Short Stochastic Analysis Course

Siddig

1 Quick Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathsf{X}, \mathcal{X})$ be a measurable space. A measurable function $X : (\Omega, \mathcal{F}) \to (\mathsf{X}, \mathcal{X})$ is a called an X valued random variable. It is customary to drop the argument notation and write X only instead of $X(\omega)$ to denote the values of X. The push forward measure $\mathbb{P} \circ X^{-1}$ is called the law or the distribution of X. Denote by $\sigma(X)$ the smallest σ-algebra σ that makes the map $X : (\Omega, \sigma) \to (\mathsf{X}, \mathcal{X})$ measurable. In fact $\sigma(X) = \{X^{-1}(S) | S \in \mathcal{X}\}$. A trivial to prove but worth noting is the fact for any probability measure μ one X on can associate a random variable whose law is μ, indeed the identity map $\mathrm{Id} : (\mathsf{X}, \mathcal{X}, \mu) \to (\mathsf{X}, \mathcal{X})$ has law μ.

Stochastic processes are random processes that take values in the path space $\mathscr{P}([0,\infty), \mathsf{X})$ which are measurable w.r.t to the σ-algebra $\mathcal{Z}$ generated by the projection maps $\pi_t : \mathscr{P}([0,\infty), \mathsf{X}) \to \mathsf{X}$ defined as $\pi_t : \omega \mapsto \omega(t)$, this is equivalent to saying that $\mathcal{Z}$ is the Borel σ algebra generated by the product topology $\mathscr{P}([0,\infty), \mathsf{X}) \cong \prod_{t\in[0,\infty)} \mathsf{X}$. We may also write $(X_t)_{t\geq 0}$ or just X_t for stochastic processes. The definition can directly be adopted for discrete times or finite intervals of time. Two processes X_t and Y_t are called indistinguishable if $\mathbb{P}(\{X_t = Y_t | \forall t \in [0,\infty)\}) = 1$ and modification if $\inf_{t\in[0,\infty)} \mathbb{P}(\{X_t = Y_t\}) = 1$. From here onwards we mostly deal with processes whose ranges at every time t is $\mathsf{X} = \mathbb{R}^d$.

Stochastic processes on the path space $\left(\mathscr{P}([0,\infty), \mathbb{R}^d), \mathcal{Z}\right)$ don't generally behave well with "continuum" operations, for example for a random process X_t the paths $t \mapsto X_t(\omega)$ needs not be measurable w.r.t Borel or Lebesgue σ-algebra and some random variables of interest such as $\sup_{t\in[0,1]} X_t$ and X_τ for a random time τ need not be measurable w.r.t $\mathcal{Z}$. A smaller path subspace but more regular is the continuous paths space $\left(\mathscr{C}([0,\infty), \mathbb{R}^d), \mathcal{A}\right)$ with $\mathcal{A} = \{S \cap \mathscr{C}([0,\infty), \mathbb{R}^d) | S \in \mathcal{Z}\}$ being the σ-subalgebra inherited from the path space and we call such processes continuous stochastic processes. A larger space that includes jumps as well is known as Skorohod space[1]. The behaviour of $\mathcal{A}$ is demonstrated by the following theorem.

Theorem 1.1 *The metric space $\mathscr{C}([0,\infty), \mathbb{R}^d)$ equipped with the metric*

$$d(f,g) = \sum_{n\geq 0} \frac{1}{2^n} \min(1, \max_{t\in[0,n]} \|f(t) - g(t)\|)$$

is a Polish space (complete separable metric space). The topology generated by the metric d coincides with the topology of uniform convergence on compact subsets of $[0,\infty)$. Furthermore, the σ-algebra generated by this topology equals $\mathcal{A}$.

A filtration $(\mathcal{F}_t)_{t\geq 0}$ is an increasing sequence of σ-subalgbras of $\mathcal{F}$. A stochastic process is called adapted w.r.t to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is measurable w.r.t $\mathcal{F}_t$ for all $t \in [0,\infty)$. X_t is called progressively measurable if the map $(s,w) \in ([0,t], \Omega) \mapsto X_s(w)$ is $\mathcal{B}([0,t]) \times \mathcal{F}_t$-measurable.

Theorem 1.2 *A progressively measurable process is adapted. A continuous[2] adapted process is progressively measurable. Two continuous modifications of each other are indistinguishable.*

A sequence of random variables $(X_n)_{n\geq 0}$ is said to converges[3] to a random variable X:

- almost surely[4] (abbreviated a.s.) if $\mathbb{P}(\{X_n \to X\}) = \mathbb{P}(\{X_n(\omega) \to X(\omega) | \omega \in \Omega\}) = 1$. This convergence is equivalent to $\mathbb{P}(\limsup \|X_n - X\| > \epsilon) = 0$ for al $\epsilon > 0$.
- in L^p (with $p \geq 1$) if $\mathbb{E}[\|X_n - X\|^p] \to 0$.

[1] Or spaces

[2] Right continuity or left continuity is enough.

[3] For a general metric spaces, separability is needed to ensure that the function $d(X_n, X)$ is measurable w.r.t to the product σ-algebra, since in the non-seprabale case the product σ-algebra is different from the Borel σ-algebra generated by the product topology.

[4] It is also known as almost everywhere (a.e.).

- in probability if $\mathbb{P}(\|X_n - X\| > \epsilon) \to 0$ for all $\epsilon > 0$.
- in distribution if $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for every continuous bounded function $f : \mathbb{R}^d \to \mathbb{R}$. This is equivalent[5] to $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ for all Borel sets A that satisfy $\mathbb{P}(X \in \partial A) = 0$.

If $X_n \to X$ in L^1 or a.s. then $X_n \to X$ in probability and convergence in probability implies convergence in distribution. Furthermore if $X_n \to X$ in probability then there exists a subsequence $\varphi(n)$ that satisfies $X_{\varphi(n)} \to X$ a.s..

Theorem 1.3

- *Continuous Mapping Theorem: Let X and $(X_n)_{n\geq 0}$ defined on a metric space. Let g be a continuous function (with an appropriate domain) and suppose that $X_n \to X$ almost surely, in probability, or in distribution, then $g(X_n) \to g(X)$ accordingly almost surely, in probability, or in distribution.*
- *Dominated Convergence Theorem: Let X and $(X_n)_{n\geq 0}$ be real-valued random variables that satisfy $\sup_n |X_n| \in L^1$ then $X_n \xrightarrow{a.s.} X$ implies $X_n \xrightarrow{L^1} X$.*
- *Uniform Integrability: Let $(X_n)_{n\geq 0}$ be a uniformly integrable (UI) family, i.e.*

$$\lim_{M\to\infty} \sup_n \mathbb{E}[|X_n| \mathbb{1}_{\{|x|\geq M\}}(X_n)] = 0$$

then $X_n \xrightarrow{p} X$ implies $X_n \xrightarrow{L^1} X$.

Theorem 1.4 (Law of Large Numbers) *Let $(X_n)_{n\geq 0}$ be an i.i.d. L^1 sequence. Then*

$$\frac{1}{n}\sum_{i=0}^{n} X_n \to \mathbb{E}[X_1]$$

where the convergence holds in all modes of convergence.

Theorem 1.5 (Central Limit Theorem) *Let $(X_n)_{n\geq 0}$ be an i.i.d L^2 sequence. Then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_n - \mathbb{E}[X_1]) \xrightarrow{d} \mathcal{N}(0, \mathrm{COV}(X_1))$$

2 Brownian Motion/Wiener Process

An $\mathbb{R}^d$ valued standard Brownian motion $(B_t)_{t\geq 0}$ is a continuous process that satisfies for every $t > s \geq 0$: $B_0 = 0$, $B_t - B_s \sim \mathcal{N}(0, (t-s)\mathrm{Id})$, and that the processes $B_t - B_s$ and B_s are independent. Below are two ways to construct Brownian motions.

Construction 1: A natural idea is to try approximating with discrete paths with small steps and apply CLT for every small interval. First we restrict ourselves to constructing Brownian motion on the interval $[0, 1]$. Let $(A_n)_{n\geq 1}$ be a sequence of i.i.d random variables that satisfy $\mathbb{E}[A_1] = 0$ and $\mathbb{E}[A_1^2] = 1$. Define the random variables

$$S_t^n = \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor} A_i + \frac{\{nt\}}{\sqrt{n}} A_{\lfloor nt\rfloor+1}$$

The process S_t^n is a piece-wise continuous process. Informally the idea is that every process $\frac{1}{\sqrt{n}}A_i$ in the sum above "represents" the increment $B_{i/n} - B_{(i-1)/n}$ (notice that they have the same mean and variance), if we double n then the process $\frac{1}{\sqrt{2n}}(A_{2i} + A_{2i+1})$ "represents" the same process $B_{i/n} - B_{(i-1)/n}$, and as $n \to \infty$ every increment is represented as averages and we hope to apply CLT to these averages. Because of the definition of S_t^n (the representatives $\frac{1}{\sqrt{n}}A_i$ and $\frac{1}{\sqrt{2n}}(A_{2i} + A_{2i+1})$ are independent) we should hope for the convergence of the laws of the processes[6] which is not directly clear. Once we prove the convergence of the laws to a probability measure on the space on continuous paths there is a random variable with this law. This convergence is reminiscent of Arzela-Ascoli theorem, in fact we have the following theorem:

[5]Portmanteau theorem.

[6]One can search Levy's construction of Brownian motion for a similar construction with converging paths. This construction shows rather properties of the space of probability measures.

Theorem 2.1 (Prokhorov) [7] *Let $(\mu_n)_{n\geq 1}$ be a sequence of Probability measures on $\mathcal{C}([0,1],\mathbb{R}^d)$. The following statements are equivalent:*

1. *The sequence $(\mu_n)_{n\geq 1}$ has a weakly*[8] *converging subsequence.*
2. *The sequence $(\mu_n)_{n\geq 1}$ is uniformlly tight, i.e. for every $\epsilon > 0$ there exists a compact subset $K \subset \mathcal{C}([0,1],\mathbb{R}^d)$ such that $\mu_n(K^c) < \epsilon$.*
3. *The sequence $(\mu_n)_{n\geq 1}$ satisfies:*
   - $\lim_{M\to\infty} \sup_n \mu_n\left(\{\|f(0)\| > M\}\right) = 0.$
   - $\lim_{\delta\to 0^+} \sup_n \mu_n\left(\{\max_{|s-t|\leq\delta} \|f(t)-f(s)\| > \epsilon\}\right) = 0.$

The first requirement in condition 3 is equivalent to $\lim_{M\to\infty}\sup_n \mathbb{P}(\|S_0^n\| > M) = 0$ which is trivial because $S_0^n = 0$. For the second requirement let $s \leq t \leq s+\delta$ and notice that when $\lfloor nt\rfloor = \lfloor ns\rfloor$ then $\delta < 1/n$ and we have $\|S_t^n - S_s^n\| = \frac{\{nt\}-\{ns\}}{\sqrt{n}}\|A_{\lfloor nt\rfloor+1}\| \leq \sqrt{\delta}\|A_{\lfloor nt\rfloor+1}\|$. If $\lfloor ns\rfloor < \lfloor nt\rfloor$ then $\|S_t^n - S_s^n\| = \|S^n_{\frac{\lfloor ns\rfloor+1}{n}} - S_s^n\| + \|S^n_{\frac{\lfloor nt\rfloor}{n}} - S^n_{\frac{\lfloor ns\rfloor+1}{n}}\| + \|S_t^n - S^n_{\frac{\lfloor nt\rfloor}{n}}\| \leq \sqrt{\delta}(\|A_{\lfloor nt\rfloor}\| + \|A_{\lfloor ns\rfloor+1}\|) + \|S^n_{\frac{\lfloor nt\rfloor}{n}} - S^n_{\frac{\lfloor ns\rfloor+1}{n}}\|$. Assume $\delta < 1$ and notice that

$$\left\{\max_{|\frac{a}{n}-\frac{b}{n}|\leq\delta} |S^n_{\frac{a}{n}} - S^n_{\frac{b}{n}}| > \epsilon\right\} \subset \bigcup_{0\leq k\leq \log_2(n\delta)} \left\{\max_{0\leq a\leq n/2^k-1} |S^n_{\frac{(a+1)2^k}{n}} - S^n_{\frac{a2^k}{n}}| > \frac{\epsilon}{10}\left(\frac{2^k}{n\delta}\right)^{1/5}\right\}$$

This can be seen by taking complement of the sets in the inclusion and using binary representation of integers. Let's further suppose that the random variables (A_i) are bounded by a constant C and employ Markov inequality to get

$$\mathbb{P}(\max_{|\frac{a}{n}-\frac{b}{n}|\leq\delta} |S^n_{\frac{a}{n}} - S^n_{\frac{b}{n}}| > \epsilon) \leq \sum_{k=0}^{\lfloor\log_2(n\delta)\rfloor} \frac{n}{2^k}\left(\frac{\epsilon}{10}\left(\frac{2^k}{n\delta}\right)^{1/5}\right)^{-4} \mathbb{E}[|S^n_{\frac{2^k}{n}} - S^n_0|^4]$$

Because of the independence of A_i we have

$$\mathbb{E}[|S^n_{\frac{2^k}{n}} - S^n_0|^4] = \mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{2^k} A_i\right)^4\right] \leq C^4\frac{2^{2k}}{n^2}$$

Putting all of this together

$$\begin{aligned}\mathbb{P}\left(\max_{|t-s|\leq\delta} \|S_t^n - S_s^n\| > \epsilon\right) \leq& \mathbb{1}_{\{C\sqrt{\delta}>\epsilon\}} + \mathbb{1}_{\{2C\sqrt{\delta}>\epsilon/3\}} + C^4\sum_{k=1}^{\lfloor\log_2(n\delta)\rfloor} \frac{n}{2^k}\left(\frac{\epsilon}{10}\left(\frac{2^k}{n\delta}\right)^{1/5}\right)^{-4}\frac{2^{2k}}{n^2} \\ <& C\epsilon^{-4}\delta\end{aligned} \tag{1}$$

with C a constant independent of n (check it). Prokhorov theorem guarantees the convergence of the laws of the processes S^n to a measure on the space of continuous functions. This measure is called *Wiener measure* $\mathbb{W}$. The identity map on the probability space $(\mathscr{C}([0,1],\mathbb{R}^d),\mathcal{A},\mathbb{W})$ is called the canonical Brownian motion, for example from CLT we have $0 < s \neq t < 1$ we have

$$\mathbb{W}\left(\{f(t)-f(s)\in A\}\right) = \frac{1}{\sqrt{(2\pi)^d|t-s|^d}}\int_A e^{-\frac{1}{2|t-s|}\|x\|^2}dx$$

Mostly $(\mathscr{C}([0,1],\mathbb{R}^d),\mathcal{A},\mathbb{W})$ is not used explicitly as the space, and we write $\mathbb{P}(B_t - B_s \in A)$ for $\mathbb{W}\left(\{f(t)-f(s)\in A\}\right)$. The Brownian motion process can be extended to the space $\mathscr{C}([0,\infty),\mathbb{R}^d)$ by concatenating many independent Brownian motions each defined on unit intervals $[k,k+1]$. Another way is to denote by $\mathbb{W}_T$ the Wiener measure on $\mathscr{C}([0,T],\mathbb{R}^d)$ (since the construction is the same) and define

$$\mathbb{W}_\infty(A) = \inf_{T\in\mathbb{N}} \mathbb{W}_T\left(\{f\restriction_{[0,T]} | f\in A\}\right)$$

The calculations we made to prove 1 form a general technique and in fact we have

[7] The theorem holds for Polish spaces, but for the sake of showing its relevance to Arezla-Ascoli we restricted the statement to space of continuous functions. The proof can utilize the diagonalization technique used to prove Arezela-Ascoli.

[8] Notice that weak convergence of measures is the weak* convergence when considering measures as elements of the dual space of continuous functions.

Theorem 2.2 *Let (X^n) be a sequence of $\mathbb{R}^d$ valued random process for which there exist $\alpha > 1, \beta > 0$ that satisfy*

$$\mathbb{E}[|X_t^n - X_s^n|^\alpha] \leq M|t-s|^{1+\beta}$$

then

$$\lim_{\delta \to 0^+} \sup_n \mathbb{P}\left(\max_{|t-s|\leq\delta} |X_t - X_s| > \epsilon\right) = 0$$

The calculation above used $\alpha = 4, \beta = 1, a = 1/5$, but any $0 < a < \beta/\alpha$ works. We will consider this again later. A similar construction in which the variables themselves converge is called Levy's construction. This convergence of the random processes S^n in distribution to the Brownian motion is known as *Donsker's Invariance Principle*. We call the filtration $\mathcal{F}_t = \sigma(\{B_s | 0 \leq s \leq t\})$ the natural filtration. In a general filtered space $B_t - B_s$ is independent of the σ-algebra $\mathcal{F}_s$.

Construction 2: Since $B_t - B_s \sim \mathcal{N}(0, |t-s|)$, we would view Brownian motions as "integration $\int_s^t \mathbb{1}_{[s,t]}(u) dB_u$ of infinitesimal normal distributions $dB_t \sim \mathcal{N}(0, dt)$". We could ask to replace $\mathbb{1}_{[s,t]}$ by a generic function f. If such an integral existed we would like to have $\mathbb{E}\left[\left(\int_s^t f(u) dB_u\right)^2\right] = \int_s^t f(u)^2 du$. This motivates the use of L^2 theory to construct a Brownian motion. We start by "imagining" that the integral is defined for an orthonormal basis (e_i) and define extend it to the whole space.

Let (e_i) be an orthonormal sequence of $L^2([0,1], \mathbb{R})$, $f \in L^2([0,1], \mathbb{R})$, and (A_i) be independent random variables defined on a common probability space Ω each distributed as $\mathcal{N}(0,1)$[9]. Define the linear operator $\Xi : L^2([0,1], \mathbb{R}) \to L^2(\Omega, \mathbb{R})$ by

$$\Xi(f) = \sum_{i=1}^{\infty} A_i \langle f, e_i \rangle$$

The operator Ξ is well-defined since $\mathbb{E}\left[\left(\sum_{i=N}^M A_i \langle f, e_i \rangle\right)^2\right] = \sum_{i=N}^M \langle f, e_i \rangle^2 \leq \|f\|_2 < \infty$, this guarantees the convergence of the sum (Cauchy sequences converge, completeness of L^2). Define the random process $B : t \mapsto \Xi(\mathbb{1}_{[0,t]})$ is our desired Brownian motion. We can show that the distribution of $B_t - B_s$ is indeed normal as follows:

$$\begin{aligned}
\mathbb{P}(B_t - B_s \leq x) &= \lim_n \mathbb{P}\left(\sum_{i=1}^n A_i \langle \mathbb{1}_{[s,t]}, e_i \rangle \leq x\right) \\
&= \lim_n \mathbb{P}\left(\mathcal{N}\left(0, \sum_{i=1}^n \langle \mathbb{1}_{[s,t]}, e_i \rangle^2\right) \leq x\right) \\
&= \lim_n \frac{1}{\sqrt{2\pi\left(\sum_{i=1}^n \langle \mathbb{1}_{[s,t]}, e_i \rangle^2\right)}} \int_{-\infty}^x \exp\left(\frac{-|y|^2}{2\sum_{i=1}^n \langle \mathbb{1}_{[s,t]}, e_i \rangle^2}\right) dy \\
&= \mathbb{P}\left(\mathcal{N}\left(0, \sum_{i=1}^{\infty} \langle \mathbb{1}_{[s,t]}, e_i \rangle^2\right) \leq x\right)
\end{aligned}$$

The first line follows because convergence in L^2 implies convergence in distribution and the last line follows from dominated convergence. However, this construction does not guarantee the continuity of B_t. We have to resort to the moments condition we saw in the previous construction.

Theorem 2.3 (Kolmogorov) *Let X_t be a random process for which there exist $\alpha > 1, \beta > 0$ that satisfy*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq |t-s|^{1+\beta}$$

then there exists a modification $\tilde{X}_t$ of the process $\tilde{X}_t$ whose paths $\tilde{X}_t(\omega)$ are continuous. Furthermore, the paths $\tilde{X}_t(\omega)$ are γ-Holder continuous for every $\gamma \in (0, \beta/\alpha)$.

3 Path Regularity

The constructions above hint towards properties of the Brownian motion. From both constructions we can see that $\mathbb{E}[|B_t - B_s|^\alpha] \leq C(\alpha)|t-s|^{\alpha/2}$. Kolmogorov theorem states that the paths of the Brownian motion are $(\alpha/2 - 1)/\alpha \approx 1/2$ for every $\alpha > 1$. It is helpful to contemplate the following theorems.

[9] These random variables represent the "imagined" evaluations we mentioned above.

Theorem 3.1 (Paley, Wiener, Zygmund) *Let $\alpha > 1/2$. Brownian motion is almost surely nowhere α-Holder continuous. In particular, almost surely nowhere differentiable.*

Theorem 3.2 (Levy)

$$\limsup_{h\to 0^+} \sup_{t\in[0,1-h]} \frac{|B_{t+h}-B_t|}{\sqrt{2h\log(1/h)}} = 1 \quad a.s.$$

Theorem 3.3 (Law of Iterated Logarithms) *For every $t \geq 0$:*

$$\limsup_{h\to 0^+} \frac{|B_{t+h}-B_t|}{\sqrt{2h\log(\log(1/h))}} = 1 = \limsup_{h\to\infty} \frac{|B_{h+t}-B_t|}{\sqrt{2h\log(\log(h))}}$$

Theorem 3.4 (Davis, slow from the right points)

$$\inf_t \limsup_{h\to 0^+} \frac{|B_{t+h}-B_t|}{\sqrt{h}} = 1 = \sup_t \liminf_{h\to 0^+} \frac{|B_{t+h}-B_t|}{\sqrt{h}}$$

Therefore, while almost surely the Brownian motion path are nowhere α-Holder for every $\alpha > 1/2$, there are almost surely Brownian motion paths which are $1/2$-Holder at least at one point. The "one point" part is important because almost surely there are no Brownian motion paths which are $1/2$-Holder everywhere. Brownian motion paths are almost surely nowhere differentiable, this can be written as

$$\begin{aligned}
&\mathbb{P}(\exists t \geq 0 \text{ such that } B_t \text{ is differentiable at } t)\\
=&\mathbb{W}(\{f : \mathscr{C}([0,\infty),\mathbb{R}^d) | \exists t \geq 0 \text{ such that } f(t) \text{ is differentiable at } t\})\\
=&0
\end{aligned}$$

The following theorem states that Wiener measure of open balls in $\mathscr{C}([0,1],\mathbb{R}^d)$ is positive.

Theorem 3.5 *Let $\varphi \in \mathscr{C}([0,1],\mathbb{R}^d)$ and $\epsilon > 0$. We have*

$$\mathbb{P}\left(\sup_{t\in[0,1]} \|B_t - \varphi(t)\| < \epsilon\right) > 0$$

Let V be a separable Banach space[10]. A random variable X on V is called Gaussian if for every bounded linear map $L : V \to \mathbb{R}$ the real valued random variable $L(X)$ is Gaussian. It is called non-degenerate if $L(X)$ is non degenerate for every $L \in V^*\backslash\{0\}$. A measure μ on V is called Gaussian if the push-forward $L_*(\mu)$ is a Gaussian measure on $\mathbb{R}$. These definitions agree with the finite dimensional definitions, furthermore they apply to Brownian motions and Wiener measure.

4 Wiener and Ito integrals

Deterministic Integrants: Let B_t be a Brownian motion and $f \in L^2([0,1],\mathbb{R})$. The two outlined constructions in section 2 hint toward defining integrals of the form $\int_0^1 f(s)dB_s$. If we impose that f is bounded and considered $f(i/n)A_i$ instead of just A_i in the first construction then the limit would be a natural candidate for the proposed integral $\int_0^1 f(s)dB_s$ (although in law). The second construction is clearer and already rigorous in this regard, the integral $\int_0^1 f(s)dB_s$ is defined simply by $\Xi(f)$ (of course the two constructions agree in law). The integral $\int_0^1 f(s)dB_s$ is called Wiener integral and it can be defined on any finite interval. An algorithmic way to define Wiener integral is by Riemann sums. We define a partition of $[0,1]$ to be a pair of sequences $P = ((t_i)_{i=0}^n, (t_i^*)_{i=0}^{n-1})$ that satisfy $0 = t_0 < t_0^* < t_1 < t_1^* < t_2 < \cdots < t_n = 1$ and define $|P| = \max_{0\leq i\leq n-1} |t_{i+1} - t_i|$. For a continuous function f on $[0,1]$ and a partition P define the corresponding Riemann sum as

$$\mathcal{R}(f,P) = \sum_{i=0}^{n-1} f(t_i^*)(B_{t_{i+1}} - B_{t_i})$$

Riemann-Stieltjes theory tells us that in order for the integral to exist for every continuous function f the integrator B_t must be of bounded variation. This is of course not the case (almost surely) for Brownian motion since absolutely continuous functions are differentiable almost everywhere. It means that that the Riemann sum above doesn't converge almost surely as we shrink the partition P arbitrarily. However, the

[10] Check Fernique's theorem

sum will converge in L^2! This is also the case in the definition using $\Xi(f)$. The steps are standard and similar to Riemann integral. For a partition P we can assume that $t_i^* = t_i$ because

$$\mathbb{E}\left[\left(\sum_{i=0}^{n-1}(f(t_i^*) - f(t_i))(B_{t_{i+1}} - B_{t_i})\right)^2\right] = \sum_{i=1}^{n-1}(f(t_i^*) - f(t_i))^2(t_{i+1} - t_i) \leq \left(\max_{|t-s|\leq|P|}|f(t) - f(s)|\right)^2$$

The quantity on the right converges to 0 as $|P| \to 0$ because a continuous function defined on a compact subset ($[0,1]$ in our case) is uniformly continuous. For two partitions $P_1 = (t_i)_{i=1}^n$ and $P_2 = (s_i)_{i=1}^m$ (notice that we ignored t^*) let $0 = u_0 < u_1 < ... < u_k = 1$ be the ordered elements of the set $\{t_i\} \cup \{s_i\}$, let $\text{ind}(i) = \max\{0 \leq j \leq i | u_j \in \{t_r\}_{r=1}^n\}$ and notice that

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{i=0}^{n-1} f(t_i)(B_{t_{i+1}} - B_{t_i}) - \sum_{i=0}^{k-1} f(u_i)(B_{u_{i+1}} - B_{u_i})\right)^2\right] \\
=&\mathbb{E}\left[\left(\sum_{i=0}^{k-1}(f(u_{\text{ind}(i)}) - f(u_i))(B_{u_{i+1}} - B_{u_i})\right)^2\right]
\end{aligned}$$

Because $|u_{\text{ind}(i)} - u_i| \leq |P_1|$ this quantity approach 0 as $|P_1| \to 0$. Similar argument can be made with P_2. This scenario is familiar to Cauchy sequences, to deal with it we take any subsequence of partitions P_n with $|P_n| \to 0$, the argument above guarantees that the sequence $(\mathcal{R}(f, P_n))_{n\geq 0}$ is Cauchy hence converges. The argument above guarantees the uniqueness of the limit because for any two sequences (P_n) and (Q_n) with $|P_n|, |Q_n| \to 0$ we can do the common refinement trick above. Now we extend this integra to L^2 functions. Let $f \in L^2([0,1], \mathbb{R}^d)$, by density of continuous functions in L^2 space there is a sequence (g_n) of continuous function such that

$$\int_{[0,1]} |f - g_n|^2 \to 0$$

Doing similar calculations to what's done above we find that $\left(\int_0^1 g_n(s)dB_s\right)_{n\geq 0}$ is a Cauchy sequence and the uniqueness of the limit follows by considering two sequences (g_n) and (h_n) and using the ideas above. Employing the definitions and the ideas used previously we have:

Theorem 4.1 (Ito Isometry) *Let $f, g \in L^2([0,1], \mathbb{R}^d)$*

$$\mathbb{E}\left[\left(\int_0^1 f(s)dB_s\right)\left(\int_0^1 g(s)dB_s\right)\right] = \int_{[0,1]} fg$$

Theorem 4.2 *Let $f \in L^2([0,1], \mathbb{R}^d)$*

$$\int_0^1 f(s)dB_s \sim \mathcal{N}\left(0, \int_{[0,1]} f^2\right)$$

For an L^2 function f the Wiener integral is define in terms of the convergence of $\mathcal{R}(f, P_n)$ in L^2 of any sequence of partitions P_n with $|P_n| \to 0$, this entails that there is a subsequence $(n_k)_{k\geq 0}$ that satisfies $\mathcal{R}(f, P_{n_k})$ converges almost surely. The impracticality is that this subsequence depends on f while for the L^2 any sequence of partitions provides convergence.

Deterministic Differentiable Integrants: If f be a differentiable (rather absolutely continuous), then the Riemann-Stieltjes exists almost surely and is given by the integration by parts formula

$$\int_0^1 f(s)dB_s = B(1)f(1) - \int_0^1 B_s f'(s)ds$$

This can be generalized to bounded variation functions but with df instead of $f'(s)ds$. We emphasize that in this case, the Ito integral can be defined as the Riemann-Stieltjes integral for almost every trajectory of the Brownian motion without the need for L^2 convergence.

Random Variables Integrants: Let A_t be a stochastic process, we would like to define $\int_0^1 A_s dB_s$. Of course, we know the recipe is to use L^2 convergence, but we quickly realize that the scenario is different here. In the definition for deterministic functions we used the simple fact that the constant random variable $f(t)$

is independent from $B_t - B_s$ for every $0 \le s < t \le 1$. This has the consequence that the choice of points t_i^*, which we previously safely ignored, affects the convergence, for example

$$\mathbb{E}\left[\sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i})\right] = 0, \quad \mathbb{E}\left[\sum_{i=0}^{n-1} B_{t_{i+1}}(B_{t_{i+1}} - B_{t_i})\right] = 1$$

It turns out that the procedure for deterministic functions above in fact works for random processes if A_s was independent of $B_t - B_s$ for $0 \le s \le t$ (more generally adapted w.r.t the filtration $\mathcal{F}$) and we always make the choice $t_i^* = t_i$. The integral defined this way is called *Ito integral.* The integral is defined by a series of approximations.

Let $A \in L^2([0,1] \times \Omega, \mathbb{R})$, that is $\mathbb{E}\left[\int_0^1 |A_s|^2 ds\right] < \infty$ and assume that A is adapted[11]. Similarly to the Wiener integral we approximate A by continuous processes, one choice is to notice that the paths $t \mapsto A_t(\omega)$ are almost surely L^2 and to define for $\epsilon, M > 0$:

$$\tilde{A}_t^{\epsilon,M} = \frac{1}{\epsilon} \int_{\min(t-\epsilon,0)}^{t} A_s \mathbb{1}_{\{|x| \le M\}}(A_s) ds$$

For almost all $\omega \in \Omega$ we can apply Lebesgue differentiation theorem $\lim_{\epsilon \to 0^+} \tilde{A}_t^{\epsilon,M}(\omega) = A_t(\omega)\mathbb{1}_{\{|x| \le M\}}(A_t)$ for almost all $t \in [0,1]$ (notice the dependence on ω). Using Dominated convergence

$$\lim_{\epsilon \to 0^+} \int_0^1 |\tilde{A}_t^{\epsilon,M}(\omega) - A_t(\omega)\mathbb{1}_{\{|x| \le M\}}(A_t)|^2 dt = 0$$

Using the dominated convergence on Ω

$$\lim_{\epsilon \to 0^+} \mathbb{E}\left[\int_0^1 |\tilde{A}_t^{\epsilon,M} - A_t \mathbb{1}_{\{|x| \le M\}}(A_t)|^2 dt\right] = 0$$

By integrability of A on $[0,1] \times \Omega$ we have

$$\lim_{M \to \infty} \mathbb{E}\left[\int_0^1 |A_t \mathbb{1}_{\{|x| \le M\}}(A_t) - A_t|^2 dt\right] = 0$$

This shows that $\tilde{A}^{\epsilon,M}$ approximates A in L^2 and finally $\tilde{A}^{\epsilon,M}$ is a bounded continuous process because of the continuity of Lebesgue integral. To define the integral for bounded continuous processes we follow the steps for Wiener integral definition and use dominated convergence in the same manner it was used a few lines above.

Theorem 4.3 *Let X_t and Y_t be L^2 bounded processes.*

- *The process $t \mapsto \int_0^t X_s dB_s$ is continuous.*
- *(Ito Isometry): $\mathbb{E}\left[\left(\int_0^1 X_s dB_s\right)\left(\int_0^1 Y_s dB_s\right)\right] = \mathbb{E}\left[\int_{[0,1]} X_s Y_s ds\right]$.*
- *(Martingale Property): For $0 \le u \le t$ we have*

$$\mathbb{E}\left[\int_0^t X_s dB_s | (B_r)_{0 \le r \le u}\right] = \int_0^u X_s dB_s$$

In particular $\mathbb{E}\left[\int_0^t X_s dB_s\right] = 0$.

Martingales is a fundamentally important notion in stochastic analysis that we are going to see again.
Example: let $0 = t_0 < \cdots < t_n = t$ be a partition of $[0,t]$. We write

$$\sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i}) = \sum_{i=0}^{n-1}\left(-\frac{1}{2}(B_{t_{i+1}} - B_{t_i})^2 + \frac{1}{2}B_{t_{i+1}}^2 - \frac{1}{2}B_{t_i}^2\right) = \frac{1}{2}B_t - \sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2$$

The quantity $\lim_n \sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2$ is called the quadratic variation of B_t and it is equal to t where it can quickly be shown that the convergence hold both a.s. and in L^2. We have

$$\int_0^t B_s dB_s = \frac{1}{2}\left(B_t^2 - t\right)$$

[11]such a process has a progressively measurable modification.

The term $-t$ appears because of the quadratic variation. If B_t was a differentiable function the quadratic variation would have been 0 and the result woul have agreed with the fundamental theorem of calculus. Symbolically we can write

$$d(B_t^2) = dt + 2B_t dB_t$$

We can further easily extend the definition to the class of processes A_t that satisfy

$$\mathbb{P}\left(\int_0^1 |A_t|^2 dt < \infty\right) = 1$$

but at the cost of losing some of the nice properties we just described such as being martingale. For this generalization, the previous approximations won't work. Let $M > 0$ and define

$$\tilde{A}_t^M = A_t \mathbb{1}_{\{|x| \leq M\}} \left(\int_0^t |A_s|^2 ds\right)$$

The process $\tilde{A}_t^M \in L^2([0,1] \times \Omega, \mathbb{R})$, thus we can define the integral of $\int_0^1 A_s dB_s$ as the $\lim_M \int_0^1 \tilde{A}_t^M dB_s$. However, the convergence is in probability! This integral isn't necessarily martingale nor does it have mean 0 necessarily. This integral is an example of *local martingales*. We have the following estimate:

Theorem 4.4 (Lenglart)

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} \left|\int_0^s A_u dB_u\right| \geq \epsilon\right) \leq \mathbb{P}\left(\int_0^t |A_s|^2 ds \geq \delta\right) + \frac{\delta}{\epsilon^2}$$

Stratonovich Integral: For a process A_s we define the Stratonovich integral as the limit of

$$\sum_{i=0}^{n-1} \left(\frac{A_{t_i} + A_{t_{i+1}}}{2}\right) (B_{t_{i+1}} - B_{t_i})$$

If the Stratonovich integral exists we denote it by

$$\int_0^t A_s \circ dB_s$$

The relation between the two integrals is

$$\int_0^t A_s \circ dB_s = \int_0^t A_s dB_s + \frac{1}{2}[A, B]_t$$

where the quadratic covariation $[A, B]_t$ will be defined below. The Ito integral has the "martingale property" which is a central topic in the theory of probability. The Stratonovich isn't a martingale but it enjoys a behaviour like standard calculus (Check Ito formulat below). For example we have

$$\int_0^t B_s \circ dB_s = \frac{1}{2} B_t^2$$

Ito Processes: Following the example above, we call a process X_t an Ito process if there exist stochastic processes μ_t and σ_t that satisfy $\int_0^1 (|\mu_t| + \sigma_t^2) dt < \infty$ a.s. and

$$dX_t = \mu_t dt + \sigma_t dB_t$$

i.e.

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$$

Such processes belong to the class of *semimartingales*. We can further define the Ito integral $\int_0^t A_s dX_s$ and prove that

$$\int_0^t A_s dX_s = \int_0^t A_s \mu_s ds + \int_0^t A_s \sigma_s dB_s$$

for processes A_s that satisfy

$$\int_0^t |A_s \sigma_s|^2 ds < \infty \quad a.s. \tag{2}$$

Such processes A are called X-integrable.

5 The process dB_t?

Although I think a mathematical write up shouldn't be too sloppy, I will exclude this section from this requirement because I feel it discusses natural ideas that are worth showing. Ito integral gives us a meaning to the "time derivative "symbol dB_t, it is whatever you integrate to get the process B_t. At this level this description is symbolic, in the sense that we can always make things formal again by writing the statements as integrals, but are used conventionally because they are short and convenient. This is probably similar to the usage of δ as a "function" and the usual notation in which Radon-Nykodyn theorem is stated. In [7] the author suggests using the hyperreal numbers to define dB_t as the finite dimensional hyperreal valued stochastic process

$$dB_t = \begin{cases} \sqrt{dt} & \text{with probability } 1/2 \\ -\sqrt{dt} & \text{with probability } 1/2 \end{cases}$$

In [2] they regard $\dot{B}_t$ as a generalized-function valued process defined by its "characteristic function" on $(\mathscr{C}([0,1], \mathbb{R}, L^2?)$ that is given by

$$C_{\dot{B}_t}(f) = \exp\left(-\|f\|_2^2/2\right)$$

6 Quadratic Variation

For a stochastic process X_t and Y_t we define their quadratic covariation process $[X, Y]_t$ as the limit in probability of

$$\sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i})$$

as the partition $0 = t_0 < \cdots < t_n = t$ diameter diminishes. We define the quadratic variation $[X]_t = [X, X]_t$. We have encountered $[B]_t$ where B is Brownian motion and argued that $[B]_t = t$. We can write the quadratic variation of a process X_t as

$$\sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 = \sum_{i=0}^{n-1} \left(X_{t_{i+1}}^2 - X_{t_i}^2 - 2X_{t_i}(X_{t_{i+1}} - X_{t_i}) \right) \xrightarrow{p} X_t^2 - X_0^2 - 2\int_0^t X_s dX_s$$

Thus, once we show that X is an X-integrable process we have proven that the quadratic variation exists for any Ito process. This amounts to verifying condition (2) which is direct because X is continuous. Furthermore, we can show, using typical approximations we have used previously, that

$$[X]_t = \int_0^t |\sigma_s|^2 ds$$

Notice that $[X]_t$ is a bounded variation (increasing) process almost surely! and that it is independent of μ. We write the relation between X and $[X]$ as

$$d(X_t^2) = 2X_t dX_t + d[X]_t$$

Consequently the square of an Ito process is an Ito process, and the integrability condition (2) can be rewritten as

$$\int_0^t |A_s|^2 d[X]_s < \infty \quad a.s.$$

We furthermore have for every X-integrable process A:

$$\left[\int_0^t A_s dX_s \right] = \int_0^t A_s^2 d[X]_s$$

For a continuous process A we have

$$\sum_{i=0}^{n-1} A_{t_i^*}(X_{t_{i+1}} - X_{t_i})^2 \xrightarrow{p} \int_0^t A_s d[X]_s$$

Since $[X, Y]_t = \frac{1}{4}\left([X + Y, X + Y] - [X - Y, X - Y]\right)$, the following integration by parts formula follows

$$\int_0^t X_s dY_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s dX_s - [X, Y]_t$$

Which can be written as

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d[X, Y]_t$$

Occasionally the term $d[X, Y]_t$ is written $dX_t dY_t$. This is an instance of Ito formula, which we now discuss.

7 Ito Formula

Ito formula is the stochastic analogue of the fundamental theorem of calculus. It allows us to compute stochastic integrals and helps us to understand the "local behaviour" of stochastic processes by expressing them as integrals (formally differentiating them). Let X_t be an Ito process and f be a $\mathscr{C}^2([0,1]\times\mathbb{R}^d,\mathbb{R})$ function. We will expand $f(t,X_t)-f(s,X_s)$ using Taylor theorem and in order to account for the "first order terms" we need to write Taylor expansion till degree 2 (since formally $\mathbb{E}[(dB_t)^2]=dt$).

$$\begin{aligned}
f(t,X_t) &= \sum_{i=0}^{n-1} f(t_{i+1},X_{t_{i+1}})-f(t_i,X_{t_i})\\
&= \sum_{i=0}^{n-1}\Big(\partial_x f(t_i,X_{t_i})(X_{t_{i+1}}-X_{t_i})+\partial_s f(t_i,X_{t_i})(t_{i+1}-t_i)\\
&+\frac{1}{2}\partial_x^2 f(t_i,X_{t_i})(X_{t_{i+1}}-X_{t_i})^2+\frac{1}{2}\partial_x\partial_s f(t_i,X_{t_i})(X_{t_{i+1}}-X_{t_i})(t_{i+1}-t_i)+\frac{1}{2}\partial_s^2 f(t_i,X_{t_i})(t_{i+1}-t_i)^2\\
&+R(t_{i+1}-t_i,X_{t_{i+1}}-X_{t_i})\Big)
\end{aligned}$$

where the remained satisfies $|R(t,x)|/\left(|t|^2+|x|^2\right)\to 0$ as $t\to 0, x\to 0$. The sum of the first term converges in probability to an Ito integral, the sum of the second term converges to a Riemann integral, and the sum of the third term converges to a Riemann-Stieltjes integral with integrator $[X]_t$ (see the Quadratic variation section). Each of the sums of the fourth and fifth terms converge to 0 almost surely, for example for every $\omega\in\Omega$ the sum of the fourth term is bounded by

$$\frac{1}{2}\max_{0\le i\le n}|X_{t_{i+1}}(\omega)-X_{t_i}(\omega)|\sum_{i=0}^{n-1}|\partial_x\partial_s f(t_i,X_{t_i}(\omega))|(t_{i+1}-t_i)\to 0\times\int_0^t|\partial_x\partial_s f(t_i,X_{t_i}(\omega))|dt$$

Similarly the sum of the fifth term converges almost surely to 0. For the sum of the six term, first notice that the limit exists in probability since the limit exist for all the other terms in the equation. Let P_j be a sequence of partitions that makes $\sum_{i=0}^{n_j-1}(X_{t^j_{i+1}}-X_{t^j_i})^2$ converges to $[X]_t$ almost surely and let $S=\{\omega\in\omega|\sum_{i=0}^{n_j-1}(X_{t^j_{i+1}}(\omega)-X_{t^j_i}(\omega))^2\xrightarrow[j\to\infty]{}[X]_t(\omega)\}$. Let $\omega\in S$ and $\epsilon>0$, let $\delta>0$ such that $|R(t,x)|<\epsilon(|t|^2+|x|^2)$ whenever $|t|^2+|x|^2<\delta$, let J large such that P_J is fine enough so that $\max_{0\le i\le n-1}|X_{t^j_{i+1}}(\omega)-X_{t^j_i}(\omega)|<\delta$ for all $j>J$. Then for every $j>J$:

$$\left|\sum_{i=0}^{n_j-1}R(t^j_{i+1}-t^j_i,X_{t^j_{i+1}}(\omega)-X_{t^j_i}(\omega))\right|\le\epsilon\sum_{i=0}^{n_j-1}\left((t^j_{i+1}-t^j_i)^2+(X_{t^j_{i+1}}(\omega)-X_{t^j_i}(\omega))^2\right)\xrightarrow[j\to\infty]{a.s.}\epsilon[X]_t(\omega)$$

since ϵ is arbitrary the limit of the sum of the six term converges to 0 in probability. Thus we have proven the Ito lemma:

$$df(t,X_t)=\partial_t f(t,X_t)dt+\partial_x f(t,X_t)dX_t+\frac{1}{2}\partial_x^2 f(t,X_t)d[X]_t$$

8 Path Integral

We have a measure $\mathbb{W}$ over the space of continuous functions, it is tempting to use it with the Lebesgue integral.

Theorem 8.1 (Feynman-Kac) *Let $x_0\in\mathbb{R}$, $u\in\mathscr{C}^2([0,T]\times\mathbb{R},\mathbb{R})$ be the solution of the heat equation (imaginary time Schrodinger equation)*

$$\begin{cases}\partial_t u=\frac{1}{2}\partial_x^2 u-V(x)u\\ u(0^+,x)=\delta(x-x_0)\end{cases}$$

with V bounded below and continuous. Let B be a Brownian motion, we have:

$$\begin{aligned}
u(t,x)=&\frac{e^{-\frac{(x_0-x)^2}{2t}}}{(2\pi t)^{1/2}}\quad\mathbb{E}\left[e^{-\int_0^t V(x_0+B_s)ds}\middle|B_t=x-x_0\right]\\
=&\frac{e^{-\frac{(x_0-x)^2}{2t}}}{(2\pi t)^{1/2}}\quad\int_{\mathscr{C}^2([0,t],\mathbb{R})}e^{-\int_0^t V(x_0+s(x-x_0)/t+(1-s/t)f(s))ds}d\mathbb{W}(f)
\end{aligned}$$

which is "equivalent" to

$$\int_{\mathbb{R}} u(t,x)F(x)dx = \mathbb{E}\left[e^{-\int_0^t V(x_0+B_s)ds}F(x_0+B_t)\right]$$

$$= \int_{\mathscr{C}^2([0,t],\mathbb{R})} e^{-\int_0^t V(x_0+f(s))ds}F(x_0+f(t))d\mathbb{W}(f)$$

This is not the usual form in which Feynman-Kac formula is presented, this form is similar to the form Kac presented in [4]. In fact this theorem can be used to show the existence of a solution of the heat equation above. The proof is an application of Ito lemma. How is it stated in the distributional sense?

9 Girsanov Theorem: Change of Measure

Let B be a Brownian motion. We have seen that $[B]_t = t$, this has the consequence that the laws $\mathcal{L}_B$ and $\mathcal{L}_{2B}$ of B_t and $2B_t$ on $\{f \in \mathscr{C}([0,1],\mathbb{R})$ are not absolutely continuous w.r.t each other. More precisely the sets $S_1 = \{f \in \mathscr{C}([0,1],\mathbb{R}) | [f]_t = t\}$ and $S_2 = \{f \in \mathscr{C}([0,1],\mathbb{R}) | [f]_t = 4t\}$ satisfy $\mathcal{L}_B(S_1) = 1 > 0 = \mathcal{L}_{2B}(S_1)$ and $\mathcal{L}_{2B}(S_2) = 1 > 0 = \mathcal{L}_B(S_2)$. In general two Ito processes X_t and Y_t with $\mathcal{L}_{[X]_t} \not\cong \mathcal{L}_{[Y]_t}$ for any t have laws that are not absolutely continuous. For the case when the drifts are different we have the following theorem

Theorem 9.1 (Girsanov) *Let $dX_t = \mu_t dt + \sigma_t dB_t$ and $Y_t = \sigma_t dB_t$. The laws of the processes X and Y are equivalent (each law absolutely continuous w.r.t to the other) given that the process*

$$Z_t = \exp\left(\int_0^t \frac{\mu_s}{\sigma_s}dB_s - \frac{1}{2}\int_0^t \left(\frac{\mu_s}{\sigma_s}\right)^2 ds\right)$$

is martingale. Furthermore, we have that for every $F : \mathscr{C}([0,1],\mathbb{R}) \to \mathbb{R}$

$$\mathbb{E}[F(X)] = \mathbb{E}[Z_1 F(Y)]$$

A condition that guarantees the martingale property og the process Z_t is the Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \left(\frac{\mu_s}{\sigma_s}\right)^2 ds\right)\right] < \infty$$

and the process Z_t is called exponential martingale because it satisfies the equation $dZ_t = \left(\frac{\mu_s}{\sigma_s}\right) Z_t dB_t$. In the statement above we didn't write the explicit Radon-Nikoym derivative of the laws of X and Y because it is messy. However, the process Z is a Radon-Nikoydm derivative in the following equivalent restatement of Girsanov's theorem. Recall that the Brownian motion notion depends on the underlying proability measure $\mathbb{P}$.

Theorem 9.2 (Girsanov) *Let B be a Brownian motion defined on a probability space Ω with measure $\mathbb{P}$. Let X, Y and Z as in Theorem 9.1. Define the measure $\mathbb{Q}$ by the relation*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_1$$

Then the the process $\tilde{B}$ defined by

$$\tilde{B}_t = \int_0^t \frac{\mu_s}{\sigma_s}ds + B_t$$

is a $\mathbb{Q}$ Bronwian motion, and the process X satisfies

$$dX_t = \sigma_t d\tilde{B}_t$$

The case $\sigma = 1$ is important as it states that any process of the form $\int_0^t \mu_s ds + B_t$ is a Brownian motion under the right measure and this right measure is absolutely continuous with the underlying measure $\mathbb{P}$. This is equivalent to saying that the laws of B_t and $\int_0^t \mu_s ds + B_t$ are equivalent. Furthermore, the law of X under $\mathbb{Q}$ is the same law of Y under $\mathbb{P}$.

Let's remark that $[X]_t = [Y]_t$ is not enough to guarantee that $\mathcal{L}_X \cong \mathcal{L}_Y$. For example, the processes $dX_t = B_t dB_t$ and $dY_t = -B_t dB_t$ have different distributions since X_t is lower bounded while Y is upper bounded. However, we have a result for a special case, let's write (W.L.O.G) $dX_t = \sigma_t^1 dB_t$ and $dY_t = \sigma_t^2 dB_t$ with the processes σ^1, σ^2 being deterministic and continuous. By the argument at the beginning of the section, for the processes X and Y to have equivalent laws we must have $\int_0^t (\sigma_s^1)^2 ds = \int_0^t (\sigma_s^2)^2 ds$ for all $t \in [0,1]$. Because of continuity $(\sigma_t^1)^2 = (\sigma_t^2)^2$ holds for every $t \in [0,1]$. Define the process $\widetilde{B}_t = \int_0^t \operatorname{sign}(\sigma_t^1 \sigma_t^2) dB_t$. Then $dY_t = \sigma_t^1 d\widetilde{B}_t$. If we prove that the process $\widetilde{B}$ is a Brownian motion (w.r.t to the filtration) then we would have shown that X and Y have the same laws. This is a consequence of *Levy's characterization of Brownian motion*. Another question we may think about is the following: Is there a function f such that $f(X_t, t)$ has a law equivalent to the Wiener measure? The answer is yes if $\sigma_t > 0$ and it is known as the *Lamperti transform* where the process $Y_t = \int_0^{X_t} \frac{1}{\sigma(x,t)} dx$ has a unit diffusion coefficient, this can be easily shown by Ito formula.

10 Change of Time

Let $f : [0,1] \to \mathbb{R}^+ \cup \{0\}$ be an increasing function. Let's study the process $B_{f(t)}$. For $0 \leq s < t \leq 1$ we have $\mathbb{E}[(B_{f(t)} - B_{f(s)})^2] = f(t) - f(s)$ and considering the L^2 limit for the quadratic variation we can prove that $[B_{f(\cdot)}]_t = f'(t)$. Thus we would hope that there exists a Brownian motion $\widetilde{B}_t$ that satisfies $B_{f(t)} = \int_0^t \sqrt{f'(s)} d\widetilde{B}_s$, which follows from Levi's characterization. That has the formal consequence that $dB_{f(t)}$ has the same law as $\sqrt{f'(t)} dB_t$ and the following general theorem.

Theorem 10.1 *(Dubins–Schwarz) Let X_t be an Ito process with 0 drift (i.e. $\int_0^t \sigma_s dB_s$). The processes X_t and $B_{[X]_t}$ have the same law.*

11 Generalizations

We have cut ourselves a lot of slack, many of the concepts have direct generalizations, martingales, local martingales, and semimartingales. We have implicitly used local times in the definition of Ito integral. We haven't used the distribution of the Brownian motion "directly".

References

[1] Davis, B. (1983). On Brownian slow points. Z. Wahrscheinlichkeitstheorie verw Gebiete 64, 359–367.

[2] Hida, T. (2003). White Noise Analysis: Part I. Theory in Progress. Taiwanese Journal of Mathematics, 7(4), 541–556.

[3] Ikeda, N., and Watanabe, S. (1981). Stochastic Differential Equations and Diffusion Processes.

[4] Kac, M. (1966). Wiener and Integration in Function Spaces.

[5] Krylov, N.V. (2002). Introduction to the Theory of Random Processes.

[6] Mörters, P., and Peres, Y. (2010). Brownian Motion. Cambridge: Cambridge University Press.

[7] Nelson, E. (1987). Radically Elementary Probability Theory.

[8] Øksendal, B. (2003). Stochastic Differential Equations.