AN OVERVIEW OF STOCHASTIC CALCULUS

1. Brownian motion

I will discuss Brownian motion briefly, leaving many interesting and technical aspects unexplored, to be taken up in future sessions. A Brownian motion is a continuous-time stochastic process, i.e., a family of random variables $\{B_t\}_{t\in[0,1]}$ (say) on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the following properties:

- $B_0 = 0$ almost surely.
- For all $0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant$, the increments $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent.
- For all $0 \leqslant s < t$, the increment $B_t - B_s$ is normally distributed with mean 0 and variance $t - s$.
- The sample paths of B are continuous almost surely.

The last point requires some clarification. Whenever we say some event happens almost surely, we *do not mean* that the set of such ω has probability 1. That would require the set of such ω to be *measurable*, which can be hairy to ensure given the uncountable nature of things. Instead, we mean that there is *some* set of ω with probability 1 on which the event happens. This is a non-issue if your probability space is *complete*, i.e., subsets of null sets are null, but completions can have other issues.

In any case, it is easiest to trust that there is a suitable construction of a probability space hosting a Brownian motion (I leave details for future speakers), and examine some consequences, but let us discuss the notion of filtration briefly. A filtration is a collection of nested σ-algebras (subsets of the global σ-algebra $\mathcal{F}$) $\{\mathcal{F}_t\}_t$ such that $\mathcal{F}_t \subseteq \mathcal{F}_s$ for $t \leqslant s$. *The idea is that $\mathcal{F}_t$ contains all the information available up to time t.* We say a process X is adapted to a filtration if $X_t \in \mathcal{F}_t$ (by which we mean that X_t is $\mathcal{F}_t$-measurable).

For Brownian motion, a natural choice is $\mathcal{F}_t = \sigma(B_s : s \leqslant t)$, the σ-algebra generated by the process up to time t. By this choice, $\mathcal{F}_0 = \{\varnothing, \Omega\}$ is the trivial algebra. Of course, any event measurable with respect to $\mathcal{F}_0$ has probability 0 or 1. But a *much more interesting* result is

m:blumenthal **Theorem 1.1** (Blumenthal's 0-1 law)**.** *For any event $A \in \mathcal{F}_0^+$, we have $\mathbb{P}(A) \in \{0, 1\}$.*

What is $\mathcal{F}_0^+$? It is every event that is known at time $\mathcal{F}_\varepsilon^+$, no matter how small $\varepsilon > 0$ is, i.e.,

$$\mathcal{F}_0^+ = \bigcap_{\varepsilon>0} \mathcal{F}_\varepsilon.$$

As an example of an event in $\mathcal{F}_0^+$ is the event that Brownian motion is differentiable at time 0. Another example is the event that

$$\bigcap_{\varepsilon>0} \{\exists 0 < t < \varepsilon : B_t = 0\},$$

that is, the Brownian motion hits zero at arbitrarily small positive times. By Blumenthal's law, both of these have probability either 0 or 1. Deciding which one is a more involved question (and perhaps a topic for future talks?)

Two interesting ideas here:

- Donsker invariance principle, stating the scaling limits of random walks to Brownian motion.

- The Komlos-Major-Tusnady theorem, a refinement of Donsker, asserting a very strong L^∞ coupling result between random walks and Brownian motions.

2. Stochastic calculus

With that brief appetizer, let us turn our attention to stochastic calculus, the focus of (half of) this seminar. Recall the usual Stieltjes integral (on $[0,1]$)

$$\int F_t dG_t := \lim \sum F_{s_i}(G_{t_{i+1}} - G_{t_i}),$$

over partitions $0 = t_0 < t_1 < \cdots < t_n = t$ of $[0,1]$, where s_i is any choice $\in [t_i, t_{i+1}]$. Comparing this with the choice $s_i = t_i$, the difference is

$$\sum (F_{s_i} - F_{t_i})(G_{t_{i+1}} - G_{t_i}),$$

with absolute value at most

$$\sum |F_{s_i} - F_{t_i}| \left|G_{t_{i+1}} - G_{t_i}\right|.$$

We say that the Stieltjes integral is well-defined only when this latter quantity goes to zero as the mesh-size goes to 0. For instance if the Holder exponents of F is p and of G is q, the last quantity is bounded by

$$\sum |s_i - t_i|^p |t_{i+1} - t_i|^q \leqslant \sum |t_{i+1} - t_i|^{p+q} \to 0,$$

as long as $p + q > 1$. To see why we require arbitrary mesh-points to work is that we want to be able to reason like

$$1 \cdot d(f(G_t)) = f'(G_t) dG_t$$

i.e. have a chain rule. If we chose our endpoint to be always t_i, we would need

$$\sum f(G_{t_{i+1}}) - f(G_{t_i}) \asymp \sum f'(G_{t_i}) \left(G_{t_{i+1}} - G_{t_i}\right).$$

But the way this goes is that we expand

$$f(G_{t_{i+1}}) - f(G_{t_i}) = f'(G_{s_i})(G_{t_{i+1}} - G_{t_i}) + \text{smaller corrections}.$$

for some $s_i \in [t_i, t_{i+1}]$. Thus, having arbitrary mesh-points would imply a chain-rule (or a substitution principle) for our integral.

Unfortunately, Brownian motion is only $\frac{1}{2}^-$ Holder (i.e., p-Holder for every $p < 1/2$), so the Stieltjes integral is not well-defined (only barely though!). Actually, for Brownian motion, the problem is worse. Even if you fix the choice of $s_i = t_i$, the integral does not exist in the sense that the limit does not exist almost surely. Let us see an example:

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

due to the identity

$$a(b-a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(b-a)^2.$$

The former telescopes to

$$\frac{1}{2}(B_1^2 - B_0^2) = \frac{1}{2}B_1^2,$$

so we are left with understanding

$$\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2.$$

The limit of this quantity, as the mesh size goes to 0, does not exist. In fact, it is known that

$$\sup_P \sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2 = \infty,$$

where the supremum is taken over all partitions of $[0, 1]$. Instead, this quantity converges in L^2 to t, that is, for any sequence of partitions P_n with mesh-size going to 0, we have

$$\lim_{n\to\infty} \sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2 = t.$$

We usually write this as $(dB_t)^2 = dt$. This is the starting point of the Ito integral, which is defined as

$$\int F_t dB_t = \lim \sum F_{t_i}(B_{t_{i+1}} - B_{t_i}),$$

where the limit is in L^2 sense (actually, to extend to a wider class of integrands, the limits may need to be taken in the weaker senses). The choice of the left endpoint also makes the integral *process* a martingale, which opens up the doors to a lot of martingale theory, some of which will (probably) be discussed in future talks.

A couple of subtopics here:

- The Girsanov theorem, about how shifts of Brownian motion by a deterministic function can be made into Brownian motions under a different measure.
- The martingale representation theorem, which says that any square-integrable martingale (with some mild measurability conditions) can be represented as an Ito integral with respect to Brownian motion.
- Some continuous martingale theory, semimartingales, and how to define the stochastic integral more broadly.
- Local times (specifically for Brownian motion or for more general processes), Tanaka's formula.

The second subtopic can be expanded into multiple talks, and is related to a lot of modern research, because of its connections to the Malliavin calculus, and in general, chaos expansions of random variables.

3. Stochastic differential equations

Once we define the stochastic integral, one can start make sense of expressions of the form

$$dX_t = a(t, X_t)dt + \sigma(t, X_t)dB_t,$$

as a process X satisfying

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

A famous example is the Black-Scholes equation, which is a stochastic differential equation for the price of a financial derivative. The equation is

$$dS_t = \mu S_t dt + \sigma S_t dB_t.$$

Let us discuss the Ito formula, which is a generalization of the chain rule to stochastic integrals (which, as we saw, is not for free as in the Stieltjes case). Basically, the observation is that to express $f(B_t)$ as an SDE, let us split this up into a telescoping sum as follows:

$$\begin{aligned} f(B_1) - f(B_0) &= \sum_{i=0}^{n-1} f(B_{t_{i+1}}) - f(B_{t_i}) \\ &= \sum_{i=0}^{n-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) + \frac{1}{2}\sum_{i=0}^{n-1} f''(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2 + \frac{1}{6}\sum_{i=0}^{n-1} f'''(B_{s_i})(B_{t_{i+1}} - B_{t_i})^3, \end{aligned}$$

for some choices s_i (which are random). The proof of the L^2 convergence of the quadratic variation can be extended to show

$$\left\| \sum_{i=0}^{n-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}) - \int_0^1 f'(B_t)dB_t \right\|_2 \to 0.$$

The first term converges in L^2 to the Ito integral by definition. If f''' bounded the last term can be shown to go to 0 in L^2 as well. This leads to the Ito formula, which written differentially, states:

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

There are several subtopics here:

- Existence and uniqueness of solutions, the notions of strong and weak solutions etc.
- Algorithms to compute these solutions, starting with Euler-Maruyama etc.
- The notion of a general Feller process (which is a general class of Markov processes), generators, semigroups, resolvents and associated functional analysis theory.
- The PDE-side of things, i.e., the Fokker-Planck equation, the Kolmogorov forward and backward equations, etc. which talk about the evolution of the density of the process.
- The reverse problem, i.e., given a PDE (of a specific type), how to interpret its solution via a stochastic differential equation. This is the Feynman-Kac formula.

Beyond these, there are some broader topics which can be the subject of talks:

- Scaling limits. How do smoother or discrete processes scale to Ito processes? A central result here is the Wong-Zakai theorem, which writes the solution to a SDE as a limit of solutions to a sequence of ordinary differential equations with smoothened noise. This idea comes up repeatedly in the modern theory of stochastic PDEs.
- Large deviations of SDEs, in particular the Freidlin-Wentzell theory, which gives a large deviation principle for the paths of a SDE. Talking about this might involve reviewing the basic theory of large deviations (which can be done heuristically, since we have seen this a couple of times in the recent past).
- Maybe a more applied talk, like applications of the theory to pricing options, in particular, risk-neutral pricing. Almost all the tools we will discuss have interesting uses there.
- Finally, a very modern direction would be to relate stochastic calculus to rough-paths theory, which defines a pathwise theory of integration, even for paths rougher than Holder 1/2, via introducing a certain second-order process that encodes the iterated integral of a path against itself. This

is particularly relevant considering the recent interest in stochastic PDEs and Hairer's theory of regularity structures.