1. Brownian Motion as a 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 parametr 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 usually 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 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 irrelevent, 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 be defined simply as a probability measure Q on $(\mathbf{X}, \mathcal{B})$.

Another point of view is that the only relevent 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 distibutions on various finite dimensional spaces that satisfy the consistency conditions. A theorem of Kolmogorov says that this not all that different. Any such consistent family arises from a Q on $(\mathbf{X}, \mathcal{B})$ which is uniquely determine 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 change the values of a single 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.

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 . *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 valuse 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 counatble 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

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

then we can conclude that

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

The limit $x_\infty(\cdot)$ will be continuous on $\mathbf{T}$ and will coincide with $x(\cdot)$ on $\mathbf{T}^0$ thereby 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

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

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

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

provided $\delta\le\alpha-\beta\gamma$. For given α,β we can pick $\gamma<\frac{\alpha}{\beta}$ 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 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. 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.

2. 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

$$\varepsilon(\ell) = P\left[\ \sup_{s,t} \frac{|x(s) - x(t)|}{|t-s|^{\alpha}} \leq \ell\ \right]$$

and showing that $\varepsilon(\ell) \to 1$ 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

$$\varepsilon(\ell) = P\left[\ \int_0^1 |x(t\,,\omega)|^2 dt \leq \ell\ \right]$$

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

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

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

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

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

Theorem 2.1 . *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*

$$\int_0^1\int_0^1 \Psi\left(\frac{|f(t)-f(s)|}{p(|t-s|)}\right)ds\,dt = B < \infty. \tag{2.1}$$

Then

$$|f(0)-f(1)| \le 8\int_0^1 \Psi^{-1}(\frac{4B}{u^2})dp(u) \tag{2.2}$$

The double integral (2.1) has a singularity on the diagonal and its finiteness depends on f, p and Ψ. The integral in (2.2) 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 2.1 and illustrate its uses by 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 2.1.

Corollary 2.2 . 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

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

For $0 \le T_1 < T_2 \le 1$ because $B_{T_1,T_2} \le B_{0,1} = B$, we can conclude from (2.1), that the modulus of continuity $\omega_f(\delta)$ satisfies

$$\omega_f(\delta) = \sup_{\substack{0\le t\le 1\\ 0\le s\le 1\\ |t-s|\le\delta}} |f(t)-f(s)| \le 8\int_0^\delta \Psi^{-1}(\frac{4B}{u^2})dp(u) \tag{2.3}$$

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

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

and

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

In particular (2.3) is now an immediate consequence.

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

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

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 2.1 with $\Psi(x) = x^\beta$, and $p(u) = u^{\frac{\gamma}{\beta}}$. Then

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

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 Tchebechev's inequality

$$P[\ B(\omega) \ge B\ ] \le \frac{C\,C_\delta}{B}.$$

On the other hand

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

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 $\frac{1}{2}$.

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. Show that

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

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: 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

$$\omega_x(\delta) \le 8\int_0^\delta \sqrt{\frac{1}{\alpha}\log\left(1+\frac{4B}{u^2}\right)} \frac{1}{2\sqrt{u}} du$$

that produces eventually a statement

$$P\left[\limsup_{\delta\to 0} \frac{\omega_x(\delta)}{\sqrt{\delta \log \frac{1}{\delta}}} \le 16 \right] = 1$$

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

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

Remark: It turns out that according to a result of Paul Levy

$$P\left[\limsup_{\delta\to 0} \frac{\omega_x(\delta)}{\sqrt{\delta \log \frac{1}{\delta}}} = \sqrt{2} \right] = 1$$

Proof of Theorem 2.1. 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

$$|f(0)-f(t_0)| \le 4\int_0^1 \Psi^{-1}\left(\frac{4B}{u^2}\right) dp(u) \tag{2.4}$$

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 (2.2). To prove (2.4) 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

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

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

Then,

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

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

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

and we are done.

3. Convergence of random walks to 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

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

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

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

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 3.1 . *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, ξ_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

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

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

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 3.2 . *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 3.3 which asserts that for any $\varepsilon > 0$, there is a compact set K_ε such that $P_n[\mathbf{X} - K_\varepsilon] \le \varepsilon$. 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 $\sup_{\omega\in K_\varepsilon} |F(\omega) - f(\omega)| \le \varepsilon$. Moreover we can assume without loss of generality that f is also bounded by 1. We can estimate

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

as well as

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

Therefore

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

and we are done.

Remark: We shall prove Lemma 3.3 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 3.3 . *Let P_n, P be as before. Assume that the random variables X_i have a finite moment of order four. Then for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \mathbf{X}$ such that*

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

for all n and

$$P[K_\varepsilon] \ge 1 - \varepsilon$$

as well.

Proof: The set

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

is a compact subset of $\mathbf{X}$ for each fixed B and α. Theorem 2.1 gives us an estimate on $Q[K_{B,\alpha}^c$ which can be made small by taking B large enough. We need to check that (1.2) holds for P_n with some constants β, α and C that do not depend on n. It clearly holds for the Brownian motion P.

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

$$E[(X_1 + X_2 + \cdots + X_k)^4] = kE[X_1^4] + 3k(k-1)\big[E[X_1^2]\big]^2 \le C_1 k + C_2 k^2 \tag{3.1}$$

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

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

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

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

which implies that

$$E\left[\ |x(t) - x(s)|^4\ \right] \leq |t-s|^4\, E[M^4] \leq C|t-s|^2 \tag{3.2}$$

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 3.2 for the end pieces that are increments over incomplete intervals and the estimate 3.1 for the piece $|x(t') - x(s')|$

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

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

Exercise : 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) are independent identically distributed Y_i has mean 0, variance 1 and a finite fourth moment. Z_i has 0 mean and arbitrarily small variance σ^2. We have $X_n(t) = Y_n(t) + Z_n(t)$ and by Kolmogorov"s inequality

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

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

4. Brownian Motion As a Martingale.

If P is the Wiener measure on $(\Omega = C[0,T], \mathcal{B})$ and $\mathcal{B}_t$ is the σ-field generated by $x(s)$ for $0 \le s \le t$, then $x(t)$ is a martingale with respect to $(\Omega, \mathcal{B}_t, P)$, i.e for each $t > s$ in $[0,T]$

$$E^P\{x(t)|\mathcal{F}_s\} = x(s) \quad \text{a.e.} \quad P \tag{4.1}$$

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

$$E^P\{x(t)^2 - t\,|\mathcal{F}_s\} = x(s)^2 - s \quad \text{a.e.} \quad P \tag{4.2}$$

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 4.1 *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

$$X_\lambda(t) = \exp\left[\ \lambda x(t) - \frac{\lambda^2}{2} t\ \right] \tag{4.3}$$

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

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

which shows that we have a Gaussian Process with independent increments with two matching moments. The proof of (4.3) is more or less the same as proving the central limit theorem. In order to prove (4.3) we assume with out loss of generality that $s = 0$ and will show that

$$E^P\left[\ \exp\left[\ \lambda x(t) - \frac{\lambda^2}{2} t\ \right] |\mathcal{B}_s\ \right] = 1 \tag{4.4}$$

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

$$\tau_{k+1,\varepsilon} = \min\left[\ \inf\left\{s : s \ge \tau_k, |x(s) - x(\tau_{k,\varepsilon})| \ge \varepsilon\right\}, t\,, \tau_{k,\varepsilon} + \varepsilon\ \right]$$

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

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

and

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

To establish (4.4) we calculate the quantity on the left hand side as

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

and show that it equals 1. Let us cosider the σ-field $\mathcal{F}_k = \mathcal{B}_{\tau_{k,\varepsilon}}$ and the quantity

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

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

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

In particular for some constant C depending on λ

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

and by induction

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

Since $\varepsilon > 0$ is arbitrary we prove one half of (4.4). Notice that in any case $\sup_\omega |q_k(\omega) - 1| \leq \varepsilon$. Hence we have the lower bound

$$q_k(\omega) \geq E^P\left[\exp\left[-C\,\varepsilon\,[\tau_{k+1,\varepsilon} - \tau_{k\;\varepsilon}]\;\right] \Big| \mathcal{F}_k \right]$$

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

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

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

$$P\left[\sup_{0\leq s\leq t} \exp\left[\lambda x(s) - \frac{\lambda^2}{2}s \right] \geq A \right] \leq \frac{1}{A}$$

and

$$\begin{aligned}
P\left[\sup_{0\leq s\leq t} x(s) \geq \ell \right] &\leq P\left[\sup_{0\leq s\leq t} [x(s) - \frac{\lambda s}{2}] \geq \ell - \frac{\lambda t}{2} \right] \\
&= P\left[\sup_{0\leq s\leq t} [\lambda x(s) - \frac{\lambda^2 s}{2}] \geq \lambda\ell - \frac{\lambda^2 t}{2} \right] \\
&\leq \exp[-\lambda\ell + \frac{\lambda^2 t}{2}]
\end{aligned}$$

Optimizing over $\lambda > 0$, we obtain

$$P\left[\sup_{0\leq s\leq t} x(s) \geq \ell \right] \leq \exp[-\frac{\ell^2}{2t}]$$

and by symmetry

$$P\left[\sup_{0\leq s\leq t} |x(s)| \geq \ell \right] \leq 2\exp[-\frac{\ell^2}{2t}]$$

The estimate is not too bad because by reflection principle

$$P\left[\sup_{0\leq s\leq t} x(s) \geq \ell \right] = 2\,P\,[\,x(t) \geq \ell\,] = \sqrt{\frac{2}{\pi\,t}} \int_\ell^\infty \exp[-\frac{x^2}{2\,t}]\,dx$$

Exercise: One can use the estimate above to prove the result of Paul Levy

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

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 $a > \sqrt{2}$

$$\sum_n P\left[\ \Delta_{\rho^n}(\omega) \ge a\sqrt{n\rho^n \log\frac{1}{\rho}}\ \right] < \infty \tag{4.5}$$

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

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

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

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

Since $a>\sqrt{2}$ we can pick $a_1>\sqrt{2}$ and $a_2>0$. For $\varepsilon>0$ sufficiently small (4.5) is easily verified.

5. 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 put down a measure P. We have the σ-fields $\mathcal{B}_t = \sigma\{x(s) : 0 \le s \le t\}$ defined for $t \le T$. The total σ-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 there mean and variance are specified as

$$E^P\big[x(t+h) - x(t)|\mathcal{B}_t\big] = h\, b(t,\omega) + o(h) \tag{5.1}$$

and

$$E^P\big[(x(t+h) - x(t))^2|\mathcal{B}_t\big] = h\, a(t,\omega) + o(h) \tag{5.2}$$

as $h \to 0$, where for each $t \ge 0$ $b(t,\omega)$ 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 anvestigate this.

(5.1) and (5.2) are infinitesimal diffrential relations and it is best to state them in integrated forms that are precise mathematical statements.

bf Definition. 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 ω on $([0,t] \times \Omega, \mathcal{B}[0,t] \times \mathcal{B}_t)$ where $\mathcal{B}[0,t]$ is the Borel σ-field on $[0,t]$.

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

1. If $f(t,x)$ is measurable function of t and x, then $f(t, x(t,\omega))$ is progressively meausrable.
2. If $f(t,\omega)$ is either left continuous (or right continuous) as function of t for every ω and if in addition $f(t\, o)$ is $\mathcal{B}_t$ measurable for every t, then f is progressively measurable.
3. There is a sub σ-field $\Sigma = \Sigma_{pm} \subset \mathcal{B}[0,T] \times \mathcal{B}_T)$ such that progressive measurability is just measurability with respect to Σ_{pm}. In particular standard operations performrd on progreesively measurable functions yield progressively measurable functions.

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

We reformulate conditions (5.1) and (5.2) as

$$M_1(t) = x(t) - x(0) - \int_0^t b(s,\omega)ds \tag{5.3}$$

and

$$M_2(t) = [M_1(t)]^2 - \int_0^t a(s\,,\omega)ds \tag{5.4}$$

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 μ 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 μ.

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 ω, then one expects P to be Gaussian, if μ is so.

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

$$Z_\lambda(t) = \exp[\lambda M_1(t) - \frac{\lambda^2}{2}\int_0^t a(s\,,\omega)ds] = \exp[\lambda[x(t) - x(0) - \int_0^t b(\,,s\omega)ds - \frac{\lambda^2}{2}\int_0^t a(s\,,\omega)ds] \tag{5.5}$$

is a martingale with respect to $(\Omega, \mathcal{B}_t, P)$ for every real λ. 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 λ. 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 (5.3)

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

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

6. Defining Diffusions (Continuation).

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 σ-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

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

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

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

Theorem 5.1 . 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 par1). $x(t\,,\omega)$ 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

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

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

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

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 anyone of the above definitions $x(t\,,\omega)$ has an almost surely continuous version satifying

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

for some constant C depending only on the 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:

1) *implies* 2). 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.
2) *implies* 3). The condition that $Z_\lambda(t)$ is a martingale can be rewritten as a whole collection of identities

$$\int_A Z_\lambda(t\,,\omega)dP=\int_A Z_\lambda(s\,,\omega)dP \tag{5.6}$$

that is valid for every $t>s$, $A\in\mathcal{F}_s$ and $\lambda\in R^n$. Both sides of (5.6) 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 λ. Since they agree when the values are real, by analyticcontinuation, they must agree for all purely imaginary values of λ as well. This is just 3).
3) *implies* 4). This part of the requires a simple Lemma. Lemma 5.2 . 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 ω a continuous function of bounded variation in t. Assume that for every t the random variable $\xi(t\,,\omega)=\sup_{0\leq s\leq t}|M(t)|\,Var_{[0,t]}\,A(t\,,\omega)$ 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 $s<t$,

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

We can subdivide $[s,t]$ 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 $M(t_{j-1})A(t_{j-1})$

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 $\varepsilon_\lambda(x)$. 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 ε_λ 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.

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

where

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

The two integrals are identical, just the roles of u and v have been interchanged.

5) *implies* 4). This is trivial because after all in 5) we are allowed to taks f to be purely a function of x.

5) *implies* 6). 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

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

6) *implies* 5). This just again reversing the steps.

6) *implies* 7). The problem here that the function $<\lambda\,,x>$ are unbounded. If we pick a function $h(x)$ of one variable to equal x in the interval $[-1.1]$ then level 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 converge to x, $|h_k(x)| \le |x|$ 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 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

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

If we now use the bound on ψ it is easy to obtain the estimate

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

This provides the necessary uniform integrability to conclude that in the limt we have a martingale. Once we have the estimate 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.

7. Stochastic Integrals.

If $y_1, \cdots, y_n$ is a martingale relative to 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

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

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 5.3 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

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

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 $\varepsilon(s\,,\omega)$ 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 $t > 0$, by the martingale property ♣

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

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

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

and clearly

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

is an $(\Omega,\mathcal{F}_t,P)$ martingale. medskipStep 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 $t>0$,

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

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

Exercise: 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. 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. 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. 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

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

8. 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

$$\beta^2(t) - \beta^2(0) = 2\int_0^t \beta(s)\,d\beta(s). \tag{8.1}$$

We have already defined integrals of the form

$$\int_0^t \beta(s)\,d\beta(s) \tag{8.2}$$

as Ito's stochastic integrals. But still a formula of the type (8.1) 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 (8.2) 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 $b(\cdot)$ 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 (8.2) 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 (8.2) 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 (8.1) 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

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

leading to Ito's formula

$$f(\beta(t)) - f(\beta(0)) = \int_0^t f'(\beta(s))d\beta(s) + \frac{1}{2}\int_0^t f''(\beta(s))ds \tag{8.3}$$

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 (8.2). We should be able to verify (8.3) 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 8.1 . 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 $[0\,,\infty) \times R^d$

$$\begin{aligned} u(t\,,x(t)) - u(0\,,x(0)) = \int_0^s u_s(s\,,x(s))ds + \int_0^t < (\nabla u)(s\,,x(s)\,,dx(s) > \\ + \frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\frac{\partial^2 u}{\partial x_i \partial x_j}(s\,,x(s))ds \end{aligned} \tag{8.4}$$

Proof: Let us define the stochastic process

$$\begin{aligned} \xi(t) = u(t\,,x(t)) - u(0\,,x(0)) - \int_0^s u_s(s\,,x(s))ds - \int_0^t < (\nabla u)(s\,,x(s)\,,dx(s) > \\ - \frac{1}{2}\int_0^t \sum_{i,j} a_{i,j}(s\,,\omega)\frac{\partial^2 u}{\partial x_i \partial x_j}(s\,,x(s))ds \end{aligned} \tag{8.5}$$

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 actuation 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

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

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 (8.5) 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 $[B(s\,,\omega), A(\sigma, \,,\omega)]$, 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

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

Now all we are left with is the following

Lemma 8.2 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

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

Exercise: Ito's formula is a local formula that is valid for almost all pathe. If u is a smooth function i.e. with one continuous t derivative and two continuous x derivatives (8.4) 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 8.3 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

$$\int_R \phi(y)\ell(t\,,y\,,\omega)\,dy = \int_0^t \phi(\beta(s))ds \tag{8.6}$$

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

Remark: One can estimate

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

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

9. Diffusions as Stocahstic Integrals.

If $(\Omega, \mathcal{F}_t, P)$ is a probability space and $\beta(\cdot)$ is a d dimensional Brownian Motion relative to it, i.e. $\beta(t)$ is a diffusion with parameters $[0, I]$ relative to $(\Omega, \mathcal{F}_t, P)$, a stochastic integral $x(t)$ of the form

$$x(t) = \int_0^t b(s, \omega) ds + \int_0^t \sigma(s, \omega) d\beta(s) \tag{9.1}$$

is a diffusion with parameters $[b, \sigma\sigma^*]$. We want to show that the converse is true. Given a diffusion $x(t)$ on some $(\Omega, \mathcal{F}_t, P)$ corresponding to $[b, a]$ and given a progressively measurable σ such that $a = \sigma\sigma^*]$, we want to show the existence of a Brownian motion $\beta(\cdot)$ on $(\Omega, \mathcal{F}_t, P)$ such that (9.1) holds. First let us remark that the converse as stated need not be true. For example if $(\Omega, \mathcal{F}_t, P)$ consists of a single point, P is the measure with mass 1 at that point, $x(t, \omega) \equiv 0$ definitely qualifies for a process corresponding to $[0, 0]$. No matter what Brownian Motion we take clearly

$$x(t) - x(0) = 0 = \int_0^t 0 \, d\beta(s) \tag{9.2}$$

so the proposition must be trivially true. Except the space is too small to support anything that is random, let alone a Brownian Motion. If we really need a Brownian Motion we have to borrow it. The way we borrow is to take a standar model of the Brownian Motion $(X, \mathcal{B}_t, Q)$ and take its product with $(\Omega, \mathcal{F}_t, P)$ as our new space. All the old previous processes are still there and replacing $\mathcal{F}_t$ with $\mathcal{F}_t \times \mathcal{B}_t$ does not destroy any of the previous martingale properties. But now we possess an extra Brownian Motion independent of everything. With the new borrowed Brownian Motion (9.2) is clearly true. One has to be careful with this sort of thing. We can only use such a totally arbitrary Brownian Motion when it does not matter what we use.

Let us describe the proof in different cases. First we assume that $a(s, \omega)$ is invertible almost surely and that σ is a square matrix with $\sigma\sigma^* = a$. Let us define

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

and

$$z(t) = \int_0^t \sigma^{-1}(s, \omega) \, dy(s)$$

One can check that $z(\cdot)$ is well defined and has parameters $[0, I]$. This involves the calculation $\sigma^{-1}a\sigma = \sigma^{-1}\sigma\sigma^*\sigma^{-1*} = I$. $z(\cdot)$ is therefore Brownian Motion and

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

The next situation is when σ is a square matrix with $\sigma = \sqrt{a}$, with perhaps a singular somewhere. We now have to borrow a Brownian Motion and ssume we have done it. Let $\pi(s\,,\omega)$ be the orthogonal projection on to the range of $a(s\,,\omega)$. The range of $\sigma(s\,,\omega)$ is the same as that of $a(\sigma\,,\omega)$ and we can construct the inverse $\tau(s\,,\omega)$ such that $\sigma(s\,,\omega)\tau(s\,,\omega) = \tau(s\,,\omega)\sigma(s\,,\omega) = \pi(s\,,\omega)$. We define $y(\cdot)$ as before. But we define $z(t)$ by

$$z(t) = \int_0^t \tau(s\,,\omega)\,dy(s) + \int_0^t [I - \pi(s\,,\omega)]d\beta(s)$$

where β is the borrowed d dimensional Brownian motion. It is only sparingly used. We note that $\sigma(s\,,\omega), \tau(s\,,\omega), \pi(s\,,\omega)$ and $[I - \pi(s\,,\omega)]$ are all symmetric.

$$\begin{aligned}
&\tau(s\,,\omega)a(s\,,\omega)\tau(s\,,\omega) + [I - \pi(s\,,\omega)][I - \pi(s\,,\omega)]\\
&\qquad\qquad = \tau(s\,,\omega)\sigma(s\,,\omega)\sigma(s\,,\omega)\tau(s\,,\omega) + [I - \pi(s\,,\omega)][I - \pi(s\,,\omega)]\\
&\qquad\qquad = \pi(s\,,\omega)\pi(s\,,\omega) + [1 - \pi(s\,,\omega)][1 - \pi(s\,,\omega)] = I
\end{aligned}$$

So z is again Brownian Motion. We can now see that

$$x(t) = \int_0^t b(s\,,\omega)ds + \int_0^t \sigma(s\,,\omega)\,dz(s)$$

We need to show that

$$\int_0^t [\tau(s\,,\omega)dz(s) - I\,dy(s)] = 0$$

A mean square calculation leads to showing

$$(\tau(s\,,\omega)\sigma(s\,,\omega) - I)a(s\,,\omega)(\tau(s\,,\omega)\sigma(s\,,\omega) - I) + \tau(s\,,\omega)(I - \pi(s\,,\omega))\tau(s\,,\omega) = 0$$

which is identically true.

We can do the same thing when $\sigma(s\,,\omega)$ is given as an $n \times k$ matrix with $\sigma\sigma^* = a$ we now have to borrow a k dimensional Brownian Motion. We define a $k \times n$ matrix $\tau(s\,,\omega)$ by $\tau(s\,,\omega) = \sigma^*(s\,,\omega)a^{-1}(s\,,\omega)$, where $a^{-1}(s\,,\omega)$ is such that $a(s\,,\omega)a^{-1}(s\,,\omega) =$

$a^{-1}(s\,,\omega)a(s\,,\omega)=\pi(s\,,\omega)$ with $\pi(s\,,\omega)$ as before. We denote by $\pi^*(s\,,\omega)$ the orthogonal projection on to the range of $\sigma^*(s\,,\omega)$. Now

$$z(t)=\int_0^t \tau(s\,,\omega)dy(s)+\int_0^t (I-\pi^*(s\,,\omega))d\beta(s)$$

The rest of the calculations go as follows.

$$\tau a\tau^*+(I-\pi)=\sigma^*a^{-1}aa^{-1}\sigma+(I-\pi^*)=\sigma^*a^{-1}\sigma+(I-\pi^*)=I$$

and

$$[I-\sigma\tau]a[I-\tau^*\sigma^*]+\tau[I-\pi^*]\tau^*=0$$

10. Diffusions as Markov Processes.

We will be intersted in defining Measures on the space $\Omega=C[[0,T];R^d]$ with the property that for some given $x_0\in R^d$

$$P[x(0)=x_0]=1 \tag{10.1}$$

and

$$Z_f(t\,,\omega)=f(x(t))-f(x(0))-\int_0^t (\mathcal{L}_s f)(x(s))ds \tag{10.2}$$

is a martingale with respect to $(\Omega\,,\mathcal{F}_t\,,P)$ for all smooth functions f, where

$$(\mathcal{L}_s f)(x)=\frac{1}{2}\sum_{i,j}a_{i,j}(s\,,x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x)+\sum_i b_i(s\,,x)\frac{\partial f}{\partial x_i}(x) \tag{10.3}$$

The data are the starting point x_0 and the coefficients $a=\{a_{i,j}(s\,,x)\}$ and $b=\{b_i(s\,,x)\}$. We are seeking a solution P defined by these properties. In the earlier notation $a(s\,,\omega)=a(s\,,x(s\,,\omega))$ and $b(s\,,\omega)=b(s\,,x(s\,,\omega))$. Instead of starting at time 0 we could start at a different time s_0 from the point x_0 and then we would be seeking P, as a measure on the space $\Omega=C[[s_0,T];R^d]$ with analogous properties. We expect to show that under reasonable hypothese on the coefficients a and b, for each s_0,x_0, $P_{s_0\,,x_0}$ exists and is unique. The solutions $\{P_{s_0\,,x_0}\}$ will all be Markov Processes with continuous paths on R^d with transition probabilities

$$p(s\,,x\,,t\,,A)=P_{s,x}[x(t)\in A]$$

satisfying the Chapman-Kolmogorov equations

$$p(s\,,x\,,u\,,A)=\int p(s\,,x\,,t\,,dy)p(t\,,y\,,u\,,A)$$

for $s_0 \le s < t < u \le T$. Moreover

$$P_{s,x}[x(t_1)\in A_1,\cdots,x(t_n)\in A_n]=\int_{A_1}\cdots\int_{A_n}p(s,x,t_1,dy_1)\cdots p(t_{n-1},y_{n-1},t_n,dy_n) \tag{10.4}$$

Our goal is to find reasonably general conditions that guarantee the existence and uniqueness. We will then study properties of the solution P_{s_0,x_0} and how they are related to the properties of the coefficients.

Martingales and Conditioning. Given $\mathcal{F}_s$ for some $s \in [0,T]$, we have the regular conditional probability distribution $Q_{s\,,\omega} = P|\mathcal{F}_t$ which has the following properties. For each ω, $Q_{s\,,\omega}$ is a probability measure on Ω and for each A, $Q_{s\,,\omega}(A)$ is $\mathcal{F}_s$ measurable. Moreover

$$Q_{s\,,\omega}[\omega' : x(t\,,\omega') = x(t\,,\omega) \text{ for } 0 \le t \le s] = 1$$

and

$$P(A)=\int Q_{s\,,\omega}(A)dP(\omega)$$

for all A.Such a Q exists and is essentially unique, i.e any two versions agree for almost all ω w.r.t P.

Lemma 10.1 If $M(t)$ is a martingale relative to $(\Omega\,,\mathcal{F}_t\,,P)$ then for almost all ω and times $t\in[0,T]$, $M(t)$ is a martingale with respect to $(\Omega\,,\mathcal{F}_t\,,Q_{s,\omega})$.

Proof: We need to check that if $A \in \mathcal{F}_{t_1}$, and $t_2 > t_1 \ge s$,

$$\int_A M(t_2)dQ_{s,\omega} = \int_A M(t_1)dQ_{s,\omega}$$

for almost all ω. Since both sides are $\mathcal{F}_s$ measurable it suffices to check that

$$\int_B \left[\int_A M(t_2)dQ_{s,\omega}\right]dP = \int_B \left[\int_A M(t_1)dQ_{s,\omega}\right]dP$$

From the properties of rcpd this reduces to

$$\int \left[\int_{A\cap B} M(t_2)dQ_{s,\omega}\right]dP = \int \left[\int_{A\cap B} M(t_1)dQ_{s,\omega}\right]dP$$

or

$$\int_{A\cap B} M(t_2)dP = \int_{A\cap B} M(t_1)dP$$

Since $A\cap B\in\mathcal{F}_{t_1}$ this follows from the martingale property of $M(t)$. Now all that is left is some technical stuff involving sets of measure 0. As of now the null set depends on $t_1\,, t_2$ and the set A. We take a countable set of rationals for $t_1\,, t_2$ and a countable set of $A's$ that generate the σ-field. One null set works for these. Any thing that works for these works for every thing by the usual bag of tricks. If we have a family $M_\alpha(t)$ of martingales indexed by α then the null set now may depend on α. If we can find a countable set of α's such that the corresponding set of $M_a(t)$'s can approximate every $M_\alpha(t)$, we can get a single null set to work for all the martingales. The family $Z_f(t)$ indexed by smooth functions f is clearly such a family.

Conditioning and Stopping Times. Let τ be a stopping time relative to the family $\mathcal{F}_t$ of σ-fields. We can apply the same reasoning to infer that for any Martingale $M(t)$ the statement, that it remains a martingale with respect to the r.c.p.d. $Q_{\tau\,,\omega}$ of P given $\mathcal{F}_\tau$ for times $t\ge\tau(\omega)$, is valid for almost all ω w.r.t. P.

Proof: The proof again requires the verification for almost all ω of the relation

$$\int_A M(t_2)dQ_{\tau\,,\omega} = \int_A M(t_1)dQ_{\tau\,,\omega}$$

on the set $\{\omega : t_2\ge t_1\ge\tau(\omega)\}$. Given $B\in\mathcal{F}_\tau$ such that $B\subset\{\omega:\tau(\omega)\le t_1\}$ we need to check

$$\int_B\Big[\int_A M(t_2)dQ_{\tau\,,\omega}\Big]dP = \int_B\Big[\int_A M(t_1)dQ_{\tau\,,\omega}\Big]dP$$

Since $B\subset\{\omega:\tau(\omega)\le t_1\}$ and $B\in\mathcal{F}_\tau$ it follows from the definition of $\mathcal{F}_\tau$ that $B\in\mathcal{F}_{t_1}$ and it amounts to verifying

$$\int_{A\cap B} M(t_2)dP = \int_{A\cap B} M(t_1)dP$$

which follows from the facts $A\cap B\in M(t_1)$ and $M(t)$ is a P-martingale. One has to again do a song and dance regarding sets of measure zero. Ultimately, this reduces to the question: is $\mathcal{F}_\tau$ countably generated? The answer is yes, and in fact, it is not hard to prove that

$$\mathcal{F}_\tau = \sigma\left\{x(t\wedge\tau(\omega)) : t\ge 0\right\}$$

which is left as an exercise.

Let us suppose that we are given some coefficients $a(t\,,x)$ and $b(t\,,x)$. For each (s,x) we can define the class $\mathcal{M}_{s_0\,,x_0}$ as the set of solutions to the *Martingale Problem* for $[a,b]$, that start from the intial position x_0 at time x_o. A restatement of the result described earlier is that the r.c.p.d. $Q_{\tau\,,\omega}$ of $P|\mathcal{F}_\tau$ is again in the class $\mathcal{M}_{s\,,\tau}$. In particular if there is a unique solution $P_{s\,,\tau}$ to the martingale problem, then the r.c.p.d. $Q_{\tau\,,\omega} = P_{\tau,x(\tau)}$. This implies that once we have proved uniqueness, the solutions are all necessarily Markov and in fact strong Markov.

11. An easy example.

In R^d let us take $a(t\,,x) = I$ and for $I, b(t\,,x)$ let us try to construct a solution to the martingale problem starting at (s_0, x_0). For simplicity let us assume that $b(t\,,x)$ is bounded uniformly. We can check that the expression

$$R_t(\omega) = \exp\Big[\int_{s_0}^t < b(s\,,x(s)\,,dx(s) > -\frac{1}{2}\int_{s_0}^t \|b(s\,,x(s))\|^2 ds\Big]$$

is a martingale with repect to $(\Omega\,,\mathcal{F}_t^{s_0}\,,Q_{s_0,x_0})$, where Q_{s_0,x_0} is the d-dimensional Brownian motion starting from x_0 at time 0. The same is true of

$$R_{\theta\,,t}(\omega) = \exp\Big[\int_{s_0}^t < \theta + b(s\,,x(s)\,,dx(s) > -\frac{1}{2}\int_{s_0}^t \|\theta + b(s\,,x(s))\|^2 ds\Big]$$

for every $\theta \in R^d$. We can write

$$R_{\theta\,,t}(\omega) = R_t(\omega) Z_{\theta\,,t}(\omega)$$

where

$$Z_{\theta\,,t}(\omega) = \exp\Big[\int_{s_0}^t < \theta\,, dx(s) > -\int_{s_0}^t < \theta\,, b(s\,,x(s)) > -\frac{1}{2}\int_{s_0}^t \|b(s\,,x(s))\|^2 ds\Big]$$

We can define a measure P_{s_0,x_0} such that for $t \geq s_0$

$$\frac{dP_{s_0,x_0}}{dQ_{s_0,x_0}}\Big|_{\mathcal{F}_t^{s_0}} = R_t(\omega)$$

Then clearly $P_{s_0\,,x_0}$ is a solution. Conversely if P is any solution, Q defined by

$$dQ = [R_t(\omega)]^{-1} dP \quad \text{on} \quad \mathcal{F}_t^{s_0}$$

is a solution for $[I\,,0]$ and is therefore the unique Brownian motion Q_{s_0,x_0}. Therefore $P = P_{s_0\,,x_0}$ defined above. So in this case we do have existence and uniqueness.

A second alternative is to try to solve the equation

$$y(t) = x(t) + \int_{s_0}^{t} b(s\,, y(s))ds$$

for $t \geq s_0$ and for Brownian paths $x(t)$ that start from x_0 at time s_0. If we can prove existence and uniqueness, the solution will define a process which solves the martingale problem. It is defined on perhaps a larger space but it is easy enough to map the Wiener measure through $y(\cdot)$ and the transformed measure is a candidate for the solution of the martingale problem. Since we have uniquenes, this must coincide with the earlier construction. If b satisfies a Lipshitz condition in x this can be carried out essentially by Picard iteration.

A third alternative is to try to solve the PDE

$$\frac{\partial u}{\partial s} + \frac{1}{2}\Delta + \sum_i b_i(s\,,x)\frac{\partial u}{\partial x_i} = 0 \tag{11.1}$$

for $s \leq t$ with the boundary condition $u(t\,,x) = f(x)$. The fundamental solution $p(s\,,x\,,t\,,y)$ can be used as transition probabilities to construct a Markov Process which again is our old P. To see this, we verify that if u is any solution of (11.1), then $u(t\,,x(t))$ is a martingale with repect to any $P_{s_0\,,x_0}$ and therefore

$$u(s_0\,,x_0) = \int f(y)p(s_0\,,x_0\,,t\,,y)dy = E^{P_{s_0\,,x_0}}\big[f(x(t)\big]$$

Since this is true for any f the fundamental solution is the same as the transition probability of the alraedy constructed Markov Process.

12. Ito's Theory of Stochastic Differential Equations.

Our goal in this lecture is to construct Markov Processes that are Diffusions in R^d corresponding to specified coefficients $a(t,x) = \{a_{i,j}(t,x)\}$ and $b(t,x) = \{b_i(t,x)\}$. Ito's method consists of starting from any $(\Omega, \mathcal{F}_t, P)$ and an adapted Brownian Motion $\beta(t,\omega) = \{\beta_i(t,\omega)\}$ relative to $(\Omega, \mathcal{F}_t, P)$, with values in R^d. That is to say β has almost surely continuous paths and

$$\exp\left[< \theta, \beta(t) > - \frac{t\|\theta\|^2}{2} \right]$$

is a martingale with respect to $(\Omega, \mathcal{F}_t, P)$ for all $\theta \in R^d$.

The basic assumption on a and b are the following.

H1. The symmetric positive semidefinite matrix $a(t,x)$ can be written as $a(t,x) = \sigma(t,x)\sigma^*(t,x)$ for some matrix $\sigma(t,x)$ that satisfies a Lipschitz condition in x.

$$\|\sigma(t,x) - \sigma(t,y)\| \le A|x-y|$$

H2. The coefficients $b_i(t,x)$ satisfy a similar condition.

$$\|b(t,x) - b(t,y)\| \le A|x-y|$$

H3. Growth conditions. For simplicity we will assume that for some constant C

$$\|\sigma(t,x)\| \le C \qquad \text{and} \qquad \|b(t,x)\| \le C$$

Note that the choice of σ is not unique. We only assume that there is a choice of σ that satisfies the Lipschitz condition. The bounds of course are really bounds on a.

Theorem 12.1 Given $s_0 \ge 0$ and an $\mathcal{F}_{s_0}$ measurable, R^d valued square integrable function $\xi_0(\omega)$, there exists an almost surely continuous progressively measurable function $\xi(t) = \xi(t,\omega)$ for $t \ge s_0$ that solves the equation

$$\xi(t) = \xi_0 + \int_{s_0}^t \sigma(s, \xi(s))d\beta(s) + \int_{s_0}^t b(s, \xi(s))ds \tag{12.1}$$

The solution is unique in the class of progressively measurable functions.

Proof: The existence and uniqueness follow very closely the standard Picard's method for constructing solutions to ODE. We define

$$\xi_0(t) \equiv \xi_0 \quad \text{for} \quad t \geq s_0$$

and define successively, for $k \geq 1$,

$$\xi_k(t) = \xi_0 + \int_{s_0}^{t} \sigma(s\,, \xi_{k-1}(s))d\beta(s) + \int_{s_0}^{t} b(s\,, \xi_{k-1}(s))ds \tag{12.2}$$

Let us remark that the iterations are well defined. They generate progressively measurable almost surely continuous functions at each stage and by induction they are well defined. In order to prove the convergence of the iteration scheme we estimate successive differences. Let us assume with out loss of generality that $s_0 = 0$ and pick a time interval $[0, T]$ in which we will prove convergence. Since T is arbitrary that will be enough. If we denote the difference $\xi_k(t) - \xi_{k-1}(t)$ by $\eta_k(t)$, we have

$$\begin{aligned}\eta_{k+1}(t) &= \int_0^t [\sigma(s\,, \xi_k(s)) - \sigma(s\,, \xi_{k-1}(s))]d\beta(s) + \int_0^t [b(s\,, \xi_k(s)) - b(s\,, \xi_{k-1}(s))]ds \\ &= \int_0^t \delta_k(s)db(s) + \int_0^t e_k(s)ds\end{aligned} \tag{12.3}$$

Because of the Lipschitz assumption

$$\|\delta_k(s)\| \leq A\|\eta_k(s)\| \qquad \text{and} \qquad \|e_k(s)\| \leq A\|\eta_k(s)\| \tag{12.4}$$

We can estimate

$$\sup_{0 \leq \tau \leq t} \|\eta_k(\tau)\| \leq \sup_{0 \leq \tau \leq t} \| \int_0^\tau \delta_k(s)db(s)\| + \int_0^t \|e_k(s)\|ds$$

By Doob's inequality for martingales, the property of stochastic integrals and (12.4)

$$\begin{aligned}E\left[\sup_{0 \leq \tau \leq t} \| \int_0^\tau \delta_k(s)db(s)\|^2\right] &\leq C_0 E\left[\ \| \int_0^t \delta_k(s)db(s)\|^2\ \right] \\ &= C_1 \int_0^t E\big[\|\delta_k(s)\|^2\big]ds \\ &\leq A^2 C_1 \int_0^t E\big[\|\eta_k(s)\|^2\big]ds\end{aligned}$$

On the other hand we can also estimate for $t \leq T$,

$$E\left[\,\left(\int_0^t \|e_k(s)\|ds\right)^2\,\right] \leq TE\left[\,\int_0^t \|e_k(s)\|^2 ds\,\right] \leq A^2 T \int_0^t E\Big[\|\eta_k(s)\|^2\Big]ds$$

Putting the two pieces together, if we denote by

$$\Delta_k(t) = E\left[\,\sup_{0\leq\tau\leq t} \|\eta_k(\tau)\|^2\,\right]$$

then, with $C_T = A^2 C_1(1+T)$,

$$\Delta_k(t) \leq C_T \int_0^t \Delta_{k-1}(s)ds$$

Clearly

$$\eta_1(t) = \int_{s_0}^t \sigma(s\,,\xi_0)d\beta(s) + \int_{s_0}^t b(s\,,\xi_0)ds$$

and

$$\Delta_1(t) \leq C_T\, t$$

By induction

$$\Delta_k(t) \leq \frac{C_T^k\, t^k}{k!}$$

From the convergence of $\sum_k [\frac{C_T^k\, T^k}{k!}]^{\frac{1}{2}}$ we conclude that

$$\sum_k E\big[\sup_{0\leq t\leq T} \|\eta_k(t)\|\big] < \infty.$$

By Fubini's theorem

$$\sum_k \sup_{0\leq t\leq T} \|\eta_k(t)\| < \infty \qquad \text{a.e.} \quad P.$$

In other words for almost all ω with respect to P,

$$\lim_{k\to\infty} \xi_k(t) = \xi(t)$$

exists uniformly in any finite time interval $[0,T]$. The limit $\xi(t)$ is easily seen to be progressively measurable solution of (12.1).

Uniqueness is a slight variation of the same method. If we have two solutions $\xi(t)$ and $\xi'(t)$, their difference $\eta(t)$ satisfies

$$\begin{aligned}\eta_(t) &= \int_0^t [\sigma(s\,,\xi(s)) - \sigma(s\,,\xi'(s))]d\beta(s) + \int_0^t [b(s\,,\xi(s)) - b(s\,,\xi'(s))]ds \\ &= \int_0^t \delta(s)db(s) + \int_0^t e(s)ds\end{aligned}$$

with

$$\|\delta(s)\| \le A\|\eta(s)\| \qquad \text{and} \qquad \|e(s)\| \le A\|\eta(s)\|$$

Just as in the proof of convergence, for the quantity

$$\Delta(t) = E\left[\sup_{0\le s\le t} \|\eta(s)\|^2\right]$$

we can now obtain

$$\Delta(t) \le C_T \int_0^t \Delta(s)ds$$

We have the obvious estimate $\Delta(t) \le C_T$, we get by iteration

$$\Delta(t) \le (C_T)^{k+1}\frac{t^k}{k!}$$

for every k. Therefore $\Delta(t) \equiv 0$ implying uniqueness.

The uniqueness is a special form of uniqueness. If two solutions of (12.1) are constructed on the same same space for the same Brownian motion with the same choice of σ then they are identical for almost all ω. This seems to leave open the possibility that somehow different choices of σ or constructions in different probability spaces could produce different results. That this is not the case is easily established. Before we return to this let us proceed with some comments.

Remark. We can start with a constant x for our initial value at some time s and construct a solution $\xi(t) = \xi(t\,;s\,,x)$ for $t \ge s$. If we define

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

then our solutions are Markov processes with transition probability $p(s\,,x\,,t\,,A)$.

The proof is based on the following argument. Because of uniqueness the solution starting from time 0 can be solved upto time s and then we can start again at time s with the initial value equal to the old solution, and we should not get anything other than the solution obtainable in a single step. In other words

$$\xi(t\,;s\,,\xi(s\,,0,x)) = \xi(t\,;0\,,x)$$

Since the solution $\xi(t\,;s\,,\xi(s\,,0,x))$ only depnds on $\xi(s\,,0,x)$ which is $\mathcal{F}^s$ measurable and increments $d\beta$ of the Brownian paths over $[s,t]$ that are independent of $\mathcal{F}_s$, the conditional distribution

$$\begin{aligned} P[\,\xi(t) \in A|\mathcal{F}_s\,] &= P[\,\xi(t\,;s\,,\xi(s)) \in A|\mathcal{F}_s\,] \\ &= P[\,\xi(t\,;s\,,z) \in A|\mathcal{F}_s\,]_{z=\xi(s)} \\ &= p(s\,,\xi(s)\,,t\,,A) \end{aligned}$$

establishing the Markov property.

Remark. A similar argument will yield the strong Markov property. We use the fact that the after a stopping time τ the future increments of the Brownian motion are still independent of the σ-field $\mathcal{F}_\tau$. There are some details to check about restarting the SDE at a stopping time. But this is left as an exercise.

Remark. If we have two solutions on two different spaces of the same equation with the same constant (i.e. non random) initial value, i.e. with the same σ and b that satisfy our assumptions, then they have the same distributions as stochastic processes. If we notice our construction, each iteration $\xi_k(t)$ was a well defined function of ξ_{k-1} and the Brownian incremets. The iteration scheme is the same in both. At each stage they are identical functions of different Brownian motions. Therefore they have the same distribution. Pass to the limit.

Remark. If $\xi(t)$ is any solution anywhere for any choice $\bar{\sigma}$ of the square root, then ξ is a diffusion corresponding to the coefficients $a = \bar{\sigma}\bar{\sigma}^*$, b and can be represented, by enlarging the space if necessary, as a solution of (12.1) with any arbitrary choice of the square root σ. In particulat if one is available with the Lipschitz property and b is also Lipschitz we are back in the old situation. Therefore if there is a Lipschitz choice available then the distribution of any solution with any choice of the square root is identical to the one coming from the Lipschitz choice. In particular the distribution of any two Lipschitz choices are identical.

13. Some Examples. A Discussion of Uniqueness.

Ornstein-Uhlenbeck Process. A stochastic differential equation of the form

$$dx(t) = \sigma\beta(t) - ax(t)dt\,; \qquad x(0) = x_0 \tag{13.1}$$

has an explicit solution

$$x(t) = e^{-at}x_0 + \sigma e^{-at}\int_0^t e^{as}d\beta(s)$$

which has a Gaussian distribution with mean $e^{-at}x_0$ and variance given by

$$\sigma^2(t) = \sigma^2 e^{-2at}\int_0^t e^{2as}ds = \frac{\sigma^2}{2a}(1 - e^{-2at})$$

This is a Markov Process with stationary Gaussian transition probablity densities:

$$p(t\,,x\,,y) = \frac{1}{\sqrt{2\pi}\,\sigma(t)}\exp\Big[-\frac{(y - e^t x)^2}{2\sigma^2(t)}\Big]$$

This is particularly interesting when $a < 0$, which is the stable case, and then

$$\lim_{t\to\infty}\sigma^2(t) = \theta = \frac{\sigma^2}{2a}$$

and

$$\lim_{t\to\infty}p(t\,,x\,,y) = \frac{1}{\sqrt{2\pi\theta}}\exp\big[-\frac{y^2}{2\theta}\big]$$

Geometric Brownian Motion: The function $x(t) = x_0\exp\big[\sigma\beta(t) + \mu t\big]$ satisfies according to Ito's formula the equation

$$dx(t) = \sigma\,x(t)d\beta(t) + (\mu + \frac{\sigma^2}{2})x(t)dt\,; \qquad x(0) = x_0$$

so that a solution of

$$dx(t) = \sigma\,x(t)d\beta(t) + \mu x(t)dt\,; \qquad x(0) = x_0$$

is provided by

$$x(t) = x_0\exp\big[\sigma\beta(t) + (\mu - \frac{\sigma^2}{2})t\big]$$

Notice the behavior

$$\frac{1}{t}\log x(t) \simeq (\mu - \frac{\sigma^2}{2}) \quad \text{a.e.}$$

as well as

$$\frac{1}{t}\log E[x(t)] \simeq \mu$$

The explanation is that the larger expectation is accounted for by certain very large values with very small probabilities.

ODE and SDE. The solution $x(t) = x_0 \exp\left[\sigma\beta(t) + (\mu - \frac{\sigma^2}{2})t\right]$ of

$$dx(t) = \sigma\, x(t) d\beta(t) + \mu x(t) dt\,; \qquad x(0) = x_0$$

is nice smooth map of Brownian paths and makes sense for all functions f

$$x(t\,,f) = x_0 \exp\left[\sigma f(t) + (\mu - \frac{\sigma^2}{2})t\right]$$

and for smooth functions as well. If we replace β by a smooth path f, it solves

$$dx(t) = \sigma\, x(t) df(t) + (\mu - \frac{\sigma^2}{2}) x(t) dt\,; \qquad x(0) = x_0$$

The Ito map satisfies the wrong equation on smooth paths. This is typical.

There are various ways of constructing a solution that correspond to a Diffusion with coefficients $a(t\,,x) = \{a_{i,j}(t\,,x)\}$ and $b(t\,,x) = \{b_i(t\,,x)\}$. For a square root σ satisfying $\sigma\sigma^* = a$ we can attempt to solve the SDE

$$dx(t) = \sigma(t\,,x(t))d\beta(t) + \beta(t\,,x(t))dt\,; x(0) = x_0$$

on the Wiener space and get a map $\beta(\cdot) \to x(\cdot)$. Such a solution if it exists will be called a **strong solution**. A **Matingale Solution** is a measure P on $\Omega = C[[0\,,\infty)\,; R^d]$ such that $P[x(0) = x_0] = 1$ and for each smooth f the expression

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

is a martingale with respect to $(\Omega\,, \mathcal{F}_t\,, P)$. If we can construct on some probability space $(\Omega\,, \mathcal{F}_t\,, \mu)$ a Brownian motion $\beta(\cdot)$ and an $x(\cdot)$ that satisfy

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

then we call $x(\cdot)$ a **Weak Solution** to the SDE. We make the following remarks.

Remark 1. A strong solution is a weak solution, and if σ is Lipschitz, then any weak solution is a strong solution. In particular two weak solutions on the same space involving the same Brownian Motion are identical.

Remark 2. The distribution P of any Weak Solution is a Martingale Solution and conversely any Martingale Solution is the distribution of some Weak Solution.

Remark 3. For a given square root σ if we deifine the $2d \times 2d$ matrix $\tilde{a}$ as the 2×2 matrix of $d \times d$ blocks $a\,,\sigma$ followed by σ^* and I, and $\tilde{b}$ as b followed by the 0 vector, then a weak solution of $\sigma\,,b$ is the same as a Martingale Solution of $\tilde{a}\,,\tilde{b}$.

Remark 4. Any two weak solutions on different probability spaces can be put on the space space with the same Brownian Motion.

This needs an explanation. What we mean is the following: Let P_1 and P_2 be two martingale solutions for $\tilde{a}\,,\tilde{b}$. Then we can construct a Q which is a martingale solution for the $3d$ dimensional problem with coordinates $x\,,y\,,z$ for $\widehat{a}\,,\widehat{b}$ where in blocks of $d \times d$ the rows of $\widehat{a}$ reads $a(t\,,x)\,,\sigma(t\,,x)\,\sigma^*(t\,,y)\,,\sigma(t\,,x)$; $\sigma(t\,,y)\,\sigma^*(t\,,x)\,,a(t\,,y)\,,\sigma(t\,,y)$ and $\sigma^*(t\,,x)\,,\sigma^*(t\,,y)\,,I$ while $\widehat{b}$ raeds $b(t\,,x)\,,b(t\,,y)\,,0$ which has the following two additional properties:

1. The distribution of x, z coordinates is P_1 and that of the y, z coordinates P_2.
2. Given the z cordinate the x and y coordinates are conditionally independent.

We start with P the Wiener measure, P_i^ω the conditional of 'x' given the Brownian Motion under P_i and define

$$Q = P(d\omega) \otimes [P_1^\omega \times P_2^\omega]$$

i.e. we build in conditional independence. We can check that Q is a Martingale Solution for the $3d$ dimensional problem.

This construction allows us to make the following remark.

Remark 5. If it is true that for some $\sigma\,,b$ any two weak solutions on the same space with the same Brownian Motion are identical, then any weak solution is a strong solution and in such a context the Martingale solution is unique.

14. Random Time Change and Uniqueness in One Dimension.

One of the properties of Martingales is Doob's stopping theorem. If $M(t)$ is a Martingale with respect to $(\Omega\,,\mathcal{F}_t\,,P)$ and $0\le\tau_1\le\tau_2\le C$ are two bounded stopping times, with the correspnding σ-fields $\mathcal{F}_{\tau_1}\subset\mathcal{F}_{\tau_2}$, then

$$E^P\,[\;M(\tau_2)|\mathcal{F}_{\tau_1}\;]=M(\tau_1)\qquad\text{a.e.}$$

In particular if τ_t is a family of bounded stopping times with $\tau_s\le\tau_t$ for $s\le t$, then with

$$N(t)=M(\tau_t)\quad\text{and}\quad\mathcal{G}_t=\mathcal{F}_{\tau_t}$$

$N(t)$ is a martingale with respect to $(\Omega\,,\mathcal{G}_t\,,P)$. If P is any Martingale solution on $\Omega=C[[0\,,\infty)\,,X]$ that corresponds to some L, then

$$f(x(t))-f(x(0))-\int_0^t(Lf)(x(s))ds$$

is a martingale with respect to $(\Omega\,,\mathcal{F}_t\,,P)$. We consider the functions $\{\tau_t\}$, defined by,

$$\int_0^{\tau_t(\omega)}\frac{ds}{V(x(s\,,\omega))}=t$$

where $V(\cdot)$ is a positive measurable function on X, satisfying

$$0<c\le V(x)\le C<\infty\tag{14.1}$$

Then it is clear that τ_t is well defined for $t\ge 0$ with $\tau_0=0$ and $\tau_s<\tau_t$ for $s<t$ and τ_t continuous in t. We can use τ_t to define a map Φ_V of $\Omega\to\Omega$ by

$$(\Phi_V\,\omega)(t)=x(\tau_t(\omega)\,,\omega)$$

Lemma 14.1 For any two functions U and V, satisfying (14.1)

$$\Phi_U\,o\,\Phi_V=\Phi_{UV}$$

Proof: The proof depends on the simple calculation

$$\frac{d\tau_t(\omega)}{dt}=V(x(\tau_t(\omega)))=V(y(t))$$

where $y(t) = x(\tau_t(\omega)) = (\Phi_V\, \omega)(t)$. If σ_t solves

$$\int_0^{\sigma_t(\omega)} \frac{ds}{U(y(s))} = t$$

or

$$\frac{d\tau_{\sigma_t}}{dt} = \frac{d\tau_s}{ds}|_{s=\sigma_t} \cdot \frac{d\sigma_t}{dt} = (VU)(y(\sigma_t)) = (VU)(x(\tau_{\sigma_t}))$$

proving the composition rule. In particular Φ_V is invertible with $\Phi_{\frac{1}{V}} = [\Phi_V]^{-1}$. The σ-field $\sigma\{y(s) : 0 \le s \le t\} \subset \mathcal{F}_{\tau_t}$, and

$$f(y(t)) - f(y(0)) - \int_0^{\tau_t} (\mathcal{L}f)(x(s))ds$$

is an $(\Omega\,, \mathcal{F}_{\tau_t}\,, P)$ martingale. By change of variables

$$\int_0^{\tau_t} (\mathcal{L}f)(x(s))ds = \int_0^t V(y(s))(\mathcal{L}f)(y(s))ds$$

Therefore

$$f(y(t)) - f(y(0)) - \int_0^t V(y(s))(\mathcal{L}f)(y(s))ds$$

is a martingale with respect to $(\Omega\,, \mathcal{F}_\tau\,, P)$. In particular $Q = \Phi_V^{-1} P$ is a Martingale solution for $\tilde{\mathcal{L}}$ defined as

$$(\tilde{\mathcal{L}}f)(x) = V(x)(\mathcal{L}f)(x)$$

The steps are reversible so that existence or uniqueness for a Martingale solution for $\mathcal{L}$ and $\tilde{\mathcal{L}}$ are equivalent so long as V satisfies (14.1).

Now when $d = 1$ we can prove existence and uniqueness of Martingale Solutions to

$$\mathcal{L} = \frac{a(x)}{2} D_x^2 + b(x) D_x$$

so long as $a\,, b$ are bounded measurable with $0 < c \le a(x) \le C < \infty$. From Girsanov Formula we can assume without loss of generality that $b \equiv 0$. By random time change we can assume that $a(x) \equiv 1$. Now we are in the Brownian motion case, and we have existence and uniqueness. Of course once we ave existence and uniqueness the Markov Property as well as the Strong Markov Property follow.

In the time dependent case it is more complicated. In one dimension we can improve the Lipschitz assumption on σ to Holder with exponent $\frac{1}{2}$.

Theorem 14.2 Assume that b is Lipschitz but σ satifies

$$|\sigma(t\,,x)-\sigma(t,\,,y)|\leq|x-y|^{\frac{1}{2}}$$

Then any two solutions

$$x_i(t)=x_0+\int_0^t\sigma(s\,,x_i(s))d\beta(s)+\int_0^t b(x_i(s))ds$$

are identical.

Proof: The proof involves the application of Ito's formula for the function $f(x_1(t)\,,x_2(t))=|x_1(t)-x_2(t)|$. Formally

$$\begin{aligned}df(x_1(t)\,,x_2(t))&=[\mathrm{sig}(x_1(t)-x_2(t))](\sigma(t\,,x_1(t))-\sigma(t\,,x_2(t)))d\beta(t)\\&\quad+\delta(x_1(t)-x_2(t))|\sigma(t\,,x_1(t))-\sigma(t\,,x_2(t))|^2dt\\&\quad+[\mathrm{sig}(x_1(t)-x_2(t))][b(t\,,x_1(t))-b(t\,,x_2(t))]dt\end{aligned}$$

We will give an argument as to why the term with δ is zero. Granting that we have by the Lipschitz condition on b,

$$E[|x_1(t)-x_2(t)|]\leq C\int_0^t E[|x_1(s)-x_2(s)|]ds$$

and this implies uniqueness. Let us approximate $|x|$ by $f_\varepsilon(x)=\sqrt{(\varepsilon^2+x^2)}$. Then

$$f_\varepsilon''(x)=\frac{\varepsilon^2}{(\varepsilon^2+x^2)^{\frac{3}{2}}}$$

and

$$|f_\varepsilon''(x_1-x_2)|\sigma(t\,,x_1)-\sigma(t\,,x_2)|^2\leq\frac{C\varepsilon^2|x_1-x_2|}{(\varepsilon^2+(x_1-x_2)^2)^{\frac{3}{2}}}\leq C\sup_u\left[\frac{u}{(1+u^2)^{\frac{3}{2}}}\right]\leq C'$$

We can let $\varepsilon\to 0$. use the dominated convergence theorem and pass to the limit to show that there is no contribution from the term with δ.

15. General comments on existence and uniqueness of the martingale solutions.

If we are given a $a(t\,,x) = \{a_{i,j}(t\,,x)\}$ and $b(t\,,x) = \{b_j(t\,,x)\}$ and are interested in proving uniqueness of martingale solutions, we specifically wish to show that the set $C_{s,x}$ of probability measures P on $\Omega_s = C[[s\,,\infty)\,;R^d\,]$ such that

$$P[x(s) = x] = 1$$

and

$$Z_f(t) = f(x(t)) - f(x(s)) - \int_s^t (\mathcal{L}_s f)(x(s)ds \tag{15.1}$$

are martingales with respect to $(\Omega_s\,,\mathcal{F}_t^s\,,P)$ for all smooth f consists of exactly one probability measure.

The existence part is simple under fairly general conditions. If a and b are smooth we can have Lipschitz σ and b and Ito's theory of SDE provides us, as we saw, both existence and uniqueness. If we only assume that a and b are just bounded and continuous we can prove existence along the following lines. We take $s = 0$ with out loss of generality and approximate a, b by smoother a_n, b_n that converge as $n \to \infty$ to a, b. The convergence can be assumed to be uniform over compact subsets of R^d, and we can also assume that a_n as well as b_n are uniformly bounded by some constant M. For some x let $P_{n,x}$ be the unique solution starting at time 0 from the point x, corresponing to a_n, b_n. We will prove that $P_{n,x}$ is a totally bounded sequence of probability measures on Ω, and that if P is any weak limit then P is a solution starting at time 0 from x for the limiting coefficients and therefore we have existence.

Lemma 15.1 The sequence P_n satisfies the following. For any $T < \infty$ and any $\varepsilon > 0$, there exists $A(T\,,\varepsilon)$ depending only on the bound M such that

$$P_n\left[\,\omega : \sup_{0\le s\le t} \frac{|x(s) - x(t)|}{|t-s|^{\frac{1}{4}}} \le A(T\,,\varepsilon)\,\right] \ge 1-\varepsilon$$

In particular the sequence is totally boundeded.

Proof: Let P be a diffusion corresponding to some a, b that are bounded by M. We remark that we can write

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

Clearly the difference $|x(t)-y(t)|$ is uniformly Lipschitz with a bound of M for the Lipschitz constant and $y(t)$ is such that

$$\exp\left[<\theta\,, y(t)-y(0)> -\frac{1}{2}\int_0^t <\theta\,, a(s\,,\omega)\theta> ds \right]$$

is a martingale with respect to $(\Omega\,, \mathcal{F}_t\,, P)$. From this we deduce the following bound

$$E^P\,[\,\exp[<\theta\,, y(t)-y(s)>]\,] \leq \exp[\frac{M(t-s)}{2}\|\theta\|^2]$$

or

$$E^P\left[\,\exp[<\theta\,, \frac{y(t)-y(s)}{\sqrt{t-s}}>]\,\right] \leq \exp[\frac{M}{2}\|\theta\|^2].$$

It is easy to conclude now that

$$E^P\left[\,\frac{|y(t)-y(s)|^4}{|t-s|^2}\,\right] \leq CM^2$$

for a universal constant C. From Garsia-Rodemich-Rumsey lemma we get our estimte and by Prohorov's theorem we get the total boundedness of the sequence P_n of probability measures.

We take a weak limit alomg a subsequence and call it P. We might as well assume that $P_n \to P$ weakly.

Lemma 15.2 The limit P is a martingale solution for a, b.

Proof: With $Z_f(t)$ as in (15.1) we need to establish

$$\int_A Z_f(t)dP = \int_A Z_f(s)dP \tag{15.2}$$

for $A \in \mathcal{F}_s$. It is sufficient to prove

$$\int \Phi(\omega)Z_f(t)dP = \int \Phi(\omega)Z_f(s)dP$$

for bounded continuous (in the topology of uniform convergence on bounded time intervals) functions Φ that are $\mathcal{F}_s$ measurable. For such a Φ clearly

$$\int \Phi(\omega)Z_f^n(t)dP_n = \int \Phi(\omega)Z_f^n(s)dP_n \tag{15.3}$$

where

$$Z_f^n(t) = f(x(t)) - f(x(s)) - \int_s^t (\mathcal{L}_s^n f)(x(s)ds \tag{15.4}$$

and

$$\mathcal{L}_s^n = \frac{1}{2}\sum_{i,j} a_{i,j}^n(s\,,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_j b_j^n(s\,,x)\frac{\partial}{\partial x_j}$$

We can let $n \to \infty$ and from the weak convergence of P_n to P, the convergence of $Z_f^n(t)$ to $Z_f(t)$, (uniformly on compact subsets of Ω), and the uniform boundedness of $Z_f^n(t)$ we can let $n \to \infty$ in (15.3) to conclude that (15.2) holds. We are done.

Uniqueness is a much harder issue. Clearly we have it in the Lipschitz case. But the uniqueness cannot be done by approximation. The following general Markovian Principle works. Assume existence.

Lemma 15.3 If there exists a family $\mu_{s\,,x\,,t}(\cdot)$ of probability measures such that, for any $P \in C_{s,x}$,

$$P[x(t) \in A] = \mu_{s\,,x\,,t}(A)$$

then P is a Markov Process with $\mu_{s\,,x\,,t}(A)$ as transition probabilities and is therefore unique.

Proof: We proved a general princilpe that the conditional probabilty distribution $P_{t,\omega}$ of any solution $P \in C_{s,x}$ give $\mathcal{F}_t$ is in $C_{t,x(t)}$ almost surely. Therefore for $s < t < u$

$$P\,[\;x(u) \in A|\mathcal{F}_t\;] = \mu_{t\,,x(t)\,,u}(A)$$

a.e. P, proving the Markov property and the lemma.

Determining $P[x(t) \in A]$ for $P \in C_{s,x}$ can be done through solving certain partial differential equations. We know that

$$u(t\,,x(t)) - u(s\,,x(s)) - \int_s^t (\frac{\partial}{\partial \sigma} + \mathcal{L}_\sigma u)(\sigma\,,x(\sigma))d\sigma$$

is a martingale. Therefore for any smooth u and $P \in C_{s,x}$,

$$u(s\,,x) = E^P\left[\; f(x(t)) + \int_s^t g(\sigma\,,x(\sigma))d\sigma \;\right] \tag{15.5}$$

where

$$g(\sigma\,,\cdot) = -(\frac{\partial u}{\partial \sigma} + \mathcal{L}_\sigma u)(\sigma\,,\cdot) \tag{15.6}$$

and

$$u(t\,,\cdot)\equiv f(\cdot) \tag{15.7}$$

The relation holds for for every smooth u and every $P\in C_{s,x}$.

Lemma 15.4 If u satisfies (15.6) and (15.7) with $g\geq 0$ and $f\geq 0$, and $C_{s,x}$ is nonempty then the maximum principle holds, i.e. $u(s\,,x)\geq 0$.

Proof: Obvious from (15.5).

We are actually interested in going in the reverse direction. Suppose either

1. (15.6) is solvable for sufficiently many g satisfying (15.7) with $f\equiv 0$

or

2. (15.6) is solvable with $g\equiv 0$ satsfying (15.7) for sufficiently many f, then

$$E^P\left[\int_s^t g(\sigma\,,x(\sigma))d\sigma\right]$$

or

$$E^P\left[\ f(x(t))\ \right]$$

are detrmined for sufficiently many g or f as the case may be. This can then be used to determine $P[x(t)\in A]$ for $P\in C_{s,x}$. What we mean by sufficiently many depends on the circumstances. We need either enough g's to recover the the measures $\{\mu_\sigma\}$ from the integrals

$$\int_s^t\int_{R^d} f(y)\mu_\sigma(dy)d\sigma$$

or enough f's to determine the measure from the integrals

$$\int_{R^d} f(y)\mu(dy)$$

If we know some thing about $P\in C_{s,x}$, like for instance

$$\mu_{s\,,x}(d\sigma,dy)=P_{s,x}[x(\sigma)\in dy]d\sigma$$

is always in some $L_p([s\,,t]\times R^d)$, then sufficiently many can be just any dense set in $L_q([s\,,t]\times R^d)$, where $\frac{1}{p}+\frac{1}{q}=1$. Similarly if we know that for any $P\in C_{s,x}$, the measure $P[x(t)\in dy]$ is in $L_p[R^d]$ it is enough to solve f from a dense subset of $L_q[R^d]$. These remarks are quite pertinent especially when the coefficients are discontinuous.

16. Time dependent diffusions in one dimension.

Let us look at $d = 1$ and consider the equation

$$\frac{\partial u}{\partial \sigma} + \frac{1}{2} a(\sigma\,, y) \frac{\partial^2 u}{\partial y^2} = g(\sigma\,, y)$$

If we insist on u being $C^{1,2}$, and g being continuous $u_t\,, u_{yy}$ and g are continuous and unless $u \equiv c$, there will be no solutions. For nonsmooth coefficients we have to deal with non-classical solutions.

Let us now illustrate the method with the problem of constructing solutions for the one dimensional problem with $0 < c \leq a(t\,, x) \leq C < \infty$ and $b \equiv 0$. We start with

Theorem 16.1 Let us consider a stochastic integral with repect the Brownian Motion on some probability space

$$\xi(t) = x_0 + \int_0^t k(s\,, \omega) d\beta(s)$$

for some $k(s\,, \omega)$ satisfying

$$0 < c \leq |k(s\,, \omega)|^2 \leq C < \infty$$

Then there is a constant M depending only on c, C and T such that

$$\left| E\left[\int_0^T g(s\,, \xi(s)) ds \right] \right| \leq M \|g\|_{L_2([0,T] \times R)}$$

Proof: The key estimate is the following: Consider a function g with compact support on $(-\infty\,, \infty) \times R$. Define

$$u(s, x) = \int_s^\infty \int_R \frac{1}{\sqrt{2C\pi(t-s)}} g(t\,, y) \exp[-\frac{(x-y)^2}{2C(t-s)}] dt dy \tag{16.1}$$

If g is smooth it is easy to verify that

$$\frac{\partial u}{\partial s} + \frac{C}{2} u_{xx} = -g(s\,, x)$$

Taking Fourier transform $\widehat{u}$ of u in x and s,

$$i\tau \widehat{u}(\tau\,, \eta) - \frac{C\eta^2}{2} \widehat{u}(\tau\,, \eta) = -\widehat{g}(\tau\,, \eta)$$

or

$$\widehat{u}(\tau\,,\eta) = \frac{1}{\frac{C\eta^2}{2} - i\tau}\widehat{g}(\tau\,,\eta)$$

and

$$\widehat{u_{xx}}(\tau\,,\eta) = \frac{\eta^2}{i\tau - \frac{C}{2}\eta^2}\widehat{g}(\tau\,,\eta)$$

Therefore using the isometry of the Fourier transform

$$\|u_{xx}\|_{L_2} \le \frac{2}{C}\|g\|_{L_2}$$

Now to prove the theorem there is no loss of generality in assuming that k is simple. With uniform bounds we can pass to the limit. We define the linear functional

$$\Lambda(g) = E\left[\int_0^T g(s\,,\xi(s))ds\right]$$

Clearly if k is simple, then ξ is piece wise Brownian Motion and the transition probability $p_{\sigma^2}(s\,,x\,,t\,,y)$ of the Brownian motion is in $L_2[[0\,,T]\times R]$ uniformly in s and x provided $0 < c \le \sigma^2 \le C < \infty$. It is now easy to get a bound

$$|\Lambda(g)| \le M\|g\|_{L-2}$$

with a constant M that depends on the number of intervals over which k is constant. We want to improve our bound to make it depend only on c, C and T. If we take g that vanishes for $t \ge T$, construct u as in (16.1), then $u(T\,,\cdot) \equiv 0$ and by Ito's formula

$$\begin{aligned}
u(0\,,x) &= -E^P\left[\int_0^T u_s + \frac{k^2(s\,,\omega)}{2}u_{xx}(s\,,\xi(s))ds\right]\\
&= -E^P\left[\int_0^T (u_s + \frac{C}{2}u_{xx})(s\,,\xi(s))ds\right] + E^P\left[\int_0^T \frac{C - k^2(s\,,\omega)}{2}u_{xx}(s\,,\xi(s))ds\right]\\
&= E^P\left[\int_0^T g(s\,,\xi(s))ds\right] + E^P\left[\int_0^T \frac{C - k^2(s\,,\omega)}{2}u_{xx}(s\,,\xi(s))ds\right]
\end{aligned}$$

or

$$E^P\left[\int_0^T g(s\,,\xi(s))ds\right] = u(0\,,x) - E^P\left[\int_0^T \frac{C - k^2(s\,,\omega)}{2}u_{xx}(s\,,\xi(s))ds\right]$$

and therefore

$$E^P\left[\int_0^T g(s\,,\xi(s))ds\right] \le |u(0\,,x)| + \frac{C-c}{2}\,E^P\left[\int_0^T |u_{xx}(s\,,\xi(s))|ds\right]$$

In other words

$$|\Lambda(g)| \le |u(0\,,x)| + \Lambda(|u_{xx}|) \tag{16.2}$$

Taking supremum in (16.2) over g with $\|g\|_{L_2} \le 1$, and denoting it by M, we get

$$M \le \sup_{g:\|g\|\le 1} |u(0\,,x)| + \frac{C-c}{2}\frac{2}{C}M = \sup_{g:\|g\|\le 1} |u(0\,,x)| + (1-\frac{c}{C})M$$

Since

$$\sup_{g:\|g\|\le 1} |u(0\,,x)| = \left[\int_0^T \int_R \left[\frac{1}{\sqrt{2\pi Ct}}\exp[-\frac{y^2}{2Ct}]\right]^2 dydt\right]^{\frac{1}{2}} = A(T\,,C) < \infty$$

we have

$$M \le \frac{C}{c}A(T\,,C) = A(T\,,C\,,c)$$

and the theorem is proved.

Remark: An immediate consequence of the estimate is that any stochastic integral of the form

$$\xi(t) = \int_0^t k(s\,,\omega)d\beta(s)$$

with

$$0 < c \le |k(s\,,\omega)|^2 \le C < \infty$$

has a distribution $q(t\,,dy)$ that has a density $q(t\,,y)dy$ in y for almost all t, with the bound

$$\int_0^T \int_R |q(t\,,y)|^2 dtdy \le [A(T\,,C\,,c)]^2$$

Remark: In particular if $p(s\,,x\,,t\,,dy)$ is the transition probability for a diffusion with smooth coefficients $a = a(t\,,x)$, and $b = 0$, with

$$0 < c \le a(t\,,x) \le C < \infty$$

it has a density $p(s\,,x\,,t\,,y)$ for almost all t and,

$$\sup_{\substack{x\\ s\le t}}\int_t^T |p(s\,,x\,,t\,,y)|^2 dtdy \le [A(T-t\,,C\,,c)]^2$$

We will now use the above theorem to prove existence as well as uniqueness of martingale solutions. We assume $0 < c \le a(t\,,x) \le C < \infty$. We can construct a_n satisfying the same bounds that are smooth and we can have $a_n \to a$ almost everywhere in t and x. We have from Lemma 15.1 the total boundeness of the measures P_n for the approximating smooth coefficients. But now the expressions

$$Z_f^n(t) = f(x(t)) - f(x(0)) - \int_0^t \frac{a_n(s\,,x(s))}{2} f_{xx}(x(s))ds$$

do not converge uniformly on compact subsets of Ω to

$$Z_f(t) = f(x(t)) - f(x(0)) - \frac{1}{2}\int_0^t a(s\,,x(s)) f_{xx}(x(s))ds$$

But, given any $\varepsilon > 0$, we can find a_n^ε and a^ε such that $a_n^\varepsilon \to a^\varepsilon$ uniformly on compact subsets of $[0\,,\infty) \times R$ and

$$\int_0^T \int_{|x|\le \ell} [|a_n^\varepsilon - a_n|^2 + |a^\varepsilon - a|^2] dx\, dt \le \delta_\varepsilon(T\,,\ell)$$

for some $\delta_\varepsilon(T\,,\ell)$ such that $\delta_\varepsilon(T\,,\ell) \to 0$ as $\varepsilon \to 0$ for each T and ℓ. Now

$$Z_f^{n,\varepsilon}(t) = f(x(t)) - f(x(0)) - \frac{1}{2}\int_0^t a_n^\varepsilon(s\,,x(s)) f_{xx}(x(s))ds$$

converges nicely to

$$Z_f^\varepsilon(t) = f(x(t)) - f(x(0)) - \frac{1}{2}\int_0^t a^\varepsilon(s\,,x(s)) f_{xx}(x(s))ds$$

and

$$\int \Phi(\omega) Z_f^{n,\varepsilon}(t) dP_n \to \int \Phi(\omega) Z_f^\varepsilon(t) dP_n$$

for for smooth f and bounded continuous $\mathcal{F}_s$ measurable functons Φ. Since we now have a bound of the form

$$
\begin{aligned}
&\sup_n \left| \int \Phi(\omega)[Z_f^{n,\varepsilon}(t) - Z_f^n(t)] dP_n \right| \\
&\le C_1 \sup_n E^{P_n}\left[\int_0^T |a_n^\varepsilon(t\,,x(s)) - a_n(t\,,x(t))| dt \right] \\
&\le 2CC_1 \sup_n P_n\left[\sup_{0\le t\le T} |x(t)| \ge \ell \right] \\
&\qquad + C_1 \int_0^T \int_{-\ell}^{\ell} |a_n^\varepsilon(t\,,x) - a^\varepsilon(t\,,x)| p_n(0\,,x\,,t\,,y) dy \\
&\le CC_1\Delta(\ell) + C_1\sqrt{\delta_\varepsilon(T\,,\ell)}A(T\,,C\,,c)
\end{aligned}
$$

with the a similar estimate for $Z_f^\varepsilon - Z_f$

$$
\left| \int \Phi(\omega)[Z_f^\varepsilon(t) - Z_f(t)] dP \right| \le CC_1\Delta(\ell) + C_1\sqrt{\delta_\varepsilon(T\,,\ell)}A(T\,,C\,,c)
$$

We can now interchange $n \to \infty$ limit and $\varepsilon \to 0$ limit and we can conclude that

$$
\lim_{n\to\infty} \int \Phi(\omega) Z_f^n(t) dP_n = \int \Phi(\omega) Z_f(t) dP
$$

for all $t \ge s$, and therefore, from

$$
\int \Phi(\omega) Z_f^n(t) dP_n = \int \Phi(\omega) Z_f^n(s) dP_n
$$

it follows that

$$
\int \Phi(\omega) Z_f(t) dP = \int \Phi(\omega) Z_f(s) dP
$$

proving that P is a martingale solution for $[a(\cdot\,,\cdot), 0]$

Now we turn to proving uniqueness. We will attempt to solve the equation (15.6) and (15.7) with a function u of the form

$$
u(s\,,x) = \int_s^T \int_R h(t\,,y)\, p_C(s\,,x\,,t\,,y) dy
$$

Then as we saw earlier

$$
\begin{aligned}
u_s(s\,,x) + \frac{a(s\,,x)}{2} u_{xx}(s\,,x) &= u_s(s\,,x) + \frac{C}{2} u_{xx}(s\,,x) + \frac{a(s\,,x) - C}{2} u_{xx}(s\,,x) \\
&= -g(s\,,x) + [Bg](s\,,x) \\
&= -([I - B]g)(s\,,x)
\end{aligned}
$$

where

$$Bg(s\,,x) = \frac{a(s\,,x) - C}{2} u_{xx}(s\,,x)$$

and

$$\|Bg\|_{L_2} \le (1 - \frac{c}{C})\|g\|_{L_2}$$

If we have two martingale solutions P^i, $i = 1,2$ in $C_{s\,,x}$ and μ_t^i are their marginal distributions at times $t \ge s$, then they have densities $q^i(t\,,y)$ for almost all t, and

$$u(s\,,x) = \int_s^T [(I-B)g](t\,,y)\,q^1(t\,,y)dy = \int_s^T [(I-B)g](t\,,y)\,q^2(t\,,y)dy$$

Since we know that

$$\int_0^T \int_R |q^i(t\,,y)|^2 dtdy < \infty$$

in order to establish that $q^1 \equiv q^2$, it sufficient to show that the set of functions of the form $(I-B)g$ as g ranges over C^∞ functions is dense. Because $\|B\| < 1$ this is indeed true.

17. Brownian Motion on the Halfline.

It is not possible to construct the Brownian on the halfline $[0\,,\infty)$. Sooner or later it will hit 0 and then immeditely would turn negative as the following lemmas show.

Lemma 17.1 For any $x \in R$, for the Brownian motion on R,

$$P_x[\ \tau_0 < \infty\ ] = 1$$

Proof: From the reflection principle for any $x > 0$,

$$P_x[\ \tau_0 \le t\ ] = 2P_x[\ x(t) \le 0\ ] \to 1 \quad \text{as} \quad t \to \infty.$$

Lemma 17.2 (Blumenthal's 0-1 Law). If $P_{0,x}$ is any diffusion constructed as the unique martingale solution starting from x, and $A \in \mathcal{F}_{0+} = \cap s > 0\mathcal{F}_s$, then $P_{0,x}(A) = 0$ or 1.

Proof: Assume that $P_{0,x}(A) > 0$. Then $Q_A(\cdot)$ defined by

$$Q_A(E) = \frac{P_{0,x}(A \cap E)}{P_{0,x}(A)}$$

is easily seen to be again a martingale solution and so must coincide with $P_{0,x}$. Hence

$$P_{0,x}(A \cap E) = P_{0,x}(E) P_{0,x}(A)$$

In particular A and E are independent. Taking E to be A, we get $P(A) = 1$.

Lemma 17.3 For the Brownian Motion P_x starting from 0, for any $\delta > 0$,

$$P_x [\ \omega : x(t) \geq 0 \quad \text{for} \quad 0 \leq t \leq \delta\ ] = 0$$

Proof: For any $\delta > 0$

$$\begin{aligned} P_x [\ \cup_{\delta>0}\{\omega : x(t) \geq 0 \quad \text{for} \quad 0 \leq t \leq \delta\}\ ] &= \lim_{\delta \to 0} P_x [\ \omega : x(t) \geq 0 \quad \text{for} \quad 0 \leq t \leq \delta\ ] \\ &\leq \lim_{\delta \to 0} P_x [\ \omega : x(\delta) \geq 0\ ] \leq \frac{1}{2} \end{aligned}$$

The set $A = \cup_{\delta>0}\{\omega : x(t) \geq 0 \quad \text{for} \quad 0 \leq t \leq \delta\}$ is in $\mathcal{F}_{0+}$ and by Lemma 17.2, $P_x(A) = 0$.

We have to do something drastic to the Brownian Motion to keep it on the halfline. We want to characterize what we could do. We want to characterize all strong Markov families $\{P_x\}$ that have continuous paths, live on the half line and behave like a normal Brownian Motion away from 0. The last property is described by the following. For any smooth f that vanishes in a neighborhood of 0,

$$X_f(t) = f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s)) ds$$

is a Martingale with respect to $(\Omega, \mathcal{F}_t, P_x)$. By approximation we can easily extend the property to functions f which are quadratic near ∞ while still vanishing near the origin. Hence such processes have two moments and in fact as many moments as we need. Any function with $f(0) = f'(0) = f''(0)$ can be approximated in the C^2 topology by functions that vanish in a neighborhood of 0. Constants are no problem. Therefore the martingale property is valid for all smooth functions f, that satisfy $f'(0) = f''(0) = 0$.

Lemma 17.4 . The function $x(t)$ is a submartingale with respect to any P_x and can be written as

$$x(t) = A(t) + M(t) \tag{17.1}$$

where $M(t)$ is a martingale and $A(t)$ is a continuous nondecreasing function of t that increases only when $x(t)$ is at 0.

Proof: Approximate x by

$$f_\varepsilon(x) = x - \varepsilon \arctan\frac{x}{\varepsilon}$$

Because

$$\frac{1}{2}f''_\varepsilon(x) = g_\varepsilon(x) = \frac{\varepsilon x}{(\varepsilon^2 + x^2)^2} \geq 0$$

$f_\varepsilon(x(t))$ is a submartingale and in the limit so is $x(t)$. Existence of moments provides enough uniform integrability. Although a general theorem wil tell us that a decomposition of the form (17.1) holds, we will do it by hand in this case. We obviously want to take

$$A(t) = \lim_{\varepsilon\to 0} A_\varepsilon(t) = \lim_{\varepsilon\to 0}\int_0^t g_e(x(s))ds$$

Let us try to control

$$\begin{aligned}
&E\big[\,[A_\varepsilon(t)]^2\big]\\
&\quad= 2E\left[\int\int_{0\leq t_1\leq t_2\leq t} g_e(x(t_1))g_e(x(t_2))dt_1dt_2\right]\\
&\quad= 2E\left[\int_{0\leq t_1\leq t} g_e(x(t_1))[f_e(x(t)) - f_\varepsilon(x(t_1))]dt_1\right]
\end{aligned}$$

If we define

$$q_\varepsilon(t) = \sup_{0\leq s\leq t}\sup_x E^{P_x}\left[\ f_\varepsilon(x(s)) - f_\varepsilon(x(0))\ \right]$$

Then

$$E\big[\,[A_\varepsilon(t)]^2\big] \leq 2\,[q_\varepsilon(t)]^2$$

Or more generally,

$$E\big[\,[A_\varepsilon(t)]^k\big] \leq k!\,[q_\varepsilon(t)]^k$$

We next estimate $q_\varepsilon(t)$.

$$\begin{aligned}
q_\varepsilon(t) &= \sup_{0\leq s\leq t}\sup_x E^{P_x}\left[\ (x(t) - \varepsilon\arctan\frac{x(t)}{\varepsilon}) - (x - \varepsilon\arctan\frac{x}{\varepsilon})\ \right]\\
&\leq \sup_{0\leq s\leq t}\sup_x E^{P_x}\left[\ |x(t) - x|\ \right]\\
&\leq \sup_{0\leq s\leq t}\sup_x \sqrt{E^{P_x}\left[\ |x(t) - x|^2\ \right]}
\end{aligned}$$

We saw that $x(t)$ is a submartingale. By a similar argument one can show easily that $x^2(t) - t$ is a supermartingale. If we approximate x^2 by $h_\varepsilon(x) = [x - \varepsilon\arctan\frac{x}{\varepsilon}]^2$

$$h''_\varepsilon(x) \to \chi_{(0,\infty)}(x)$$

and is uniformly bounded. Therefore

$$x^2(t) - x^2(0) - \int_0^t \chi_{(0,\infty)}(x(s))ds$$

is a martingale. Therefore

$$\begin{aligned} E^{P_x}\left[\ |x(t)-x|^2\ \right] &= E^{P_x}\left[\ x^2(t) - 2x(t)x(0) + x^2(0)\ \right] \\ &\leq E^{P_x}\left[\ x^2(0) + t - 2x^2(0) + x^2(0)\ \right] \\ &= t \end{aligned}$$

Providing us the estimate

$$q_\varepsilon(t) \leq k! t^{\frac{k}{2}}$$

18. Brownian Motion on the Halfline (Continued).

We will develop two methods for the construction of Brownian motions on the halfline with sticky boundary condition. The reflected Brownian motion exists as the family of distributions $\{P_x^0\}$ obtained from the Brownian motion measures $\{P_x\}$, by the map $P_x^0 = P_x\Phi^{-1}$ where Φ maps $C[[0\,,\infty);R]$ into $C[[0\,,\infty);R_+]$ by $\beta(\cdot) \to |\beta(\cdot)|$. Relative to any $(\Omega_+\,,\mathcal{F}_t\,,P_x^0)$ there is a local time $A(t)$ with the following properties:

1. $A(t)$ is nondecreasing and the support of the measure $dA(t)$ is contained in the set $\{t : x(t) = 0\}$.

2. For any smooth function f

$$f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s))ds - f'(0)A(t) \tag{18.1}$$

is a martingale relative to $(\Omega_+\,,\mathcal{F}_t\,,P_x^0)$.

3. The process $x(t)$ spends no time on the boundary 0, i.e. for anr $x \in R_+$,

$$\int_0^t \chi_{\{0\}}(x(s))ds = 0 \quad \text{a.e.} \quad P_x^0 \tag{18.2}$$

We define a new increasing function

$$B(t) = \lambda^{-1}A(t) + t$$

where $\lambda > 0$ is a positive constant. $B(t)$ is a continuous strictly increasing function of t for any choice of $\lambda > 0$. For almost all ω the decomposition of B into

$$dB = \lambda^{-1}\, dA + dt$$

is its Lebesgue decomposition.

$$\text{support } dA = \{t : x(t) = 0\}$$

and, because of (18.2) we can take

$$\text{support } dt = \{t : x(t) > 0\}$$

We now conclude that the Radon-Nikodym derivatives are given by

$$\frac{dA}{dB} = \lambda\, \chi_{\{0\}}(x(s))$$

and

$$\frac{dt}{dB} = \chi_{(0,\infty)}(x(s))$$

We define τ_t as the solution of

$$B(\tau_t) = t$$

and define

$$y(t) = x(\tau_t).$$

Then

$$
\begin{aligned}
f(y(t))-f(y(0)) - &\int_0^t \frac{1}{2} f''(y(s))\chi_{(0,\infty)}(y(s))ds - \lambda f'(0)\int_0^t \chi_{\{0\}}(y(s))ds \\
&= f(x(\tau_t)) - f(y(0)) - \int_0^t \frac{1}{2} f''(x(\tau_s))\chi_{(0,\infty)}(x(\tau_s))ds \\
&\quad - \lambda f'(0)\int_0^t \chi_{\{0\}}(x(\tau_s))ds \\
&= f(x(\tau_t)) - f(x(0)) - \int_0^{\tau_t} \frac{1}{2} f''(x(\tau_{B(s)}))\chi_{(0,\infty)}(x(\tau_{B(s)}))dB(s) \\
&\quad - \lambda f'(0)\int_0^{\tau_t} \chi_{\{0\}}(x(\tau_{B(s)}))dB(s) \quad \text{(by cahange of variables} \quad s \to B(s)\,) \\
&= f(x(\tau_t)) - f(x(0)) - \int_0^{\tau_t} \frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s))dB(s) \\
&\quad - \lambda f'(0)\int_0^{\tau_t} \chi_{\{0\}}(x(s))dB(s) \\
&= f(x(\tau_t)) - f(x(0)) - \int_0^{\tau_t} \frac{1}{2} f''(x(s))ds - \lambda f'(0)\int_0^{\tau_t} dA(s) \\
&= f(x(\tau_t)) - f(x(0)) - \int_0^{\tau_t} \frac{1}{2} f''(x(s))ds - \lambda f'(0)A(\tau_t)
\end{aligned}
$$

is a martingale with respect to $(\Omega\,, \mathcal{F}_{\tau_t}\,, P_x^0)$. Since the σ-field $\sigma\{y(s) : 0 \le s \le t\} \subset \mathcal{F}_{\tau_t}$ we conclude that the distributions $\{P_x^\lambda\}$ of $y(\cdot)$ have the property:

$$
f(y(t)) - f(y(0)) - \int_0^t \frac{1}{2} f''(y(s))\chi_{(0,\infty)}(y(s))ds - \lambda f'(0)\int_0^t \chi_{\{0\}}(y(s))ds
$$

are $(\Omega_+\,, \mathcal{F}_t\,, P_x^\lambda)$ martingales. Speeding up the clock at the boundary so that the local time at the boundary turns into real time converts the reflected case to the sticky case. Conversely if we stop the clock when the process is at the boundary, any sticky case will become the reflected case.

Let us cosider the sticky case and define the function

$$
B(t) = \int_0^t \chi_{(0,\infty)}(x(s))ds.
$$

We then define τ_t by

$$
B(\tau_t) = t
$$

and $y(\cdot)$ by

$$
y(t) = x(\tau_t)
$$

To begin we need a lemma.

Lemma 18.1 Relative to any P_x^λ, the function $B(t)$ is almost surely strictly increasing in t. In other words, although the process sticks at the boundary it never spends a positive 'interval 'of time at the boundary.

Proof: The proof amounts to showing that if we start at the boundary, then

$$P_0^\lambda\,[\,\inf\{t : x(t) > 0\} = 0\,] = 1$$

Let us define

$$\tau = \inf\{t : x(t) > 0\}$$

Although τ is not quite a stopping time, it almost is, in the sense that $\tau + \varepsilon$ is a stopping time for every $\varepsilon > 0$. By working with $\tau + \varepsilon$ and letting ε go to 0 at the end the strong Markov property is seen to hold for τ. By Blumenthal's $0-1$ law,

$$P_x\,[\,\tau = 0\,] = 0 \quad \text{or} \quad 1$$

If it is 1 we are done. If it is 0, at the end of this time τ, the process is still at 0 but now 'knows' that it should get out. Clearly a violation of the strong Markov property.

Now we return to our main goal. We know that

$$f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s))ds - \lambda f'(0)\int_0^t \chi_{\{0\}}(x(s))ds$$

is a martingale. with respect to $(\Omega\,,\mathcal{F}_t\,,P_x^\lambda)$. Therefore for f satisfying the boundary condition $f'(0) = 0$,

$$\begin{aligned} f(x(\tau_t)) - f(x(0)) - \int_0^{\tau_t} \frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s))ds \\ = f(y(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(y(s))ds \end{aligned}$$

is a martingale and we are done.

Calculation: Let us try to calculate

$$p_\lambda(t) = P_0^\lambda\,[\,x(t) = 0\,]$$

We try to calculate

$$p_\lambda(x\,,t) = P_x^\lambda\,[\;x(t) = 0\;]$$

by solving the equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$$

with the boundary condition

$$\lambda u_x(0) = \frac{1}{2}u_{xx}(0)$$

and the initial condition

$$u(0\,,x) = \chi_{\{0\}}(x)$$

The Laplace transform

$$v(\sigma\,,x) = \int_0^\infty e^{-\sigma t}u(t\,,x)dt$$

solves

$$\sigma v - \frac{1}{2}v_{xx} = 0 \quad \text{for} \quad x > 0$$

with the boundary condition

$$\sigma v(0) - \lambda v_x(0) = 1$$

Clearly

$$v_\sigma(x) = a\exp[-\sqrt{2\sigma}\,x]$$

with

$$a\left[\;\sigma + \lambda\sqrt{2}\sigma\;\right] = 1$$

or

$$\alpha = a(\sigma\,,\lambda) = \left[\;\sigma + \lambda\sqrt{2\sigma}\;\right]^{-1}$$

Hence

$$\int_0^\infty p_\lambda(t)e^{-\sigma t}dt = \left[\;\sigma + \lambda\sqrt{2\sigma}\;\right]^{-1}$$

This can be explicitly inverted to yield

$$p_\lambda(t) = \int_0^\infty \sqrt{\frac{2}{\pi t}}e^{-\frac{x^2}{2t}-2\lambda x}dx = \int_0^\infty \sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}-2\lambda x\sqrt{t}}dx$$

Let P_0^0 be the reflected Brownian Motion starting from 0. The distribution of the local time process $A(t)$ can be found exactly.

Theorem 18.2 The process $A(t)$ has the same distribution as that of the process

$$M(t) = \sup_{0\le s\le t} \beta(t)$$

of the maximum of a Brownian motion starting from 0.

Proof: By Tanaka formula

$$x(t) = |\beta(t)| = \int_0^t \text{ sign } (\beta(s))d\beta(s) + A(t) = z(t) + A(t)$$

where $z(\cdot)$ is again a Browninan Motion process and $A(t)$ is the local time process. We will establish that

$$A(t) = -\inf_{0\le s\le t} z(s) = \sup_{0\le s\le t} [-z(s)]$$

In fact let f, g and A be arbitrary continuous functions with $f(t) \equiv g(t) + A(t)$, $f(0) = g(0) = A(0) = 0$, $f \ge 0$ and $A(\cdot)$, nondecreasing and increasing only when $f(t) = 0$, i.e. support of dA is contained in $\{t : f(t) = 0\}$. Then f and A are uniquely determined by g and

$$A(t) = \sup_{0\le s\le t} [-g(s)] \tag{18.3}$$

It is easy to see that with the choice (18.3) for $A(t)$, and $f(t) = g(t) + A(t)$ we get $f(t) \ge 0$ as well as $\{\text{support of dA}\} \subset \{\text{t} : \text{f(t)} = 0\}$. We will now prove uniqueness. Let

$$f_i(t) = g(t) + A_i(t);\quad i = 1, 2$$

with

$$\{\text{support of dA}_\text{i}\} \subset \{\text{t} : \text{f}_\text{i}(\text{t}) = 0\}\ \ \text{i} = 1, 2. \tag{18.4}$$

We have

$$f_1(t) - f_2(t) = A_1(t) - A_2(t).$$

Since $A_1(t) - A_2(t)$ is a function of bounded variation, using (18.4)

$$\begin{aligned}
[A_1(t) - A_2(t)]^2 &= \int_0^t [f_1(s) - f_2(s)][dA_1(s) - dA_2(s)] \\
&= -\int_0^t f_1(s)dA_2(s) - \int_0^t f_2(s)dA_1(s) \\
&\le 0
\end{aligned}$$

giving us uniqueness.

In particular we have

$$P_0^0\,[\,A(t) \ge \ell\,] = P_0\left[\sup_{0\le s\le t} \beta(s) \ge \ell\right] = \int_\ell^\infty \sqrt{\frac{2}{\pi t}} \exp[-\frac{x^2}{2\,t}]dx$$

19. Convergence of Markov Chains and markov Processes.

Let $f(t)$ be a right continuous function with left limits on $[0,T]$. The modulus of continuity of f is the function

$$\omega_f(h) = \sup_{\substack{0\le s\le t\le T\\ |s-t|\le h}} |f(s)-f(t)|$$

Then $\omega_f(h)$ is $\downarrow$ as $h\downarrow$ and $\delta_f = \omega_f(0+) = \lim_{\delta\downarrow 0}\omega_f(\delta)$ is the size of the largest jump of f. For any $\varepsilon>0$ let us define successive times $\{\tau_j\}$, $j=0,1,\cdots,N=N_f(\varepsilon)$, as

$$\tau_0 = 0,\ \tau_j = \{\inf t : t\ge \tau_{j-1} \& \, |f(t)-f(\tau_{j-1})|\ge\varepsilon\}$$

We get a finite sequence $0=\tau_0<\tau_1<\cdots<\tau_N$ and τ_{N+1} exceeds T. Let us define

$$\Delta_f(\varepsilon) = \inf_{1\le j\le N}[\tau_j-\tau_{j-1}]$$

If $|s-t|\le\Delta_f(\varepsilon)$, then s and t are in some pair of adjacent intervals $[\tau_{j-2},\tau_{j-1}],[\tau_{j-1},\tau_j]$ and it is easily seen that

$$\begin{aligned}|f(s)-f(t)| &\le |f(\tau_{j-2})-f(s)| + |f(\tau_{j-2})-f(\tau_{j-1}-0)|\\ &\quad + |f(\tau_{j-1}-0)-f(\tau_{j-1})| + |f(\tau_{j-1})-f(t)|\\ &\le 3\varepsilon+\delta\end{aligned}$$

In particular if $\Delta_f(\varepsilon)\ge h$ then $\omega_f(h)\le 3\varepsilon+\delta$. If f is random then

$$P[\,\omega_f(h)\ge 3\varepsilon+\delta\,]\le P[\,\Delta_f(\varepsilon)\ge h\,]$$

If P is a strong Markov Process such that there is an estimate of the form

$$P_{s\,,x}\left[\,\sup_{s\le t\le s+h}|x(t)-x|\ge\varepsilon\,\right]\le\phi_\varepsilon(h)$$

with $\phi_e(h)\to 0$ as $h\downarrow 0$, then

$$\sup_{j,\omega} P\left[\,\tau_j-\tau_{j-1}\le h\,|\mathcal{F}_{\tau_{j-1}}\,\right]\le\phi_\varepsilon(h).$$

We now get an estimate on N.

$$\begin{aligned}\sup_{j\,,\omega} E^P\left[\,e^{-(\tau_j-\tau_{j-1})}\,|\mathcal{F}_{\tau_{j-1}}\,\right] &\le e^{-h}(1-\phi_e(h))+\phi_e(h)\\ &= 1-(1-\phi_e(h))e^{-h}\le\rho_\varepsilon<1\end{aligned}$$

by choosing h sufficiently small.

$$E^P \left[\; e^{-\tau_k} \chi_{\tau_k < T} \; \right] \le \rho_\varepsilon^k$$

and therefore

$$P[N \ge k] \le P \left[\; \tau_k \le T \; \right] \le e^T \rho_\varepsilon^k$$

Finally we have

$$P \left[\; \Delta_{x(\cdot)}(\varepsilon) \le h \; \right] \le \inf_{k \ge 1} [k \phi_e(h) + \rho_\varepsilon^k] = \psi_\varepsilon(h)$$

where $\psi_e(h) \to 0$ as $h \to 0$ for any fixed ε. We have proved the following theorem.

Theorem 19.1 In order to prove that a family of Markov Process $\{P^{(n)}\}$ with values in R^d having right continuous paths with left limits is compact relative to weak convergence in the uniform topology wih the limit being supported on the set of continuous functions, the following estimates are sufficient:

$$\lim_{A \to \infty} \limsup_n P^{(n)}[\, |x(0)| \ge A \,] = 0 \tag{19.1}$$

$$\lim_{h \to 0} \limsup_{n \to \infty} \sup_{s\,,x} P^{(n)}_{s\,,x} \left[\sup_{s \le t \le s+h} |x(s) - x| \ge \varepsilon \right] = 0 \tag{19.2}$$

The maximum jump goes to zero, i.e. for any $\delta > 0$,

$$\lim_{n \to \infty} P^{(n)} \left[\; \delta_{x(\cdot)} \ge \delta \; \right] = 0 \tag{19.3}$$

We will illustrate by considering some examples. Let us construct an approximation to the reflected or sticky Brownian Motion by moving the trajectory to $a > 0$ after waiting for an exponential time at 0, with b being the expected value of the exponential time. This way we get a process $\{P_x^{a,b}\}$. Let us assume that $a \to 0$ and $b \to 0$ in such a way that

$$\lim_{\substack{a \to 0 \\ b \to 0}} \frac{a}{b} = \rho$$

for some $0 \le \rho \le \infty$. Then

Theorem 19.2 The limit

$$\lim_{\substack{a \to 0 \\ b \to 0}} P_x^{a\,,b} = P_x^\rho$$

exists and is the following process. If $\rho = 0$ we get the process where the Brownian motion is absorbed at 0. If $\rho = \infty$ we get the reflected Brownian Motion. If $0 < \rho < \infty$ we get the sticky case with the boundary condition $\frac{1}{2}u_{xx}(0) = \rho\, u_x(0)$.

Proof: The starting point is fixed at x. Therefore (19.1) is trivially valid. The size of the maximum jump is a that goes to 0 so that (19.3) is also trivial. Let us concentrate on (19.2). Since the process is homogeneous in time we can always take $s = 0$. Let $\varepsilon > 0$ be given. For any x and ε let τ be the exit time from $(x-\varepsilon, x+\varepsilon) \cap [0\,,\infty)$. We need to show that for each $\varepsilon > 0$,

$$\limsup_{h\to 0} \limsup_{\substack{a\to 0\\ b\to 0}} \sup_x P_x^{a,b}[\ \tau \le h\ ] = 0$$

We can assume that $a \le \varepsilon$. We consider two cases.

Case 1. $x \ge \frac{\varepsilon}{2}$. Consider a function f_ε which is 0 at x and 1 outside $(x - \frac{\varepsilon}{4}\,, x + \frac{\varepsilon}{4})$ and is smooth. Then

$$f(x(t)) - \int_0^t \frac{1}{2} f''(x(s))ds$$

is a martingale and if we take $\tau' \le \tau$ to be the exit time from $(x - \frac{\varepsilon}{2}\,, x + \frac{\varepsilon}{2})$

$$E^{P_x^{a\,,b}}[f(x(\tau' \wedge h)) - C_\varepsilon(\tau' \wedge h)] \le 0$$

where C_ε is a bound on f'' that can be made to depend only on ε.

$$P_x^{a\,,b}[\ \tau \le h\ ] \le P_x^{a\,,b}[\ \tau' \le h\ ] \le E^{P_x^{a\,,b}}[\ f(x(\tau' \wedge h))\ ] \le C_\varepsilon E^{P_x^{a\,,b}}[\ \tau' \wedge h\ ] \le C_\varepsilon h$$

Case 2. $x \le \frac{\varepsilon}{2}$. In this case τ is the same as the exit time from $[0\,, x+\varepsilon)$ and $\tau \ge \tau'$ where τ' is the exit time from $[0\,,\varepsilon)$. Since $x \le \frac{\varepsilon}{2}$, before it can exit from $[0\,, x+\varepsilon\,)$ it must reach $\frac{\varepsilon}{2}$ and exit again from the interval $(0\,,\varepsilon)$ which is an interval of the form $(x - \frac{\varepsilon}{2}\,, x + \frac{\varepsilon}{2})$ with $x = \frac{\varepsilon}{2}$. For the last event, we have already estimated the probability under case 1. Therefore even in this case

$$P_x^{a\,,b}[\ \tau \le h\ ] \le C_\varepsilon h$$

and we are done.

Now we turn to the identification of the limits. We deal with the three cases seperartely.

Case 1. $\rho = 0$. It suffices to show that for any $t < \infty$ and $\ell > 0$,

$$P_0^{a\,,b}[\ \inf\{s : x(s) \ge \ell\} \le t\ ] \to 0$$

proving that the limiting process never gets out of 0. We can infact try to calculate

$$P_0^{a,b}[\,\inf\{s : x(s) \geq \ell\} \leq t\,] = F(a\,,b\,,\ell\,,t)$$

exactly. If $\ell \leq a$, we need just one jump and

$$F(a\,,b\,,\ell\,,t) = 1 - e^{-\frac{t}{b}}$$

If $\ell > a$ we can solve it by computing the Laplace transform

$$\int_0^\infty e^{-\lambda t} F(a\,,b\,,\ell\,,dt) = \psi(a\,,b\,,\ell\,,\lambda)$$

as the value at 0 of solution to

$$\frac{1}{2}\psi''(x) = \lambda\psi(x)$$

with the 'boundary' conditions

$$\frac{1}{2}\psi''(0) = \frac{1}{b}[\psi(a) - \psi(0)]\,;\psi(\ell) = 1$$

Clearly

$$\psi(x) = Ae^{\sqrt{2\lambda}x} + Be^{-\sqrt{2\lambda}x}$$

for some constants A and B to be determined so that

$$Ae^{\sqrt{2\lambda}\ell} + Be^{-\sqrt{2\lambda}\ell} = 1$$

and

$$\lambda b(A+B) = A(e^{\sqrt{2\lambda}\,a} - 1) + B(e^{-\sqrt{2\lambda}\,a} - 1) \tag{19.4}$$

With out actually solving, we can pass to the limit and obtain in the limit

$$A + B = 0 \tag{19.5}$$

if $\frac{a}{b} \to 0$. Since $\psi(a\,,b\,,\ell\,,\lambda) = A + B$ this proves that

$$\lim_{\frac{a}{b}\to 0} \psi(a\,,b\,,\ell\,,\lambda) = 0$$

and this implies

$$\lim_{\frac{a}{b}\to 0} F(a\,,b\,,\ell\,,t) = 0$$

for all $\ell > 0$ and $t < \infty$.

Case 2. $\rho = \infty$. The crucial step in this case is to prove that for any $t < \infty$,

$$\lim_{\ell \to 0} \limsup_{\frac{a}{b} \to \infty} E^{P_x^{a,b}} \left[\int_0^t \chi_{[0,\ell)}(x(s)) ds \right] = 0, \tag{19.6}$$

guaranteeing that the process in the limit does not spend any time at 0. Let us construct a test function $f(x)$ on $[0, \infty)$ with these properties.

$$f(0) = 0\ ; \quad f'(0) > 0\ ; \quad f''(0) = \infty\ ; \quad f''(x) \geq -1\ ; \quad 0 \leq f \leq 1$$

The function f can be easily constructed and can be approximated by smooth functions by changing the function slightly near 0. The processes

$$f_n(x(t)) - f_n(x(0)) - \int_0^t f_n''(x(s))\chi_{(0,\infty)}(x(s))ds - \frac{f_n(a) - f_n(0)}{b} \int_0^t \chi_{\{0\}}(x(s))ds$$

are martingales, providing a uniform bound

$$E^{P_x^{a,b}} \left[\int_0^t f_n''(x(s))\chi_{(0,\infty)}(x(s))ds + \frac{f_n(a) - f_n(0)}{b} \int_0^t \chi_{\{0\}}(x(s))ds \right] \leq 1$$

If we let $n \to \infty$ and use Fatou's lemma, we get

$$E^{P_x^{a,b}} \left[\int_0^t f''(x(s))\chi_{(0,\infty)}(x(s))ds + \frac{f(a) - f(0)}{b} \int_0^t \chi_{\{0\}}(x(s))ds \right] \leq 1$$

Since $f'(0) > 0$, $f''(0) = \infty$ and $\frac{b}{a} \to \infty$ it is not hard to see that

$$\lim_{\ell \to 0} \liminf_{\frac{a}{b} \to \infty} \inf_{0 \leq x \leq \ell} \left[f''(x)\chi_{(0,\infty)}(x) + \frac{f(a) - f(0)}{b} \chi_{\{0\}}(x) \right] = \infty$$

and this is enough to establish (19.6)

We now complete the argument in case 2. Let $f(x)$ be a smooth function satisfying the boundary condition $f'(0) = 0$. We want to prove that with respect to any limit Q

$$f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s))ds \tag{19.7}$$

is a Martingale. It is enough to prove that it is a submartingale, because we can change the sign and show that it is a supermartingale as well. We can replace f by $f + \varepsilon g$ with

$g'(0) > 0$ and let $\varepsilon \to 0$. We can therefore assume with out loss of generality that $f'(0) > 0$ and try to establish that the expression in (19.7) is a submartingale. We know that

$$f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s))ds - \frac{f(a)-f(0)}{b}\int_0^t \chi_{\{0\}}(x(s))ds \quad (19.8)$$

is a martingale. Because $f'(0) > 0$, $\frac{f(a)-f(0)}{b} > 0$ for small a and we can assert that

$$f(x(t)) - f(x(0)) - \int_0^t \frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s))ds$$

is a submartingale with respect to $P_x^{a,b}$. The estimate (19.6) allows us to pass to the limit and conclude that the above expression remains a submartingale in the limit and the same estimate allows us to repalce $\chi_{(0,\infty)}$ by 1.

Case 3. $0 < \rho < \infty$. We start with f satisfying $\frac{1}{2}f''(0) = \rho f'(0)$. In the expression for (4.1) if we replace

$$\frac{1}{2} f''(x(s))\chi_{(0,\infty)}(x(s)) + \frac{f(a)-f(0)}{b}\chi_{\{0\}}(x(s))$$

by $\frac{1}{2}f''(x(s))$ the error is

$$\left[\ \frac{f(a)-f(0)}{b} - \rho f'(0)\ \right]\int_0^t \chi_{\{0\}}(x(s))ds \le t\left[\ \frac{a}{b}\frac{f(a)-f(0)}{a} - \rho f'(0)\ \right] = o(1)$$

as $a \to 0$ and $b \to 0$ with $\frac{a}{b} \to \rho$.

We now turn to the general question of approximating Markov Processes by Markov Chains. Let us for simplicity consider the one dimensional case with time homogeneous transitions. Assume that for $h > 0$ we have a transition probability $\pi_h(x, dy)$ from $R \to R$. h is the time step and we have a Markov Chain $X_n^{(h)}$ with π_h as transition probabilities. We make a process out of it by defining $X_h(t) = X^{(h)}_{[\frac{t}{h}]}$ so we get a continuous parameter process moving only by jumps at times that are integral multiples of h. Let us make the assumption that

$$\lim_{h\to 0}\frac{1}{h}\int_R [f(y) - f(x)]\pi_h(x, dy) = (\mathcal{L}f)(x) \quad (19.9)$$

exists for all smooth functions with compact support, with the limit being uniform for x in any compact subset of R. $\mathcal{L}$ is assumed to be an operator of the form

$$(\mathcal{L}f)(x) = \frac{1}{2}a(x)f''(x) + b(x)f'(x)$$

with continuous coefficients $a(x)$ and $b(x)$, that may be unbounded. We assume that for the coefficients $a(\cdot)\,, b(\cdot)$ there is a unique martingale solution P_x starting from any point x at time 0. In particular, it is assumed that although the coefficients are unbounded, there is no explosion. Let $P_x^{(h)}$ denote the distribution of $X^{(h)}(t)$ that starts from x. We want to prove that $P_x^h \to P_x$ weakly as $h \to 0$. First we prove a lemma.

Lemma 19.3 The condition (19.9) is equivalent to the following set of conditions:

$$\lim_{h\to 0} \frac{1}{h}\pi_h(x\,, \{y : |x-y| \geq \varepsilon\}) = 0 \tag{19.10}$$

$$\lim_{h\to 0} \frac{1}{h}\int_{|y-x|\leq 1}(y-x)\pi_h(x\,, dy) = b(x) \tag{19.11}$$

$$\lim_{h\to 0} \frac{1}{h}\int_{|y-x|\leq 1}(y-x)^2\pi_h(x\,, dy) = a(x) \tag{19.12}$$

with all limits holding uniformly over x in compact subsets of R.

Proof: Assuming (19.10) the domain of integration in (19.11) and (19.12) can be limited to $\{y : |y-x| \leq \varepsilon\}$ for any $\varepsilon > 0$, and Taylor expansion of $f(y)$ around x is all that is needed to go from (19.11) and (19.12) to (19.9). If we pick $f = 1$ in a small interval and zero outside a slightly larger interval then $(\mathcal{L}f)(x)$ is zero in the small interval. The convergence implied by (19.9) is seen to imply (19.10). The choices of $f(y) = y$ and $f(y) = y^2$ in a neighbourhood of x are easily seen to imply (19.11) and (19.12).

Theorem 19.4 Under the assumption that (19.9) holds $P_x^{(h)} \to P_x$ as $h \to 0$.

Proof: Step 1. Let us construct a smooth cutoff function $\phi(x)$ which is equal to 1 on $|x| \leq 1$ and 0 on $|x| \geq 2$ and $0 \leq \phi \leq 1$ everywhere. Consider

$$\pi_h^\ell(x\,, dy) = \phi(\frac{x}{\ell})\pi_h(x\,, dy) + (1 - \phi(\frac{x}{\ell}))\delta_x(dy)$$

where δ_x is the degenerate measure at x. Then

$$\lim_{h\to 0}\int[f(y) - f(x)]\pi_h^\ell(x\,, dy) = \phi(\frac{x}{\ell})(\mathcal{L}f)(x) \tag{19.13}$$

uniformly in x for smooth functions f. In particular for any smooth function f

$$f(x(nh)) - f(x(0)) - \sum_{j=1}^{n} g_h^\ell(x((j-1)h))$$

is a martingale with respect to $P_x^{h,\ell}$ which is the process corresponding to the Markov Chain with transition probabilities π_h^ℓ. Here

$$g_h^\ell(x) = \int [f(y) - f(x)]\pi_h^\ell(x\,, dy),$$

and from (19.13) we obtain a bound

$$|g_h^\ell(x)| \le C_f^\ell h$$

implying that

$$f(x((nh)) - f(x(0)) + nhC_f^\ell$$

is a submartingale. Consider a smooth function satisfying $f(y) = 1$ for $|y| \le \frac{1}{2}$, $f(y) = 0$ for $|y| \ge 1$ and $0 \le f(y) \le 1$ for all y. Let τ be the exit time from the interval $(-1\,,1)$. Then for any path with $|x(0)| \le \frac{1}{2}$,

$$\inf\{t : |x(t) - x(0)| \ge 2\} \ge \tau.$$

From the submartingale property we have

$$E\left[\ f(x((nh\wedge\tau)) - 1 + nhC_f^\ell\ \right] \ge 0$$

or

$$E\,[\ 1 - f(x((nh\wedge\tau))\ ] \le nhC_f^\ell$$

Therefore

$$\begin{aligned} P\,[\ \inf\{t : |x(t) - x(0)| \ge 2\} \le kh\ ] &\le P\,[\ \tau \le kh\ ] \\ &\le P\,[\ 1 - f(x(((kf\wedge\tau)) = 1\ ] \\ &\le khC_f^\ell \end{aligned}$$

where P refers to the chain starting from any point inside $|x| \le \frac{1}{2}$. Using a finite number of such functions, for any x inside $|x| \le 2\ell$, we can get a uniform bound on the exit time from an interval $(x-2\,, x+2)$ around x. On the other hand if we start from any x with $|x| \ge 2\ell$ the path does not move. Therefore we have the following estimate

$$\sup_x P_x^{h,\ell}\left[\ \sup_{0\le s\le t} |x(s) - x| \ge 2\ \right] \le C^\ell\, t\, h$$

We can modify the argument be rescaling the function f and we will get

$$\sup_x P_x^{h,\ell}\left[\sup_{0\le s\le t}|x(s)-x|\ge\varepsilon\right]\le C_\varepsilon^\ell\, t\, h$$

The size of the larges jump is easily estimated.

$$\begin{aligned}P\left[\sup_{0\le j\le n-1}|x((j+1)h)-x(jh)|\ge\delta\right]&\le n\sup_x \pi_h^\ell(x\,,\{y:|x-y|\ge\delta\})\\&\le n\,o(h)=o(1)\end{aligned}$$

provided $nh=O(1)$. This proves the compactness of the family $\{P_x^{h,\ell}\}$ as $h\to 0$ for fixed ℓ and x.

Step 2. If Q is any limit point of $P_x^{h,\ell}$ as $h\to 0$, we now show that Q is a martingale solution for

$$\mathcal{L}_\ell=\phi(\frac{x}{\ell})\mathcal{L}=\phi(\frac{x}{\ell})\left[\frac{a(x)}{2}D_x^2+b(x)D_x\right]$$

This is relatively simple. We consider the functionals

$$Z_f^h(nh)=f(x(nh))-f(x(0))-h\sum_{j=0}^{n-1}\frac{1}{h}\int[f(y)-f(x(jh))]\pi_h^\ell(x(jh\,,dy)$$

that are $P_x^{h,\ell}$ martingales and converge as $h\to 0$ to

$$Z_f^h(t)=f(x(t))-f(x(0))-\int_0^t(\mathcal{L}_\ell f)(x(s))ds$$

proving that Q is a martingale solution for $\mathcal{L}_\ell$.

Step 3. We may not have uniqueness for $\mathcal{L}_\ell$. But in any case until exit from the interval $|x|\le\ell$ there is no difference between $\mathcal{L}$ and $\mathcal{L}_\ell$. Therefore any weak limit Q of $P_x^{h,\ell}$ as $h\to 0$ must coincide with P_x the unique soulution to $\mathcal{L}$ starting from x on the σ-field $\mathcal{F}_{\tau_\ell}$ that corresponds to stopping time τ_ℓ, which is the exit time from the interval $|x|\le\ell$. The set

$$B_t^\ell=\left\{x(\cdot):\sup_{0\le s\le t}|x(s)|\ge\ell\right\}$$

is a closed set and therefore

$$\limsup_{h\to 0}P_x^{h,\ell}\left[\,B_t^\ell\,\right]\le Q\left[\,B_t^\ell\,\right]=P_x\left[\,B_t^\ell\,\right]$$

From the assumption that the solution to $\mathcal{L}$ does not explode, we can now assert

$$\limsup_{\ell\to\infty}\limsup_{h\to 0}P_x^{h,\ell}\left[\,B_t^\ell\,\right]\le\limsup_{\ell\to\infty}P_x\left[\,B_t^\ell\,\right]=0$$

allowing us to interchange the order of the two limits $\ell\to\infty$ and $h\to 0$ and we are done.

20. Reflected Processes in Higher Dimensions.

We will quickly describe some multidimensional generalizations of reflected Brownian Motion. Let G be a smooth region in R^d and $a = \{a_{i,j}(x)\}$, $b = \{b_i(x)\}$, coefficients that are 'nice ', i.e. a is smooth and positive definite and b is smooth. We want to construct a solution and we need to describe what happens when the path reaches the boundary. We will deal exclusively with the the reflected case and just make some comments at the end regarding other possibilities. Reflection is a bad choice for the name, but in reality the process gets kicked in, in some direction pointing to the interior as soon as it reaches the boundary. So we have a direction $J(b)$ pointing to the interior at every point $b \in B = \partial G$. We want to show that given a, b, G and J, there is a unique family of solutions $\{P_x : x \in G \cup B\}$ on $\Omega = C[[0\,,\infty); G \cup B]$ with the following properties.

1. $P_x[x(0) = x] = 1$

2. $P_x\Big[\int_0^t \chi_B(x(s))ds = 0\Big] = 1$

3. For any smooth function f that satisfies $< J(b)\,, (\nabla f)(b) > \geq 0$ on B,

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

is a submartingale with respect to $(\Omega\,, \mathcal{F}_t\,, P_x)$.

The question of existence is a question of nonexplosion as well. To avoid the problem of dealing with this issue let us assume that our domain G is bounded. Then the question is purely local. If we start from $x \in G$ we know what happens until we reach the boundary. We do not see it. P_x is just the same as the solution with no boundary until the exit time from G. We therefore need to construct local solutions when we start on or near the boundary. This is carried out in several steps.

Step 1. Make a change of coordinates so that a boundary point b becomes 0 and the boundary becomes $x_1 = 0$, a straightline near that point. This will reduce the problem to a half space. The direction J on $B = \{x : x_1 = 0\}$ can be described by $(1, J_2(x_2, \cdots, x_d), \cdots, J_d(x_2, \cdots, x_d))$.

Step 2. Now make another change of coordinates of a special type, $x_1 \to x_1$, $x_i \to x_i - x_1\, J_i(x_2\,, \cdots, x_d)$ for $2 \leq i \leq d$. The bounadry remains the same, but the new direction J is just $(1, 0, \cdots, 0)$, the inward normal.

Step 3. By a Girsanov formula which can be extended to this case we can assume that $b = 0$.

Step 4. In the current coordinate system $a_{1,1}(x)$ is a strictly positive function and we can do a random time change using τ_t defined by

$$\int_0^{\tau_t} a_{1,1}(x(s))ds = t$$

to reduce it to $a_{1,1} \equiv 1$. At this point if $f = f(x_1)$ is a function of x_1 only then

$$(\mathcal{L}f)(x) = \frac{1}{2}f''(x_1)$$

so that the process $x_1(t)$ is in fact the one dimensional reflected Brownian Motion.

Step 5. We can find a square root $\sigma(x)$ for $a(x)$ such that $a(x) = \sigma(s)\sigma^*(x)$ with $\sigma_{1,1}(x) \equiv 1$ and $\sigma_{1,j}(x) \equiv 0$ for $2 \leq j \leq d$. The stochastic differential equations for $x(t)$ now look like

$$dx_1(t) = d\beta(t) + A(t)$$

which is the decomposition of the reflected one dimensional Brownian motion and is already solved.

$$dx_j(t) = \sigma_{j,1}(x_1(t), x_2(t), \cdots, x_d(t))dx_1(t) + \sum_{2\leq k\leq d} \sigma_{j,k}(x_1(t), x_2(t), \cdots, x_d(t))dx_j(t)$$

which can be solved by iteration for $x_2(\cdot), \cdots, x_d(\cdot)$ because the boundary has no effect on them directly.

Comments: We may try to stick to the boundary a little bit. This is dealt the same way as in one dimension. We can obtain it by random time change from the reflected case using the local time on the boundary. The holding rate ρ can now be a function $\rho(b)$ defined on B. The local time $A(t)$ in the reflected case can be used to construct the time change

$$\int_0^{\tau_t} \lambda(x(s))dA(s) + \tau_t = t$$

where $\lambda(b) = [\rho(b)]^{-1}$. Finally a new phenomenon that can happen is that the path might diffuse on the boundary which amounts to kick having a random tangential component. Imagine in the case of a halfspace, being kicked form the boundary point $(0, y)$, to the interior point $(\delta\, , y + \delta J(y) + \sqrt{\delta}\xi)$ where ξ is a gaussian random vector with mean 0 and covariance matrix $D(y)$. The boundary condition then becomes

$$(\mathcal{B}f)(b) = \frac{\partial f}{\partial x_1} + \sum_{j=2}^{d} J_j(y)\frac{\partial f}{\partial x_j} + \frac{1}{2}\sum_{i,j=1}^{d} D_{i,j}(y)\frac{\partial^2 f}{\partial x_i \partial x_j} = 0$$

at $b = (0, y) \in B$. Here y refers to the cordinates x_2 through x_d. Of course this can happen in the sticky situation as well and the boundary condition then is

$$(\mathcal{L}f)(b) = \rho(b)(\mathcal{B}f)(b)$$

THE END