdot)}{Z(s,\cdot)}\chi_A(\chi(t,\cdot))|\mathscr{F}_s)] \ldots \\
&= E^{Q_x}[\chi_B E^{P_x}(\exp[\int\limits_s^t \langle b, dX\rangle - \frac{1}{2}\int\limits_0^t |b|^2]\chi_A(X(t,\cdot))|\mathscr{F}_s)]
\end{aligned}$$

$$= E^{Q_x}[\chi_B E^{P_{X(s)}}(\exp[\int_0^{t-s}\langle b, dX\rangle \frac{1}{2}\int_0^{t-s}|b|^2]X_A(X(t-s))]$$

(by Markov property of Brownian motion)

$$= E^{Q_x}(\chi_B E^{Q^{t-s}_{X(s)}}(\chi_A(\chi(t-s))) \quad (\text{since } \frac{dQ^{t-s}_{X(s)}}{dP_{X(s)}} = Z(t-s,\cdot)$$

$$= E^{Q_x}(X_B E^{Q_{X(s)}}\chi_A(X(t-s,\cdot))$$

□ **142**

The result follows from definition.

Let $b : [0,\infty] \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded measurable function, $P_{s,x}$ the probability measure corresponding to the Brownian motion starting at time s at the point x. Define, for $t \geq s$,

$$Z_{s,t}(w) = \exp[\int_s^t \langle b(\sigma, X(\sigma,w)), dX(\sigma,w)\rangle - \frac{1}{2}\int_s^t |b(\sigma, X(\sigma,w))|^2 d\sigma]$$

Exercise. (i) $Z_{s,t}$ is a martingale relative to $(\mathscr{F}^s_t, P_{s,x})$.

(ii) Define $Q^t_{s,x}$ by $Q^t_{s,x}(A) = \int_A Z_{s,t} dP_{s,x}$, $\forall A \in \mathscr{F}^s_t$.

Show that $Q^t_{s,x}$ is a probability measure on $\mathscr{F}^s_t$.

(iii) $Q^t_{s,x}$ is a consistent family.

(iv) $Q_{s,x}$ defined on $\underset{t\geq s}{U}\mathscr{F}^s_t$ by $Q_{s,x}|\mathscr{F}^s_t = Q^t_{s,x}$ is a finitely additive set function which is countably additive.

(v) The family $\{Q_{s,x} : 0 \leq s < \infty, x \in \mathbb{R}^d\}$ is an inhomogeneous Markov family, i.e.

$$Q_{s,x}(X(t,\cdot) \in A|\mathscr{F}^s_\sigma) = Q_{\sigma,X(\sigma,\cdot)}(X(t,\cdot) \in A), \forall s < \sigma < t, A \in \mathscr{F}^s_t.$$

[Hint: Repeat the arguments of the previous section with obvious modifications]. 143

Proposition . *Let τ be a stopping time, $\tau \geq s$. Then*

$$Q_{s,x}[X(t,\cdot) \in A | \mathscr{F}^s_\tau] = Q_{\tau, X_\tau(\cdot)}(X(t,\cdot) \in A) \text{ on } \tau(w) < t,$$
$$= \chi_A(X(t,\cdot)) \text{ on } \tau(w) \geq t.$$

Proof. Let $B \in \mathscr{F}^s_\tau$ and $B \subset \{\tau < t\}$ so that $B \in \mathscr{F}^s_t$.

$$E^{Q_{s,x}}(\chi_B \chi_A(X_t)) = E^{P_{s,x}}(Z_{s,t}\chi_B\chi_A(X_t))$$
$$= E^{P_{s,x}}[E^{P_{s,x}}(Z_{\tau,t}Z_{s,\tau}\chi_B\chi_A(X_t)|\mathscr{F}^s_\tau)]$$

(since Z satisfies the multiplicative property)

$$= E^{P_{s,x}}(Z_{s,\tau}\chi_B E^{P_{s,x}}(Z_{\tau,t}\chi_A(X_t)|\mathscr{F}^s_\tau)]$$

(since $Z_{s,\tau}$ is $\mathscr{F}^s_\tau$-measurable)

(*) $$= E^{P_{s,x}}[Z_{s,\tau}X_B E^{P_{\tau},X_\tau}(Z_{\tau,t}\chi_A(X_t))]$$

(by strong Markov property).

Now
$$\left.\frac{dQ_{s,x}}{dP_{s,x}}\right|_{\mathscr{F}^s_t} = Z_{s,t}.$$

so that the optional stopping theorem,

$$\left.\frac{dQ_{s,x}}{dP_{s,x}}\right|_{\mathscr{F}^s_\tau} = Z_{s,} \quad \text{on} \quad \{\tau < t\}, \ \forall x.$$

144 Putting this in (*) we get

$$E^{Q_{s,x}}[X_B\chi_A(X_t)] = E^{P_{s,x}}[Z_{s,\tau}\chi_B E^{Q_{\tau,X_\tau}(\chi_A X_t)}].$$

Observe that
$$\chi_B E^{Q_{\tau,X_\tau}}(\chi_A(X_t))$$
is $\mathscr{F}^s_\tau$-measurable to conclude the first part of the proof. For part (ii) observe that
$$X_t^{-1}(A) \cap \{\tau \geq t\} \cap \{\tau \leq k\}$$

is in $\mathscr{F}_k^s$ if $k > s$, so that

$$X_t^{-1}(A) \cap \{\tau \geq t\} \in \mathscr{F}_\tau^s.$$

Therefore

$$E^{Q_{s,x}}(X_t \in A \cap (\tau \geq t))|\mathscr{F}_\tau^s) = \chi_A(X_t)\chi_{\{\tau \geq t\}},$$

or

$$E^{Q_{s,x}}[(X_t \in A)|\mathscr{F}_\tau^s] = \chi_A(X_t) \quad \text{if} \quad \tau \geq t.$$

□

Proposition . *Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded measurable function, $f : \mathbb{R}^d \to \mathbb{R}$ any continuous bounded function. If*

$$\frac{\partial u}{\partial s} + \frac{1}{2}\Delta u + \langle b(s,x), \Delta u\rangle = 0, \quad 0 \leq s \leq t,$$
$$u(t,x) = f(x)$$

has a solution u, then

$$u(s,x) = \int_\Omega f(X_t) dQ_{s,x}.$$

Remark. b is called the *drift*. If $b = 0$ and $s = 0$ then we recover the result obtained earlier. With the presence of the drift term, the result is the same except that instead of $P_{s,x}$ one has to use $Q_{s,x}$ to evaluate the expectation. 145

Proof. Let

$$Y(\sigma,\cdot) = \int_s^\sigma \langle b(\theta, X(\theta,\cdot)), dX(\theta,\cdot)\rangle - \frac{1}{2}\int_s^\sigma |b(\theta, X(\theta,\cdot)|^2 d\theta.$$

Step 1. $(X(\sigma,\cdot) - X(s,\cdot), Y(\sigma,\cdot))$ is a $(d+1)$-dimensional Itô process with parameters

$$(0,0,\ldots,0,-\frac{1}{2}|b(\sigma, X(\sigma,\cdot))|^2) \quad \text{and}$$

d terms

$$a = \begin{bmatrix} I_{d\times d} & b_{d\times 1} \\ b^*_{1\times d} & b \end{bmatrix}$$

Let $\lambda = (\lambda_1, \ldots, \lambda_d)$. We have to show that

$$\exp[\mu Y + \frac{\mu}{2}\int_s^\sigma |b(\sigma, X(\sigma,\cdot)|^2 d\sigma + \langle \lambda, X(\sigma,\cdot) - X(s,\cdot)\rangle -$$

$$-\frac{1}{2}\int_s^\sigma \langle \lambda, \lambda\rangle + 2\mu\langle \lambda, b + \mu^2 |b(\sigma,\cdot)|^2 d\sigma]$$

is a martingale, i.e. that

$$\exp[\langle \lambda, X(\sigma,\cdot) - X(s,\cdot)\rangle + \mu\int_s^\sigma \langle b, dX\rangle - \frac{1}{2}\int_s^\sigma |\lambda + b\mu|^2 d\rho].$$

is a martingale; in other words that

$$\exp[\int_s^\sigma \langle \lambda + b\mu, dX\rangle - \frac{1}{2}\int_s^\sigma |\lambda + b\mu|^2 d\rho]$$

146 is a martingale. But this is obvious because

$$Z(\sigma,\cdot) = \int_s^\sigma \langle \lambda + \mu b, dX\rangle$$

is a stochastic integral and hence an Itô process with parameters $(0, |\lambda + \mu b|^2)$. (Refer to the section on vector-valued Itô process).

Step 2. Put $\phi(\sigma, X(\sigma,\cdot), Y(\sigma,\cdot)) = u(\sigma, X(\sigma,\cdot))e^{Y(\sigma,\cdot)}$. By Itô formula,

$$d\phi = e^Y \frac{\partial u}{\partial t}dt + e^Y\langle \nabla u, dX\rangle + ue^Y dY + \frac{1}{2}\sum a_{ij}\frac{\pi^2\phi}{\partial z_i \partial z_j},$$

where $z = (x, y)$, or

$$d\phi = e^Y[\frac{\partial u}{\partial t}dt + \langle \nabla u, dX\rangle + u\langle b, dX\rangle - \frac{\mu}{2}|b|^2 dt + \frac{1}{2}\nabla u dt + \langle b, \nabla u\rangle dt +$$

$$+\frac{1}{2}u|b|^2 dt]$$
$$= e^Y[\langle \nabla u, dX\rangle + u\langle b, dX\rangle].$$

Therefore ϕ is an Itô process and hence a martingale. Therefore

$$E(\phi(t,\cdot)) = E(\phi(s,\cdot))$$
$$u(s,x) = E^{P_{s,x}}[(f(X(t))e^{\int_s^t \langle b, dX\rangle - \frac{1}{2}\int\limits_0^t |b|^2 d\theta}]$$
$$= E^{Q_{s,x}}[f(X(t))],$$

which proves the theorem.

□

Alternate Proof. **147**

Exercise . Let $Y(\sigma,\cdot)$ be progressively measurable for $\sigma \geq s$. Then $Y(\sigma,\cdot)$ is a martingale relative to $(Q_{s,x}, \mathscr{F}_t^s)$ if and only if $Y(\sigma)Z_{s,\sigma}$ is a martingale relative to $(P_{s,x}, \mathscr{F}_t^s)$.

Now for any function θ which is progressively measurable and bounded,

$$\exp[\int\limits_s^t \langle \theta, dX\rangle - \frac{1}{2}\int\limits_s^t |\theta|^2 d\sigma]$$

is a martingale relative to $(\Omega, \mathscr{F}_t^s, P_{s,x})$. In particular let θ be replaced by $\theta + b(\sigma, w(\sigma))$. After some rearrangement one finds that X_t is an Itô process with parameters b, I relative to $Q_{s,x}$. Therefore

$$u(t, X_t) - \int\limits_s^t \left(\frac{\partial u}{\partial \sigma} + \langle b, \nabla u\rangle + \frac{1}{2}\nabla u\right) d\sigma$$

is a martingale relative to $Q_{s,x}$. But

$$\frac{\partial u}{\partial \sigma} + \langle b, \nabla u\rangle + \frac{1}{2}\nabla u = 0.$$

Therefore

$$E^{Q_{s,x}}(u(t, X(t)) = u(s, X).$$

We have defined $Q_{s,x}$ by using the notion of the Radon-Nikodym derivative. We give one more relation between P and Q.

Theorem . *Let* $T : C([s, \infty), \mathbb{R}^d) \to C([s, \infty), \mathbb{R}^d)$ *be given by*

$$TX = Y \quad \text{where} \quad Y(t) = X(t) - \int_s^t b(\sigma, X(\sigma)) d\sigma.$$

(b is as before). Then

$$Q_{s,x} T^{-1} = P_{s,x}.$$

148 *Proof.* Define $Y(t, w) = X(t, Tw)$ where X is a Brownian motion. We prove that Y is a Brownian motion with respect to $Q_{s,x}$. Clearly Y is progressively measurable because T is $(\mathscr{F}_t - \mathscr{F}_t)$-measurable for every t, i.e. $T^{-1}(\mathscr{F}_t) \subset \mathscr{F}_t$ and X is progressively measurable. Clearly $Y(t, w)$ is continuous $\forall\ w$. We have only to show that $Y(t_2) - Y(t_1)$ is $Q_{s,x}$-independent of $\mathscr{F}^s_{t_1}$ and has distribution $N(0; (t_2 - t_1)I)$ for each $t_2 > t_1 \geq s$. But we have checked that

$$\exp[\langle \theta, X_t - x \rangle - \frac{1}{2} |\theta|^2 (t - s) - \int_s^t \langle \theta, b \rangle d\sigma]$$

is a martingale relative to $Q_{s,x}$. Therefore

$$E^{Q_{s,x}}(\exp\langle \theta, Y_{t_2} - Y_{t_1} \rangle | \mathscr{F}^s_{t_1}) = \exp(\frac{1}{2} |\theta|^2 (t_2 - t_1)),$$

showing that $Y_{t_2} - Y_{t_1}$ is independent of $\mathscr{F}^s_{t_1}$ and has normal distribution $N(0; (t_2 - t_1)I)$. Thus Y is a Brownian motion relative to $Q_{s,x}$. Therefore

$$Q_{s,x} T^{-1} = P_{s,x}.$$

□

21. Integral Equations

Definition. A function $b : R^d \to R^d$ is said to be locally Lipschitz if given any $x_0 \in R^d$ there exists an open set U_0 containing x_0 such that $b|_{U_0}$ is Lipschitz. **149**

Exercise 1. b is locally Lipschitz iff $b|_K$ is Lipschitz for every compact set K i.e. iff $b|_K$ is Lipschitz for every closed sphere K.

Exercise 2. Every locally Lipschitz function is continuous.

Theorem . *Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be locally Lipschitz and $X : [0, \infty) \to \mathbb{R}^d$ continuous. Then*

(i) *the equation*

$$Y(t) = X(t) + \int_0^t b(Y(s))ds \quad (*)$$

has a continuous solution near 0*, i.e. there exists an $\epsilon > 0$ and a continuous function $Y : [0, \epsilon] \to R^d$ such that the above equation is satisfied for all t in* $[0, \epsilon]$.

(ii) *(Uniqueness) If Y_1, Y_2 are continuous solutions of the above equation in* $[0, T]$*, then*

$$Y_1 = Y_2 \quad on \quad [0, T].$$

Proof. (ii) (Uniqueness) Let $f(t) = |Y_1(t) - Y_2(t)|$. As Y_1, Y_2 are continuous, there exists a $k > 0$ such that $|Y_1(t)|$, $|Y_2(t)| \leq k$ for all t in

$[0,T]$. Choose C such that $|b(x)-b(y)| \leq C|x-y|$ for $|x|$, $|y| \leq k$ and
150 $C \sup\limits_{0\leq t\leq T} f(t)$. Then $f(t) \leq C$ and $f(t) \leq C\int\limits_0^t f(s)ds$ so that $f(t) \leq \frac{(ct)^n}{n!}$ for all $n = 1,2,3,\ldots$. Thus $Y_1(t) = Y_2(t)$, proving uniqueness.

(i) (Existence) We can very well assume that $X(0) = 0$. Let $a = \inf\{t : |X(t)| \geq 1\}$,

$$M > \sup\{|b(x)| : |x| \leq 2\}, \quad \alpha = \inf\{a, \frac{1}{M}\},$$

$C \neq 0$, a Lipschitz constant, so that $|b(x)-b(y)| \leq C|x-y|$ for all $|x|$, $|y| \leq 2$. Define the iterations $Y_0, Y_1, \ldots$ by

$$Y_0(t) = X(t), \quad Y_{n+1}(t) = X(t) + \int\limits_0^t b(Y_n(s))ds$$

for all $t \geq 0$. By induction, each Y_n is continuous. By induction again, $|Y_n(t) - X(t)| \leq Mt$ for all n, $0 \leq t \leq \alpha$. Again, by induction $|Y_{n+1}(t) - Y_n(t)| \leq \frac{M}{C}\frac{(Ct)^{n+1}}{(n+1)!}$ for $0 \leq t \leq \alpha$. Again, by induciton $|Y_{n+1}(t) - Y_n(t)| \leq \frac{M}{C}\frac{(Ct)^{n+1}}{(n+1)!}$ for $0 \leq t \leq \alpha$. Thus $Y_n(t)$ converges uniformly on $[0,\alpha]$ to a continuous function $Y(t)$ which is seen to satisfy the integral equation. □

Remark. Let $X : [-\delta, \infty) \to \mathbb{R}^d$ be continuous where $\delta > 0$. Then a similar proof guarantees that the equation (*) has a solution in $[-\epsilon, \epsilon]$ for some $\epsilon > 0$.

Define $B(X) = \sup\{t : (*) \text{ has a solution in } [0,t]\}$. The theorem above implies that $0 < B(X) \leq \infty$. $B(X)$ is called the *exploding time*.

Remark . If b is, in addition, either bounded or globally Lipschitz, $B(X) = \infty$ for every continuous $X : [0,\infty) \to \mathbb{R}^d$.

151 **Example.** Let $b(y) = y^2$, $X(t) = 0$. The equation

$$Y(t) = x_0 + \int\limits_0^t b(y(s))ds$$

with $x_0 > 0$ has a solution

$$Y(t) = \frac{1}{\frac{1}{x_0} - t}, \ \forall t < \frac{1}{x_0};$$

the solution explodes at $t = x_0^{-1}$.

Proposition . *If*

$$B(w) < \infty, \quad \textit{then} \quad \operatorname*{Lt}_{t \uparrow B(w)} |y(t)| = +\infty.$$

Proof. Suppose that $\varliminf_{t \to B(w)} |y| = R < \infty$. Let (t_n) be a sequence increasing to $B(w)$ such that $|y(t_n)| \leq R + 1$, $\forall n$. Let

$$\tau_n = \inf\{t \geq t_n : |y(t) - y(-_n)| \geq 1\}.$$

Then

$$\begin{aligned} 1 &= |y(\tau_n) - y(t_n)| \\ &\leq w(\tau_n) - w(t_n)| + (\tau_n - t_n) \sup |b(\lambda)| \ldots, \ (1) \\ &\quad \lambda \in S(y(t_n), 1). \end{aligned}$$

Since (t_n) is bounded, we can choose a constant M such that

$$|w(t_n) - w(t)| < \frac{1}{2} \quad \text{if} \quad |t - t_n| \leq M.$$

Then using (1),

$$\tau_n - t_n \geq \inf\{M, (2 \sup |b(\lambda)|)^{-1} \quad \text{where} \quad \lambda \in S(Y(t_n); 1)$$

Therefore

$$\tau_n - t_n \geq \inf(M, (2 \sup |b(\lambda)|)^{-1}, \lambda \in S(0; R + 2)) = \alpha(\text{say}) \forall\ n.$$

Chose n such that $\tau_n > B(w) > t_n$. Then y is bounded in $[t_n, B(w)]$ and hence it is bounded in $[0, B(w))$. From the equation

$$y(t) = X(t) + \int_0^t b(y(s))ds$$

one then gets that $\underset{t \to B(w)}{\mathrm{Lt}}\, y(t)$ exists. But this is clearly a contradiction since in such a case the solution exists in $[0, B(w) + \epsilon)$ for suitable ϵ, contradicting the definition of $B(w)$. Thus

$$\varlimsup_{t \to B(w)} |y(t)| = +\infty$$

and hence

$$\lim_{t \to B(w)} |y(t)| = +\infty.$$

$\square$

Corollary . *If b is locally Lipschitz and bounded, then $B(X) = \infty$ for all X in $C([0, \infty), \mathbb{R}^d)$.*

Proof. Left as an exercise. $\square$

Proposition . *Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be locally Lipschitz and bounded. Define $T : C([0, \infty), \mathbb{R}^d) \to C([0, \infty), \mathbb{R}^d)$ by $TX = Y$ where*

$$Y(t) = X(t) + \int_0^t b(Y(s))ds.$$

Then T is continuous.

Proof. Let X, $X^* : [0, \infty) \to \mathbb{R}^d$ be continuous, $K > 0$ be given. Let $Y_0, Y_1, \ldots, Y_0^*, Y_1^*, \ldots$ be the iterations for X, X^* respectively. Then $|Y_n(t) - X(t)| \leq K\|b\|_\infty$ for $0 \leq t \leq K$, $n = 0, 1, 2, 3, \ldots$, so that we can
153 find R such that $|Y_n(t)|$, $Y_n^*(t)| \leq R$ for $0 \leq t \leq k$, $n = 0, 1, 2, \ldots$, Let $C \geq 1$ be any Lipschitz constant for the function b on $|x| \leq R$. Then

$$|Y_n(t) - Y_n^*(t)| \leq \sup_{0 \leq t \leq K} |X(t) - X^*(t)| \cdot (1 + Ct + \frac{(Ct)^2}{2!} +$$

$$+\cdots+\frac{(Ct)^n}{n!}) \quad \text{for} \quad 0 \leq t \leq K, \quad n = 0, 1, 2, 3, \ldots.$$

A b is bounded, Y_n converges uniformly to Y on $[0, K]$. Letting $n \to \infty$, we get

$$\sup_{0 \leq t \leq K} |(TX)t - TX^*)t| \leq e^{ck} \sup_{0 \leq t \leq K} |X(t) - X^*(t)|, \ldots (2)$$

where c depends on $\sup\limits_{0\leq t\leq K} |X(t)|, \sup\limits_{0\leq t\leq K} |X^*(t)|$. The proof follows by (2). □

22. Large Deviations

LET P_ϵ BE THE Brownian motion starting from zero scaled to Brown- **154**
ian motion corresponding to the operator $\epsilon\dfrac{\Delta}{2}$. More precisely, let

$$P_\epsilon(A) = P\left(\frac{A}{\sqrt{\epsilon}}\right)$$

where P is the Brownian motion starting at time 0 at the point 0.

Interpretation 1. Let $\{X_t : t \geq 0\}$ be Brownian motion with $X(0) = x$. Let $Y(t) = X(\epsilon t)$, $\forall t \geq 0$. Then P_ϵ is the measure induced by the process $Y(t)$. This amounts to stretching the time or scaling time.

Interpretation 2. Let $Y(t, \cdot) = \sqrt{\epsilon} X(t, \cdot)$. In this case also P_ϵ is the measure induced by the process $Y(t, \cdot)$. This amounts to 'looking at the process from a distance' or scaling the length.

Exercise. Make the interpretations given above precise.
(Hint: Calculate (i) the probability that $X(\epsilon t) \in A$, and (ii) the probability that $\sqrt{\epsilon}X(t,) \in A$).

Problem. Let

$$I(w) = \frac{1}{2}\int_0^1 |\dot{w}(t)|^2 dt$$

if $w(0) = 0$, w absolutely continuous on $[0, 1]$. Put $I(w) = \infty$ otherwise.

We would like to evaluate

$$\int_{\Omega} e^{\frac{F(w)}{\epsilon}} dP_\epsilon(w)$$

155 for small values of ϵ. Here $F(w) : C[0,1] \to \mathbb{R}$ is assumed to the a bounded and continuous function.

Theorem . *Let C be any closed set in $C[0,1]$ and let G be any open set in $C[0,1]$. Then*

$$\limsup_{\epsilon\to 0} \in \log P_\epsilon(C) \leq -\int_{w\in C} I(w),$$

$$\liminf_{\epsilon\to 0} \epsilon \log P_\epsilon(G) \geq -\inf_{w\in G} I(w).$$

Here $P_\epsilon(G) = P_\epsilon(\pi^{-1}G)$ where $\pi : C[0,\infty) \to C[0,1]$ is the canonical projection.

Significance of the theorem . If

1. $$dP_\epsilon = e^{\frac{-I(w)}{\epsilon}},$$

   then
   $$P_{\epsilon(A)} = \int_A e^{\frac{-I(w)}{\epsilon}} dP_\epsilon$$
   is asymptotically equivalent to
   $$\exp[-\frac{1}{\epsilon} \inf_{w\in A} I(w)].$$

2. If A is any set such that
   $$\inf_{w\in A^0} I(w) = \inf_{w\in \overline{A}} I(w),$$
   then by the theorem
   $$\operatorname*{Lt}_{\epsilon\to 0} \log P_\epsilon(A) = \inf_{w\in A} I(w).$$

156 **Proof of the theorem.**

Lemma 1. *Let $w_0 \in \Omega$ with $I(w_0) = \ell < \infty$. If $S = S(w_0;\delta)$ is any sphere of radius δ with centre at w_0 then* $\varliminf_{\epsilon\to 0} \epsilon \log P_\epsilon(S) \geq -I(w_0)$.

Proof.

$$
\begin{aligned}
P(S) &= \int \chi^{(w)}_{S(w_0,\delta)} dP_\epsilon\\
&= \int \chi^{(w-w_0)}_{S(0;)} dP_\epsilon\\
&= \int \chi_{S(0;\delta)}(\lambda w) dP\left(\frac{w}{\sqrt{\epsilon}}\right), \quad \text{where} \quad \lambda(w) = w - w_0,\\
&= \int \chi_{S(0;\delta)}(\lambda(\sqrt{\epsilon} w)) dP(w)\\
&= \int \chi_{S(0,\delta)}(\sqrt{\epsilon} w) \exp\left[\int_0^1 \langle w_0, dX\rangle - \frac{1}{2}\int_0^1 |\dot{w}_0|^2 d\sigma\right] dP(w)\\
&= \int \chi_{S(0;\delta)}(\sqrt{\epsilon} w) \exp\left[\int_0^1 \langle \dot{w}_0, dX\rangle - I(w_0)\right] dP(w)\\
&= \int \chi_{S(0;\delta)}(w) \exp\left[-\frac{1}{\epsilon}\int_0^1 \langle \dot{w}_0, dX\rangle - \frac{1}{\epsilon} I(w_0)\right] dP_\epsilon(w)\\
&= e^{\frac{-I(w_0)}{\epsilon}} P_\epsilon(S(0;\delta)) \frac{1}{P_\epsilon(S(0;\delta))} \int_{S(0;\delta)} \exp\left[-\frac{1}{\epsilon}\int_0^1 \langle \dot{w}_0, dX\rangle\right] dP_\epsilon\\
&\geq e^{\frac{-I(w_0)}{\epsilon}} P_\epsilon(S(0;\delta)) e\left[-\frac{1}{\epsilon}\frac{1}{P_\epsilon(S(0;\delta))}\int_{S(0;\delta)}\int_0^1 \langle \dot{w}_0, dX\rangle dP_\epsilon\right]\\
&\qquad \text{by Jensen's inequality,}\\
&= e^{\frac{-I(w_0)}{\epsilon}} P_\epsilon(S(0,\delta)) e^0 (\text{use } P_\epsilon(w) = P_\epsilon(-w) \text{ if } w \in S(0;\delta))\\
&= e^{\frac{-I(w_0)}{\epsilon}} P_\epsilon(S(0,\delta)).
\end{aligned}
$$

157

Therefore

$$P_\epsilon(S(w_0;\delta)) \geq e^{\frac{-I(w_0)}{\epsilon}} P(S(0;\frac{\delta}{\sqrt{\epsilon}}))$$

or,

$$\epsilon \log P_\epsilon(S(w_0;\delta)) \geq -I(w_0) + \epsilon \log P(S(0;\frac{\delta}{\sqrt{\epsilon}}));$$

let $\epsilon \to 0$ to get the result. Note that the Lemma is trivially satisfied if $I(w_0) = +\infty$. $\square$

Proof of Part 2 of the theorem.

Let G be open, $w_0 \in G$; then there exists $\delta > 0$ with $S(w_0, \delta) \subset G$. By Lemma 1

$$\varliminf_{\epsilon \to 0} \in \log P_\epsilon(G) \geq \varliminf_{\epsilon \to 0} \in \log P_\epsilon(S(w_0;\delta)) \geq -I(w_0).$$

Since w_0 is arbitrary, we get

$$\varliminf \in \log P_\epsilon(G) \geq -\inf\{I(w_0) : w_0 \in G\}.$$

For part 1 we need some more preliminaries.

Lemma 2. *Let* $(w_n) \in C[0,1]$ *be such that* $w_n \to w$ *uniformly on* $[0,1]$, $I(w_n) \leq \alpha < \infty$. *Then* $I(w) < \alpha$, *i.e. I is lower semi-continuous.*

Proof.

Step 1. w is absolutely continuous. Let $\{(x_i', x_i'')\}_{i=1}^n$ be a collection of mutually disjoint intervals in $[0,1]$. Then

$$\sum_{i=1}^n |w_m(x_i') - w_m(x_i'')| \leq \sum_{i=1|}^n |x_i'' - x_i'|^{1/2} [\int\limits_{x_i'}^{x_i''} |w_m|^2]^{1/2}$$

(by Hölder's inequality)

$$\leq \left(\sum_{i=1}^{n} \int_{x_i'}^{x_i''} |w_m|^2\right)^{1/2} \left(\sum_{i=1}^{n} |x_i'' - x_i'|\right)^{1/2} \quad \text{(again by Hölder)}$$

$$\leq \sqrt{(2\alpha)}(\sum |x_i'' - x_i'|)^{1/2}.$$

158

Letting $m \to \infty$ we get the result.

Step 2. Observe that $w_m(0) = 0 S_0 w(0) = 0$. Therefore

$$|\frac{w_n(x+h) - w_n(x)}{h}|^2 = |\frac{1}{h} \int_x^{x+h} w_n dt|^2 \leq \frac{1}{h^2} \left(\int_x^{x+h} |w_n| dt\right)^2$$

$$\leq \frac{1}{h} \int_x^{x+h} |w_n|^2 dt.$$

Hence

$$\int_0^{1-h} |\frac{w_n(x+h) - w_n(x)}{h}|^2 dx \leq \frac{1}{h} \int_0^{1-h} dx \int_0^h |(\dot{w}_n(x+t))|^2 dt$$

$$\leq \frac{1}{h} \int_0^h dt \int_0^{1-h} |\dot{w}_n(x+t)|^2 dx$$

$$\leq \frac{1}{2} 2 \int_0^h dt = 2\alpha$$

letting $n \to \infty$, we get

$$\int_0^{1-h} |\frac{w(x+h) - w(x)}{h}|^2 dx \leq 2\alpha.$$

Let $h \to 0$ to get $I(w) \leq \alpha$, completing the proof.

□

159 **Lemma 3.** *Let C be closed and put $C^\delta = \bigcup_{w\in C} S(w;\delta)$; then*

$$\lim_{\delta\to 0}(\inf_{w\in C^\delta} I(w)) = \inf_{w\in C} I(w).$$

Proof. If $\delta_1 < \delta_2$, then $C^{\delta_1} \subset C^{\delta_2}$ so that $\inf_{w\in C^\delta} I(w)$ is decreasing. As $C^\delta \supset C$ for each δ,

$$\lim_{\delta\to 0}(\inf_{w\in C^\delta} I(w)) \le \inf_{w\in C} I(w)$$

Let $\ell = \lim_{\delta\to 0}(\inf_{w\in C^\delta} I(w))$. Then there exists $w_\delta \in C^\delta$ such that $I(w_\delta) \to \ell$, and therefore $(I(w_\delta))$ is a bounded set bounded by α (say).

Claim.

$$|w_\delta(t_1) - w_\delta(t_2)| = |\int_{t_1}^{t_2} w_\delta dt| \le \sqrt{|(|t_1 - t_2|)(\int |w_\delta|^2)^{1/2}}$$
$$\le \sqrt{(2\alpha|t_1 - t_2|)}.$$

The family (w_δ) is therefore equicontinuous which, in view of the fact that $w_\delta(0) = 0$, implies that it is uniformly bounded and the claim follows from Ascoli's theorem. Hence every subfamily of (w_δ) is equicontinuous. By Ascoli's theorem there exists a sequence $\delta_n \to 0$ such that $w\delta_n \to w$ uniformly on $[0, 1]$. It is clear that $w \in C$. By lower semicontinuity of $I(w)$,

$$\lim_{\delta\to 0} \inf_{w\in C^\delta} I(w) \ge \inf_{w\in C}$$

completing the proof.

□

160 **Proof of Part 1 of the theorem.** Let X be continuous in $[0, 1]$. For each n let X_n be a piecewise linear version of X based on n equal intervals, i.e.

X_n is a polygonal function joining the points $(0, X(0))$, $(1/n, X(1/n)), \ldots,$ $(1, X(1))$.

$$P_\epsilon(\|X_n - X\| \geq \delta), (\|\cdot\| = \|\cdot\|_\infty)$$
$$\leq P\left(\bigcup_n \sup_{i\leq j\leq n} \cdot \sup_{\frac{j-1}{n}\leq t\leq \frac{j}{n}} |X_r(t) - X_r\left(\frac{j-1}{n}\right)| \geq \frac{\delta}{2\sqrt{d}}\right),$$
$$\text{where} \quad X = (X_1, \ldots, X_d).$$

$\leq ndP_\epsilon(\sup\limits_{0\leq t\leq 1/n} |X(t) - X(0)| \geq \dfrac{\delta}{2\sqrt{d}}$ (Markov property; here X is one-dimensional).

$$\leq ndP_\epsilon\left(\sup_{0\leq t\leq 1/n} |X_t| \geq \frac{\delta}{2\sqrt{d}}\right) \quad \text{(since } X(0) = 0\text{)}$$
$$\leq 2nd\ P_\epsilon(\sup_{0\leq t\leq 1/n} X_t \geq \frac{\delta}{2\sqrt{d}})$$
$$= 2dn\ P(\sup_{0\leq t\leq 1/n} X_t \geq \frac{\delta}{2\sqrt{\epsilon d}})$$
$$= 4dn\ P(X(1/n) \geq \frac{\delta}{2\sqrt{\epsilon d}}) \quad \text{(by the reflection principle)}$$
$$= 4dn \int\limits_{\delta\sqrt{n}/2\sqrt{\epsilon d}}^{\infty} \frac{1}{\sqrt{2\pi/n}} e^{-ny^2/2} dy$$
$$= 4d \int\limits_{\delta\sqrt{n}/2\sqrt{\epsilon d}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

Now, for every $a > 0$,

$$a\int\limits_a^\infty e^{-x^2/2} dx \leq \int\limits_a^\infty x e^{-x^2/2} dx = e^{-a^2/2}.$$

Thus

$$P_\epsilon(\|X_n - X\| \geq \delta) \leq \frac{4dn\, e^{-n\delta^2/(8\epsilon d)}}{\delta\sqrt{n/2}\,\sqrt{\epsilon}d} = C_1(n)\frac{\sqrt{\epsilon}}{\sqrt{\delta}}e^{-n\delta^2/(8\epsilon d)},$$

where C_1 depends only on n. We have now

$$\begin{aligned} P_\epsilon(X_n \in C^\delta) &\leq P_\epsilon(I(X_n) \geq \ell_\delta) \text{ where } \ell_\delta = \inf\{I(w)w \in C^\delta\}. \\ &= P\left(\frac{1}{2}\sum_{j=0}^{n-1} n|X\left(\frac{j+1}{n}\right) - X\left(\frac{j}{n}\right)|^2 \geq \ell_\delta\right) \\ &= P\left(Y_1^2 + Y_2^2 + \cdots + Y_{nd}^2 \geq \frac{2\ell_\delta}{\epsilon}\right), \end{aligned}$$

where $Y_1 = \sqrt{n}(X_1(1/n) - X_1(0))$ etc. are *independent* normal random variables with mean 0 and variance 1. Therefore,

$$\begin{aligned} &P(Y_1^2 + \cdots + Y_{nd}^2 \geq \frac{2\ell_\delta}{\epsilon}) \\ &= \int\limits_{y_1^2+\cdots+y_{nd}^2} \frac{2\ell_\delta}{\epsilon} e^{-(y_1^2+\cdots+y_{nd}^2)^{1/2}} dy_1 \ldots dy_{nd}. \\ &= C(n) \int\limits_{\sqrt{(2\ell_\delta/\epsilon)}}^{\infty} e^{-r^2/2} r^{nd-1} dr, \end{aligned}$$

using polar coordinates, i.e.

$$P(Y_1^2 + Y_2^2 + \cdots + Y_{nd}^2 \geq \frac{2\ell_\delta}{\epsilon}) = C'(n) \int\limits_{(\ell_\delta/\epsilon)}^{\infty} e^{-s} s^{\frac{nd}{2}-1} ds$$

(change the variable from r to $s = \dfrac{r^2}{2}$). An integration by parts gives

$$\int\limits_{\alpha}^{\infty} e^{-s} s^k ds = e^{-\alpha}(\alpha^k + \frac{k!}{(k-1)!}\alpha^{k-2} + \cdots).$$

Using this estimate (for n even) we get

$$P((Y_1^2 + \cdots + Y_{nd}^2) \geq \frac{2\ell_\delta}{\epsilon}) \leq C_2(n) e^{-\ell_\delta/\epsilon} (\frac{\ell_\delta}{\epsilon})^{\frac{nd}{2}-1},$$

where C_2 depends only on n. Thus,

$$\begin{aligned}
P_\epsilon(C) &\le P_\epsilon(\|X_n - X\| \ge \delta) + P_\epsilon(X_n \notin C^\delta)\\
&\le C_1(n)\sqrt{\left(\frac{\epsilon}{\delta}\right)} e^{-n\delta^2/(8\in d)} + C_2(n) e^{-\ell_\delta/\epsilon}\left(\frac{\ell_\delta}{\epsilon}\right)^{\frac{nd}{2}-1}\\
&\le 2\max[C_1(n)\sqrt{\left(\frac{\epsilon}{\delta}\right)} e^{-n\delta^2/(8\in d)}, C_2(n) e^{-\ell_\delta/\epsilon}\left(\frac{\ell_\delta}{\epsilon}\right)^{\frac{nd}{2}-1}\\
&\in \log P_\epsilon(C) \le \epsilon \log 2 + \epsilon \max[\log(C_1(n)\sqrt{\left(\frac{\epsilon}{\delta}\right)} e^{-n\delta^2/(8\in d)}\\
&\quad \log C_2(n) e^{-\ell_\delta/\epsilon}(\frac{\ell_\delta}{\epsilon})^{\frac{nd}{2}-1}]
\end{aligned}$$

162

Let $\epsilon \to 0$ to get

$$\overline{\lim} \in \log P_\epsilon(C) \le \max\left\{\frac{-n\delta^2}{8d}, \frac{-\ell_\delta}{1}\right\}.$$

Fix δ and let $n \to \infty$ through even values to get

$$\overline{\lim} \in \log P_\epsilon(C) \le -\ell_\delta.$$

Now let $\delta \to 0$ and use the previous lemma to get

$$\overline{\lim_{\epsilon \to 0}} \in \log P_\epsilon(C) \le -\int_{w \in C} I(w).$$

Proposition . *Let ℓ be finite; then $\{w : I(w) \le \ell\}$ is compact in Ω.*

Proof. Let (w_n) be any sequence, $I(w_n) \le \ell$. Then

$$|w_n(t_1) - w_n(t_2)| \le \sqrt{(}\ell|t_1 - t_2|)$$

and since $w_n(0) = 0$, we conclude that $\{w_n\}$ is equicontinuous and uniformly bounded. ◻

Assumptions. Let Ω be any separable metric space, $\mathscr{F}$ = Borel σ-field on Ω. For every $\epsilon > 0$ let P_ϵ be a probability measure. Let $I : \Omega \to [0, \infty]$ be any function such that

(i) I is lower semi-continuous.

(ii) $\forall$ finite ℓ, $\{w : I(w) \leq \ell\}$ is compact. **163**

(iii) For every closed set C in Ω,

$$\lim_{\epsilon\to 0}\sup \in \log P_\epsilon(C) \leq -\inf_{w\in C} I(w).$$

(iv) For every open set G in Ω

$$\lim_{\epsilon\to 0}\inf \in \log P_\epsilon(G) \geq -\inf_{w\in G} I(w).$$

Remark. Let $\Omega = C[0,1]$, P_ϵ the Brownian measure corresponding to the scaling ϵ. If $I(w) = \dfrac{1}{2}\int\limits_0^1 |w|^2 dt$ if $w(0) = 0$ and ∞ otherwise, then all the above assumptions are satisfied.

Theorem . *Let $F : \Omega \to \mathbb{R}$ be bounded and continuous. Under the above assumptions the following results hold.*

(i) *For every closed set C in Ω*

$$\lim_{\epsilon\to 0}\sup \epsilon \log \int\limits_C \exp \frac{F(w)}{\epsilon} dP_\epsilon \leq \sup_{w\in C}(F(w) - I(w)).$$

(ii) *For every open set G in Ω*

$$\lim_{\epsilon\to 0}\inf \in \log \int\limits_G \exp \frac{F(w)}{\epsilon} dP_\epsilon \geq \sup_{w\in G}(F(w) - I(w)).$$

In particular, if $G = \Omega = C$, then

$$\lim_{\epsilon\to 0} \in \log \int\limits_\Omega \exp \frac{F(w)}{\epsilon} dP_\epsilon = \sup_{w\in\Omega}(F(w) - I(w)).$$

Proof. Let G be open, $w_0 \in G$. Let $\delta \to 0$ be given. Then there exists a neighbourhood N of w_0, $F(w) \geq F(w_0) - \delta$, $\forall w$ in N. Therefore

$$\int_G \exp \frac{F(w)}{\epsilon} dP_\epsilon \geq \int_N \exp \frac{F(w)}{\epsilon} dP_\epsilon \geq e^{\frac{F(w_0)-\delta}{\epsilon}} P_\epsilon(N).$$

164 Therefore

$$\epsilon \log \int_G \exp \frac{F(w)}{\epsilon} dP_\epsilon \geq F(w_0) - \delta + \epsilon \log P_\epsilon(N).$$

Thus

$$\begin{aligned} \underline{\lim} \log \int_G \exp \frac{F(w)}{\epsilon} dP_\epsilon &\geq F(w_0) - \delta + \underline{\lim}\, \epsilon \log P_\epsilon(N). \\ &\geq F(w_0) - \delta - \inf_{w \in N} I(w) \geq F(w_0) - I(w_0) - \delta. \end{aligned}$$

Since δ and w_0 are arbitrary ($w_0 \in G$) we get

$$\underline{\lim} \in \log \int_G \exp \frac{F(w)}{\epsilon} dP_\epsilon \geq \sup_{w \in G} (F(w) - I(w)).$$

This proves Part (ii) of the theorem. □

Proof of Part (i).

Step 1. Let C be compact; $L = \sup_{w \in G}(F(w) - I(w))$. If $L = -\infty$ it follows easily that

$$\lim_{\epsilon \to 0} \sup \in \log \int_C e^{F/\epsilon} dP_\epsilon \leq -\infty.$$

(Use the fact that F is bounded). Thus without any loss, we may assume L to be finite. Let $w_0 \in C$; then there exists a neighbourhood N of w_0 such that $F(w) \leq F(w_0) + \delta$ and by lower semi-continuity of I,

$$I(w) \geq I(w_0) - \delta, \ \forall w \in N(w_0).$$

By regularity, there exists an open set G_{w_0} containing w_0 such that $G_{w_0}\overline{G}_{w_0}N(w_0)$. Therefore 165

$$\int\limits_{\overline{G}_{w_0}} \exp\frac{F(w)}{\epsilon} dP_\epsilon \exp\left(\frac{F(w_0)+\delta}{\epsilon}\right) P_\epsilon(\overline{G}_{w_0}).$$

Therefore

$$\begin{aligned}
\lim_{\epsilon\to 0}\sup \in \log \int\limits_{\overline{G}_{w_0}} \exp\frac{F(w)}{\epsilon} dP_\epsilon &\leq F(w_0) + \delta + \epsilon \overline{\lim_{\epsilon\to 0}} P_\epsilon(\overline{G}_{w_0}) \\
&\leq F(w_0) + \delta - \inf_{w\in\overline{G}_{w_0}} I(w) \\
& F(w_0) + \delta - I(w_0) + \delta \\
&\leq L + 2\delta.
\end{aligned}$$

Let $K_\ell = \{w : I(w) \leq \ell\}$. By assumption, K_ℓ is compact. Therefore, for each $\delta > 0$, there exists an open set G_δ containing $K_\ell \cap C$ such that

$$\lim_{\epsilon\to 0}\sup \in \log \int\limits_{G_\delta} e^{\frac{F(w)}{\epsilon}} dP_\epsilon \leq L + 2\delta.$$

Therefore

$$\lim_{\epsilon\to 0}\sup \in \log \int\limits_{G_\delta\cap C} e^{\frac{F(w)}{\epsilon}} dP_\epsilon \leq L + 2\delta,$$

$$\int\limits_{G_\delta^c\cap C} e^{\frac{F(w)}{\epsilon}} dP_\epsilon \leq e^{M/\epsilon} P(C \cap G_\delta^c).$$

Therefore

$$\begin{aligned}
\lim_{\epsilon\to 0}\sup \in \log \int\limits_{G_\delta^c\cap C} e^{\frac{F(w)}{\epsilon}} dP_\epsilon &\leq M + \lim_{\epsilon\to 0}\sup \epsilon \log P_\epsilon(C_\delta^c \cap C) \\
&\leq M - \inf_{w\in C\cap G_\delta^c} I(w).
\end{aligned}$$

166 Now
$$G_\delta^c \subset K_\ell^c \cap C^c.$$

Therefore
$$C \cap G_\delta^c \subset C \cap K_\ell^c$$
if $w \in C \cap G_\delta^c$, $w \notin K_\ell$. Therefore $I(w) > \ell$. Thus
$$\limsup_{\epsilon\to 0} \in \log \int\limits_{G_\delta^c \cap C} e^{F(w)/\epsilon} dP_\epsilon \leq M - \ell \leq L \leq L + 2\delta.$$

This proves that
$$\limsup_{\epsilon\to 0} \in \log \int\limits_C \exp \frac{F(w)}{\epsilon} dP_\epsilon \leq L + 2\delta.$$

Since C is compact there exists a finite number of points $w_1, \ldots, w_n$ in C such that
$$C \subset \bigcup_{i=1}^n G_{w_i}$$

Therefore
$$\begin{aligned}\overline{\lim} \in \log \int\limits_C \exp \frac{F(w)}{\epsilon} dP_\epsilon &\leq \overline{\lim}\, \epsilon \log \int\limits_{\bigcup_{i=1}^n G_{w_i}} e^{F(w)/\epsilon} dP_\epsilon \\ &\leq \overline{\lim}(\epsilon \log n \operatorname*{Max}_{1\leq i \leq} \int\limits_{G_{w_i}} \exp \frac{F(w)}{\epsilon} dP_\epsilon) \\ &\leq L + 2\delta.\end{aligned}$$

Since δ is arbitrary.
$$\overline{\lim} \in \log \int\limits_C \exp \frac{F(w)}{\epsilon} dP_\epsilon \leq \sup_{w\in C}(F(w) - I(w)).$$

The above proof shows that given a compact set C, and $\delta > 0$ there 167
exists an open set G containing C such that
$$\overline{\lim} \in \log \int\limits_G \exp \frac{F(w)}{\epsilon} dP_\epsilon \leq L + 2\delta.$$

Step 2. Let C be any arbitrary closed set in Ω. Let

$$L = \sup_{w \in C}(F(w) - I(w)).$$

Since F is bounded there exists an M such that $|F(w)| \leq M$ for all w. Choose ℓ so large that $M - \ell \leq L$. Since δ is arbitrary

$$\lim_{\epsilon \to 0} \sup \in \log \int_C \exp \frac{F(w)}{\epsilon} dP_\epsilon \leq \sup_{w \in C}(F(w) - I(w))$$

We now prove the above theorem when P_ϵ is replaced by Q_x^ϵ. Let P_x^ϵ be the Brownian motion starting at time $t = 0$ at the space point x corresponding to the scaling ϵ. Precisely stated, if

$$\tau_\epsilon : C([0, \infty); \mathbb{R}^d) \to C([0, \infty); \mathbb{R}^d)$$

is the map given by $(\tau_\epsilon w)(t) = w(\epsilon t)$, then $P_x^\epsilon \underset{\text{def}}{=} P_x \tau_\epsilon^{-1}$. Note $T_1'' \tau_\epsilon = T_\epsilon$ and T_ϵ is given by

$$T_\epsilon w = y \text{ where } y(t) = w(\epsilon t) + \int_0^t b(y(s)) ds.$$

Hence

$$P_x T_\epsilon^{-1} = P_x (T_1, \tau_\epsilon)^{-1} = P_x \tau_\epsilon^{-1} T_1^{-1} = P_x^\epsilon T_1^{-1};$$

either of these probability measures is denoted by Q_x^ϵ.

Theorem . *Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be bounded measurable and locally Lipschitz. Define*

$$I(w) : \frac{1}{2} \int_0^1 |X(t) - b(X(t))|^2 dt$$

168 *If $w \in C([0, \infty); \mathbb{R}^d)$, $w(0) = x$ and x absolutely continuous. Put $I(w) = \infty$ otherwise. If C is closed in $C[(0, 1]; \mathbb{R}^d)$, then*

$$\overline{\lim_{\epsilon \to 0}} \in \log Q_x^\epsilon(C) \leq - \inf_{w \in C} I(w).$$

If G is open in $C([0,1];\mathbb{R}^d)$, then

$$\varliminf_{\epsilon\to 0} \in \log Q^\epsilon_x(G) \geq -\inf_{w\in G} I(w).$$

As usual $Q^\epsilon_x(C) = Q^\epsilon_x \pi^{-1}(C)$ where

$$\pi : C([0,\infty);\mathbb{R}^d) \to C([0,1];\mathbb{R}^d)$$

is the canonical projection.

Remark. If $b = 0$ we have the previous case.

Proof. Let T be the map $x(\cdot) \to y(\cdot)$ where

$$y(t) = x(t) + \int_0^t b(y(s))ds.$$

Then

$$Q^\epsilon_x = P^\epsilon_x(T^{-1}).$$

If C is closed

$$Q^\epsilon_x(C) = P^\epsilon_x(T^{-1}C).$$

The map T is continuous. Therefore $T^{-1}(C)$ is closed. Thus

$$\begin{aligned}
&\limsup_{\epsilon\to 0} \in \log Q_x(C) = \limsup_{\epsilon\to 0} \in \log P^\epsilon_x(T^{-1}C)\\
(*)\qquad &\leq -\inf_{w\in T^{-1}(C)} \frac{1}{2}\int_0^1 |X|^2 dt \quad \text{(see Exercise 1 below)}\\
&= -\inf_{w\in C} \frac{1}{2}\int_0^1 |T^{-1}w|^2 dt.
\end{aligned}$$

Now **169**

$$y(\cdot) \xrightarrow{T^{-1}} y(t) - \int_0^t b(y(s))ds.$$

Therefore

$$(T^{-1}y) = y - b(y(s)).$$

Therefore

$$\lim_{\epsilon\to 0}\sup \in \log Q^{\epsilon}_x(C) \leq - \inf_{w\in C} I(w).$$

The proof when G is one is similar. □

Exercise 1. Replace P_ϵ by P^{ϵ}_x and I by I_x where

$$I_x(w) = \frac{1}{2}\int_0^1 |w|^2,\ w(0) = x, w \text{ absolutely continuous},$$
$$= \infty \text{ otherwise}.$$

Check that (*) holds, i.e.

$$\lim_{\epsilon\to 0}\sup \in \log P^{\epsilon}_x(C) \leq - \inf_{w\in C} I_x(w),\ \text{if } C \text{ is closed},$$

and

$$\lim_{\epsilon\to 0}\inf \in \log P^{\epsilon}_x(G) \geq - \inf_{w\in G} I_x(w).$$

Let G be a bounded open set in $\mathbb{R}^n$, with a smooth boundary $\Gamma = \partial G$. Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be a smooth C^∞ function such that

(i) $\langle b(x), n(x)\rangle 0$, $\forall x \in \partial G$ where $n(x)$ is the unit inward normal.

(ii) there exists a point $x_0 \in G$ with $b(x_0) = 0$ and $|b(x)| > 0$, $\forall x$ in $G - \{x_0\}$.

170 (iii) for any x in G the solution

$$\xi(t) = x + \int_0^t b(\xi(s))ds,$$

of the vector field starting from x converges to x_0 as $t \to +\infty$.

Remark. (a) (iii) is usually interpreted by saying that "x_0 is stable".

(b) By (i) and (ii) every solution of (iii) takes all its values in G and ultimately stays close to x_0.

Let $\epsilon > 0$ be given; $f : \partial G \to \mathbb{R}$ be any continuous bounded function. Consider the system

$$L_\epsilon u_\epsilon = \frac{1}{2}\Delta u_\epsilon + b(x) \cdot \Delta u_\epsilon = 0 \text{ in } G$$
$$u_\epsilon = f \text{ on } \partial G.$$

We want to study $\lim\limits_{\epsilon \to 0} u_\epsilon(x)$. Define

$$I_0^T(X(t)) = \frac{1}{2}\int\limits_0^T |X(t) - b(X(t))|^2 dt; \; X : [0, T] \to \mathbb{R}^d$$

whenever X is absolutely continuous, $= \infty$ otherwise.

Remark. Any solution of (iii) is called an integral curve. For any curve X on $[0, T]$, I_0^T gives a measure of the deviation of X from being an integral curve. Let

$$V_T(x, y) = \inf\{I_0^T(X) : X(0) = x; \; X(T) = y\}$$

and

$$V(x, y) = \inf\{V_T(x, y) : T > 0\}.$$

V has the following properties.

(i) $V(x, y) \leq V(x, z) + V(z, y) \; \forall x, y, z.$ **171**

(ii) Given any x, $\exists \delta \to 0$ and $C > 0$ such that for all y with $|x - y| \leq \delta$.

$$V(x, y) \leq C|x - y|$$

Proof. Let $X(t) = \dfrac{t(y - x)}{|y - x|} + x$.

Put

$$T = |y - x|, \; X(0) = x, \; X(T) = y,$$

$$I_0^T(X(t)) = \frac{1}{2}\int\limits_0^T \left|\frac{y-x}{T} - b(X + \frac{S}{T}(y-x))\right|^2 ds.$$

Then

$$I_0^T \leq \frac{1}{2}\int\limits_0^T 2\left(\frac{|y-x|^2}{T^2} + ||b||_\infty^2\right) ds,$$

where

$$||b||_\infty = \sup_{|\lambda - x| \leq |y-x|} b(\lambda),$$

or,

$$I_0^T \leq (1 + ||b||_\infty^2)|y-x|.$$

As a consequence of (ii) we conclude that

$$V(x,y) \leq \left(1 + \sup_{|\lambda - x \leq |y-x|} |b(\lambda)|^2\right)|y-x|,$$

i.e. V is locally Lipschitz.

The answer to the problem raised is given by the following. □

Theorem .

$$\lim_{\epsilon \to 0} u_\epsilon(x) = f(y_0)$$

where y_0 is assumed to be such that $y_0 \in \partial G$ and

$$V(x_0, y_0) < V(x,y),\ \forall y \in \partial G,\ y \neq y_0.$$

172

We first proceed to get an equivalent statement of the theorem. Let P_x^ϵ be the Brownian measure corresponding to the starting point x, and corresponding to the scaling ϵ. Then there exists a probability measure Q_x^ϵ such that

$$\frac{dQ_x^\epsilon}{dP_x^\epsilon}\Big|\mathscr{F}_t = Z(t)$$

where

$$Z(t,\cdot) = \exp \int_0^t \langle b^*(X(s)),\ dX(s)\rangle - \frac{1}{2}\int_0^t b^*(X(s))ds;$$

b^* is any bounded smooth function such that $b^* = b$ on G. Further we have the integral representation

$$u_\epsilon(x) = \int_{\partial G} f(X(\tau))dQ_x^\epsilon$$

where τ is the exit time of G, i.e.

$$\tau(w) = \inf\{t : w(t) \notin G\}.$$

$$|u_\epsilon(x) - f(y_0)| = |\int_{\partial G}(f(X(\tau)) - f(y_0))dQ_x^\epsilon|$$

$$\leq |\int_{N\cap\partial G} (f(X(\tau)) - f(Y_0))dQ_x^\epsilon| +$$

$$+|\int_{N^c\cap\partial G} (f(X(\tau)) - f(Y_0))dQ_x^\epsilon|$$

(N is any neighbourhood of y_0).

$$\leq Q_x^\epsilon(X(\tau) \in N \cap \partial G) \sup_{\lambda\in N\cap\partial G} |f(\lambda) - f(y_0)| +$$

$$+2\|f\|_\infty Q_x^\epsilon(X(\tau) \in N^c \cap \partial G).$$

173

Since f is continuous, to prove the theorem it is sufficient to prove the

Theorem .

$$\lim_{\epsilon\to 0} Q_x(X(\tau) \in N^c \cap \partial G) = 0$$

for every neighbourhood N of y_0.

Let N be any neighbourhood of y_0. Let

$$V = V(x_0, y_0), V' = \inf_{y \in N^c \cap \partial G} V(x, y).$$

By definition of y_0 and the fact that $N^c \cap \partial G$ is compact, we conclude that $V' > V$. Choose $\eta = \eta(N) > 0$ such that $V' = V + \eta$. For any $\delta > 0$ let $D = S(x_0; \delta) = \{y : |y - x_0| < \delta\}$, $\partial D = \{y : |y - x_0| = \delta\}$.

Claim. We can choose a δ_2 such that

(i) $V(x, y) \geq V + \dfrac{3\eta}{4}, \forall x \in \partial D_2, y \in N^c\ \partial G.$

(ii) $V(x, y_0) \leq V + \dfrac{\eta}{4}, \forall x \in \partial D_2.$

Proof. (i) $V(x_0, y) \geq V + \eta$, $\forall y \in N^c\ \partial G$. Therefore

$$\begin{aligned} V + \eta &\leq V(x_0, y) \leq V(x_0, x) + V(x, y) \\ &\leq C|x - x_0| + V(x, y). \end{aligned}$$

Choose C such that $C|x - x_0| \leq \dfrac{\eta}{4}$. Thus

$$V + \frac{3\eta}{4} \leq V(x, y) \text{ if } C|x - x_0| \leq \frac{\eta}{4},\ \forall y \in N^c\ \partial G.$$

174 C depends only on x_0. This proves (i).

(ii) $|V(x_0, y_0) - V(x, y_0)| \leq V(x_0, x) \leq C|x_0 - x| \leq \dfrac{\eta}{4}$ if x is close to x_0. Thus

$$V(x, y_0) \leq V(x_0, y_0) + \frac{\eta}{4} = V + \frac{\eta}{4}$$

if x is close to x_0. This can be achieved by choosing δ_2 very small. ∎

Claim (iii). We can choose $\delta_1 < \delta_2$ such that for points x_1, x_2 in ∂D_1 there is a path $X(\cdot)$ joining x_1, x_2 with $X(\cdot) \in D_2 - D_1$, i.e. it never penetrates D_1; and

$$I(X) \leq \frac{\eta}{8}.$$

Proof. Let $C = \sup\{|b(\lambda)|^2 : |\lambda - x_0| \leq \delta_2\}$. Choose $X(\cdot)$ to be any path on $[0, T]$, taking values in D_2 with $X(0) = x_1$; $X(T) = x_2$ and such that $|X| = 1$ (i.e. the path has unit speed). Then

$$\begin{aligned} I_0^T(X) &\leq \int_0^T (|X|^2 + C)dt \leq CT + \int_0^T |X|dt \\ &= (c+1)T = (C+1)|x_2 - x_1|. \end{aligned}$$

Choose δ_1 small such that $(C+1)|x_2 - x_1| \leq \dfrac{\eta}{8}$.

Let $\Omega\delta_1 = \{w : w(t) \in \overline{G} - D_1, \forall t \geq 0\}$, i.e. $\Omega\delta_1$ consists of all trajectories in $\overline{G}$ that avoid D_1. ◻

Claim (iv).

$$\inf_{T>0} \inf_{\substack{X \in \Omega_{\delta_1}, X(0) \in \partial D_2 \\ X(T) \in N^c \cap \partial G}} I_0^T(X(\cdot)) \geq V + \frac{3\eta}{4}$$

175

Proof. Follows from Claim (i) and (ii). ◻

Claim (v).

$$\inf_{T>0} \inf_{\substack{X \in \Omega_{\delta_1}, X(0) \in \partial D_2 \\ X(T) = y_0}} I_0^T(X(\cdot)) \leq V + \frac{3\eta}{8}.$$

Proof. By (ii) $V(x, y_0) \leq V + \dfrac{\eta}{4} \forall x \in \partial D_2$, i.e.

$$\inf_{T>0} \inf_{X(0)=x, X(T)=y_0} I_0^T(X(\cdot)) \leq V + \frac{\eta}{4}.$$

Let $\epsilon > 0$ be arbitrary. Choose T and $X(\cdot)$ such that $I_0^T(X) \leq V + \dfrac{\eta}{4} + \epsilon$ with $X(0) = x$, $X(T) = y_0$, $X(\cdot) \in \overline{G}$. If $X \in \Omega_{\delta_1}$ define $Y = X$. If $X \notin \Omega_{\delta_1}$ define Y as follows:

Let t_1 be the first time that X enters D_1 and t_2 the last time that it gets out of D_1. Then $0 < t_1 \leq t_2 < T$. Let X^* be a path on $[0, s]$ such that (by Claim (iii)) $I_0^s(X^*) \leq \dfrac{\eta}{8}$ with $X^*(0) = X(t_1)$ and $X^*(s) = X(t_2)$. Define Y on $[0, T - (t_2 - t_1) + s]$ $[T - (t_2 - t_1) + s, \infty)$ by

$$\begin{aligned} Y(t) = X(t) \text{ on } [0, t_1] &= X^*(t - t_1) \text{ on } [t_1, t_1 + s] \\ &= X(t - t_1 - s + t_1), \text{ on } [t_1 + s, T - (t_2 - t_1) + s] \\ &= X(t_2), \text{ for } t \geq T - (t_2 - t_1) + s. \end{aligned}$$

Then

$$\begin{aligned} I_0^{T - t_2 + t_1 + s} &= \frac{1}{2}\int_0^{t_1} |X - b(X(s))|^2 ds + \frac{1}{2}\int_0^{s} |X^* - b(X^*(s))|^2 ds \\ &\quad + \frac{1}{2}\int_{t_2}^{T} |X(s) - b(X(s))|^2 ds \\ &\leq V + \frac{\eta}{4} + \epsilon + \frac{\eta}{8} \end{aligned}$$

176 by choice of X and X^*. As $Y \in \Omega_{\delta_1}$, we have shown that

$$\inf_{T > 0} \inf_{\substack{X \in \Omega_{\delta_1}, X(0) \in \partial D_1 \\ X(T) = y_0}} I_0^T(X(\cdot)) \leq V + \frac{3\eta}{8} + \epsilon.$$

Since ϵ is arbitrary we have proved (v). ◻

Lemma . *I_0^T is lower semi-continuous for every finite T.*

Proof. This is left as an exercise as it involves a repetiti on of an argument used earlier. ◻

Lemma . *Let $X_n \in \Omega_{\delta_1}$. If $T_n \to \infty$ then $I_0^{T_n}(X_n) \to \infty$.*

This result says that we cannot have a trajectory which starts outside of a deleted ball for which I remains finite for arbitrary long lengths of time.

Proof. Assume the contrary. Then there exists a constant M such that $I_0^{T_n}(X_0) \le M$, $\forall n$. Let $T < \infty$, so that $M_T = \sup_n I_0^T(X_n) < \infty$.

Define $X_n^T = X_n|_{[0,T]}$. □

Claim. $\{X_n^T\}_{n=1}$ is an equicontinuous family. **177**

Proof.

$$|X_n^T(x_2) - X_n^T(x_1)|^2 = |\int_{x_1}^{x_2} X_n^T(t)dt|^2$$
$$\le |x_2 - x_1|^2 \int_{x_1}^{x_2} |X_n^T|^2 dt$$
$$\le 2|x_2 - x_1|^2 \int^{x_2} |X_n^T - b(X_n^T)|^2 ds + \int_0^T b(X_n^T)^2 ds]$$
$$\le 2|x_2 - x_1|^2 [2M_T + T\|b\|_\infty^2].$$

Thus, $\{X_n^T\}_n$ is an equicontinuous family. Since $\overline{G}$ is bounded, $\{X_n^T\}_n$ is uniformly bounded. By Arzela-Ascoli theorem and a "diagonal procedure" there exists a subsequence X_{n_k} and a continuous function to X uniformly on compact subsets of $[0, \infty)$. As $X_{n_k}(\cdot) \in \overline{G} - D_1$, $X \in \overline{G} - D_1$. Let $m \ge n$. $I_0^{T_n}(X_m) \le M$. $X_n \to X$ uniformly on $[0, T_n]$. By lower semicontinuity $I_0^T(X) \le M$. Since this is true for every T we get on letting T tend to ∞, that

$$\frac{1}{2} \int_0^\infty |X - b(X(s))|^2 ds \le M.$$

Thus we can find a sequence $a_1 < b_1 < a_2 < b_2 < \ldots$ such that

$$I_{a_n}^{b_n}(X(\cdot)) = \frac{1}{2} \int_{a_n}^{b_n} |X(t) - b(X(t))|^2 dt$$

converges to zero with $b_n - a_n \to \infty$. Let $Y_n(t) = X(t + a_n)$. Then

$$I_0^{b_n - a_n}(Y_n) \to 0 \quad \text{with} \quad b_n - a_n \to +\infty, \quad Y_n \in \Omega_{\delta_1}.$$

178 Just as X was constructed from X_n, we can construct Y from Y_n such that $Y_n \to Y$ uniformly on compact subsets of $[0, \infty)$.

$$I_0^{b_n - a_n}(Y) \leq \inf_{m \geq n} I_0^{b_n - a_n}(Y_m) = 0$$

(by lower semi-confirmity of I_0^T). Thus $I_0^{b_n - a_n}(Y) = 0$, $\forall n$, showing that

$$\int\limits_0^\infty Y(t) - b(Y(t))|^2 dt = 0$$

Thus Y satisfies $Y(\cdot) = x + \int\limits_0^t b(Y(s))ds$ with $Y(t) \in \overline{G} - \partial D_1$, $\forall t$.

Case (i). $Y(t_0) \in G$ for some t_0. Let $Z(t) = Y(t + t_0)$ so that Z is an integral curve starting at a point of G and remaining away from D_1 contradicting the stability condition.

Case (ii). $Y(t_0) \notin G$ for any t_0, i.e. $Y(t) \in \partial G$ for all t. Since $Y(t) = b(Y(t))\langle Y(t), n(Y(t))\rangle$ is strictly positive. But $Y(t) \in \partial G$ and hence $\langle Y(t), n(Y(t))\rangle = 0$ which leads to a contradiction. Thus our assumption is incorrect and hence the lemma follows. ◻

Lemma . *Let $x \in \partial D_2$ and define*

$$E = \{X(t) \textit{ exits from } G \textit{ before hitting } D_1 \textit{ and it exits from } N\}$$
$$F = \{X(t) \textit{ exists from } G \textit{ before hitting } D_1 \textit{ and it exits from } N^c\}$$

Then

$$\frac{Q_x^\epsilon(F)}{Q_x^\epsilon(E)} \leq \exp\left(-\frac{3\eta}{8\epsilon} + 0\left(\frac{1}{\epsilon}\right)\right) \to 0 \textit{ uniformly in } x(x \in \partial D_2).$$

179 **Significance.** $Q_x^\epsilon(E)$ and $Q_x^\epsilon(F)$ are both small because diffusion is small and the drift is large. The lemma says that $Q_x^\epsilon(E)$ is relatively much larger than $Q_x^\epsilon(F)$.

Proof. $Q^\epsilon_x(E) \geq Q^\epsilon_x\{X(t)$ exists from G before hitting D_1, and exists in N before time $T\}$, $= Q_x(B) \geq \exp[-\dfrac{1}{\epsilon}\inf I^T_0(X(\cdot))]$ where the infimum is taken over the interior of B,

$$\geq \exp\left[-\frac{1}{\epsilon}\left(V+\frac{3\eta}{8}\right)+0\left(\frac{1}{\epsilon}\right)\right].$$

Similarly,

$$Q^\epsilon_x(F) \leq \exp\left[-\frac{1}{\epsilon}\left(V+\frac{3\eta}{4}\right)+0\left(\frac{1}{\epsilon}\right)\right].$$

Therefore

$$\frac{Q^\epsilon_x(F)}{Q^\epsilon_x(E)} \leq \exp\left[-\frac{3\eta}{8\epsilon}+0\left(\frac{1}{\epsilon}\right)\right] \to 0 \quad \text{as} \quad \epsilon \to 0.$$

We now proceed to prove the main theorem. Let

$$\begin{aligned}
\tau_0 &= 0,\\
\tau_1 &= \text{first time } \partial D_1 \text{ is hit},\\
\tau_2 &= \text{next time } \partial D_2 \text{ is hit},\\
\tau_3 &= \text{next time } \partial D_1 \text{ is hit},\\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}$$

and so on. Observe that the particle can get out of G only between the time intervals τ_{2n} and τ_{2n+1}. Let $E_n = \{$between τ_{2n} and τ_{2n+1} the path exits from G for the first time and that it exits in $N\}$, $F_n = \{$between τ_{2n} **180**
and τ_{2n+1} the path exits from G for the first time and that it exists in $N^c\}$.

$$Q^\epsilon_x(X(\tau) \in N) + Q_x(X(\tau) \in N^c) = 1.$$

Also

$$Q^\epsilon_x(X(\tau) \in N^c) = \sum_{n=1}^\infty Q^\epsilon_x(F_n),$$

$$Q^\epsilon_x(X(\tau) \in N) = \sum_{n=1}^\infty Q^\epsilon_x(E_n),$$

$$
\begin{aligned}
\sum_{n=1}^{\infty} Q_x^\epsilon(F_n) &= \sum_{n=1}^{\infty} E^{Q_x^\epsilon}(Q_x^\epsilon(F_n|\mathscr{F}_{\tau_{2n}})) \\
\leq \sum_{n=1} E^{Q_x^\epsilon}[\chi_{(\tau>\tau_{2n})} \sup_{x\in\partial D_2} Q_x^\epsilon(F)] &\quad \text{(by the Strong Markov property)} \\
\leq 0(\epsilon)\sum_{n=1}^{\infty} E^{Q_x^\epsilon}[\chi_{(\tau>\tau_{2n})} \inf_{x\in\partial D_2} Q_x^\epsilon(E)] &\quad \left(\text{as } \frac{Q_x^\epsilon(F)}{Q_x^\epsilon(E)} \to 0\right) \\
\leq 0(\epsilon)\sum_{n=1}^{\infty} Q_x^\epsilon(E_n) &= 0(\epsilon) Q_x(X(\tau)\in N^c).
\end{aligned}
$$

Therefore

$$
Q_x^\epsilon(\chi(\tau)\in N)\to 1, \quad Q_x(X(\tau)\in N^c)\to 0.
$$

□

Exercise. Suppose $b(x)=\nabla u(x)$ for some $u\in C^1(G\cup\partial G, R)$. Assume that $u(x_0)=0$ and $u(x)<0$ for $x\neq x_0$. Show that

$$
V(x_0,y) = -2u(y).
$$

[Hint: For any trajectory X with $X(0)=x_0$,

$$
X(T)=y, I_0^T(X) = \frac{1}{2}\int_0^T |X+\nabla u(X)|^2 dt - 2\int_0^T \nabla u(X)\cdot X(t)dt \geq -2u(y)
$$

181 so that $V(x_0,y)\geq -2u(y)$. For the other inequality, let u be a solution of $X(t)+\nabla u(X(t)=0$ on $[0,\infty)$ with $X(0)=y$. Show that because $\dfrac{duX(s)}{ds}0$ for $X(s)\neq 0$ and x_0 is the only zero of u, $\lim_{t\to\infty}X(t)=x_0$. Now conclude that $V(x_0,y)\leq -u(y)$].

23. Stochastic Integral for a Wider Class of Functions

WE SHALL NOW define the stochastic integral for a wider class of functions. 182

Let $\theta : [0,) \times \Omega \to \mathbb{R}^d$ be any progressively measurable function such that for every t

$$\int\limits_0^t |\theta(s,w)|^2 ds < \infty, \quad \text{a.e.}$$

Define, for every finite $L \geq 0$,

$$\theta_L(s,w) = \begin{cases} \theta(s,w), & \text{if } \int\limits_0^s |\theta(t,w)|^2 dt < L < \infty, \\ 0, & \text{if } \int\limits_0^s |\theta(t,w)|^2 dt \geq L. \end{cases}$$

We can write $\theta_L(s,w) = \theta(s,w)\chi_{[0,L)}(\phi(s,w))$ where

$$\phi(s,w) = \int\limits_0^s |\theta(t,w)|^2 dt$$

is progressively measurable. Hence $\theta_L(s,w)$ is progressively measurable. It is clear that $\int\limits_0^T |\theta_L(s,w)|^2 ds \leq L$, a.e. $\forall T$. Therefore

$$E(\int\limits_0^T |\theta_L(s,w)|^2 ds) \leq L.$$

Thus the stochastic integral $\xi_L(t,w) = \int\limits_0^t \langle \theta_L(s,w), dX(s,w)\rangle$ is well defined.

The proofs of the next three lemmas follow closely the treatment of Stochastic integration given earlier.

Lemma 1. *Let τ be a bounded, progressively measurable, continuous*
183 *function. Let τ be any finite stopping time. If $\theta(s,w) = 0$, $\forall (s,w)$ such that $0 \leq s \leq \tau(w)$ then $\int\limits_0^t \langle \theta(s,w), dX(s,w)\rangle = 0$ for $0 \leq t \leq \tau(w)$.*

Proof. Define $\theta_n(s,w) = \theta(\frac{[ns]}{n}, w)$. θ_n is progressively measurable and by definition of the stochastic integral of θ_n,

$$\int\limits_0^t \langle \theta_n(s,w), dX(s,w)\rangle = 0, \quad \forall t, \quad 0 \leq t \leq \tau(w).$$

Now

$$E\left(\int\limits_0^t |\theta_n(s,w) - \theta(s,w)|^2 ds\right)$$

$$= E\left(\int\limits_0^t |\theta\left(\frac{[ns]}{n} w\right) - \theta(s,w)|^2 ds\right) \to 0 \text{ as } n \to \infty$$

and

$$\int\limits_0^t \langle \theta_n(s,w), dX(s,w)\rangle \to \int\limits_0^t \langle \theta(s,w), dX(s,w)\rangle$$

in probability. Therefore

$$\int_0^t \langle \theta, dX \rangle = 0 \text{ if } 0 \leq t \leq \tau(w).$$

□

Lemma 2. *If θ is progressively measurable and bounded the assertion of Lemma 1 still holds.*

Proof. Let

$$\theta_n(t,w) = \frac{1}{n} \int_{(t-1/n)V_0}^{t} \theta(s,w) ds.$$

Then

$$E\left(\int_0^T |\theta_n(t,w) - \theta(t,w)|^2 dt\right) \to 0 \quad \text{(Lebesgue's theorem)}.$$

θ_n is continuous and boundd, $\theta_n(t,w) = 0$ for $0 \leq t \leq \tau(w)$. By lemma 1 **184**

$$\int_0^t \langle \theta_n(s,w), dX(s,w) \rangle = 0$$

if $0 \leq t \leq \tau(w)$. This proves the result. □

Lemma 3. *Let θ be progressively measurable such that, for all t,*

$$E\left(\int_0^t |\theta(s,w)|^2 ds\right) < \infty.$$

If $\theta(s,w) = 0$ for $0 \leq s \leq \tau(w)$, then

$$\int_0^t \langle \theta(s,w), dX(s,w) \rangle = 0 \quad \textit{for} \quad 0 \leq t \leq \tau(w).$$

Proof. Define

$$\theta_n = \begin{cases} \theta, & \text{if } |\theta| \leq n, \\ 0, & \text{if } |\theta| > n. \end{cases}$$

Then

$$\int\limits_0^c \langle \theta_n, dX \rangle = 0, \text{ if } 0 \leq t \leq \tau(w), \quad \text{(Lemma 2) and}$$

$E(\int\limits_0^t |\theta_n - \theta|^2 ds) \to 0$. The result follows. □

Lemma 4. *Let θ be progressively measurable such that $\forall t$,*

$$\int\limits_0^t |\theta(s, w)|^2 ds < \infty \quad a.e.$$

Then $\underset{L \to \infty}{\mathrm{Lt}}\ \xi_L(t, w)$ exists a.e.

185 *Proof.* Define

$$\tau_L(w) = \inf \left\{ s : \int\limits_0^s |\theta(\sigma, w)|^2 d\sigma \geq L \right\};$$

clearly τ_L is a stopping time. If $L_1 \leq L_2$, $\tau_{L_1}(w) \leq \tau_{L_2}(w)$ and by assumptions of the lemma $\tau_L \uparrow \infty$ as $L \uparrow \infty$. If

$$L_1 \leq L_2, \quad \theta_{L_1}(s, w) = \theta_{L_2}(s, w) \quad \text{for} \quad 0 \leq s \leq \tau_{L_1}(w).$$

Therefore by Lemma 3,

$$\xi_{L_2}(t, w) = \xi_{L_1}(t, w)$$

if $0 \leq t \leq \tau_{L_1}(w)$. Therefore as soon as L is large enough such that $t \leq \tau_L(w)$, $\xi_L(t, w)$ remains constant (as a function of L). Therefore $\underset{L \to \infty}{\mathrm{Lt}}\ \xi_L(t, w)$ exists a.e. □

Definition. The stochastic integral of θ is defined by

$$\int_0^t \langle \theta(s,w), dX(s,w)\rangle = \operatorname*{Lt}_{L\to\infty} \xi_L(t,w).$$

Exercise. Check that the definition of the stochastic integral given above coincides with the previous definition in case

$$E\left(\int_0^t |\theta(s,w)|^2 ds\right) < \infty, \ \forall t.$$

Lemma . *Let θ be a progressively measurable function, such that*

$$\int_0^t |\theta(s,w)|^2 ds < \infty, \ \forall t.$$

If $\xi(t,w)$ denotes the stochastic integral of θ, then **186**

$$P\left(\sup_{0\le t\le T} |\xi(t,\cdot)| \ge \epsilon\right) \le P\left(\int_0^T |\theta|^2 ds \ge L\right) + \frac{L^2}{\epsilon^2}.$$

Proof. Let $\tau_L = \inf\{t : \int_0^t |\theta|^2 ds \ge L\}$. If $T < \tau_L(w)$, then $\theta_L(s,w) = \theta(s,w)$. Also

$$\xi_L(t,w) = \xi(t,w) \quad \text{for} \quad 0 \le t \le T.$$

$\square$

Claim.

$$\left\{w : \sup_{0\le t\le T} |\xi(t,w)| \ge \epsilon\right\}$$
$$\left\{w : \sup_{0\le t\le T} |\xi_L(t,w)| \ge \epsilon\right\} \cup \{w : \tau_L(w) \le T\}$$

For, if w is not contained in the right side, then

$$\sup_{0\le t\le T} |\xi_L(t,w)|^2 < \epsilon \quad \text{and} \quad |\tau_L(w)| > T.$$

If $\tau_L > T$, $\xi_L(t,w) = \xi(t,w)\ \forall t \le T$. Therefore

$$\sup_{0\le t\le T} |\xi_L(t,w)| = \sup_{0\le t\le T} |\xi(t,w)|$$

Therefore $w \notin$ left side. Since

$$\{w : \tau_L(w) > T\} = \left\{ w : \int\limits_0^T |\theta|^2 ds \ge L \right\}$$

we get

$$\begin{aligned} &P\left(\sup_{0\le t\le T} |\xi(t,\cdot)| \ge \epsilon\right)\\ &\le P\left(\int\limits_0^T |\theta|^2 ds \ge L\right) + P\left(\sup_{0\le t\le T} |\xi_L(t,\cdot)| \ge \epsilon\right)\\ &\le P\left(\int\limits_0^T |\theta|^2 ds \ge L\right) + \frac{L^2}{\epsilon^2}, \end{aligned}$$

187 by Kolmogorov's inequality. This proves the result.

Corollary . *Let θ_n and θ be progressively measurable functions such that*

(a) $\int\limits_0^t |\theta_n(s,w)|^2 ds < \infty$, $\int\limits_0^t |\theta(s,w)|^2 ds < \infty$, $\forall t$;

(b) $\mathop{\mathrm{Lt}}\limits_{n\to\infty} \int\limits_0^t |\theta_n(s,w) - \theta(s,w)|^2 ds = 0$ *in probability.*

If $\xi_n(t,w)$ and $\xi(t,w)$ denote, respectively the stochastic integrals of θ_n and θ, then $\sup\limits_{0\le t\le T} |\xi_n(t,w) - \xi(t,w)|$ converges to zero in probability.

Proof. Let $\tau_{n,L}(w) = \inf\{t : \int\limits_0^t |\theta_n|^2 ds \geq L\}$; replacing θ by $\theta_n - \theta$ and ξ by $\xi_n - \xi$ in the previous lemma, we get

$$P\left(\sup_{0\leq t\leq T} |\xi_n(t,\cdot) - \xi(t,\cdot)| \geq \epsilon\right)$$
$$\leq \frac{L^2}{\epsilon^2} + P\left(\int\limits_0^T |\theta_n(s,\cdot) - \theta(s,\cdot)|^2 ds \geq L\right).$$

Letting $n \to \infty$, we get

$$\lim_{n\to\infty} P\left(\sup_{0\leq t\leq T} |\xi_n(t,\cdot) - \xi(t,\cdot)| \geq \epsilon\right) \frac{L^2}{\epsilon^2}.$$

As L is arbitrary we get the desired result. □

Proposition . *Let θ be progressively measurable such that*

$$\int\limits_0^t |\theta(s,w)|^2 ds < \infty, \quad \forall t \text{ and } \forall\, w.$$

Then **188**

$$(*) \qquad Z(t,\cdot) = \exp\left[\int\limits_0^t \langle\theta(s,\cdot), dX(s,\cdot)\rangle - \frac{1}{2}\int\limits_0^t |\theta(s,\cdot)|^2 ds\right]$$

is a super martingale satisfying

(a) $E(Z(t,\cdot)) \leq 1$;

(b) $\operatorname*{Lt}_{t\to 0} E(Z(t,\cdot)) = 1$.

Proof. Let (θ_n) be a sequence of bounded progressively measurable functions such that

$$\int\limits_0^t |\theta_n - \theta|^2 ds \to 0\ \forall t, \quad \forall w.$$

(For example we may take $\theta_n = \theta$ if $|\theta| < n$, $= 0$ otherwise). Then $(*)$ is a martingale when θ is replaced by θ_n. This martingale satisfies $E(Z_n(t, \cdot)) = 1$, and $Z_n(t, \cdot) \to Z(t, \cdot)$ pointwise (a) now follows from Fatou's lemma:

$$\begin{aligned}\lim_{t \to 0} E(Z(t, \cdot)) &\geq E\left(\varliminf_{t \to 0} Z(t, \cdot)\right)\\ &= E(1) = 1.\end{aligned}$$

Therefore $\underset{t \to 0}{\mathrm{Lt}}\, E(Z(t, \cdot)) = 1$. This proves (b). $\square$

24. Explosions

Exercise. Let $R > 0$ be given. Put $b_R = b\phi_R$ where $\phi_R = 1$ on $|x| \geq R$, 189
$\phi_R = 0$ if $|x| \geq R+1$; ϕ_R is C^∞. Show that $b_R = b$ on $|x| \leq R$, b_R is bounded on $\mathbb{R}^d$ and b_R is globally Lipschitz.

Let $\Omega_T = \{w \in \Omega : B(w) > T\}$. Let $S^T = \Omega_T \to C[0,T]$ be the map $S^T w = y(\cdot)$ where $y(t) = w(t) + \int\limits_0^t b(y(s))ds$ on $[0,T]$. Unless otherwise specifie $b : \mathbb{R}^d \to \mathbb{R}^d$ is assumed to be locally Lipschitz. Define the measure Q_x^T on (Ω, T) by

$$Q_x^T(A) = P_x\{w : S^T w \in A, B(w) > T\},$$

where P_x is the probability measure corresponding to Brownian motion.

Theorem .

$$Q_x^T(A) = \int\limits_A Z(T,\cdot)dP_x, \quad \forall A \in \mathscr{F}_T,$$

where

$$Z(T,\cdot) = \exp\left[\int\limits_0^T \langle b, dX\rangle - \frac{1}{2}\int\limits_0^T |b(X(s,\cdot))|^2 ds\right].$$

Remark. If b is bounded or if b satisfies a global Lipschitz condition then $B(w) = \infty$, so that $\Omega_T = \Omega$ and Q_x^T are probability measures.

Proof. Let $0 \leq R < \infty$. For any w in Ω, let y be given by

$$y(t) = w(t) + \int\limits_0^t b(y(\sigma))d\sigma.$$

Define $\sigma_R(w) = \inf\{t : |y(t)| \geq R$ and let b_R be as in the Exercise. Then the equation

$$y_R(t) = w(t) + \int_0^t b_R(y_R(\sigma))d\sigma$$

190 has a global solution. Denote by $S_R : \Omega \to \Omega$ the map $w \to y_R$. If $Q_{R,x}$ is the measure induced by S_R, then

$$\frac{dQ_{R,x}}{dP_x}\Big|\mathscr{F}_t = Z_r(t) = \exp\left(\int_0^t \langle b_R, dX\rangle - \frac{1}{2}\int_0^t |b_R|^2 ds\right).$$

Let $\tau_R(w) = \inf\{t : |w(t)| > R\}$. τ_R is a stopping time satisfying $\tau_R S_R = \sigma_R$. By the optional stopping theorem.

$$\frac{dQ_{R,x}}{dP_x}\Big|\mathscr{F}_{\tau_R\wedge T} = Z_R(\tau_R \wedge T) = Z(\tau_R \wedge T). \tag{1}$$

□

Claim. $Q_{R,x}((\tau_R > T) \cap A) = Q_x^T((\tau_R > T) \cap A),\ \forall A$ in $\mathscr{F}_T$.

Proof.

$$\begin{aligned}
\text{Right side } &= P_x\{w : B(w) > T, S^T(w) \in A \cap (\tau_R > T)\}\\
&= P_x\{w : B(w) > T, y \in A, \sup_{0\leq t\leq T} |y(t)| < R\}\\
&= P_x\{w : y \text{ is defined at least upto time } T,\\
&\qquad y \in A, \sup_{0\leq t\leq T} |y(t)| > R\}\\
&= P_x\{w : y_R \in A, \sup_{0\leq t\leq T} |y_R(t)| < R\}\\
&= P_x\{w : S_R(w) \in A, \tau_R S_R(w) > T\}\\
&= Q_{R,x}\{(\tau_R > T) \cap A\}
\end{aligned}$$

(by definition). As Ω is an increasing union of $\{\tau_R > T\}$ for R increasing,

$$Q_x^T(A) = \operatorname*{lt}_{R\to+\infty} Q_x^T((\tau_R > T) \cap A),\ \forall A \text{ in } \mathscr{F}_T,$$

$$= \operatorname*{lt}_{R\to\infty} Q_{R,x}((\tau_R > T)\cap A) \quad \text{(by claim)}$$

$$= \operatorname*{lt}_{R\to\infty} \int\limits_{(\tau_R\wedge T)\cap A} \exp\left(\int\limits_0^{\tau_R\wedge T} \langle b, dX\rangle - \frac{1}{2}\int\limits_0^{\tau_R\wedge T} |b|^2 ds\right) dP_x \quad \text{(by (1))}$$

$$= \int\limits_A \exp\left(\int\limits_0^T \langle b, dX\rangle - \frac{1}{2}\int\limits_0^T |b|^2 ds\right) dP_x$$

$$= \int\limits_A Z(T) dP_x.$$

□ **191**

Theorem . *Suppose* $b : \mathbb{R}^d \to \mathbb{R}^d$ *is locally Lipschitz; let* $L = \dfrac{\Delta}{2} + b.\nabla$.

(i) *If there exists a smooth function* $u : \mathbb{R}^d \to (0,\infty)$ *such that* $u(x) \to \infty$ *as* $|x| \to \infty$ *and* $Lu \leq cu$ *for some* $c > 0$ *then* $P_x\{w : B(w) < \infty\} = 0$, *i.e. for almost all w there is no explosion.*

(ii) *If there exists a smooth bounded function* $u : \mathbb{R}^d \to (0,\infty)$ *such that* $Lu \geq cu$ *for some* $c > 0$, *then* $P_x\{w : B(w) < \infty\} > 0$, *i.e. there is explosion.*

Corollary . *Suppose, in particular, b satisfies* $|\langle b(x), x\rangle| \leq A + B|x|^2$ *for some constants A and B; then* $P_x(w : B(w) < \infty) = 0$.

Proof. Take $u(x) = 1 + |x|^2$ and use part (1) of the theorem. □

Proof of theorem. Let b_R be as in the Exercise and let $L_R = \dfrac{\Delta}{2} + b_R \cdot \nabla$; then $L_R u(x) \leq cu(x)$ if $|x| \leq R$.

Claim. $u(X(t))e^{-ct}$ is a supermartingale upto time τ_R relative to Q_x^R,

$$d\left(u(X(t))e^{-ct} \exp\left(\int\limits_0^t \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^t |b_R|^2 ds\right)\right)$$

$$e^{-ct\int\limits_0^t \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^t |b_R|^2 ds} \times$$

$$\times\{-cudt + \langle \nabla u, dX\rangle + u(x)[\langle b_R, dX\rangle - \frac{|b_R|^2}{2}dt] + b_R udt + \frac{1}{2}|b_R|^2 udt\}$$

$$= \exp(-ct + \int\limits_0^t \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^t |b_R|^2 ds)$$

$$\cdot[L_R - c)u + \langle \nabla u, dX\rangle + u\langle b_R, dX\rangle].$$

192 Therefore

$$u(X(t))e^{-ct}E^{e^0 \int\limits^t \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^t |b_R|^2 ds}$$

$$-\int\limits_0^t \exp\left(-cs + \int\limits_0^s \langle b_R, dX\rangle - \int\limits_0^s |b_R|^2 ds\right) \cdot (L_R - c)u(X(s))ds$$

is a Brownian stochastic integral. Therefore

$$u(X(\tau_R \wedge t)) \exp\left(-c(\tau_R \wedge t) + \int\limits_0^{\tau_R\wedge t} \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^{\tau_R\wedge t} |b_R|^2 ds\right) -$$

$$-\int\limits_0^{\tau_R\wedge t} \exp\left(-cs + \int\limits_0^s \langle b_R, dX\rangle - \frac{1}{2}\int\limits_0^s |b_R|^2 ds\right)(L_R - c)u(X(s))ds$$

is a martingale relative to P_x, $\mathscr{F}_{\tau_R\wedge t}$. But $b_R(x) = b(x)$ if $|x| \le R$. Therefore

$$u(X(\tau_R \wedge t)) \exp\left(-c(\tau_R \wedge t) + \int\limits_0^{\tau_R\wedge t} \langle b, dX\rangle - \frac{1}{2}\int\limits_0^{\tau_R\wedge t} |b|^2 ds\right) -$$

$$-\int\limits_0^{\tau_R\wedge t} \exp\left(-cs + \int \langle b, dX\rangle - \int\limits_0^s |b|^2 ds\right)(L_R - c)u(X(s))ds$$

is a martingale relative to $\mathscr{F}_{\tau_R\wedge t}$. But $(L_R - c)u \leq 0$ in $[0, \tau_R]$.

193 Therefore

$$u(X(\tau_R \wedge t)) \exp(-c(\tau_R \wedge t) + \int_0^{\tau_R\wedge t} \langle b, dX\rangle - \frac{1}{2}\int_0^{\tau_R\wedge t} |b|^2 ds)$$

is a supermartingale relative to $\mathscr{F}_{\tau_R\wedge t}, P_x$. Therefore $u(X(\tau_R \wedge t)e^{-c(\tau_R\wedge t)}$ is a supermartingale relative to Q_x^R (optional sampling theorem). Therefore

$$E^{Q_x^R}(u(X(t_R \wedge t))e^{-c(\tau_R\wedge t)} \leq u(x);$$

letting $t \to \infty$, we get, using Fatou's lemma,

$$E_x^{Q^R}(u(X(\tau_R)e^{-c\tau_R}) \leq u(x).$$

Therefore

$$E^{Q_x^R}(e^{-c\tau_R}) \leq \frac{u(x)}{\inf\limits_{|y|=R} |u(y)|}.$$

Thus

$$E^{P_x}(e^{-c\sigma_R}) \leq \frac{u(x)}{\inf\limits_{|y|=R} |u(y)|}$$

(by change of variable). Let $R \to \infty$ to get $\underset{R\to\infty}{\mathrm{Lt}} \int e^{-c\sigma_R} dP_x = 0$, i.e. $P_x\{w : B(w) < \infty\} = 0$.

Sketch of proof for Part (ii).

By using the same technique as for Part (i), show that $u(X(t))e^{-ct}$ is a submartingale upto time τ_R relative to Q_x^R, so that

$$E^{P_x}(e^{-c\sigma_R}) \geq \frac{u(x)}{\sup\limits_{|y|=R} |u(y)|} \geq \frac{u(x)}{||u||_\infty} > 0;$$

let $R \to \infty$ to get the result. 194

Exercise. Show that if $L = \dfrac{1}{2}\dfrac{\partial^2}{\partial x} + x^3 \dfrac{\partial}{\partial x}$, there is explosion. (Hint: take $u = e^{\tan^{-1}(x^2)}$ and show that $Lu \geq u$).

25. Construction of a Diffusion Process

Problem. Given $a : [0,\infty) \times \mathbb{R}^d \to S_d^+$, bounded measurable and $b :$ 195
$[0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ bounded measurable, to find $(\Omega, \mathscr{F}_t, P, X)$ where Ω is a space, $\{\mathscr{F}_t\}_{t \geq 0}$ an increasing family of σ-algebras on Ω, P a probability measure on the smallest σ-algebra containing all the $\mathscr{F}_t$'s. and $X : [0,t) \times \Omega \to \mathbb{R}^d$, a progressively measurable function such that $X(t,w) \in I[b(t,X_t), a(t,X_t)]$.

Let $\Omega = C[0,\infty); \mathbb{R}^n)$, $\beta(t,\cdot) = n$-dimensional Brownian motion, $\mathscr{F}_t = \sigma\{\beta(s) : 0 \leq s \leq t\}$, P the Brownian measure on Ω and a and b as given in the problem. We shall show that thee problem has a solution, under some special conditions on a and b.

Theorem . *Assume that there exists $\sigma : [0,\infty) \times \mathbb{R}^d \to M_{d\times n}$ ($M_{d\times n}$= set of all $d \times n$ matrices over the reals) such that $\sigma\sigma^* = a$. Further let*

$$\sum_{i,j} |\sigma_{ij}(t,x)| \leq C, \quad \sum_j |b_j(t,x)| \leq C,$$
$$\sum_{i,j} |\sigma_{ij}(t,x_1) - \sigma_{ij}(t,x_2)| \leq A|x_1 - x_2|,$$
$$\sum_j |b_j(t,x_1) - b_j(t,x_2)| \leq A|x_1 - x_2|.$$

Then the equation

$$\text{(1)} \qquad \xi(t,\cdot) = x + \int\limits_0^t \langle \sigma(s,\xi(s,\cdot)), d\beta(s,\cdot)\rangle + \int\limits_0^t b(s,\xi(s,\cdot))ds$$

196 *has a solution. The solution $\xi(t,w) : [0,\infty)\times\Omega \to \mathbb{R}^d$ can be taken to be such that $\xi(t,\cdot)$ is progressively measurable and such that $\xi(t,\cdot)$ is continuous for a, a.e. If ξ, η are progressively measurable, continuous (for a.a.e) solutions of equation (l), then $\xi = n$ for a.a.w.*

Proof. The proof proceeds in several steps. □

Lemma 1. *Let Ω be any space with $(\mathscr{F}_t)_{t\geq 0}$ an increasing family of σ-algebras. If $0 \leq T \leq \infty$ then there exists a σ-algebra $\mathscr{A}_0 \subset \mathscr{A} = \mathscr{B}[0,T)\times\mathscr{F}_T$ such that a function $f : [0,T]\times\Omega \to \mathbb{R}$ is progressively measurable if and only if f is measurable with respect to $\mathscr{A}_0$.*

Proof. Let $\mathscr{A}_0 = \{A \in \mathscr{A} : \chi_A$ is progressively measurable$\}$. Clearly $[0,T]\times\Omega \in \mathscr{A}_0$, and if $A \in \mathscr{A}_0$, $A^c \in \mathscr{A}_c$. Thus $\mathscr{A}_0$ is an algebra. As increasing limits (decreasing limits) of progressively measurable functions are progressively measurable, $\mathscr{A}_0$ is a monotone class and hence a σ-algebra. □

Let $f : [0,T]\times\Omega \to \mathbb{R}$ be progressively measurable; in fact, $f^+ = \dfrac{f+1=f|}{2}$, $f^- = \dfrac{f-|f|}{2}$. Let $g = f^+$. Then

$$g_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n}\chi_{g^{-1}[\frac{i-1}{2^n},\frac{i}{2^n})} + n\chi_{g^{-1}[n,)}$$

is progressively measurable. Hence nVg_n is progressively measurable,
197 i.e. $n\chi_{g^{-1}[n,\infty)}$ is progressively measurable. Similarly $\phi_{g^{-1}[\frac{i-1}{2^n},\frac{i}{2^n})}$ is progressively measurable, etc. Therefore, by definition, g_n is measurable with respect to $\mathscr{A}_0$. As $g = f^+$ is the pointwise limit of g_n, f^+ is measurable with respect to $\mathscr{A}_0$. Similarly f^- is $\mathscr{A}_0$-measurable. Thus f is $\mathscr{A}_0$-measurable.

Let $f : [0,T] \times \Omega \to \mathbb{R}$ be measurable with respect to $\mathscr{A}_0$. Again, if $g = f^+$

$$g_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{g^{-1}[\frac{i-1}{2^n}, \frac{i}{2^n})} + n\chi_{g^{-1}[n,\infty)}$$

is $\mathscr{A}_0$-measurable. Since $g^{-1}[n,\infty), \ldots g^{-1}[\dfrac{i-1}{2^n}, \dfrac{i}{2^n}) \in \mathscr{A}_0$. So g_n is progressively measurable. Therefore g is progressively measurable. Hence f is progressively measurable. This completes the proof of the Lemma.

To solve (1) we use the standard iteration technique.

Step 1. Let $\xi_0(t,w) = x$,

$$\xi_n(t,w) = x + \int_0^t \langle \sigma(s, \xi_{n-1}(s,w)), d\beta(s,w)\rangle + \int_0^t b(s, \xi_{n-1}(s,w)) ds.$$

By induction, it follows that $\xi_n(t,w)$ is progressively measurable.

Step 2. Let $\Delta_n(t) = E(|\xi_{n+1}(t) - \xi_n(t)|^2)$. If $0 \leq t \leq T$, $\Delta_n(t) \leq C^* \int_0^t \Delta_{n-1}(s) ds$ and $\Delta_0(t) \leq C^* t$, where C^* is a constant depending only on T.

Proof.

$$\begin{aligned}
\Delta_0(t) &= E(|\xi(t) - x|^2) \\
&= E\left(|\int_0^t \langle (s,x), d\beta(s,x)\rangle + \int_0^t b(s,x) ds|^2 \right) \\
&\leq 2E\left(|\int_0^t \langle \sigma(s,x), d\beta(s,x)\rangle|^2 \right) + \\
&\quad + 2E\left(|\int_0^t b(s,x) ds|^2 \right) \quad \text{(use the fact that } |x+y|^2 \\
&\quad \leq 2(|x|^2 + |y|^2)\ \forall x, y \in \mathbb{R}^d)
\end{aligned}$$

$$= 2E\left(\int_0^t Tr\ \sigma\sigma^* ds\right) = 2E\left(|\int_0^t b(s,x)ds|^2\right).$$

198 or

$$\Delta_0(t) \leq 2E\left(\int_0^t \mathrm{tr}\,\sigma\sigma^* ds\right) + 2E\left(t\int_0^t |b(s,x)|^2 ds\right)$$

(Cauchy-Schwarz inequality)

$$\leq 2nd\ C^2(1+T)t.$$

$$\Delta_n(t) = E(|\xi_{n+1}(t) - \xi_n(t)|^2)$$

$$= E\left(|\int_0^t \langle\sigma(s,\xi_n(s,w)) - \sigma(s,\xi_{n-1}(s,w))d\beta\rangle + \right.$$

$$\left. + \int_0^t b(s,\xi_n(s,w)) - b(s,\xi_{n-1}(s,w))ds|^2\right)$$

$$\leq 2E\left(|\int_0^t \langle\sigma(s,\xi_n(s,w)) - (s,\xi_{n-1}(s,w)), d\beta(s,w)\rangle|^2\right) +$$

$$+ 2E(|\int_0^t (b(s,\xi_n(s,w)) - b(s,\xi_{n-1}(s,w))ds|^2)$$

$$\leq 2E(\int_0^t \mathrm{tr}[(\sigma(s,\xi_n(s,w)) - \sigma(s,\xi_{n-1}(s,w))]\times$$

$$\times [\sigma^*(s,\xi_n(s,w)) - \sigma^*(s,\xi_{n-1}(s,w))]ds] +$$

$$+ 2E\left(t\int_0^t |b(s,\xi_n(s,w)) - b(s,\xi_{n-1}(s,w))|^2 ds\right)$$

$$\leq 2dn\ A^2 \int_0^t \Delta_{n-1}(s)ds + 2tA^2 n \int_0^t \Delta_{n-1}(s)ds$$

$$\leq 2dn\, A^2(1+T) \int_0^T \Delta_{n-1}(s)ds.$$

This proves the result. □ 199

Step 3. $\Delta_n(t) \leq \dfrac{(C^* t)^{n+1}}{(n+1)!}$ $\forall n$ in $0 \leq t \leq T$, where

$$C^* = \max\{2nd\, C^2(1+T), \quad \text{and} \quad A^2(1+T)\}.$$

Proof follows by induction on n.

Step 4. $\xi_n|_{[0,T]\times\Omega}$ is Cauchy in $L^2([0,T]\times\Omega, B([0,T]\times\Omega), \mu\times P)$, where μ is the Lebesgue measure on $[0,T]$.

Proof. $\Delta_n(t) \leq \dfrac{(C^* t)^{n+1}}{(n+1)!}$ implies that

$$\|\xi_{n+1} - \xi_n\|_2^2 \leq \frac{(C^* T)^{n+2}}{(n+2)!}.$$

Here $\|\cdot\|_2$ is the norm in $L^2([0,T]\times\Omega)$. Thus

$$\sum_{n=1}^{\infty} \|\xi_{n+1} - \xi_n\|_2 < \infty, \quad \text{proving Step (4).}$$

□

Step 5. (4) implies that $\xi_n|_{[0,T]\times\Omega}$ is Cauchy in $L^2([0,T]\times\Omega, \mathscr{A}_0, \mu\times P)$ where $\mathscr{A}_0$ is as in Lemma 1. Thus $\xi_n|_{[0,T]\times\Omega}$ converges to $\overline{\xi}_T$ in $L^2([0,T]\times\Omega)$ where $\overline{\xi}_T$ is progressively measurable.

Step 6. If $\xi_n|_{[0,T_2]\times\Omega}\overline{\xi}_{T_2}$ in $L^2([0,T_2]\times\Omega)$ and 200

$$\xi_n|_{[0,T_1]\times\Omega}\overline{\xi}_{T_1} \quad \text{in} \quad L^2([0,T_1]\times\Omega),$$

then $\overline{\xi}_{T_2|_{[0,T_1]\times\Omega}} = \overline{\xi}_{T_1}$ a.e. on $[0,T_1]\times\Omega$, $T_1 < T_2$.

This follows from the fact that if $\xi_n \to \xi$ in L^2, a subsequence of (ξ_n) converges pointwise a.e. to ξ.

Step 7. Let $\overline{\xi}$ be defined on $[0,\infty)\times\Omega$ by $\overline{\xi}|_{[0,T]\times\Omega}=\overline{\xi}_T$. We now show that

$$\overline{\xi}(t,w)=x+\int_0^t\langle\sigma(s,\overline{\xi}(s,\cdot)),d\beta(s,\cdot)\rangle+\int_0^t b(s,\overline{\xi}(s,\cdot))ds.$$

Proof. Let $0\leq t\leq T$. By definition,

$$\xi_n(t,w)=x+\int_0^t\langle\sigma(s,\xi_{n-1}(s,\cdot)),d\beta(s,\cdot)\rangle+\int_0^t b(s,\xi_{n-1}(s,\cdot))ds.$$

$$E\left[\left(\int_0^t\langle(\sigma(s,\xi_n(s,\cdot))-\sigma(s,\overline{\xi}(s,w))),d\beta(s,w)\rangle\right)^2\right]$$

$$=E(\int_0^T\operatorname{tr}[(\sigma(s,\xi_n(s,w))-\sigma(s,\overline{\xi}(s,w)))(\sigma(s,\xi_n(s,w))-\sigma(s,\overline{\xi}(s,w)))^*ds$$

$$\leq dn\,A\int_0^T\int_\Omega|\xi_n(s,w)-\overline{\xi}(s,w)|^2ds\to 0\quad\text{as}\quad n\to\infty$$

(by Lipschitz condition on σ).

Therefore

$$\int_0^t\langle\sigma(s,\xi_{n-1}(s,w),d\beta(s,w)\rangle\to\int_0^t\langle\sigma(s,\overline{\xi}(s,w)),d\beta(s,w)\rangle$$

201 in $L^2(\Omega,P)$. Similarly,

$$\int_0^t b(s,\xi_n(s,w))ds\to\int_0^t b(s,\overline{\xi}(s,w))ds,\quad\text{in}\quad L^2.$$

Thus we get

$$(*)\qquad\overline{\xi}(t,w)=x+\int_0^t\langle\sigma(s,\overline{\xi}(s,w)),d\beta(s,w)\rangle+$$

$$+\int_0^t b(s,\bar{\xi}(s,w))ds \quad \text{a.e. in} \quad t,w.$$

$\square$

Step 8. Let $\xi(t,w) \equiv$ the right hand side of $(*)$ above. Then $\xi(t,w)$ is almost surely continuous because the stochastic integral of a bounded progressively measurable function is almost surely continuous. The result follows by noting that $[0,\infty) = \bigcup_{n=1}^{\infty}[0,n]$ and a function on $[0,\infty)$ is continuous iff it is continuous on $[0,n]$, $\forall n$.

Step 9. Replacing $\bar{\xi}$ by ξ in the right side of $(*)$ we get a solution

$$\xi(t,w) = x + \int_0^t \langle\sigma(s,\xi), d\beta\rangle + \int_0^t b(s,\xi(s,w))ds$$

that is a.s. continuous $\forall t$ and a.e.

Uniqueness. Let ξ and η be two progressively measurable a.s. continuous functions satisfying (1). As in Step 3,

$$\begin{aligned}
E(|\xi(t,w)-x|^2) &\leq 2(E(\int_0^t \operatorname{tr}\sigma\sigma^* ds) + 2E(t\int_0^t b|^2 ds) \\
&\leq 2E(\int_0^T \operatorname{tr}\sigma\sigma^* ds + 2E(T\int_0^T |b|^2 ds), \quad \text{if} \quad 0 \leq t \leq T \\
&< \infty.
\end{aligned}$$

202

Thus $E(|\xi(t,w)|^2)$ is bounded in $0 \leq t \leq T$. Therefore

$$\begin{aligned}
\phi(t) &= E(|\xi(t,w)-\eta(t,w)|^2) \\
&\leq 2E(|\xi(t,w)|^2) + 2E(|\eta(t,w)|^2)
\end{aligned}$$

and so $\phi(t)$ is bounded in $0 \leq t \leq T$. But

$$\phi(t) \leq 2dn\, A^2(1+T) \int\limits_0^t \phi(s)ds$$

as in Step 2; using boundedness of $\phi(t)$ in $0 \leq t \leq T$ we can find a constant C such that

$$\phi(t) \leq Ct \quad \text{and} \quad \phi(t) \leq C \int\limits_0^t \phi(s)ds, \quad 0 \leq t \leq T.$$

By iteration $\phi(t) \leq \dfrac{(Ct)^n}{n!} \leq \dfrac{(CT)^n}{n!}$. Therefore

$$\phi = 0 \quad \text{on} \quad [0, T],$$

i.e. $\xi(t, w) = \eta(t, w)$ a.e. in $[0, T]$. But rationals being dense in $\mathbb{R}$ we have

$$\xi = \eta \quad \text{a.e. and} \quad \forall t.$$

It is now clear that $\xi \in I[b, a]$.

Remark. The above theorem is valid for the equation

$$\xi(t, w) = x_0 + \int\limits_{t_0}^t \langle \sigma(s,), d\beta \rangle + \int\limits_{t_0}^t b(s, \xi)ds, \quad \forall t \geq t_0.$$

This solution will be denoted by ξ_{t_0, x_0}.

203 **Proposition .** *Let $\phi : C[(0, \infty); \mathbb{R}^n) \to C([t_0, \infty); \mathbb{R}^d)$ be the map sending w to ξ_{t_0,x_0}, P the Brownian measure on $C([0, \infty); \mathbb{R}^n)$. Let $P_{t_0,x_0} = P\phi^{-1}$ be the measure induced on $C([t_0, \infty); \mathbb{R}^d)$. Define $X : [t_0, \infty) \times C)[t_0, \infty); \mathbb{R}^d)$ by $X(t, w) = w(t)$. Then X is an Itö process relative to $(C([t_0, \infty); \mathbb{R}^d), {}_{t_0}, P_{t_0,x_0})$ with parameters*

$$[b(t, X_t), a(t, X_t)].$$

The proof of the proposition follows from

Exercise. Let $(\overline{\Omega}, \overline{\mathscr{F}}_t, \overline{P})$, $(\Omega, \mathscr{F}_t, P)$ be any two measure spaces with X, Y progressively measurable on Ω, $\overline{\Omega}$ respectively. Suppose $\lambda : \overline{\Omega} \to \Omega$ is such that λ is $(\overline{\mathscr{F}}_t, \mathscr{F}_t)$-measurable for all t, and $\overline{P}\lambda^{-1} = P$. Let $X(t, \overline{w}) = Y(t, \lambda w)$, $\forall \overline{w} \in \overline{\Omega}$. Show that

(a) If X is a martingale, so is Y.

(b) If $X \in I[b(t, X_t), a(t, X_t)]$ then
$Y \in I[b(t, Y_t), a(t, Y_t)]$.

Lemma . *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be (Ω, P)-measurable, $\sum$ a sub - σ - algebra of $\mathscr{F}$. Let $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ be such that X is $\sum$-measurable and Y is $\sum$-independent. If $g(w) = f(X(w), Y(w))$ with $E(g(w)) < \infty$, then*

$$E(g|\sum)(w) = E(f(x, Y)|_{x=X(w)},$$

i.e.

$$E(f(X, Y)|_{\sum})(w) = \int\limits_{\Omega} f(X(w), Y(w'))dP(w').$$

Proof. Let A and B be measurable subsets in $\mathbb{R}$. The result is trivially **204**
verified if $f = X_{A\times B}$. The set

$$\mathscr{A} = \{F \in \mathbb{R} \ : \ \text{the result is true for } X_F\}$$

is a monotone class containing all measurable rectangles. Thus the Lemma is true for all characteristic functions. The general result follows by limiting procedures. ◻

26. Uniqueness of Diffusion Process

IN THE LAST section we proved that 205

$$\xi(t,w) = x_0 + \int_{t_0}^{t} \langle \sigma(s, \xi(s,w)), d\beta(s,w) + \int_{t_0}^{t} b(s,(s,w)ds$$

has a solution under certain conditions on b and σ where $\sigma\sigma^* = a$. The measure $P_{t_0,x_0} = P\xi^{-1}_{t_0,x_0}$ was constructed on $(C([t_0,\infty);\mathbb{R}^d), \mathscr{F}_{t_0})$ so that the map $X(t,w) = w(t)$ is an Itô process with parameters b and a. We now settle the uniqueness question, about the diffusion process.

Theorem . *Let*

(i) $a : [0,\infty) \times \mathbb{R}^d \to S^+_d$ *and* $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ *be bounded measurable functions;*

(ii) $\Omega = C([0,);\mathbb{R}^d)$*;*

(iii) $X : [0,\infty) \times \Omega \to \mathbb{R}^d$ *be defined by* $X(t,w) = w(t)$*;*

(iv) $X_t = \sigma\{X(s) : 0 \leq s \leq t\}$*;*

(v) *P be any probability measure on*

$$= \sigma\left(\bigcup_{t\geq 0} X_t\right)$$

such that $P\{X(0) = x_0\} = 1$ and X is an Itô process relative to (Ω, X_t, P) with parameters $b(t, X_t)$ and $a(t, X_t)$;

(vi) *$\sigma : [0, \infty) \times \mathbb{R}^d \to M_{d\times n}$ be a bounded measurable map into the set of all real $d \times n$ matrices such that $\sigma\sigma^* = a$ on $[0, \infty) \times \mathbb{R}^d$.*

206 *Then there exists a generalised n-dimensional Brownian motion β on $(\overline{\Omega}, \sum_t, Q)$ and a progressively measurable a.s. continuous map $\xi : [0, \infty) \times \overline{\Omega} \to \mathbb{R}^d$ satisfying the equation*

$$\xi(t, \overline{w}) = x_0 + \int\limits_0^t \langle \sigma(s, \int(s, \overline{w})), d\beta(s, \overline{w})\rangle + \int\limits_0^t b(s, \xi(s, \overline{w}))ds \tag{1}$$

with $Q\xi^{-1} = P$, where $\xi : \overline{\Omega} \to \Omega$ is given by $(\xi(\overline{w}))(t) = \xi(t, \overline{w})$.

Roughly speaking, any Itô process can be realised by means of a diffusion process governed by equation (1) with $\sigma\sigma^* = a$.

Proof. **Case (i).** Assume that there exist constants m, $M > 0$ such that $mI \leq a(t, x) \leq MI$ and σ is a $d \times d$ matrix satisfying $\sigma\sigma^* = a$. In this case we can identify $(\overline{\Omega}, \sum_t, Q)$ with $(\Omega, \mathscr{F}_t, P)$. Since $D(t, \cdot)$ is an Itô process,

$$\exp\langle\theta, X(t)\rangle - \int\limits_0^t \langle\theta, b(s, X(s, \cdot))\rangle ds - \frac{1}{2}\int\limits_0^t \langle\theta, a(s, X(s, \cdot))\theta\rangle ds$$

is a $(\Omega, \mathscr{F}_t, P)$-martingale. Put

$$Y(t, w) = X(t, w) - \int\limits_0^t b(s, X(s, w))ds - x_0.$$

Clearly $Y(t, w)$ is an Itô process corresponding to the parameters

$$[0, a(s, X_s)],$$

so that

$$\exp\langle\theta, Y(t,w)\rangle - \frac{1}{2}\int_0^t \langle\theta, a(s, X(s,\cdot))\theta\rangle ds$$

is a $(\Omega, \mathscr{F}_t, P)$-martingale. The conditions $m \leq a \leq M$ imply that σ^{-1} exists and is bounded. Let

$$\eta(t) = \int_0^t \sigma^{-1} dY = \int_0^t \sigma^{-1}(s, X(s,\cdot)) dY(s,\cdot),$$

so that (by definition of a stochastic integral) η is a $(\Omega, \mathscr{F}_t, P)$-Itô process with parameters zero and $\sigma^{-1}a(\sigma^{-1})^* = 1$. Thus η is a Brownian motion relative to $(\Omega, \mathscr{F}_t, P)$. Now by change of variable formula for stochastic integrals, **207**

$$\begin{aligned}\int_0^t \sigma d\eta &= \int_0^t \sigma\sigma^{-1} dY\\ &= Y(t) - Y(0) = Y((t),\end{aligned}$$

since $Y(0) = 0$. Thus

$$X(t) = x_0 + \int_0^t \sigma(s, X(s,\cdot))d + \int_0^t b(s, X(s,\cdot))ds.$$

Taking $Q = P$ we get the result.

Case (ii). $a = 0, b = 0, x_0 = 0, \sigma = 0$ where $\sigma \in M_{d\times n}$. Let $(\Omega^*, \mathscr{F}_t^*, P^*)$ be an n-dimensional Brownian motion. Define

$$(\overline{\Omega}, \sum_t, Q) = (\Omega \times \Omega^*, \mathscr{F}_t \times \mathscr{F}_t^*, P \times P^*).$$

If β is the n-dimensional Brownian motion on $(\Omega^*, \mathscr{F}_t^*, P^*)$, we define $\overline{\beta}$ on $\overline{\Omega}$ by $\overline{\beta}(t, w, w^*) = \beta(t, w^*)$. It is easy to verify that $\overline{\beta}$ is an n-dimensional Brownian motion on $(\overline{\Omega}, \sum_t, Q)$. Taking $\xi(t, w, w^*) = x_0$ we get the result. □

Before we take up the general case we prove a few Lemmas.

Lemma 1. *Let $\sigma : \mathbb{R}^n \to \mathbb{R}^d$ be linear $\sigma\sigma^* = a : \mathbb{R}^d \to \mathbb{R}^d$; then there exists a linear map which we denote by $\sigma^{-1} : \mathbb{R}^d \to \mathbb{R}^n$ such that $\sigma^{-1} a \sigma^{-1*} = \pi_{N_\sigma^\perp}$, where π denotes the projection and N_σ null space of σ.* **208**

Proof. Let R_σ = range of σ. Clearly $\sigma : N_\sigma^\perp \to R$ is an isomorphism. Let $\tau : R_\sigma \to N_\sigma^\perp$ be the inverse. We put

$$\sigma^{-1} = \tau \oplus 0 : R_\sigma \oplus R_\sigma^\perp \to N_\sigma^\perp \oplus N_\sigma.$$

□

Lemma 2. *Let X, Y be martingales relative to $(\Omega, \mathscr{F}_t, P)$ and $(\overline{\Omega}, \overline{\mathscr{F}}_t, \overline{P})$ respectively. Then Z given by*

$$Z(t, w, \overline{w}) = X(t, w)Y(t, \overline{w})$$

is a martingale relative to

$$(\Omega \times \overline{\Omega}, \mathscr{F}_t \times \overline{\mathscr{F}}_t, P \times \overline{P}).$$

Proof. From the definition it is clear that for every $t > s$

$$\int\limits_{A\times\overline{A}} Z(t, w, \overline{w}) d(P \times \overline{P})|_{\mathscr{F}_s \times \overline{\mathscr{F}_s}} = \int\limits_{A\times\overline{A}} Z(s, w, \overline{w}) d(P \times \overline{P})$$

if $A \in \mathscr{F}_s$ and $\overline{A} \in \overline{\mathscr{F}}_s$. The general case follows easily. □

As a corrollary to Lemma 2, we have

Lemma 3. *Let X be a d-dimensional Itô process with parameters b and a relative to $(\Omega, \mathscr{F}_t, P)$ and let Y be a $\overline{d}$-dimensional Itô process relative to $(\overline{\Omega}, \overline{\mathscr{F}}_t, \overline{P})$ relative to $\overline{b}$ and $\overline{a}$. Then $Z(t, w, \overline{w}) = (X(t, w), Y(t, \overline{w}))$ is a $(d + d)$-dimensional Itô process with parameters $B = (b, \overline{b})$, $A = \begin{bmatrix} a & 0 \\ 0 & \overline{a} \end{bmatrix}$ relative to $(\Omega \times \overline{\Omega}, \mathscr{F}_t \times \overline{\mathscr{F}}_t, P \times \overline{P})$.*

209 **Lemma 4.** *Let X be an Itô process relative to $(\Omega, \mathscr{F}_t, P)$ with parameters* 0 *and a. If θ is progressively measurable such that $E(\int_0^t |\theta|^2, ds) < \infty$, $\forall t$ and $\theta a\theta^*$ is bounded, then $\int_0^t \langle\theta, dX\rangle \in I[0, \theta a\theta^*]$.*

Proof. Let θ_n be defined by

$$\theta_n^i = \begin{cases} \theta^i, & \text{if } |\theta| \leq n, \\ 0, & \text{otherwise;} \end{cases}$$

Then $\int_0^t \langle\theta_n, dX\rangle \in I[0, \theta_n a\theta_n^*]$. Therefore

$$X_n(t) = \exp(\lambda \int_0^t \langle\theta_n, dX\rangle - \frac{\lambda^2}{2} \int_0^t \langle\theta_n, a\theta_n\rangle ds$$

is a martingale converging pointwise to

$$X(t) = \exp\left(\lambda \int_0^t \langle\theta, dX\rangle - \frac{\lambda^2}{2} \int_0^t \langle\theta, a\theta\rangle ds\right).$$

To prove that $\int_0^t \langle\theta, dX\rangle$ is an Itô process we have only to show that $X_n(t)$ is uniformly integrable. Without loss of generality we may assume that $\lambda = 1$. Let $[0, T]$ be given

$$\begin{aligned} E(X_n^2(t, w)) &= E\left(\exp\left[2\int_0^t \langle\theta_n, dX\rangle - \int_0^t \langle\theta_n, a\theta_n\rangle ds\right]\right) \\ &= E\left(\exp\left[2\int_0^t \langle\theta_n, dX\rangle - 2\int_0^t \langle\theta_n, a\theta_n\rangle ds\right.\right. \\ &\qquad \left.\left. + \int_0^t \langle\theta_n, a\theta_n\rangle ds\right]\right). \end{aligned}$$

$$\leq e^T \sup_{0\leq t\leq T} \langle \theta_n, a\theta_n\rangle.$$

210 But $\langle \theta, a\theta^*\rangle$ is bounded and therefore $\langle \theta_n, a\theta_n\rangle$ is uniformly bounded in n. Therefore (X_n) are uniformly integrable. Thus $X(t,\cdot)$ is a martingale.

Case (iii). Take $d = 1$, and assume that

$$\int_0^t a^{-1}\chi_{(a>0)} ds < \infty, \forall t$$

with $a > 0$; let σ = +ve squareroot of a. Define $1/\sigma = 1/\sigma$ if $\sigma > 0$, and $1/\sigma = 0$ if $\sigma = 0$. Let

$$Y(t) = X(t) - x_0 - \int_0^t b(s, X(s))ds.$$

Denote by Z the one-dimensional Brownian motion on $(\Omega^*, \mathscr{F}_t^*, P^*)$ where $\Omega^* = C([0,\infty), R)$. Now

$$Y \in I[0, a(s, X(s,\cdot))],\ Z \in I[0,1].$$

By Lemma 3,

$$(Y, Z) \in I\left((0,0); \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}\right).$$

If

$$\eta(t, w, w^*) = \int_0^t \langle\left(\frac{1}{\sigma(s, X(s,\cdot))}\chi_{(\sigma>0)}, \chi_{\sigma=0}, d(Y,Z)\rangle\right)$$

then Lemma 4 shows that

$$\eta \in I[0,1].$$

Therefore η is a one-dimensional Brownian motion on $\overline{\Omega} = (\Omega \times \Omega^*, \mathscr{F}_t \times \mathscr{F}_t^*, P \times P^*)$. Put

$$\overline{Y}(t, w, w^*) = Y(t, w) \quad \text{and} \quad \overline{X}(t, w, w^*) = X(t, w);$$

211 then

$$\int_0^t \sigma d\eta = \int_0^t \sigma \frac{1}{\sigma} \chi_{(\sigma>0)} dY + \int_0^t \sigma \chi_{(\sigma=0)} dZ$$
$$= \int_0^t \chi_{(\sigma>0)} dY.$$

Since

$$E\left(\left(\int_0^t \chi_{(\sigma=0)} dY\right)^2\right) = E\left(\int_0^t \sigma^2 \chi_{(\sigma=0)} ds\right) = 0,$$

it follows that

$$\int_0^t \sigma d\eta = \int_0^t dY = Y(t) = \overline{Y}(t, w, w^*).$$

Thus,

$$\overline{X}(t, w, w^*) = x_0 + \int_0^t \sigma(s, \overline{X}(s, w, w^*) d\eta +$$
$$+ \int_0^t b(s, \overline{X}(s, w, w^*) ds$$

with $\overline{X}(t, w, w^*) = X(t, w)$. Now

$$(P \times P^*)\overline{X}^{-1}(A) = (P \times P^*)(A \times \Omega^*) = P(A).$$

Therefore

$$(P \times P^*)\overline{X}^{-1} = P \quad \text{or} \quad Q\overline{X}^{-1} = P.$$

Case (iv). (General Case). Define

$$Y(t,\cdot) = X(t,\cdot) - x_0 - \int_0^t b(s, X(s,\cdot))ds.$$

Therefore $Y \in I[0, a(s, X(s,\cdot))]$ relative to $(\Omega, \mathscr{F}_t, P)$. Let Z be the **212**
n-dimensional Brownian motion on $(\Omega^*, \mathscr{F}_t^*, P^*)$ where

$$\Omega^* = C([0,\infty); \mathbb{R}_n).$$

$$(Y, Z) \quad I\left[(0,0); \begin{bmatrix} a(s, X_s), & 0 \\ 0 & I \end{bmatrix}\right]$$

Let σ be a $d \times n$ matrix such that $\sigma\sigma^* = a$ on $[0,\infty) \times \mathbb{R}^d$. Let

$$\begin{aligned}
\eta(t, w, w^*) &= \int_0^t \sigma^{-1}(s, X(s,w))dY(s,w) + \int_0^t r_{N_\sigma}(s, Z(s,w^*))dZ(s,w^*) \\
&= \int_0^t \langle (\sigma^{-1}(s, X(s,w)), \pi_{N_\sigma}(s, Z(s,w^*))), d(Y,Z) \rangle.
\end{aligned}$$

Therefore η is an Ito process with parameters zero and

$$\begin{aligned}
A &= (\sigma^{-1}, \pi_N) \begin{pmatrix} a & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \sigma^{-1^*} \\ \pi_{N_\sigma^*} \end{pmatrix} \\
&= \sigma^{-1} a (\sigma^{-1})^* + \pi_{N_\sigma} \pi_{N_\sigma^*}. \\
&= \pi_{N_\sigma} + \pi_{N_\sigma} \quad \text{(for any projection } PP^* = PP = P) \\
&= I_{\mathbb{R}^n}.
\end{aligned}$$

Therefore η is n-dimensional Brownian motion on

$$(\overline{\Omega}, \overline{\mathscr{F}_t}, \overline{P}) = (\Omega \times \Omega^*, \mathscr{F}_t \times \mathscr{F}_t^*, P \times P^*).$$

$$\int_0^t \sigma(s, X(s,w))d\eta(s, w, w^*)$$

$$= \int_0^t \sigma(s, \overline{X}(s, w, w^*))d\eta(s, w, w^*), \text{ where } \overline{X}(s, w, w^*) = X(s, w),$$

$$= \int_0^t \sigma\sigma^{-1}dY + \int_0^t \sigma\pi_{N_\sigma}dZ.$$

$$= \int_0^t \pi_{R_\sigma}dY, \text{ since } \sigma\sigma^{-1} = \pi_{R_\sigma} \text{ and } \sigma\pi_{N_\sigma} = 0,$$

$$= \int_0^t (1 - \pi_{R_\sigma})dY.$$

□ **213**

Claim. $\int_0^t \pi_{R_\sigma}dY = 0.$

For

$$E\left[\left(\int_0^t \pi_{R_\sigma}dY\right)^2\right] = \int_0^t a\pi_{R_\sigma}ds = \int_0^t \sigma\sigma^*\pi_{R_\sigma}ds$$

$$= \int_0^t \sigma(0)ds = 0.$$

Therefore we obtain

$$\int_0^t \sigma(s, X(s, w))d\eta(s, w, w^*) = \int_0^t dY = Y(t) - Y(0) = Y(t)$$

putting $\overline{Y}(t, w, w^*) = Y(t, w)$, one gets

$$\overline{X}(t, w, w^*) = x_0 + \int_0^t \sigma(s, \overline{X}(s, w, w^*))d\eta(s, w, w^*)$$

$$+\int_0^t b(s, \overline{X}(s, w, w^*))ds.$$

As in Case (iii) one shows easily that

$$(P \times P^*)\overline{X}^{-1} = P.$$

This completes the proof of the theorem.

Corollary . *Let $a : [0, \infty) \times \mathbb{R}^d \to S_d^+$, and $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be bounded, progressively measurable functions. If for some choice of a Lipschitz function $\sigma : [0, \infty) \times \mathbb{R}^d \to M_{d\times n}$, $\sigma\sigma^* = a$ then the Itô process corresponding to $[b, a)$ is unique.*

214 To state the result precisely, let P_1 and P_2 be two probability measures on $C([0, \infty); \mathbb{R}^d)$ such that $X(t, w) = w(t)$ is an Itô process with parameters b and a. Then $P_1 = P_2$.

Proof. By the theorem, there exists a generalised n-dimensional Brownian motion β_i on $(\Omega_i, \sum_t^i, Q_i)$ and a map $\xi_i : \Omega_i \to \Omega$ satisfying (for $i = 1, 2$)

$$i^{(t,w)} = x_0 + \int_0^t \sigma(s, \xi_i(s, w))d\beta_i(s, w) + \int_0^t b(s, \xi_i(s, w))ds.$$

and $P_i = Q_i\xi_i^{-1}$.

Now σ is Lipschitz so that ξ_i is unique but we know that the iterations converge to a solution. As the solution is unique the iterations converge to ξ_i. Each iteration is progressively measurable with respect to

$_t^i = \sigma\{\beta_i(s); 0 \leq s \leq t\}$ so that ξ_i is also progressively

measurable with respect to $\mathscr{F}_t^i$. Thus we can restate the result as follows: There exists $(\Omega_i, \mathscr{F}_t^i, Q_i)$ and a map $\xi_i : \Omega_i \to \Omega$ satisfying

$$\xi_i(t, w) = x_0 + \int_0^t \sigma(s, \xi_i(s, w))d\beta_i(s, w)$$

$$+ \int_0^t b(s, \xi_i(s, w))ds,$$

and $P_i = Q_i \xi_i^{-1}$.

$(\Omega_i, \mathscr{F}_t^i, Q_i, \beta_i)$ can be identified with the standard Brownian motion $(\Omega^*, \mathscr{F}_t^*, Q, \beta)$. Thus $P_1 = Q\xi^{-1} = P_2$, completing the proof. $\square$

27. On Lipschitz Square Roots

Lemma . *Let* $f : \mathbb{R} \to \mathbb{R}$ *be such that* $f(x) \geq 0$, $f(x) \in C^2$ *and* $|f''(x)| \leq$ **215**
A *on* $(-\infty, \infty)$*; then*

$$|f'(x)| \leq \sqrt{f(x)}\sqrt{2A}.$$

Proof.

$$\begin{aligned} 0 \leq f(y) = f(x) + (y-x)f'(x) + \frac{(y-x)^2}{2} f''(\xi) \\ \leq f(x) + Zf'(x) + \frac{Z^2}{2} f''(\xi) \end{aligned}$$

where $Z = y - x$, or $f(y) \leq f(x) + Zf'(x) + \dfrac{AZ^2}{2}$. Therefore

$$\frac{AZ^2}{2} + Zf'(x) + f(x) \geq 0, \quad \forall Z \in \mathbb{R}$$

$$|f'(x)|^2 \leq 2A\ f(x).$$

So

$$|f'(x)| \leq \sqrt{2A} f(x).$$

□

Note. If we take $f(x) = x^2$, we note that the constant is the best possible.

Corollary . *If* $f \geq 0$, $|f''| \leq A$*, then*

$$|\sqrt{(f(x_1))} - \sqrt{(f(x_2))} \leq \sqrt{(A/2)}|x_1 - x_2|.$$

Proof. Let $\epsilon > 0$, then $\sqrt{(f(x)+\epsilon)}$ is a smooth function.

$$(\sqrt{(f(x)+\epsilon)})' = \frac{f'(x)}{2\sqrt{(f(x)+\epsilon)}} = \frac{(f(x)+\epsilon)'}{2\sqrt{(f(x)+\epsilon)}}.$$

Therefore

$$|(\sqrt{(f(x)+\epsilon)})'| \leq \sqrt{(2A/2)} \leq \sqrt{(A/2)},$$

or

$$|\sqrt{(f(x_1)+\epsilon)} - \sqrt{(f(x_2)+\epsilon)}| \leq \sqrt{(A/2)}|x_1 - x_2|.$$

216

Let $\epsilon \to 0$ to get the result.

We now consider the general case and give conditions on the matrix a so that σ defined by $\sigma\sigma^* = a$ is Lipschitz. □

Theorem . *Let $a : \mathbb{R}^n \to S_d^+$ be continuous and bounded C^2-function such that the second derivative is uniformly bounded, i.e. $\|D_s D_r a_{ij}\| \leq M$, where M is independent of i, j, r, s; ($D_r \equiv \dfrac{d}{dx_r}$). If $\sigma : \mathbb{R}^n \to S_d^+$ is the unique positive square root of a, then*

$$\|\sigma(x_1) - \sigma(x_2)\| \leq A|x_1 - x_2|, \ \forall x_1, x_2, A = A(M,d).$$

Proof.

Step 1. Let $A \in S_d^+$ be strictly positive such that $\|I - A\| < 1$. Then

$$\begin{aligned}\sqrt{A} &= \sqrt{(I-(I-A))}\\ &= \sum_{r=0}^{\infty} \frac{C_r}{r!}(I-A)^r,\end{aligned}$$

so that on the set $\{A : \|I - A\| < 1\}$ the map $A \to \sqrt{A}$ is C^∞ (in fact analytic).

Now assume that A is any positive definite matrix. Let $\lambda_1, \ldots, \lambda_n$ be the eigen values so that $\lambda_j > 0$, $j = 1, 2 \ldots n$. Therefore $I - \epsilon A$ is aymmetric with eigen values $1 - \epsilon\lambda_j$. By choosing ϵ sufficiently small we can make

$$||I - \epsilon A|| = \max\{1 - \epsilon\lambda_1, \ldots, 1 - \epsilon\lambda_n\} < 1.$$

Fixing such an ϵ we observe that

$$\sqrt{A} = \frac{1}{\sqrt{\epsilon}}\sqrt{(\epsilon A)} = \frac{1}{\sqrt{\epsilon}}\sqrt{(I - (I - \epsilon A))}.$$

217

So the map $A \to \sqrt{A}$ is smooth on the set of symmetric positive definite matrices.

Step 2. Let $n = 1$, $\sigma(t_0) = \sqrt{(a(t_0)}$ where $a(t_0)$ is positive definite. Assume $a(t_0)$ to be diagonal so that $\sigma(t_0)$ is also diagonal.

$$\sum_j \sigma_{ij}(t)\sigma_{jk}(t) = a_{ik}(t).$$

Differentiating with respect to t at $t = t_0$ we get

$$\sum_j \sigma_{ij}(t_0)\sigma'_{jk}(t_0) + \sum_j \sigma'_{ij}(t_0)\sigma_{jk}(t_0) = a'_{ik}(t_0)$$

or

$$\sqrt{a_{ii}}(t_0)\sigma'_{ik}(t_0) + \sqrt{a_{kk}}(t_0)\sigma'_{ik}(t_0) = a'_{ik}(t_0)$$

or

$$\sigma'_{ik}(t_0) = \frac{a'_{ik}(t_0)}{\sqrt{(a_{ii}(t_0))} + \sqrt{(a_{kk}(t_0))}}.$$

Since the second derivatives are bounded by $4M$ and $a_{ii} - 2a_{ij} + a_{jj} \geq 0$, we get

$$\begin{aligned}|a'_{ii}(t) + 2a'_{ij}(t) + a'_{jj}(t)| &\leq \sqrt{(8M)}\sqrt{(a_{ii}(t) + 2a_{ij}(t) + a_{jj}(t))}\\ &\leq \sqrt{(8M)}\sqrt{2}\sqrt{(a_{ii} + a_{jj})}(t)\end{aligned}$$

or

(1) $$|a'_{ii}(t) + 2a'_{ij}(t) + a'_{jj}(t)| \leq 4\sqrt{M}(\sqrt{a_{ii}} + \sqrt{a_{jj}}).$$

Since a is non-negative definite,

$$|a'_{ii}(t)| \leq \sqrt{(2M)}\sqrt{(a_{ii}(t))},\ \forall i.$$

substituting this in (1), we get **218**

$$|a'_{ij}(t)| \leq 4\sqrt{M}(\sqrt{a_{ii}} + \sqrt{a_{jj}}),$$

and hence

$$|\sigma'_{ij}(t_0)| \leq 4\sqrt{M}.$$

Step 3. Let $a(t_0)$ be positive definite and σ its positive definite square root. There exists a constant unitary matrix α such that $\alpha a(t_0)\alpha^{-1} = b(t_0)$ is a diagonal positive definite matrix. Let $\lambda(t_0)$ be the positive square root of $b(t_0)$ so that

$$\lambda(t_0) = \alpha\sigma(t_0)\alpha^{-1}.$$

Therefore $\sigma'(t_0) = (\alpha^{-1}\lambda'\alpha)(t_0)$ where $(\sigma'(t_0))_{ij} = \sigma'_{ij}(t_0)$ and

$$a''(t_0) = (\alpha^{-1}b''\alpha)(t_0).$$

Since α is unitary.

$$\|\lambda\| = \|\sigma\|, \|a''\| = \|b''\|, \|\lambda'\| = \|\sigma'\|.$$

By hypothesis, $\|b''\| = \|a''\| \leq C(d) \cdot M$. Therefore

$$\|\lambda'\| \leq 4\sqrt{(MC(d))},$$

i.e.

$$\|\sigma'\| \leq 4\sqrt{(MC(d))}.$$

Thus $\|\sigma(t_1) - \sigma(t_2)\| \leq |t_1 - t_2|C(M, d)$.

Step 4. Let $a : \mathbb{R} \to S^+_d$ and σ be the unique non-negative definite square root of a. For each $\epsilon > 0$ let $a_\epsilon = a + \epsilon I$, σ_ϵ = unique positive square root of a_ϵ. Then by step 3,

$$\|\sigma_\epsilon(t_1) - \sigma_\epsilon(t_2)\| \leq C(M, d)|t_1 - t_2|.$$

If a is diagonal then it is obvious that $\sigma_\epsilon \to \sigma$ as $\epsilon \to 0$. In the general case reduce a to the diagonal form and conclude that $\sigma_\epsilon \to \sigma$.

219 Thus

$$\|\sigma(t_1) - \sigma(t_2)\| \leq C(M, d)|t_1 - t_2|.$$

Step 5. Let $a : \mathbb{R}^n \to S_d^+$ and $\sigma^2 = a$, with $\|D_r D_s a_{ij}\| \leq M$, $\forall x$, i, j; r, $s \times \epsilon \mathbb{R}^n$. Choose x_1, $x_2 \in \mathbb{R}^n$. Let $x_1 = y_1, y_2, \ldots, y_{n+1} = x_2$ be $(n + 1)$ points such that y_i and y_{i+1} differ almost in one coordinate. By Step 4, we have

$$(*) \qquad \|\sigma(y_i) - \sigma(y_{i+1})\| \leq C|y_i - y_{i+1}|.$$

The result follows easily from the fact that

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{and} \quad \|x\|_2 = (x_1 + \cdots + x_n)^{1/2}$$

are equivalent norms.

This completes the proof of the theorem.

$\square$

28. Random Time Changes

LET 220

$$L = \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j \frac{\partial}{\partial x_j}$$

with $a : \mathbb{R}^b \to S_d^+$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ bounded measurable funcitons. Let $X(t,\cdot)$, given by $X(t,w) = w(t)$ for (t,w) in $[0,\infty) \times C([0,\infty) : \mathbb{R}^d)$ be an Itô process corresponding to $(\Omega, \mathscr{F}_t, Q)$ with parameters b and a where $\Omega = C([0,\infty); \mathbb{R}^d)$. For every constant $c > 0$ define

$$L_c \equiv c\left[\frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j \frac{\partial}{\partial x_j}\right]$$

Define Q_c by $Q_c = PT_c^{-1}$ where $(T_c w)(t) = w(ct)$. Then one can show that X is an Itô process corresponding to $(\Omega, \mathscr{F}_t, Q_c)$ with parameters cb and ca [Note: We have done this in the case where $a_{ij} = \delta_{ij}$].

Consider the equation

$$\frac{\partial u}{\partial t} = L_c u \quad \text{with} \quad u(0,x) = f(x).$$

This can be written as $\dfrac{\partial u}{\partial \tau} = Lu$ with $u(0,x) = f(x)$ when $\tau = ct$. Thus changing time in the differential equation is equivalent to stretching time in probablistic language.

So far we have assumed that c is a constant. Now we shall allow c to be a function of x.

Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be any bounded measurable function such that

$$0 < C_1 \leq \phi(x) < C_2 < \infty, \quad \forall x \in \mathbb{R}^d$$

and suitable constants C_1 and C_2. If

$$L \equiv \left[\frac{1}{2}\sum a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum b_j \frac{\partial}{\partial x_j}\right]$$

221 we define

$$L_\phi = \phi L \equiv \phi\left[\frac{1}{2}\sum a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum b_j \frac{\partial}{\partial x_j}\right].$$

In this case we can say that the manner which time changes depends on the position of the particle.

Define $T_\phi : \Omega \to \Omega$ by

$$(T_\phi w)(t) = w(\tau_t(w))$$

where $\tau_t(w)$ is the solution of the equation

$$\int_0^{\tau_t} \frac{ds}{\phi(w(s))} = t.$$

As $C_1 \leq \phi \leq C_2$ it is clear that $\tau_t \frac{1}{C_1} \leq t \leq \tau_t \frac{1}{C_2}$. When $\phi \equiv c$ a constant, then $\tau_t = ct$ and T_ϕ coincides with T_c.

As

$$0 < C_1 \leq \phi \leq C_2 < \infty, \quad \int_0^{\lambda} \frac{1}{\phi(w(s))} ds$$

is continuous and increases strictly from 0 to ∞ as λ increases, so that τ_t exists, is unique, and is a continuous function of t for each fixed w.

Some properties of T_ϕ.

(i) If l is the constant function taking the value 1 then it is clear that T_l = identity.

222 (ii) Let ϕ and ψ be two measurable funcitons such that $0 < a \leq \phi(x)$, $\psi(x) \leq b < \infty$, $\forall x \in \mathbb{R}^d$. Then $T_\phi \circ T = T_{\phi\psi} = T_\psi \circ T_\phi$.

Proof. Fix w. Let τ_t be given by

$$\int\limits_0^{\tau_t} \frac{1}{\phi(w(s))} ds = t.$$

Let $w^*(t) = w(\tau_t)$ and let σ_t be given by

$$\int\limits_0^{\sigma_t} \frac{1}{\phi(w^*(s))} ds = t.$$

Let $w^{**}(t) = w^*(\sigma_t) = w(\tau_{\sigma_t})$. Therefore

$$\begin{aligned} ((T_\psi \circ T_\phi)w)(t) &= (T_\phi w^*)(t) = w^*(\sigma_t) \\ &= w^{**}(t) = w(\tau_{\sigma_t}). \end{aligned}$$

Hence to prove the property (ii) we need only show that

$$\int\limits_0^{\tau_{\sigma_t}} \frac{1}{\phi(w(s))} \frac{1}{\psi(w(s))} ds = t.$$

Since

$$\int\limits_0^{\tau_t} \frac{1}{\phi(w(s))} ds = t, \quad \frac{dt}{d\tau_t} = \frac{1}{\phi(w(\tau_t))}$$

and

$$\frac{dt}{d\sigma_t} = \frac{1}{\psi(w^*(\sigma_t))} = \frac{1}{\psi(w(\tau_{\sigma_t}))}$$

Therefore

$$\begin{aligned} \frac{d\tau_{\sigma_t}}{dt} = \frac{d\tau_{\sigma_t}}{d\sigma_t} - \frac{d_{\sigma_t}}{dt} &= \phi(w(\tau_{\sigma_t}))\phi(w^*(\sigma_t)) \\ &= \phi(w(\tau_{\sigma_t})\psi(w(\tau_{\sigma_t})) \end{aligned}$$

$$= (\phi\psi)(w(\tau_{\sigma_t})).$$

223

Thus

$$\int_0^{\tau_{\sigma_t}} \frac{1}{(\phi\psi)(w(s))} ds = t.$$

This completes the proof. □

(iii) From (i) and (ii) it is clear that $T_\phi^{-1} = T_{\phi-1}$ where $\phi^{-1} = \dfrac{1}{\phi}$.

(iv) (τ_t) is a stopping time relative to τ_t. i.e.

$$\left\{ w : \int_0^{\lambda} \frac{1}{\phi(w(s))} ds \geq r \right\} \in \lambda \text{ for each } \lambda \geq 0.$$

(v) $T_\phi(w)(t) = w(\tau_t w) = X_{\tau_t}(w)$.

Thus T_ϕ is $(\mathscr{F}_t - \mathscr{F}_{\tau_t})$-measurable, i.e. $T_\phi^{-1}(\mathscr{F}_t) \subset \mathscr{F}_{\tau_t}$.

Since $X(t)$ is an Itô process, with parameters b, a, $\forall f \in C_0^\infty(\mathbb{R}^d)$, $f(X(t)) - \int_0^t (Lf)(X(s)) ds$ is a martingale relative to $(\Omega, \mathscr{F}_t, P)$. By the optional sampling theorem

$$f(X_{\tau_t}) - \int_0^{\tau_t} (Lf)(X(s)) ds$$

is a martingale relative to $(\Omega, \mathscr{F}_{\tau_t}, P)$, i.e.

$$f(X_{\tau_t}) - \int_0^{t} (Lf)(X(\tau_s)) d\tau_s$$

is a martingale relative to $(\Omega, \mathscr{F}_{\tau_t}, P)$. But $\dfrac{d\tau_s}{dt} = \phi$. Therefore

$$f(X(\tau_t)) - \int_0^t (Lf)(X_{\tau_s})\phi(X_{\tau_s})ds$$

is a martingale. 224

Put $Y(t) = X_{\tau_t}$ and appeal to the definition of L_ϕ to conclude that

$$f(Y(t)) - \int_0^t (L_\phi f)(Y(s))ds$$

is a martingale. $Y(t, w) = X_{\tau_t}(w) = (T_\phi w)(t)$. Let $\overline{\mathscr{F}}_t = \sigma\{Y(s) : 0 \leq s \leq t\}$. Then $\overline{\mathscr{F}}_t = T_\phi^{-1}(\mathscr{F}_t) \subset \mathscr{F}_{\tau_t}$. Thus

$$f(Y(t)) - \int_0^t (L_\phi f)(Y(s))ds$$

is a martingale relative to $(\Omega, \overline{\mathscr{F}}_t, P)$. Define $Q = PT_\phi^{-1}$ so that

$$f(X(t)) - \int_0^t (L_\phi f)(X(s))ds$$

is an $(\Omega, \mathscr{F}_t, Q)$-martingale, i.e. Q is an Itô process that corresponds to the operator ϕL. Or, PT_ϕ^{-1} is an Itô process that corresponds to the operator ϕL.

We have now proved the following theorem.

Theorem . *Let* $\Omega = C([0, \infty); \mathbb{R}^d); X(t, w) = w(t)$*;*

$$L = \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum b_j \frac{\partial}{\partial x_j}.$$

Suppose that $X(t)$ *is an Itô process relative to* $(\Omega, \mathscr{F}_t, P)$ *that corresponds to the operator* L*. Let* $0 \leq C_1 \leq \phi \leq C_2$ *where* $\phi : \mathbb{R}^d \to \mathbb{R}$ 225
is measurable. If $Q = PT_\phi^{-1}$*, then* $X(t)$ *is an Itô process relative to* $(\Omega, \mathscr{F}_t, Q)$ *that corresponds to the operator* ϕL*.*

As $0 < C_1 \leq \phi \leq C_2$, we get $0 < 1/C_2 \leq 1/\phi < 1/C_1$ with $T_{\phi-1} \circ T_\phi = I$. We have thus an obvious corollary.

Corollary . *There exists a probability measure P on Ω such that X is an Itô process relative to $(\Omega, \mathscr{F}_t, P)$ that corresponds to the operator L if and only if there exists a probability measure Q on Ω such that X is an Ito process relative to $(\Omega, \mathscr{F}_t, Q)$ that corresponds to the operator ϕL.*

Remark. If $C_2 \geq \phi \geq C_1 > 0$ then we have shown that existence and uniqueness of an Itô process for the operator L guarantees existence and uniqueness of the Itô process for the operator ϕL. The solution is no longer unique if we relax the strict positivity on C_1 as is illustrated by the following example.

Let $\phi \equiv a(x) = |x|^\alpha \wedge 1$ where $0 < \alpha < 1$ and let $L = \dfrac{1}{2} a \dfrac{\partial^2}{\partial x}$. Define δ_0 on $\{C([0,\infty); \mathbb{R})\}$ by

$$\delta_0(A) = \begin{cases} 1, & \text{if} \quad \theta \in A, \ \forall A \epsilon, \\ 0, & \text{if} \quad \theta \notin A, \end{cases}$$

where θ is the zero function on $[0,\infty)$.

Claim. δ_0 is an Itô process with parameters 0 and a. For this it is enough to show that, $\forall f \in C_0^\infty(\mathbb{R})$

$$f(X(t)) - \int\limits_0^t (Lf)(X(s))ds$$

226 is a martingale, using $a(0) = 0$, it follows easily that

$$\int\limits_A \int\limits_0^t (Lf)(X(\sigma))d\sigma d\delta_0 = 0$$

$\forall$ Borel set A of $C([0,\infty); \mathbb{R})$ and $\int\limits_A f(X(t))d\delta_0 = 0$ if $\theta \notin A$ and

$$\int\limits_A f(X(t))d\delta_0 = f(0)$$

if $\theta \in A$, and this is true $\forall t$, showing that $X(t,w) = w(t)$ is an Itô process relative to δ_0 corresponding to the operator L.

Next we shall define T_a (as in the theorem); we note that T_a cannot be defined everywhere (for example $T_a(\theta)$ is not defined). However T_a is defined a.e. P where $P = P_0$ is the Brownian motion.

$$E^P\left(\int_0^t \frac{1}{|X(s)|^\alpha} ds\right) = \int_0^t \int_0^\infty \frac{1}{y^\alpha} \frac{1}{\sqrt{(2\pi s)}} e^{\frac{-y}{2s}} dy\, ds < \infty$$

since $0 < \alpha < 1$. Thus by Fubini's theorem,

$$\int_0^t \frac{1}{|w(s)|^\alpha} ds < \infty \quad \text{a.e.}$$

Taking $t = 1, 2, 3\ldots$, there exists a set Ω^* such that $P(\Omega^*) = 1$ and

$$\int_0^t \frac{1}{|w(s)|^\alpha} ds < \infty, \quad \forall t, \quad \forall w \in \Omega^*$$

Observe that

$$\int_0^t \frac{1}{|w(s)|^\alpha} ds < \infty$$

implies that

$$\int_0^t \frac{1}{|w(s)|^\alpha \wedge 1} ds < \infty,$$

for **227**

$$\int_0^t \frac{ds}{|w(s)|^\alpha \wedge 1} =$$

$$= \int_{[0,t]\{|w(s)|^\alpha > 1\}} \frac{ds}{|w(s)|^\alpha \wedge 1} + \int_{\{|w(s)|^\alpha \le 1\}[0,t]} \frac{ds}{|w(s)|^\alpha}, \ldots$$

$$\leq m\{(|w(s)|^\alpha > 1)[0,t]\} + \int\limits_0^t \frac{1}{|w(s)|^\alpha} ds < \infty$$

$$(m = \text{Lebesgue measure})$$

Thus T_a is defined on the whole of Ω^*. Using the same argument as in the theorem, it can now be proved that X is an Itô process relative to Q corresponding to the operator L. Finally, we show that $Q\{\theta\} = 0$. $Q\{\theta\} = PT_a^{-1}\{\theta\}$. Now: $T_a^{-1}\{\theta\} = \text{empty}$. For, let $w \in T_a^{-1}\{\theta\}$. Then $w(\tau_t) = 0$, $\forall t$, $w \in \Omega^*$. Since $|\tau_t - \tau_s| \leq |t - s|$, one finds that τ_t is a continuous function of t. Further $\tau_1 > 0$, and $w = 0$ on $[0, \tau_1]$ gives

$$\int\limits_0^{\tau_1} \frac{1}{|w(s)|^\alpha \wedge 1} ds = \infty.$$

This is false unless $T_a^{-1}\{\theta\} = \text{empty}$. Thus $Q\{\theta\} = 0$ and Q is different from δ_0.

29. Cameron - Martin - Girsanov Formula

LET US REVIEW what we did in Brownian motion with drift. **228**

Let $(\Omega, \mathscr{F}_t, P)$ be a d-dimensional Brownian motion with

$$P\{w : w(0) = x\} = 1.$$

Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded measurable function and define

$$Z(t) = \exp\left(\int_0^t \langle b, dx\rangle - \frac{1}{2}\int_0^t |b|^2 ds\right).$$

Then we see that $Z(t, \cdot)$ is an $(\Omega, \mathscr{F}_t, P)$-martingale. We then had a probability measure Q given by the formula

$$\frac{dQ}{dP}\Big|_{\mathscr{F}_t} = Z(t, \cdot).$$

We leave it as an exercise to check that in effect X is an Itô process relative to Q with parameters b and I. In other words we had made a transition from the operator $\Delta/2$ to $\Delta/2 + b \cdot \nabla$. We now see whether such a relation also exists for the more general operator L.

Theorem . *Let $a : \mathbb{R}^d \to S_d^+$ be bounded and measurable such that $a \geq CI$ for some $C > 0$. Let $b : \mathbb{R}^d \to \mathbb{R}^d$ be bounded, $\Omega = ([0, \infty); \mathbb{R}^d)$, $X(t, w) = w(t)$, P any probability measure on Ω such that X is an Itô* *process relative to* $(\Omega, \mathscr{F}_t, P)$ *with parameters* $[0, a]$. *Define* Q^t *on* $\mathscr{F}_t$ *by the rule*

$$\frac{dQ^t}{dP}\bigg|_{\mathscr{F}_t} = Z(t, \cdot) = \exp\left[\int_0^t \langle a^{-1}b, dX\rangle - \frac{1}{2}\int_0^t \langle b, a^{-1}b\rangle ds\right].$$

229 *Then*

(i) $\{0^t\}t \geq 0$ *is a consistent family.*

(ii) *there exists a measure* Q *on* $\sigma(\|\mathscr{F}_t)$*:*

$$Q\Big|_{\mathscr{F}_t} = Q^t.$$

(iii) $X(t)$ *is an Itô process relative to* $(\Omega, \mathscr{F}_t, Q)$ *with parameters* $[b, a]$, *i.e. it corresponds to the operator*

$$\frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j \frac{\partial}{\partial x_j}.$$

Proof. (i) Let $A(t) = \int_0^t \langle a^{-1}b, dX\rangle$. Then $A \in I[0, \langle b, a^{-1}b\rangle]$.

Therefore $Z(t)$ is a martingale relative to $(\Omega, \mathscr{F}_t, P)$ hence $\{Q^t\}_{t\geq 0}$ is a consistent family.

(ii) Proof as in the case of Brownian motion.

(iii) We have to show that

$$\exp[\langle\theta, X(t, \cdot)\rangle - \langle\theta, \int_0^t b ds\rangle - \frac{1}{2}\int_0^t \langle\theta, a\theta\rangle ds]$$

is a martingale relative to $(\Omega, \mathscr{F}_t, Q)$.

Now for any function θ which is progressively measurable and bounded

$$\exp[\int_0^t \langle\theta, dX\rangle - \frac{1}{2}\int_0^t \langle\theta, a\theta\rangle ds]$$

230 is an $(\Omega, \mathscr{F}_t, P)$-martingale. Replace θ by $\theta(w) = \theta + (a^{-1}b)(\chi(s, w))$, where θ now is a constant vector. Then

$$\exp[\int_0^t \langle\theta + a^{-1}b, dX\rangle - \frac{1}{2}\int_0^t \langle\theta + a^{-1}b, a\theta\rangle ds$$

is an $(\Omega, \mathscr{F}_t, P)$-martingale, i.e.

$$\exp[\langle\theta, X(t)\rangle - \frac{1}{2}\int_0^t \langle\theta, a\theta\rangle ds - \frac{1}{2}\int_0^t \langle a^{-1}b, a\theta\rangle - \frac{1}{2}\int_0^t \langle\theta, b\rangle]$$

is an $(\Omega, \mathscr{F}_t, Q)$-martingale, and

$$\begin{aligned}\langle a^{-1}b, a\theta\rangle &= \langle a^* a^{-1}b, \theta\rangle \\ &= \langle aa^{-1}b, \theta\rangle \quad \text{(since } a = a^*\text{)} \\ &= \langle b, \theta\rangle.\end{aligned}$$

Thus

$$\exp[\langle\theta, X(t)\rangle - \frac{1}{2}\int_0^t \langle\theta, a\theta\rangle ds - \int_0^t \langle\theta, b\rangle ds]$$

is an $(\Omega, \mathscr{F}_t, Q)$-martingale, i.e. X is an Itô process relative to $(\Omega, \mathscr{F}_t, Q)$ with parameters $[b, a]$. This proves the theorem. □

We now prove the converse part.

Theorem . *Let*

$$L_1 = \frac{1}{2}\sum i, ja_{ij}\frac{\partial^2}{\partial x_i \partial x_j}$$

and

$$L_2 \equiv \frac{1}{2}\sum_{i,j} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j \frac{\partial}{\partial x_j},$$

where $a : \mathbb{R}^d \to S_d^+$ is bounded measurable such that $a \geq CI$ for some $C > 0$; $b : \mathbb{R}^d \to \mathbb{R}^d$ is bounded and measurable. Let $\Omega = C([0,\infty); \mathbb{R}^d)$
231 *with $\mathscr{F}_t$ as usual. Let θ be a probability measure on $\sigma(\cup \mathscr{F}_t)$ and X a progressively measurable function such that X is an Itô process relative to $(\Omega, \mathscr{F}_t, Q)$ with parameters $[b, a]$ i.e. X corresponds to the operator L_2. Let*

$$Z(t) = \exp[-\int_0^t \langle a^{-1}b, dX\rangle + \frac{1}{2}\int_0^t \langle b, a^{-1}b\rangle ds].$$

Then

(i) *$Z(t)$ is an $(\Omega, \mathscr{F}_t, Q)$-martingale.*

(ii) *If P^t is defined on $\mathscr{F}_t$ by*

$$\frac{dP^t}{dQ}\Big|_{\mathscr{F}_t} = Z(t),$$

Then there exists a probability measure P on $\sigma(\cup \mathscr{F}_t)$ such that

$$P\Big|_{\mathscr{F}_t} = P^t$$

(iii) *X is an Itô process relative to $(\Omega, \mathscr{F}_t, P)$ corresponding to parameters $[0, a]$, i.e. X corresponds to the operator L_1.*

Proof. (i) Let

$$A(t) = \int_0^t \langle -a^{-1}b, dX\rangle.$$

Then $A(t)$ is an Itô process with parameters $[\langle -a^{-1}b, b\rangle, \langle a^{-1}b, b\rangle]$.

Thus

$$\exp[A(t) - \int_0^t \langle -a^{-1}b, b\rangle ds - \frac{1}{2}\int_0^t \langle a^{-1}b, b\rangle ds]$$

is an $(\Omega, \mathscr{F}_t, Q)$-martingale, i.e. $Z(t)$ is an $(\Omega, \mathscr{F}_t, Q)$ martingale.

(ii) By (i), P^t is a consistent family. The proof that there exists a probability measure P is same as before. **232**

Since X is an Itô process relative to Q with parameters b and a,

$$\exp[\int_0^t \langle \theta, dX\rangle - \int_0^t \langle \theta, b\rangle ds - \frac{1}{2}\int_0^t \langle \theta, a\theta\rangle ds]$$

is a martingale relative to Q for every bounded measurable θ. Replace θ by $\theta(w) = \theta - (a^{-1}b)(X(s, w))$ where θ now is a constant vector to get

$$\exp[\langle \theta, X(t)\rangle - \int_0^t \langle a^{-1}b, dX\rangle - \int_0^t \langle \theta, b\rangle + \int_0^t \langle a^{-1}b, b\rangle ds -$$
$$-\frac{1}{2}\int_0^t \langle \theta - a^{-1}b, a\theta - b\rangle ds]$$

is an $(\Omega, \mathscr{F}_t, Q)$ martingale, i.e.

$$\exp[\langle \theta, X\rangle - \int_0^t \langle a^{-1}b, dX\rangle - \int_0^t \langle \theta, b\rangle ds + \int_0^t \langle a^{-1}b, b\rangle ds -$$
$$-\frac{1}{2}\int_0^t \langle \theta, a\theta\rangle ds - \frac{1}{2}\int_0^t \langle a^{-1}b, b\rangle ds + \frac{1}{2}\int_0^t \langle \theta, b\rangle ds +$$
$$+\int_0^t \langle a^{-1}b, a\theta\rangle ds]$$

is an $(\Omega, \mathscr{F}_t, Q)$ martingale. Let $\theta \in \mathbb{R}^d$, so that

$$\exp[\langle \theta, X\rangle - \frac{1}{2}\int_0^t \langle \theta, b\rangle ds - \frac{1}{2}\int_0^t \langle \theta, a\theta\rangle ds + \frac{1}{2}\int_0^t \langle a^{-1}b, a\theta\rangle ds]Z(t)$$

is an $(\Omega, \mathscr{F}_t, Q)$-matringale and

$$\langle a^{-1}b, a\rangle = \langle b, \theta\rangle \quad (\text{since } a = a^*).$$

233 Therefore

$$\exp[\langle \theta, X\rangle - \frac{1}{2}\int_0^t \langle \theta, a\theta\rangle ds]Z(t)$$

is an $(\Omega, \mathscr{F}_t, Q)$ martingale.

Using the fact that $\left.\dfrac{dP}{dQ}\right|_{\mathscr{F}_t} = Z(t)$, we conclude that

$$\exp[\langle \theta, X\rangle - \frac{1}{2}\int_0^t \langle \theta, a\theta\rangle ds]$$

is a martingale relative to $(\Omega, \mathscr{F}_t, P)$, i.e. $X \in I[0, a]$ relative to

$$(\Omega, \mathscr{F}_t, P).$$

This proves the theorem. ◻

Summary. We have the following situation

$$L_1, \Omega, \mathscr{F}_t,\ \Omega = C([0,\infty); \mathbb{R}^d),\ L_2,\ \Omega,\ \mathscr{F}_t.$$

P a probability measure such that X is an Itô Process relative to P corresponding to the operator L_1.	$\Longrightarrow$	X is an Itô process relative to a probability measure Q corresponding to L_2. Q is given by $\left.\dfrac{dQ}{dP}\right\|_{\mathscr{F}_t} = Z(t, \cdot)$
X is an Itô process relative to P corresponding to L_1 where $\left.\dfrac{dP}{dQ}\right\|_{\mathscr{F}_t} = \dfrac{1}{Z(t,\cdot)}$	$\Longleftarrow$	X is an Itô process relative to Q corresponding to L_2.

Thus existence and uniqueness for any system guarantees the existence and uniqueness for the other system.

Application. (Exercise). 234

Take $d = 1$, $a : \mathbb{R} \to \mathbb{R}$ bounded and measurable with $0 < C_1 \leq a < C_2 < \infty$. Let $L = \dfrac{a}{2}\dfrac{\partial^2}{\partial x^2} + b\dfrac{\partial}{\partial x}$. Show that there exists a unique probability masure P on $\Omega = C([0, \infty); \mathbb{R})$ such that $X(t)$ is Itô relative to P corresponding to L. $(X(t, w) \equiv w(t))$ for any given starting point.

30. Behaviour of Diffusions for Large Times

LET $L_2 = \Delta/2 + b \cdot \nabla$ WITH $b : \mathbb{R}^d \to \mathbb{R}^d$ measurable and bounded 235
on each compact set. We assume that there is no explosion. If P_x is the d-dimensional Brownian measure on $\Omega = C([0, \infty); \mathbb{R}^d)$ we know that there exists a probability measure Q_x on Ω such that

$$\left.\frac{dQ_x}{dP_x}\right|_t = \exp\left[\int_0^t \langle b, dX\rangle - \frac{1}{2}\int_0^t |b|^2 ds\right]$$

Let K be any compact set in $\mathbb{R}^d$ with non-empty interior. We are interested in finding out how often the trajectories visit K and whether this 'frequency' depends on the starting point of the trajectory and the compact set K.

Theorem . *Let K be any compact set in $\mathbb{R}^d$ having a non-empty interior. Let*

$$\begin{aligned} E_\infty^K &= \{w : w \textit{ revisits } K \textit{ for arbitrarily large times}\} \\ &= \{w : \textit{ there exists a sequence } t_1 < t_2 < \cdot < \infty \\ &\qquad \textit{with } t_n \to \infty \textit{ such that } w(t_n) \in K\} \end{aligned}$$

Then,

$$\begin{aligned} &\textit{either } Q_x(E_\infty^K) = 0, \ \forall x, \ \textit{and } \forall K, \\ &\textit{or} \quad Q_x(E_\infty^K) = 1, \ \forall x, \ \textit{and } \forall K. \end{aligned}$$

Remark. 1. In the first case $\lim\limits_{t\to+\infty} |X(t)| = +\infty$ a.e. Q_x, $\forall x$, i.e. almost all trajectories stay within K only for a short period.

236 These trajectories are called *transient*. In the second case almost all trajectories visit K for arbitrary large times. Such trajectories are called *recurrent*.

2. If $b = 0$ then $Q_x = P_x$. For the case $d = 1$ or $d = 2$ we know that the trajectories are recurrent. If $d \geq 3$ the trajectories are transient.

Proof.

Step 1. We introduce the following sets.

$$\begin{aligned}
E_0^K &= \{w : X(t, w) \in K \text{ for some } t \geq 0\},\\
E_{t_0}^K &= \{w : X(t, w) \in K \text{ for some } t \geq t_0\},\ 0 \leq t_0 < \infty.
\end{aligned}$$

Then clearly

$$E_\infty^K = \bigcap_{n=1}^{\infty} E_n^K = \bigcap_{t_0 \geq 0} E_{t_0}^K.$$

Let

$$\begin{aligned}
\psi(x) &= Q_x(E_\infty^K),\ F = \chi_{E_\infty^K}.\\
E^{Q_x}(F|\mathscr{F}_t) &= E^{Q_x}(\chi_{E_\infty^K}|\mathscr{F}_t) = Q_{X(t)}(E_\infty^K)\\
&\qquad \text{by the Markov property,}\\
&= \psi(X(t)) \text{ a.e. } Q_x.
\end{aligned}$$

Next we show that $\psi(X(t))$ is a martingale relative to Q_x. For, if $s < t$,

$$\begin{aligned}
&E^{Q_x}(\psi(X(t))|\mathscr{F}_s)\\
&= E^{Q_x}(E^{Q_x}(F|\mathscr{F}_t)|\mathscr{F}_s)\\
&= E^{Q_x}(F|\mathscr{F}_s)\\
&= \psi(X(s)).
\end{aligned}$$

237 Equating the expectations at time $t = 0$ and time t one gets

$$\begin{aligned}\psi(x) &= \int\limits_{\Omega} \psi(X(t))dQ_x\\ &= \int \psi(y)q(t,x,y)dy,\end{aligned}$$

where $q(t,x,A) = Q_x(X_t \in A)$, $\forall A$ Borel in $\mathbb{R}^d$.

We assume for the present that $\psi(x)$ is continuous (This will be shown in Lemma 4 in the next section). By definition $0 \le \psi \le 1$.

Step 2. $\psi(x) = 1$, $\forall x$ or $\psi(x) = 0$, $\forall x$.

Suppose that $\psi(x_0) = 0$ for some x_0. Then

$$0 = \psi(x_0) = \int \psi(y)q(t,x_0,y)dy.$$

As $q > 0$ a.e. and $\psi \ge 0$ we conclude that $\psi(y) = 0$ a.e. (with respect to Lebesgue measure). Since ψ is continuous ψ must vanish identically.

If $\psi(x_0) = 1$ for some x_0, we apply the above argument to $1 - \psi$ to conclude that $\psi = 1$, $\forall x$. We now show that the third possibility $0 < \psi(x) < 1$ can never occur.

Since K is compact and ψ is continuous,

$$0 < a = \inf_{y\in K} \psi(y) \le \sup_{y\in K} \psi(y) = b < 1.$$

From an Exercise in the section on martingales it follows that

$$\psi(X(t)) \to \chi_{E^K_\infty} \quad \text{a.e.} \quad Q_x \quad \text{as} \quad t \to +\infty.$$

Therefore $\lim\limits_{t\to\infty} \psi(X(t))(1-\psi(X(t))) = 0$ a.e. Q_x. Now

$$\begin{aligned}\psi(x_0) = Q_{x_0}(E^K_\infty) &= Q_{x_0}\{w : w(t) \in K \text{ for arbitrary large time}\}\\ &\le Q_{x_0}\{w : a \le \psi(X(t,w)) \le b \text{ for arbitrarily large times}\}\\ &\le Q_{x_0}\{w : (1-b)a \le \psi(X(t))[1-\psi(X(t)] \le b(1-a)\\ &\qquad \text{for arbitrarily large times}\}\end{aligned}$$

$$= 0.$$

238

Thus $\psi(x) = 0$ identically, which is a contradiction. Thus for the given compact set K,

$$\begin{aligned} &\text{either} \quad Q_x(E_\infty^K) = 0, \ \forall x, \\ &\text{or} \qquad Q_x(E_\infty^K) = 1, \ \forall x. \end{aligned}$$

Step 3. If $Q_x(E_\infty^{K_0}) = 1$ for some compact set $K_0(\overset{\circ}{K}_0 \neq \emptyset)$ and $\forall x$, then $Q_x(E_\infty^K) = 1$, $\forall$ compact set K with non-empty interior.

We first given an intuitive argument. Suppose $Q_x(E_\infty^{K_0}) = 1$, i.e. almost all trajectories visit K_0 for arbitrarily large times. Each time a trajectory hits K_0, it has some chance of hitting K. Since the trajectory visits K_0 for arbitrarily large times it will visit K for arbitrarily large times. We now give a precise arguent. Let

$$\begin{aligned} \tau_0 &= \inf\{t : X(t) \in K_0\} \\ \tau_1 &= \inf\{t \geq t_0 + 1 \quad X(t) \in K_0\} \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \\ \tau_n &= \inf\{t \geq t_{n-1} + 1 : X(t) \in K_0\} \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \end{aligned}$$

Clearly $\tau_0 < \tau_1 < \ldots <$ and $\tau_n \geq n$.

$$\begin{aligned} Q_x(E_n^K) &\geq Q_x\{X(t) \in K \text{ for } t \geq \tau_n\} \\ &\geq Q_x\{X(t) \in K \text{ for } t \in \bigcup_{j=n}^{\infty} [\tau_j, \tau_j + 1]\} \\ &= 1 - Q_x \left\{ \bigcap_{j \geq n} X(t) \notin K \text{ for } t \in [\tau_j, \tau_j + 1] \right\} \\ &\geq 1 - Q_x \left\{ \bigcap_{j \geq n} X(\tau_j + 1) \in K \right\} \end{aligned}$$

239 We claim that

$$Q_x(\bigcap_{j\geq n} X(\tau_j+1)\notin K)=0,$$

so that $Q_x(E_n^K)=1$ for every n and hence $Q_x(E_\infty^K)=1$, completing the proof of the theorem.

Now

$$q(1,x,K)\geq q(1,x,\overset{\circ}{K})>0,\quad \forall x,\ \overset{\circ}{K}\quad \text{interior of}\quad K.$$

It is clear that if $x_n\to x$, then

$$\underline{\lim}\, q(1,x_n,\overset{\circ}{K})\geq q(1,x,\overset{\circ}{K}).$$

Let $d=\inf\limits_{x\in K_0} q(1,x,\overset{\circ}{K})$. Then there exists a sequence x_n in K_0 such that $d=\mathop{\mathrm{Lt}}\limits_{n\to\infty} q(1,x_n,\overset{\circ}{K})$. K_0 being compact, there exists a subsequence y_n of x_n with $y_n\to x$ in K_0, so that

$$d=\lim_{n\to\infty} q(1,x,\overset{\circ}{K})=\underline{\lim}\, q(1,y_n,\overset{\circ}{K})\geq q(1,x,\overset{\circ}{K})>0.$$

Thus

$$\inf_{x\in K_0} q(1,x,K)\geq d>0.$$

Now

$$\begin{aligned}
&Q_x\left(\prod_{j=n}^{N} X(\tau_j+1)\notin K|\mathscr{F}_{\tau_N}\right)\\
&=\prod_{j=n}^{N-1}\chi(X(\tau_j+1)\notin K)Q_x(X(\tau_N+1)\in K|\mathscr{F}_{\tau_N})\quad\text{because}\\
&\qquad \tau_j+1\leq\tau_N\quad\text{for}\quad j<N,\\
&=\prod_{j=n}^{N-1}(\chi_{X(\tau_j+1)}\notin K)Q_{X(\tau_N)}(X(1)\notin K)\quad\text{by the strong}
\end{aligned}$$

Markov property,

$$= \prod_{j=1}^{N-1} q(1, X(\tau_N), K^c)\chi_{(X(\tau_j+1)\notin K)}.$$

240

Therefore

$$\begin{aligned}
&Q_x\left(\bigcap_{j=n}^{N} X(\tau_j+1)\notin K\right)\\
&= E^{Q_x}(Q_x(\bigcap_{j=n}^{N} X(\tau_j+1)\notin K|_{\tau_N})))\\
&= E^{Q_x}\left(\prod_{j=n}^{N-1}(\chi_{[X(\tau_j+1)\notin K]})q(1, X(\tau_N), K^c)\right)
\end{aligned}$$

Since K_0 is compact and $X(\tau_N)\in K_0$,

$$q(1, X(\tau_N), K^c) = 1 - q(1, X(\tau_N), K) \le 1 - d$$

Hence

$$Q_x\left(\bigcap_{j=n}^{N} X(\tau_j+1)\notin K\right) \le (1-d)Q_x\left(\bigcap_{j=n}^{N-1} X(\tau_j+1)\notin K\right).$$

Iterating, we get

$$Q_x\left(\bigcap_{j=n}^{N} X(\tau_j+1)\notin K\right) \le (1-d)^{N-n+1},\ \forall N.$$

241 Let $N\to\infty$ to get

$$Q_x\left(\bigcap_{j=n} X(\tau_j+1)\notin K\right) = 0,$$

since $0\le 1-d<1$. Thus the claim is proved and so is the theorem. $\square$

Corollary . *Let K be compact, $\overset{\circ}{K} \neq \emptyset$. Then $Q_x(E^K_\infty) = 1$ if and only if $Q_x(E^K_0) = 1$, $\forall x$.*

Proof. Suppose $Q_x(E^K_\infty) = 1$; then $Q_x(E^K_0) = 1$ because $E^K_\infty E^K_0$. Suppose $Q_x(E^K_0) = 1$, then

$$\begin{aligned} Q_x(E^K_n) &= E^{Q_x}(E^{Q_x}(\chi_{E^K_n} | \mathscr{F}_n)) \\ &= E^{Q_x}(Q_{X(n)}(E^K_0)) \\ &= E^{Q_x}(1) \\ &= 1, \ \forall n. \end{aligned}$$

Therefore $Q_x(E^K_\infty) = 1$. ◻

Remark. If $Q_x(E^K_\infty) = 0$ then it need not imply that

$$Q_x(E^K_0) = 0.$$

Example. Take $b = 0$ and $d = 3$. LEt $K = S_1 = \{x \in \mathbb{R}^3$ such that $|x| \leq 1\}$. Define

$$\psi(n) = \begin{cases} 1, & \text{for} \quad |x| \leq 1, \\ \frac{1}{|x|}, & \text{for} \quad |x| \geq 1. \end{cases}$$

$P_x(E^K_0) \neq$ constant but $P_x(E^K_\infty) = 0$. In fact, $P_x(E^K_0) = \psi(x)$ (Refer Dirichlet Problem).

31. Invariant Probability Distributions

Definition. Let $\{P_x\}_{x \in \mathbb{R}^d}$ be a family of Markov process on **242**

$$\Omega = C([0, \infty); \mathbb{R}^d)$$

indexed by the starting points x, with homogeneous transition probability $p(t, x, A) = P_x(X_t \in A)$ for every Borel set A in $\mathbb{R}^d$. A probability measure μ on the Borel field of $\mathbb{R}^d$ is called an *invariant distribution* if, $\forall A$ Borel in $\mathbb{R}^d$.

$$\int\limits_{\mathbb{R}^d} p(t, x, A) d\mu(x) = \mu(A).$$

We shall denote $dp(t, x, y)$ by $p(t, x, dy)$ or $p(t, x, y)dy$ if it has a density.

Proposition . *Let $L_2 = \Delta/2 + b \cdot \nabla$ with no explosion. Let Q_x be the associated measure. If $\{Q_x\}$ has an invariant measure μ then the process is recurrent.*

Proof. It is enough to show that if K is a compact set with non-empty interior then

$$Q_x(E^K_\infty) = 1$$

for some x. Also $Q_x(E^K_t) \geq Q_x(X_t \in K) = q(t, x, K)$. Therefore

$$\mu(K) = \int q(t, x, K) d\mu(x) \leq \int Q_x(E^K_t) d\mu(x).$$

Now, $0 \leq Q_x(E_t^K) \leq 1$ and $Q_x(E_t^K)$ decreases to $Q_x(E_\infty^K)$ as $t \to \infty$. Therefore by the dominated convergence theorem

$$\mu(K) \leq \int Q_x(E_\infty^K) d\mu(x).$$

243

If the process were transient, then $Q_x(E_\infty^{S_n}) = 0$, $\forall n$, where $S_n = \{x \in \mathbb{R}^d : |x| \leq n\}$, i.e. $\mu(S_n) = 0$, $\forall n$. Therefore $\mu(\mathbb{R}^d) = 0$, which is false. Thus the process is recurrent.

The converse of this proposition is *not* true as is seen by the following example.

Let $L = \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}$ so that we are in a one-dimensional situation (Brownian motion). Then

$$p(t, x, K) = \int\limits_K \frac{1}{\sqrt{(2\pi t)}} e^{\frac{-(x-y)^2}{2t}} dy \leq \frac{1}{\sqrt{(2\pi t)}} \lambda(K),$$

where λ denotes the Lebesgue measure on $\mathbb{R}$. If there exists an invariant distribution μ, then

$$\mu(K) = \int p(t, x, K) d\mu(x) \leq \frac{1}{\sqrt{(2\pi t)}} \lambda(K) \int d\mu(x) = \frac{\lambda(K)}{\sqrt{(2\pi t)}}$$

Letting $t \to \infty$, we get $\mu(K) = 0$ $\forall$ compact K, giving $\mu = 0$, which is false. □

Theorem . *Let $L = \Delta/2 + b \cdot \nabla$ with no explosion. Assume b to be C^∞. Define the formal adjoint L^* of L by $L^* = \Delta/2 - \nabla \cdot b$ (i.e. $L^* u = \dfrac{1}{2}\Delta u - \nabla \cdot (bu)$). Suppose there exists a smooth function ϕ(C^2 - would do) such that $L^*\phi = 0$, $\phi \geq 0$, $\inf \phi dx = 1$. If one defines μ by the rule $\mu(A) = \int\limits_A \phi(y) dy$, then μ is an invariant distribution relative to the family $\{Q_x\}$.*

244 *Proof.* We assume the following result from the theory of partial differential equations.

If $f \in C_0^\infty(G)$ where G is a bounded open set with a smooth boundary ∂G and $f \geq 0$, then there exists a smooth function $U_G : [0,\infty)\times\overline{G} \to [0,\infty)$ such that

$$\begin{aligned}
&\frac{\partial U_G}{\partial t} = LU_G \quad \text{on} \quad (0,\infty)\times G,\\
&U_G(0,x) = f(x) \quad \text{on} \quad \{0\}\times\overline{G},\\
&U_G(t,x) = 0, \quad \forall x \in G.
\end{aligned}$$

Let $t > 0$. As U_G, ϕ are smooth and G is bounded, we have

$$\frac{\partial}{\partial t}\int\limits_G U_G(t,x)\phi(x)dx = \int\limits_G \frac{\partial}{\partial t}U_G\phi ds = \int\limits_G \phi LU_G dx$$

Using Green's formula this can be written as

$$\frac{\partial}{\partial t}\int\limits_G U_G(t,x)\phi(x)dx = \int\limits_G U_G L^*\phi - \frac{1}{2}\int\limits_{\partial G}\left[\phi\frac{\partial U_G}{\partial n} - U_G\frac{\partial}{\partial n}\right]dS + \\ + \int\limits_{\partial G}\langle b\cdot n\rangle U_G(t,x)\phi(x)dS$$

Here n is assumed to be the inward normal to ∂G. So,

$$\frac{\partial}{\partial t}\int\limits_G U_G(t,x)\phi(x)dx = -\frac{1}{2}\int\limits_{\partial G}(x)\frac{\partial U_G}{\partial n}(t,x)dS$$

(Use the equation satisfied by ϕ and the conditions on U_G). Now $U_G(t,x) \geq 0$, $\forall x \in G$, $U_G(t,x) = 0$, $\forall x$ in ∂G, so that

$$\frac{\partial U_G}{\partial n}(t,x) \geq 0.$$

This means that **245**

$$\frac{\partial}{\partial t}\int\limits_G U_G(t,x)\phi(x)dx \leq 0,\ \forall t > 0,$$

i.e. $\int\limits_G U_G(t,x)\phi(x)dx$ is a monotonically decresing function of t. Therefore

$$\begin{aligned}\int\limits_G U_G(t,x)\phi(x)dx &\leq \int\limits_G U_G(0,x)\phi(x)dx\\ &= \int\limits_G f(x)\phi(x)dx\\ &= \int\limits_{\mathbb{R}^d} f(x)\phi(x)dx.\end{aligned}$$

Next we prove that if $U : [0,\infty)\times\mathbb{R}^d \to [0,\infty)$ is such that $\frac{\partial U}{\partial t} = LU$, $\forall t > 0$ and $U(0,x) = f(x)$, then

$$\int\limits_{\mathbb{R}^d} U(t,x)\phi(x)dx \leq \int\limits_{\mathbb{R}^d} f(x)\phi(x)dx.$$

The solution $U_G(t,x)$ can be obtained by using Itô calculus and is given by

$$U_G(t,x) = \int f(X(t))\chi_{\{\tau_{G>t}\}}dQ_x.$$

We already know that

$$U(t,x) = \int f(X(t))dQ_x.$$

Therefore

$$\int U(t,x)\phi(x)dx = \iint f(X(t))\phi(x)DQ_x dx.$$

Now

$$\begin{aligned}&\iint f(X(t))\chi_{\{\tau_G>t\}}dQ_x\phi(x)dx\\ &\qquad\int U_G(t,x)\phi(x)dx \leq \int\limits_{\mathbb{R}^d} f(x)\phi(x)dx.\end{aligned}$$

246 Letting G increase to $\mathbb{R}^d$ and using Fatou's lemma, we get

$$\iint f(X(t))\phi(x)dQ_x dx \leq \int f(x)\phi(x)dx$$

$\square$

This proves the assertion made above.

Let

$$\mu(A) = \int_A \phi(X)dx,$$

$$\nu(A) = \int Q_x(X_t \in A)d\mu(x) = \int q(t, x, A)d\mu(x).$$

Let $f \in C_0^\infty(G)$, $f \geq 0$, where G is a bounded open set with smooth boundary. Now

$$\begin{aligned}\int f(y)d\nu(y) &= \iint f(y)q(t, x, y)d\mu(x)dy\\ &= \iint f(X(t))dQ_x d\mu(x)\\ &= \int U(t, x)d\mu(x)\\ &= \int U(t, x)\phi(x)dx\\ &\leq \int f(x)\phi(x)dx = \int f(x)d\mu(x).\end{aligned}$$

Thus, $\forall f \geq 0$ such that $f \in C_0^\infty$,

$$\int f(x)d\nu(x) \leq \int f(x)d\mu(x).$$

This implies that $\nu(A) \leq \mu(A)$ for every Borel set A. (Use mollifier 247
s and the dominated convergence theorem to prove the above inequality for χ_A when A is bounded). Therefore $\nu(A^c) \leq \mu(A^c)$, or $1 - \mu(A) \leq 1 -$

$\mu(A)$, since μ, ν are both probability measures. This gives $\mu(A) = \nu(A)$, i.e.

$$\mu(A) = \int q(t, x, A) d\mu(x), \quad \forall t,$$

i.e. μ is an invariant distribution.

We now see whether the converse result is true or not. Suppose there exists a probability measure μ on $\mathbb{R}^d$ such that

$$\int Q_x(X_t \in A) d\mu(x) = \mu(A), \quad \forall A \text{ Borel in } \mathbb{R}^d \text{ and } \forall t.$$

The question we have in mind is whether $\mu(A) = \int\limits_A \phi dx$ for some smooth ϕ satisfying $L^* \phi = 0$, $\phi \geq 0$, $\int \phi(x) dx = 1$. To answer this we proceed as follows.

By definition $\mu(A) = \int q(t, x, A) d\mu(x)$. Therefore

$$\begin{aligned} &\iint f(X(t)) dQ_x d\mu(x) \\ &= \iint f(y) q(t, x, y) dy \, d\mu(x) \\ &= \int f(y) d\mu(y) \forall f \in C_0^\infty(\mathbb{R}^d) \|f\|_\infty \leq 1. \end{aligned} \tag{1}$$

Since X is an Itô process relative to Q_x with parameters b and I,

$$f(X(t)) - \int_0^t (Lf)(X(s)) ds$$

248 is a martingale. Equating the expectations at time $t = 0$ and time t we obtain

$$E^{Q_x}(f(X(t)) = f(x) + E^{Q_x}\left(\int_0^t (Lf)(X(s)) ds\right)$$

Integrating this expression with respect to μ gives

$$\iint f(X(t)) dQ_x d\mu(x) = \int f(x) d\mu(x) \iint\limits_{\mathbb{R}^d} \int_0^t (Lf)(X(s)) ds \, dQ_x d\mu.$$

Using (1), we get

$$0 = \int_{\mathbb{R}^d} \int_{\Omega} \int_0^t (Lf)(X(s)) dQ_x ds \, d\mu(x)$$

Applying equation (1) to the function Lf we then get

$$\begin{aligned} 0 &= \int_{\mathbb{R}^d} \int_0^t (Lf)(y) d\mu(y) ds \\ &= t \int_{\mathbb{R}^d} (Lf)(y) d\mu(y), \quad \forall t > 0. \end{aligned}$$

Thus

$$0 = \int_{\mathbb{R}^d} (Lf)(y) d\mu(y), \quad \forall f \in C_0^\infty(\mathbb{R}^d).$$

In the language of distributions this just means that $L^* \mu = 0$.

From the theory of partial differential equations it then follows that there exists a smooth function ϕ such that $\forall A$ Borel in $\mathbb{R}^d$,

$$\mu(A) = \int_A \phi(y) dy$$

with $L^* \phi = 0$. As $\mu \geq 0$, $\phi \geq 0$ and since

$$\mu(\mathbb{R}^d) = 1, \quad \int_{\mathbb{R}^d} \phi(x) dx = 1.$$

249

We have thus proved the following (converse of the previous) theorem.

Theorem . *Let μ be an invariant distribution with respect to the family $\{Q_x\}$ with $b : \mathbb{R}^d \to \mathbb{R}^d$ being C^∞. Then there exists a $\phi \in L'(\mathbb{R}^d)$, $\phi \geq 0$, ϕ smooth such that*

$$L^* \phi = 0, \quad \int \phi(y) dy = 1$$

and such that

$$\mu(A) = \int_A \phi(y)dy, \quad \forall A \quad \textit{Borel in} \quad \mathbb{R}^d.$$

Theorem (Uniqueness). *Let ϕ_1, ϕ_2 be smooth on $\mathbb{R}^d$ such that*

$$\phi_1, \phi_2 \geq 0, 1 = \int_{\mathbb{R}^d} \phi_1 dy = \int_{\mathbb{R}^d} \phi_2 dy, L^*\phi_1 = 0 = L^*\phi_2.$$

Then $\phi_1 = \phi_2$.

Proof. Let $f(x) = \phi_1(x) - \phi_2(x)$,

$$\mu_i(A) = \int_A \phi_i(x)dx, \quad i = 1, 2.$$

Then μ_1, and μ_2 are invariant distributions. Therefore

$$\begin{aligned} \int q(t,x,y)\phi_i(x)dx &= \int q(t,x,y)d\mu_i(x) \\ &= \phi_i(y), \quad \text{(a.e.)}, \quad i = 1, 2. \end{aligned}$$

Taking the difference we obtain

$$\int q(t,x,y) f(x)dx = f(y), \quad \text{a.e.}$$

Now

$$\begin{aligned} \int |f(y)\, dy &= \int |\int q(t,x,y) f(x)dx| dy \\ &\leq \iint q(t,x,y)|f(x)|dx\, dy \\ &= \int |f(x)|dx \int q(t,x,y)dy \\ &= \int |f(\underline{x})|dx. \end{aligned}$$

250

Thus

$$(*) \qquad \iint |f(x)|q(t,x,y)dx\,dy = \int |\int q(t,x,y)f(x)dx|dy \ \forall t.$$

We show that f does not change sign, i.e. $f \geq 0$ a.e. or $f \leq 0$ a.e. The result then follows from the fact that $\int f(x)dx = 0$. Now

$$|\int q(1,x,y)f(x)dx| \leq \int q(1,x,y)|f(x)|dx$$

and $(*)$ above gives

$$\int |\int q(1,x,y)f(x)dx|dy = \iint q(1,x,y)|f(x)|dx\,dy.$$

Thus

$$|\int q(1,x,y)f(x)dx| = \int q(1,x,y)|f(x)|dx \text{ a.e. } y,$$

i.e.

$$|\int_{E^-} q(1,x,y)f(x)dx + \int_{E^-} q(1,x,y)f(x)dx|$$
$$= \int_{E^+} q(1,x,y)f(x)dx - \int_{E^-} q(1,x,y)f(x)dx \text{ a.e. } y,$$

where

$$E^+ = \{x : f(x) > 0\}, \quad E^- = \{x : f(x) < 0\}, \quad E^0 = \{x : f(x) = 0\}.$$

Squaring both sides of the above equality, we obtain

$$(**) \qquad \left(\int_{E^+} q(1,x,y)f(x)dx\right)\left(\int_{E^-} q(1,x,y)f(x)dx\right) = 0, \quad \text{a.e.} \quad y.$$

251

Let A be a set of positive Lebesgue measure; then $p(1, x, A) = P_x(X(1) \in A) > 0$. Since Q_x is equivalent to P_x on Ω we have $Q_x(X(1) \in A) = q(1, x, A) > 0$. Therefore $q(1, x, y) > 0$ a.e. y for each x. By Fubini's theorem $q(1, x, y) > 0$ a.e. x, y. Therefore for almost all y, $q(1, x, y) > 0$ for almost all x. Now pick a y such that $(**)$ holds for which $q(1, x, y) > 0$ a.e. x.

We therefore conclude from $(**)$ that either

$$\int\limits_{E^+} q(1, x, y) f(x) dx = 0, \quad \text{in which case} \quad f \leq 0 \quad \text{a.e.,}$$

or

$$\int\limits_{E^-} q(1, x, y) f(x) dx = 0, \quad \text{in which case} \quad f \geq 0 \quad \text{a.e.}$$

Thus f does not change its sign, which completes the proof. $\square$

Remark. The only property of the operator L we used was to conclude $q > 0$. We may therefore expect a similar result for more general operators.

Theorem . *Let $L^*\phi = 0$ where $\phi \geq 0$ is smooth and $\int \phi(x) dx = 1$. Let K be any compact set. Then*

$$\sup_{x \in K} \int |q(t, x, y) - \phi(y)| dy \to 0 \quad as \quad t \to +\infty.$$

Lemma 1. *Let b be bounded and smooth. For every $f : \mathbb{R}^d \to \mathbb{R}^d$ that is bounded and measurable let $u(t, x) = E^{Q_x}(f(X(t))$. Then for every fixed t, $u(t, x)$ is a continuous function of x. Further, for $t \geq \epsilon > 0$,*

$$\begin{aligned} |u(t, x) - \int u(t - \epsilon, y) \frac{1}{\sqrt{(2\pi\epsilon)^d}} \exp - \frac{|x - y|^2}{2\epsilon} dy| \\ \leq ||f||_\infty \sqrt{(e^{ct}(e^{c\epsilon} - 1))}, \end{aligned}$$

252 *where c is a constant depending only on $||b||_\infty$.*

Proof. Let

$$(1)\qquad \begin{aligned}(T_t f)(x) = E^{Q_x}(f(X(t)) = E^{P_x}(f(X(t))Z(\epsilon,t))+\\ + E^{P_x}(f(X(t))(Z(t) - Z(\epsilon,t))),\end{aligned}$$

where

$$Z(t) = \exp\left[\int_0^t \langle b^2, dx\rangle - \frac{1}{2}\int_0^t |b|^2 ds\right],$$

$$Z(\epsilon,t) = \exp\left[\int_\epsilon^t \langle b, dx\rangle - \frac{1}{2}\int_\epsilon^t |b(X(s))|^2 ds\right].$$

$$\begin{aligned}E^{P_x}(f(X(t))Z(\epsilon,t)) &= E^{P_x}(E^{P_x}(f(X(t))Z(\epsilon,t)|_\epsilon))\\ &= E^{P_x}(E^{P_X}\epsilon)(f(X(t-\epsilon))Z(t-\epsilon)))\\ &\qquad \text{(by Markov property),}\\ &= E^{P_x}(u(t-\epsilon, X(\epsilon)).\end{aligned}$$

$$(2)\qquad = \int u(t-\epsilon, y)\frac{1}{(\sqrt{(2\pi\epsilon)})^d}\exp\left[-\frac{|(x-y)|^2}{2\epsilon}\right]dy.$$

Now

$$\begin{aligned}&(E^{P_x}(|Z(t) - Z(\epsilon,t)|)) =\\ &\quad = E^{P_x}(|Z(\epsilon)Z(\epsilon,t) - Z(\epsilon,t)|))^2\\ &\quad = E^{P_x}(Z(\epsilon,t)Z(\epsilon) - 1|))^2\\ &\quad \le (E^{P_x}((Z(\epsilon)-1)^2))(E^{P_x}(Z^2(\epsilon,t)))\\ &\qquad\quad \text{(by Cauchy Schwarz inequality),}\\ &\quad \le E^{P_x}(Z^2(\epsilon) - 2Z(\epsilon) + 1)E^{P_x}(Z^2(\epsilon,t))\\ &\quad \le E^{P_x}(Z^2(\epsilon) - 1)E^{P_x}(Z^2(\epsilon,t)), \text{(since } E^{P_x}(Z(\epsilon)) = 1),\\ &\quad \le E^{P_x}(Z^2(\epsilon) - 1)E^{P_x}(\exp(2\int_\epsilon^t \langle b, dX\rangle - \frac{2}{2}\int_\epsilon^t |b|^2 ds + \int_\epsilon^t |b|^2 ds))\end{aligned}$$

$$\leq E^{P_x}(Z^2(\epsilon) - 1)e^{ct},$$

253 using Cauchy Schwarz inequality and the fact that

$$E^{P_x}(\exp(2 \int_\epsilon^t \langle b, dX \rangle - \frac{2^2}{2} \int_\epsilon^t |b|^2 ds)) = 1.$$

Thus

$$E^{P_x}(|Z(t) - Z(\epsilon, t)||^2 \leq (e^{c\epsilon} - 1)e^{ct}$$

where c depends only on $||b||_\infty$. Hence

$$\begin{aligned} |E^{P_x}(f(X(t))(Z(t) - Z(\epsilon, t))| &\leq \|f\|_\infty E^{P_x}(|Z(t) - Z(\epsilon, t)|) \\ &\leq \|f\|_\infty \sqrt{((e^{c\epsilon} - 1)e^{ct})}. \end{aligned} \tag{3}$$

Substituting (2) and (3) in (1) we get

$$\begin{aligned} |u(t, x) - \int u(t - \epsilon y) \cdot \frac{1}{(\sqrt{(2\pi\epsilon)^d})} \exp\left[\frac{-|x - y|^2}{2\epsilon}\right] dy \\ \leq \|f\|_\infty \sqrt{((e^{c\epsilon} - 1)e^{ct})} \end{aligned}$$

Note that the right hand side is independent of x and as $\epsilon \to 0$ the right hand side converges to 0. Thus to show that $u(t, x)$ is a continuous function of x (t fixed), it is enough to show that

$$\int u(t - \epsilon, y) \frac{1}{(\sqrt{(2\pi\epsilon)^d}} \exp\left[\frac{-|x - y|^2}{2\epsilon}\right] dy$$

254 is a continuous function of x; but this is clear since u is bounded. Thus for any fixed $tu(t, x)$ is continuous. □

Lemma 2. *For any compact set $K \subset R^d$, for r large enough so that*

$$K \subset \{x : |x| < r\}, \quad x \to Q_x(\tau_r \leq t)$$

is continuous on K for each $t \geq 0$, where

$$\tau_r(w) = \inf\{s : |w(s)| \geq r\}.$$

Proof. $Q_x(\tau_r \leq t)$ depends only on the coefficient $b(x)$ on $|x| \leq r$. So modifying, if necessary, outside $|x| \leq r$, we can very well assume that $|b(x)| \leq M$ for all x. Let

$$\tau_r^\epsilon = \inf\{s : s \geq \epsilon, |w(s)| \geq r\}.$$
$$Q_x(\tau_r^\epsilon \leq t) = E^{Q_x}(u(X(\epsilon))),$$

where

$$u(x) = Q_x(\tau_r \leq t - \epsilon)$$

As b and u are bounded, for every fixed $\epsilon > 0$, by Lemma 1, $Q_x(\tau_r \leq t)$ is a continuous function of x. As

$$|Q_x(\tau_r^\epsilon \leq t) - Q_x(\tau_r \leq t)| \leq Q_x(\tau_r \leq \epsilon),$$

to prove the lemma we have only to show that

$$\operatorname*{limit}_{\epsilon \to 0} \sup_{x \in K} Q_x(\tau_r \leq \epsilon) = 0$$

Now

$$Q_x(\tau_r \leq \epsilon) = \int\limits_{\{\tau_r \leq \epsilon\}} Z()dP_x$$
$$\leq (\int (Z(\epsilon))^2 dP_x)^{1/2} \cdot \sqrt{P_x(\tau_r \leq \epsilon)},$$

by Cauchy-Schwarz inequality. The first factor is bounded because b is **255**
bounded. The second factor tends to zero uniformly on K because

$$\sup_{x \in K} P_x(\tau_r \leq \epsilon) \leq P(\sup_{0 \leq s \leq \epsilon} |w(s)| > \delta)$$

where

$$\delta = \inf_{\substack{y \in K \\ |x| = r.}} |(x - y)|.$$

$\square$

Lemma 3. *Let K be compact in $\mathbb{R}^d$. Then for fixed t, $Q_x(\tau_r \leq t)$ monotically decreses to zero as $r \to \infty$ and the convergence is uniform on K.*

Proof. Let $f_r(x) = Q_x(\tau_r \leq t)$. As $\{\tau_r \leq t\}$ decreases to the null set, $f_r(x)$ decreases to zero. As K is compact, there exists an r_0 such that for $r \geq r_0$, $f_r(x)$ is continuous on K, by Lemma 2. Lemma 3 is a consequence of Dini's theorem. ◻

Lemma 4. *Let $b : \mathbb{R}^r \to \mathbb{R}^d$ be smooth (not necessarily bounded). Then $E^{Q_x}(f(X(t)))$ is continuous in x for every fixed t, f being any bounded measurable function.*

Proof. Let b_r be any bounded smooth function on R^d such that $b_r \equiv b$ on $|x| \leq r$ and Q_x^r the measure corresponding to b_r. Then by Lemma 1, $E^{Q_x}(f(X(t)))$ is continuous in x for all r. Further,

$$|E^{Q_x^r}(f(X(t))) - E^{Q_x}(f(X(t)))| \leq 2\|f\|_\infty \cdot Q_x(\tau_r \leq t).$$

256 The result follows by Lemma 3. ◻

Lemma 5. *With the hypothesis as the same as in Lemma 1, (S_1) is an equicontinuous family, where*

$$S_1 = \{f : \mathbb{R}^d \to \mathbb{R}, f \text{ bounded measurable}, \|f\|_\infty \leq 1\}$$

Proof. For any f in S_1, let $U(x) = U(t, x) \equiv E^{Q_x}(f(X(t)))$ and

$$U_\epsilon(x) = U_\epsilon(t, x) = \int U(t - \epsilon, y) \frac{1}{(\sqrt{(2\pi\epsilon)^d})} \exp\left[\frac{-|x - y|^2}{2\epsilon}\right] dy.$$

By Lemma 1,

$$\begin{aligned} |U(x) - U_\epsilon(x)| &\leq (((e^{c\epsilon} - 1)\epsilon^{ct}))^{1/2} \\ |U(x) - U(y)| &\leq |U(x) - U_\epsilon(x)| + |U_\epsilon(y) - U(y)| + |U_\epsilon(x) - U_\epsilon(y)| \\ &\leq 2\sqrt{((e^{c\epsilon} - 1)e^{ct})} + |U_\epsilon(x) - U_\epsilon(y)|. \end{aligned}$$

The family $\{U_\epsilon : f \in S_1\}$ is equicontinuous because every U occuring in the expression for U_ϵ is bounded by 1, and the exponential factor is uniformly continuous. Thus the right hand side is very small if ϵ is small and $|x - y|$ is small. This proves the lemma. ◻

Lemma 6. *Let b be smooth and assume that there is no explosion (b is not necessarily bounded). Then (S_1) is an equi-continuous family* $\forall t > 0$.

Proof. Let $r > 0$ be given. Define $b_r \in C^\infty$ such that $b_r = 0$ on $|x| > r+1$, $b_r = b$ on $|x| \leq r$, $b_r : \mathbb{R}^d \to \mathbb{R}$. By Lemma 2, we have that

$$\{E^{Q^r_x}(f(X(t))) : f \in S_1\}$$

is equicontinuous, where Q^r_x is the probability measure corresponding **257**
to the function b_r.

$$E^{Q_x}(f(X(t))\chi_{\{\tau_r>t\}})E^{Q^r_x}(f(X(t))\chi_{\{\tau_r>t\}}). \tag{1}$$

Therefore

$$\begin{aligned}
&|E^{Q_x}(f(X(t))) - E^{Q^r_x}(f(X(t))|\\
&= |E^{Q_x}(f(X(t))\chi_{\{\tau_r>t\}}) + E^{Q_x}(f(X(t))\chi_{\{\tau_r\leq t\}})\\
&- E^{Q^r_x}(f(X(t))\chi_{\{\tau_r>t\}}) - E^{Q^r_x}(f(X(t)))\chi_{\{\tau_r\leq t\}})\\
&= |E^{Q_x}(f(X(t)\chi_{\{\tau_r\leq t\}}) - E^{Q^r_x}(f(X(t))\chi_{\{\tau_r\leq t\}})|\\
&\leq \|f\|_\infty(E^{Q_x}(\chi_{\{\tau_r\leq t\}}) + E^{Q^r_x}(\chi_{\{\tau_r\leq t\}})\\
&\leq l[E^{Q_x}(\chi_{(\tau_r\leq t)}) + E^{Q_x}(\chi_{(\tau_r\leq t)})](\text{use (1) with } f = 1)\\
&= 2E^{Q_x}(\chi_{(\tau_r\leq t)}).
\end{aligned}$$

Thus

$$\sup_{x\in K}\sup_{\|f\|_\infty\leq 1}|E^{Q_x}(f(X(t)) - E^{Q^r_x}(f(X(t))| \leq 2\sup_{x\in K}(\chi_{\{\tau_r\leq t\}}).$$

By Lemma 3,

$$\sup_{x\in K} E^{0_x}(\tau_r \leq t) \to 0$$

for every compact set K as $n \to \infty$, for every fixed t.

The equicontinuity of the family (S_1) now follows easily. For fixed **258**
x_0, put $u_r(x) = E^{Q^r_x}(f(X(t)))$ and $u(x) = E^{Q_x}(f(X(t)))$ and let $K = s[x_0, 1] = \{x : |x - x_0| \leq 1\}$. Then

$$|u(x) - u(x_0)| \leq |u(x) - u_r(x)| + |u(x_0) - u_r(x_0)| + |u_r(x) - u_r(x_0)|$$

$$\leq 2 \sup_{y \in K} E^{Q_y}(\chi_{(\tau_r \leq |t)}) + |u_r(x) - u_r(x_0)|$$

By the previous lemma $\{u_r\}$ is an equicontinuous family and since $\sup_{y\in K} E^{Q_y}(\chi_{(\tau_r \leq 1)}) \to 0$, $\{u : \|f\|_\infty \leq 1\}$ is equicontinuous at x_0. This proves the Lemma. ◻

Lemma 7. $T_r \circ T_s = T_{t+s}$, $\forall s, t \geq 0$.

Remark. This property is called the *semigroup property*.

Proof.

$$\begin{aligned} &T_r(T_s f)(x) \\ &= \iint f(z) q(s, y, z) q(t, x, y) dy\, dz. \end{aligned}$$

Thus we have only to show that

$$\int q(t, x, y) q(s, y, A) dy = q(t + s, x, A).$$

$$\begin{aligned} q(t + s, x, A) &= E^{Q_x}(X(t + s) \in A) \\ &= E^{Q_x}(X(t + s) \in A|_t)) \\ &= E^{Q_x}(E^{Q_x} X(t)(X(s) \in A))), \\ &\qquad \text{by Markov property} \\ &= E^{Q_x}(q(s, X(t), A)) \\ &= \int q(t, x, y) q(s, y, A) dy, \end{aligned}$$

which proves the result.

259 As a trivial consequence we have the following. ◻

Lemma 8. *Let $\epsilon > 0$ and let S_1 be the unit ball in $B(\mathbb{R}^d)$. Then $\bigcup_{t\geq\epsilon} T_t(S_1)$ is equicontinuous.*

Proof. $\bigcup_{t\geq\epsilon>0} T_t(S_1) = T(\bigcup_{t\geq 0} T_t(S_1))$ (by Lemma 7) $T_\epsilon(S_1)$.

The result follows by Lemma 6. □

Lemma 9. *Let $u(ttx) = E^{Q_x}(f(X(t)))$ with $\|f\|_\infty \leq 1$. Let $\epsilon > 0$ be given and K any compact set. Then there exists a $T_0 = T_0(\epsilon, K)$ such that $\forall T \geq T_0$ and $\forall x_1, x_2 \in K$,*

$$|u(T, x_1) - u(T, x_2)| \leq \epsilon.$$

Proof. Define $q^*(t, x_1, x_2, y_1, y_2) = q(t, x_1, y_1)q(t, x_2, y_2)$ and let $Q_{(x_1,x_2)}$ be the measure corresponding to the operator

$$L = \frac{1}{2}(\Delta_{x_1} + \Delta_{x_2}) + b(x_1)\cdot\nabla_{x_1} + b(x_2)\cdot\nabla_{x_2}$$

i.e., for any $u : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$,

$$Lu = \frac{1}{2}\sum_{i=1}^{2d} \frac{\partial^2 u}{\partial x_i^2} + \sum_{i=1}^{t} b_i(x_1, \ldots, x_d)\frac{\partial u}{\partial x_i^2} + \\ + \sum_{i=1}^{d} b_i(x_{d+1,\ldots,x_{2d}})\frac{\partial u}{\partial x_{i+d}}.$$

Then $Q_{(x_1,x_2)}$ will be a measure on $C([0,\infty);\mathbb{R}^d\times\mathbb{R}^d)$. We claim that $Q_{(x_1,x_2)} = Q_{x_1} \times Q_{x_2}$. Note that

$$C([0,\infty);\mathbb{R}^d\times\mathbb{R}^d) = C([0,\infty);\mathbb{R}^d) \times C[(0,\infty);\mathbb{R}^d)$$

and since $C([0,\infty);\mathbb{R}^d)$ is a second countable metric space, the Borel field of $C([0,\infty)\mathbb{R}^d\times\mathbb{R}^d)$ is the σ-algebra generated by 260

$$\mathscr{B} = (C([0,\infty);\mathbb{R}^d)) \times \mathscr{B}(C[0,\infty);\mathbb{R}^d).$$

By going to the finite-dimensional distributions one can check that $P_{(x_1,x_2)} = P_{x_1} \times P_{x_2}$.

$$\frac{dQ_{(x_1,x_2)}}{dP_{(x_1,x_2)}}\Big|_{\mathscr{F}_t} = \exp\left[\int_0^t \langle b^{(1)}, dX_1\rangle - \frac{1}{2}\int_0^t |b^{(1)}|^2 ds\right]\times$$

$$\times \exp\left[\int\limits_0^t \langle b^{(2)}, dX_2\rangle - \frac{1}{2}\int\limits_0^t |b^{(2)}|^2 ds\right],$$

where

$$b^{(1)}(x_1 \ldots x_d) = b(x_1 \ldots, x_d) \cdot b^{(2)}(x_{d+1} \ldots x_{2d}) = b(x_{d+1}, \ldots, x_{2d}),$$

so that $Q_{(x_1,x_2)} = Q_{x_1} \times Q_{x_2}$.

It is clear that if ϕ defined an invariant measure for the process Q_x, i.e.

$$\int\limits_A \phi(x)dx = \int \phi(y)Qy(X_t \in A)dy,$$

then $\phi(y_1)\phi(y_2)$ defines an invariant measure for the process $Q_{(x_1,x_2)}$. Thus the process $Q_{(x_1,x_2)}$ is recurrent.

Next we show that $u(T-t, X_1(t))$ is a martingale $(0 \leq t \leq T)$ for any fixed T on $C([0,T];\mathbb{R}^d)$.

$$\begin{aligned}
&E^{Q_x}(u(T-t, X(t)|\mathscr{F}_s)) \\
&= [\int u(T-t, y)q(t-s, x, dy)]_{x=X(s)} \\
&= [\iint f(z)q(T-t, y, dz)q(t-s, x, dy)]_{x=X(s)} \\
&= [\int f(z)q(T-s, x, dz)]_{x=X(s)} \\
&= u(T-s, X(s)), \quad s < t.
\end{aligned}$$

261 It now follows that $u(T-t, X_1(t))$ is a martingale on $C([0,\infty);\mathbb{R}^d) \times C([0,\infty);\mathbb{R}^d)$. Hence $u(T-t, X_1(t)) - u(T-t, X_2(t))$ is a martingale relative to $Q_{(X_1,x_2)}$.

Let $V = S(0, \delta/2) \subset \mathbb{R}^d \times \mathbb{R}^d$ with $\delta < 1/4$. If $(x_1, x_2) \in V$, then

$$|x_1 - x_2| \leq |(x_1, 0) - (0,0)| + |(0,0) - (0, x_2)| < \delta.$$

□

Claim 1. $Q_{(x_1,x_2)}(\tau_V \leq T) \to 1$ as $T \to \infty$, where τ_V is the exit time from $R^d - V$.

Proof. If w is any trajectory starting at some point in V, then $\tau_V = 0 \leq T$, $\forall T$. If w starts at some point outside V then, by the recurrence property, w has to visit a ball with centre 0 and radius $\delta/2$; hence it must get into V at some finite time. Thus $\{\tau_V \leq T\} \uparrow$ to the whole space as $T \uparrow \infty$. Next we show that the convergence is uniform on compact sets.

If x_1, $x_2 \in K$, $(x_1, x_2) \in K \times K$ (a compact set). Put $g_T(x_1, x_2) = Q_{(x_1,x_2)}(\tau_V \leq T)$. Then $g_T(x_1, x_2) \geq 0$ and $g_T(x_1, x_2)$ increases to 1 as T tends to ∞.

$$\begin{aligned} g_T(x_1, x_2) &= Q_{(x_1,x_2)}(\tau_V \leq T) \\ &\quad Q_{(x_1,x_2)}(\tau_V^1 \leq T), \end{aligned}$$

where 262

$$\tau_V^1 = \inf\{t \geq 1 : (x_1, x_2) \in V\}.$$

Therefore

$$\begin{aligned} g_T(x_1, x_2) &\geq E^Q(x_1, x_2)(E^Q(x_1, x_2)((\tau_V^1 \leq T)|_1)) \\ &= E^Q(x_1, x_2)(Q_{(X_1(1),X_2(1))}\{\tau_V^1 \leq T\}) \\ &= E^Q(x_1, x_2)(\psi_T(X_1(1), X_2(1))), \end{aligned}$$

where ψ_T is a bounded non-negative function. Thus, if

$$\begin{aligned} h_T(x_1, x_2) &= Q_{(x_1,x_2)}(\tau_V^1 \leq T) = \\ &= E^Q(x_1, x_2)(\psi_T(X_1(1), X_2(1))), \end{aligned}$$

then by Lemma 4, h_T is continuous for each T, $g_T \geq h_T$ and h_T increases to 1 as $T \to \infty$. Therefore, h_T converges uniformly (and so does g_T) on compact sets.

Thus given $\epsilon > 0$ chose $T_0 = T_0(\epsilon, K)$ such that if $T \geq T_0$,

$$\sup_{x_2 \in K} \sup_{x_1 \in K} Q_{(x_1,x_2)}(\tau_V \geq T - 1) \leq \epsilon.$$

By Doob's optional stopping theorem and the fact that

$$u(T-t, X_1(t)) - u(t-t, X_2(t))$$

is a martingale, we get, on equating expectations,

$$\begin{aligned}
&|u(T, x_1) - u(T, x_2)| \\
&\quad = |E^{Q_{(x_1,x_2)}}[u(T-0, X_1(0) - u(T-0, X_2(0)]| \\
&\quad = |E^{Q_{(x_1,x_2)}}[u(T-(\tau_v \wedge (T-1)), X_1(T-(\tau_v \wedge (T-1))- \\
&\qquad - u(T-(\tau_v \wedge T(-1)), X_2(T-(\tau_v \wedge (T-1)]| \\
&\quad | \int\limits_{\{\tau_v \geq T-1\}} [u(1, X_1(1)) - u(1, X_2(1))] dQ_{(x_1,x_2)} + \\
&\quad + \int\limits_{\{\tau_v < (T-1)\}} [u(T-\tau_v, X_1(T-\tau_v)) - u(T-\tau_v), X_2(T-\tau_v)) dQ_{(x_1,x_2)}|.
\end{aligned}$$

263

Therefore

$$\begin{aligned}
&|u(T, x_1) - u(T, x_2)| \\
&\leq \int\limits_{\{\tau_v \geq (T-1)\}} |[u(1, X_1(1)) - u(1, X_2(1))]| dQ_{(x_1,x_2)} + \\
&+ | \int\limits_{\{\tau_v < (T-1)\}} [u(T-\tau_v, X_1(T-\tau_v)) - u(T-\tau_v, X_2(T-\tau_v)) dQ_{(x_1,x_2)}| \\
&\leq 2\epsilon + | \int\limits_{\{\tau_v < (T-1)\}} [u(T-\tau_v, X_1(T-\tau_v)) - u(T-\tau_v, X_2(T-\tau_v))] dQ_{(x_1,x_2)}|,
\end{aligned}$$

since u is bounded by 1.

The second integration is to be carried out on the set $\{T - v \geq 1\}$. Since $\bigcup\limits_{t \geq 1} T_t(S_1)$ is equicontinuous we can choose a $\delta > 0$ such that whenever $x_1, x_2 \in K$ such that $|x_1 - x_2| < \delta$

$$|u(t, x_1) - u(t, x_2)| \leq \epsilon, \quad \forall t \geq 1.$$

264

Thus $|u(T, x_1) - u(T, x_2)| \leq 3\epsilon$ whenever $x_1, x_2 \in K$ and $T \geq T_0$. This proves the Lemma. □

Corollary to Lemma 9. $\sup\limits_{x_1, x_2 \in K} \int |q(t, x_1, y)|dy$ *converges to* 0 *as* $t \to \infty$.

Proof. Since the dual of L^1 is L^∞, we have

$$\int |q(t, x_1, y) - q(t, x_2, y)|dy$$
$$= \sup_{\|f\|_\infty \leq 1} |\int [q(t, x_1, y) - q(t, x_2, y)] f(y)dy|$$

and the right side converges to 0 as $t \to \infty$, by Lemma 9.

We now come to the proof of the main theorem stated before Lemma 1. Now

$$\int |q(t, x, y) - \phi(y)|dy$$
$$= \int |q(t, x, y) - \int \phi(x^1)q(t, x^1, y)dx^1|dy$$
$$\text{(by invariance property)}$$
$$= \int |\int q(t, x, y)\phi(x^1)dx^1 - \int \phi(x^1)q(t, x^1, y)dx^1|dy$$
$$\left(\text{since } \int \phi(x^1)dx^1 = 1\right)$$
$$\leq \iint |q(t, x, y) - q(t, x^1, y)|\phi(x^1)dx^1dy \quad (\text{since } \phi \geq 0)$$
$$= \int \phi(x^1)dx^1 \int |q(t, x, y) - q(t, x^1, y)|dy$$

Since
$$\int \phi(x^1)dx^1 = \operatorname*{Lt}_{n\to\infty} \int\limits_{|x^1| \leq n} \phi(x^1)dx^1,$$

choose a compact set L K such that $\int\limits_{\mathbb{R}^d - L} \phi(x^1)dx^1 < \epsilon$. Then 265

$$\int \phi(x^1)dx^1 \int |q(t, x, y) - q(t, x^1, y)|dy$$

$$= \int_L \phi(x^1)dx^1 \int |q(t,x,y) - q(t,x^1,y)|dy +$$
$$+ \int_{\mathbb{R}^d - L} \phi(x^1)dx^1 \int |q(t,x,y) - q(t,x^1,y)|dy$$
$$\leq \iint_L \phi(x^1)dx^1 \int |q(t,x,y) - q(t,x^1,y)|dv + 2\epsilon.$$

Chose t_0 such that whenever $t \geq t_0$,

$$\int |q(t,x,y) - q(t,x^1,y)|dy \leq \frac{\epsilon}{1 + \int_L \phi(x^1)dx^1}$$

$\forall x$, x_1 in L. (Corollary to Lemma 9). Then

$$\int |q(t,x,y) - \phi(y)|dy \leq 3\epsilon,$$

if $t \geq t_0 \forall x \in K$ completing the proof of the theorem. □

32. Ergodic Theorem

Theorem . *Let* $f : \mathbb{R}^d \to R$ *be bounded and measurable with* $\|f\|_\infty \leq 1$. **266**
If ϕ *is an invariant distribution for the family* $\{Q_x\}$, $x \in \mathbb{R}^d$ *then*

$$\lim_{\substack{t_1 \to \infty \\ 0 \leq t_2 - t_1 \to \infty}} E^{Q_x}(f(X(t_1))f(X(t_2))) = [\int f(y)\phi(y)dy]^2$$

Proof.

$$\begin{aligned} & E^{Q_x}[(f(X(t_1)f(X(t_2))] \\ & = E^{Q_x}(E^{Q_x}[f(X(t_1))f(X(t_2))|\mathscr{F}_{t_1}]) \\ & = E^{Q_x}(f(X(t_1))(E^{Q_x}[f(X(t_2))|\mathscr{F}_{t_1}])) \\ & = E^{Q_x}(f(X(t_1))\int f(y)q(t_2 - t_1, X(t_1), y))dy), t_2 > t_1 \\ & \qquad\qquad \text{(by Markov property)}, \end{aligned}$$

$$= \int f(z)q(t_1, x, z)dz \int f(y)q(t_2 - t_1, z, y)dy \tag{1}$$

does any bounded an measurable f. By theorem of § 31,

$$\sup_{x \in K} |\int f(y)[q(t, x, y) - \phi(y)]dy| \to 0$$

as $t \to +\infty$. We can therefore write (1) in the form

$$\begin{aligned} & E^{Q_x}[(f(X(t_1))f(X(t_2))] = \\ & \quad = (\int f(z)q(t_1, x, x)dz) \int f\phi + \int f(z)q(t_1, x, z)A(t_2 - t_1, z)dz, \end{aligned}$$

where $A(t_2-t_1, z)$ converges to 0 (uniformly on compact sets as) $t_2-t_1 \to +\infty$.

267 To prove the theorem we have therefore only to show that

$$\int f(z)q(t_1, x, z)A(t_2 - t_1, z)dz \to 0$$

as $t_1 \to +\infty$ and $t_2 - t_1 \to \infty$ (because $\int f(z)q(t_1, x, z)dz \to \int f\phi$). Now

$$\begin{aligned} &|\int f(z)q(t_1, x, z)A(t_2 - t_1, z)dz| \\ &\leq ||f||_\infty \int q(t_1, x, z)|A(t_2 - t_1, z)|dz \\ &\leq \int q(t_1, x, z)|A(t_2 - t_1, z)|dz \end{aligned} \tag{2}$$

Let K be any compact set, then

$$\int_K q(t_1, x, z)dz = \int \chi_K q(t_1, x, z)dz \to \int \chi_K \phi(z)dz$$

at $t_1 \to \infty$. Given $\epsilon > 0$, let K be compact so that

$$|\int \chi_{K^c}\phi(z)dz| \leq \epsilon;$$

then $|\int \chi_{K^c} q(t_1, x, z)dz| \leq 2\epsilon$ if $t_1 \gg 0$. Using (2) we therefore get

$$\begin{aligned} &|\int f(z)q(t_1 x, z)A(t_2 - t_1, z)dz| \\ &\leq \int_K q(t_1, x, z)|A(t_2 - t_1, z)|dz + \int_{K^c} q(t_1, x, z)|A(t_2 - t_1, z)|dz \\ &\leq \int_K q(t_1, x, z)|A(t_2 - t_1, z)|dz + 2\int_{K^c} q(t_1, x, z)dz, \\ &\qquad \text{since } |A(t_2 - t_1, z)| \leq 2, \\ &\leq \int_K q(t_1, x, z)|A(t_2 - t_1, z)|dz + 2\epsilon, \ \text{ if } t_1 \gg 0. \end{aligned}$$

268 The theorem now follows from the fact that

$$\lim_{t_2 - t_1 \to \infty} \sup_{z \in K} |A(t_2 - t_1, z) = 0.$$

$\square$

Weak Ergodic Theorem.

$$\lim_{t\to\infty} E^{Q_x}\left[|\frac{1}{t}\int_0^t f(X(s))ds - f(x)\phi(x)dx| > \epsilon\right] = 0.$$

Proof.

$$E^{Q_x}\left[|\frac{1}{t}\int_0^t f(X(s))ds - \int f(x)\phi(x)dx| > \epsilon\right]$$
$$\leq \frac{1}{\epsilon^2}E^{Q_x}\left[|\frac{1}{t}\int_0^t f(X(s))ds - \int f(y)\phi(y)dy|^2\right],$$

by Tchebychev's inequality. We show that the right side $\to 0$ as $t \to \infty$. Now

$$E^{Q_x}\left[|\frac{1}{t}\int_0^t f(X(s))ds - \int f\phi|^2\right]$$
$$= E^{Q_x}[|\frac{1}{t^2}\int_0^t\int_0^t f(X(\sigma_1))f(X(\sigma_2))d\sigma_1\, d\sigma_2 + (\int f\phi dy)$$
$$- 2\frac{1}{t}\int_0^t f(X(\sigma))d\sigma\int f\phi dy]$$

Also

$$\sup_{x\in K} |E^{Q_x}[f(X(t)) - \int f(y)\phi(y)dy]|$$

$$= \sup_{x \in K} | \int q(t,x,y) f(y) dy - \int f(y)\phi(y) dy|$$
$$\leq \|f\|_\infty \sup_{x \in K} \int |q(t,x,y) - \phi(y)| dy;$$

269 the right hand side tends to 0 as t tends to $+\infty$. Consider

$$|E^{Q_x}(\frac{1}{t}\int_0^t f(X(\sigma))d\sigma - \int f(y)\phi(y)dy)|$$
$$= |E^{Q_x}\left(\frac{1}{t}\int_0^T f(X(\sigma))d\sigma - \int f(y)\phi(y)dy + \frac{1}{t}\int_0^t f(X(\sigma))d\sigma\right), 0 \leq T \leq t,$$
$$\leq \frac{1}{t}|\int_0^T E^{Q_x} f(X(\sigma))d\sigma - T\int f(y)\phi(y)dy| +$$
$$+ |E^{Q_x}\left(\frac{1}{t}\int_T^t f(X(\sigma))d\sigma - \left(\frac{t-T}{t}\right)\int f(y)\phi(y)dy\right)|.$$

Given $\epsilon > 0$ choose T large so that

$$|E^{Q_x}(f(X(\sigma)) - \int f(y)\phi(y)dy| \leq \epsilon, \quad (\sigma \geq T).$$

Then

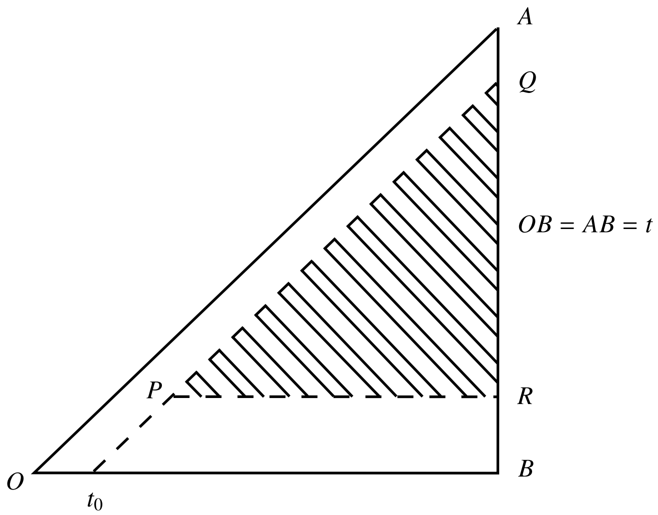
$$|E^{Q_x}\Big(\frac{1}{t}\int_0^t f(X(\sigma))d\sigma - \int f(y)\phi(y)dy| \leq$$
$$\leq |\frac{1}{t}\int_0^T E^{Q_x}[f(X(\sigma))] - \frac{T}{t}\int f(y)\phi(y)dy]| + \frac{t-T}{t}\epsilon$$
$$\leq 2\epsilon$$

provided t is large. Thus

$$\lim_{t \to +\infty} E^{Q_x}[\frac{1}{t}\int_0^t f(X(\sigma))d\sigma] = \int f\phi dy.$$

To prove the result we have therefore only to show that

$$\lim_{t\to+\infty} E^{Q_x}\left(\frac{1}{t^2}\int_0^t\int_0^t f(X(\sigma_1))f(X(\sigma_2))d\sigma_1 d\sigma_2\right) = \left[\int f(y)\phi(y)dy\right]^2$$



270

POR is the region $\sigma_2 \geq t_0$, $\sigma_1 - \sigma_2 \geq t_0$.
Let I

$$= E^{Q_x}\left(\frac{1}{t}\int_0^t\int_0^t f(X(\sigma_1))f(X(\sigma_2))d\sigma_1 d\sigma_2\right) - \left(\int f\phi dy\right)$$

$$= \frac{2}{t^2}\int\left[E^{Q_x}(f(X(\sigma_1))f(X(\sigma_2))) - \left(\int f(y)\phi(y)dy\right)^2\right]d\sigma_1 d\sigma_2$$

$$0 \leq \sigma_2 \leq \sigma_1 \leq t.$$

Then

$$|I| \leq \frac{2}{t^2}\int_{\Delta PQR} |E^{Q_x}(f(X(\sigma_1))f(X(\sigma_2))) - \left(\int f(y)\phi(y)dy\right)^2 |d\sigma_1 d\sigma_2$$

$$+\frac{2}{t^2}\cdot 2\|f\|_\infty^2 \quad [\text{area of } OAB - \text{area of } PQR]$$

By the Ergodic theorem the integrand of the first term on the right can be made less than $\epsilon/2$ provided t_0 is large (see diagram). Therefore

$$\begin{aligned}
|I| &\leq \frac{\epsilon}{2}\cdot\frac{2}{t^2}\text{ area of } PQR + \frac{4}{t^2}\|f\|_\infty^2\left[\frac{t^2}{2} - \left(\frac{(t-2t_0)}{2}\right)^2\right] \\
&\leq \frac{\epsilon}{2} + \frac{2\|f\|_\infty^2}{t^2}[4tt_0 - 4t_0^2]. \\
&< \epsilon
\end{aligned}$$

271 if t is large. This completes the proof of the theorem. □

33. Application of Stochastic Integral

LET b BE A bounded function. For every Brownian measure P_x on 272
$\Omega = C([0,\infty); \mathbb{R}^d)$ we have a probability measure Q_x on $(\Omega, \mathscr{F})$.

Problem. Let $q(t,x,A) = Q_x(X_t \in A) \cdot q(t,x,\cdot)$ is a probability measure on $\mathbb{R}^d$. We would like to know if $q(t,x,\cdot)$ is given by a density function on $\mathbb{R}^d$ and study its properties.

Step (i). $q(t,x,\cdot)$ is absolutely continuous with respect to the Lebesgue measure.

For, $p(t,x,A) = P_x(X(t) \in A)$ is given by a density function. Therefore $p(t,x,\cdot) \gg m_d$ (Lebesgue measure). Since

$$
\begin{aligned}
&Q_x \ll P_x \text{ on } \mathscr{F}_t,\\
&q(t,z,\cdot) \leq M_d \text{ on } \mathscr{F}_t.
\end{aligned}
$$

Step (ii). Let $q(t,x,y) \geq 0$ be the density function of $q(t,x,\cdot)$ and write $p(t,x,y)$ for the density of $p(t,x,\cdot)$. Let $1 < \alpha < \infty$. Put

$$
\begin{aligned}
r(t,x,y) &= \frac{q(t,x,y)}{p(t,x,y)}.\\
\int_{\mathbb{R}^d} q^\alpha dy &= \int r^\alpha p^\alpha dy\\
&= \int r^\alpha p^{1/\alpha} P^{\frac{\alpha-1}{\alpha}} dy \leq \left(\int r^{\alpha^2} p dy\right)^{1/\alpha} \times
\end{aligned}
$$

$$\times\left(\int p^{\alpha+1}dy\right)^{\alpha-1/\alpha}$$

273 **Step (iii).**

$$\begin{aligned}
Q_x(X(t)\in A) &= \int q(t,x,y)dy\\
&= \int r(t,x,y)p(t,x,y)dy\\
&= \int r(t,x,y)P_x(X_t\in dy).
\end{aligned}$$

Therefore

$$\frac{dQ_x}{dP_x}\Big|_t^t = r(t,x,y)$$

Therefore

$$\begin{aligned}
&\left(\int r^{\alpha^2}pdy\right)^{1/\alpha} = \left\|\frac{dQ_x}{dP_x}\Big|_t^t\right\|^{\alpha}_{\alpha^2,P_x}\\
&\leq \left\|\frac{dQ_x}{dP_x}\Big|_t\right\|^{\alpha}_{\alpha^2,P_x}, \quad \text{since} \quad \mathscr{F}_t^t\subset\mathscr{F}_t,\\
&\quad \{E^{P_x}[Z(t)^{\alpha^2}]\}^{1/\alpha}\\
&= \{E^{P_x}[\exp(\alpha^2\int_0^t\langle b,dX\rangle - \frac{\alpha^2}{2}\int_0^t|b|^2ds)]\}^{1/\alpha}\\
&= \{E^{P_x}[\exp(\alpha^2\int_0^t\langle b,dX\rangle - \frac{\alpha^4}{2}\int_0^t|b|^2ds + \frac{\alpha^4-\alpha^2}{2}\int_0^t|b|^2ds)]\}^{1/\alpha},
\end{aligned}$$

i.e.,

$$\left(\int r^{\alpha^2}pdy\right)^{1/\alpha} \leq \left\{E^{P_x}\left[\exp\left(\frac{\alpha^4-a^2}{2}ct+\alpha^2\int_0^t\langle b,dX\rangle-\frac{\alpha^4}{2}\int_0^t|b|^2ds\right)\right]\right\}^{1/\alpha}$$

where c is such that $|b|^2\leq c$. Using Schwarz inequality we then get

$$\left(\int r^{\alpha^2}pdy\right)^{1/\alpha} \leq \left[\exp\left(\frac{\alpha^4-\alpha^2}{2}ct\right)\right]^{1/\alpha}.$$

274 Hence

$$\int q^{\alpha} dy \leq \left(\exp\left[\frac{\alpha^4 - \alpha^2}{2} ct\right]\right)^{1/\alpha} \left(\int P^{\alpha+1} dy\right)^{\alpha - 1/\alpha}$$

Significance. Pure analytical objects like $q(t, x, y)$ can be studied using stochastic integrals.

Appendix

Language of Probability

275 **Definition.** A probability space is a measure space $(\Omega, \mathscr{B}, P)$ with $P(\Omega) = 1$. P is called a *probability measure* or simply a *probability*. Elements of $\mathscr{B}$ are called *events*. A measurable function $X : (\Omega, \mathscr{B}) \to R^d$ is called d-dimensional *random variable*. Given the random variable X, define $F : \mathbb{R}^d \to \mathbb{R}$ by

$$F((a_1, \ldots a_n)) = P\{w : X_i(w) < a_i, \text{ for } i = 1, 2, \ldots, d\}$$

where $X = (X_1, X_2, \ldots, X_d)$. Then F is called the *distribution function* of the random variable X. For any random variable X, $\int X\, dP = (\int X_1 dP, \ldots, \int X_d dP)$, if it exists, is called *mean* of X or *expectation* of X and is denoted by $E(X)$. Thus $E(X) = \int X dP = \mu$. $E(X^n)$, where $X^n = (X_1^n, X_2^n, \ldots, X_d^n)$ is called the n^{th} *moment* about zero. $E((X - \mu)^n)$ is called the n^{th} *central moment*. The 2nd central moment is called *variance* and is denoted by σ^2 we have the following.

Tchebyshev's Inequality.

Let X be a one-dimensional random variable with mean and variance μ. Then for every $\epsilon > 0$, $P\{w : |X(w) - \mu| \geq \epsilon\} \leq \sigma^2/\epsilon^2$.

Generalised Tchebyshev's Inequality. Let $f : \mathbb{R} \to \mathbb{R}$ be measurable
276 such that $f(u) = f(-u)$, f is strictly positive and increasing on $(0, \infty)$. Then for any random variable $X : \Omega \to R$,

$$P(w : |X(w)| > \epsilon) \leq \frac{E(f(X))}{f(\epsilon)}$$

for every $\epsilon > 0$.

For any random variable $X : \Omega \to R^d$, $\phi(t) = E(e^{itX}) : R^d \to C$ is called the *characteristic function* of X. Here $t = (t_1, \ldots, t_d)$ and $tX = t_1X_1 + t_2X_2 + \cdots + t_dX_d$.

Independence. Events $E_1, \ldots, E_n$ are called *independent* if for every $\{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$ we have

$$P(E_{i_1} \cap \ldots \cap E_{i_k}) = P(E_{i_1}) P(E_{i_2}) \ldots P(E_{i_k}).$$

An arbitrary collection of events $\{E_\alpha : \alpha \in I\}$ is called independent if every finite sub-collection is independent. Let $\{\mathscr{F}_\alpha : \alpha \in I\}$ be a collection of sub-σ-algebras of $\mathscr{B}$. This collection is said to be independent if for every collection $\{E_\alpha : \alpha \in I\}$, where $E_\alpha \in \mathscr{F}_\alpha$, of events is independent. A collection of random variables $\{X_\alpha : \alpha \in I\}$ is said to be independent if $\{\sigma(X_\alpha) : \alpha \in I\}$ is independent where $\sigma(X_\alpha)$ is the σ-algebra generated by X_α.

Theorem . *Let $X_1, X_2, \ldots, X_n$ be random variables with $F_{X_1}, \ldots, F_{X_n}$ as their distribution functions and let F be distribution function of $X = (X_1, \ldots, X_n)$, $\phi_{X_1}, \ldots, \phi_{X_n}$ the characteristic functions of $X_1, \ldots, X_n$ and ϕ that of $X = (X_1, \ldots, X_n)$. $X_1, \ldots, X_n$ are independent if and only if $F((a_1, \ldots, a_n)) = F_{X_1}(a_1) \ldots F_{X_n}(a_n)$ for all $a_1, \ldots, a_n$, iff $\phi((t_1, \ldots, t_n))$ $= \phi_{X_1}(t_1) \ldots \phi_{X_n}(t_n)$ for all $t_1, \ldots, t_n$.* **277**

Conditioning.

Theorem . *Let $X : (\Omega, \mathscr{B}, P) \to \mathbb{R}^d$ be a random variable, with $E(X)$ finite, i.e. if $X = (X_1, \ldots, X_d)$, $E(X_i)$ is finite for each i. Let $\mathscr{C}$ be a sub-σ-algebra of $\mathscr{B}$. Then there exists a random variable $Y : (\Omega, \mathscr{C}) \to \mathbb{R}^d$ such that $\int\limits_C Y dP = \int\limits_C X dP$ for every C in $\mathscr{C}$.*

If Z is any random variable with the same properties then $Y = Z$ almost everywhere (P).

Definition. Any such Y is called the *conditional expectation of X with respect to $\mathscr{C}$* and is denoted by $E(X|\mathscr{C})$.

If $X = \chi_A$, the characteristic function of A in $\mathscr{B}$, then $E(\chi_A|\mathscr{C})$ is also denoted by $P(A|\mathscr{C})$.

Properties of conditional expectation.

1. $E(1|\mathscr{C}) = 1$.

2. $E(aX+bY|\mathscr{C}) = aE(X|\mathscr{C})+bE(Y|\mathscr{C})$ for all real numbers a, b and random variables X, Y.

3. If X is a one-dimensional random variable and $X \geq 0$, then $E(X|\mathscr{C}) \geq 0$.

4. If Y is a bounded $\mathscr{C}$-measurable real valued random variable and X is a one-dimensional random variable, then

$$E(YX|\mathscr{C}) = YE(X|\mathscr{C}).$$

5. If $\mathscr{D} \subset \mathscr{C} \subset \mathscr{B}$ are σ-algebras, then

$$E(E(X|\mathscr{C})|\mathscr{D}) = E(X|\mathscr{D}).$$

278 6. $\int\limits_{\Omega} |E(X|\mathscr{D})| d(P|\mathscr{D}) \leq \int\limits_{\Omega} E(X|\mathscr{C}) d(P|\mathscr{C})$.

Exercise 1. Let $(\Omega, \mathscr{B}, P)$ be a probability space, $\mathscr{C}$ a sub-σ-algebra of $\mathscr{B}$. Let $X(t, \cdot)Y(t, \cdot) : \Omega \to R$ be measurable with respect to $\mathscr{B}$ and $\mathscr{C}$ respectively where t ranges over the real line. Further let $E(X(t, \cdot)|\mathscr{C}) = Y(t, \cdot)$ for each t. If f is a simple $\mathscr{C}$-measurable function then show that

$$\int\limits_{C} X(f(w), w) d(P|\mathscr{C}) = \int\limits_{C} Y(f(w)w) dP$$

for every C in $\mathscr{C}$.

[**Hint.** Let $A_1, \ldots, A_n$ be a $\mathscr{C}$-measurable partition such that f is constant on each A_i. Verify the equality when C is replaced by $C \cap A_i$.]

Exercise 2. Give conditions on X, Y such that exercise 1 is valid for all bounded $\mathscr{C}$-measurable functions and prove your claim.

The next lemma exhibits conditioning as a projection on a Hilbert space.

Lemma . *Let $(\Omega, \mathscr{B}, P)$ be any probability space $\mathscr{C}$ a sub-σ-algebra of $\mathscr{B}$. Then*

(a) *$L^2(\Omega, \mathscr{C}, P)$ is a closed subspace of $L^2(\Omega, \mathscr{B}, P)$.*

(b) *If $\pi : L^2(\Omega, \mathscr{B}, P) \to L^2(\Omega, \mathscr{C}, P)$ is the projection,then $\pi(f) = E(f|\mathscr{C})$.*

Proof. (a) is clear, because for any $f \in L^1(\Omega, \mathscr{C}, P)$

$$\int\limits_{\Omega} f d(P|\mathscr{C}) = \int\limits_{\Omega} f dP$$

(use simple function $0 \leq s_1 \leq \ldots \leq f$, if $f \geq 0$) and $L^2(\Omega, \mathscr{C}, P)$ **279**
is complete.

(b) To prove this it is enough to verify it for characteristic functions because both π and $f \to E(f|\mathscr{C})$ are linear and continuous.

Let $A \in \mathscr{B}$, $C \in \mathscr{C}$ then $\pi(\chi_C) = \chi_C$. As π is a projection

$$\int \pi(\chi_A)\overline{\chi}_C d(P|\mathscr{B}) = \int \chi_A \overline{\pi(\chi_C)} d(P|\mathscr{B}),$$

i.e.

$$\int\limits_{C} \pi(\chi_A) d(P|\mathscr{B}) = \int\limits_{C} X_A d(P|\mathscr{B}).$$

Since $\pi(\chi_A)$ is $\mathscr{C}$-measurable,

$$\int\limits_{C} \pi(\chi_A) d(P|\mathscr{B}) = \int\limits_{C} \pi(\chi_A) d(P|\mathscr{C})$$

Therefore

$$\int\limits_{C} \pi(\chi_A) d(P|\mathscr{C}) = \int\limits_{C} \chi_A d(P|\mathscr{B}), \ \forall C \text{ in } \mathscr{C}.$$

Hence

$$\pi(\chi_A) = E(\chi_A|\mathscr{C}).$$

□

Kolmogorov's Theorem.

Statement. Let A be any nonempty set and for each finite ordered subset $(t_1, t_2, \ldots, t_n)$ of A [i.e. $(t_1, \ldots, t_n)$ an ordered n-tuple with t_i in A], let
280 $P_{(t_1,\ldots,t_n)}$ be a probability on the Borel sets in $R^{dn} = R^d \times R^d \times \cdots R^d$. Assume that the family $P_{(t_1,\ldots,t_n)}$ satisfies the following two conditions

(i) Let $\tau : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ be any permutation and $f_\tau : R^{dn} \to R^{dn}$ be given by

$$f_\tau((x_1, \ldots, x_n)) = (x_{\tau(1)}, \ldots, x_{\tau(n)}).$$

We have

$$\underset{(t_{\tau(1)},\ldots,t_{\tau(n)})}{P}(E) = P_{(t_1,\ldots,t_n)}(f_\tau^{-1}(E))$$

for every Borel set E of R^{dn}. In short, we write this condition as $P_{\tau t} = P_t \tau^{-1}$.

(ii) $\underset{(t_1,\ldots,t_n)}{P}(E) = P_{(t_1,t_2,\ldots,t_n,t_{n+1},\ldots t_{n+m})}(E \times R^{dm})$ for all Borel sets E of R^{dn} and this is true for all $t_1, \ldots, t_n, t_{n+1}, \ldots, t_{n+m}$ of A.

Then, there exists a probability space $(\Omega, \mathscr{B}, P)$ and a collection of random variable $\{X_t : t \in A\} : (\Omega, \mathscr{B}) \to R^d$ such that

$$\underset{(t_1,\ldots,t_n)}{P}(E) = P\{w : (X_{t_1}(w), \ldots, X_{t_n}(w)) \in E\}$$

for all Borel sets E of R^{dn}.

Proof. Let $\Omega = \pi\{R_t^d : t \in A\}$ where $R_t^d = R^d$ for each t. Define $X_t : \Omega \to R^d$ to be the projection given by $X_t(w) = w(t)$. Let $\mathscr{B}_0$ be the algebra generated by $\{X_t : t \in A\}$ and $\mathscr{B}$ the σ-algebra generated by $\{X_t : t \in A\}$. Having got Ω and $\mathscr{B}$ we have to construct a probability P on $(\Omega, \mathscr{B})$ satisfying the conditions of the theorem.

281 Given $t_1, \ldots, t_n$ define

$$\pi_{(t_1,\ldots,t_n)} : \Omega \to R^d \times R^d \times \cdots \times R^d (n \text{ times})$$

by

$$\pi_{(t_1,\ldots,t_n)}(w) = (w(t_1), \ldots, w(t_n)).$$

It is easy to see that every element of $\mathscr{B}_0$ is $\pi^{-1}_{(t_1,\ldots,t_n)}(E)$ for suitable $t_1, \ldots, t_n$ in A and a suitable Borel set E of R^{dn}. Define P on $\mathscr{B}_0$ by $P(\pi^{-1}_{(t_1,\ldots,t_n)}(E)) = P_{(t_1,\ldots,t_n)}(E)$. Conditions (1) and (2) ensure that P is a well-defined function on $\mathscr{B}_0$ and that, as $P_{(t_1,\ldots,t_n)}$ are measures, P is finitely additive on $\mathscr{B}_0$. $\square$

Claim. Let $C_1 \supset C_2 \supset \ldots \supset C_n \supset \ldots$ be a decreasing sequence in $\mathscr{B}_0$ with $\lim\limits_{n\to\infty} P(C_n) \geq \delta > 0$. Then $\cap C_n$ is non-empty. Once the claim is proved, by Kolmogorov's theorem on extension of measures, the finitely additive set function P can be extended to a measure P on $\mathscr{B}$. One easily sees that P is a required probability measure.

Proof of the Claim. As $C_n \in \mathscr{B}_0$, we have

$$C_n = \pi^{-1}_{(t_1^{(n)},\ldots,t_{k(n)}^{(n)})}(E_n) \text{ for suitable } t_i^{(n)} \text{ in } A$$

and Borel set E_n in $R^{dk(n)}$. Let

$$T_n = (t_1^{(n)}, \ldots, t_{k(n)}^{(n)}) \quad \text{and} \quad A_n = \{t^{(n)}, \ldots, t_{k(n)}^{(n)}\}.$$

We can very well assume that A_n is increasing with n. Choose a compact **282**
subset E'_n of E_n such that

$$P_{T_n}(E_n - E'_n) \leq \delta/2^{n+1}.$$

If $C'_n = \pi^{-1}_{T_n}(E'_n)$, then $P(C_n - C'_n) \leq \delta/2^{n+1}$. If $C''_n = C'_1 \cap C'_2 \cap \ldots \cap C'_n$ then $C''_n \subset C'_n \subset C_n$, C''_n is decreasing and

$$P(C''_n) \geq P(C_n) - \sum_{i=1}^{n} P(C_i - C'_i) \geq \delta/2.$$

We prove $\cap C''_n$ is not empty, which proves the claim.

Choose w_n in C''_n. As $\pi_{T_1}(w_n)$ is in the compact set E_1 for all n, choose a subsequence

$$n_1^{(1)}, n_2^{(1)}, \ldots \text{ of } 1, 2, \ldots \text{ such that } \pi_{T_1}(w_{n_k}(1))$$

converges as $k \to \infty$. But for finitely many $n_k^{(1)}$'s, $\pi_{T_2}(\omega_{n_m}(1))$ is in the compact set E_2'. As before choose a subsequence $n_k^{(2)}$ of $n_k^{(1)}$ such that $\pi_{T_1}(\omega_{n_k}(2))$ converges as $k \to \infty$. By the diagonal process obtain a subsequence, w_n^* of w_n such that $\pi_{T_m}(w_n^*)$ converges as $n \to \infty$ for all m. Thus, if t is in

$$\bigcup_{m=1}^{\infty} A_m, \quad \text{then} \quad \underset{n\to\infty}{\text{limit}}\, w_n^*(t) = x_t$$

exists. Define w by $w(t) = 0$ if $t \in A \bigcup\limits_{m=1}^{\infty} A_m$, $w(t) = x_t$ if $t \in \bigcup\limits_{m=1}^{\infty} A_m$. One easily sees that $w \in \bigcap\limits_{n=1}^{\infty} C_n''$, completing the proof of the theorem.

283 **Martingales.**

Definition. Let $(\Omega, \mathscr{F}, P)$ be a probability space, $(T, \leq)$ a totally ordered set. Let $(\mathscr{F}_t)_{t \in T}$ be an increasing family of sub-σ-algebras of $\mathscr{F}$. A collection $(X_t)_{t \in T}$ of random variables on Ω is called a *martingale* with respect to the family $(\mathscr{F}_t)_{t \in T}$ if

(i) $E(|X_t|) < \infty$, $\forall t \in T$;

(ii) X_t is $\mathscr{F}_t$-measurable for each $t \in T$;

(iii) $E(X_t | \mathscr{F}_s) = X_s$ a.s. for each s, t in T with $t \geq s$. (Markov property).

If instead of (iii) one has

(iii)′ $E(X_t | \mathscr{F}_s) \geq (\leq) X_s$ a.s.,

then $(X_t)_{t \in T}$ is called a *submartingale* (respectively *supermartingale*).

From the definition it is clear that $(X_t)_{t \in T}$ is a submartingale if and only if $(-X_t)_{t \in T}$ is a supermartingale, hence it is sufficient to study the properties of only one of these. T is usually any one of the following sets

$$[0, \infty), N, Z, \{1, 2, \ldots, n\}, [0, \infty] \quad \text{or} \quad N \cup \{\infty\}.$$

Examples. (1) Let $(X_n)_{n=1,2\ldots}$ be a sequence of independent random variables with

$$E(X_n) = 0.$$

Then $Y_n = X_1 + \cdots + X_n$ is a martingale with respect to $(\mathscr{F}_n)_{n=1,2,\ldots}$ **284**
where

$$\mathscr{F}_n = \sigma\{Y_1, \ldots, Y_n\} = \sigma\{X_1, \ldots, X_n\}.$$

Proof. By definition, each Y_n is $\mathscr{F}_n$-measurable.

$$\begin{aligned}
&E(Y_n) = 0.\\
&E((X_1 + \cdots + X_n + X_{n+1} + \cdots + X_{n+m})|\sigma\{X_1, \ldots, X_n\})\\
&= X_1 + \cdots + X_n + E((X_{n+1} + \cdots + X_{n+m})|\sigma\{X_1, \ldots, X_n\})\\
&= Y_n + E(X_{n+1} + \cdots + X_{n+m}) = Y_n.
\end{aligned}$$

□

(2) Let $(\Omega, \mathscr{F}, P)$ be a probability space, Y a random variable with $E(|Y|) < \infty$. Let $\mathscr{F}_t \subset \mathscr{F}$ be a σ-algebra such that $\forall t \in [0, \infty)$

$$\mathscr{F}_t \subset \mathscr{F}_s \quad \text{if} \quad t \leq s.$$

If $X_t = E(Y|\mathscr{F}_t)$, X_t is a martingale with respect to $(\mathscr{F}_t)$.

Proof. (i) By definition, X_t is $\mathscr{F}_t$-measurable.

(ii) $E(X_t) = E(Y)$ (by definition) $< \infty$.

(iii) if $t \geq s$,

$$E(X_t|\mathscr{F}_s) = E(E(Y|\mathscr{F}_t)|\mathscr{F}_s) = E(Y|\mathscr{F}_s) = X_s$$

□

Exercise 1. $\Omega = [0, 1]$, $\mathscr{F} = \sigma$-algebra of all Borel sub sets of Ω, $P =$ Lebesgue measure.

Let $\mathscr{F}_n =$-algebra generated by the sets

$$\Big[0, \frac{1}{2^n}\Big)\Big[\frac{1}{2^n}, \frac{2}{2^n}\Big); \ldots, \Big[\frac{2^n - 1}{2^n}, 1\Big].$$

Let $f \in L'[0,1]$ and define **285**

$$X_n(w) = 2^n \left(\sum_{j=1}^{2^n-1} \chi_{[\frac{j-1}{2^n}, \frac{j}{2^n})} \int\limits_{j-1/2^n}^{j/2^n} f\, dy + \chi_{[\frac{2^n-1}{2^n}, 1]} \int\limits_{2^n-1/2^n}^{1} f\, dy \right)$$

Show that (X_n) is a martingale relative to $(\mathscr{F}_n)$.

Exercise. Show that a submartingale or a supermartingale $\{X_s\}$ is a martingale iff $E(X_s) =$ constant.

Theorem . *If $(X_t)_{t\in T}$, $(Y_t)_{t\in T}$ are supermartingales then*

(i) $(aX_t + bY_t)_{t\in T}$ *is a supermartingale,* $\forall a, b \in \mathbb{R}^+ = [0, \infty)$.

(ii) $(X_t \wedge Y_t)_{t\in T}$ *is a supermartingale.*

Proof. (i) Clearly $Z_t = aX_t + bY_t$ is $\mathscr{F}_t$-measurable and $E(|Z_t|) \leq aE(|X_t|) + bE(|Y_t|) < \infty$.

$$\begin{aligned} E(aX_t + bY_t|\mathscr{F}_s) &= aE(X_t|\mathscr{F}_s) + bE(Y_t|\mathscr{F}_s) \\ &\leq aX_s + bY_s = Z_s, \quad \text{if} \quad t \geq s. \end{aligned}$$

(ii) Again $X_t \wedge Y_t$ is $\mathscr{F}_t$-measurable and $E(|X_t \wedge Y_t|) < \infty$,

$$E(X_t \wedge Y_t|\mathscr{F}_s) \leq E(X_t|\mathscr{F}_s) \leq X_s.$$

286 Similarly

$$E(X_t \wedge Y_t|\mathscr{F}_s) \leq E(Y_t|\mathscr{F}_s) \leq Y_s, \quad \text{if} \quad t \geq s.$$

Therefore

$$E(X_t \wedge Y_t|\mathscr{F}_s) \leq X_s \wedge Y_s.$$

□

Jensen's Inequality. Let X be a random variable in $(\Omega, \mathscr{B}, P)$ with $E(|X|) < \infty$ and let $\phi(x)$ be a convex function defined on the real line such that $E(|\phi_0 X|) < \infty$. Then

$$\phi(E(X|\mathscr{C})) \leq E(\phi_0 X|\mathscr{C}) \quad \text{a.e.}$$

where $\mathscr{C}$ is any sub-σ-algebra of $\mathscr{B}$.

Proof. The function ϕ being convex, there exist sequences $a_1, a_2, \ldots a_n, \ldots, b_1, b_2, \ldots$ of real numbers such that $\phi(x) = \sup\limits_n (a_n x + b_n)$ for each x. Let $L_n(x) = a_n x + b_n$. Then

$$L_n(E(X|\mathscr{C})) = E(L_n(X)|\mathscr{C}) \leq E(\phi(X)|\mathscr{C})$$

for all n so that

$$\phi(E(X|\mathscr{C})) \leq E(\phi(X)|\mathscr{C}).$$

$\square$

Exercise. (a) If $\{X_t : t \in T\}$ is a martingale with respect to $\{\mathscr{F}_t : t \in T\}$ and ϕ is a convex function on the real line such that $E(|\phi(X_t)|) < \infty$ for every t, then $\{\phi(X_t)\}$ is a sub martingale.

(b) If $(X_t)_{t\in T}$ is a submartingale and $\phi(x)$ is a convex function and nondecreasing and if $E(|\phi_0 X_t|) < \infty$, $\forall t$ then $\{\phi(X_t)\}$ is a submartingale. (Hint: Use Jensen's inequality).

Definition. Let $(\Omega, \mathscr{B}, P)$ be a probability space and $(\mathscr{F}_t)_{t\in[0,\infty)}$ an in- **287**
creasing family of sub-σ-algebras of $\mathscr{F}$. Let $(X_t)_{t\in[0,\infty)}$ be a family of random variables on Ω such that X_t is $\mathscr{F}_t$-measurable for each $t \geq 0$. (X_t) is said to be *progressively measurable* if

$$X : [0, t] \times \Omega \to \mathbb{R} \quad \text{defined by} \quad X(s, w) = X_s(w)$$

is measurable with respect to the σ-algebra $\mathscr{B}[0, t] \times \mathscr{F}_t$ for every t.

Stopping times. Let us suppose we are playing a game of chance, say, tossing a coin. The two possible outcomes of a toss are H (Heads) and T (Tails). We assume that the coin is unbiased so that the probability of getting a head is the same as the probability of getting a tail. Further suppose that we gain +1 for every head and lose 1 for every tail. A game of chance of this sort has the following features.

1. A person starts playing with an initial amount N and finishes with a certain amount M.

2. Certain rules are specified which allow one to decide when to stop playing the game. For example, a person may not have sufficient money to play all the games, in which case he may decide to play only a certain number of games.

It is obvious that such a game of chance is fair in that it is neither ad-
288 vantageous nor disadvantageous to play such a game and on the average M will equal N, the initial amount. Furthermore, the stopping rules that are permissible have to be reasonable. The following type of stopping rule is obviously unreasonable.

Rule. If the first toss is a tail the person quits at time 0 and if the first toss is a head the person quits at time $t = 1$.

This rule is unreasonable because the decision to quit is made on the basis of a future event, whereas if the game is fair this decision should depend only on the events that have already occured. Suppose, for example, 10 games are played, then the quitting times can be 0, $1, 2, \ldots, 10$. If $\xi_1, \ldots, \xi_{10}$ are the outcomes ($\xi_i = +1$ for H, $\xi_i = -1$ for T) then the quitting time at the 5th stage (say) should depend only on $\xi_1, \ldots, \xi_4$ and not any of $\xi_5, \ldots, \xi_{10}$. If we denote $\xi = (\xi_1, \ldots, \xi_{10})$ and the quitting time τ as a function of ξ then we can say that $\{\xi : \tau = 5$ depends only $\xi_1, \ldots, \xi_4\}$. This leads us to the notion of stopping times.

Definition. Let $(\Omega, \mathscr{F}, P)$ be a probability space, $(\mathscr{F}_t)_{t \in [0,\infty)}$ an increasing family of sub-σ-algebras of $\mathscr{F}$. $\tau : \Omega \to [0, \infty]$ is called a *stopping time* or *Markov time* (or a random variable independent of the future) if

$$\{w : \tau(w) \leq t\} \in \mathscr{F}_t \quad \text{for each} \quad t \geq 0.$$

Observe that a stopping time is a measurable function with respect to $\sigma(\cup \mathscr{F}_t) \subset \mathscr{F}$.

289 **Examples.** 1. τ = constant is a stopping time.

2. For a Brownian motion (X_t), the hitting time of a closed set is stopping time.

Exercise 2. Let $\mathscr{F}_{t+} \underset{\text{Def}}{\equiv} \bigcap\limits_{s>t} \mathscr{F}_s \equiv \mathscr{F}_t$.

[If this is satisfied for every $t \geq 0$, $\mathscr{F}_t$ is said to be *right continuous*]. If $\{\tau < t\} \in \mathscr{F}_t$ for each $t \geq 0$, then τ is a stopping time. (Hint: $\{\tau \leq t = \bigcap\limits_{n=k}^{\infty} \{\tau < t + 1/n\}$ for every k).

We shall denote by $\mathscr{F}_\infty$ the σ-algebra generated by $\bigcup\limits_{t \in T} \mathscr{F}_t$. If τ is a stopping time, we define

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty : A \cap \{\tau \leq t\} \in \mathscr{F}_t, \forall t \geq 0\}$$

Exercise 3. (a) Show that $\mathscr{F}_\tau$ is a σ-algebra. (If $A \in \mathscr{F}_\tau$,

$$A^c \cap \{\tau \leq t\} = \{t \leq t\} - A \cap \{\tau \leq t\}).$$

(b) If $\tau = t$ (constant) show that $\mathscr{F}_\tau = \mathscr{F}_t$.

Theorem . *Let τ and σ be stopping times. Then*

(i) *$\tau + \sigma$, $\tau v \sigma$, $\tau \wedge \sigma$ are all stopping times.*

(ii) *If $\sigma \leq \tau$, then $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$.*

(iii) *τ is $\mathscr{F}_\tau$-measurable.*

(iv) *If $A \in \mathscr{F}_\sigma$, then $A \cap \{\sigma = \tau\}$ and $A \cap \{\sigma \leq \tau\}$ are in $\mathscr{F}_{\sigma \wedge \tau} \subset \mathscr{F}_\sigma \cap \mathscr{F}_\tau$. In particular, $\{\tau < \sigma\}$, $\{\tau = \sigma\}$, $\{\tau > \sigma\}$ are all in* **290**
$\mathscr{F}_\tau \cap \mathscr{F}_\sigma$.

(v) *If τ' is $\mathscr{F}_\tau$-measurable and $\tau' \geq \tau$, then τ' is a stopping time.*

(vi) *If $\{\tau_n\}$ is a sequence of stopping times, then $\underline{\lim}\, \tau_n$. $\overline{\lim}\, \tau_n$ are also stopping times provided that $\mathscr{F}_{t+} = \mathscr{F}_t$, $\forall t \geq 0$.*

(vii) *If $\tau_n \downarrow \tau$, then $\mathscr{F}_\tau = \bigcap\limits_{n=1}^{\infty} \mathscr{F}_{\tau_n}$ provided that $\mathscr{F}_{t+} = \mathscr{F}_t$, $\forall t \geq 0$.*

Proof. (i)

$$\{\sigma + \tau\} > t\} = \{\sigma + \tau > t, \tau \leq t, \sigma \leq t\} \cup \{\tau > t\} \cup \{\sigma > t\};$$
$$\{\sigma + \tau > t, \sigma \leq t\} = \tau \leq \mathscr{A}$$

$$= \bigcup_{\substack{r\in\mathscr{Q}\\ 0\le r\le t}} \{\sigma > r > t-\tau, \tau \le t, \sigma \le t\}$$

$$(\mathscr{Q} = \text{ set of rationals})$$

$$\{\sigma > r > t-\tau, \tau \le t, \sigma \le t\} = \{t \ge \sigma > r\} \cap \{t \ge \tau > t-r\}$$
$$= \{\sigma \le t\} \cap \{\sigma \le r\}^c \cap \{\tau \le t\} \cap \{\tau \le t-r\}^c.$$

The right side is in $\mathscr{F}_t$. Therefore $\sigma + \tau$ is a scopping time.

$$\{\tau V \sigma \le t\} = \{\tau \le t\} \cap \{\sigma \le t\}$$
$$\{\tau \wedge \sigma > t\} = \{\tau > t\} \cap \{\sigma > t\}$$

(ii) Follows from (iv).

(iii) $\{\tau \le t\}\{\tau \le s\} = \{\tau \le t \wedge s\} \in \mathscr{F}_{t\wedge s} \subset \mathscr{F}_s, \forall s \ge 0.$

(iv) $A \cap \{\sigma < \tau\} \cap \{\sigma \wedge \tau \le t\} = [A \cap \{\sigma \le t < \tau\}]$

$U[A \cap \underset{\substack{r\in\mathscr{Q}\\ 0\le r\le t}}{U} \{\sigma \le < \tau\} \cap \{\tau \le t\}] \in \mathscr{F}_t.$

291 $A \cap \{\sigma \le \tau\} \cap \{\sigma \wedge \tau \le t\} = A \cap \{\sigma \le \tau\} \cap \{\sigma \le t\}.$

It is now enough to show that $(\sigma \le \tau) \in \mathscr{F}_\sigma$; but this is obvious because $(\tau < \sigma) = (\sigma \le \tau)^c$ is in $\mathscr{F}_{\sigma\wedge\tau} \subset \mathscr{F}_\sigma$. Therefore $A \cap \{\sigma \le \tau\} \in \mathscr{F}_{\sigma\wedge\tau}$ and (iv) is proved.

(v) $\{\tau' \le t\} = \{\tau' \le t\} \cap \{\tau \le t\} \in \mathscr{F}_t$ as $(\tau' \le t) \in \mathscr{F}_\tau$. Therefore τ' is a stopping time.

(vi) $\underline{\lim}\, \tau_n \equiv \sup_n \inf_{k\ge n} \tau_k$

$$= \sup_n \inf_\ell \inf\{\tau_n, \tau_{n+1}, \ldots, \tau_{n+\ell}\}.$$

By (i), $\inf\{\tau_n, \tau_{n+1}, \ldots, \tau_{n+\ell}\}$ is a stopping time. Thus we have only to prove that if $\tau_n \uparrow \tau$ or $\tau_n \downarrow \tau$ where τ_n are stopping times, then τ is a stopping time. Let $\tau_n \uparrow \tau$. Then $\{\tau \le t\} = \bigcap_{n=1}^{\infty} \{\tau_n \le t\}$ so that τ is a stopping time. Let $\tau_n \downarrow \tau$. Then

$$\{\tau \ge t\} = \bigcap_{n=1}^{\infty} \{\tau_n \ge t\}.$$

By Exercise 3, τ is a stopping time. That $\overline{\lim}\, \tau_n$ is a stopping time is proved similarly.

(vii) Since $\tau \le \tau_n$, $\forall n$, $\mathscr{F}_\tau \subset \bigcap_{n=1}^{\infty} \mathscr{F}_{\tau_n}$. Let $A \in \bigcap_{n=1}^{\infty} \mathscr{F}_{\tau_n}$. Therefore $A \cap (\tau_n < t) \in \mathscr{F}_t$, $\forall n$. $A \cap (\tau < t) = \bigcap_{m=1}^{\infty} (A \cap (\tau_m < t)) \in \mathscr{F}_t$. Therefore $A \in \mathscr{F}_\tau$.

□

Optional Sampling Theorem. (Discrete case). *Let $\{X_1, \dots, X_k\}$ be a martingale relative to $\{\mathscr{F}_1, \dots, \mathscr{F}_k\}$. Let $\{\tau_1, \dots, \tau_p\}$ be a collection of stopping times relative to $\{\mathscr{F}_1, \dots, \mathscr{F}_k\}$ such that $\tau_1 \le \tau_2 \le \dots \le \tau_p$* **292**
a.s. and each τ_i takes values in $\{1, 2, \dots, k\}$. Then $\{X_{\tau_1}, \dots, X_{\tau_p}\}$ is a martingale relative to $\{\mathscr{F}_{\tau_1}, \dots, \mathscr{F}_{\tau_p}\}$ where for any stopping time τ, $X_\tau(\omega) = X_{\tau(w)}(\omega)$.

Proof. It is easy to see that each X_{τ_i} is a random variable. In fact $X_{\tau_m} = \sum_{i=1}^{k} X_i \chi_{\{\tau_m = i\}}$. Let $\tau \in \{1, 2, \dots, k\}$. Then

$$E(|X_\tau|) \le \sum_{j=1}^{k} \int |X_j| dP < \infty.$$

Consider

$$(X_{\tau_j} \le t) \cap (\tau_j \le s) = \bigcap_{\ell \le s} (X_\ell \le t) \in \mathscr{F}_s.$$

Then $(X_{\tau_j} \le t)$ is in $\mathscr{F}_{\tau_j}$, i.e. X_{τ_j} is $\mathscr{F}_{\tau_j}$-measurable. Next we show that

(*) $$E(X_{\tau_j} | \mathscr{F}_{\tau_k}) \le X_{\tau_k}, \quad \text{if} \quad j \ge k.$$

(*) is true if and only if

$$\int_A X_{\tau_j} dP \le \int_A X_{\tau_k} dP \quad \text{for every} \quad A \in \mathscr{F}_{\tau_k}.$$

The theorem is therefore a consequence of the following □

Lemma . *Let* $\{X_1, \ldots, X_k\}$ *be a supermartingale relative to*

$$\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}.$$

If τ *and* σ *are stopping times relative to* $\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}$ *taking values in* $\{1, 2, \ldots, k\}$ *such that* $\tau \leq \sigma$ *then*

$$\int_A X_\tau dP \geq \int_A X_\sigma dP \quad \text{for every} \quad A \in \mathscr{F}_\tau.$$

293 *Proof.* Assume first that $\sigma - \tau \leq 1$. Then

$$\begin{aligned}\int_A (X_\tau - X_\sigma) dP &= \sum_{j=1}^{k} \int_{[A\cap(\tau=j)\cap(\tau<\sigma)]} (X_\tau - X_\sigma) dP \\ &= \sum_{j=1}^{k} \int_{[A\cap(\tau=j)]} (X_j - X_{j+1}) dP\end{aligned}$$

$A \in \mathscr{F}_i$. Therefore $A \cap (\tau = j) \in \mathscr{F}_j$. By supermartingale property

$$\int_{[A\cap(\tau=j)]} (X_j - X_{j+1}) dP \geq 0.$$

Therefore

$$\int_A (X_\tau - X_\sigma) dP \geq 0.$$

Consider now the general case $\tau \leq \sigma$. Define $\tau_n = \sigma \wedge (\tau + n)$. Therefore $\tau_n \geq \tau$. τ_n is a stopping time taking values in $\{1, 2, \ldots, k\}$,

$$\tau_{n+1} \geq \tau_n, \quad \tau_{n+1} - \tau_n \leq 1, \quad \tau_k = \sigma.$$

Therefore $\int\limits_A X_{\tau_n} dP \geq \int\limits_A X_{\tau_{n+1}} dP$, $\forall A \in \mathscr{F}_{\tau_n}$. If $A \in \mathscr{F}_\tau$ then $A \in \mathscr{F}_{\tau_n}$, $\forall n$. Therefore

$$\int_A X_{\tau_1} dP \geq \int_A X_{\tau_2} dP \geq \ldots \geq \int_A X_{\tau_k} dP, \quad \forall A \in \mathscr{F}_\tau.$$

Now $\tau_1 - \tau \leq 1$. $\tau \leq \tau_1$. Therefore

$$\int\limits_A X_\tau dP \geq \int\limits_A X_{\tau_1} dP \geq \int\limits_A X_\sigma dP.$$

This completes the proof. □

N.B. The equality in (*) follows by applying the argument to **294**

$$\{-X_1, \ldots, -X_k\}.$$

Corollary 1. *Let* $\{X_1, X_2, \ldots, X_k\}$ *be a super-martingale relative to*

$$\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}.$$

If τ *is any stopping time, then*

$$E(X_k) \leq E(X_\tau) \leq E(X_1).$$

Proof. Follows from the fact that $\{X_1, X_\tau, X_k\}$ is a supermartingale relative to $\{\mathscr{F}_1, \mathscr{F}_\tau, \mathscr{F}_k\}$. □

Corollary 2. *If* $\{X_1, X_2, \ldots, X_k\}$ *is a super-martingale relative to*

$$\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}$$

and τ *is any stopping time, then*

$$E(X_\tau) \leq E(|X_1|) + 2E(X_k^-) \leq 3 \sup_{1 \leq n \leq k} E(|X_n|)$$

where for any real x, $x^- = \dfrac{|x| - x}{2}$.

Proof. $X_k^- = \dfrac{|X_k| - X_k}{2}$, so $2E(X_k^-) = E(|X_k|) - E(X_k)$.

By theorem $\{X_\tau \wedge 0, X_k \wedge 0\}$ is a super-martingale relative to $\{\mathscr{F}_\tau, \mathscr{F}_k\}$. Therefore $E(X_k \wedge 0 | \mathscr{F}_\tau) \leq E(X_\tau \wedge 0)$. Hence

$$E(X_k^-) \geq E(X_\tau^-) = \frac{E(|X_\tau|) - E(X_\tau)}{2}.$$

Therefore

$$\begin{aligned} E(|X_\tau|) &\leq 2E(X_k^-) + E(X_\tau) \\ &\leq 2E(X_k^-) + E(X_1) \leq 3 \sup_{1\leq n\leq k} E(|X_n|). \end{aligned}$$

□

Theorem . *Let $(\Omega, \mathscr{F}, P)$ be a probability space and $(\mathscr{F}_t)_{t\geq 0}$ on increasing family of sub-σ-algebras of $\mathscr{F}$. Let τ be a finite stopping time, and $(X_t)_{t\geq 0}$ a progressively measurable family (i.e. $X : [0,\infty)\times\Omega \to \mathbb{R}$ defined by $X(t,w) = X_t(w)$ is progressively measurable). If $X_\tau(w) = X_{\tau(w)}(w)$, then X_τ is $\mathscr{F}_\tau$-measurable.*

295

Proof. We show that $\{w : X(\tau(w), w) \leq t, \tau(w) \leq s\} \in \mathscr{F}_t$ for every t. Let $\Omega_s = \{w : \tau(w) \leq s\}$; $\Omega_s \in \mathscr{F}_s$ and hence the σ-algebra induced by $\mathscr{F}_s$ on Ω_s is precisely

$$\{A \cap \Omega_s : A \in \mathscr{F}_s\} = \{A \in \mathscr{F}_s : A \subset \Omega_s\}.$$

Since $\tau(w)$ is measurable,

$$w \to (\tau(w), w) \quad \text{of} \quad \Omega_s \to [0, s] \times \Omega_s$$

is $(\mathscr{F}_s, \mathscr{B}[0, s] \times \mathscr{F}_s)$-measurable. Since X is progressively measurable,

$$[0, s] \times \Omega_s \xrightarrow{X} \mathbb{R} \quad \text{is measurable.}$$

Therefore $\{w : X(\tau(w), w) \leq t, \tau(w) \leq s\} \in \sigma$-algebra on Ω_s. Therefore X_τ is $\mathscr{F}_\tau$ measurable.

The next theorem gives a condition under which $(X_t)_{t\geq 0}$ is progressively measurable. □

Theorem . *If X_t is right continuous in t, $\forall w$ and X_t is $\mathscr{F}_t$-measurable, $\forall t \geq 0$ then $(X_t)_{t\geq 0}$ is progressively measurable.*

Proof. Define

$$X_n(t, w) = X\left(\frac{[nt] + 1}{n}, w\right) \cdot \frac{[nt] + 1}{n} \downarrow t.$$

Then

$$\operatorname*{Lt}_{n\to\infty} X_n(t,w) = X(t,w) \quad \text{(by right continuity)}$$

296

Step 1. Suppose T is rational, $T = m/n$ where $m \geq 0$ is an integer. Then

$$\begin{aligned}&\{(t,w) : 0 \leq t < T,\ X_n(t,w) \leq \alpha\}\\ &= \bigcup_{0\leq i\leq m-1} \left\{\Big[\frac{i}{n}, \frac{i+1}{n}\Big) X \frac{X^{-1}_{i+1}}{n}(-\infty, \alpha]\right\}\end{aligned}$$

Thus if $T = m/n$, $X_n|_{[0,T)\times\Omega}$ is $\mathscr{B}[0,T] \times \mathscr{F}_T$-measurable. Now $T = \dfrac{km}{kn}$. Letting $k \to \infty$, by right continuity of $X(t)$ one gets $X|_{[0,)\times\Omega}$ is $[0,T] \times \mathscr{F}_T$-measurable. As $X(T)$ is $\mathscr{F}_T$-measurable, one gets $X|_{[0,T]\times\Omega}$ is $[0,T] \times \mathscr{F}_T$-measurable.

Step 2. Let T be irrational. Choose a sequence of rationals S_n increasing to T.

$$\begin{aligned}&\{(t,w) : 0 \leq t \leq T,\ X(t,w) \leq \alpha\}\\ &= \bigcup_{n=1}^{\infty} \{(t,w) : 0 \leq t \leq S_n, X(t,w) \leq \alpha\} \cup \{T\} \times X_T^{-1}(-\infty, \alpha]\end{aligned}$$

The countable union is in $\mathscr{B}[0,T] \times \mathscr{F}_T$ by Step 1. The second member is also in $\mathscr{B}[0,T] \times \mathscr{F}_T$ as $X(T)$ is $\mathscr{F}_T$-measurable. Thus $X|_{[0,T]\times\Omega}$ is $\mathscr{B}_{[0,T]} \times \mathscr{F}_T$-measurable when T is irrational also.

$\square$

Remark. The technique used above is similar to the one used for proving that a right continuous function $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable.

Theorem . *Let $\{X_1, \ldots, X_k\}$ be a supermartingale and $\lambda \geq 0$. Then* **297**

$$\begin{aligned}(1)\quad \lambda P(\sup_{1\leq n\leq k} X_n \geq \lambda) &\leq E(X_1) - \int\limits_{\left\{\substack{\sup X_n < \lambda\\ 1\leq n\leq k}\right\}} X_k dP\\ &\leq E(X_1) + E(X_k^-).\end{aligned}$$

$$(2)\quad \lambda P(\inf_{1\le n\le k} X_n \le -\lambda) \le -\int\limits_{\{\inf X_n \le -\lambda\}} X_k dP$$
$$\le E(X_k^-).$$

Proof. Define

$$\begin{aligned}\tau(w) &= \inf\{n : X_n \ge \lambda\} \quad \text{if} \quad \sup X_n \ge \lambda,\\ &= k, \quad \text{if} \quad \sup_n X_n < \lambda.\end{aligned}$$

Clearly $\tau \ge 0$ and τ is a stopping time. If $\tau < k$, then $X_\tau(w) \ge \lambda$ for each w.

$$\begin{aligned}E(X_\tau) &= \int\limits_{(\sup X_n \ge \lambda)} X_\tau dP + \int\limits_{(\sup X_n < \lambda)} X_\tau dP\\ &\ge \lambda P(\sup X_n \ge \lambda) + \int\limits_{(\sup X_n < \lambda)} X_k dP.\end{aligned}$$

Therefore

$$E(X_1) \ge \lambda P(\sup X_n \ge \lambda) + \int\limits_{(\sup X_n < \lambda)} X_k dP,$$

$$\lambda P(\sup X_n \ge \lambda) \le E(X_1) - \int\limits_{(\sup X_n < \lambda)} X_k dP \le E(X_1) + E(X_k^-)$$

The proof of (2) is similar if we define

$$\tau(w) = \begin{cases}\inf\{n : X_n \le -\lambda\}, & \text{if } \inf X_n \le -\lambda,\\ k, & \text{if } \inf X_n > -\lambda.\end{cases}$$

$\square$

298 **Kolmogorov's Inequality** (Discrete Case). *Let* $\{X_1, \ldots, X_k\}$ *be a finite sequence of independent random variables with mean* 0. *Then*

$$P\left(\sup_{1\le n\le k} (|X_1 + \cdots + X_n| \ge \lambda) \le \frac{1}{\lambda^2} E((X_1 + X_2 + \cdots + X_k)^2)\right)$$

Proof. If $S_n = X_1 + \cdots + X_n$, $n = 1, 2, \ldots, k$, then $\{S_1, \ldots, S_k\}$ is a martingale with respect to $\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}$ where $\mathscr{F}_n = \sigma\{X_1, \ldots, X_n\}$. Therefore $S_1^2, \ldots, S_k^2$ is a submartingale (since $x \to x^2$ is convex). By the previous theorem,

$$\lambda^2 P\{\inf -S_n^2 \leq -\lambda^2\} \leq E((-S_K^2)^-)$$

Therefore

$$\begin{aligned} P\{\sup |S_n| \geq \lambda\} &\leq \frac{E((-S_k^2)^-)}{\lambda^2} = \frac{E(S_k^2)}{\lambda^2} \\ &= \frac{1}{\lambda^2} E((X_1 + X_2 + \cdots + X_k)^2). \end{aligned}$$

□

Kolmogorov's Inequality (Continuous case). *Let $\{X(t) : t \geq 0\}$ be a continuous martingale with $E(X(0)) = 0$. If $0 < T < \infty$, then for any $\epsilon > 0$*

$$P\Big\{w : \sup_{0 \leq s \leq T} |X(s, w)| \geq \epsilon\Big) \leq \frac{1}{\epsilon^1} E((X(T))^2).$$

Proof. For any positive integer k define $Y_0 = X(0)$,

$$\begin{aligned} Y_1 &= X\left(\frac{T}{2^k}\right) - X(0),\ Y_2 = X\left(\frac{2T}{2^k}\right) - X\left(\frac{T}{2^k}\right), \ldots, Y_2 k \\ &= X\left(\frac{2^k T}{2^k}\right) - X\left(\frac{(2^k - 1)}{2^k} T\right). \end{aligned}$$

By Kolmogorov inequality for the discrete case, for any $\delta > 0$.

$$P\left(\sup_{0 \leq n \leq 2^k} |X\left(\frac{nT}{2^k}\right)| > \delta\right) \leq \frac{1}{\delta^2} E((X(T))^2).$$

By continuity of $X(t)$, $A_k = \{w : \sup_{0 \leq n \leq 2^k} |X\left(\frac{nT}{2^k}\right)| > \delta\}$ increases to **299**
$\{\sup_{0 \leq s \leq T} |X(s)| > \delta\}$ so that one gets

$$(1) \qquad P\left(\sup_{0 \leq s \leq T} |X(s)| > \delta\right) \leq \frac{1}{\delta^2} E((X(T))^2).$$

Now

$$P\left(\sup_{0\le s\le T}|X(s)|\ge\epsilon\right)\le\operatorname*{limit}_{m\to\infty}P\left(\sup_{0\le s\le T}|X(s)|>\epsilon-\frac{1}{m}\right)$$

$$\le\operatorname*{limit}_{m\to\infty}\frac{1}{(\epsilon-1/m)^2}E((X(T))^2),\quad\text{by (1).}$$

$$=1/\epsilon^2E((X(T))^2).$$

This completes the proof. □

Optional Sampling Theorem (Countable case). *Let $\{X_n : n\ge 1\}$ be a supermartingale relative to $\{\mathscr{F}_n : n\ge 1\}$. Assume that for some $X_\infty\in L^1$, $X_n\ge E(X_\infty|\mathscr{F}_n)$. Let σ, τ be stopping times taking values in $N\cup\{\infty\}$, with $\sigma\le\tau$. Define $X_\tau=X_\infty$ on $\{\sigma=\infty\}$ and $X_\tau=X_\infty$ on $\{\sigma=\infty\}$. Then $E(X_\tau|\mathscr{F}_\sigma)\le X_\sigma$.*

Proof. We prove the theorem in three steps.

Step 1. Let $X_\infty=0$ so that $X_n\ge 0$. Let $\tau_k=\tau\wedge k$, $\sigma_k=\tau\wedge k$. By optional sampling theorem for discrete case $E(X_{\tau_k})\le E(X_k)\le E(X_1)$. By Fatou's lemma, $E(X_\tau)<\infty$. Again by optional sampling theorem for the discrete case,

$$E(X_{\tau_k}|\mathscr{F}_{\sigma_k})\le X_{\sigma_k}\ldots,(0).$$

300 Let $A\in\mathscr{F}_\sigma$. Then $A\cap\{\sigma\le k\}\in\mathscr{F}_{\sigma_k}$, and by (0)

$$\int\limits_{A\cap\{\tau\le k\}}X_\tau dP\le\int\limits_{A\cap(\tau\le k)}X_{\tau_k}dP\le\int\limits_{A\cap(\sigma\le k)}X_{\sigma_k}dP\le\int\limits_{A\cap(\sigma\le k)}X_\sigma dP.$$

Letting $k\to\infty$,

$$\int\limits_{A\cap(\tau\ne\infty)}X_\tau dP\le\int\limits_{A\cap(\sigma\ne\infty)}X_\sigma dP. \tag{1}$$

Clearly

$$\int\limits_{A\cap(\tau=\infty)}X_\tau dP=\int\limits_A X_\infty dP=\int\limits_{A\cap(\sigma=\infty)}X_\sigma dP \tag{2}$$

By (1) and (2), $\inf_A X\,dP \le \int_A X_\sigma dP$, proving that

$$E(X_\tau|\mathscr{F}_\sigma) \le X_\sigma.$$

Step 2. Suppose $X_n = E(X_\infty|\mathscr{F}_n)$. In this case we show that $X_\tau = E(X_\infty|\mathscr{F}_\tau)$ for every stopping time so that $E(X_\tau|\mathscr{F}_\sigma) = X_\sigma$. If $A \in \mathscr{F}_\tau$, then

$$\int_{(\tau\le k)} X_\tau dP = \int_{A\cap(\tau\le K)} X_\infty dP \quad \text{for every} \quad k.$$

Letting $k \to \infty$,

$$\int_{A\cap(\tau\ne\infty)} X_\tau dP = \int_{A\cap(\tau\ne\infty)} X_\infty dP, \tag{1}$$

$$\int_{A\cap(\tau=\infty)} X_\tau dP = \int_A X_\infty dP = \int_{A\cap(\tau=\infty)} X_\infty dP \tag{2}$$

The assertion follows from (1) and (2).

Step 3. Let X_n be general. Then **301**

$$X_n = X_n - E(X_\infty|\mathscr{F}_n) + E(X_\infty|\mathscr{F}_n).$$

Apply Step (1) to $Y_n = X_n - E(X_\infty|\mathscr{F}_n)$ and Step (2) to

$$Z_n = E(X_\infty|\mathscr{F}_n)$$

to complete the proof.

□

Uniform Integrability.

Definition. Let $(\Omega, \mathscr{B}, P)$ be any probability space, $L^1 = L^1(\Omega, \mathscr{B}, P)$. A family $H \subset L^1$ is called *uniformly integrable* if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\int_{(|X|\ge\delta)} |X|dP < \epsilon$ for all X in H.

Note. Every uniformly integrable family is a bounded family.

Proposition . *Let X_n be a sequence in L^1 and let $X_n \to X$ a.e. Then $X_n \to X$ in L^1 iff $\{X_n : n \geq 1\}$ is uniformly integrable.*

Proof. is left as an exercise. ◻

As $\{X_n : n \geq 1\}$ is a bounded family, by Fatou's lemma $X \in L^1$. Let $\epsilon > 0$ be given. By Egoroff's theorem there exists a set F such that $P(F) < \epsilon$ and $X_n \to X$ uniformly on F.

$$
\begin{aligned}
\int |X_n - X| dP &\leq \|X_n - X\|_{\infty, \Omega - F^+} \int\limits_F |X_n - X| dP \\
&\leq \|X_n - X\|_{\infty, \Omega - F} + \int\limits_F |X_n| dP + \int\limits_F |X| dP \\
&\leq \|X_n - X\|_{\infty, \Omega - F} + \int\limits_{F \cap (|X_n| \geq \delta)} |X_n| dP + \int\limits_{F \cap (|X| \geq \delta)} |X| dP + \\
&\quad + \int\limits_{F \cap \{|X_n| \leq \delta\}} |X_n| dP + \int\limits_{F \cap (|X| \leq \delta)} X dP \\
&\leq \|X_n - X\|_{\infty, \Omega - F} + \int\limits_{(|X_n| \geq \delta)} |X_n| dP + \int\limits_{(|X| \geq \delta)} |X| dP + 2\delta\epsilon
\end{aligned}
$$

302

The result follows by uniform integrability of $\{X, X_n : n \geq 1\}$.

Corollary . *Let $\mathscr{C}$ be any sub-σ-algebra of $\mathscr{B}$. If $X_n \to X$ a.e. and X_n is uniformly integrable, then $E(X_n|\mathscr{C}) \to E(X|\mathscr{C})$ in $L^1(\Omega, \mathscr{C}, P)$.*

Proposition . *Let $H \subset L^1$. Suppose there exists an increasing convex function $G : [0, \infty) \to [0, \infty)$ such that*

$$
\operatorname*{limit}_{t \to \infty} \frac{G(t)}{t} = \infty \quad \textit{and} \quad \sup_{X \in H} E(G(|X|)) < \infty.
$$

Then the family H is uniformly integrable.

Example. $G(t) = t^2$ is a function satisfying the conditions of the theorem.

Proof. (of the proposition). Let

$$M = \sup_{X \in H} E(G(|X|)).$$

Let $\epsilon > 0$ be given. Choose $\delta > 0$ such that

$$\frac{G(t)}{t} \geq \frac{M}{\epsilon} \quad \text{for} \quad t \geq \delta.$$

Then for X in H

$$\int_{(|X| \geq \delta)} |X| dP \leq \frac{\epsilon}{M} \int_{(|X| \geq \delta)} G(|X|) dP \leq \frac{\epsilon}{M} \int_{G} G(|X|) dP \leq \epsilon$$

□

Remark. The converse of the theorem is also true.

Exercise. Let H be a bounded set in L^∞, i.e. there exists a constant M such that $\|X\|_\infty \leq M$ for all X in H. Then H is uniformly integrable.

Up Crossings and Down Crossings. **303**

Definition. Let $a < b$ be real numbers; let $s_1, s_2, \ldots, s_k$ be also given reals. Define $i_1, i_2, \ldots, i_k$ as follows.

$$i_1 = \begin{cases} \inf\{n : s_n < a\}, \\ k, \text{ if no } s_i < a; \end{cases}$$

$$i_2 = \begin{cases} \inf\{n > i_1 : s_n > b\}, \\ k, \text{ if } s_n \leq b \text{ for each } n > i_1; \end{cases}$$

$$i_3 = \begin{cases} \inf\{n > i_2 : s_n < a\}, \\ k, \text{ if } s_n \geq a \text{ for each } n > i_2; \end{cases}$$

and so on

Let $t_1 = s_{i_1}$, $t_2 = s_{i_2}, \ldots$. If (t_1, t_2), $(t_3, t_4), \ldots$, (t_{2p-1}, t_{2p}) are the only non-empty intervals and $(t_{2p+1}, t_{2p+2}), \ldots$ are all empty, then p is called the ******** of the sequence $s_1, \ldots, s_k$ for the interval $[a, b]$ and is denoted by $U(s_1, \ldots, s_k; [a, b])$.

Note. U (the up crossing) always takes values in $\{0, 1, 2, 3, \ldots\}$.

Definition. For any subset S of reals define

$$U(S; [a, b]) = \sup\{U(F; [a, b]) : F \text{ is a finite subset of } S\}$$

The number of down crossings is defined by

$$D(S; [a, b]) = U(-S; [-b, -a]).$$

For any real valued function f on any set S we define

$$U(f, S, [a, b]) = U(f(S), [a, b]).$$

If the domain of S is known, we usually suppress it.

304 **Proposition .** *Let $a_1, a_2, \ldots$ be any sequence of real numbers and $S = \{a_1, a_2, \ldots\}$. If $U(S, [a, b]) < \infty$ for all $a < b$, then these sequence $\{a_n\}$ is a convergent sequence.*

Proof. It is clear that if $T \subset S$ then $U(T, [a, b]) \leq U(S, [a, b])$. If the sequence were not convergent, then we can find a and b such that $\liminf a_n < a < b < \limsup a_n$. Choose $n_1 < n_2 < n_3 \ldots$; $m_1 < m_2 < \ldots$ such that $a_{n_i} < a$ and $a_{m_i} > b$ for all i. If $T = \{a_{n_1}, a_{m_1}, a_{n_2}, a_{m_2}, \ldots\}$, then $U(S; [a, b]) \geq U(T; [a, b]) = \infty$; a contradiction. ◻

Remark. The converse of the proposition is also true.

Theorem . *(Doob's inequalities for up crossings and down crossings). Let $\{X_1, \ldots, X_k\}$ be a submartingale relative to $\{\mathscr{F}_1, \ldots, \mathscr{F}_k\}$ $a < b$. Define $U(w, [a, b]) = U(X_1(w), \ldots, X_k(w); [a, b])$ and similarly define $D(w, [a, b])$. Then*

(i) *U, D are measurable functions;*

(ii) $E(U(\cdot,[a,b])) \leq \dfrac{E((X_k - a) + 1) - E((X_1 - a)^+)}{b-a}$;

(iii) $E(D(\cdot,[a\cdot b])) \leq E((X_k - b)^+)/(b-a)$.

Proof. (i) is left as an exercise.

(ii) Define $Y_n = (X_n - a)^+$; there are submartingales. Then clearly $Y_n \leq 0$ if and only if $X_n \leq a$ and $Y_n \geq b-a$ iff $X_n \geq b$, so that

$$UY_1(w),\ldots,Y_k(w);[0,b-a]) = U(X_1(w),\ldots,X_k(w);[a,b])$$

Define

$$\tau_1 = 1$$

$$\tau_2 = \begin{cases} \inf\{n : Y_n = 0\} \\ k, \text{ if each } Y_n = 0 \end{cases}$$

$$\tau_3 = \begin{cases} \inf\{n > \tau_2 : Y_n > b-a, \\ k, \text{ if } Y_n < b-a \text{ for each } n > \tau_2; \end{cases}$$

$$\tau_{k+1} = k.$$

As $\{Y_1,\ldots,Y_k\}$ is a submartingale, by optional sampling theorem $Y_{\tau_1},\ldots,Y_{\tau_{k+1}}$ is also a submartingale. Thus

(1) $$E(Y_{\tau_2} - Y_{\tau_1}) + E(Y_{\tau_4} - Y_{\tau_3}) + \cdots \geq 0.$$

Clearly

$$[(Y_{\tau_3} - Y_{\tau_2}) + (Y_{\tau_5} - Y_{\tau_4}) + \cdots](w) \geq (b-a) \cup (Y_1(w),\ldots Y_k(w);$$
$$[0,b-a]) = (b-a) \cup (w,[a,b]).$$

Therefore

(2) $$E(Y_{\tau_3} - Y_{\tau_2}) + E(Y_{\tau_5} - Y_{\tau_4}) + \cdots \geq (b-a)E(U(\cdot,[a,b])).$$

By (1) and (2),

$$E(Y_k - Y_1) \geq (b-a)E(U(\cdot, [a,b]))$$

giving the result.

(iii) Let $Y_n = (X_n - a)^+$ so that

$$D(Y_1(w), \ldots Y_k(w); [0, b-a]) = D(X_1(w), \ldots, X_k(w); [a,b])$$

306 Define

$$\begin{aligned}
\tau_1 &= 1;\\
\tau_2 &= \begin{cases} \inf\{n : Y_n \geq b-a\},\\ k, \ \text{if each } Y_n < b-a; \end{cases}\\
\tau_3 &= \begin{cases} \inf\{n > \tau_2 : Y_n = 0\},\\ k, \ \text{if each } Y_n > 0 \text{ for each } n > \tau_2; \end{cases}\\
\tau_{k+1} &= k.
\end{aligned}$$

By optional sampling theorem we get

$$0 \geq E(Y_{\tau_2} - Y_{\tau_3}) + E(Y_{\tau_4} - Y_{\tau_5}) + \cdots .$$

Therefore

$$0 \geq (b-a)E(D(Y_1, \ldots, Y_k; [0, b-a])) + E((b-a) - Y_k).$$

Hence

$$\begin{aligned}
E(D(\cdot, [a,b])) &\leq E((X_k - a)^+ - (b-a))/(b-a)\\
&\leq \frac{E((X_k - b)^+)}{(b-a)}, \ \text{for } (c-a)^+ - (b-a) \leq (c-b)^+\\
&\quad \text{for all } c.
\end{aligned}$$

□

Corollary . *Let $\{X_1, \ldots, X_k\}$ be a supermartingale. U, D as in theorem. Then*

(i) $E(D(\cdot, [a, b])) \leq \dfrac{E(X_1 \wedge b) - E(X_k \wedge b)}{b - a}$.

(ii) $E(U(\cdot, [a, b])) \leq \dfrac{E((X_k - b)^-)}{b - a}$.

Proof. (i) $E(D(\cdot, [a, b])) = E(U(-X_1(w), \ldots, -X_k(w), [-b, -a])$

$$\leq \frac{E((-X_k + b)^+ - (-X_1 + b)^+)}{b - a}, \quad \text{by above theorem,}$$

$$\leq \frac{E((b \wedge X_k) - (b \wedge X_1))}{b - a}, \quad \text{since for}$$

since for all a, b, c, $(b - c)^+ - (b - a)^+ \leq (b \wedge a) - (b \wedge c)$.

(ii) $E(U(\cdot, [a, b])) =$

$= E(D(-X_1(w), \ldots, -X_k(w); [-b, -a]))$

$\leq \dfrac{E((-X_k + a)^+)}{b - a}$, by theorem,

$\leq \dfrac{E((X_k - b)^-)}{b - a}$, 307

(since $(-X_k + a)^+ \leq (X_k - b)^-$,

□

Theorem . *Let $\{X_n : n = 1, 2, \ldots\}$ be a supermartingale relative to $\{\mathscr{F}_n : n = 1, 2, \ldots\}$. Let $(\Omega, \mathscr{F}, P)$ be complete.*

(i) *If $\sup\limits_n E(X_n^-) < \infty$, then X_n converges a.e. to a random variable denoted by X_∞.*

(ii) *if $\{X_n : n \geq 1\}$ is uniformly integrable, then also X_∞ exists. Further, $\{X_n : n = 1, 2, \ldots, n = \infty\}$ is a supermartingale with the natural order.*

(iii) *if $\{X_n : n \geq 1\}$ is a martingale, then $\{X_n : n \geq 1, n = \infty\}$ is a martingale.*

Proof. (i) Let $U(w[a,b]) = U(X_1(w), X_2(w), \ldots, [a,b])$. By the corollary to Doob's inequalities theorem,

$$E(U(\cdot,[a,b]) \leq \sup_n E((X_n - b)^-) < \infty$$

308 for all $a < b$. Allowing a, b to vary over the rationals alone we find that the sequence X_n is convergent a.e.

(ii) $\operatorname{Sup}_n E(X_n^-) \leq \sup_n E(|X_n|) < \infty$ so that X_∞ exists. As $X_n \to X_\infty$ in L^1 we get that $\{X_n : n \geq 1, n = \infty\}$ is a supermartingale.

(iii) follows from (ii).

□

Proposition . *Let $\{X_t : t \geq 0\}$ be a supermartingale relative to $\{\mathscr{F}_t : t \geq 0\}$. $I = [r,s]$, $a < b$ and S any countable dense subset. Let $U(w, S \cap I, [a,b]) = U(\cdot, \{X_t(w) : t \in S \cap I\}, [a,b])$. Then*

$$E(U(\cdot, S \cap I, [a,b)) \leq \frac{E((X_s - b)^-)}{b-a}.$$

Proof. Let $S \cap I$ be an increasing union of finite sets F_n: then

$$E(U(\cdot, F_n, [a,b])) \leq \frac{E((X_{\max F_n} - b)^-)}{b-a} \leq \frac{E((X_s - b)^-)}{b-a}.$$

The result follows by Fatou's lemma. □

Exercise . If further X_t is continuous i.e. $t \to X_t(w)$ is continuous for each w, then prove that

$$E(U(\cdot, I, [a,b])) \leq \frac{E((X_s - b)^-)}{b-a}$$

Theorem . *Let $(\Omega, \mathscr{F}, P)$ be complete and $\{X_t : t \geq 0\}$ a continuous supermartingale.*

309 (i) *If $\sup_{t \geq 0} E(X_t^-) < \infty$, then X_t converges a.e. to a random variable X_∞.*

(ii) *If* $\{X_t : t \geq 0\}$ *is uniformly integrable then also* X_∞ *exists and* $\{X_t : t \geq 0, t = \infty\}$ *is a supermartingale.*

Proof. (i) $E(U(\cdot, [0,n], [a,b])) \leq E((X_n - b)^-)/(b-a)$ so that

$$\operatorname*{limit}_{n\to\infty} E(U(\cdot, [0,n], [a,b])) \leq \sup_{0\leq s} \frac{E((X_s - b)^-)}{b-a}$$

for all $a < b$. Thus $\{X_t(w) : t > 0\}$ converges a.e. whose limit in denoted by X_∞ which is measurable.

(ii) As $E(X_t^-) \leq E(|X_t|)$ by (i) X_∞ exists, the other assertion is a consequence of uniform integrability.

□

Corollary . *Let* $\{X_t : t \geq 0\}$ *be a continuous uniformly integrable martingale. Then* $\{X_t : 0 \leq t \leq \infty\}$ *is also a martingale.*

Exercise. Let $\{X_t : t \geq 0\}$ be a continuous martingale such that for some Y with $0 \leq Y \leq 1$ $E(Y|\mathscr{F}_t) = X_t$ show that $X_t \to Y$ a.e.

Lemma . *Let* $(\Omega, \mathscr{F}, P)$ *be a probability space,* $\mathscr{F}_1 \supset \mathscr{F}_2 \supset \mathscr{F}_3 \ldots$ *be sub-*σ*-algebras. Let* $X_1, X_2, \ldots$ *be a real valued functions measurable with respect to* $\mathscr{F}_1, \ldots, \mathscr{F}_n, \ldots$ *respectively. Let*

(i) $E(X_{n-1}|\mathscr{F}_n) \leq X_n$

(ii) $\sup\limits_n E(X_n) < \infty$.

Then $\{X_n : n \geq 1\}$ *is uniformly integrable.* **310**

Proof. By (i) $E(X_n)$ is increasing. By (ii) given $\epsilon > 0$, we can find n_0 such that if $n \geq n_0$ then $E(X_n) \leq E(X_{n_0}) + \epsilon$. For and $\delta > 0$,

$$n \geq n_0 \int\limits_{(|X_n|\geq\delta)} |X_n| dP$$

$$= E(X_n) + \int\limits_{(X_n\leq -\delta)} -X_n dP - \int\limits_{(X_n<\delta)} X_n dP$$

$$\leq \int\limits_{(X_n \leq -\delta)} -X_{n_0} dP - \int\limits_{(X_n < \delta)} X_{n_0} dP + E(X_n) \quad \text{by (i)}$$

$$\leq \epsilon + \int\limits_{(X_n \geq \delta)} X_{n_0} dP - \int\limits_{(X_n \leq -\delta)} X_{n_0} dP \quad (\text{because } E(X_n) \leq E(X_{n_0}) + \epsilon)$$

$$\leq \epsilon + \int\limits_{(|X_n| \geq \delta)} |X_{n_0}| dP$$

Thus to show uniform integrability we have only to show $P(|X_n| \geq \delta) \to 0$ uniformly in n as $\delta \to \infty$. Now

$$\begin{aligned} E(|X_n|) &= E(X_n + 2X_n^-) \\ &\leq E(X_n) + 2E(|X_1|) \quad \text{by (i)} \\ &\leq M < \infty \text{ for all } n \text{ by (ii)} \end{aligned}$$

The result follows as $P(|X_n| \geq \delta) \leq M/\delta$. □

Optional Sampling Theorem. (Continuous case).

Let $\{X_t : t \geq 0\}$ be a right continuous supermartingale relative to $\{\mathscr{F}_t : t \geq 0\}$. Assume there exists an $X_\infty \in L'(\Omega, \mathscr{F}, P)$ such that $X_t \geq$
311 *$E(X_\infty | \mathscr{F}_t)$ for $t \geq 0$. For any stopping time τ taking values in $[0, \infty]$, let $X_\tau = X_\infty$ on $\{\tau = \infty\}$. Then*

(i) *X_τ is integrable.*

(ii) *If $\sigma \leq \tau$ are stopping times, then*

$$E(X_\tau | \mathscr{F}_\sigma) \leq X_\sigma.$$

Proof. Define

$$\sigma_n = \frac{[2^n \sigma] + 1}{2^n}, \quad \tau_n = \frac{[2^n \sigma] + 1}{2^n}.$$

Then σ_n, τ_n are stopping times, $\sigma_n \leq \tau_n$, $\sigma \leq \sigma_n$, $\tau \leq \tau_n$. σ_n, τ_n take values in $D_n = \{\infty, 0, 1/2^n, 2/2^n, \ldots, 1/2^n, \ldots\}$ so that we have $E(X_{\tau_n} | \mathscr{F}_{\sigma_n}) \leq X_{\sigma_n}$. Thus if, $A \in \mathscr{F}_\sigma \subset \mathscr{F}_{\sigma_n}$, then

$$(*) \qquad \int\limits_A X_{\tau_n} dP \leq \int\limits_A X_{\sigma_n} dP$$

As $\sigma_1 \geq \sigma_2 \geq \ldots$, by optional sampling theorem for the countable case, we have

$$E(X_{\sigma_{n-1}} | \mathscr{F}_{\sigma_n}) \leq X_{\sigma_n}.$$

Further

$$\mathscr{F}_{\sigma_1} \supset \mathscr{F}_{\sigma_2} \supset \ldots; \quad E(X_{\sigma_n} | \mathscr{F}_0) \leq X_0.$$

By the lemma $\{X_{\sigma_n}\}$, $\{X_{\tau_n}\}$ are uniformly integrable families. By right continuity $X_{\sigma_n} \to X_\sigma$ pointwise and $X_{\tau_n} \to X_\tau$ pointwise. Letting $n \to \infty$ in (*) we get the required result. □

Lemma (Integration by Parts). *Let $M(t, \cdot)$ be a continuous progressively measurable martingale and $A(t, w) : [0, \infty) \times \Omega \to \mathbb{R}$ be of bounded variation for each w. Further, assume that $A(t, w)$ is $\mathscr{F}_t$-measurable for each t. Then* **312**

$$Y(t, \cdot) = M(t, \cdot)A(t, \cdot) = \int_0^t M(s, \cdot)dA(s, \cdot)$$

is a martingale if

$$E(\sup_{0 \leq s \leq t} |M(s, \cdot)| \, ||A(\cdot)||_t) < \infty$$

for each t, where $||A(w)||_t$ is the total variation of $A(s, w)$ in $[0, t]$.

Proof. By hypothesis,

$$\sum_{i=0}^{n} M(s, \cdot)(A(s_{i+1}, \cdot) - A(s_i, \cdot))$$

converges to

$$\int_s^t M(u, \cdot)dA(u, \cdot) \text{ in } L^1 \text{ as } n \to \infty$$

and as the norm of the partition $s = s_0 < s_1 < \ldots < s_{n+1} = t$ converges to zero. Hence it is enough to show that

$$E([M(t, \cdot)A(t, \cdot) - \sum_{i=0}^{n} M(s_{i+1}, \cdot)(A(s_{i+1}, \cdot) - A(s_i, \cdot))] | \mathscr{F}_s)$$

$$= M(s, \cdot)A(s, \cdot).$$

But the left side $= E(M(s_{n+1}, \cdot)A(s_{n+1}, \cdot)-$

$$- \sum_{i=0}^{n} M(s_{i+1}, \cdot)(A(s_{i+1}, \cdot) - A(s_i, \cdot))|\mathscr{F}_s)$$
$$= M(s, \cdot)A(s, \cdot).$$

313 Taking limits as $n \to \infty$ and observing that

$$\sup_{0 \leq i \leq n} |(s_{i+1} - s_i)| \to 0,$$

we get

$$E(M(t, \cdot)A(t, \cdot) - \int_0^t M(u, \cdot)dA(u, \cdot)|\mathscr{F}_s)$$
$$= M(s, \cdot)A(s, \cdot) - \int_0^s M(u, \cdot)dA(u, \cdot).$$

$\square$

Bibliography

[1] BILLINGSLEY, P. *Convergence of probability measures*, John Wiley, New York, (1968). **314**

[2] DOOB, J.L. *Stochastic Processes,* John Wiley, New York (1952).

[3] DYNKIN, E.B. Markov Processes and Problems in Analysis, *Proc. Int. Cong of Math. Stockholm*, (1962) 36 - 58.

[4] DYNKIN, E.B. *Markov Processes,* Vols. 1, 2, Springer-Verlag, Berlin (1965).

[5] FRIEDMAN, AVNER. *Stochastic Differential Equations and Applications*, vols. 1, 2, Academic Press, New York (1976).

[6] FRIEDMAN, AVNER. *Partial Differential Equations of Parabolic type*, Prentice Hall, Englewood Cliffs (1964).

[7] GIHMAN, I.I. and A.V. SKOROHOD. *Stochastic Differential Equations,* Springer-Verlag, New York (1973).

[8] ITÔ, K. On Stochastic Differential Equations, *Memoirs Amer. Math. Soc. 4* (1951).

[9] ITÔ, K. and H.P. McKEAN, Jr. *Diffusion Processes and their sample paths,* Springer-Verlag, Berlin (1965).

[10] KUNITA, H. and S. WATANABE. On Square Integrable Martingales, *Nagoya Math. Journal.* 30 (1967) 209 - 245.

[11] LÉVY, P. *Processus Stochastic et Mouvement Brownien,* Gaulhier-Villars, Paris (1948).

315 [12] MEYER, P.A. *Probability and Potentials,* Blaisdell, Wallham (1966).

[13] NEVEU, J. *Discrete Parameter Martingales,* North Holland, Amsterdam (1975).

[14] PARTHASARATHY, K.R. *Probability Measures on Metric Spaces,* Academic Press, New York (1967).

[15] PROTTER, M.H. and H.F. WEINBERGET. *Maximum Principles in Differential Equations,* Printice Hall, Englewood Cliffs (1967).

[16] STROOCK, D.W. and S.R.S. VARADHAN. *Diffusion Process in Several Dimension,* Springer-Verlag, New York (to appear)

[17] WILLIAMS, D. Brownian motions and diffusions as Markov Processes. *Bull. Lond. Math. Soc.* 6 (1974) 257 - 303.