# Chapter 1

# Brownian Motion

## 1.1 Stochastic Process

A stochastic process can be thought of in one of many equivalent ways. We can begin with an underlying probability space $(\Omega, \Sigma, P)$ and a real valued stochastic process can be defined as a collection of random variables $\{x(t, \omega)\}$ indexed by the parameter set $\mathbf{T}$. This means that for each $t \in \mathbf{T}$, $x(t, \omega)$ is a measurable map of $(\Omega, \Sigma) \to (\mathbf{R}, \mathcal{B}_0)$ where $(\mathbf{R}, \mathcal{B}_0)$ is the real line with the usual Borel $\sigma$-field. The parameter set often represents time and could be either the integers representing discrete time or could be $[0, T]$, $[0, \infty)$ or $(-\infty, \infty)$ if we are studying processes in continuous time. For each fixed $\omega$ we can view $x(t, \omega)$ as a map of $\mathbf{T} \to \mathbf{R}$ and we would then get a *random function* of $t \in \mathbf{T}$. If we denote by $\mathbf{X}$ the space of functions on $\mathbf{T}$, then a stochastic process becomes a measurable map from a probability space into $\mathbf{X}$. There is a natural $\sigma$-field $\mathcal{B}$ on $\mathbf{X}$ and measurability is to be understood in terms of this $\sigma$-field. This natural $\sigma$-field, called the Kolmogorov $\sigma$-field, is defined as the smallest $\sigma$-field such that the projections $\{\pi_t(f) = f(t)\,;\, t \in \mathbf{T}\}$ mapping $\mathbf{X} \to \mathbf{R}$ are measurable. The point of this definition is that a random function $x(\cdot, \omega) : \Omega \to \mathbf{X}$ is measurable if and only if the random variables $x(t, \omega) : \Omega \to \mathbf{R}$ are measurable for each $t \in \mathbf{T}$.

The mapping $x(\cdot, \cdot)$ induces a measure on $(\mathbf{X}, \mathcal{B})$ by the usual definition

$$Q(A) = P\big[\omega : x(\cdot, \omega) \in A\big] \tag{1.1}$$

for $A \in \mathcal{B}$. Since the underlying probability model $(\Omega, \Sigma, P)$ is irrelevant, it can be replaced by the *canonical* model $(\mathbf{X}, \mathcal{B}, Q)$ with the special choice of $x(t, f) = \pi_t(f) = f(t)$. A stochastic process then can then be defined simply as a probability measure $Q$ on $(\mathbf{X}, \mathcal{B})$.

Another point of view is that the only relevant objects are the joint distributions of $\{x(t_1, \omega), x(t_2, \omega), \cdots, x(t_k, \omega)\}$ for every $k$ and every finite subset $F = (t_1, t_2, \cdots, t_k)$ of $\mathbf{T}$. These can be specified as probability measures $\mu_F$ on $\mathbf{R}^k$. These $\{\mu_F\}$ cannot be totally arbitrary. If we allow different permutations of the same set, so that $F$ and $F'$ are permutations of each other then $\mu_F$ and $\mu_{F'}$ should be related by the same permutation. If $F \subset F'$, then we can obtain the joint distribution of $\{x(t\,,\omega)\,;\,t \in F\}$ by projecting the joint distribution of $\{x(t\,,\omega)\,;\,t \in F'\}$ from $\mathbf{R}^{k'} \to \mathbf{R}^k$ where $k'$ and $k$ are the cardinalities of $F'$ and $F$ respectively. A stochastic process can then be viewed as a family $\{\mu_F\}$ of distributions on various finite dimensional spaces that satisfy the consistency conditions. A theorem of Kolmogorov says that this is not all that different. Any such consistent family arises from a $Q$ on $(\mathbf{X}\,,\mathcal{B})$ which is uniquely determined by the family $\{\mu_F\}$.

If $\mathbf{T}$ is countable this is quite satisfactory. $\mathbf{X}$ is the the space of sequences and the $\sigma$-field $\mathcal{B}$ is quite adequate to answer all the questions we may want to ask. The set of bounded sequences, the set of convergent sequences, the set of summable sequences are all measurable subsets of $\mathbf{X}$ and therefore we can answer questions like, does the sequence converge with probability 1, etc. However if $\mathbf{T}$ is uncountable like $[0, T]$, then the space of bounded functions, the space of continuous functions etc, are not measurable sets. They do not belong to $\mathcal{B}$. Basically, in probability theory, the rules involve only a countable collection of sets at one time and any information that involves the values of an uncountable number of measurable functions is out of reach. There is an intrinsic reason for this. In probability theory we can always change the values of a random variable on a set of measure 0 and we have not changed anything of consequence. Since we are allowed to mess up each function on a set of measure 0 we have to assume that each function has indeed been messed up on a set of measure 0. If we are dealing with a countable number of functions the 'mess up' has occured only on the countable union of these invidual sets of measure 0, which by the properties of a measure is again a set of measure 0. On the other hand if we are dealing with an uncountable set of functions, then these sets of measure 0 can possibly gang up on us to produce a set of positive or even full measure. We just can not be sure.

Of course it would be foolish of us to mess things up unnecessarily. If we can clean things up and choose a nice version of our random variables we should do so. But we cannot really do this sensibly unless we decide first what nice means. We however face the risk of being too greedy and it may not be possible to have a version as nice as we seek. But then we can always change our mind.

## 1.2 Regularity

Very often it is natural to try to find a version that has continuous trajectories. This is equivalent to restricting $\mathbf{X}$ to the space of continuous functions on $[0, T]$ and we are trying to construct a measure $Q$ on $\mathbf{X} = C[0\,,T]$ with the natural $\sigma$-field $\mathcal{B}$. This is not always possible. We want to find some sufficient conditions on the finite dimensional distributions $\{\mu_F\}$ that guarantee that a choice of $Q$ exists on $(\mathbf{X}\,,\mathcal{B})$.

**Theorem 1.1. (Kolmogorov's Regularity Theorem)** *Assume that for any pair $(s,t) \in [0\,,T]$ the bivariate distribution $\mu_{s,t}$ satisfies*

$$\int\int |x-y|^\beta \mu_{s,t}(dx\,,dy) \le C|t-s|^{1+\alpha} \tag{1.2}$$

*for some positive constants $\beta, \alpha$ and $C$. Then there is a unique $Q$ on $(\mathbf{X}\,,\mathcal{B})$ such that it has $\{\mu_F\}$ for its finite dimensional distributions.*

*Proof.* Since we can only deal effectively with a countable number of random variables, we restrict ourselves to values at diadic times. Let us, for simplicity, take $T=1$. Denote by $\mathbf{T}_n$ time points $t$ of the form $t=\frac{j}{2^n}$ for $0 \le j \le 2^n$. The countable union $\cup_{j=0}^\infty \mathbf{T}_j = \mathbf{T}^0$ is a countable dense subset of $\mathbf{T}$. We will construct a probability measure $Q$ on the space of sequences corresponding to the values of $\{x(t) : t \in \mathbf{T}^0\}$, show that $Q$ is supported on sequences that produce uniformly continuous functions on $\mathbf{T}^0$ and then extend them automatically to $\mathbf{T}$ by continuity and the extension will provide us the natural $Q$ on $C[0\,,1]$. If we start from the set of values on $\mathbf{T}_n$, the $n$-th level of diadics, by linear iterpolation we can construct a version $x_n(t)$ that agrees with the original variables at these diadic points. This way we have a sequence $x_n(t)$ such that $x_n(\cdot) = x_{n+1}(\cdot)$ on $\mathbf{T}_n$. If we can show

$$Q\big[x(\cdot) : \sup_{0\le t\le 1} |x_n(t) - x_{n+1}(t)| \ge 2^{-n\gamma}\big] \le C2^{-n\delta} \tag{1.3}$$

then we can conclude that

$$Q\big[x(\cdot) : \lim_{n\to\infty} x_n(t) = x_\infty(t) \ \text{ exists } \ \text{uniformly } \ \text{on } [0\,,1]\big] = 1 \tag{1.4}$$

The limit $x_\infty(\cdot)$ will be continuous on $\mathbf{T}$ and will coincide with $x(\cdot)$ on $\mathbf{T}^0$ there by establishing our result. Proof of (1.3) depends on a simple observation. The difference $|x_n(\cdot) - x_{n+1}(\cdot)|$ achieves its maximum at the mid point of one of the diadic intervals determined by $\mathbf{T}_n$ and hence

$$\begin{aligned}
&\sup_{0\le t\le 1} |x_n(t) - x_{n+1}(t)| \\
&\qquad \le \sup_{1\le j\le 2^n} |x_n(\frac{2j-1}{2^{n+1}}) - x_{n+1}(\frac{2j-1}{2^{n+1}})| \\
&\qquad \le \sup_{1\le j\le 2^n} \max\big\{|x(\frac{2j-1}{2^{n+1}}) - x(\frac{2j}{2^{n+1}})|, |x(\frac{2j-1}{2^{n+1}}) - x(\frac{2j-2}{2^{n+1}})|\big\}
\end{aligned}$$

and we can estimate the left hand side of (1.3) by

$$
\begin{aligned}
&Q\big[x(\cdot) : \sup_{0\le t\le 1} |x_n(t) - x_{n+1}(t)| \ge 2^{-n\gamma}\big] \\
&\le Q\big[\sup_{1\le i\le 2^{n+1}} |x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})| \ge 2^{-n\gamma}\big] \\
&\le 2^{n+1} \sup_{1\le i\le 2^{n+1}} Q\big[|x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})| \ge 2^{-n\gamma}\big] \\
&\le 2^{n+1} 2^{n\beta\gamma} \sup_{1\le i\le 2^{n+1}} E^Q\big[|x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})|^\beta\big] \\
&\le C 2^{n+1}\, 2^{n\beta\gamma}\, 2^{-(1+\alpha)(n+1)} \\
&\le C 2^{-n\delta}
\end{aligned}
$$

provided $\delta \le \alpha - \beta\gamma$. For given $\alpha, \beta$ we can pick $\gamma < \alpha\beta$ and we are done. □

An equivalent version of this theorem is the following.

**Theorem 1.2.** *If $x(t\,,\omega)$ is a stochastic process on $(\Omega\,,\Sigma\,,P)$ satisfying*

$$E^P\left[|x(t) - x(s)|^\beta\right] \le C|t-s|^{1+\alpha}$$

*for some positive constants $\alpha$, $\beta$ and $C$, then if necessary , $x(t,\omega)$ can be modified for each $t$ on a set of measure zero, to obtain an equivalent version that is almost surely continuous.*

As an important application we consider Brownian Motion, which is defined as a stochastic process that has multivariate normal distributions for its finite dimensional distributions. These normal distributions have mean zero and the variance covariance matrix is specified by $Cov(x(s), x(t)) = \min(s,t)$. An elementary calculation yields

$$E|x(s) - x(t)|^4 = 3|t-s|^2$$

so that Theorem 1.1 is applicable with $\beta = 4$, $\alpha = 1$ and $C = 3$.

To see that some restriction is needed, let us consider the Poisson process defined as a process with independent increments with the distribution of $x(t) - x(s)$ being Poisson with parameter $t - s$ provided $t > s$. In this case since

$$P[x(t) - x(s) \ge 1] = 1 - \exp[-(t-s)]$$

we have, for every $n \ge 0$,

$$E|x(t) - x(s)|^n \ge 1 - \exp[-|t-s|] \simeq C|t-s|$$

and the conditions for Theorem 1.1 are never satisfied. It should not be, because after all a Poisson process is a counting process and jumps whenever the event that it is counting occurs and it would indeed be greedy of us to try to put the measure on the space of continuous functions.

*Remark* 1.1. The fact that there cannot be a measure on the space of continuous functions whose finite dimensional distributions coincide with those of the Poisson process requires a proof. There is a whole class of nasty examples of measures $\{Q\}$ on the space of continuous functions such that for every $t \in [0\,,1]$

$$Q\big[\omega : x(t\,,\omega) \ \text{ is a rational } \ \text{number } \big] = 1$$

The difference is that the rationals are dense, whereas the integers are not. The proof has to depend on the fact that a continuous function that is not identically equal to some fixed integer must spend a positive amount of time at nonintegral points. Try to make a rigorous proof using Fubini's theorem.

## 1.3 Garsia, Rodemich and Rumsey inequality.

If we have a stochastic process $x(t\,,\omega)$ and we wish to show that it has a nice version, perhaps a continuous one, or even a Holder continuous or differentiable version, there are things we have to estimate. Establishing Holder continuity amounts to estimating

$$\epsilon(\ell) = P\big[\sup_{s,t} \frac{|x(s) - x(t)|}{|t-s|^\alpha} \le \ell\big]$$

and showing that $\epsilon(\ell) \to 1$ as $\ell \to \infty$. These are often difficult to estimate and require special methods. A slight modification of the proof of Theorem 1.1 will establish that the nice, continuous version of Brownian motion actually satisfies a Holder condition of exponent $\alpha$ so long as $0 < \alpha < \frac{1}{2}$.

On the other hand if we want to show only that we have a version $x(t\,,\omega)$ that is square integrable, we have to estimate

$$\epsilon(\ell) = P\big[\int_0^1 |x(t\,,\omega)|^2 dt \le \ell\big]$$

and try to show that $\epsilon(\ell) \to 1$ as $\ell \to \infty$. This task is somewhat easier because we could control it by estimating

$$E^P\big[\int_0^1 |x(t\,,\omega)|^2\, dt\big]$$

and that could be done by the use of Fubini's theorem. After all

$$E^P\big[\int_0^1 |x(t\,,\omega)|^2\, dt\big] = \int_0^1 E^P\big[|x(t\,,\omega)|^2\,\big]\, dt$$

Estimating integrals are easier that estimating suprema. Sobolev inequality controls suprema in terms of integrals. Garsia, Rodemich and Rumsey inequality is a generalization and can be used in a wide variety of contexts.

**Theorem 1.3.** *Let $\Psi(\cdot)$ and $p(\cdot)$ be continuous strictly increasing functions on $[0\,,\infty)$ with $p(0) = \Psi(0) = 0$ and $\Psi(x) \to \infty$ as $x \to \infty$. Assume that a continuous function $f(\cdot)$ on $[0\,,1]$ satisfies*

$$\int_0^1 \int_0^1 \Psi\left(\frac{|f(t)-f(s)|}{p(|t-s|)}\right) ds\, dt = B < \infty. \tag{1.5}$$

*Then*

$$|f(0)-f(1)| \le 8 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u) \tag{1.6}$$

The double integral (1.5) has a singularity on the diagonal and its finiteness depends on $f, p$ and $\Psi$. The integral in (1.6) has a singularity at $u = 0$ and its convergence requires a balancing act between $\Psi(\cdot)$ and $p(\cdot)$. The two conditions compete and the existence of a pair $\Psi(\cdot)\,, p(\cdot)$ satisfying all the conditions will turn out to imply some regularity on $f(\cdot)$.

Let us first assume Theorem 1.3 and illustrate its uses with some examples. We will come back to its proof at the end of the section. First we remark that the following corollary is an immediate consequence of Theorem 1.3.

**Corollary 1.4.** *If we replace the interval $[0\,,1]$ by the interval $[T_1\,,T_2]$ so that*

$$B_{T_1,T_2} = \int_{T_1}^{T_2} \int_{T_1}^{T_2} \Psi\left(\frac{|f(t)-f(s)|}{p(|t-s|)}\right) ds\, dt$$

*then*

$$|f(T_2)-f(T_1)| \le 8 \int_0^{T_2-T_1} \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u)$$

*For $0 \le T_1 < T_2 \le 1$ because $B_{T_1,T_2} \le B_{0,1} = B$, we can conclude from (1.5), that the modulus of continuity $\varpi_f(\delta)$ satisfies*

$$\varpi_f(\delta) = \sup_{\substack{0 \le s,t \le 1 \\ |t-s| \le \delta}} |f(t)-f(s)| \le 8 \int_0^{\delta} \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u) \tag{1.7}$$

*Proof.* (of Corollary). If we map the interval $[T_1\,,T_2]$ into $[0\,,1]$ by $t' = \frac{t-T_1}{T_2-T_1}$ and redefine $f'(t) = f(T_1 + (T_2-T_1)t)$ and $p'(u) = p((T_2-T_1)u)$, then

$$\begin{aligned}
&\int_0^1 \int_0^1 \Psi\Big[\frac{|f'(t)-f'(s)|}{p'(|t-s|)}\Big] ds\, dt \\
&\qquad = \frac{1}{(T_2-T_1)^2} \int_{T_1}^{T_2} \int_{T_1}^{T_2} \Psi\Big[\frac{|f(t)-f(s)|}{p(|t-s|)}\Big] ds\, dt \\
&\qquad = \frac{B_{T_1,T_2}}{(T_2-T_1)^2}
\end{aligned}$$

and

$$
\begin{aligned}
|f(T_2) - f(T_1)| &= |f'(1) - f'(0)| \\
&\leq 8 \int_0^1 \Psi^{-1}\left(\frac{4B_{T_1,T_2}}{(T_2-T_1)^2\, u^2}\right) dp'(u) \\
&= 8 \int_0^{(T_2-T_1)} \Psi^{-1}\left(\frac{4B_{T_1,T_2}}{u^2}\right) dp(u)
\end{aligned}
$$

In particular (1.7) is now an immediate consequence. □

Let us now turn to Brownian motion or more generally processes that satisfy

$$
E^P\left[|x(t) - x(s)|^\beta\right] \leq C|t-s|^{1+\alpha}
$$

on $[0\,,1]$. We know from Theorem 1.1 that the paths can be chosen to be continuous. We will now show that the continuous version enjoys some additional regularity. We apply Theorem 1.3 with $\Psi(x) = x^\beta$, and $p(u) = u^{\frac{\gamma}{\beta}}$. Then

$$
\begin{aligned}
&E^P\left[\int_0^1\int_0^1 \Psi\left(\frac{|x(t)-x(s)|}{p(|t-s|)}\right) ds\, dt\right] \\
&= \int_0^1\int_0^1 E^P\left[\frac{|x(t)-x(s)|^\beta}{|t-s|^\gamma}\right] dsdt \\
&\leq C\int_0^1\int_0^1 |t-s|^{1+\alpha-\gamma}\, dsdt \\
&= C\, C_\delta
\end{aligned}
$$

where $C_\delta$ is a constant depending only on $\delta = 2 + \alpha - \gamma$ and is finite if $\delta > 0$. By Fubini's theorem, almost surely

$$
\int_0^1\int_0^1 \Psi\left(\frac{|x(t)-x(s)|}{p(|t-s|)}\right) ds\, dt = B(\omega) < \infty
$$

and by Tchebychev's inequality

$$
P\big[B(\omega) \geq B\big] \leq \frac{C\, C_\delta}{B}.
$$

On the other hand

$$
\begin{aligned}
8\int_0^h (\frac{4B}{u^2})^{\frac{1}{\beta}} du^{\frac{\gamma}{\beta}} &= 8\frac{\gamma}{\beta}(4B)^{\frac{1}{\beta}} \int_0^h u^{\frac{\gamma-2}{\beta}-1} du \\
&= 8\frac{\gamma}{\gamma-2}(4B)^{\frac{1}{\beta}} h^{\frac{\gamma-2}{\beta}}
\end{aligned}
$$

We obtain Holder continuity with exponent $\frac{\gamma-2}{\beta}$ which can be anything less than $\frac{\alpha}{\beta}$. For Brownian motion $\alpha = \frac{\beta}{2} - 1$ and therefore $\frac{\alpha}{\beta}$ can be made arbitrarily close to $\frac{1}{2}$.

*Remark* 1.2. With probability 1 Brownian paths satisfy a Holder condition with any exponent less than $\frac{1}{2}$.

It is not hard to see that they do not satisfy a Holder condition with exponent $\frac{1}{2}$

*Exercise* 1.1. Show that

$$P\Big[\sup_{0\le s,\, t\le 1}\frac{|x(t)-x(s)|}{\sqrt{|t-s|}}=\infty\Big]=1.$$

Hint: The random variables $\frac{x(t)-x(s)}{\sqrt{|t-s|}}$ have standard normal distributions for any interval $[s,t]$ and they are independent for disjoint intervals. We can find as many disjoint intervals as we wish and therefore dominate the Holder constant from below by the supremum of absolute values of an arbitrary number of independent Gaussians.

*Exercise* 1.2. (Precise modulus of continuity). The choice of $\Psi(x)=\exp[\alpha\, x^2]$ with $\alpha<\frac{1}{2}$ and $p(u)=u^{\frac{1}{2}}$ produces a modulus of continuity of the form

$$\varpi_x(\delta)\le 8\int_0^\delta\sqrt{\frac{1}{\alpha}\log\big[1+\frac{4B}{u^2}\big]}\frac{1}{2\sqrt{u}}du$$

that produces eventually a statement

$$P\Big[\limsup_{\delta\to 0}\frac{\varpi_x(\delta)}{\sqrt{\delta\log\frac{1}{\delta}}}\le 16\Big]=1.$$

*Remark* 1.3. This is almost the final word, because the argument of the previous exercise can be tightened slightly to yield

$$P\Big[\limsup_{\delta\to 0}\frac{\varpi_x(\delta)}{\sqrt{\delta\log\frac{1}{\delta}}}\ge\sqrt{2}\Big]=1$$

and according to a result of Paul Lévy

$$P\Big[\limsup_{\delta\to 0}\frac{\varpi_x(\delta)}{\sqrt{\delta\log\frac{1}{\delta}}}=\sqrt{2}\Big]=1.$$

*Proof.* (of Theorem 1.3.) Define

$$I(t)=\int_0^1\Psi\left(\frac{|f(t)-f(s)|}{p(|t-s|)}\right)ds$$

and

$$B=\int_0^1 I(t)\,dt$$

There exists $t_0 \in (0\,,1)$ such that $I(t_0) \le B$. We shall prove that

$$|f(0) - f(t_0)| \le 4 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u) \tag{1.8}$$

By a similar argument

$$|f(1) - f(t_0)| \le 4 \int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u)$$

and combining the two we will have (1.6). To prove 1.8 we shall pick recursively two sequences $\{u_n\}$ and $\{t_n\}$ satisfying

$$t_0 > u_1 > t_1 > u_2 > t_2 > \cdots > u_n > t_n > \cdots$$

in the following manner. By induction, if $t_{n-1}$ has already been chosen, define

$$d_n = p(t_{n-1})$$

and pick $u_n$ so that $p(u_n) = \frac{d_n}{2}$. Then

$$\int_0^{u_n} I(t)\, dt \le B$$

and

$$\int_0^{u_n} \Psi\left(\frac{|f(t_{n-1}) - f(s)|}{p(|t_{n-1} - s|)}\right) ds \le I(t_{n-1})$$

Now $t_n$ is chosen so that

$$I(t_n) \le \frac{2B}{u_n}$$

and

$$\Psi\left(\frac{|f(t_n) - f(t_{n-1})|}{p(|t_n - t_{n-1}|)}\right) \le 2\frac{I(t_{n-1})}{u_n} \le \frac{4B}{u_{n-1}\, u_n} \le \frac{4B}{u_n^2}$$

We now have

$$|f(t_n) - f(t_{n-1})| \le \Psi^{-1}\left(\frac{4B}{u_n^2}\right) p(t_{n-1} - t_n) \le \Psi^{-1}\left(\frac{4B}{u_n^2}\right) p(t_{n-1}).$$

$$p(t_{n-1}) = 2p(u_n) = 4[p(u_n) - \frac{1}{2}p(u_n)] \le 4[p(u_n) - p(u_{n+1})]$$

Then,

$$|f(t_n) - f(t_{n-1})| \le 4\Psi^{-1}\left(\frac{4B}{u_n^2}\right)[p(u_n) - p(u_{n+1})] \le 4 \int_{u_{n+1}}^{u_n} \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u)$$

Summing over $n = 1, 2, \cdots$, we get

$$|f(t_0) - f(0)| \le 4 \int_0^{u_1} \Psi^{-1}\left(\frac{4B}{u^2}\right) p(du) \le 4 \int_0^{u_1} \Psi^{-1}\left(\frac{4B}{u^2}\right) p(du)$$

and we are done. $\square$

*Example* 1.1. Let us consider a stationary Gaussian process with

$$\rho(t) = E[X(s)X(s+t)]$$

and denote by

$$\sigma^2(t) = E[(X(t) - X(0))^2] = 2(\rho(0) - \rho(t)).$$

Let us suppose that $\sigma^2(t) \leq C|\log t|^{-a}$ for some $a > 1$ and $C < \infty$. Then we can apply Theorem 1.3 and establish the existence of an almost sure continuous version by a suitable choice of $\Psi$ and $p$.

On the other hand we will show that, if $\sigma^2(t) \geq c|\log t|^{-1}$, then the paths are almost surely unbounded on every time interval. It is generally hard to prove that some thing is unbounded. But there is a nice trick that we will use. One way to make sure that a function $f(t)$ on $t_1 \leq t \leq t_2$ is unbounded is to make sure that the measure $\mu_f(A) = \text{LebMes}\,\{t : f(t) \in A\}$ is not supported on a compact interval. That can be assured if we show that $\mu_f$ has a density with respect to the Lebsgue measure on $\mathbf{R}$ with a density $\phi_f(x)$ that is real analytic, which in turn will be assured if we show that

$$\int_{-\infty}^{\infty} |\widehat{\mu}_f(\xi)| e^{\alpha|\xi|}\, d\xi < \infty$$

for some $\alpha > 0$. By Schwarz's inequality it is sufficient to prove that

$$\int_{-\infty}^{\infty} |\widehat{\mu}_f(\xi)|^2 e^{\alpha|\xi|}\, d\xi < \infty$$

for some $\alpha > 0$. We will prove

$$\int_{-\infty}^{\infty} E\left[\left|\int_{t_1}^{t_2} e^{i\,\xi\,X(t)}\, dt\right|^2\right] e^{\alpha\xi}\, d\xi < \infty$$

for some $\alpha > 0$. Sine we can replace $\alpha$ by $-\alpha$ , this will control

$$\int_{-\infty}^{\infty} E\left[\left|\int_{t_1}^{t_2} e^{i\,\xi\,X(t)}\, dt\right|^2\right] e^{\alpha|\xi|}\, d\xi < \infty$$

and we can apply Fubini's theorem to complete the proof.

$$\begin{aligned}
&\int_{-\infty}^{\infty} E\Big[\Big|\int_{t_1}^{t_2} e^{i\,\xi\,X(t)}\,dt\Big|^2\Big] e^{\alpha\xi}\,d\xi \\
&= \int_{-\infty}^{\infty} E\Big[\int_{t_1}^{t_2}\int_{t_1}^{t_2} e^{i\,\xi\,(X(t)-X(s))}\,ds\,dt\Big] e^{\alpha\xi}\,d\xi \\
&= \int_{-\infty}^{\infty}\int_{t_1}^{t_2}\int_{t_1}^{t_2} E\Big[e^{i\,\xi\,(X(t)-X(s))}\Big]\,ds\,dt\,e^{\alpha\xi}\,d\xi \\
&= \int_{-\infty}^{\infty}\int_{t_1}^{t_2}\int_{t_1}^{t_2} e^{-\frac{\sigma^2(t-s)\xi^2}{2}}\,ds\,dt\,e^{\alpha\xi}\,d\xi \\
&= \int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{\sqrt{2\pi}}{\sigma(t-s)} e^{\frac{\alpha^2}{2\sigma^2(t-s)}} \\
&\le \int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{\sqrt{2\pi}}{\sigma(t-s)} e^{\frac{\alpha^2|\log|(t-s)||}{2c}}\,ds\,dt \\
&< \infty
\end{aligned}$$

provided $\alpha$ is small enough.

## 1.4 Brownian Motion as a Martingale

$P$ is the Wiener measure on $(\Omega,\,\mathcal{B})$ where $\Omega = C[0,T]$ and $\mathcal{B}$ is the Borel $\sigma$-field on $\Omega$. In addition we denote by $\mathcal{B}_t$ the $\sigma$-field generated by $x(s)$ for $0 \le s \le t$. It is easy to see tha $x(t)$ is a martingale with respect to $(\Omega,\,\mathcal{B}_t,\,P)$, i.e for each $t > s$ in $[0,T]$

$$E^P\{x(t)|\mathcal{B}_s\} = x(s) \quad \text{a.e.} \quad P \tag{1.9}$$

and so is $x(t)^2 - t$. In other words

$$E^P\{x(t)^2 - t\,|\mathcal{F}_s\} = x(s)^2 - s \quad \text{a.e.} \quad P \tag{1.10}$$

The proof is rather straight forward. We write $x(t) = x(s) + Z$ where $Z = x(t) - x(s)$ is a random variable independent of the past history $\mathcal{B}_s$ and is distributed as a Gaussian random variable with mean 0 and variance $t - s$. Therefore $E^P\{Z|\mathcal{B}_s\} = 0$ and $E^P\{Z^2|\mathcal{B}_s\} = t - s$ a.e $P$. Conversely,

**Theorem 1.5. Lévy's theorem.** *If $P$ is a measure on $(C[0,T],\,\mathcal{B})$ such that $P[x(0) = 0] = 1$ and the the functions $x(t)$ and $x^2(t) - t$ are martingales with respect to $(C[0,T],\,\mathcal{B}_t,\,P)$ then $P$ is the Wiener measure.*

*Proof.* The proof is based on the observation that a Gaussian distribution is determined by two moments. But that the distribution is Gaussian is a consequence of the fact that the paths are almost surely continuous and not part of our assumptions. The actual proof is carried out by establishing that for each real number $\lambda$

$$X_\lambda(t) = \exp\big[\lambda x(t) - \frac{\lambda^2}{2}t\big] \tag{1.11}$$

is a martingale with respect to $(C[0,T], \mathcal{B}_t, P)$. Once this is established it is elementary to compute

$$E^P\big[\exp\big[\lambda(x(t)-x(s))\big]|\mathcal{B}_s\big] = \exp\big[\frac{\lambda^2}{2}(t-s)\big]$$

which shows that we have a Gaussian Process with independent increments with two matching moments. The proof of (1.11) is more or less the same as proving the central limit theorem. In order to prove (2.5) we can assume with out loss of generality that $s=0$ and will show that

$$E^P\big[\exp\big[\lambda x(t) - \frac{\lambda^2}{2}t\big]\big] = 1 \tag{1.12}$$

To this end let us define successively $\tau_{0,\epsilon} = 0$,

$$\tau_{k+1,\epsilon} = \min\big[\inf\big\{s : s \ge \tau_{k,\epsilon}, |x(s) - x(\tau_{k,\epsilon})| \ge \epsilon\big\}, t\,, \tau_{k,\epsilon}+\epsilon\big]$$

Then each $\tau_{k,\epsilon}$ is a stopping time and eventually $\tau_{k,\epsilon} = t$ by continuity of paths. The continuity of paths also guarantees that $|x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})| \le \epsilon$. We write

$$x(t) = \sum_{k\ge 0}[x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})]$$

and

$$t = \sum_{k\ge 0}[\tau_{k+1,\epsilon} - \tau_{k,\epsilon}]$$

To establish (1.12) we calculate the quantity on the left hand side as

$$\lim_{n\to\infty} E^P\big[\exp\big[\sum_{0\le k\le n}\big[\lambda[x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})] - \frac{\lambda^2}{2}[\tau_{k+1,\epsilon} - \tau_{k,\epsilon}]\big]\big]\big]$$

and show that it is equal to 1. Let us cosider the $\sigma$-field $\mathcal{F}_k = \mathcal{B}_{\tau_{k,\epsilon}}$ and the quantity

$$q_k(\omega) = E^P\big[\exp\big[\lambda[x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})] - \frac{\lambda^2}{2}[\tau_{k+1,\epsilon} - \tau_{k,\epsilon}]\big]\Big|\mathcal{F}_k\big]$$

Clearly, if we use Taylor expansion and the fact that $x(t)$ as well as $x(t)^2 - t$ are martingales

$$\begin{aligned}|q_k(\omega) - 1| &\le CE^P\big[\big[|\lambda|^3|x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})|^3 + \lambda^2|\tau_{k+1,\epsilon} - \tau_{k,\epsilon}|^2\big]\Big|\mathcal{F}_k\big] \\ &\le C_\lambda\,\epsilon\,E^P\big[\big[|x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})|^2 + |\tau_{k+1,\epsilon} - \tau_{k,\epsilon}|\big]\big|\mathcal{F}_k\big] \\ &= 2C_\lambda\,\epsilon\,E^P\big[|\tau_{k+1,\epsilon} - \tau_{k,\epsilon}|\big|\mathcal{F}_k\big]\end{aligned}$$

In particular for some constant $C$ depending on $\lambda$

$$q_k(\omega) \le E^P\big[\exp\big[C\,\epsilon\,[\tau_{k+1,\epsilon} - \tau_{k,\epsilon}]\big]\big|\mathcal{F}_k\big]$$

and by induction

$$\limsup_{n\to\infty} E^P\big[\exp\big[\sum_{0\le k\le n}\big[\lambda[x(\tau_{k+1,\epsilon})-x(\tau_{k,\epsilon})]-\frac{\lambda^2}{2}[\tau_{k+1,\epsilon}-\tau_{k,\epsilon}]\big]\big]\big]$$
$$\le \exp[C\,\epsilon\, t\,]$$

Since $\epsilon > 0$ is arbitrary we prove one half of (1.12). Notice that in any case $\sup_\omega |q_k(\omega)-1| \le \epsilon$. Hence we have the lower bound

$$q_k(\omega) \ge E^P\big[\exp\big[-C\,\epsilon\,[\tau_{k+1,\epsilon}-\tau_{k\ \epsilon}]\big]\Big|\mathcal{F}_k\big]$$

which can be used to prove the other half. This completes the proof of the theorem. □

*Exercise* 1.3. Why does Theorem 1.5 fail for the process $x(t) = N(t) - t$ where $N(t)$ is the standard Poisson Process with rate 1?

*Remark* 1.4. One can use the Martingale inequality in order to estimate the probability $P\{\sup_{0\le s\le t}|x(s)| \ge \ell\}$. For $\lambda > 0$, by Doob's inequality

$$P\big[\sup_{0\le s\le t}\exp\big[\lambda x(s) - \frac{\lambda^2}{2}s\big] \ge A\big] \le \frac{1}{A}$$

and

$$P\big[\sup_{0\le s\le t} x(s) \ge \ell\big] \le P\big[\sup_{0\le s\le t}[x(s) - \frac{\lambda s}{2}] \ge \ell - \frac{\lambda t}{2}\big]$$
$$= P\big[\sup_{0\le s\le t}[\lambda x(s) - \frac{\lambda^2 s}{2}] \ge \lambda\ell - \lambda^2 t2\big]$$
$$\le \exp[-\lambda\ell + \frac{\lambda^2 t}{2}]$$

Optimizing over $\lambda > 0$, we obtain

$$P\big[\sup_{0\le s\le t} x(s) \ge \ell\big] \le \exp[-\frac{\ell^2}{2t}]$$

and by symmetry

$$P\big[\sup_{0\le s\le t} |x(s)| \ge \ell\big] \le 2\exp[-\frac{\ell^2}{2t}]$$

The estimate is not too bad because by reflection principle

$$P\big[\sup_{0\le s\le t} x(s) \ge \ell\big] = 2\,P\big[x(t) \ge \ell\big] = \sqrt{\frac{2}{\pi\, t}}\int_\ell^\infty \exp[-\frac{x^2}{2\,t}]\,dx$$

*Exercise* 1.4. One can use the estimate above to prove the result of Paul Lévy

$$P\Big[\limsup_{\delta\to 0}\frac{\sup_{\substack{0\le s,t\le 1\\ |s-t|\le\delta}}|x(s)-x(t)|}{\sqrt{\delta\log\frac{1}{\delta}}}=\sqrt{2}\Big]=1$$

We had an exercise in the previous section that established the lower bound. Let us concentrate on the upper bound. If we define

$$\Delta_\delta(\omega)=\sup_{\substack{0\le s,t\le 1\\ |s-t|\le\delta}}|x(s)-x(t)|$$

first check that it is sufficient to prove that for any $\rho<1$, and $a>\sqrt{2}$

$$\sum_n P\Big[\Delta_{\rho^n}(\omega)\ge a\sqrt{n\rho^n\log\frac{1}{\rho}}\Big]<\infty \tag{1.13}$$

To estimate $\Delta_{\rho^n}(\omega)$ it is sufficient to estimate $\sup_{t\in I_j}|x(t)-x(t_j)|$ for $k_\epsilon\rho^{-n}$ overlapping intervals $\{I_j\}$ of the form $[t_j, t_j+(1+\epsilon)\rho^n\,]$ with length $(1+\epsilon)\rho^n$. For each $\epsilon>0$, $k_\epsilon=\epsilon^{-1}$ is a constant such that any interval $[s\,,t]$ of length no larger than $\rho^n$ is completely contained in some $I_j$ with $t_j\le s\le t_j+\epsilon\rho^n$. Then

$$\Delta_{\rho^n}(\omega)\le\sup_j\Big[\sup_{t\in I_j}|x(t)-x(t_j)|+\sup_{t_j\le s\le t_j+\epsilon\rho^n}|x(s)-x(t_j)|\Big]$$

Therefore, for any $a=a_1+a_2$,

$$\begin{aligned}
&P\left[\Delta_{\rho^n}(\omega)\ge a\sqrt{n\rho^n\log\frac{1}{\rho}}\,\right]\\
&\qquad\le P\left[\sup_j\sup_{t\in I_j}|x(t)-x(t_j)|\ge a_1\sqrt{n\rho^n\log\frac{1}{\rho}}\right]\\
&\qquad\quad+P\left[\sup_j\sup_{t_j\le s\le t_j+\epsilon\rho^n}|x(s)-x(t_j)|\ge a_2\sqrt{n\rho^n\log\frac{1}{\rho}}\right]\\
&\qquad\le 2k_\epsilon\rho^{-n}\left[\exp[-\frac{a_1^2\,n\rho^n\log\frac{1}{\rho}}{2(1+\epsilon)\rho^n}]+\exp[-\frac{a_2^2\,n\rho^n\log\frac{1}{\rho}}{2\epsilon\rho^n}]\right]
\end{aligned}$$

Since $a>\sqrt{2}$, we can pick $a_1>\sqrt{2}$ and $a_2>0$. For $\epsilon>0$ sufficiently small (1.13) is easily verified.

# Chapter 2

# Diffusion Processes

## 2.1 What is a Diffusion Process?

When we want to model a stochastic process in continuous time it is almost impossible to specify in some reasonable manner a consistent set of finite dimensional distributions. The one exception is the family of Gaussian processes with specified means and covariances. It is much more natural and profitable to take an evolutionary approach. For simplicity let us take the one dimensional case where we are trying to define a real valued stochastic process with continuous trajectories. The space $\Omega = C[0,T]$ is the space on which we wish to construct the measure $P$. We have the $\sigma$-fields $\mathcal{B}_t = \sigma\{x(s) : 0 \leq s \leq t\}$ defined for $t \leq T$. The total $\sigma$-field $\mathcal{B} = \mathcal{B}_T$. We try to specify the measure $P$ by specifying approximately the conditional distributions $P[x(t+h)-x(t) \in A|\mathcal{B}_t]$. These distributions are nearly degenerate and and their mean and variance are specified as

$$E^P\big[x(t+h) - x(t)|\mathcal{B}_t\big] = h\, b(t,\omega)) + o(h) \tag{2.1}$$

and

$$E^P\big[(x(t+h) - x(t))^2|\mathcal{B}_t\big] = h\, a(t,\omega)) + o(h) \tag{2.2}$$

as $h \to 0$, where for each $t \geq 0$ $b(t,\omega)$ and $a(t,\omega)$ are $\mathcal{B}_t$ measurable functions. Since we insist on continuity of paths, this will force the distributions to be nearly Gaussian and no additional specification should be necessary. We will devote the next few lectures to investigate this.

Equations (2.1)and (2.2) are infinitesimal differential relations and it is best to state them in integrated forms that are precise mathematical statements.

We need some definitions.

**Definition 2.1.** *We say that a function $f : [0,T] \times \Omega \to R$ is progressively measurable if, for every $t \in [0,T]$ the restiction of $f$ to $[0,t] \times \Omega$ is a measurable function of $t$ and $\omega$ on $([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{B}_t)$ where $\mathcal{B}[0,t]$ is the Borel $\sigma$-field on $[0,t]$.*

The condition is somewhat stronger than just demanding that for each $t$, $f(t, \omega)$ is $\mathcal{B}_t$ is measurable. The following facts are elementary and left as exercises.

*Exercise* 2.1. If $f(t, x)$ is measurable function of $t$ and $x$, then $f(t, x(t, \omega))$ is progressively meausrable.

*Exercise* 2.2. If $f(t, \omega)$ is either left continuous (or right continuous) as function of $t$ for every $\omega$ and if in addition $f(t\, omega)$ is $\mathcal{B}_t$ measurable for every $t$, then $f$ is progressively measurable.

*Exercise* 2.3. There is a sub $\sigma$-field $\Sigma = \Sigma_{pm} \subset \mathcal{B}[0, T] \times \mathcal{B}_T)$ such that progressive measurability is just measurability with respect to $\Sigma_{pm}$. In particular standard operations performed on progreesively measurable functions yield progressively measurable functions.

We shall always insist that the functions $b(\cdot\,, \cdot)$ and $a(\cdot\,, \cdot)$ be progressively measurable. Let us suppose in addition that they are bounded functions. The boundedness will be relaxed at a later stage.

We reformulate conditions 2.1 and 2.2 as

$$M_1(t) = x(t) - x(0) - \int_0^t b(s, \omega) ds \tag{2.3}$$

and

$$M_2(t) = [M_1(t)]^2 - \int_0^t a(s\,, \omega)) ds \tag{2.4}$$

are martingales with respect to $(\Omega), \mathcal{B}_t, P)$.

We can then define a Diffusion Process corresponding to $a, b$ as a measure $P$ on $(\Omega), \mathcal{B})$ such that relative to $(\Omega), \mathcal{B}_t, P)$ $M_1(t)$ and $M_2(t)$ are martingales. If in addition we are given a probability measure $\mu$ as the initial distribution, i.e.

$$\mu(A) = P[x(0) \in A]$$

then we can expect $P$ to be determined by $a, b$ and $\mu$.

We saw already that if $a \equiv 1$ and $b \equiv 0$, with $\mu = \delta_0$, we get the standard Brownian Motion. $a = a(t, x(t))$ and $b = b(t, x(t))$, we expect $P$ to be a Markov Process, because the infinitesimal parameters depend only on the current position and not on the past history. If there is no explicit dependence on time, then the Markov Process can be expected to have stationary transition probabilities. Finally if $a(t, x) = a(t)$ is purely a function of $t$ and $b(t, \omega)) = b_1(t) + \int_0^t c(t\,, s) x(s) ds$ is linear in $\omega$), then one expects $P$ to be Gaussian, if $\mu$ is so.

Because the pathe are continuous the same argument that we provided earlier can be used to establish that

$$\begin{aligned} Z_\lambda(t) &= \exp[\lambda M_1(t) - \frac{\lambda^2}{2} \int_0^t a(s\,, \omega) ds] \\ &= \exp[\lambda[x(t) - x(0) - \int_0^t b(s\,, \omega) ds] - \frac{\lambda^2}{2} \int_0^t a(s\,, \omega) ds] \end{aligned} \tag{2.5}$$

is a martingale with respect to $(\Omega), \mathcal{B}_t, P)$ for every real $\lambda$. We can also take for our definition of a Diffusion Process corresponding to $a, b$ the condition that $Z_\lambda(t)$ be a martingale with respect to $(\Omega), \mathcal{B}_t, P)$ for every $\lambda$. If we do that we did not have to assume that the paths were almost surely continuous. $(\Omega, \mathcal{B}_t, P)$ could be any space suppporting a stochastic process $x(t\,,\omega)$ such that the martingale property holds for $Z_\lambda(t)$. If $C$ is an upper bound for $a$, it is easy to check with $M_1(t)$ defined by equation (2.5)

$$E^P\left[\exp[\lambda[M_1(t) - M_1(s]\right] \leq \exp[\frac{\lambda^2 C}{2}]$$

The lemma of Garsia, Rodemich and Rumsey will guarantee that the paths can be chosen to be continuous.

Let $(\Omega, \mathcal{F}, P)$ be a Probability space. Let $\mathbf{T}$ be the interval $[0, T]$ for some finite $T$ or the infinite interval $[0, \infty)$. Let $\mathcal{F}_T \subset \mathcal{F}$ be sub $\sigma$-fields such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s, t \in \mathbf{T}$ with $s < t$. We can assume with out loss of generality that $\mathcal{F} = \vee_{t\in\mathbf{T}}\mathcal{F}_t$. Let a stochastic process $x(t\,,\omega)$ with values in $R^n$ be given. Assume that it is progressively measurable with respect to $(\Omega\,, \mathcal{F}_t)$. We can easily gneralize the ideas described in the previous section to diffusion processe with values in $R^n$. Given a positive semidefinite $n \times n$ matrix $a = a_{i,j}$ and an $n$-vector $b = b_j$, we define the operator

$$(\mathcal{L}_{a,b} f)(x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{i,j} \partial^2 f \partial x_i \partial x_j(x) + \sum_{j=1}^{n} \partial f \partial x_j(x)$$

If $a(t\,,\omega) = a_{i,j}(t\,,\omega)$ and $b(t\,,\omega) = b_j(t\,,\omega)$ are progresssively measurable functions we define

$$(L_{t\,,\omega} f)(x) = (L_{a(t\,,\omega), b(t\,,\omega)} f)(x)$$

**Theorem 2.1.** *The following defintions are equivalent.* $x(t\,,\omega)$ *is a diffusion process correponding to bounded progressively measurable functions* $a(\cdot\,,\cdot), b(\cdot\,,\cdot)$ *with values in the space of symmetric positive semidefinite* $n \times n$ *matrices, and* $n$*-vectors if*

1. $x(t\,,\omega)$ *has an almost surely continuous version and*

$$y_i(t\,,\omega) = x_i(t\,,\omega) - x_i(0\,,\omega) - \int_0^t b(s\,,\omega) ds$$

*and*

$$z_{i,j}(t\,,\omega) = y_i(t\,,\omega)\, y_j(t\,,\omega) - \int_0^t a_{i,j}(s\,,\omega) ds$$

*are* $(\Omega, \mathcal{F}_t, P)$ *martingales.*

2. *For every* $\lambda \in R^n$

$$Z_\lambda(t\,,\omega) = \exp\left[ < \lambda\,, y(t\,,\omega) > -\frac{1}{2}\int_0^t < \lambda, a(s\,,\omega)\lambda > ds\right]$$

*is an* $(\Omega, \mathcal{F}_t, P)$ *martingale.*

3. *For every* $\lambda \in R^n$

$$X_\lambda(t\,,\omega) = \exp\left[i < \lambda\,, y(t\,,\omega) + \frac{1}{2}\int_0^t < \lambda,\; a(s\,,\omega)\lambda > ds\right]$$

*is an* $(\Omega, \mathcal{F}_t, P)$ *martingale.*

4. *For every smooth bounded function* $f$ *on* $R^n$ *with atleast two bounded continuous derivatives*

$$f(x(t\,,\omega)) - f((x(0\,,\omega)) - \int_0^t (\mathcal{L}_{s,\omega} f)(x(s\,,\omega))ds$$

*is an* $(\Omega, \mathcal{F}_t, P)$ *martingale.*

5. *For every smooth bounded function* $f$ *on* $\mathbf{T} \times R^n$ *with atleast two bounded continuous* $x$ *derivatives and one bounded continuous* $t$ *derivative*

$$f(t\,, x(t\,,\omega)) - f(0\,,(x(0\,,\omega)) - \int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,, x(s\,,\omega))ds$$

*is an* $(\Omega, \mathcal{F}_t, P)$ *martingale.*

6. *For every smooth bounded function* $f$ *on* $\mathbf{T} \times R^n$ *with atleast two bounded continuous* $x$ *derivatives and one bounded continuous* $t$ *derivative*

$$\exp\Big[f(t\,, x(t\,,\omega)) - f(0\,,(x(0\,,\omega)) - \int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,, x(s\,,\omega))ds$$
$$- \frac{1}{2}\int_0^t < (\nabla f)(s\,, x(s\,,\omega)), a(s\,,\omega)\,(\nabla f)(s\,, x(s\,,\omega)) > ds\Big]$$

*is an* $(\Omega, \mathcal{F}_t, P)$ *martingale.*

7. *Same as (6) except that* $f$ *is replaced by* $g$ *of the form*

$$g(t\,,x) =< \lambda,\; x > + f(t\,,x)$$

*where* $f$ *is as in (6) and* $\lambda \in R^n$ *is arbitrary.*

*Under any one of the above definitions,* $x(t\,,\omega)$ *has an almost surely continuous version satifying*

$$P\left[\sup_{0\le s\le t} |y(s\,,\omega) - y(0\,,\omega)| \ge \ell\right] \le 2n\exp[\frac{-\ell^2}{Ct}]$$

*for some constant* $C$ *depending only on the dimension* $n$ *and the upper bound for* $a$*. Here*

$$y_i(t\,,\omega) = x_i(t\,,\omega) - x_i(0\,,\omega) - \int_0^t b_i(s\,,\omega)ds$$

*Proof.* <u>(1) *implies* (2)</u>. This was essentially the content of Theorem and the comments of the previous section. Also we saw that the exponential inequality is a consequence of Doob's inequality.
<u>(2) *implies* (3)</u>. The condition that $Z_\lambda(t)$ is a martingale can be rewritten as a whole collecction of identities

$$\int_A Z_\lambda(t\,,\omega)dP = \int_A Z_\lambda(s\,,\omega)dP \tag{2.6}$$

that is valid for every $t > s$, $A \in \mathcal{F}_s$ and $\lambda \in R^n$. Both sides of eqation (2.6) are well defined when $\lambda \in R^n$ is replaced by $\lambda \in \mathbf{C}^n$, with complex components and define entire functions of the $n$ complex variables $\lambda$. Since they agree when the values are real, by analytic continuation, they must agree for all purely imaginary values of $\lambda$ as well. This is just (3).
<u>(3) *implies* (4)</u>. This part of the proof requires a simple lemma.

**Lemma 2.2.** *Let $M(t\,,\omega)$ be a martingale relative to $(\Omega\mathcal{F}_t, P)$ which has almost surely continuous trajectories and $A(t\,,\omega)$ be a progressively measurable process that is for almost all $\omega$ a continuous function of bounded variation in $t$. Assume that for every $t$ the random variable $\xi(t\,,\omega) = \sup_{0\le s\le t}|M(t)|\,Var_{[0,t]}\,A(t\,,\omega)$ has a finite expectation. Then*

$$\eta(t) = M(t)A(t) - M(0)A(0) - \int_0^T M(s)dA(s)$$

*is again a martingale relative to $(\Omega, \mathcal{F}_t, P)$.*

*Proof.* (of lemma.) We need to prove that for every $s < t$,

$$E^P\left[M(t)A(t) - M(s)A(s) - \int_s^t M(u)dA(u)\big|\mathcal{F}_s\right] = 0 \qquad \text{a.e.}$$

We can subdivide the interval $[s,t]$ into subintervals with end points $s = t_0 < t_1 < \cdots < t_N = t$, and approximate $\int_s^t M(u)dA(u)$ by $\sum_{j=1}^N M(t_j)[A(t_j) - A(t_{j-1})]$. The fact that $A$ is continuous and $\xi(t)$ is integrable makes the approximation work in $L_1(P)$ so that

$$\begin{aligned}
E^P\left[\int_s^t M(u)dA(u)\big|\mathcal{F}_s\right] &= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N M(t_j)[A(t_j) - A(t_{j-1})]\big|\mathcal{F}_s\right]\\
&= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N [M(t_j)A(t_j) - M(t_j)A(t_{j-1})]\big|\mathcal{F}_s\right]\\
&= \lim_{N\to\infty} E^P\left[\sum_{j=1}^N [M(t_j)A(t_j) - M(t_{j-1})A(t_{j-1})]\big|\mathcal{F}_s\right]\\
&= E^P\left[M(t)A(t) - M(s)A(s)\right]
\end{aligned}$$

and we are done. We used the martingale property in going from the second line to the third when we replaced $M(t_j)A(t_{j-1})$ by $M(t_{j-1})A(t_{j-1})$ $\square$

Now we return to the proof of the theorem. Let us apply the above lemma with $M_\lambda(t) = X_\lambda(t)$ and

$$A_\lambda(t) = \exp[i \int_0^t < \lambda\,, b(s) > ds - \frac{1}{2}\int_0^t < \lambda\,, a(s)\lambda > ds].$$

Then a simple computation yields

$$\begin{aligned} M_\lambda(t)A_\lambda(t) - M_\lambda(0)A_\lambda(0) - \int_0^t M_\lambda(s)dA_\lambda(s) \\ = e_\lambda(x(t) - x(0)) - 1 - \int_0^t (\mathcal{L}_{s,\omega}e_\lambda)((x(s) - x(0))ds \end{aligned}$$

where $e_\lambda(x) = \exp[i < \lambda\,, x >]$. Multiplying by $\exp[i < \lambda\,, x(0) >]$, which is essentially a constant, we conclude that

$$e_\lambda(x(t)) - e_\lambda(x(0)) - \int_0^t (\mathcal{L}_{s,\omega}e_\lambda)((x(s))ds$$

is a martingale. The above expression is just what we had to prove, except that our $f$ is special namely, the exponentials $e_\lambda(x)$. But by linear combinations and limits we can easily pass from exponentials to arbitray smooth bounded functions with two bounded derivatives. We first take care of infinitely diffrentiable functions with compact support by Fourier integrals and then approximate twice differentiable functions with those.

(4) *implies* (3). The steps can be retraced. We start with the martingales defined by (4) in the special case of $f$ being $e_\lambda$ and choose

$$A_\lambda(t) = \exp[-i \int_0^t < \lambda\,, b(s) > ds + \frac{1}{2}\int_0^t < \lambda\,, a(s)\lambda > ds]$$

and do the computations to get back to the martingales of type (3).

(4) *implies* (5). This is basically a computation. If $f(t\,, x)$ can be approximated by smooth function and so we may assume with out loss of generality more

derivatives.

$$
\begin{aligned}
&E^P[f(t\,,x(t)) - f(s\,,x(s))|\mathcal{F}_s] \\
&\quad = E^P[f(t\,,x(t)) - f(t\,,x(s))|\mathcal{F}_s] + E^P[f(t\,,x(s)) - f(s\,,x(s))|\mathcal{F}_s] \\
&\quad = E^P[\int_s^t (\mathcal{L}_{u,\omega} f(t\,,\cdot))(x(u))du|\mathcal{F}_s] + E^P[\int_s^t \frac{\partial f}{\partial u}(u\,,x(s))du|\mathcal{F}_s] \\
&\quad = E^P[\int_s^t (\mathcal{L}_{u,\omega} f(u\,,\cdot))(x(u))du|\mathcal{F}_s] \\
&\qquad + E^P[\int_s^t (\mathcal{L}_{u,\omega}[f(t\,,\cdot) - f(u\,,\cdot)])(x(u))]du|\mathcal{F}_s] \\
&\qquad\quad + E^P[\int_s^t \frac{\partial f}{\partial u}(u\,,x(u))du|\mathcal{F}_s] \\
&\qquad\qquad + E^P[\int_s^t [\frac{\partial f}{\partial u}(u\,,x(s)) - \frac{\partial f}{\partial u}(u\,,x(u))]du|\mathcal{F}_s] \\
&\quad = E^P[\int_s^t [\frac{\partial f}{\partial u} + (\mathcal{L}_{u,\omega} f)](u\,,x(u))du|\mathcal{F}_s] + J
\end{aligned}
$$

where

$$
\begin{aligned}
J &= E^P[\int_s^t (\mathcal{L}_{u,\omega}[f(t\,,\cdot) - f(u\,,\cdot)])(x(u))du|\mathcal{F}_s] \\
&\quad + E^P[\int_s^t [\frac{\partial f}{\partial u}(u\,,x(s)) - \frac{\partial f}{\partial u}(u\,,x(u))]du|\mathcal{F}_s] \\
&= E^P[\int_s^t \int_u^t (\frac{\partial f}{\partial v}\mathcal{L}_{u,\omega} f)(v\,,x(u))du\,dv|\mathcal{F}_s] \\
&\quad - E^P[\int_s^t \int_s^u (\mathcal{L}_{v,\omega}\frac{\partial f}{\partial u})(u\,,(x(v))du\,dv|\mathcal{F}_s] \\
&= E^P\Big[\int\int_{s\le u\le v\le t} (\mathcal{L}_{u,\omega}\frac{\partial f}{\partial v})(v\,,(x(u))du\,dv \\
&\quad - \int\int_{s\le v\le u\le t} (\mathcal{L}_{v,\omega}\frac{\partial f}{\partial u})(u\,,(x(v))du\,dv\Big] \\
&= 0.
\end{aligned}
$$

The two integrals are identical, just the roles of $u$ and $v$ have been interchanged.
<u>(5) *implies* (4)</u>. This is trivial because after all in (5) we are allowed to take $f$ to be purely a function of $x$.
<u>(5) *implies* (6)</u>. This is again the lemma on multiplying a martingale by a function of bounded variation. We start with a function of the form $\exp[f(t\,,x)]$ and the martingale

$$\exp[f(t\,,x(t))] - \exp[f(0\,,x(0))] - \int_0^t (\frac{\partial e^f}{\partial s} + \mathcal{L}_{s,\omega} e^f)(s\,,x(s))ds$$

and use

$$A(t) = \exp\big[ -\int_0^t (\frac{\partial f}{\partial s} + \mathcal{L}_{s,\omega} f)(s\,,x(s))ds \\ - \frac{1}{2}\int_0^t < (\nabla f)(s\,,x(s))\,,a(s)(\nabla f)(s\,,x(s)) > ds\big]$$

<u>(6) *implies* (5)</u>. This just again reversing the steps.
<u>(6) *implies* (7)</u>. The problem here is that the function $< \lambda\,, x >$ are unbounded. If we pick a function $h(x)$ of one variable to equal $x$ in the interval $[-1.1]$ and then levels off smoothly we get easily a smooth bounded function with bounded derivatives that agrees with $x$ in $[-1,1]$. Then the sequence $h_(x) = kh(\frac{x}{k})$ clearly converges to $x$, $|h_k(x)| \leq |x|$ and more over $|h'_k(x)|$ is uniformly bounded in $x$ and $k$ and $|h''_k(x)|$ goes to 0 uniformly in $k$. We approximate $< \lambda\,, x >$ by $\sum_j \lambda_j h_k(x_j)$ and consider the martingales

$$\exp\left[\sum_j \lambda_j h_k(x_j(t)) - \sum_j \lambda_j h_k(x_j(0)) - \int_0^t \psi_k^\lambda(s)ds\right]$$

where

$$\psi_k^\lambda(s) = \int_0^t \sum_j \lambda_j b_j(s\,,\omega)h'_k(x_j(s))ds + \frac{1}{2}\int_0^t \sum_j a_{j,j}(s\,,\omega)h''_k(x_j(s))ds \\ + \frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\lambda_i\lambda_j h'_i(x_i(s)h'_j(x_j(s)ds$$

and converges to

$$\psi^\lambda(s) = \int_0^t \sum_j \lambda_j b_j(s\,,\omega)ds + \frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\lambda_i\lambda_j ds$$

as $k \to \infty$. By Fatous's lemma the limit of nonnegative martingales is always a supermartingale and therefore in the limit

$$\exp\left[< \lambda\,, x(t) - x(0) > - \int_0^t \psi^\lambda(s)ds\right]$$

is a supermartingale. In particular

$$E^P\left[\exp[< \lambda\,, x(t) - x(0) > - \int_0^t \psi^\lambda(s)ds]\right] \leq 1$$

If we now use the bound on $\psi$ it is easy to obtain the estimate

$$E^P[\exp[< \lambda\,, x(t) - x(0) >] \leq C_\lambda$$

This provides the necessary uniform integrability to conclude that in the limt we have a martingale. Once we have the estimate, it is easy to see that we can approximate $f(t\,,x)+ <\lambda\,,x>$ by $f(t\,,x)+\sum_j \lambda_j h_k(x_j)$ and pass to the limit, thus obtaining (7) from (6). Of course (7) implies both (2) and (6). Also all the exponential estimates follow at this point. Once we have the estimates there is no difficulty in obtainig (1) from (3). We need only take $f(x)=x_i$ and $x_i x_j$ that can be justified by the estimates. Some minor manipulation is needed to obtain the results in the form presented. $\square$

## 2.2 Random walks and Brownian Motion

Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables with mean 0 and variance 1. The partial sums $S_k$ are defined by $S_0 = 0$ and for $k \geq 1$

$$S_k = X_1 + X_2 + \cdots + X_k$$

We rescale and interpolate to define stochastic processes $X_n(t) : 0 \leq t \leq 1$ by

$$X_n\big(\frac{k}{n}\big) = \frac{S_k}{\sqrt{n}}$$

for $0 \leq k \leq n$ and for $1 \leq k \leq n$ and $t \in [\frac{k-1}{n}, \frac{k}{n}]$

$$X_n(t) = (nt-k+1)X_n\big(\frac{k}{n}\big) + (k-nt)X_n\big(\frac{k-1}{n}\big)$$

Let $P_n$ denote the distribution of the process $X_n(\cdot)$ on $\mathbf{X} = C[0\,,1]$ and $P$ the distribution of Brownian Motion, or the Wiener measure as it is often called. We want to explore the sense in which

$$\lim_{n\to\infty} P_n = P$$

**Lemma 2.3.** *For any finite collection $0 \leq t_1 < t_2 < \cdots < t_m \leq 1$ of $m$ time points the joint distribution of $(x(t_1), \cdots, x(t_m))$ under $P_n$ converges, as $n \to \infty$, to the corresponding distribution under $P$.*

*Proof.* We are dealing here basically with the central limit theorem for sums independent random variables. Let us define $k_n^i = [nt_i]$ and the increments

$$\xi_n^i = \frac{S_{k_n^i} - S_{k_n^{i-1}}}{\sqrt{n}}$$

for $i = 1, 2, \cdots, m$ with the convention $k_n^0 = 0$. For each $n$, $\xi_n^i$ are $m$ mutually independent random variables and their distributions converge as $n \to \infty$ to Gaussians with 0 means and variances $t_i - t_{i-1}$ respectively. We take $t_0 = 0$. This is of course the same distribution for these increments under Brownian Motion. The interpolation is of no consequence, because the difference between the end points is exactly some $\frac{X_i}{\sqrt{n}}$. So it does not really matter if in the definition of $X_n(t)$ if we take $k_n = [nt]$ or $k_n = [nt]+1$ or take the interpolated value. We can state this convergensce in the form

$$\lim_{n\to\infty} E^{P_n}\big[f(x(t_1), x(t_2), \cdots, x(t_m))\big] = E^P\big[f(x(t_1), x(t_2), \cdots, x(t_m))\big]$$

for every $m$, any $m$ time points $(t_1, t_2, \cdots, t_m)$ and any bounded continuous function $f$ on $\mathbf{R}^m$. $\square$

These measures $P_n$ are on the space $\mathbf{X}$ of bounded continuous functions on $[0\,,1]$. The space $\mathbf{X}$ is a metric space with $d(f,g) = \sup_{0\le t\le 1} |f(t)-g(t)|$ as the distance between two continuous functions. The main theorem is

**Theorem 2.4.** *If $F(\cdot)$ is a bounded continuous function on $\mathbf{X}$ then*

$$\lim_{n\to\infty} \int_{\mathbf{X}} F(\omega)dP_n = \int_{\mathbf{X}} F(\omega)dP$$

*Proof.* The main difference is that functions depending on a finite number of coordinates have been replaced by functions that are bounded and continuous, but otherwise arbitrary. The proof proceeds by approximation. Let us assume Lemma 2.5 which asserts that for any $\epsilon > 0$, there is a compact set $K_\epsilon$ such that $\sup_n P_n[\mathbf{X} - K_\epsilon] \le \epsilon$ and $P[\mathbf{X} - K_\epsilon] \le \epsilon$. From standard approximation theory (i.e. Stone-Weierstrass Theorem) the continuous function $F$, which we can assume to be bounded by 1, can be approximated by a function $f$ depending on a finite number of coordinates such that $\sup_{\omega\in K_\epsilon} |F(\omega)-f(\omega)| \le \epsilon$. Moreover we can assume without loss of generality that $f$ is also bounded by 1. We can estimate

$$|\int_{\mathbf{X}} F(\omega)dP_n - \int_{\mathbf{X}} f(\omega)dP_n| \le \int_{K_\epsilon} |F(\omega) - f(\omega)|dP_n + 2P_n[K_\epsilon^c] \le 3\epsilon$$

as well as

$$|\int_{\mathbf{X}} F(\omega)dP - \int_{\mathbf{X}} f(\omega)dP| \le \int_{K_\epsilon} |F(\omega) - f(\omega)|dP + 2P[K_\epsilon^c] \le 3\epsilon$$

Therefore

$$|\int_{\mathbf{X}} F(\omega)\,dP_n - \int_{\mathbf{X}} F(\omega)\,dP| \le 6\epsilon + |\int_{\mathbf{X}} f(\omega)\,dP_n - \int_{\mathbf{X}} f(\omega)\,dP|$$

and we are done. $\square$

*Remark* 2.1. We shall prove Lemma 2.5 under the additional assuption that the underlying random variables $X_i$ have a finite 4-th moment. See the exercise at the end to remove this condition.

**Lemma 2.5.** *Let $P_n, P$ be as before. Assume that the random variables $X_i$ have a finite moment of order four. Then for any $\epsilon > 0$ there exists a compact set $K_\epsilon \subset \mathbf{X}$ such that*

$$P_n[K_\epsilon] \ge 1 - \epsilon$$

*for all $n$ and*

$$P[K_\epsilon] \geq 1 - \epsilon$$

*as well.*

*Proof.* The set

$$K_{B,\alpha} = \{f : f(0) = 0,\ |f(t) - f(s)| \leq B|t - s|^\alpha\}$$

is a compact subset of $\mathbf{X}$ for each fixed $B$ and $\alpha$. Theorem 1.3 can be used to give us a uniform estimate on $P_n[K^c_{B,\alpha}]$ which can be made small by taking $B$ large enough. We need only to check that the condition (1.2) holds for $P_n$ with some constants $\beta, \alpha$ and $C$ that do not depend on $n$. Such an estimate clearly holds for the Brownian motion $P$.

If $\{X_i\}$ are independent identically distributed random variables with zero mean, an elementary calculation yields

$$E[(X_1 + X_2 + \cdots + X_k)^4] = kE[X_1^4] + 3k(k-1)\big[E[X_1^2]\big]^2 \leq C_1 k + C_2 k^2 \quad (2.7)$$

Let us try to estimate $E[(X_n(t) - X_n(s))^4]$. If $|t - s| \leq \frac{2}{n}$ we can estiamte

$$|X_n(t) - X_n(s)| \leq M|t - s|$$

where $M$ is the maximum slope. There are atmost three intervals involved and

$$E[M^4] \leq n^2\, E\big[[\max |X_i|, |X_2|, |X_3|]^4\big] \leq C\, n^2$$

which implies that

$$E^{P_n}\big[|x(t) - x(s)|^4\big] \leq |t - s|^4\, E[M^4] \leq C|t - s|^2 \quad (2.8)$$

If $|t - s| > \frac{2}{n}$ we can find $t', s'$ such that $ns', nt'$ are integers, $|t - t'| \leq \frac{1}{n}$ and $|s - s'| \leq \frac{1}{n}$. Applying the estimate (2.8) for the end pieces that are increments over incomplete intervals and the estimate (2.7) for the piece $|x(t') - x(s')|$, we get

$$E^{P_n}[|x(t) - x(s)|^4] \leq C\, n^{-2} + \frac{C}{n}|t' - s'| + C|t' - s'|^2$$

Since both $|t - s|$ and $|t' - s'|$ are atleast $\frac{1}{n}$ we obtain (1.2). □

*Exercise* 2.4. To extend the result to the case where only the second moment exists, we do truncation and write $X_i = Y_i + Z_i$. The pairs $\{(Y_i, Z_i) : 1 \leq i \leq n\}$ are mutually independent identically distributed random vectors. We can asume that both $Y_i$ and $Z_i$ have mean 0. We can fix it so that $Y_i$ has variance 1 and a finite fourth moment. $Z_i$ can be forced to have an arbitrarily small variance $\sigma^2$. We have $X_n(t) = Y_n(t) + Z_n(t)$ and by Kolmogorov's inequality

$$P\big[\sup_{0 \leq t \leq 1} |Z_n(t)| \geq \delta\big] \leq \delta^{-2} E\big[[Z_n(1)]^2\big] = \delta^{-2}\sigma^2$$

which can be made small uniformly in $n$ if $\sigma^2$ is small enough. Complete the proof.

# Chapter 3

# Stochastic Integration

## 3.1 Stochastic Integrals

If $y_1, \ldots, y_n$ is a martingale relative to the $\sigma$-fields $\mathcal{F}_j$, and if $e_j(\omega)$ are random functions that are $\mathcal{F}_j$ measurable, the sequence

$$z_j = \sum_{k=0}^{j-1} e_k(\omega)[y_{k+1} - y_k]$$

is again a martingale with respect to the $\sigma$-fields $\mathcal{F}_j$, provided the expectations are finite. A computation shows that if

$$a_j(\omega) = E^P[(y_{j+1} - y_j)^2 | \mathcal{F}_j]$$

then

$$E^P[z_j^2] = \sum_{k=0}^{j-1} E^P\left[a_k(\omega)|e_k(\omega)|^2\right]$$

or more precisely

$$E^P\left[(z_{j+1} - z_j)^2 | \mathcal{F}_j\right] = a_j(\omega)|e_j(\omega)|^2 \quad \text{a.e. P}$$

Formally one can write

$$\delta z_j = z_{j+1} - z_j = e_j(\omega)\delta y_j = e_j(\omega)(y_{j+1} - y_j)$$

$z_j$ is called a martingale transform of $y_j$ and the size of $z_n$ measured by its mean square is exactly equal to $E^P\left[\sum_{j=0}^{n-1} |e_j(\omega)|^2 \, a_j(\omega)\right]$. The stochastic integral is just the continuous analog of this.

**Theorem 3.1.** *Let $y(t)$ be an almost surely continuous martingale relative to $(\Omega, \mathcal{F}_t, P)$ such that $y(0) = 0$ a.e. $P$, and*

$$y^2(t) - \int_0^t a(s\,,\omega) ds$$

*is again a martingale relative to* $(\Omega, \mathcal{F}_t, P)$*, where* $a(s\,,\omega)ds$ *is a bounded progressively measurable function. Then for progressively measurable functions* $e(\cdot\,,\cdot)$ *satisfying, for every* $t > 0$*,*

$$E^P\left[\int_0^t e^2(s)a(s)ds\right] < \infty$$

*the stochastic integral*

$$z(t) = \int_0^t e(s)dy(s)$$

*makes sense as an almost surely continuous martingale with respect to* $(\Omega, \mathcal{F}_t, P)$ *and*

$$z^2(t) - \int_0^t e^2(s)a(s)ds$$

*is again a martingale with respect to* $(\Omega, \mathcal{F}_t, P)$*. In particular*

$$E^P\left[z^2(t)\right] = E^P\left[\int_0^t e^2(s)a(s)ds\right] \tag{3.1}$$

*Proof.*
**Step 1.** The statements are obvious if $e(s)$ is a constant.

**Step 2.** Assume that $e(s)$ is a simple function given by

$$e(s\,,\omega) = e_j(\omega) \qquad \text{for } t_j \le s < t_{j+1}$$

where $e_j(\omega)$ is $\mathcal{F}_{t_j}$ measurable and bounded for $0 \le j \le N$ and $t_{N+1} = \infty$. Then we can define inductively

$$z(t) = z(t_j) + e(t_j\,,\omega)[y(t) - y(t_j)]$$

for $t_j \le t \le t_{j+1}$. Clearly $z(t)$ and

$$z^2(t) - \int_0^t e^2(s\,,\omega)a(s\,,\omega)ds$$

are martingales in the interval $[t_j, t_{j+1}]$. Since the definitions match at the end points the martingale property holds for $t \ge 0$.

Step 3. If $e_k(s\,,\omega)$ is a sequence of uniformly bounded progressively measurable functions converging to $e(s\,,\omega)$ as $k \to \infty$ in such a way that

$$\lim_{k\to\infty}\int_0^t |e_k(s)|^2 a(s)ds = 0$$

for every $t > 0$, because of the relation (3.1)

$$\lim_{k,k'\to\infty} E^P\left[|z_k(t) - z_{k'}(t)|^2\right] = \lim_{k,k'\to\infty} E^P\left[\int_0^t |e_k(s) - e_{k'}(s)|^2 a(s)ds\right] = 0.$$

Combined with Doob's inequality, we conclude the existence of a an almost surely continuous martingale $z(t)$ such that

$$\lim_{k\to\infty} E^P\left[\sup_{0\le s\le t}|z_k(s)-z(s)|^2\right]=0$$

and clearly

$$z^2(t)-\int_0^t e^2(s)a(s)ds$$

is an $(\Omega,\mathcal{F}_t,P)$ martingale.

**Step 4.** All we need to worry now is about approximating $e(\cdot\,,\cdot)$. Any bounded progressively measurable almost surely continuous $e(s\,,\omega)$ can be approximated by $e_k(s\,,\omega)=e(\frac{[ks]\wedge k^2}{k}\,,\omega)$ which is piecewise constant and levels off at time $k$. It is trivial to see that for every $t>0$,

$$\lim_{k\to\infty}\int_0^t |e_k(s)-e(s)|^2 a(s)\,ds=0$$

**Step 5.** Any bounded progressively measurable $e(s\,,\omega)$ can be approximated by continuous ones by defining

$$e_k(s\,,\omega)=k\int\limits_{(s-\frac{1}{k})\vee 0}^{s} e\,(u\,,\omega)du$$

and again it is trivial to see that it works.

**Step 6.** Finally if $e(s\,,\omega)$ is un bounded we can approximate it by truncation,

$$e_k(s\,,\omega)=f_k(e(s\,,\omega))$$

where $f_k(x)=x$ for $|x|\le k$ and 0 otherwise.

This completes the proof of the theorem. □

If we have a continuous diffusion process $x(t\,,\omega)$ defined on $(\Omega,\mathcal{F}_t,P)$, corresponding to coefficients $a(t\,,\omega)$ and $b(t\,,\omega)$, then we can define stochastic integrals with respect to $x(t)$. We write

$$x(t\,,\omega)=x(0\,,\omega))+\int_o^t b(s\,,\omega)ds+y(t\,,\omega))$$

and the stochastic integral $\int_0^t e(s)dx(s)$ is defined by

$$\int_0^t e(s)dx(s)=\int_0^t e(s)b(s)ds+\int_0^t e(s)dy(s)$$

For this to make sense we need for every $t$,

$$E^P\big[\int_0^t |e(s)b(s)|ds\big] < \infty \quad \text{and} \quad E^P\big[\int_0^t |e(s)|^2 a(s)ds\big] < \infty$$

If we assume for simplicity that $e$ is bounded then $eb$ and $e^2a$ are uniformly bounded functions in $t$ and $\omega$. It then follows, that for any $\mathcal{F}_0$ measurable $z(0)$, that

$$z(t) = z(0) + \int_0^t e(s)dx(s)$$

is again a diffusion process that corresponds to the coefficients $be, ae^2$. In particular all of the equivalent relations hold good.

*Exercise* 3.1. If $e$ is such that $eb$ and $e^2a$ are bounded, then prove directly that the exponentials

$$\exp\big[\lambda(z(t) - z(0)) - \lambda\int_0^t e(s)b(s)ds - \frac{\lambda^2}{2}\int_0^t a(s)e^2(s)ds\big]$$

are $(\Omega, \mathcal{F}_t, P)$ martingales.

We can easily do the mutidimensional generalization. Let $y(t)$ be a vector valued martingale with $n$ components $y_1(t), \cdots, y_n(t)$ such that

$$y_i(t)y_j(t) - \int_o^t a_{i,j}(s\,,\omega)ds$$

are again martingales with respect to $(\Omega, \mathcal{F}_t, P)$. Assume that the progressively measurable functions$\{a_{i,j}(t\,,\omega)\}$ are symmetric and positive semidefinite for every $t$ and $\omega$ and are uniformly bounded in $t$ and $\omega$. Then the stochastic integral

$$z(t) = z(0) + \int_0^t < e(s), dy(s) = z(0) + \sum_i \int_0^t e_i(s)dy_i(s)$$

is well defined for vector velued progressively measurable functions $e(s\,,\omega)$ such that

$$E^P\big[\int_0^t < e(s)\,, a(s)e(s) > ds\big] < \infty$$

In a similar fashion to the scalar case, for any diffusion process $x(t)$ corresponding to $b(s\,,\omega) = \{b_i(s\,,\omega)\}$ and $a(s\,,\omega) = \{a_{i,j}(s\,,\omega)\}$ and any $e(s\,,\omega)) = \{e_i(s\,,\omega)\}$ which is progressively measurable and uniformly bounded

$$z(t) = z(0) + \int_0^t < e(s)\,, dx(s) >$$

is well defined and is a diffusion corresponding to the coefficients

$$\tilde{b}(s\,,\omega) =< e(s\,,\omega)\,, b(s\,,\omega) > \quad \text{and} \quad \tilde{a}(s\,,\omega) =< e(s\,,\omega)\,, a(s\,,\omega)e(s\,,\omega) >$$

It is now a simple exercise to define stocahstic integrals of the form

$$z(t) = z(0) + \int_0^t \sigma(s\,,\omega)dx(s)$$

where $\sigma(s\,,\omega)$ is a matrix of dimension $m \times n$ that has the suitable properties of boundedness and progressive measurability. $z(t)$ is seen easily to correspond to the coefficients

$$\tilde{b}(s) = \sigma(s)b(s) \quad \text{and} \quad \tilde{a}(s) = \sigma(s)a(s)\sigma^*(s)$$

The analogy here is to linear transformations of Gaussian variables. If $\xi$ is a Gaussian vector in $R^n$ with mean $\mu$ and covariance $A$, and if $\eta = T\xi$ is a linear transformation from $R^n$ to $R^m$, then $\eta$ is again Gaussian in $R^m$ and has mean $T\mu$ and covariance matrix $TAT^*$.

*Exercise* 3.2. If $x(t)$ is Brownian motion in $R^n$ and $\sigma(s\,,\omega)$ is a progreessively measurable bounded function then

$$z(t) = \int_0^t \sigma(s\,,\omega)dx(s)$$

is again a Brownian motion in $R^n$ if and only if $\sigma$ is an orthogonal matrix for almost all $s$ (with repect to Lebesgue Measure) and $\omega$ (with respect to $P$)

*Exercise* 3.3. We can mix stochastic and ordinary integrals. If we define

$$z(t) = z(0) + \int_0^t \sigma(s)dx(s) + \int_0^t f(s)ds$$

where $x(s)$ is a process corresponding to $b(s), a(s)$, then $z(t)$ corresponds to

$$\tilde{b}(s) = \sigma(s)b(s) + f(s) \quad \text{and} \quad \tilde{a}(s) = \sigma(s)a(s)\sigma^*(s)$$

The analogy is again to affine linear transformations of Gaussians $\eta = T\xi + \gamma$.

*Exercise* 3.4. Chain Rule. If we transform from $x$ to $z$ and again from $z$ to $w$, it is the same as makin a single transformation from $z$ to $w$.

$$dz(s) = \sigma(s)dx(s) + f(s)ds \quad \text{and} \quad dw(s) = \tau(s)dz(s) + g(s)ds$$

can be rewritten as

$$dw(s) = [\tau(s)\sigma(s)]dx(s) + [\tau(s)f(s) + g(s)]ds$$

## 3.2 Ito's Formula.

The chain rule in ordinary calculus allows us to compute

$$df(t\,,x(t)) = f_t(t\,,x(t))dt + \nabla f(t\,,x(t)).dx(t)$$

We replace $x(t)$ by a Brownian path, say in one dimension to keep things simple and for $f$ take the simplest nonlinear function $f(x) = x^2$ that is independent of $t$. We are looking for a formula of the type

$$\beta^2(t) - \beta^2(0) = 2\int_0^t \beta(s)\, d\beta(s) \tag{3.2}$$

We have already defined integrals of the form

$$\int_0^t \beta(s)\, d\beta(s) \tag{3.3}$$

as Ito's stochastic integrals. But still a formula of the type (3.2) cannot possibly hold. The left hand side has expectation $t$ while the right hand side as a stochastic integral with respect to $\beta(\cdot)$ is mean zero. For Ito's theory it was important to evaluate $\beta(s)$ at the back end of the interval $[t_{j-1}, t_j]$ before multiplying by the increment $(\beta(t_j) - \beta(t_{j-1})$ to keeep things progressively measurable. That meant the stochastic integral (3.3) was approximated by the sums

$$\sum_j \beta(t_{j-1})(\beta(t_j) - \beta(t_{j-1})$$

over successive partitions of $[0, t]$. We could have approximated by sums of the form

$$\sum_j \beta(t_j)(\beta(t_j) - \beta(t_{j-1}).$$

In ordinary calculus, because $\beta(\cdot)$ would be a continuous function of bounded variation in $t$, the difference would be negligible as the partitions became finer leading to the same answer. But in Ito calculus the differnce does not go to 0. The difference $D_\pi$ is given by

$$\begin{aligned} D_\pi &= \sum_j \beta(t_j)(\beta(t_j) - \beta(t_{j-1}) - \sum_j \beta(t_{j-1}(\beta(t_j) - \beta(t_{j-1}) \\ &= \sum_j (\beta(t_j) - \beta(t_{j-1})(\beta(t_j) - \beta(t_{j-1}) \\ &= \sum_j (\beta(t_j) - \beta(t_{j-1})^2 \end{aligned}$$

An easy computation gives $E[D_\pi] = t$ and $E[(D_\pi - t)^2] = 3\sum_j (t_j - t_{j-1})^2$ tends to 0 as the partition is refined. On the other hand if we are neutral and approximate the integral (3.3) by

$$\sum_j \frac{1}{2}(\beta(t_{j-1}) + \beta(t_j))(\beta(t_j) - \beta(t_{j-1})$$

then we can simplify and calculate the limit as

$$\lim \sum_j \frac{\beta(t_j)^2 - \beta(t_{j-1})^2}{2} = \frac{1}{2}(\beta^2(t) - \beta^2(0))$$

This means as we defined it (3.3) can be calculated as

$$\int_0^t \beta(s)\, d\beta(s) = \frac{1}{2}(\beta^2(t) - \beta^2(0)) - \frac{t}{2}$$

or the correct version of (3.2) is

$$\beta^2(t) - \beta^2(0) = \int_0^t \beta(s)\, d\beta(s) + t$$

Now we can attempt to calculate $f(\beta(t)) - f(\beta(0))$ for a smooth function of one variable. Roughly speaking, by a two term Taylor expansion

$$\begin{aligned} f(\beta(t)) - f(\beta(0)) &= \sum_j [f(\beta(t_j)) - f(\beta(t_{j-1}))] \\ &= \sum_j f'(\beta(t_{j-1})(\beta(t_j)) - \beta(t_{j-1})) \\ &\qquad\qquad + \frac{1}{2}\sum_j f''(\beta(t_{j-1})(\beta(t_j)) - \beta(t_{j-1}))^2 \\ &\qquad\qquad + \sum_j O|\beta(t_j)) - \beta(t_{j-1})|^3 \end{aligned}$$

The expected value of the error term is approximately

$$E\big[\sum_j O|\beta(t_j)) - \beta(t_{j-1})|^3\big] = \sum_j O|t_j - t_{j-1}|^{\frac{3}{2}} = o(1)$$

leading to Ito's formula

$$f(\beta(t)) - f(\beta(0)) = \int_0^t f'(\beta(s)) d\beta(s) + \frac{1}{2}\int_0^t f''(\beta(s)) ds \tag{3.4}$$

It takes some effort to see that

$$\sum_j f''(\beta(t_{j-1})(\beta(t_j)) - \beta(t_{j-1}))^2 \to \int_0^t f''(\beta(s)) ds$$

But the idea is, that because $f''(\beta(s))$ is continuous in $t$, we can pretend that it is locally constant and use that calculation we did for $x^2$ where $f''$ is a constant.

While we can make a proof after a careful estimation of all the errors, in fact we do not have to do it. After all we have already defined the stochastic integral (3.3). We should be able to verify (3.4) by computing the mean square of the difference and showing that it is 0.

In fact we will do it very generally with out much effort. We have the tools already.

**Theorem 3.2.** *Let $x(t)$ be a Diffusion Process with values on $R^d$ corresponding to $[b,a]$, a collection of bounded, progressively measurable coefficients. For any smooth function $u(t\,,x)$ on $[0\,,\infty)\times R^d$*

$$u(t\,,x(t))-u(0\,,x(0))=\int_0^s u_s(s\,,x(s))ds+\int_0^t <(\nabla u)(s\,,x(s))\,,dx(s)> \\ +\frac{1}{2}\int_0^t\sum_{i,j}a_{i,j}(s\,,\omega)\frac{\partial^2 u}{\partial x_i\partial x_j}(s\,,x(s))ds$$

*Proof.* Let us define the stochastic process

$$\xi(t)=u(t\,,x(t))-u(0\,,x(0))-\int_0^s u_s(s\,,x(s))ds-\int_0^t <(\nabla u)(s\,,x(s))\,,dx(s)> \\ -\frac{1}{2}\int_0^t\sum_{i,j}a_{i,j}(s\,,\omega)\frac{\partial^2 u}{\partial x_i\partial x_j}(s\,,x(s))ds \tag{3.5}$$

We define a $d+1$ dimensional process $y(t)=\{u(t\,,x(t)),x(t)\}$ which is also a diffusion, and has its parameters $[\tilde{b},\tilde{a}]$. If we number the extra coordinate by 0, then

$$\tilde{b}_i=\begin{cases}[\frac{\partial u}{\partial s}+\mathcal{L}_{s,\omega}u](s\,,x(s)) & \text{if } i=0\\ b_i(s\,,\omega) & \text{if } i\geq 1\end{cases}$$

$$\tilde{a}_{i,j}=\begin{cases}< a(s\,,\omega)\nabla u\,,\nabla u> & \text{if } i=j=0\\ [a(s\,,\omega)\nabla u]_i & \text{if } j=0, i\geq 1\\ a_{i,j}(s\,,\omega) & \text{if } i,j\geq 1\end{cases}$$

The actual computation is interesting and reveals the connection between ordinary calculus, second order operators and Ito calculus. If we want to know the parametrs of the process $y(t)$, then we need to know what to subtract from $v(t\,,y(t))-v(0\,,y(0))$ to obtain a martingale. But $v(t,\,,y(t))=w(t\,,x(t))$, where $w(t,x)=v(t\,,u(t\,,x)\,,x)$ and if we compute

$$\begin{aligned}(\frac{\partial w}{\partial t}+\mathcal{L}_{s,\omega}w)(t\,,x)&=v_t+v_u[u_t+\sum_i b_i u_{x_i}+\sum_i b_i v_{x_i}+\frac{1}{2}\sum_{i,j}a_{i,j}u_{x_i,x_j}]\\ &\quad+v_{u,u}\frac{1}{2}\sum_{i,j}a_{i,j}u_{x_i}u_{x_j}+\sum_i v_{u,x_i}\sum_j a_{i,j}u_{x_j}\\ &\quad+\frac{1}{2}\sum_{i,j}a_{i,j}v_{x_i,x_j}\\ &=v_t+\tilde{\mathcal{L}}_{t,\omega}v\end{aligned}$$

with

$$\tilde{\mathcal{L}}_{t,\omega} v = \sum_{i \geq 0} \tilde{b}_i(s\,,\omega) v_{y_i} + \frac{1}{2} \sum_{i,j \geq 0} \tilde{a}_{i,j}(s\,,\omega) v_{y_i,y_j}$$

We can construct stochastic integrals with respect to the $d+1$ dimensional process $y(\cdot)$ and $\xi(t)$ defined by (3.5) is again a diffusion and its parameters can be calculated. After all

$$\xi(t) = \int_0^t < f(s\,,\omega)\,, dy(s) > + \int_0^t g(s\,,\omega) ds$$

with

$$f_i(s\,,\omega) = \begin{cases} 1 & \text{if } i = 0 \\ -(\nabla u)_i(s\,, x(s)) & \text{if } i \geq 1 \end{cases}$$

and

$$g(s\,,\omega) = -[\frac{\partial u}{\partial s} + \frac{1}{2} \sum_{i,j} a_{i,j}(s\,,\omega) \frac{\partial^2 u}{\partial x_i \partial x_j}](s\,, x(s))$$

Denoting the parameters of $\xi(\cdot)$ by $[B(s\,,\omega), A(s\,,\omega)]$, we find

$$\begin{aligned} A(s\,,\omega) &=< f(s\,,\omega)\,, \tilde{a}(s\,,\omega) f(s\,,\omega) > \\ &=< a\nabla u\,, \nabla u > -2 < a\nabla u\,, \nabla u > + < a\nabla u\,, \nabla u > \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} B(s\,,\omega) &=< \tilde{b}\,, f > + g = \tilde{b}_0(s\,,\omega) - < b(s\,,\omega)\,, \nabla u(s\,, x(s)) > \\ &\qquad - [\frac{\partial u}{\partial s}(s\,,\omega) + \frac{1}{2} \sum_{i,j} a_{i,j}(s\,,\omega) \frac{\partial^2 u}{\partial x_i \partial x_j}(s\,, x(s))] \\ &= 0 \end{aligned}$$

Now all we are left with is the following

**Lemma 3.3.** *If $\xi(t)$ is a scalar process corresponding to the coefficients $[0,0]$ then*

$$\xi(t) - \xi(0) \equiv 0 \qquad \text{a.e.}$$

*Proof.* Just compute

$$E[(\xi(t) - \xi(0))^2] = E[\int_0^t 0\, ds] = 0$$

□

This concludes the proof of the theorem. □

*Exercise* 3.5. Ito's formula is a local formula that is valid for almost all paths. If $u$ is a smooth function i.e. with one continuous $t$ derivative and two continuous $x$ derivatives (3.4) must still be valid a.e. We cannot do it with moments, because for moments to exist we need control on growth at infinity. But it should not matter. Should it?

**Application: Local time in one dimension. Tanaka Formula.**

If $\beta(t)$ is the one dimensional Brownian Motion, for any path $\beta(\cdot)$ and any $t$, the occupation meausre $L_t(A\,,\omega)$ is defined by

$$L_t(A,\omega) = m\{s : 0 \le s \le t \;\;\&\;\; \beta(s) \in A\}$$

**Theorem 3.4.** *There exists a function $\ell(t\,,y,\,\omega)$ such that, for almost all $\omega$,*

$$L_t(A,\omega) = \int_A \ell(t\,,y\,,\omega)\,dy$$

*identically in $t$.*

*Proof.* Formally

$$\ell(t\,,y\,,\omega) = \int_0^t \delta(\beta(s) - y)ds$$

but, we have to make sense out of it. From Ito's formula

$$f(\beta(t)) - f(\beta(0)) = \int_0^t f'(\beta(s))\,d\beta(s) + \frac{1}{2}\int_0^t f''(\beta(s))ds$$

If we take $f(x) = |x - y|$ then $f'(x) =$ sign $x$ and $\frac{1}{2}f''(x) = \delta(x - y)$. We get the 'identity'

$$|\beta(t) - y| - |\beta(0) - y| - \int_0^t \text{ sign } \beta(s)d\beta(s) = \int_0^t \delta(\beta(s) - y)ds = \ell(t\,,y\,,\omega)$$

While we have not proved the identity, we can use it to define $\ell(\cdot\,,\cdot\,,\cdot)$. It is now well defined as a continuous function of $t$ for almost all $\omega$ for each $y$, and by Fubini's theorem for almost all $y$ and $\omega$.

Now all we need to do is to check that it works. It is enough to check that for any smooth test function $\phi$ with compact support

$$\int_R \phi(y)\ell(t\,,y\,,\omega)\,dy = \int_0^t \phi(\beta(s))ds \tag{3.6}$$

The function

$$\psi(x) = \int_R |x - y|\phi(y)\,dy$$

is smooth and a straigt forward calculation shows

$$\psi'(x) = \int_R \text{ sign } (x - y)\phi(y)\,dy$$

and

$$\psi''(x) = -2\phi(x)$$

It is easy to see that (3.6) is nothing but Ito's formuls for $\psi$. $\square$

1641

*Remark* 3.1. One can estimate

$$E\Big[\int_0^t [\text{ sign } (\beta(s) - y) - \text{ sign } (\beta(s) - z)]d\beta(s)\Big]^4 \leq C|y - z|2$$

and by Garsia- Rodemich- Rumsey or Kolmogorov one can conclude that for each $t$, $\ell(t\,,y\,,\omega)$ is almost surely a continuous function of $y$.

*Remark* 3.2. With a little more work one can get it to be jointly continuous in $t$ and $y$ for almost all $\omega$.