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 (Ω, Σ, 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 σ-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 ω 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 σ-field $\mathcal{B}$ on $\mathbf{X}$ and measurability is to be understood in terms of this σ-field. This natural σ-field, called the Kolmogorov σ-field, is defined as the smallest σ-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 (Ω, Σ, 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 μ_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 μ_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 σ-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 σ-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 β,α 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

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

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

$$\begin{aligned}
&Q\big[x(\cdot) : \sup_{0\le t\le 1}|x_n(t) - x_{n+1}(t)| \ge 2^{-n\gamma}\big] \\
&\le Q\big[\sup_{1\le i\le 2^{n+1}}|x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})| \ge 2^{-n\gamma}\big] \\
&\le 2^{n+1}\sup_{1\le i\le 2^{n+1}} Q\big[|x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})| \ge 2^{-n\gamma}\big] \\
&\le 2^{n+1}2^{n\beta\gamma}\sup_{1\le i\le 2^{n+1}} E^Q\big[|x(\frac{i}{2^{n+1}}) - x(\frac{i-1}{2^{n+1}})|^\beta\big] \\
&\le C2^{n+1}\,2^{n\beta\gamma}\,2^{-(1+\alpha)(n+1)} \\
&\le C2^{-n\delta}
\end{aligned}$$

provided $\delta \le \alpha - \beta\gamma$. For given α, β 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*

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

for some positive constants α, β 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 α 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 Ψ. 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_δ 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

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

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. □

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 $\sigma^2(t) \le C|\log t|^{-a}$ 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 Ψ and p.

On the other hand we will show that, if $\sigma^2(t) \ge c|\log t|^{-1}$, 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 $t_1 \le t \le t_2$ 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 μ_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 α 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.

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

provided α is small enough.

Chapter 2

Stochastic Integration.

2.1 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 σ-field on Ω. In addition we denote by $\mathcal{B}_t$ the σ-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 $[0,T]$

$$E^P\{x(t)|\mathcal{B}_s\} = x(s) \quad \text{a.e.} \quad P \tag{2.1}$$

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

$$E^P\{x(t)^2 - t\,|\mathcal{F}_s\} = x(s)^2 - s \quad \text{a.e.} \quad P \tag{2.2}$$

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 2.1. 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 λ

$$X_\lambda(t) = \exp\big[\lambda x(t) - \frac{\lambda^2}{2}t\big] \tag{2.3}$$

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 (2.3) is more or less the same as proving the central limit theorem. In order to prove (2.3) we can assume with out loss of generality that $s = 0$ and will show that

$$E^P\big[\exp\big[\lambda x(t) - \frac{\lambda^2}{2}t\big]\big] = 1 \tag{2.4}$$

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

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

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 $|x(\tau_{k+1,\epsilon}) - x(\tau_{k,\epsilon})| \leq \epsilon$. We write

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

and

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

To establish (2.4) we calculate the quantity on the left hand side as

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

and show that it is equal to 1. Let us cosider the σ-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

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

In particular for some constant C depending on λ

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

and by induction

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

Since $\epsilon > 0$ is arbitrary we prove one half of (2.4). 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 2.1. Why does Theorem 2.1 fail for the process $x(t) = N(t) - t$ where $N(t)$ is the standard Poisson Process with rate 1?

Remark 2.1. 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

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

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 2.2. 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 $a > \sqrt{2}$

$$\sum_n P\big[\Delta_{\rho^n}(\omega) \geq a\sqrt{n\rho^n \log\frac{1}{\rho}}\big] < \infty \tag{2.5}$$

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 ρ^n is completely contained in some I_j with $t_j \leq s \leq t_j + \epsilon\rho^n$. Then

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

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

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

Since $a > \sqrt{2}$, we can pick $a_1 > \sqrt{2}$ and $a_2 > 0$. For $\epsilon > 0$ sufficiently small (2.5) is easily verified.

2.2 Brownian Motion as a Markov Process.

Brownian motion is a process with independent increments, the increment over any interval of length t has the Gaussian distribution with density

$$q(t,y) = \frac{1}{(2\pi t)^{\frac{d}{2}}}e^{-\frac{\|y\|^2}{2t}}$$

It is therefore a Markov process with transition probability

$$p(t,x,y) = q(t,y-x) = \frac{1}{(2\pi t)^{\frac{d}{2}}}e^{-\frac{\|y-x\|^2}{2t}}$$

The operators

$$(T_t f)(x) = \int f(y)p(t,x,y)dy$$

satisfy $T_tT_s = T_sT_t = T_{t+s}$, i.e the semigroup property. This is seen to be an easy consequence of the Chapman-Kolmogorov equations

$$\int p(t,x,y)p(s,y,z)dy = p(t+s,x,z)$$

The infinitesimal generator of the semigroup

$$(Af)(x) = \lim_{t\to 0} \frac{T_t f - f}{t}$$

is easily calculated as

$$\begin{aligned}(Af)(x) &= \lim_{t\to 0} \frac{1}{t} \int [f(x+y) - f(x)]q(t,y)dy \\ &= \lim_{t\to 0} \frac{1}{t} \int [f(x+\sqrt{t}y) - f(x)]q(t,y)dy \\ &= \frac{1}{2}(\Delta f)(x)\end{aligned}$$

by expanding f in a Taylor series in $\sqrt{t}$. The term that is linear in y integrates to 0 and the quadratic term leads to the Laplace operator. The differential equation

$$\frac{dT_t}{dt} = T_t A = A T_t$$

implies that $u(t,x) = (T_t f)(x)$ satisfies the heat equation

$$u_t = \frac{1}{2}\Delta u$$

and

$$\frac{d}{dt}\int f(y)p(t,x,y)dy = \int \frac{1}{2}(\Delta f)(y)p(t,x,y)dy$$

In particular if E_x is expectation with respect to Brownian motion starting from x,

$$E_x[f(x(t)] - f(x) = E_x\left[\int_0^t \frac{1}{2}(\Delta f)(x(s))ds\right]$$

By the Markov property

$$E_x\left[f(x(t) - f(x(s)) - \int_s^t \frac{1}{2}(\Delta f)(x(\tau))d\tau \Big| \mathcal{F}_s\right] = 0$$

or

$$f(x(t) - f(x(0)) - \int_0^t \frac{1}{2}(\Delta f)(x(\tau))d\tau$$

is a Martingale with respect to Brownian Motion.

It is just one step from here to show that for functions $u(t,x)$ that are smooth

$$u(t,x(t)) - u(0,x(0)) - \int_0^t [\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u](s,x(s))ds \tag{2.6}$$

is a martingale. There are in addition some natural exponential Martingales associated with Brownian motion. For instance for any $\lambda \in R^d$,

$$\exp[< \lambda, x(t) - x(0) > -\frac{1}{2}\|\lambda\|^2 t]$$

is a martingale. More generally for any smooth function $u(t,x)$ that is bounded away from 0,

$$u(t,x(t))\exp\left[-\int_0^t\left[\frac{\frac{\partial u}{\partial t}+\frac{1}{2}\Delta u}{u}\right](s,x(s))ds\right] \tag{2.7}$$

is a martingale. In particular if

$$\frac{\partial u}{\partial t}+\frac{1}{2}\Delta u+v(t,x)u(t,x)=0$$

then

$$u(t,x(t))\exp[\int_0^t v(s,x(s))ds]$$

is a Martingale, which is the Feynman-Kac formula. To prove (2.7) from (2.6), we make use of the following elementary lemma.

Lemma 2.2. *Suppose $M(t)$ is almost surely continuous martingale with respect to $(\Omega,\mathcal{F}_t,P)$ and $A(t)$ is a progressively measurable function, which is almost surely continuous and of bounded variation in t. Then, under the assumption that $\sup_{0\le s\le t}|M(s)|Var_{0,t}A(\cdot,\omega)$ is integrable,*

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

is again a Martingle.

Proof. The main step is to see why

$$E[M(t)A(t)-M(0)A(0)-\int_0^t M(s)dA(s)]=0$$

Then the same argument, repeated conditionally will prove the martingale property.

$$\begin{aligned}
E[M(t)A(t)-M(0)A(0)] &= \lim\sum_j E[M(t_j)A(t_j)-M(t_{j-1})A(t_{j-1})]\\
&= \lim\sum_j E[M(t_j)A(t_{j-1})-M(t_{j-1})A(t_{j-1})]\\
&\quad+\lim\sum_j E[M(t_j)[A(t_j)-A(t_{j-1})]]\\
&= \lim\sum_j E[M(t_j)[A(t_j)-A(t_{j-1})]]\\
&= E\Big[\int_0^t M(s)dA(s)\Big]
\end{aligned}$$

The limit is over the partition $\{t_j\}$ becoming dense in $[0,t]$ and ones uses the integrability of $\sup_{0\le s\le t}|M(s)|Var_{0,t}A(\cdot)$ and the dominated convergence theorem to complete the proof. ∎

Now, to go from (2.6) to (2.7), we choose

$$M(t) = u(t, x(t)) - u(0, x(0)) - \int_0^t [\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u](s, x(s))ds$$

and

$$A(t) = \exp\left[- \int_0^t \left[\frac{\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u}{u} \right](s, x(s))ds \right]$$

2.3 Stochastic Integrals

If $y_1, \ldots, y_n$ is a martingale relative to the σ-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 σ-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

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

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 2.3. *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*

$$E^P\big[z^2(t)\big] = E^P\big[\int_0^t e^2(s)a(s)ds\big] \tag{2.8}$$

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

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

where $e_j(\omega)$ is $\mathcal{F}_{t_j}$ measurable and bounded for $0 \leq j \leq N$ 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 $t_j \leq t \leq t_{j+1}$. 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 $t \geq 0$.

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 (2.8)

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

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

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

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

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

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. □

Suppose we have an almost surely continuous process $x(t\,,\omega)$ defined on some $(\Omega, \mathcal{F}_t, P)$, and progressively measurable functions $b(s,\omega), a(s,\omega)$ with $a \ge 0$, such that

$$x(t\,,\omega) = x(0\,,\omega) + \int_0^t b(s\,,\omega)ds + y(t\,,\omega)$$

where $y(t,\omega)$ and

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

are martingales with respect to $(\Omega, \mathcal{F}_t, P)$. The stochastic integral $z(t) = \int_0^t e(s)dx(s)$ is defined by

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

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 eb and e^2a are uniformly bounded functions in t and ω. 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 an almost surely continuous process such that

$$z(t) = z(0) + \int_0^t b'(s,\omega)ds + y'(t,\omega)$$

where $y'(t)$ and

$$y'(t)^2 - \int_0^t a'(s,\omega)ds$$

are martingales with $b' = eb$ and $a' = e^2a$.

Exercise 2.3. 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 ω and are uniformly bounded in t and ω. 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 ξ is a Gaussian vector in R^n with mean μ and covariance A, and if $\eta = T\xi$ is a linear transformation from R^n to R^m, then η is again Gaussian in R^m and has mean $T\mu$ and covariance matrix TAT^*.

Exercise 2.4. 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 σ is an orthogonal matrix for almost all s (with repect to Lebesgue Measure) and ω (with respect to P)

Exercise 2.5. 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 2.6. 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

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

2.4 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

$$\beta^2(t) - \beta^2(0) = 2\int_0^t \beta(s)\, d\beta(s) \tag{2.9}$$

We have already defined integrals of the form

$$\int_0^t \beta(s)\, d\beta(s) \tag{2.10}$$

as Ito's stochastic integrals. But still a formula of the type (2.9) 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 (2.10) 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_π 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 (2.10) 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 (2.10) 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 (2.9) 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

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

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

$$f(\beta(t)) - f(\beta(0)) = \int_0^t f'(\beta(s))d\beta(s) + \frac{1}{2}\int_0^t f''(\beta(s))ds \tag{2.11}$$

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 (2.10). We should be able to verify (2.11) 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 2.4. *Let $x(t)$ be an almost surely continuous process with values on R^d such that*

$$y_i(t) = x_i(t) - x_i(0) - \int_0^t b_i(s,\omega)ds \tag{2.12}$$

and

$$y_i(t)y_j(t) - \int_0^t a_{i,j}(s,\omega)ds \tag{2.13}$$

are martingales for $1 \le i,j \le d$. For any smooth function $u(t\,,x)$ on $[0\,,\infty)\times R^d$

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

Proof. Let us define the stochastic process

$$\begin{aligned} \xi(t) = &u(t\,,x(t)) - u(0\,,x(0)) - \int_0^s u_s(s\,,x(s))ds \\ &- \int_0^t < (\nabla u)(s\,,x(s))\,,dx(s) > -\frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\frac{\partial^2 u}{\partial x_i \partial x_j}(s\,,x(s))ds \end{aligned} \tag{2.14}$$

We define a $d+1$ dimensional process $\tilde{x}(t) = \{u(t\,,x(t)), x(t)\}$ which is again a process with almost surely continuous paths satisfying relations analogous to (2.12) and (2.13) with $[\tilde{b}, \tilde{a}]$. If we number the extra coordinate by 0, then

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

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

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

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

with

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

We can construct stochastic integrals with respect to the $d+1$ dimensional process $y(\cdot)$ and $\xi(t)$ defined by (2.14) is again an almost surely continuous process 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

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

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 2.5. *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 2.7. 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 (2.11) 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 2.6. *There exists a function $\ell(t\,, y,\, \omega)$ such that, for almost all ω,*

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

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 ω for each y, and by Fubini's theorem for almost all y and ω.

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

$$\int_R \phi(y)\ell(t\,,y\,,\omega)\,dy = \int_0^t \phi(\beta(s))ds \tag{2.15}$$

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 (2.15) is nothing but Ito's formuls for ψ. □

Remark 2.2. 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 2.3. With a little more work one can get it to be jointly continuous in t and y for almost all ω.

Chapter 3

Stochastic Differential Equations.

3.1 Existence and Uniqueness.

One of the ways of constructing a Diffusion process is to solve the stochastic differential equation

$$dx(t) = \sigma(t, x(t)) \cdot d\beta(t) + b(t, x(t))dt \ ; \ x(0) = x_0 \tag{3.1}$$

where $x_0 \in R^d$ is either nonrandom or measurable with respect to $\mathcal{F}_0$. This is of course written as a stochastic integral equation

$$x(t) = x(0) + \int_0^t \sigma(s, x(s)) \cdot d\beta(s) + \int_0^t b(s, x(s))ds \tag{3.2}$$

If $\sigma(s, x)$ and $b(s, x)$ satisfy the following conditions

$$|\sigma(s, x)| \leq C(1 + |x|) \ ; \ |b(s, x)| \leq C(1 + |x|) \tag{3.3}$$

$$|\sigma(s, x) - \sigma(s, y)| \leq C|x - y| \ ; \ |b(s, x) - b(s, y)| \leq C|x - y| \tag{3.4}$$

by a Picard type iteration scheme one can prove existence and uniqueness.

Theorem 3.1. *Given σ, b that satisfy (3.3) and (3.4), for given x_0 which is $\mathcal{F}_0$ measurable, there is a unique solution $x(t)$ of (3.2), with in the class of progressively measurable almost surely continuous solutions.*

Proof. Define iteratively

$$x_0(t) \equiv x_0$$

$$x_n(t) = x_0 + \int_0^t \sigma(s, x_{n-1}(s)) \cdot d\beta(s) + \int_0^t b(s, x_{n-1}(s))ds \tag{3.5}$$

If we denote the difference $x_n(t) - x_{n-1}(t)$ by $z_n(t)$, then

$$z_{n+1}(t) = \int_0^t [\sigma(s, x_n(s)) - \sigma(s, x_{n-1}(s))] \cdot d\beta(s) + \int_0^t [b(s, x_n(s)) - b(s, x_{n-1}(s))] ds$$

If we limit ourselves to a finite interval $0 \le t \le T$, then

$$E\big[\big|\int_0^t [\sigma(s, x_n(s)) - \sigma(s, x_{n-1}(s))] \cdot d\beta(s)\big|^2\big] \le CE\big[\int_0^t |z_n(s)|^2 ds\big]$$

and

$$E\big[\big|\int_0^t [b(s, x_n(s)) - b(s, x_{n-1}(s))] ds\big|^2\big] \le CTE\big[\int_0^t |z_n(s)|^2 ds\big]$$

Therefore

$$E\big[|z_{n+1}(t)|^2\big] \le C_T E\big[\int_0^t |z_n(s)|^2 ds\big]$$

With the help of Doob's inequality one can get

$$\Delta_{n+1}(t) = E\big[\sup_{0 \le s \le t} |z_{n+1}(s)|^2\big] \le C_T E\big[\int_0^t |z_n(t)|^2 dt\big] \le C_T \int_0^t \Delta_n(s) ds$$

By induction this yields

$$\Delta_n(t) \le A \frac{C_T^n t^n}{n!}$$

which is sufficient to prove the existence of an almost sure uniform limit $x(t)$ of $x_n(t)$ on bounded intervals $[0, T]$. The limit $x(t)$ is clearly a solution of (3.2). Uniqueness is essentially the same proof. For the difference $z(t)$ of two solutions one quickly establishes

$$E\big[|z(t)|^2\big] \le C_T E\big[\int_0^t |z(s)|^2 ds\big]$$

which suffices to prove that $z(t) = 0$. □

Once we have uniqueness one should thing of $x(t)$ as a map of x_0 and the Brownian increments $d\beta$ in the interval $[0, t]$. In particular $x(t)$ is a map of $x(s)$ and the Brownian increments over the interval $[s, t]$. Since $x(s)$ is $\mathcal{F}_s$ measurable, we can conclude that $x(t)$ is a Markov process with transition probability

$$p(s, x, t, A) = P[x(t; s, x) \in A]$$

where $x(t; s, x)$ is the solution of (3.2) for $t \ge s$, initialised to start with $x(s) = x$.

It is easy to see, by an application of Itô's lemma that

$$\begin{aligned}M(t) =& u(t, x(t)) - u(s, x(s)) - \int_s^t \big[\frac{\partial u}{\partial s}(s, x(s)) \\ &+ \frac{1}{2}\sum_{i,j} a_{i,j}(s, x(s))\frac{\partial^2 u}{\partial x_i \partial x_j}(s, x(s)) + \sum_i b_i(s, x(s))\frac{\partial u}{\partial x_i}(s, x(s))\big]ds\end{aligned}$$

is a martingale, where $a = \sigma\sigma^*$, i.e.

$$a_{i,j}(s, x) = \sum_k \sigma_{i,k}(s, x)\sigma_{k,j}(s, x)$$

The process $x(t)$ is then clearly the Diffusion process associated with

$$L_s = \frac{1}{2}\sum_{i,j} a_{i,j}(s, x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i(s, x)\frac{\partial}{\partial x_i}$$

Remark 3.1. We might consider equations of the form

$$dx(t) = \sigma(t, \omega, x(t)) \cdot d\beta(t) + b(t, \omega, x(t))dt$$

where $\sigma(t, \omega, x)$ and $b(t, \omega, x)$ are progressively measurable, bounded (or have linear growth in x) and satisfy a Lipschitz condition in x. There will be a unique solution. But, in general, it will not be Markov if σ and b depend on ω.

3.2 Smooth dependence on parameters.

If σ and b depend smoothly on an additional parameter θ then we will show that the solution $x(t) = x(t, \theta, \omega)$ will depend smoothly on the parameter. The idea is to start with the solution

$$x(t, \theta, \omega) = x_0(\theta, \omega) + \int_0^t \sigma(s, x(s, \theta), \theta) \cdot d\beta(s) + \int_0^t b(s, x(s, \theta), \theta)ds$$

Differentiating with respect to θ, and denoting by Y the derivative, we get

$$\begin{aligned}Y(t, \theta) = y_0(\theta) &+ \int_0^t \big[\sigma_x(s, x(s, \theta), \theta)Y(s, \theta) + \sigma_\theta(s, x(s, \theta), \theta)\big] \cdot d\beta(s) \\ &+ \int_0^t \big[b_x(s, x(s, \theta), \theta)Y(s, \theta) + b_\theta(s, x(s, \theta), \theta)\big]ds\end{aligned}$$

We look at (x, Y) as an enlarged system satisfying

$$\begin{aligned}dx =& \sigma \cdot d\beta + bdt \\ dY =& [\sigma_x Y + \sigma_\theta] \cdot d\beta + [b_x Y + b_\theta]dt\end{aligned}$$

We can repeat this process with higher order derivatives. Denoting by $Y^{(n)}$ the mixed derivatives of order n of $x(t,\theta,\omega)$ with respect to θ, we obtain equations of the form

$$dY^{(n)}(t,\theta,\omega) = \sigma_x Y^{(n)} \cdot d\beta(t) + b_x Y^{(n)} \cdot dt + c^{(n)}(t,\theta,\omega) \cdot d\beta(t) + d^{(n)}(t,\theta,\omega)dt \tag{3.6}$$

where $c^{(n)}$ and $d^{(n)}$ are polynomials in $Y^{(j)}$ with $j \leq n-1$ with coefficients that are functions of t,θ,ω that have finite moments of all orders. Because σ_x and b_x are bounded, we can prove the existence of the solution $Y^{(n)}(t)$ to the linear SDE, with moment estimates. We can in fact go through the iteration scheme in such a manner that the approximations to $Y^{(n)}$ are the derivatives with respect to θ of the corresponding approximation of $x(t,\theta,\omega)$. The limits $Y^{(n)}$ can be shown to be the derivatives $\nabla_\theta^n x(t,\theta,\omega)$. We therefore arrive at the following Theorem.

Theorem 3.2. *Let $\sigma(t,x,\theta)$, $b(t,x,\theta)$ and $x_0(\theta,\omega)$ satisfy*

$$E[\|\nabla_\theta^n x_0(\theta,\omega)\|^p] \leq C_{n,p}$$
$$\|\nabla_\theta^n \sigma\| \leq C_n, \|\nabla_\theta^n b\| \leq C_n;$$
$$\|\sigma(t,x,\theta) - \sigma(t,y,\theta)\| \leq C\|x-y\|$$
$$\|b(t,x,\theta) - b(t,y,\theta)\| \leq C\|x-y\|$$

Then the solution $x(t,\omega,\theta)$ of

$$x(t,\omega,\theta) = x_0(\theta,\omega) + \int_0^t \sigma(t,x(s,\omega),\theta) \cdot d\beta(s) = \int_0^t b(t,x(s,\omega),\theta) \cdot ds$$

has derivatives in L_p of all orders with respect to θ and $Y^{(n)}(t) = \nabla_\theta^n x(t,\omega,\theta)$ has moments of all orders and satisfies the SDE (3.6).

Corollary 3.3. *If $x(t)$ is viewed as a function of the starting point x, then one can view x as the parameter and conclude that if the coefficients have bounded derivatives of all orders then the solution $x(t)$ is almost surely an infinitely differentiable function of its starting point.*

Proof. One needs to observe that if a stochastic process $\xi(x,\omega)$ on R^d has derivatives in L_p of all orders for some $p > 1$, then it is in fact C_∞ in the classical sense for almost all ω. This is a consequence of Sobolev's lemma and Fubini's theorem. □

Remark 3.2. Since smoothness is a local property, if σ and b have at most linear growth, the solution exists for all time with out explosion, and then one can modify the coefficients outside a bounded domain with out changing much. This implies that with out uniform bounds on derivatives the solutions $x(t)$ will still depend smoothly on the initial point, but the derivatives may not have moment estimates.

Remark 3.3. This means one can view the solution $u(t,x)$ of the equation

$$du(t,x) = \sigma(u(t,x)) \cdot d\beta + b(u(t,x))dt;\ u(0,x) = x$$

as random flow $u(t) : R^d \to R^d$. The flow as we saw is almost surely smooth.

3.3 Itô and Stratonovich Integrals.

In the definition of the stochastic integral

$$\eta(t) = \int_0^t f(s)dx(s)$$

we approximated it by sums of the form

$$\sum_j f(t_{j-1})[x(t_j) - x(t_{j-1})]$$

always sticking the increments in the future. This allowed the integrands to be more or less arbitrary, so long as it was measurable with respect to the past. This meshed well with the theory of martingales and made estimation easier. Another alternative, symmetric with respect to past and future, is to use the approximation

$$\sum_j \frac{[f(t_{j-1}) + f(t_j)]}{2}[x(t_j) - x(t_{j-1})]$$

It is not clear when this limit exists. When it exists it is called the Stratonovich integral and is denoted by $\int f(s) \circ dx(s)$. If $f(s) = f(s, x(s))$, then the difference between the two integrals can be explicitly calculated.

$$\int_0^t f(s, x(s)) \circ \, dx(s) = \int_0^t f(s, x(s)) \cdot dx(s) + \frac{1}{2}\int_0^t a(s)ds$$

where

$$\int_0^t a(s)ds = \lim \sum_j [f(t_j, x(t_j)) - f(t_{j-1}, x(t_{j-1}))][x(t_j) - x(t_{j-1})]$$

If $x(t)$ is just Brownian motion in R^d, then $a(s) = (\nabla \cdot f)(s, x(s))$. More generally if

$$\lim \sum_j [x_i(t_j) - x_i(t_{j-1})][x_k(t_j) - x_k(t_{j-1})] = \int_0^t a_{i,k}(s)ds$$

then

$$a(s) = \sum_{i,k} f_{i,k}(s, x(s))a_{i,k}(s) = Tr[(Df)(s, x(s))a(s)]$$

Solutions of

$$dx(t) = \sigma(t, x(t)) \cdot d\beta(t) + b(t, x(t))dt$$

can be recast as solutions of

$$dx(t) = \sigma(t, x(t)) \circ d\beta(t) + \tilde{b}(t, x(t))dt$$

with b and $\tilde{b}$ related by

$$b_i(t, x) = \tilde{b}_i(t, x) + \frac{1}{2}\sum \sigma_{j,k}(t, x)\frac{\partial}{\partial x_j}\sigma_{i,k}(t, x)$$

To see the relevance of this, one can try to solve

$$dx(t) = \sigma(t, x(t)) \cdot d\beta_j(t) + b(t, x(t))dt$$

by approximating $\beta(t)$ by a piecewise linear approximation $\beta^{(n)}(t)$ with derivative $f^{(n)}(t)$. Then we will have just ODE's

$$\frac{dx^{(n)}(t)}{dt} = \sigma(t, x^{(n)}(t))f^{(n)}(t) + b(t, x^{(n)}(t))$$

where $f^{(n)}(\cdot)$ are piecewise constant. An elementary calculation shows that over an interval of constancy $[t, t+h]$,

$$\begin{aligned} x_i^{(n)}(t+h) &= x_i^{(n)}(t) + \sigma(t, x^{(n)}(t)) \cdot Z_h + b_i(t, x^{(n)}(t))h \\ &\quad + \frac{1}{2} < Z_h, c_i(t, x^{(n)}(t))Z_h > + o((Z_h)^2) \end{aligned}$$

where

$$c_i(t, x) = \sum \sigma_{j,k}(t, x) \frac{\partial}{\partial x_j} \sigma_{i,k}(t, x)$$

and Z_h is a Gaussian with mean 0 and variance hI while

$$\beta^{(n)}(t+h) = \beta_n(t) + Z_h$$

It is not hard to see that the limit of $x^{(n)}(\cdot)$ exists and the limit solves

$$dx(t) = \sigma(t, x(t)) \cdot d\beta(t) + b(t, x(t))dt + \frac{1}{2}c(t, x(t))dt$$

or

$$dx(t) = \sigma(t, x(t)) \circ d\beta(t) + b(t, x(t))dt$$

It is convenient to consider a vector field

$$X = \sum_i \sigma_i(x) \frac{\partial}{\partial x_j}$$

and its square

$$X^2 = \sum_{i,j} \sigma_i(x)\sigma_j(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_j c_j(x) \frac{\partial}{\partial x_j}$$

where

$$c_j(x) = \sum_i \sigma_i(x) \frac{\partial \sigma_j(x)}{\partial x_i}$$

Then the solution of

$$dx(t) = \sigma(t, x(t)) \circ d\beta(t) + b(t, x(t))dt = \sum X_i(t, x(t)) \circ d\beta_i(t) + Y(t, x(t))dt$$

is a Diffusion with generator

$$L_t = \frac{1}{2}\sum X_i(t)^2 + Y(t)$$

When we change variables the vector fields change like ordinary first order calculus and

$$\widehat{L}_t = \frac{1}{2}\sum \widehat{X}_i(t)^2 + \widehat{Y}(t)$$

and the Stratonovich solution

$$dx(t) = \sigma(t, x(t)) \circ d\beta(t) + b(t, x(t))dt$$

transforms like

$$dF(x(t)) = DF \cdot dx(t) = (DF)(x(t))[\sigma(t, x(t)) \circ d\beta(t) + b(t, x(t))dt]$$

The Itô corrections are made up by the difference between the two integrals.

Remark 3.4. Following up on remark (3.2), for each $t > 0$, the solution actually maps $R^d \to R^d$ as a diffeomorphism. To see this it is best to view this through Stratonovich equations. Take $t = 1$. If the forward flow is therough vector fileds $X_i(t), Y(t)$, the reverse flow is through $-X_i(1-t), Y(1-t)$ and the reversed noise is $\hat{\beta}(t) = \beta(1) - \beta(1-t)$. One can see by the piecewise linear approximations that these are actually inverses of each other.