Lectures on Lévy Processes, Stochastic Calculus and Financial Applications, Ovronnaz September 2005

2 Lecture 2: Sample Paths, Jumps and Stochastic Integration

2.1 Càdlàg Paths

A function $f : \mathbb{R}^+ \to \mathbb{R}^d$ is *càdlàg* if it is *continue à droite et limité à gauche*, i.e. right continuous with left limits. Such a function has only jump discontinuities.

Define $f(t-) = \lim_{s \uparrow t} f(s)$ and $\Delta f(t) = f(t) - f(t-)$. If f is càdlàg, $\{0 \leq t \leq T, \Delta f(t) \neq 0\}$ is at most countable.

If the filtration satisfies the "usual hypotheses" of right continuity and completion, then every Lévy process has a càdlàg modification which is itself a Lévy process.

From now on, we will always make the following assumptions:-

- $(\Omega, \mathcal{F}, P)$ will be a fixed probability space equipped with a filtration $(\mathcal{F}_t, t \geq 0)$ which satisfies the "usual hypotheses".
- Every Lévy process $X = (X(t), t \geq 0)$ will be assumed to be $\mathcal{F}_t$-adapted and have càdlàg sample paths.
- $X(t) - X(s)$ is independent of $\mathcal{F}_s$ for all $0 \leq s < t < \infty$.

2.2 The Jumps of A Lévy Process - Poisson Random Measures

The *jump process* $\Delta X = (\Delta X(t), t \geq 0)$ associated to a Lévy process is defined by

$$\Delta X(t) = X(t) - X(t-),$$

for each $t \geq 0$.

Theorem 2.1 *If N is a Lévy process which is increasing (a.s.) and is such that $(\Delta N(t), t \geq 0)$ takes values in $\{0, 1\}$, then N is a Poisson process.*

Proof. Define a sequence of stopping times recursively by $T_0 = 0$ and $T_n = \inf\{t > T_{n-1}; N(t + T_{n-1}) - N(T_{n-1})) \neq 0\}$ for each $n \in \mathbb{N}$. It follows from (L2) that the sequence $(T_1, T_2 - T_1, \ldots, T_n - T_{n-1}, \ldots)$ is i.i.d.

By (L2) again, we have for each $s, t \geq 0$,

$$\begin{aligned} P(T_1 > s + t) &= P(N(s) = 0, N(t + s) - N(s) = 0) \\ &= P(T_1 > s)P(T_1 > t) \end{aligned}$$

From the fact that N is increasing (a.s.), it follows easily that the map $t \to P(T_1 > t)$ is decreasing and by a straightforward application of stochastic continuity (L3) we find that the map $t \to P(T_1 > t)$ is continuous at $t = 0$. Hence there exists $\lambda > 0$ such that $P(T_1 > t) = e^{-\lambda t}$ for each $t \geq 0$. So T_1 has an exponential distribution with parameter λ and

$$P(N(t) = 0) = P(T_1 > t) = e^{-\lambda t},$$

for each $t \geq 0$.

Now assume as an inductive hypothesis that $P(N(t) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$, then

$$P(N(t) = n + 1) = P(T_{n+2} > t, T_{n+1} \leq t) = P(T_{n+2} > t) - P(T_{n+1} > t).$$

$$\text{But } T_{n+1} = T_1 + (T_2 - T_1) + \cdots + (T_{n+1} - T_n)$$

is the sum of $(n+1)$ i.i.d. exponential random variables, and so has a gamma distribution with density $f_{T_{n+1}}(s) = e^{-\lambda s}\dfrac{\lambda^{n+1}s^n}{n!}$ for $s > 0$.

The required result follows on integration. □

The following result shows that ΔX is not a straightforward process to analyse.

Lemma 2.1 *If X is a Lévy process, then for fixed $t > 0, \Delta X(t) = 0$ (a.s.).*

Proof. Let $(t(n), n \in \mathbb{N})$ be a sequence in $\mathbb{R}^+$ with $t(n) \uparrow t$ as $n \to \infty$, then since X has càdlàg paths, $\lim_{n\to\infty} X(t(n)) = X(t-)$. However, by (L3) the sequence $(X(t(n)), n \in \mathbb{N})$ converges in probability to $X(t)$, and so has a subsequence which converges almost surely to $X(t)$. The result follows by uniqueness of limits. □

Much of the analytic difficulty in manipulating Lévy processes arises from the fact that it is possible for them to have

$$\sum_{0\leq s\leq t} |\Delta X(s)| = \infty \quad \text{a.s.}$$

and the way in which these difficulties is overcome exploits the fact that we always have

$$\sum_{0\leq s\leq t} |\Delta X(s)|^2 < \infty \quad \text{a.s.}$$

We will gain more insight into these ideas as the discussion progresses.

Rather than exploring ΔX itself further, we will find it more profitable to count jumps of specified size. More precisely, let $0 \leq t < \infty$ and $A \in \mathcal{B}(\mathbb{R}^d - \{0\})$. Define

$$\begin{aligned} N(t,A) &= \#\{0 \leq s \leq t; \Delta X(s) \in A\} \\ &= \sum_{0\leq s\leq t} \chi_A(\Delta X(s)). \end{aligned}$$

Note that for each $\omega \in \Omega, t \geq 0$, the set function $A \to N(t,A)(\omega)$ is a counting measure on $\mathcal{B}(\mathbb{R}^d - \{0\})$ and hence

$$\mathbb{E}(N(t,A)) = \int N(t,A)(\omega)dP(\omega)$$

is a Borel measure on $\mathcal{B}(\mathbb{R}^d - \{0\})$. We write $\mu(\cdot) = \mathbb{E}(N(1,\cdot))$.

We say that $A \in \mathcal{B}(\mathbb{R}^d - \{0\})$ is *bounded below* if $0 \notin \bar{A}$.

Lemma 2.2 *If A is bounded below, then $N(t,A) < \infty$ (a.s.) for all $t \geq 0$.*

Proof. Define a sequence of stopping times $(T_n^A, n \in \mathbb{N})$ by $T_1^A = \inf\{t > 0; \Delta X(t) \in A\}$, and for $n > 1, T_n^A = \inf\{t > T_{n-1}^A; \Delta X(t) \in A\}$. Since X has càdlàg paths, we have $T_1^A > 0$ (a.s.) and $\lim_{n\to\infty} T_n^A = \infty$ (a.s.). Indeed if either of these were not the case, then the set of all jumps in A would have an accumulation point, and this is not possible if X is càdlàg. Hence, for each $t \geq 0$,

$$N(t,A) = \sum_{n\in\mathbb{N}} \chi_{\{T_n^A \leq t\}} < \infty \text{ a.s.} \qquad \square$$

Be aware that if A fails to be bounded below, then Lemma 2.2 may no longer hold, because of the accumulation of large numbers of small jumps.

The following result should at least be plausible, given Theorem 2.1 and Lemma 2.2.

Theorem 2.2 *1. If A is bounded below, then* $(N(t,A), t \geq 0)$ *is a Poisson process with intensity* $\mu(A)$.

2. If $A_1, \ldots, A_m \in \mathcal{B}(\mathbb{R}^d - \{0\})$ *are disjoint, then the random variables* $N(t, A_1), \ldots, N(t, A_m)$ *are independent.*

It follows immediately that $\mu(A) < \infty$ whenever A is bounded below, hence the measure μ is σ-finite.

The main properties of N, which we will use extensively in the sequel, are summarised below:-.

1. For each $t > 0, \omega \in \Omega, N(t, \cdot)(\omega)$ is a counting measure on $\mathcal{B}(\mathbb{R}^d - \{0\})$.
2. For each A bounded below, $(N(t,A), t \geq 0)$ is a Poisson process with intensity $\mu(A) = \mathbb{E}(N(1,A))$.
3. The *compensator* $(\tilde{N}(t,A), t \geq 0)$ is a martingale-valued measure where $\tilde{N}(t,A) = N(t,A) - t\mu(A)$, for A bounded below, i.e.

   For fixed A bounded below, $(\tilde{N}(t,A), t \geq 0)$ is a martingale.

2.3 Poisson Integration

Let f be a Borel measurable function from $\mathbb{R}^d$ to $\mathbb{R}^d$ and let A be bounded below then for each $t > 0, \omega \in \Omega$, we may define the *Poisson integral* of f as a random finite sum by

$$\int_A f(x) N(t, dx)(\omega) = \sum_{x \in A} f(x) N(t, \{x\})(\omega).$$

Note that each $\int_A f(x) N(t, dx)$ is an $\mathbb{R}^d$-valued random variable and gives rise to a càdlàg stochastic process, as we vary t.

Now since $N(t, \{x\}) \neq 0 \Leftrightarrow \Delta X(u) = x$ for at least one $0 \leq u \leq t$, we have

$$\int_A f(x) N(t, dx) = \sum_{0 \leq u \leq t} f(\Delta X(u)) \chi_A(\Delta X(u)). \tag{2.1}$$

In the sequel, we will sometimes use μ_A to denote the restriction to A of the measure μ. in the following theorem, Var stands for variance.

Theorem 2.3 *Let A be bounded below, then*

1. *$\left(\int_A f(x)N(t,dx), t \geq 0\right)$ is a compound Poisson process, with characteristic function*

$$\mathbb{E}\left(\exp\left\{i\left(u, \int_A f(x)N(t,dx)\right)\right\}\right) = \exp\left[t\int_{\mathbb{R}^d}(e^{i(u,x)} - 1)\mu_{f,A}(dx)\right]$$

for each $u \in \mathbb{R}^d$, where $\mu_{f,A}(B) = \mu(A \cap f^{-1}(B))$, for each $B \in \mathcal{B}(\mathbb{R}^d)$.

2. *If $f \in L^1(A, \mu_A)$, then*

$$\mathbb{E}\left(\int_A f(x)N(t,dx)\right) = t\int_A f(x)\mu(dx).$$

3. *If $f \in L^2(A, \mu_A)$, then*

$$Var\left(\left|\int_A f(x)N(t,dx)\right|\right) = t\int_A |f(x)|^2\mu(dx).$$

Proof. - part of it!

1. For simplicity, we will prove this result in the case where $f \in L^1(A, \mu_A)$. First let f be a simple function and write $f = \sum_{j=1}^n c_j \chi_{A_j}$ where each $c_j \in \mathbb{R}^d$. We can assume, without loss of generality, that the A_j's are disjoint Borel subsets of A. By Theorem 2.2, we find that

$$\begin{aligned}
\mathbb{E}\left(\exp\left\{i\left(u, \int_A f(x)N(t,dx)\right)\right\}\right) &= \mathbb{E}\left(\exp\left\{i\left(u, \sum_{j=1}^n c_j N(t, A_j)\right)\right\}\right) \\
&= \prod_{j=1}^n \mathbb{E}\left(\exp\left\{i\left(u, c_j N(t, A_j)\right)\right\}\right) \\
&= \prod_{j=1}^n \exp\left\{t\left(e^{i(u,c_j)} - 1\right)\mu(A_j)\right\} \\
&= \exp\left\{t\int_A (e^{i(u,f(x))} - 1)\mu(dx)\right\}.
\end{aligned}$$

Now for an arbitrary $f \in L^1(A, \mu_A)$, we can find a sequence of simple functions converging to f in L^1 and hence a subsequence which converges to f almost surely. Passing to the limit along this subsequence in the above yields the required result, via dominated convergence.

(2) and (3) follow from (1) by differentiation. □

It follows from Theorem 2.3 (2) that a Poisson integral will fail to have a finite mean if $f \notin L^1(A, \mu)$.

For each $f \in L^1(A, \mu_A), t \geq 0$, we define the *compensated Poisson integral* by

$$\int_A f(x)\tilde{N}(t, dx) = \int_A f(x)N(t, dx) - t\int_A f(x)\mu(dx).$$

A straightforward argument shows that $\left(\int_A f(x)\tilde{N}(t, dx), t \geq 0\right)$ is a martingale and we will use this fact extensively in the sequel. Note that by Theorem 2.3 (2) and (3), we can easily deduce the following two important facts:

$$\mathbb{E}\left(\exp\left\{i\left(u, \int_A f(x)\tilde{N}(t, dx)\right)\right\}\right) = \exp\left\{t\int_{\mathbb{R}^d}(e^{i(u,x)} - 1 - i(u, x))\mu_{f,A}(dx)\right\}, \tag{2.2}$$

for each $u \in \mathbb{R}^d$, and for $f \in L^2(A, \mu_A)$,

$$\mathbb{E}\left(\left|\int_A f(x)\tilde{N}(t, dx)\right|^2\right) = t\int_A |f(x)|^2\mu(dx). \tag{2.3}$$

2.4 Processes of Finite Variation

We begin by introducing a useful class of functions. Let $\mathcal{P} = \{a = t_1 < t_2 < \cdots < t_n < t_{n+1} = b\}$ be a partition of the interval $[a, b]$ in $\mathbb{R}$, and define its mesh to be $\delta = \max_{1 \leq i \leq n} |t_{i+1} - t_i|$. We define the *variation* $\mathrm{Var}_{\mathcal{P}}(g)$ of a càdlàg mapping $g : [a, b] \to \mathbb{R}^d$ over the partition $\mathcal{P}$ by the prescription

$$\mathrm{Var}_{\mathcal{P}}(g) = \sum_{i=1}^{n} |g(t_{i+1}) - g(t_i)|.$$

If $V(g) = \sup_{\mathcal{P}} \mathrm{Var}_{\mathcal{P}}(g) < \infty$, we say that g has *finite variation on* $[a, b]$. If g is defined on the whole of $\mathbb{R}$ (or $\mathbb{R}^+$), it is said to have *finite variation* if it has finite variation on each compact interval.

It is a trivial observation that every non-decreasing g is of finite variation. Conversely if g is of finite variation, then it can always be written as the difference of two non-decreasing functions (to see this, just write $g = \frac{V(g)+g}{2} - \frac{V(g)-g}{2}$, where $V(g)(t)$ is the variation of g on $[a, t]$).

Functions of finite variation are important in integration, for suppose that we are given a function g which we are proposing as an integrator, then as a minimum we will want to be able to define the Stieltjes integral $\int_I f dg$, for all continuous functions f (where I is some finite interval). In fact a necessary and sufficient condition for obtaining such an integral as a limit of Riemann sums is that g has finite variation.

A stochastic process $(X(t), t \geq 0)$ is of *finite variation* if the paths $(X(t)(\omega), t \geq 0)$ are of finite variation for almost all $\omega \in \Omega$. The following is an important example for us.

Example *Poisson Integrals*

Let N be a Poisson random measure with intensity measure μ and let $f : \mathbb{R}^d \to \mathbb{R}^d$ be Borel measurable. For A bounded below, let $Y = (Y(t), t \geq 0)$ be given by $Y(t) = \int_A f(x)N(t, dx)$, then Y is of finite variation on $[0, t]$ for each $t \geq 0$. To see this, we observe that for all partitions $\mathcal{P}$ of $[0, t]$, we have

$$\text{Var}_{\mathcal{P}}(Y) \leq \sum_{0 \leq s \leq t} |f(\Delta X(s))| \chi_A(\Delta X(s)) < \infty \quad \text{a.s.} \tag{2.4}$$

where $X(t) = \int_A x N(t, dx)$, for each $t \geq 0$.

In fact, a necessary and sufficient condition for a Lévy process to be of finite variation is that there is no Brownian part (i.e. $a = 0$ in the Lévy-Khinchine formula), and $\int_{|x|<1} |x| \nu(dx) < \infty$.

2.5 The Lévy-Itô Decomposition

This is the first key result of this lecture.

First, note that for A bounded below, for each $t \geq 0$

$$\int_A x N(t, dx) = \sum_{0 \leq u \leq t} \Delta X(u) \chi_A(\Delta X(u))$$

is the sum of all the jumps taking values in the set A up to the time t. Since the paths of X are càdlàg, this is clearly a finite random sum. In particular, $\int_{|x| \geq 1} x N(t, dx)$ is the sum of all jumps of size bigger than one. It is a compound Poisson process, has finite variation but may have no finite moments. Conversely it can be shown that $X(t) - \int_{|x| \geq 1} x N(t, dx)$ is a Lévy process having finite moments to all orders.

Now lets turn our attention to the small jumps. We study compensated integrals, which we know are martingales. Introduce the notation

$$M(t, A) = \int_A x \tilde{N}(t, dx)$$

, for $t \geq 0$ and A bounded below. For each $m \in \mathbb{N}$, let

$$B_m = \left\{ x \in \mathbb{R}^d, \frac{1}{m+1} < |x| \leq \frac{1}{m} \right\}$$

and for each $n \in \mathbb{N}$, let $A_n = \bigcup_{m=1}^n B_m$. It can be shown that

$$\int_{|x|<1} x\tilde{N}(t, dx) = L^2 - \lim_{n\to\infty} M(t, A_n),$$

and hence it is a martingale. Moreover, on taking limits in (2.2), we get

$$\mathbb{E}\left(\exp i \left(u, \int_{|x|<1} x\tilde{N}(t, dx) \right) \right) = \exp \left\{ t \int_{|x|<1} (e^{i(u,x)} - 1 - i(u, x))\mu(dx) \right\}.$$

Consider

$$B_a(t) = X(t) - bt - \int_{|x|<1} x\tilde{N}(t, dx) - \int_{|x|\geq 1} xN(t, dx),$$

where $b = \mathbb{E}\left(X(1) - \int_{|x|\geq 1} xN(1, dx) \right)$. The process B_a is a centred martingale with continuous sample paths. With a little more work, we can show that $\mathrm{Cov}(B_a^i(t)B_a^j(t)) = a^{ij}t$. Using Lévy's characterisation of Brownian motion (see below) we have that B_a is a Brownian motion with covariance a. Hence we have:

Theorem 2.4 (The Lévy-Itô Decomposition) *If X is a Lévy process, then there exists $b \in \mathbb{R}^d$, a Brownian motion B_a with covariance matrix a in $\mathbb{R}^d$ and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}^d - \{0\})$ such that for each $t \geq 0$,*

$$X(t) = bt + B_a(t) + \int_{|x|<1} x\tilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx) \tag{2.5}$$

Note that the three processes in (2.5) are all independent.

An interesting by-product of the Lévy-Itô decomposition is the Lévy-Khintchine formula, which follows easily by independence in the Lévy-Itô decomposition:-

Corollary 2.1 *If X is a Lévy process, then for each $u \in \mathbb{R}^d, t \geq 0$,*

$$\mathbb{E}(e^{i(u,X(t))}) =$$
$$\exp\left(t \left[i(b, u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d - \{0\}} (e^{i(u,y)} - 1 - i(u, y)\chi_B(y))\mu(dy) \right] \right) \tag{2.6}$$

so the intensity measure μ is the Lévy measure for X.

The process $\int_{|x|<1} x\tilde{N}(t, dx)$ is the *compensated sum of small jumps*. The compensation takes care of the analytic complications in the Lévy-Khintchine formula in a probabilistically pleasing way, since it is an L^2-martingale.

The process $\int_{|x|\geq 1} xN(t, dx)$ describes the "large jumps" - it is a compound Poisson process, but may have no finite moments.

e.g. A Lévy process has finite variation iff its Lévy-Itô decomposition takes the form

$$\begin{aligned} X(t) &= \gamma t + \int_{x\neq 0} xN(t, dx) \\ &= \gamma t + \sum_{0\leq s\leq t} \Delta X(s), \end{aligned}$$

where $\gamma = b - \int_{|x|<1} x\nu(dx)$.

H.Geman, D.Madan and M.Yor have proposed a nice financial interpretation for the jump terms in the Lévy-Itô decomposition:- where the intensity measure is infinite, the stock price manifests "infinite activity" and this is the mathematical signature of the jitter arising from the interaction of pure supply shocks and pure demand shocks. On the other hand, where the intensity measure is finite, we have "finite activity", and this corresponds to sudden shocks that can cause unexpected movements in the market, such as a terrorist atrocity or a major earthquake.

If a pure jump Lévy process (no Brownian part) has finite activity then it has finite variation. The converse is false.

The first three terms on the rhs of (2.5) have finite moments to all orders, so if a Lévy process fails to have a moment, this is due entirely to the "large jumps"/"finite activity" part. In fact:

$$\mathbb{E}(|X(t)|^n) < \infty \text{ for all } t > 0 \text{ if and only if } \int_{|x|\geq 1} |x|^n \nu(dx) < \infty.$$

Semimartingales A stochastic process X is a *semimartingale* if it is an adapted process such that for each $t \geq 0$,

$$X(t) = X(0) + M(t) + C(t),$$

where $M = (M(t), t \geq 0)$ is a local martingale and $C = (C(t), t \geq 0)$ is an adapted process of finite variation. In particular

Every Lévy process is a semimartingale.

To see this, use the Lévy-Itô decomposition to write

$$M(t) = B_a(t) + \int_{|x|<1} x\tilde{N}(t, dx) \text{ - a martingale,}$$

$$C(t) = bt + \int_{|x|\geq 1} xN(t, dx).$$

2.6 Stochastic Integration

We next give a rather rapid account of stochastic integration in a form suitable for application to Lévy processes.

Let $X = M + C$ be a semimartingale. The problem of stochastic integration is to make sense of objects of the form

$$\int_0^t F(s)dX(s) = \int_0^t F(s)dM(s) + \int_0^t F(s)dC(s).$$

The second integral can be well-defined using the usual Lebesgue-Stieltjes approach. The first one cannot - indeed if M is a continuous martingale of finite variation, then M is a.s. constant.

Refer to the martingale part of the Lévy-Itô decomposition (2.4). Define a "martingale-valued measure" by

$$M(t, E) = B(t)\delta_0(E) + \tilde{N}(t, E - \{0\}),$$

for $E \in \mathcal{B}(\mathbb{R}^d)$, where $B = (B(t), t \geq 0)$ is a one-dimensional Brownian motion. The following key properties then hold:-

- $M((s,t], E) = M(t, E) - M(s, E)$ is independent of $\mathcal{F}_s$, for $0 \leq s < t < \infty$.
- $\mathbb{E}(M((s,t], E)) = 0$.
- $\mathbb{E}(M((s,t], E)^2) = \rho((s,t], E)$
  where $\rho((s,t], E) = (t - s)(\delta_0(E) + \nu(E - \{0\}))$.

We're going to unify the usual stochastic integral with the Poisson integral, by defining:

$$\int_0^t \int_E F(s,x)M(ds,dx) = \int_0^t G(s)dB(s) + \int_0^t \int_{E-\{0\}} F(s,x)\tilde{N}(ds,dx).$$

where $G(s) = F(s, 0)$. Of course, we need some conditions on the class of integrands:-

Fix $E \in \mathcal{B}(\mathbb{R}^d)$ and $0 < T < \infty$ and let $\mathcal{P}$ denote the smallest σ-algebra with respect to which all mappings $F : [0,T] \times E \times \Omega \to \mathbb{R}$ satisfying (1) and (2) below are measurable.

1. For each $0 \leq t \leq T$, the mapping $(x, \omega) \to F(t, x, \omega)$ is $\mathcal{B}(E) \otimes \mathcal{F}_t$ measurable,

2. For each $x \in E, \omega \in \Omega$, the mapping $t \to F(t, x, \omega)$ is left continuous.

We call $\mathcal{P}$ the *predictable σ-algebra*. A $\mathcal{P}$-measurable mapping $G : [0, T] \times E \times \Omega \to \mathbb{R}$ is then said to be *predictable*. The definition clearly extends naturally to the case where $[0, T]$ is replaced by $\mathbb{R}^+$.

Note that by (1), if G is predictable then the process $t \to G(t, x, \cdot)$ is adapted, for each $x \in E$. If G satisfies (1) and is left continuous then it is clearly predictable.

Define $\mathcal{H}_2(T, E)$ to be the linear space of all equivalence classes of mappings $F : [0, T] \times E \times \Omega \to \mathbb{R}$ which coincide almost everywhere with respect to $\rho \times P$ and which satisfy the following conditions:

- F is predictable,
-
$$\int_0^T \int_E \mathbb{E}(|F(t, x)|^2)\rho(dt, dx) < \infty.$$

It can be shown that $\mathcal{H}_2(T, E)$ is a real Hilbert space with respect to the inner product $< F, G >_{T,\rho} = \int_0^T \int_E \mathbb{E}((F(t, x), G(t, x)))\rho(dt, dx)$.

Define $S(T, E)$ to be the linear space of all simple processes in $\mathcal{H}_2(T, E)$, where F is *simple* if for some $m, n \in \mathbb{N}$, there exists $0 \leq t_1 \leq t_2 \leq \cdots \leq t_{m+1} = T$ and there exists a family of disjoint Borel subsets $A_1, A_2, \ldots, A_n$ of E with each $\nu(A_i) < \infty$ such that

$$F = \sum_{j=1}^{m} \sum_{k=1}^{n} F_k(t_j)\chi_{(t_j, t_{j+1}]}\chi_{A_k},$$

where each $F_k(t_j)$ is a bounded $\mathcal{F}_{t_j}$-measurable random variable. Note that F is left continuous and $\mathcal{B}(E) \otimes \mathcal{F}_t$ measurable, hence it is predictable. An important fact is that

$$S(T, E) \text{ is dense in } \mathcal{H}_2(T, E),$$

One of Itô's greatest achievements was the definition of the stochastic integral $I_T(F)$, for F simple, by separating the "past" from the "future" within the Riemann sum:-

$$I_T(F) = \sum_{j=1}^{m} \sum_{k=1}^{n} F_k(t_j)M((t_j, t_{j+1}], A_k), \tag{2.7}$$

so on each time interval $[t_j, t_{j+1}]$, $F_k(t_j)$ encapsulates information obtained by time t_j, while M gives the innovation into the future $(t_j, t_{j+1}]$.

Lemma 2.3 *For each $T \geq 0, F \in S(T, E)$,*

$$\mathbb{E}(I_T(F)) = 0, \qquad \mathbb{E}(I_T(F)^2) = \int_0^T \int_E \mathbb{E}(|F(t,x)|^2)\rho(dt, dx).$$

Proof. $\mathbb{E}(I_T(F)) = 0$ is a straightforward application of linearity and independence. The second result is quite messy - we lose nothing important by just looking at the Brownian case, with $d = 1$. So let $F(t) = \sum_{j=1}^m F(t_j)\chi_{(t_j, t_{j+1}]}$, then $I_T(F) = \sum_{j=1}^m F(t_j)(B(t_{j+1}) - B(t_j))$, and

$$I_T(F)^2 = \sum_{j=1}^m \sum_{p=1}^m F(t_j)F(t_p)(B(t_{j+1}) - B(t_j))(B(t_{p+1}) - B(t_p)).$$

Now fix j and split the second sum into three pieces - corresponding to $p < j, p = j$ and $p > j$. When $p < j, F(t_j)F(t_p)(B(t_{p+1}) - B(t_p)) \in \mathcal{F}_{t_j}$ which is independent of $B(t_{j+1}) - B(t_j)$,

$$\begin{aligned} & \mathbb{E}[F(t_j)F(t_p(B(t_{j+1}) - B(t_j))(B(t_{p+1}) - B(t_p))] \\ = \; & \mathbb{E}[F(t_j)F(t_p)(B(t_{p+1}) - B(t_p))]\mathbb{E}(B(t_{j+1}) - B(t_j)) = 0. \end{aligned}$$

Exactly the same argument works when $p > j$. What remains is the case $p = j$, and by independence again,

$$\begin{aligned} \mathbb{E}(I_T(F)^2) &= \sum_{j=1}^m \mathbb{E}(F(t_j)^2)\mathbb{E}(B(t_{j+1}) - B(t_j))^2 \\ &= \sum_{j=1}^m \mathbb{E}(F(t_j)^2)(t_{j+1} - t_j). \qquad \square \end{aligned}$$

We deduce from Lemma 2.3 and Exercise 4.1, that I_T is a linear isometry from $S(T, E)$ into $L^2(\Omega, \mathcal{F}, P)$, and hence it extends to an isometric embedding of the whole of $\mathcal{H}_2(T, E)$ into $L^2(\Omega, \mathcal{F}, P)$. We continue to denote this extension as I_T and we call $I_T(F)$ the *(Itô) stochastic integral* of $F \in \mathcal{H}_2(T, E)$. When convenient, we will use the Leibniz notation $I_T(F) = \int_0^T \int_E F(t, x)M(dt, dx)$.

So for predictable F satisfying $\int_0^T \int_E \mathbb{E}(|F(s, x)|^2)\rho(ds, dx) < \infty$, we can find a sequence $(F_n, n \in \mathbb{N})$ of simple processes such that

$$\int_0^T \int_E F(t, x)M(dt, dx) = \lim_{n \to \infty} \int_0^T \int_E F_n(t, x)M(dt, dx).$$

The limit is taken in the L^2-sense and is independent of the choice of approximating sequence.

The following theorem summarises some useful properties of the stochastic integral.

Theorem 2.5 *If $F, G \in \mathcal{H}_2(T, E)$ and $\alpha, \beta \in \mathbb{R}$, then :*

1. $I_T(\alpha F + \beta G) = \alpha I_T(F) + \beta I_T(G)$.
2. $\mathbb{E}(I_T(F)) = 0, \qquad \mathbb{E}(I_T(F)^2) = \int_0^T \int_E \mathbb{E}(|F(t,x)|^2)\rho(dt, dx)$.
3. $(I_t(F), t \geq 0)$ *is* $\mathcal{F}_t$*-adapted.*
4. $(I_t(F), t \geq 0)$ *is a square-integrable martingale.*

Proof. (1) and (2) are easy.

For (3), let $(F_n, n \in \mathbb{N})$ be a sequence in $S(T, E)$ converging to F; then each process $(I_t(F_n), t \geq 0)$ is clearly adapted. Since each $I_t(F_n) \to I_t(F)$ in L^2 as $n \to \infty$, we can find a subsequence $(F_{n_k}, n_k \in \mathbb{N})$ such that $I_t(F_{n_k}) \to I_t(F)$ a.s. as $n_k \to \infty$, and the required result follows.

(4) Let $F \in S(T, E)$ and (without loss of generality) choose $0 < s = t_l < t_{l+1} < t$. Then it is easy to see that $I_t(F) = I_s(F) + I_{s,t}(F)$ and hence $\mathbb{E}_s(I_t(F)) = I_s(F) + \mathbb{E}_s(I_{s,t}(F))$ by (3). However,

$$\begin{aligned}
\mathbb{E}_s(I_{s,t}(F)) &= \mathbb{E}_s\left(\sum_{j=l+1}^{m}\sum_{k=1}^{n} F_k(t_j) M((t_j, t_{j+1}], A_k)\right) \\
&= \sum_{j=l+1}^{n}\sum_{k=1}^{n} \mathbb{E}_s(F_k(t_j))\mathbb{E}(M((t_j, t_{j+1}], A_k)) = 0.
\end{aligned}$$

The result now follows by the continuity of $\mathbb{E}_s$ in L^2. Indeed, let $(F_n, n \in \mathbb{N})$ be a sequence in $S(T, E)$ converging to F; then we have

$$\begin{aligned}
||\mathbb{E}_s(I_t(F)) - \mathbb{E}_s(I_t(F_n))||_2 &\leq ||I_t(F) - I_t(F_n)||_2 \\
&= ||F - F_n||_{T,\rho} \to 0 \text{ as } n \to \infty.
\end{aligned}$$

□

We can extend the stochastic integral $I_T(F)$ to integrands in $\mathcal{P}_2(T, E)$. This is the linear space of all equivalence classes of mappings $F : [0, T] \times E \times \Omega \to \mathbb{R}$ which coincide almost everywhere with respect to $\rho \times P$, and which satisfy the following conditions:

- F is predictable.
- $P\left(\int_0^T \int_E |F(t,x)|^2 \rho(dt, dx) < \infty\right) = 1$.

If $F \in \mathcal{P}_2(T, E), (I_t(F), t \geq 0)$ is always a local martingale, but not necessarily a martingale.

2.6.1 Poisson Stochastic Integrals

Let A be an arbitrary Borel set in $\mathbb{R}^d - \{0\}$ which is bounded below, and introduce the compound Poisson process $P = (P(t), t \geq 0)$, where each $P(t) = \int_A xN(t, dx)$. Let K be a predictable mapping, then generalising equation (2.1), we define

$$\int_0^T \int_A K(t, x)N(dt, dx) = \sum_{0 \leq u \leq T} K(u, \Delta P(u))\chi_A(\Delta P(u)), \tag{2.8}$$

as a random finite sum.

In particular, if H satisfies the square-integrability condition given above, we may then define, for each $1 \leq i \leq d$,

$$\int_0^T \int_A H^i(t, x)\tilde{N}(dt, dx) = \int_0^T \int_A H^i(t, x)N(dt, dx) - \int_0^T \int_A H^i(t, x)\nu(dx)dt.$$

The definition (2.8) can, in principle, be used to define stochastic integrals for a more general class of integrands than we have been considering. For simplicity, let $N = (N(t), t \geq 0)$ be a Poisson process of intensity 1 and let $f : \mathbb{R} \to \mathbb{R}$, then we may define

$$\int_0^t f(N(s))dN(s) = \sum_{0 \leq s \leq t} f(N(s-) + \Delta N(s))\Delta N(s).$$

2.6.2 Lévy-type stochastic integrals

We take $E = \hat{B} - \{0\} = \{x \in \mathbb{R}^d; 0 < |x| < 1\}$ throughout this subsection. We say that an $\mathbb{R}^d$-valued stochastic process $Y = (Y(t), t \geq 0)$ is a *Lévy-type stochastic integral* if it can be written in the following form for each $1 \leq i \leq d, t \geq 0$,

$$\begin{aligned} Y^i(t) &= Y^i(0) + \int_0^t G^i(s)ds + \int_0^t F^i_j(s)dB^j(s) + \int_0^t \int_{|x|<1} H^i(s, x)\tilde{N}(ds, dx) \\ &+ \int_0^t \int_{|x|\geq 1} K^i(s, x)N(ds, dx), \end{aligned} \tag{2.9}$$

where for each $1 \leq i \leq d, 1 \leq j \leq m, t \geq 0, |G^i|^{\frac{1}{2}}, F^i_j \in \mathcal{P}_2(T), H^i \in \mathcal{P}_2(T, E)$ and K is predictable. B is an m-dimensional standard Brownian motion and N is an independent Poisson random measure on $\mathbb{R}^+ \times (\mathbb{R}^d - \{0\})$ with compensator $\tilde{N}$ and intensity measure ν, which we will assume is a Lévy measure. We will assume that the random variable $Y(0)$ is $\mathcal{F}_0$-measurable, and then it is clear that Y is an adapted process.

We can often simplify complicated expressions by employing the notation of *stochastic differentials* to represent Lévy-type stochastic integrals. We then write (2.9) as

$$dY(t) = G(t)dt + F(t)dB(t) + H(t,x)\tilde{N}(dt,dx) + K(t,x)N(dt,dx).$$

When we want to particularly emphasise the domains of integration with respect to x, we will use an equivalent notation

$$dY(t) = G(t)dt+F(t)dB(t)+\int_{|x|<1} H(t,x)\tilde{N}(dt,dx)+\int_{|x|\geq 1} K(t,x)N(dt,dx).$$

Clearly Y is a semimartingale.

Let $M = (M(t), t \geq 0)$ be an adapted process which is such that $MJ \in \mathcal{P}_2(t,A)$ whenever $J \in \mathcal{P}_2(t,A)$ (where $A \in \mathcal{B}(\mathbb{R}^d)$ is arbitrary) . For example, it is sufficient to take M to be adapted and left-continuous.

For these processes we can define an adapted process $Z = (Z(t), t \geq 0)$ by the prescription that it have the stochastic differential

$$dZ(t) = M(t)G(t)dt+M(t)F(t)dB(t)+M(t)H(t,x)\tilde{N}(dt,dx)+M(t)K(t,x)N(dt,dx),$$

and we will adopt the natural notation,

$$dZ(t) = M(t)dY(t).$$

Example (Lévy Stochastic Integrals)

Let X be a Lévy process with characteristics (b,a,ν) and Lévy-Itô decomposition given by equation (2.5):

$$X(t) = bt + B_a(t) + \int_{|x|<1} x\tilde{N}(t,dx) + \int_{|x|\geq 1} xN(t,dx),$$

for each $t \geq 0$. Let $L \in \mathcal{P}_2(t)$ for all $t \geq 0$ and in (2.9), choose each $F^i_j = a^i_j L, H^i = K^i = x^i L$. Then we can construct processes with the stochastic differential

$$dY(t) = L(t)dX(t) \tag{2.10}$$

We call Y a *Lévy stochastic integral.*

In the case where X has finite variation, the Lévy stochastic integral Y can also be constructed as a Lebesgue-Stieltjes integral, and this coincides (up to a set of measure zero) with the prescription (2.10).

Example: The Ornstein Uhlenbeck Process (OU Process)

$$Y(t) = e^{-\lambda t}y_0 + \int_0^t e^{-\lambda(t-s)}dX(s) \tag{2.11}$$

where $y_0 \in \mathbb{R}^d$ is fixed is a Markov process. The condition

$$\int_{|x|>1} \log(1+|x|)\nu(dx) < \infty$$

is necessary and sufficient for it to be stationary. There are important applications to volatility modelling in finance which have recently been developed by Ole Barndorff-Nielsen and Neil Sheppard. Intriguingly, every self-decomposable random variable can be naturally embedded in a stationary Lévy-driven OU process.

If X is a standard Brownian motion then each $Y(t)$ is Gaussian with mean $e^{-\lambda t}y_0$ and variance $\frac{1}{2\lambda}(1-e^{-2\lambda t}I)$. In this case the OU process is a good model of the physical phenomenon of Brownian motion as it includes the viscous drag of the medium on the particle as well as random fluctuations. This is more easily seen by writing the differential form of (2.11) - the "Langevin equation":-

$$dY(t) = -\lambda Y(t)dt + dX(t),$$

which is a stochastic differential equation (see below).

Stochastic integration has a plethora of applications including filtering, stochastic control and infinite dimensional analysis. Here's some motivation from finance. Suppose that $X(t)$ is the value of a stock at time t and $F(t)$ is the number of stocks owned at time t. Assume for now that we buy and sell stocks at discrete times $0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = T$. Then the total value of the portfolio at time T is:

$$V(T) = V(0) + \sum_{j=0}^{n} F(t_j)(X(t_{j+1}) - X(t_j)),$$

and in the limit as the times become infinitesimally close together, we have the stochastic integral

$$V(T) = V(0) + \int_0^t F(s)dX(s).$$

Stochastic integration against Brownian motion was first developed by Wiener for sure functions. Itô's groundbreaking work in 1944 extended this to random adapted integrands. The generalisation of the integrator to arbitrary martingales was due to Kunita and Watanabe in 1967 and the further step to allow semimartingales was due to P.A.Meyer and the Strasbourg school in the 1970s.

2.7 Itô's Formula

We begin with the easy case - Itô's formula for Poisson stochastic integrals of the form

$$W^i(t) = W^i(0) + \int_0^t \int_A K^i(t,x)N(dt,dx) \tag{2.12}$$

for $1 \leq i \leq d$, where $t \geq 0$, A is bounded below and each K^i is predictable. Itô's formula for such processes takes a particularly simple form.

Lemma 2.4 *If W is a Poisson stochastic integral of the form (2.12) then for each $f \in C(\mathbb{R}^d)$, and for each $t \geq 0$, with probability one, we have*

$$f(W(t)) - f(W(0)) = \int_0^t \int_A [f(W(s-) + K(s,x)) - f(W(s-))]N(ds,dx).$$

Proof. Let $Y(t) = \int_A xN(dt,dx)$ and recall that the jump times for Y are defined recursively as $T_0^A = 0$ and for each $n \in \mathbb{N}, T_n^A = \inf\{t > T_{n-1}^A; \Delta Y(t) \in A\}$. We then find that,

$$\begin{aligned} f((W(t)) - f(W(0)) &= \sum_{0 \leq s \leq t} f(W(s)) - f(W(s-)) \\ &= \sum_{n=1}^{\infty} f(W(t \wedge T_n^A)) - f(W(t \wedge T_{n-1}^A)) \\ &= \sum_{n=1}^{\infty} [f(W(t \wedge T_n^A -)) + K(t \wedge T_n^A, \Delta Y(t \wedge T_n^A)) - f(W(t \wedge T_n^A -))] \\ &= \int_0^t \int_A [f(W(s-) + K(s,x)) - f(W(s-))]N(ds,dx). \quad \square \end{aligned}$$

The celebrated Itô formula for Brownian motion is probably well-known to you so I'll briefly outline the proof. Let $(\mathcal{P}_n, n \in \mathbb{N})$ be a sequence of partitions of the form $\mathcal{P}_n = \{0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_{m(n)}^{(n)} < t_{m(n)+1}^{(n)} = T\}$. and suppose that $\lim_{n\to\infty} \delta(\mathcal{P}_n) = 0$, where the mesh, $\delta(\mathcal{P}_n) = \max_{0 \leq j \leq m(n)} |t_{j+1}^{(n)} - t_j^{(n)}|$. As a preliminary - you need the following:-

Lemma 2.5 *If $W_{kl} \in \mathcal{H}_2(T)$ for each $1 \leq k, l \leq m$, then*

$$L^2-\lim_{n\to\infty}\sum_{j=0}^{n} W_{kl}(t_j^{(n)})(B^k(t_{j+1}^{(n)})-B^k(t_j^{(n)}))(B^l(t_{j+1}^{(n)})-B^l(t_j^{(n)})) = \sum_{k=1}^{m}\int_0^T W_{kk}(s)ds.$$

The proof is similar to that of Lemma 2.3 - but you will need the Gaussian moment $\mathbb{E}(B(t)^4) = 3t^2$. □

Now let M be a Brownian integral with drift of the form

$$M^i(t) = \int_0^t F_j^i(s)dB^j(s) + \int_0^t G^i(s)ds, \tag{2.13}$$

where each $F_j^i, (G^i)^{\frac{1}{2}} \in \mathcal{P}_2(t)$, for all $t \geq 0, 1 \leq i \leq d, 1 \leq j \leq m$.

For each $1 \leq i \leq j$, we introduce the *quadratic variation process* denoted as $([M^i, M^j](t), t \geq 0)$ by

$$[M^i, M^j](t) = \sum_{k=1}^{m}\int_0^T F_k^i(s)F_k^j(s)ds.$$

We will explore quadratic variation in greater depth in the sequel. The following slick method of proving Itô's formula is due to Kunita.

Theorem 2.6 (Itô's Theorem 1) *If $M = (M(t), t \geq 0)$ is a Brownian integral with drift of the form (2.13), then for all $f \in C^2(\mathbb{R}^d), t \geq 0$, with probability 1, we have*

$$f(M(t))-f(M(0)) = \int_0^t \partial_i f(M(s))dM^i(s)+\frac{1}{2}\int_0^t \partial_i\partial_j f(M(s))d[M^i, M^j](s).$$

Proof. Let $(\mathcal{P}_n, n \in \mathbb{N})$ be a sequence of partitions of $[0, t]$ as above. By Taylor's theorem, we have, for each such partition (where we suppress the index n).

$$\begin{aligned} f(M(t)) - f(M(0)) &= \sum_{k=0}^{m} f(M(t_{k+1})) - f(M(t_k)) \\ &= J_1(t) + \frac{1}{2}J_2(t), \end{aligned}$$

where

$$J_1(t) = \sum_{k=0}^{m} \partial_i f(M(t_k))(M^i(t_{k+1}) - M^i(t_k)),$$

$$J_2(t) = \sum_{k=0}^{m} \partial_i \partial_j f(N_{ij}^k)(M^i(t_{k+1}) - M^i(t_k))(M^j(t_{k+1}) - M^j(t_k)),$$

and where the N_{ij}^k's are each $\mathcal{F}(t_{k+1})$-adapted $\mathbb{R}^d$-valued random variables satisfying $|N_{ij}^k - M(t_k)| \leq |M(t_{k+1}) - M(t_k)|$.

We write each $J_2(t) = K_1(t) + K_2(t)$, where

$$K_1(t) = \sum_{k=0}^{m} \partial_i \partial_j f(M(t_k))(M^i(t_{k+1}) - M^i(t_k))(M^j(t_{k+1}) - M^j(t_k)),$$

$$K_2(t) = \sum_{k=0}^{m} [\partial_i \partial_j f(N_{ij}^k) - \partial_i \partial_j f(M(t_k))](M^i(t_{k+1}) - M^i(t_k))(M^j(t_{k+1}) - M^j(t_k)).$$

Now take limits as $n \to \infty$. It turns out that $K_2(t) \to 0$, in probability and the result follows. $\square$

Itô's formula for general Lévy-type stochastic integrals is obtained essentially by combining the Poisson and Brownian results and making sure you take good care of the compensators for small jumps. You should be able to guess the right result.

To give a precise statement, consider a Lévy-type stochastic integral of the form

$$dY(t) = G(t)dt + F(t)dB(t) + H(t,x)\tilde{N}(dt,dx) + K(t,x)N(dt,dx). \quad (2.14)$$

Theorem 2.7 (Itô's Theorem 2) *If Y is a Lévy-type stochastic integral of the form (2.14), then for each $f \in C^2(\mathbb{R}^d), t \geq 0$, with probability 1, we have*

$$\begin{aligned} f(Y(t)) - f(Y(0)) &= \int_0^t \partial_i f(Y(s-)) dY_c^i(s) + \frac{1}{2} \int_0^t \partial_i \partial_j f(Y(s-)) d[Y_c^i, Y_c^j](s) \\ &+ \int_0^t \int_{|x| \geq 1} [f(Y(s-) + K(s,x)) - f(Y(s-))] N(ds,dx) \\ &+ \int_0^t \int_{|x|<1} [f(Y(s-) + H(s,x)) - f(Y(s-))] \tilde{N}(ds,dx) \\ &+ \int_0^t \int_{|x|<1} \big[f(Y(s-) + H(s,x)) - f(Y(s-)) \\ &- H^i(s,x) \partial_i f(Y(s-)) \big] \nu(dx) ds. \end{aligned}$$

Tedious but straightforward algebra yields the following form, which is important since it extends to general semimartingales:-

Theorem 2.8 (Itô's Theorem 3) *If Y is a Lévy-type stochastic integral of the form (2.14), then for each $f \in C^2(\mathbb{R}^d), t \geq 0$, with probability 1, we have*

$$\begin{aligned} f(Y(t)) - f(Y(0)) &= \int_0^t \partial_i f(Y(s-))dY^i(s) + \frac{1}{2}\int_0^t \partial_i\partial_j f(Y(s-))d[Y_c^i, Y_c^j](s) \\ &+ \sum_{0\leq s\leq t}[f(Y(s)) - f(Y(s-)) - \Delta Y^i(s)\partial_i f(Y(s-))]. \end{aligned}$$

Here Y_c denotes the *continuous* part of Y defined by $Y_c^i(t) = \int_0^t G^i(s)ds + \int_0^t F_j^i(s)dB^j(s)$.

Note that a special case of Itô's formula yields the following "classical" chain rule for differentiable functions f, when the process Y is of finite variation:

$$\begin{aligned} f(Y(t)) - f(Y(0)) &= \int_0^t \partial_i f(Y(s-))dY^i(s) + \\ &+ \sum_{0\leq s\leq t}[f(Y(s)) - f(Y(s-)) - \Delta Y^i(s)\partial_i f(Y(s-))]. \end{aligned}$$

A form of Itô's formula may even be established for fractional Brownian motion, which is not a semimartingale.

2.8 Quadratic Variation and Itô's Product Formula

We extend the definition of quadratic variation to the more general case of Lévy-type stochastic integrals $Y = (Y(t), t \geq 0)$ of the form (2.14). So for each $t \geq 0$ we define a $d \times d$ matrix-valued adapted process $[Y, Y] = ([Y, Y](t), t \geq 0)$ by the following prescription for its (i, j)th entry $(1 \leq i, j \leq d)$,

$$[Y^i, Y^j](t) = [Y_c^i, Y_c^j](t) + \sum_{0\leq s\leq t}\Delta Y^i(s)\Delta Y^j(s). \tag{2.15}$$

Each $[Y^i, Y^j](t)$ is almost surely finite, and we have

$$\begin{aligned} [Y^i, Y^j](t) &= \sum_{k=1}^m \int_0^T F_k^i(s)F_k^j(s)ds + \int_0^t\int_{|x|<1} H^i(s,x)H^j(s,x)N(ds,dx) \\ &+ \int_0^t\int_{|x|\geq 1} K^i(s,x)K^j(s,x)N(ds,dx), \end{aligned} \tag{2.16}$$

so that we clearly have each $[Y^i, Y^j](t) = [Y^j, Y^i](t)$. Note that the integral over small jumps in this case is always a.s. finite (Why ?)

It is easy to show that for each $\alpha, \beta \in \mathbb{R}$ and $1 \leq i, j, k \leq d, t \geq 0$,

$$[\alpha Y^i + \beta Y^j, Y^k](t) = \alpha[Y^i, Y^k](t) + \beta[Y^j, Y^k](t).$$

The importance of $[Y, Y]$ is that it measures the deviation in the stochastic differential of products from the usual Leibniz formula. The following result makes this precise

Theorem 2.9 (Itô's Product Formula) *If Y^1 and Y^2 are real-valued Lévy-type stochastic integrals of the form (2.14), then for all $t \geq 0$, with probability one, we have that*

$$\begin{aligned} Y^1(t)Y^2(t) &= Y^1(0)Y^2(0) + \int_0^t Y^1(s-)dY^2(s) \\ &+ \int_0^t Y^2(s-)dY^1(s) + [Y^1, Y^2](t). \end{aligned}$$

Proof. We consider Y^1 and Y^2 as components of a vector $Y = (Y^1, Y^2)$ and we take f in Theorem 2.8 to be the smooth mapping from $\mathbb{R}^2$ to $\mathbb{R}$ given by $f(x^1, x^2) = x^1 x^2$.

By Theorem 2.8, we then obtain, for each $t \geq 0$, with probability one,

$$\begin{aligned} Y^1(t)Y^2(t) &= Y^1(0)Y^2(0) + \int_0^t Y^1(s-)dY^2(s) \\ &+ \int_0^t Y^2(s-)dY^1(s) + [Y_c^1, Y_c^2](t) \\ &+ \sum_{0 \leq s \leq t} [Y^1(s)Y^2(s) - Y^1(s-)Y^2(s-) \\ &- (Y^1(s) - Y^1(s-))Y^2(s-) - (Y^2(s) - Y^2(s-))Y^1(s-)], \end{aligned}$$

from which the required result easily follows. □

We can learn much about the way our Itô formulae work by writing the product formula in differential form:-

$$d(Y^1(t)Y^2(t)) = Y^1(t-)dY^2(t) + Y^2(t-)dY^1(t) + d[Y^1, Y^2](t).$$

By equation (2.16), we see that the term $d[Y^1, Y^2](t)$, which is sometimes called an *Itô correction*, arises as a result of the following formal product relations between differentials:-

$$dB^i(t)dB^j(t) = \delta^{ij}dt \quad ; \quad N(dt, dx)N(dt, dy) = N(dt, dx)\delta(x - y),$$

for $1 \leq i, j \leq m$, with all other products of differentials vanishing and if you have little previous experience of this game, these relations are a very valuable guide to intuition.

For completeness, we will give another characterisation of quadratic variation which is sometimes quite useful. We recall the sequence of partitions $(\mathcal{P}_n, n \in \mathbb{N})$, with mesh tending to zero which were introduced earlier.

Theorem 2.10 *If X and Y are real-valued Lévy-type stochastic integrals of the form (2.14), then for each $t \geq 0$, with probability one, we have*

$$[X, Y](t) = \lim_{n \to \infty} \sum_{j=0}^{m_n} (X(t_{j+1}^{(n)}) - X(t_j^{(n)}))(Y(t_{j+1}^{(n)}) - Y(t_j^{(n)})),$$

where the limit is taken in probability.

Proof. By polarisation, it is sufficient to consider the case $X = Y$. Using the identity

$$(x - y)^2 = x^2 - y^2 - 2y(x - y)$$

for $x, y \in \mathbb{R}$, we deduce that

$$\begin{aligned}\sum_{j=0}^{m_n} (X(t_{j+1}^{(n)}) - X(t_j^{(n)}))^2 &= \sum_{j=0}^{m_n} X(t_{j+1}^{(n)})^2 - \sum_{j=0}^{m_n} X(t_j^{(n)})^2 \\ &- 2\sum_{j=0}^{m_n} X(t_j^{(n)})(X(t_{j+1}^{(n)}) - X(t_j^{(n)})),\end{aligned}$$

and the required result follows from Itô's product formula (theorem 2.9). □

Many of the results of this lecture extend from Lévy-type stochastic integrals to arbitrary semimartingales. In particular, if F is a simple process and X is a semimartingale we can again use Itô's prescription to define

$$\int_0^t F(s)dX(s) = \sum F(t_j)(X(t_{j+1}) - X(t_j)),$$

and then pass to the limit to obtain more general stochastic integrals. Itô's formula can be established in the form given in Theorem 2.8 and the quadratic variation of semimartingales defined as the correction term in the corresponding Itô product formula.

Although stochastic calculus for general semimartingales is not the subject of these lectures, we do require one result - the famous Lévy characterisation of Brownian motion.

Theorem 2.11 (Lévy's characterisation) *Let $M = (M(t), t \geq 0)$ be a continuous centered martingale, which is adapted to a given filtration $(\mathcal{F}_t, t \geq 0)$. If $[M_i, M_j](t) = a_{ij}t$ for each $t \geq 0, 1 \leq i, j \leq d$ where $a = (a_{ij})$ is a positive definite symmetric matrix, then M is an $\mathcal{F}_t$-adapted Brownian motion with covariance a.*

Proof. Fix $u \in \mathbb{R}^d$ and define the process $(Y_u(t), t \geq 0)$ by $Y_u(t) = e^{i(u,M(t))}$, then by Itô's formula, we obtain

$$\begin{aligned} dY_u(t) &= iu^j Y_u(t) dM_j(t) - \frac{1}{2} u^i u^j Y_u(t) d[M_i, M_j](t) \\ &= iu^j Y_u(t) dM_j(t) - \frac{1}{2}(u, au) Y_u(t) dt. \end{aligned}$$

Upon integrating from s to t, we obtain

$$Y_u(t) = Y_u(s) + iu^j \int_s^t Y_u(\tau) dM_j(\tau) - \frac{1}{2}(u, au) \int_s^t Y_u(\tau) d\tau.$$

Now take conditional expectations of both sides with respect to $\mathcal{F}_s$, and use the conditional Fubini Theorem to obtain

$$\mathbb{E}(Y_u(t)|\mathcal{F}_s) = Y_u(s) - \frac{1}{2}(u, au) \int_s^t \mathbb{E}(Y_u(\tau)|\mathcal{F}_s) d\tau.$$

Hence $\mathbb{E}(e^{i(u,M(t)-M(s))}|\mathcal{F}_s) = e^{-\frac{1}{2}(u,au)(t-s)}$. □

2.9 Stochastic Differential Equations

Using Picard iteration one can show the existence of a unique solution to

$$\begin{aligned} dY(t) &= b(Y(t-))dt + \sigma(Y(t-))dB(t) + \qquad (2.17) \\ &+ \int_{|x|<c} F(Y(t-), x)\tilde{N}(dt, dx) + \int_{|x|\geq c} G(Y(t-), x)N(dt, dx), \end{aligned}$$

which is a convenient shorthand for the system of SDE's:-

$$\begin{aligned} dY^i(t) &= b^i(Y(t-))dt + \sigma^i_j(Y(t-))dB^j(t) + \qquad (2.18) \\ &+ \int_{|x|\leq c} F^i(Y(t-), x)\tilde{N}(dt, dx) + \int_{|x|>c} G^i(Y(t-), x)N(dt, dx), \end{aligned}$$

where each $1 \leq i \leq d$. The conditions under which this holds are:-

(1) **Lipschitz Condition**

There exists $K_1 > 0$ such that for all $y_1, y_2 \in \mathbb{R}^d$,

$$|b(y_1) - b(y_2)|^2 + ||a(y_1, y_1) - 2a(y_1, y_2) + a(y_2, y_2)|| \qquad (2.19)$$
$$+ \int_{|x|<c} |F(y_1, x) - F(y_2, x)|^2 \nu(dx) \leq K_1 |y_1 - y_2|^2.$$

(2) **Growth Condition**

There exists $K_2 > 0$ such that for all $y \in \mathbb{R}^d$,

$$|b(y)|^2 + ||a(y, y)|| + \int_{|x|<c} |F(y, x)|^2 \nu(dx) \leq K_2(1 + |y|^2). \qquad (2.20)$$

(3) **Big Jumps Condition**

G is jointly measurable and $y \to G(y, x)$ is continuous for all $|x| \geq 1$.

Here, $|| \cdot ||$ is the matrix seminorm $||a|| = \sum_{i=1}^{d} |a_i^i|$, and $a(x, y) = \sigma(x)\sigma(y)^T$.

We also impose the *standard initial condition* $Y(0) = Y_0$ (a.s.) for which Y_0 is independent of $(\mathcal{F}_t, t > 0)$. Solutions of SDEs are Markov processes and, in the case where there are no jumps, diffusion processes.

A special case of considerable interest is

$$dY(t) = L(Y(t-))dX(t).$$

You can check that the conditions given above boil down to the single requirement that L be globally Lipshitz, in order to get existence and uniqueness.

2.10 Stochastic Exponentials

For convenience we take $d = 1$ and consider the problem of finding an adapted process $Z = (Z(t), t \geq 0)$ which has a stochastic differential

$$dZ(t) = Z(t-)dY(t),$$

where Y is a Lévy-type stochastic integral.

The solution of this problem is obtained as follows. We take Z to be the *stochastic exponential* (sometimes called *Doléans-Dade exponential* after its discoverer), which is denoted as $\mathcal{E}_Y = (\mathcal{E}_Y(t), t \geq 0)$ and defined as

$$\mathcal{E}_Y(t) = \exp\left\{Y(t) - \frac{1}{2}[Y_c, Y_c](t)\right\} \prod_{0 \leq s \leq t} (1 + \Delta Y(s))e^{-\Delta Y(s)}, \qquad (2.21)$$

for each $t \geq 0$.

We will need the following assumption:

$$\text{(SE)} \quad \inf\{\Delta Y(t), t > 0\} > -1, \text{ (a.s.)}.$$

Proposition 2.1 *If Y is a Lévy-type stochastic integral of the form (2.14) and (SE) holds, then each $\mathcal{E}_Y(t)$ is almost surely finite.*

Proof. We must show that the infinite product in (2.21) converges almost surely. We write

$$\prod_{0 \leq s \leq t} (1 + \Delta Y(s)) e^{-\Delta Y(s)} = A(t) + B(t),$$

where $A(t) = \prod_{0 \leq s \leq t} (1 + \Delta Y(s)) e^{-\Delta Y(s)} \chi_{\{|\Delta Y(s)| \geq \frac{1}{2}\}}$ and
$B(t) = \prod_{0 \leq s \leq t} (1 + \Delta Y(s)) e^{-\Delta Y(s)} \chi_{\{|\Delta Y(s)| < \frac{1}{2}\}}$.

Now since Y is càdlàg, $\#\{0 \leq s \leq t; |\Delta Y(s)| \geq \frac{1}{2}\} < \infty$ (a.s.), and so $A(t)$ is a finite product. Using the assumption (SE), we have

$$B(t) = \exp\left\{ \sum_{0 \leq s \leq t} [\log(1 + \Delta Y(s)) - \Delta Y(s)] \chi_{\{|\Delta Y(s)| < \frac{1}{2}\}} \right\}.$$

We now employ Taylor's theorem to obtain the inequality

$$\log(1 + y) - y \leq K y^2,$$

where $K > 0$, which is valid whenever $|y| < \frac{1}{2}$. Hence

$$\begin{aligned} \left| \sum_{0 \leq s \leq t} [\log(1 + \Delta Y(s)) - \Delta Y(s)] \chi_{\{|\Delta Y(s)| < \frac{1}{2}\}} \right| &\leq \sum_{0 \leq s \leq t} |\Delta Y(s)|^2 \chi_{\{|\Delta Y(s)| < \frac{1}{2}\}} \\ &< \infty \quad \text{a.s.} \end{aligned}$$

and we have our required result. □

Of course (SE) ensures that $\mathcal{E}_Y(t) > 0$ (a.s.).

Note

(i) The stochastic exponential is, in fact the unique solution of the stochastic differential equation $dZ(t) = Z(t-)dY(t)$, with initial condition $Z(0) = 1$ (a.s.).

(ii) The restrictions (SE) can be dropped and the stochastic exponential extended to the case where Y is an arbitrary (real valued or even complex-valued) càdlàg semimartingale, but the price we have to pay is that $\mathcal{E}_Y$ may then take negative values.

The following alternative form of (2.21) is quite useful :

$$\mathcal{E}_Y(t) = e^{S_Y(t)},$$

$$\begin{aligned}
\text{where } dS_Y(t) &= F(t)dB(t) + \left(G(t) - \frac{1}{2}F(t)^2\right)dt \\
&+ \int_{|x|\geq 1} \log(1 + K(t,x))N(dt,dx) + \int_{|x|<1} \log(1 + H(t,x))\tilde{N}(dt,dx) \\
&+ \int_{|x|<1} (\log(1 + H(t,x)) - H(t,x))\nu(dx)ds. \qquad (2.22)
\end{aligned}$$

Theorem 2.12

$$d\mathcal{E}_Y(t) = \mathcal{E}_Y(t-)dY(t).$$

Proof. We apply Itô's formula to (2.22) to obtain for each $t \geq 0$,

$$\begin{aligned}
d\mathcal{E}_Y(t) &= \mathcal{E}_Y(t-)\left(F(t)dB(t) + G(t)dt + \int_{|x|<1} (\log(1 + H(t,x)) - H(t,x))\nu(dx)dt\right) \\
&+ \int_{|x|\geq 1} [\exp\{S_Y(t-) + \log(1 + K(t,x))\} - \exp(S_Y(t-))]N(dt,dx) \\
&+ \int_{|x|<1} [\exp\{S_Y(t-) + \log(1 + H(t,x))\} - \exp(S_Y(t-))]\tilde{N}(dt,dx) \\
&+ \int_{|x|<1} [\exp\{S_Y(t-) + \log(1 + H(t,x))\} - \exp(S_Y(t-)) \\
&- \log(1 + H(t,x)) \exp S_Y(t-)]\nu(dx)dt) \\
&= \mathcal{E}_Y(t-)[F(t)dB(t) + G(t)dt + K(t,x)N(dt,dx) + H(t,x)\tilde{N}(dt,dx)],
\end{aligned}$$

as required. □

Examples

1. If $Y(t) = \sigma B(t)$ where $\sigma > 0$ and $B = (B(t), t \geq 0)$ is a standard Brownian motion, then

$$\mathcal{E}_Y(t) = \exp\left\{\sigma B(t) - \frac{1}{2}\sigma^2 t\right\}.$$

2. If $Y = (Y(t), t \geq 0)$ is a compound Poisson process, so that each $Y(t) = X_1 + \cdots + X_{N(t)}$, where $(X_n, n \in \mathbb{N})$ are i.i.d. and N is an independent Poisson process, we have

$$\mathcal{E}_Y(t) = \prod_{j=1}^{N(t)} (1 + X_j),$$

for each $t \geq 0$.

If X and Y are Lévy-type stochastic integrals, you can check that

$$\mathcal{E}_X(t)\mathcal{E}_Y(t) = \mathcal{E}_{X+Y+[X,Y]}(t),$$

for each $t \geq 0$.

Let X be a real valued Lévy process with characteristics (b, σ, ν) and associated Lévy-Itô decomposition given by (2.5). For applications to finance, it is useful to know whether the stochastic exponential $\mathcal{E}_X(t)$ can be rewritten as the exponential $\exp(X_1(t))$ of a Lévy process X_1 and vice versa.

Suppose that $\mathcal{E}_X(t) > 0$, then by (2.22) we have $\mathcal{E}_X(t) = \exp(S_X(t))$, where for each $t \geq 0$,

$$\begin{aligned} S_X(t) &= \sigma B(t) + \int_{|x|\geq 1} \log(1+x) N(t, dx) + \int_{|x|<1} \log(1+x) \tilde{N}(t, dx) \\ &+ \left[b - \frac{1}{2}\sigma^2 + \int_{|x|<1} (\log(1+x) - x)\nu(dx) \right] t. \end{aligned} \tag{2.23}$$

Comparing (2.5) with (2.23) we deduce the following:

Theorem 2.13 *If X is a Lévy process with each $\mathcal{E}_X(t) > 0$, then $\mathcal{E}_X(t) = \exp(X_1(t))$ for each $t \geq 0$ where X_1 is the Lévy process with characteristics (b_1, σ_1, ν_1) given by*

$$\nu_1 = \nu \circ f^{-1}, \text{ where } f(x) = \log(1+x),$$

$$b_1 = b - \frac{1}{2}\sigma^2 + \int_{\mathbb{R}-\{0\}} [\log(1+x)\chi_{\hat{B}}(\log(1+x)) - x\chi_{\hat{B}}(x)]\nu(dx),$$

$$\sigma_1 = \sigma.$$

Conversely, there exists a Lévy process with characteristics (b_2, σ_2, ν_2) such that $\exp(X(t)) = \mathcal{E}_{X_2}(t)$ for all $t \geq 0$ wherein

$$\nu_2 = \nu \circ g^{-1}, \text{ where } g(x) = e^x - 1,$$

$$b_2 = b + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}-\{0\}} [(e^x - 1)\chi_{\hat{B}}(e^x - 1) - x\chi_{\hat{B}}(x)]\nu(dx),$$

$$\sigma_2 = \sigma.$$