Stochastic Calculus for Discontinuous Processes

Richard F. Bass
Department of Mathematics
University of Connecticut

October 2, 1998

1. Preliminaries.

Let $\{\mathcal{F}_t\}$ be a right continuous complete filtration. We define the *predictable σ-field* $\mathcal{P}$ as the σ-field on $\Omega \times [0,\infty)$ generated by all processes of the form

$$H(\omega, s) = \sum_{i=1}^{n} K_i(\omega) 1_{(a_i,b_i]}(s), \tag{1.1}$$

where each K_i is $\mathcal{F}_{a_i}$-measurable and bounded. A process that is measurable with respect to $\mathcal{P}$ is called predictable. One can show that left continuous processes, i.e., processes whose paths are left continuous, are measurable with respect to $\mathcal{P}$. If a process has paths that are right continuous with left limits, we say the process is *rcll* (or 'càdlàg' in French). If X_t is a process that is right continuous with left limits, we set $X_{t-} = \lim_{s\to t, s<t} X_s$ and $\Delta X_t = X_t - X_{t-}$. Thus ΔX_t is the size of the jump of X_t at time t.

We say a stopping time T is *predictable* (also known as *previsible*) if there exist stopping times T_n increasing to T with $T_n < T$ on the set $(T < \infty)$. An example is $T = \inf\{t : B_t = 1\}$, where B_t is a Brownian motion; in this case we can take $T_n = \inf\{t : B_t = 1 - 1/n\}$.

A stopping time T is *totally inaccessible* if for all predictable stopping times S we have $\mathbb{P}(T = S) = 0$. An example is $T = \inf\{t : P_t = 1\}$, where P_t is a Poisson process. To see this, suppose S is predictable. Let M be a large integer. It is easy to see that $S \wedge M$ is predictable. If we show $\mathbb{P}(T = S \wedge M) = 0$ for each M, then $\mathbb{P}(S = T) = 0$. So we may suppose S is bounded. Take S_n increasing to S. Note $P_t = 0$ if $t < T$ and $P_T = 1$. The process $P_t - t$ is a martingale, so by optional stopping, $\mathbb{E}\,(P_{S_n \wedge T}) = \mathbb{E}\,(S_n \wedge T) \uparrow \mathbb{E}\,(S \wedge T) = \mathbb{E}\,(P_{S\wedge T}) = \mathbb{P}(S \geq T)$. On the other hand $\mathbb{E}\,(P_{S_n\wedge T}) = \mathbb{P}(S_n \geq T) \to \mathbb{P}(S > T)$. So $\mathbb{P}(S > T) = \mathbb{P}(S \geq T)$, and hence $\mathbb{P}(S = T) = 0$.

Note that if T is predictable, then $1_{[0,T(\omega))} = \lim 1_{[0,T_n(\omega)]}$. But $1_{[0,T_n(\omega)]}$ is a left continuous process, hence $\mathcal{P}$ measurable.

Recall the *Doob-Meyer decomposition.* Suppose X_t is a supermartingale with paths that are right continuous with left limits and that the collection of random variables $\{X_T : T \text{ a stopping time}\}$ is uniformly integrable. There exists a martingale M_t and a predictable increasing process A_t such that $X_t = M_t - A_t$. The decomposition is unique.

Suppose $a(t)$ is a deterministic right continuous nondecreasing function of t. We have the following formula.

Lemma 1.1. *If $a(0) = 0$, then*

$$a(t)^2 = \int_0^t [(a(t) - a(s)) + (a(t) - a(s-))]da(s).$$

Proof. Since an increasing function has at most countably many discontinuities, we may write $a(t) = a^c(t) + \sum_{i=1}^\infty a_i(t)$, where a^c is continuous and nondecreasing, and each a_i is constant except for a single jump. Since a is the limit, uniformly on $[0, t]$, of $a^c + \sum_{i=1}^n a_i$ as $n \to \infty$, it suffices to prove our result when there are only finitely many jumps. Since we can approximate a^c uniformly on $[0, t]$ by step functions, it suffices to prove our result when a is of the form $\sum_{j=1}^m b_j$, where each b_j is constant except for a single jump. But for such a we can prove our result by a direct calculation. □

We will need the following.

Lemma 1.2. *Suppose A_t is a predictable nondecreasing process with $A(0) = 0$. Suppose X and Y are uniformly integrable rcll processes, not necessarily adapted, such that $\mathbb{E}\, X_S \le \mathbb{E}\, Y_S$ for every stopping time S, $X_{\infty-} = Y_{\infty-} = 0$, and $\mathbb{E} \int_0^\infty Y_{s-} dA_s < \infty$. Then*

$$\mathbb{E} \int_0^\infty X_{s-} dA_s \le \mathbb{E} \int_0^\infty Y_{s-} dA_s.$$

Proof. Suppose first that $A_s(\omega)$ is either identically zero as a function of s or else has a single jump of size b. Let $S = \inf\{t : \Delta A_s = b\}$; note $S = \infty$ on those paths that have no jump. Since A_t is predictable, S is predictable, and there exist stopping times S_n increasing up to S. Then $\mathbb{E}\, X_{S_n} \le \mathbb{E}\, Y_{S_n}$ and taking a limit, $\mathbb{E}\, X_{S-} \le \mathbb{E}\, Y_{S-}$. Multiplying by b, this is the same as $\mathbb{E} \int_0^\infty X_{s-} dA_s \le \mathbb{E} \int_0^\infty Y_{s-} dA_s$.

By linearity, the inequality holds if A_t is the sum of finitely many processes of the above form. So it remains to show that every such predictable A can be written as the limit of sums of finitely many such processes. Since increasing processes have countably many jumps, it suffices to consider continuous A_s. If A_s is continuous, let $S_1 = \inf\{t : A_s \ge \varepsilon\}$, $S_{i+1} = \inf\{t > S_i : A_t - A_{S_i} \ge \varepsilon\}$ for $i = 1, 2, \ldots$. If $A_s^i = \varepsilon 1_{(t \ge S_i)}$, then taking the limit first as $n \to \infty$ of $\sum_{i=1}^n A_s^i$, and then the limit as $\varepsilon \to 0$, we approximate A_s. □

2. Decomposition of martingales.

Suppose A_t is a bounded increasing process. Then trivially A_t is a submartingale, and by the Doob-Meyer decomposition (applied to $-A_t$) there exists a predictable increasing process $\widetilde{A}_t$ such that $A_t - \widetilde{A}_t$ is a martingale. We call $\widetilde{A}_t$ the *compensator* of A_t.

Lemma 2.1. *If A_t is bounded by K, then $\mathbb{E}\,\widetilde{A}_\infty^2 \le 2K^2$.*

Proof. We fix t_0, let $X_t = A_\infty - A_t$ and $Y_t = K$ for $t < t_0 + 1$, and $X_t = Y_t = 0$ for $t \ge t_0 + 1$. By Lemma 1.1,

$$\widetilde{A}_{t_0}^2 = \int_0^{t_0} [(\widetilde{A}_{t_0} - \widetilde{A}_s) + [(\widetilde{A}_{t_0} - \widetilde{A}_{s-})] d\widetilde{A}_s \le 2 \int_0^{t_0} (\widetilde{A}_{t_0} - \widetilde{A}_{s-}) d\widetilde{A}_s.$$

Applying Lemma 1.2,

$$\mathbb{E}\,\widetilde{A}_{t_0}^2 \le 2K\mathbb{E} \int_0^{t_0} d\widetilde{A}_s \le 2K^2.$$

We now let $t_0 \to \infty$ and use Fatou's lemma to finish the proof. □

In particular, $\widetilde{A}_t$ is integrable.

A key proposition is the following orthogonality lemma.

Lemma 2.2. *Suppose A_t is a bounded increasing process, and $M_t = A_t - \widetilde{A}_t$, where $\widetilde{A}_t$ is defined as above. Suppose N_t is a square integrable martingale such that $\Delta N_t \Delta M_t = 0$ for all t. Then $\mathbb{E}\, M_\infty N_\infty = 0$.*

Proof. If $H = K 1_{(a,b]}$ with K $\mathcal{F}_a$-measurable, then using the fact that M_t is of bounded variation, we have (this is a Lebesgue-Stieltjes integral here)

$$\mathbb{E} \int_0^\infty H_s dM_s = \mathbb{E}\,[K(M_b - M_a)] = \mathbb{E}\,[K\mathbb{E}\,[M_b - M_a \mid \mathcal{F}_a]\,] = 0.$$

Since N_{s-} is predictable, by linearity and taking limits, $\mathbb{E} \int_0^\infty N_{s-} dM_s = 0$. By hypothesis, $\mathbb{E} \int_0^\infty \Delta N_s dM_s = 0$, so $\mathbb{E} \int_0^\infty N_s dM_s = 0$. On the other hand, by conditioning on $\mathcal{F}_s$ we see

$$\mathbb{E}\, M_\infty N_\infty = \mathbb{E} \int_0^\infty N_\infty dM_s = \mathbb{E} \int_0^\infty N_s dM_s.$$

So $\mathbb{E}\, M_\infty N_\infty = 0$. □

If we apply the above to $N_{t \wedge T}$, we have $\mathbb{E}\, M_\infty N_T = 0$. If we then condition on $\mathcal{F}_T$,

$$\mathbb{E}\, M_T N_T = \mathbb{E}\, N_T \mathbb{E}\,[M_\infty \mid \mathcal{F}_T]\,] = \mathbb{E}\,[N_T M_\infty] = 0.$$

Let M_t be a square integrable rcll martingale, so that $\mathbb{E}\, M_\infty^2 < \infty$. By Doob's inequality we have $\mathbb{E}\,[\sup_{0\le s<\infty} M_s^2] < \infty$. For each i, let $T_{i1} = \inf\{t : |\Delta M_t| \in [2^i, 2^{i+1})\}$, $T_{i2} = \inf\{t > T_{i1} : |\Delta M_t| \in [2^i, 2^{i+1})\}$, and so on. Since M_t is right continuous with left limits, $T_{ij} \to \infty$ as $j \to \infty$. So M_t has at most countably many jumps. We order them as $S_1, S_2, \ldots$ Let $A_i(t) = \Delta M_{S_i} 1_{(t\ge S_i)}$ and $M_i(t) = A_i(t) - \widetilde{A}_i(t)$.

It is worth noting that from the general theory of processes, if S_i is totally inaccessible, then $\widetilde{A}_i(t)$ will be continuous, while if S_i is predictable, then by virtue of the fact that M_t is a martingale, we see that $\widetilde{A}_i(t)$ is identically zero.

Theorem 2.3. *Each M_i is square integrable. If $M_t^c = M_t - \sum_{i=1}^\infty M_i(t)$, then M^c is square integrable. The martingale M^c is orthogonal to each M_i.*

Proof. We have already seen that each M_i is square integrable. By the orthogonality lemma, the M_i are mutually orthogonal and also $M_t - \sum_{i=1}^n M_i(t)$ is orthogonal to $M_1, \ldots, M_n$ for each n.

We will show $M_t - \sum_{i=1}^n M_i(t)$ converges in L^2. Then by Doob's inequality, we see that $\sup_s (M_s - \sum_{i=1}^n M_i(s))$ will converge in L^2. A subsequence will converge a.s., so the limit will have no jumps, and consequently must be a continuous process. Moreover it will be orthogonal to each M_i. Write $S_n(t) = \sum_{i=1}^n M_i(t)$. By the orthogonality,

$$\mathbb{E}\, M_\infty^2 = \mathbb{E}\Big[M_\infty - S_n(\infty) + S_n(\infty)\Big]^2 = \mathbb{E}\Big[M_\infty - S_n(\infty)\Big]^2 + \mathbb{E}\, S_n(\infty)^2 \ge \mathbb{E}\, S_n(\infty)^2.$$

So, using orthogonality again, the series $\mathbb{E}\sum_{i=1}^n M_i(\infty)^2$ converges. Then

$$\mathbb{E}\Big[\Big(M_\infty - S_n(\infty)\Big) - \Big(M_\infty - S_m(\infty)\Big)\Big]^2 = \mathbb{E}\Big[\sum_{i=n+1}^m M_i(\infty)\Big]^2 = \sum_{i=n+1}^m \mathbb{E}\, M_i(\infty)^2.$$

This can be made small by taking n and m large. Hence $M(\infty) - S_n(\infty)$ is a Cauchy sequence in L^2. □

3. Stochastic integrals.

If $a(t)$ is a function of bounded variation and $0 = s_0 \le s_1 \le \cdots \le s_n = t$ is a partition of $[0,t]$, note

$$\begin{aligned} a(t)^2 &= \sum_{i=1}^n a(s_i)^2 - a(s_{i-1})^2 = \sum (a(s_i) + a(s_{i-1}))(a(s_i) - a(s_{i-1})) \\ &= \sum (2a(s_{i-1}) + a(s_i) - a(s_{i-1}))(a(s_i) - a(s_{i-1})). \end{aligned}$$

Passing to a limit, we have

$$a(t)^2 = \int_0^t (2a(s-) + \Delta a(s))\, da(s) = 2\int_0^t a(s-)\, da(s) + \sum_s (\Delta a(s))^2. \tag{3.1}$$

If M_t is a square integrable martingale, then M_t^2 is a submartingale by Jensen's inequality. By the Doob-Meyer decomposition, there exists a predictable increasing process, denoted $\langle M\rangle_t$, such that $M_t^2 - \langle M\rangle_t$ is a martingale. Let us define

$$[M]_t = \langle M^c\rangle_t + \sum_{s\le t} |\Delta M_s|^2.$$

Here M^c is the continuous part of the martingale M_t as defined in Theorem 2.3.

Proposition 3.1. *$M_t^2 - [M]_t$ is a martingale.*

Proof. By the orthogonality it is easy to see that $\langle M\rangle_t = \langle M^c\rangle_t + \sum_i \langle M_i\rangle_t$. Since $M_t^2 - \langle M\rangle_t$ is a martingale, we need only show $[M]_t - \langle M\rangle_t$ is a martingale, or that $\sum_i \langle M_i\rangle_t - \sum_i \sum_{s\le t} |\Delta M_i(s)|^2$ is a martingale.

By (3.1), we see that

$$M_i(t)^2 = 2\int_0^t M_i(s-)\, dM_s + \sum_{s\le t} |\Delta M_i(s)|^2. \tag{3.2}$$

It is easy to check, by approximating by a Riemann sum and using the fact that M_i is a martingale, that the integral on the right in (3.2) is a martingale. So $M_i^2(t) - \sum_{s\le t} |\Delta M_s|^2$ is a martingale. Since $M_i^2(t) - \langle M_i\rangle_t$ is a martingale, that completes the proof. $\square$

If $H(\omega, s)$ is as in (1.1) and M is square integrable, define the stochastic integral by

$$N_t = \int_0^t H_s\, dM_s = \sum_{i=1}^n K_i [M_{b_i\wedge t} - M_{a_i\wedge t}].$$

Just as in [PTA], pp. 43-44, the left hand side will be a martingale and just as in the proof of [PTA], Section I.5, with [ ] instead of $\langle\ \rangle$, $N_t^2 - [N]_t$ is a martingale.

If H is $\mathcal{P}$-measurable and $\mathbb{E}\int_0^\infty H_s^2 d[M]_s < \infty$, approximate H by integrands H_s^n of the form (1.1), define N_t^n as the stochastic integral of H^n with respect to M_t. By the same proof as in [PTA], Section I.5, the martingales N_t^n converge in L^2. We call the limit $N_t = \int_0^t H_s\, dM_s$. The stochastic integral is a martingale, and $[N]_t = \int_0^t H_s^2 d[M]_s$.

4. Ito's formula.

Suppose $X_t = M_t + A_t$, where M_t is a square integrable martingale and A_t is a process of bounded variation with total variation integrable. We will state and prove Ito's formula in this case. The extension to semimartingales without the integrability conditions will be done later and is easy. Define $\langle X^c\rangle_t$ to be $\langle M^c\rangle_t$.

Theorem 4.1. *Suppose X_t is as above and f is C^2 on $\mathbb{R}$ with bounded first and second derivatives. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d\langle X^c\rangle_s \qquad (4.1)$$
$$+ \sum_{s\le t}[f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s].$$

Proof. The proof will be given in several steps.

Step 1: If X_t is continuous, see [PTA], pp. 46-48.

Step 2: Suppose X_t has a single jump at time T and there exists $N > 0$ such that $|\Delta M_T| + |\Delta A_T| \le N$ a.s.

Let $C_t = \Delta M_T 1_{(t\le T)}$; write $C_t^+ = (\Delta M_T)^+ 1_{(t\le T)}$, let $\widetilde{C}^+$ be the compensator for C^+, define C^- and $\widetilde{C}^-$ similarly, and set $\widetilde{C}_t = \widetilde{C}_t^+ - \widetilde{C}_t^-$. If we replace M_t by $M_t - C_t + \widetilde{C}_t$ and A_t by $A_t + C_t - \widehat{C}_t$, we may assume that M_t is continuous.

Let $B_t = \Delta X_T 1_{(t\ge T)}$. Let $\widehat{X}_t = X_t - B_t$. If we set $\widehat{A}_t = A_t - B_t$, then $\widehat{X}_t = M_t + \widehat{A}_t$. Then $\widehat{X}_t$ is a continuous process that agrees with X_t up to time T. By Step 1, we have

$$\begin{aligned}
f(\widehat{X}_t) &= f(\widehat{X}_0) + \int_0^t f'(\widehat{X}_s)d\widehat{X}_s + \frac{1}{2}\int_0^t f''(\widehat{X}_s)d\langle M\rangle_s \\
&= f(\widehat{X}_0) + \int_0^t f'(\widehat{X}_{s-})d\widehat{X}_s + \frac{1}{2}\int_0^t f''(\widehat{X}_{s-})d\langle M\rangle_s \\
&= f(\widehat{X}_0) + \int_0^t f'(\widehat{X}_{s-})d\widehat{X}_s + \frac{1}{2}\int_0^t f''(\widehat{X}_{s-})d\langle M\rangle_s \\
&\quad + \sum_{s\le t}[f(\widehat{X}_s) - f(\widehat{X}_{s-}) - f'(\widehat{X}_{s-})\Delta\widehat{X}_s].
\end{aligned}$$

Since $\widehat{X}_t$ agrees with X_t up to time T, the above holds with $\widehat{X}_t$ replaced by X_t for $t < T$. At time T the left hand side has a jump of size $f(X_T) - f(X_{T-})$. The right hand side will jump $f'(X_{T-})\Delta X_T + f(X_T) - f(X_{T-}) - f'(X_{T-})\Delta X_T$, which is the same. So in this case, (4.1) holds.

Step 3: There exist times $T_1 < T_2 < \cdots$ with $T_n \to \infty$, $|\Delta M_{T_i}|$ and $|\Delta A_{T_i}|$ are bounded by N_i for some $N_i > 0$ for each i, and X_t is continuous except at one of the times $T_1, T-2, \ldots$

Let $X'_t = X_{(t-T_i)\vee 0}$, and similarly for A'_t and M'_t. If we apply Step 2 to X' at time $T_i + t$, we have for $t > T_i$

$$f(X_t) = f(X_{T_i}) + \int_{T_i}^t f'(X_{s-})\,dX_s + \frac{1}{2}\int_{T_i}^t f''(X_{s-})\,d\langle X^c\rangle_s$$
$$+ \sum_{T_i<s\le t}[f(X_s) - f(X_{s-}) - f(X_{s-})\Delta X_s].$$

So for any t we have

$$f(X_t) = f(X_{T_i \wedge t}) + \int_{T_i \wedge t}^{t} f'(X_{s-})\, dX_s + \frac{1}{2} \int_{T_i \wedge t}^{t} f''(X_{s-})\, d\langle X^c \rangle_s$$
$$+ \sum_{T_i \wedge t < s \le t} [f(X_s) - f(X_{s-}) - f(X_{s-}) \Delta X_s].$$

Summing over i, we have (4.1) for each t.

Step 4: The general case.

As in the paragraphs preceding Theorem 2.3, we order the jumps of X_t as $S_1, S_2, \dots,$ so that for each i there is an N_i with $|\Delta M_{S_i}| + |\Delta A_{S_i}| \le N_i$. Let M be decomposed into M^c and M_i as in Theorem 2.3 and let $A_t^c = A_t - \sum_{i=1}^{\infty} \Delta A_{S_i} 1_{(t \le S_i)}$. Since A_t is of bounded variation, then A^c will be finite and continuous. Let $X_t^n = M_t^n + A_t^n$, where $M_t^n = M_t^c + \sum_{i=1}^{n} M_i(t)$ and $A_t^n = A_t^c + \sum_{i=1}^{n} \Delta A_{S_i} 1_{(t \le S_i)}$. Then M^n converges to M in L^2 and A^n converges to A in L^1.

We now examine the various terms. Uniformly in t over compact intervals, X_t^n converges to X_t in probability. Since $\int_0^t d\langle M^c \rangle_s < \infty$, by dominated convergence

$$\int_0^t f'(X_{s-}^n) d\langle M^c \rangle_s \to \int_0^t f'(X_{s-}) d\langle M^c \rangle_s$$

in probability.

We write

$$\begin{aligned}
&\int_0^t f'(X_{s-}^n) dA_s^n - \int_0^t f'(X_{s-}) dA_s \\
&\qquad = \Big[\int_0^t f'(X_{s-}^n) dA_s^n - \int_0^t f'(X_{s-}) dA_s^n \Big] \\
&\qquad\quad + \Big[\int_0^t f'(X_{s-}) dA_s^n - \int_0^t f'(X_{s-}) dA_s \Big] = I_1 + I_2.
\end{aligned}$$

The absolute value of I_1 is less than

$$\int_0^t |f'(X_{s-}^n) - f'(X_{s-})|\, |dA_s^n| \le \int_0^t |f'(X_{s-}^n) - f'(X_{s-})|\, |dA_s|;$$

this tends to 0 as $n \to \infty$ by dominated convergence.

$$I_2| \le \|f'\|_\infty \int_0^t |dA_s^n - dA_s|,$$

which also tends to 0.

We also write

$$\begin{aligned}\int_0^t f'(X^n_{s-})dM^n_s - \int_0^t f'(X_{s-})dM_s \\ &= \Big[\int_0^t f'(X^n_{s-})dM^n_s - \int_0^t f'(X_{s-})dM^n_s\Big] \\ &\quad + \Big[\int_0^t f'(X_{s-})dM^n_s - \int_0^t f'(X_{s-})dM_s\Big] = I_3 + I_4.\end{aligned}$$

The L^2 norm of I_3 is bounded by

$$\mathbb{E}\int_0^t |f'(X^n_{s-}) - f'(X_{s-})|^2 d[M^n]_s \le \mathbb{E}\int_0^t |f'(X^n_{s-}) - f'(X_{s-})|^2 d[M]_s,$$

which goes to 0 by dominated convergence. And

$$I_4 = \int_0^t f'(X_{s-}) \sum_{i=n+1}^\infty dM_i(s),$$

so the L^2 norm of I_4 is less than

$$\|f'\|^2_\infty \sum_{i=n+1}^\infty \mathbb{E}\,[M_i]_\infty \le \|f'\|^2_\infty \sum_{i=n+1}^\infty \mathbb{E}\,M_i(\infty)^2,$$

which goes to 0 as $n \to \infty$.

Finally,

$$\begin{aligned}I_5 &= \sum_{s\le t}[f(X_s - f(X_{s-}) - f'(X_{s-})\Delta X_s] - \sum_{s\le t}[f(X^n_s) - f(X^n_{s-}) - f'(X^n_{s-})\Delta X^n_s] \\ &= \sum_{i=n+1}^\infty [f(X_{S_i}) - f(X_{S_i-}) - f'(X_{S_i-})\Delta X_{S_i}]1_{(t\ge S_i)}.\end{aligned}$$

This is bounded in absolute value by

$$\sum_{i=n+1}^\infty \|f''\|_\infty |\Delta X_{S_i}|^2 1_{(|\Delta X(S_i)|\le 1)} + \sum_{i=n+1}^\infty 2\|f'\|_\infty |\Delta X_{S_i}| 1_{(|\Delta X(S_i)|>1)}.$$

If $|\Delta X_{S_i}| \le 1$, then

$$|\Delta X_{S_i}|^2 \le 2|\Delta M_{S_i}|^2 + 2|\Delta A_{S_i}|,$$

while if $|\Delta X_{S_i}| > 1$,

$$|\Delta X_{S_i}| \le |\Delta M_{S_i}|^2 + |\Delta A_{S_i}|.$$

Since $\sum_{i=1}^\infty |\Delta M_{S_i}|^2 \le [M]_\infty < \infty$ and $\sum_{i=1}^\infty |\Delta A_{S_i}| < \infty$, then I_5 tends to 0 as $n \to \infty$.

This completes the proof. □

The following corollary is very useful.

Corollary 4.2. *If X and Y are semimartingales of the above form,*

$$X_tY_t = X_0Y_0 + \int_0^t X_{s-}dY_s + \int_0^t Y_{s-}dX_s + [X,Y]_t.$$

Here $[X,Y]_t$ is defined by polarization.

Proof. Apply Theorem 4.1 with $f(x) = x^2$ (this f does not have bounded first derivative, but that may be circumvented by an approximation argument.) Since in this case $f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s = \Delta X_s^2$, we obtain

$$X_t^2 = X_0^2 + 2\int_0^t X_{s-}dX_s + [X]_t.$$

We obtain the corollary by polarization. □

5. The reduction theorem.

Let M be an adapted process. If there exist stopping times increasing to ∞ such that each process $M_{t\wedge T_n}$ is a uniformly integrable martingale, we say M is a *local martingale*. If each $M_{t\wedge T_n}$ is square integrable, we say M is locally square integrable. We say a stopping time T *reduces* a process M if $M_{t\wedge T}$ is a uniformly integrable martingale.

Lemma 5.1. *(a) The sum of two local martingales is a local martingale.*

(b) If S and T reduce M, then so does $S \vee T$.

(c) If there exist times $T_n \to \infty$ such that $M_{t\wedge T_n}$ is a local martingale for each n, then M is a local martingale.

Proof. (a) If the sequence S_n reduces M and the sequence T_n reduces N, then $S_n \wedge T_n$ will reduce $M + N$.

(b) $M_{t\wedge(S\vee T)}$ is bounded in absolute value by $|M_{t\wedge T}| + |M_{t\wedge S}|$. Both $\{|M_{t\wedge T}|\}$ and $\{|M_{t\wedge S}|\}$ are uniformly integrable families of random variables.

(c) Let S_{nm} be a family of stopping times reducing $M_{t\wedge T_n}$ and let $S'_{nm} = S_{nm} \wedge T_n$. Renumber the stopping times into a single sequence $R_1, R_2, \ldots$ and let $H_k = R_1 \vee \cdots \vee R_k$. Note $H_k \uparrow \infty$. To show that H_k reduces M, we need to show that R_i reduces M and use (b). But $R_i = S'_{nm}$ for some m, n, so $M_{t\wedge R_i} = M^{T_n}_{t\wedge S_{nm}}$ is a uniformly integrable martingale, where here $M^{T_n}_t = M_{t\wedge T_n}$. □

Let M be a local martingale with $M_0 = 0$. We say that a stopping time T *strongly reduces* M if T reduces M and the martingale $\mathbb{E}[|M_T| \mid \mathcal{F}_s]$ is bounded on $[0, T)$.

Lemma 5.2. *(a) If T strongly reduces M and $S \le T$, then S strongly reduces M.*

(b) If S and T strongly reduce M, then so does $S \vee T$.

(c) $\mathbb{E}\,[\mathbb{E}\,[\,\cdot \mid \mathcal{F}_T] \mid \mathcal{F}_S] = \mathbb{E}\,[\,\cdot \mid \mathcal{F}_{S\wedge T}]]$.

Proof. (a) Let $Y_s = \mathbb{E}\,[|M_T| \mid \mathcal{F}_s]$. Note $\mathbb{E}\,[|M_S| \mid \mathcal{F}_s] \le Y_{S\wedge s}$, hence S strongly reduces M.

(b) It suffices to show that $\mathbb{E}\,[|M_{S\vee T}| \mid \mathcal{F}_t]$ is bounded for $t < T$, since by symmetry the same will hold for $t < S$. For $t < T$ this expression is bounded by

$$\mathbb{E}\,[|M_T| \mid \mathcal{F}_t] + \mathbb{E}\,[|M_S| 1_{(S>T)} \mid \mathcal{F}_t].$$

The first term is bounded since T strongly reduces M. The second term is bounded for $t < T$ by $\mathbb{E}\,[|M_S|1_{(t<S)} \mid \mathcal{F}_t] = \mathbb{E}\,[|M_S| \mid \mathcal{F}_t]1_{(t<S)}$, which in turn is bounded since S strongly reduces M.

(c) Let Y_∞ be a bounded random variable and let $Y_t = \mathbb{E}\,[Y_\infty \mid \mathcal{F}_t]$. $Z_t = Y_{t\wedge T}$ is a martingale with $Z_\infty = Y_T$. So

$$\mathbb{E}\,[\mathbb{E}\,[Y_\infty \mid \mathcal{F}_T] \mid \mathcal{F}_s] = \mathbb{E}\,[Y_T \mid \mathcal{F}_S] = \mathbb{E}\,[Z_\infty \mid \mathcal{F}_S] = Z_S = Y_{S\wedge T} = \mathbb{E}\,[Y_\infty \mid \mathcal{F}_{S\wedge T}].$$

□

Lemma 5.3. *If M is a local martingale with $M_0 = 0$, then there exist stopping times $T_n \uparrow \infty$ that strongly reduce M.*

Proof. Let $R_n \uparrow \infty$ be a sequence reducing M. Let

$$S_{nm} = R_n \wedge \inf\{t : \mathbb{E}\,[|M_{R_n}| \mid \mathcal{F}_t] \ge m\}.$$

Arrange the stopping times S_{nm} into a single sequence U_n and let $T_n = U_1 \vee \cdots \vee U_n$. In view of the preceding lemmas, we need to show U_i strongly reduces M, or that S_{nm} does for each n and m.

Let $Y_t = \mathbb{E}\,[|M_{R_n}| \mid \mathcal{F}_t]$. Y is bounded by m on $[0, S_{nm})$. We write

$$\begin{aligned}
\mathbb{E}\,[|M_{S_{nm}}|1_{(t<S_{nm})}] &= \mathbb{E}\,[|\mathbb{E}\,[|M_{R_n}|1_{(t<S_{nm})} \mid \mathcal{F}_{S_{nm}}] \mid \mathcal{F}_t] \\
&\le \mathbb{E}\,[\mathbb{E}\,[|M_{R_n}|1_{(t<S_{nm})} \mid \mathcal{F}_{S_{nm}}] \mid \mathcal{F}_t] \\
&= \mathbb{E}\,[|M_{R_n}|1_{(t<S_{nm})} \mid \mathcal{F}_{S_{nm}\wedge t}] \\
&= Y_{S_{nm}\wedge t}1_{(t<S_{nm})} \\
&= Y_t 1_{(t<S_{nm})} \le m.
\end{aligned}$$

□

Our main theorem of this section is the following.

Theorem 5.4. *Suppose M is a local martingale. Then there exist stopping times $T_n \uparrow \infty$ such that $M_{t\wedge T_n} = U_t^n + V_t^n$, where U^n is a square integrable martingale and V^n is a process of bounded variation whose total variation is integrable. Moreover, $U_t = U_T$ and $V_t = V_T$ for $t \geq T$.*

Proof. It suffices to prove that if M is a local martingale with $M_0 = 0$ and T strongly reduces M, then $M_{t\wedge T}$ can be written as $U + V$ with U and V of the described form. Thus we may assume $M_t = M_T$ for $t > T$, $|M_T|$ is integrable, and $\mathbb{E}\,[|M_T| \mid \mathcal{F}_t]$ is bounded, say by K, on $[0, T)$.

Let $C_t = M_T 1_{(t\geq T)} = M_t 1_{(t\geq T)}$ and $X_t = M_t 1_{(t<T)}$. Define $\widetilde{C}$ as in the proof of Theorem 4.1, let $V = C - \widetilde{C}$, and let $U = X + \widetilde{C}$. Then V is a martingale of bounded variation, and the expectation of the total variation is at most $2\mathbb{E}\,|M_T|$. We need to show U is square integrable. Since X_t is bounded by K, it suffices to show $\widetilde{C}$ is square integrable.

Let $C_t^+ = M_t^+ 1_{(t\geq T)}$, $X_t^+ = M_t^+ 1_{(t<T)}$, and let $\widetilde{C}_t^+$ be the compensator of $\widetilde{C}_t^+$. We define C^- and X^- similarly. Note $\mathbb{E}\,[\widetilde{C}_\infty^+ - \widetilde{C}_{S_n}^+] = \mathbb{E}\,X_{S_n}^+$ for every stopping time S_n. If S is predictable and the S_n increase up to S, we then have $\mathbb{E}\,[\widetilde{C}_\infty^+ - \widetilde{C}_{S-}^+] = \mathbb{E}\,X_{S-}^+$. So by Lemmas 1.1 and 1.2,

$$
\begin{aligned}
\mathbb{E}\,\widetilde{C}^+\infty^2 &\leq 2\mathbb{E}\int_0^\infty [\widetilde{C}_\infty - \widetilde{C}_{s-}]d\widetilde{C}_s \\
&= 2\mathbb{E}\int_0^\infty X_{s-}^+ d\widetilde{C}_s \leq 2K\int_0^\infty \widetilde{C}_s \\
&= 2K\mathbb{E}\,C_\infty^+ \leq 2K^2.
\end{aligned}
$$

We have a similar expression for C^-, and combining shows C is square integrable. $\square$

6. Semimartingales.

We define a *semimartingale* to be a process of the form $X_t = X_0 + M_t + A_t$, where X_0 is finite, M_t is a local martingale, and A_t is a process whose paths have bounded variation on $[0, t]$ for each t.

If M_t is a local martingale, let T_n be a sequence of stopping times as in Theorem 5.4. We set $M_{t\wedge T_n}^c = (U^n)_t^c$ for each n and $[M]_{t\wedge T_n} = \langle M^c\rangle_{t\wedge T_n} + \sum_{s\leq t\wedge T_n} \Delta M_s^2$. It is easy to see that these definitions are independent of the decomposition of M into $U^n + V^n$ and of which sequence of stopping times T_n reducing M we choose. We define $\langle X^c\rangle_t = \langle M^c\rangle_t$ and similarly $[X]_t$.

We say an adapted process H is *locally bounded* if there exist stopping times $S_n \uparrow \infty$ and constants K_n such that on $[0, S_n]$ the process H is bounded by K_n. If X_t is a semimartingale and H is a locally bounded predictable process, define $\int_0^t H_s\,dX_s$ as follows. Let $X_t = X_0 + M_t + A_t$. If $R_n = T_n \wedge S_n$, where the T_n are as in Theorem

5.4 and the S_n are as in the definition of locally bounded, set $\int_0^{t\wedge R_n} H_s\, dM_s$ to be the usual stochastic integral. Define $\int_0^{t\wedge R_n} H_s\, dA_s$ to be the usual Lebesgue-Stieltjes integral. Define the stochastic integral with respect to X as the sum of these two. Since $R_n \uparrow \infty$, this defines $\int_0^t H_s\, dX_s$ for all t. One needs to check that the definition does not depend on the decomposition of X into M and A nor on the choice of stopping times R_n. See Meyer (1976) for details.

We now state the general Ito formula.

Theorem 6.1. *Suppose X is a semimartingale and f is C^2. Then*

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-})\, dX_s + \frac{1}{2}\int_0^t f''(X_{s-}) d\langle X^c\rangle_s + \sum_{s\le t}[f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s].$$

Proof. (See Meyer,(1976) for details.) First suppose f has bounded first and second derivatives. Let T_n be stopping times strongly reducing M_t, let $S_n = \inf\{t : \int_0^t |dA_s| \ge n\}$, let $R_n = T_n \wedge S_n$, and let $X_t^n = X_{t\wedge R_n} - \Delta A_{R_n}$. Since the total variation of A_t is bounded on $[0, R_n)$, it follows that X^n is a semimartingale which is the sum of a square integrable martingale and a process whose total variation is integrable. We apply Theorem 4.1 to this process. X_t^n agrees with X_t on $[0, R_n)$. As in the proof of Theorem 4.1, by looking at the jump at time R_n, both sides of Ito's formula jump the same amount at time R_n, and so Ito's formula holds for X_t on $[0, R_n]$. If we now only assume that f is C^2 but approximate f by C^2 functions that have bounded first and second derivatives, it is not hard to see that Ito's formula holds for f without the assumption of bounded derivatives. We finally let $n \to \infty$. $\square$

The proof of the following corollary is similar to the proof of Ito's formula.

Corollary 6.2. *If $X_t = (X_t^1, \ldots, X_t^d)$ is a process taking values in $\mathbb{R}^d$ such that each component is a semimartingale, and f is a C^2 function on $\mathbb{R}^d$, then*

$$f(X_t) = f(X_0) + \int_0^t \sum_{i=1}^d f_i(X_{s-})\, dX_s^i + \tfrac{1}{2}\int_0^t \sum_{i,j=1}^d f_{ij}(X_{s-}) d\langle (X^i)^c, (X^j)^c\rangle_s + \sum_{s\le t}[f(X_s) - f(X_{s-}) - \sum_{i=1}^d f_i(X_{s-})\Delta X_s^i],$$

where f_i and f_{ij} denote first and second partial derivatives.

7. Exponential of a semimartingale.

A function with finite total variation is *purely discontinuous* if it has no continuous part, i.e., $a(t) = \sum_{s\le t} \Delta a(s)$.

Theorem 7.1. *Let X_t be a semimartingale. Define*

$$Z_t = Z_0 \exp\left(X_t - \tfrac{1}{2}\langle X^c\rangle_t\right) \prod_{0\le s\le t}(1+\Delta X_s)e^{-\Delta X_s}.$$

Then Z_t is a semimartingale, the product is a process of bounded variation whose paths are purely discontinuous, and Z_t satisfies

$$Z_t = Z_0 + \int_0^t Z_{s-}\, dX_s.$$

Proof. Since the product of finitely many functions of bounded variation which are purely discontinuous will give a function of the same type and there are only finitely many jumps of X_t of size larger than $1/2$ or less than $-1/2$ in every finite time interval, it suffices to consider

$$V_t' = \prod_{0\le s\le t}(1+\Delta X_s)e^{-\Delta X_s}1_{(|\Delta X_s|\le 1/2)}.$$

Note

$$\log V_t' = \sum_{s\le t}(\log(1+\Delta X_s) - \Delta X_s)1_{(|\Delta X_s|\le 1/2)},$$

which is bounded in absolute value by a constant times $\sum_{s\le t}\Delta X_s^2 < \infty$. It follows that $V_t' = \exp(\log V_t')$ is a purely discontinuous process, and consequently V is also.

We apply the multivariate version of Ito's formula (Corollary 6.2). Let $f(x,y) = e^x y$ and let $Z_t = f(K_t, V_t)$ where $K_t = X_t - \frac{1}{2}\langle X^c\rangle_t$. We obtain

$$\begin{aligned} Z_t - Z_0 &= \int_0^t Z_{s-}dK_s + \int_0^t e^{K_{s-}}dV_s + \frac{1}{2}\int_0^t Z_{s-}d\langle K^c\rangle_t \\ &\quad + \sum_{s\le t}[Z_s - Z_{s-} - Z_{s-}\Delta K_s - e^{-K_{s-}}\Delta V_s] \\ &= I_1 + I_2 + I_3 + I_4. \end{aligned}$$

In I_1 replace dK_s by $dX_s - \frac{1}{2}d\langle X^c\rangle_s$ and in I_3 replace K^c by X^c. Since V_t is purely discontinuous, then $I_2 = \sum_{s\le t} e^{K_{s-}}\Delta V_s$. To simplify I_4, note than $Z_s = Z_{s-}(1+\Delta X_s)$ and $Z_{s-}\Delta K_s = Z_{s-}\Delta X_s$. □

8. The Girsanov theorem.

Let $\mathbb{P}$ and $\mathbb{Q}$ be two equivalent probability measures. Let M_∞ be the density of $\mathbb{Q}$ with respect to $\mathbb{P}$ and let $M_t = \mathbb{E}[M_\infty \mid \mathcal{F}_t]$. The martingale M_t is uniformly integrable. Once a nonnegative martingale hits zero, it is easy to see that it must be zero from then on. Since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, then $M_\infty > 0$ a.s., and so M_t never equals zero, a.s. It is easy to see that M_T is the density of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{F}_T$.

Let L_t be the local martingale defined by

$$L_t = \int_0^t \frac{1}{M_{s-}} dM_s,$$

so that M is the exponential of L.

Theorem 8.1. *Suppose X is a local martingale with respect to $\mathbb{P}$. Then $X_t - D_t$ is a local martingale with respect to $\mathbb{Q}$, where*

$$D_t = \int_0^t \frac{1}{M_s} d[X, M]_s = \int_0^t \frac{M_{s-}}{M_s} d[X, L]_s.$$

Proof. We need to show that $M(X - D)$ is a martingale with respect to $\mathbb{P}$. We see that

$$d(M(X - D)) = (X - D)_{s-} dM_s + M_{s-} dX_s - M_{s-} dD_s + d[M, X - D]_s.$$

The first two terms on the right are local martingales with respect to $\mathbb{P}$. Since D is of bounded variation, $[M, D] = \sum \Delta M_s \Delta D_s = \int_0^t \Delta M_s \, dD_s$. So $M(X - D)$ is a local martingale plus $[M, X]_s - \int_0^t M_s \, dD_s$. Substituting in the formula for D shows that $M(X - D)$ is a local martingale. $\square$

9. Poisson point processes.

Let (S, λ) be an arbitrary measure space (letting $S = \mathbb{R}$ and λ be Lebesgue measure will usually do). For each $\omega \in \Omega$ let $\mu(\omega, dt, dz)$ be a measure $[0, \infty) \times S$. The random measure μ is a *Poisson point process* if (i) for each set $A \subset S$ with $\lambda(A) < \infty$ the process $\mu([0, t] \times A)$ is a Poisson process with parameter $\lambda(A)$ and (ii) if $A_1, \ldots, A_n$ are disjoint subsets of S with $\lambda(A_i) < \infty$, then the processes $\mu([0, t] \times A_i)$ are independent.

The model to keep in mind is where $S = \mathbb{R}$ and λ is Lebesgue measure. At time t there may be a point with value z. The number of points with values in a set A is a Poisson process with parameter $\lambda(A)$. Since $\lambda(\mathbb{R}) = \infty$, there are infinitely many points in every time interval.

Define a non-random measure ν by $\nu([0, t] \times A) = t\lambda(A)$. If $\lambda(A) < \infty$, then $\mu([0, t] \times A) - \nu([0, t] \times A)$ is the same as a Poisson process minus its mean, hence is a (square integrable) martingale.

We can define a stochastic integral with respect to the point process $\mu - \nu$ as follows. Suppose $H(\omega, s, z)$ is of the form

$$H(\omega, s, z) = \sum_{i=1}^n K_i(\omega) 1_{(a_i, b_i]}(s) 1_{A_i}(z), \tag{9.1}$$

where for each i the random variable K_i is bounded and $\mathcal{F}_{a_i}$-measurable and $A_i \subset S$ with $\lambda(A_i) < \infty$. We define

$$N_t = \int_0^t H(\omega, s, z)\, d(\mu - \nu)(ds, dz) = \sum_{i=1}^n K_i(\mu - \nu)(([a_i, b_i) \cap [0, t]) \times A_i).$$

Let us assume without loss of generality that the A_i are disjoint. By linearity it is easy to see that N_t is a martingale. It is also easy to see that $N^c = 0$ and

$$[N]_t = \int_0^t H(\omega, s, z)^2 d\mu(ds, dz). \tag{9.2}$$

Since $\langle N \rangle_t$ must be predictable and all the jumps of N are totally inaccessible, it follows from the general theory of processes that $\langle N \rangle_t$ is continuous. Since $[N]_t - \langle N \rangle_t$ is a martingale, we conclude

$$\langle N \rangle_t = \int_0^t H(\omega, s, z)^2 d\nu(ds, dz). \tag{9.4}$$

Suppose $H(\omega, s, z)$ is predictable process in the following sense: H is measurable with respect to the σ-field generated by all processes of the form (9.1). Suppose also that $\mathbb{E} \int_0^\infty H(s, z)^2 d\nu(ds, dz) < \infty$. Take processes H^n of the form (9.1) converging to H in the space L^2 with norm $(\mathbb{E} \int_0^\infty H^2 d\nu)^{1/2}$. The corresponding $N_t^n = \int_0^t H^n(s, z) d(\mu - \nu)$ are easily seen to be a Cauchy sequence in this L^2 space, and the limit N_t we call the *stochastic integral of H with respect to $\mu - \nu$*. As in the continuous case, we may prove that $\mathbb{E}\, N_t^2 = \mathbb{E}\, [N]_t = \mathbb{E}\, \langle N \rangle_t$, and it follows from this, (9.2), and (9.3) that

$$[N]_t = \int_0^t H(s, z)^2 \mu(ds, dz), \qquad \langle N \rangle_t = \int_0^t H(s, z)^2 \nu(ds, dz). \tag{9.4}$$

One may think of the stochastic integral as follows: if μ has a point at time t with value z, then N_t jumps at this time t and the size of the jump is $H(t, z)$.

Now consider a stochastic differential equation that has a jump component. Look at

$$dX_t = \sigma(X_t)\, dW_t + b(X_t)\, dt + F(X_{t-}, z)\, d(\mu - \nu), \qquad X_0 = x_0. \tag{9.5}$$

This means

$$X_t = x_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t b(X_s) ds + \int_0^t F(X_{s-}, z)(\mu - \nu)(ds, dz).$$

Here W_t is a standard Brownian motion on $\mathbb{R}$. If σ and b are bounded and Lipschitz, $\int \sup_x |F(x,z)|^2 \lambda(dz) < \infty$ and

$$\int |F(x,z) - F(y,z)|^2 \lambda(dz) \le c_1 |x-y|^2$$

for all x, y, then the standard Picard iteration procedure works to prove there exists a solution to (9.5) and that solution is pathwise unique.

10. Lévy processes.

A *Lévy process* is a rcll process with stationary and independent increments. This means that the law of $X_t - X_s$ depends only on $t-s$ and $X_t - X_s$ is independent of $\sigma(X_r; r \le s)$.

We want to examine the structure of Lévy processes. We have three examples already: a Poisson process, a Brownian motion, and the deterministic process $X_t = t$. It turns out all Lévy processes can be built up out of these as building blocks.

Theorem 10.1. *Suppose m is a measure on $\mathbb{R}$ with $m(\{0\}) = 0$ and $\int (1 \wedge x^2) m(dx) < \infty$. There exists a Lévy process X_t such that*

$$\mathbb{E}\, e^{iuX_t} = \exp\Big(t\Big\{iub - \sigma^2 u^2/2 + \int_{\mathbb{R}} [e^{iux} - 1 - iux 1_{(|x|\le 1)}] m(dx)\Big\}\Big). \tag{10.1}$$

The above equation is called the *Lévy-Khintchine formula*. If we let

$$m(dx) = \frac{1+x^2}{x^2} m'(dx)$$

and

$$b = b' + \int_{(|x|\le 1)} \frac{x^3}{1+x^2} m(dx) - \int_{(|x|>1)} \frac{x}{1+x^2} m(dx),$$

then we can also write

$$\mathbb{E}\, e^{iuX_t} = \exp\Big(t\Big\{iub' - \sigma^2 u^2/2 + \int_{\mathbb{R}} \Big[e^{iux} - 1 - \frac{iux}{1+x^2}\Big] \frac{1+x^2}{x^2} m'(dx)\Big\}\Big).$$

Both expressions for the Lévy-Khintchine formula are in common use.

Proof of Theorem 10.1. Let P_t be a Poisson process with parameter λ. P_t is a Lévy process and the characteristic function is $\exp(t\lambda(e^{iu} - 1))$. So the characteristic function of aP_t is $\exp(t\lambda(e^{iua} - 1))$ and the characteristic function of $aP_t - a\lambda t$ is $\exp(t\lambda(e^{iua} - 1 - iua))$. If we let m be the measure on $\mathbb{R}$ defined by $m(dx) = \lambda \delta_a(dx)$, where $\delta_a(dx)$ is point mass at a, then the characteristic function for aP_t is

$$\exp(t \int_{\mathbb{R}} [e^{iux} - 1]\, m(dx) \tag{10.2}$$

and the one for $aP_t - a\lambda t$ is

$$\exp(t \int_{\mathbb{R}} [e^{iux} - 1 - iux]\, m(dx). \tag{10.3}$$

Now let X_t^i be Lévy processes of the form $a_i P_t^i - a_i \lambda_i t$, where the P_t^i are independent Poisson processes with parameter λ_i. Clearly a finite sum of independent Lévy processes is a Lévy process, and so $X_t = \sum_{i=1}^n X_t^i$ is a Lévy process. Moreover the characteristic function of a sum of independent random variables is the product of the characteristic functions, so the characteristic function of X_t is given by (10.3), with $m(dx) = \sum_{i=1}^n \lambda_i \delta_{a_i}(dx)$.

Recall that if φ is the characteristic function of a random variable Y, then $\varphi'(0) = i\mathbb{E}\, Y$ and $\varphi''(0) = -\mathbb{E}\, Y^2$. If X_t is as in the paragraph above, then clearly $\mathbb{E}\, X_t = 0$, and from what we just said, we see that

$$\mathbb{E}\, X_t^2 = t \int x^2\, m(dx).$$

Now let $m(dx)$ be a finite measure with compact support giving no mass to the point 0, let $m_n(dx)$ be purely atomic measures converging weakly to $m(dx)$ with support contained in the support of m and giving no mass to 0, and let X_t^n be Lévy processes corresponding to the measures m_n. Since $\mathbb{E}\,(X_t^n)^2 = \int x^2 m_n(dx)$ is uniformly bounded, we see that X_t^n converges weakly to a random variable whose characteristic function is given by (10.3). Using the bounds on the second moment, it is not hard to see that the limit, call it X_t, is again a Lévy process.

Next let $m(dx)$ be a measure supported on $(0,1]$ with $\int x^2 m(dx) < \infty$. Let $m_n(dx)$ be the measure m restricted to $(2^{-n}, 2^{-n+1}]$. Let X_t^n be independent Lévy processes whose characteristic functions are given by (10.3) with m replaced by m_n. Note $\mathbb{E}\, X_t^n = 0$ for all n and by the independence

$$\mathbb{E}\Big(\sum_{n=0}^N X_t^n\Big)^2 = \sum_{n=0}^N \mathbb{E}\,(X_t^n)^2 = \sum_{n=0}^N \int x^2\, m_n(dx) = \int_{2^{-N}}^1 x^2\, m(dx).$$

By our assumption on m, this shows that $\sum_{n=0}^N X_t^n$ converges in L^2 for each t. Call the limit X_t. X_t has independent and stationary increments, and by Doob's inequality and the L^2 convergence, X_t will have rcll paths. If we do a similar procedure for m restricted to $[-1,0)$ and add the two Lévy processes, we end up with a Lévy process corresponding to m.

If m is a measure supported in $(1,\infty)$ with $m(\mathbb{R}) < \infty$, we do a similar procedure starting with Lévy processes whose characteristic functions are of the form (10.2). We let $m_n(dx)$ be the restriction of m to $(2^n, 2^{n+1}]$, let X_t^n be the corresponding Lévy process, and form $X_t = \sum_{n=0}^\infty X_t^n$. Since $m(\mathbb{R}) < \infty$, for each t_0, the number of times t less than

t_0 at which one of the X_t^n jumps is finite. This shows X_t is rcll, and it is easy to then see that X_t is a Lévy process. We do a similar procedure for m supported on $(-\infty, 1)$ and add the two processes together.

Finally, suppose $\int x^2 \wedge 1\, m(dx) < \infty$. Let X_t^1 be a Lévy process with characteristic function given by (10.2) with m replaced by the restriction of m to $[-1,1]^c$, let X_t^2 be a Lévy process with characteristic function given by (10.3) with m replaced by the restriction of m to $[-1,1]$, let $X_t^3 = bt$, and let X_t^4 be σ times a Brownian motion. Suppose the X^i are all independent. Then their sum will be a Lévy process whose characteristic function is given by (10.1). $\square$

It is clear from the construction that if $m(A) < \infty$ and $N_t(A) = \sum_{s \le t} 1_{(\Delta X_s \in A)}$, that $N_t(A)$ is a Poisson process with intensity or parameter $m(A)$. A key step in the construction was the centering of the Poisson processes to get Lévy processes with characteristic functions given by (10.3). Without the centering one is forced to work only with characteristic functions given by (10.2).

We now work towards showing that every Lévy process is of the form given by (10.1). We start with the following lemma.

Lemma 10.2. *If X_t is a Lévy process and T is a bounded stopping time, then $X_{T+t} - X_T$ is a Lévy process with the same law as $X_t - X_0$ and independent of $\mathcal{F}_T$.*

Proof. Suppose first that T takes only finitely many values $t_1, \ldots, t_n$. If f is a bounded continuous function,

$$\begin{aligned} \mathbb{E}\, f(X_{T+t} - X_T) &= \sum_{i=1}^n \mathbb{E}\,[f(X_{t_i+t} - X_{t_i}); T = t_i] \\ &= \sum \mathbb{E}\,[f(X_{t_i+t} - X_{t_i})] \mathbb{P}(T = t_i) \\ &= \mathbb{E}\,[f(X_t - X_0)] \sum \mathbb{P}(T = t_i) \\ &= \mathbb{E}\,[f(X_t - X_0)]. \end{aligned}$$

We used the independent increments property in the second equality and the stationary increments property in the third equality.

If T does not necessarily have only finitely many values, let $T_n = \inf\{k/2^n : k/2^n \le T < (k+1)/2^n\}$, apply the above to T_n, and let $n \to \infty$.

Since f is an arbitrary bounded continuous function, this shows the law of $X_{T+t} - X_T$ is the same as the law of $X_t - X_0$. A similar argument with $\prod_{i=1}^n f_i(X_{T+t_{i+1}} - X_{T+t_i})$ shows that as processes the law of the process $X_{T+t} - X_T$ is the same as the law of $X_t - X - 0$.

If we repeat the procedure above starting with $\mathbb{E}\,[f(X_{T+t}-X_T);A]$, where $A\in\mathcal{F}_T$, we see that $\mathbb{E}\,[f(X_{T+t}-X_T);A]=\mathbb{E}\,[f(X_{T+t}-X_T)]\mathbb{P}(A)$, which implies independence of $X_{T+t}-X_T$ from $\mathcal{F}_T$. □

Lemma 10.3. *If X_t is a Lévy process with bounded jumps, then X_t has moments of all orders.*

Proof. Suppose the jumps of X_t are bounded in absolute value by K. Since X_t is rcll, there exists $M>K$ such that $\mathbb{P}(\sup_{s\le t}|X_s|>2M)\le(1/2)$.

Let $T_1=\inf\{t:|X_t|\ge M\}$ and $T_{i+1}=\inf\{t>T_i:|X_t-X_{T_i}|>M\}$. Observe $\sup_{s\le T_1}|X_s|\le K+M\le 2M$. We have

$$\begin{aligned}\mathbb{P}(\sup_{s\le t}|X_s|>2(i+1)M)&\le\mathbb{P}(T_{i+1}\le t)\le\mathbb{P}(T_i\le t,T_{i+1}-T_i\le t)\\&=\mathbb{P}(\sup_{s\le t}|X_{T_i+s}-X_{T_i}|>2M,T_i\le t)\\&=\mathbb{P}(\sup_{s\le t}|X_s|>2M)\mathbb{P}(T_i\le t)\\&\le(1/2)\mathbb{P}(T_i\le t).\end{aligned}$$

By induction, $\mathbb{P}(\sup_{s\le t}|X_s|>2iM)\le 2^{-i}$, and the lemma now follows immediately. □

Theorem 10.4. *Suppose X_t is a Lévy process. Then there exists a measure m on $\mathbb{R}-\{0\}$ with $\int 1\wedge x^2\,m(dx)>\infty$ such that the characteristic function of X_t is given by (10.1).*

Proof. Let f be a strictly increasing function on $\mathbb{R}$ such that $f(x)=x$ if $|x|\le 1$ and f is bounded by 2. Let

$$X'_t=X_t+\sum_{s\le t}[f(\Delta X_s)-\Delta X_s]1_{(|\Delta X_s|>1)}.$$

Since X_t is rcll, then there are only finitely many jumps with size larger than 1 in any finite time interval, so the sum is finite. X'_t is again a Lévy process, this time with bounded jumps; if we show the desired representation for X'_t, it is not hard to see that this gives the desired representation for X_t. Now let $X''_t=X'_t-\mathbb{E}\,X'_t$. Again, it suffices to look at X''. So without loss of generality we may suppose X_t has jumps bounded by 2 and that it has mean 0.

Let $\{I_i\}$ be an ordering of the intervals $(1,2],[-2,-1),(1/2,1],[-1,-1/2),\ldots$ Let

$$\overline{X}^i_t=\sum_{s\le t}\Delta X_s 1_{(\Delta X_s\in I_i)},\qquad X^i_t=\overline{X}^i_t-\mathbb{E}\,\overline{X}^i_t.$$

We want to show that the X^i are all independent and also that X^i is independent of $X_t-\sum_{j=1}^{i-1}X^j_t$. We will prove that the random variable X^i_t is independent of X^j_t. To show that they are independent as processes and to do the other cases are similar. Let φ_i be the characteristic function of X_1^i. Then $M_t^i = \exp(iuX_t^i - t\varphi_i(u))$ is a martingale, since

$$\mathbb{E}\,[M_t^i \mid \mathcal{F}_s] = M_s^i \mathbb{E}\,[\exp(iu(X_t - X_s)) - (t-s)\varphi_i(u)] = M_s^i,$$

using the independence and stationarity of the increments. $\exp(-t\varphi_i(u)$ is a process of bounded variation, so has no martingale part. By Ito's formula, we see $\exp(iuX_t^i)$ is a semimartingale, and the continuous part of the martingale part is 0. So M_t^i has no continuous part. We let $M_t^j = \exp(ivX_t^j - t\varphi_j(v))$ and see that M_t^j is also a martingale. By Lemma 2.2, M_t^i is orthogonal to M_t^j, since M^j does not have jumps in common with M^i. Since $M_0^i = M_0^j = 1$, then $\mathbb{E}\,[M_t^i M_t^j] = 1$, or

$$\mathbb{E}\,[\exp(iuX_t^i)\exp(ivX_t^j)] = \exp(t\varphi_i(u))\exp(t\varphi_j(v)) = \mathbb{E}\,[\exp(iuX_t^i)]\mathbb{E}\,[\exp(ivX_t^j)].$$

This proves the independence.

Since X_t has bounded jumps, then by Lemma 10.2 X_t has second moments. By the independence and the fact that all the X^i have mean 0,

$$\sum_{i=1}^{\infty} \mathbb{E}\,(X_t^i)^2 \le \mathbb{E}\left[(X_t - \sum_{i=1}^{\infty} X_t^i)^2\right] + \mathbb{E}\left[(\sum_{i=1}^{\infty} X_t^i)^2\right] = \mathbb{E}\,(X_t)^2 < \infty.$$

Hence

$$\mathbb{E}\,[\sum_{i=m}^{n} X_t^i]^2 = \sum_{i=m}^{n} \mathbb{E}\,(X_t^i)^2$$

tends to 0 as $n, m \to \infty$, and thus $X_t - \sum_{i=1}^{N} X_t^i$ converges in L^2. The limit, X_t^c, say, will be a Lévy process independent of all the X_t^i. Moreover, X^c has no jumps, i.e., it is continuous. By the stationary and independent increments properties and the Lindeberg-Feller theorem, we conclude that X_t^c is Gaussian. Therefore X_t^c is a Brownian motion.

To complete the proof, it remains to show that X_t^i has a characteristic function of the form (10.3) for m supported on I_i. The fact that $\mathbb{E}\,(X_t)^2 < \infty$ will imply $\int_{[-2,2]} x^2 m(dx) < \infty$. Fix i. Let $D_n = \{k/2^n : k/2^n \in I_i\}$. Thus the D_n are finite subsets of I_i increasing to a countable dense subset of I_i. Let $\overline{N}_t^{in}$ be the process that jumps $k/2^n$ if X_t^i has a jump in the interval $[k/2^n, (k+1)/2^n)$. Let $N_t^{in} = \overline{N}_t^{in} - \mathbb{E}\,\overline{N}_t^{in}$. As $n \to \infty$, $N_t^{in} \to X_t^i$, so it suffices to show that N_t^{in} has a representation of the form (10.3) with m supported in D_n.

By the fact that we are dealing with Lévy processes, the process $\sum_{s \le t} 1_{(\Delta \overline{N}_t^{in} = k/2^n)}$ is a Poisson process. Let the intensity be λ_{kn}. If we set $m_{in}(dx) = \sum_{k=0}^{\infty} \lambda_{kn}\delta_{k/2^n}(dx)$, then we see that $\overline{N}_t^{in}$ has a representation of the form (10.2). It follows that N_t^{in} has a characteristic function given by A(10.3) for this m_{in}. The proof is complete. $\square$