12. Semimartingales and their quadratic variation

In the previous section we proved the Itô formula (11.22) for the stochastic processes of the form $\{f(B_t): t \geqslant 0\}$. The formula takes the following differential form

$$\mathrm{d}f(B_t) = f'(B_t)\mathrm{d}B_t + \frac{1}{2}f''(B_t)\mathrm{d}t \tag{12.1}$$

As it turns out, in order to simplify manipulations with stochastic processes, it is convenient to single out the class of processes that are generally expressed in this way.

Definition 12.1 (Generalized diffusion) *A stochastic process $\{X_t: t \geqslant 0\}$ is a* generalized diffusion*, a.k.a. a* (continuous) semimartingale*, if the underlying probability carries a standard Brownian motion $\{B_t: t \geqslant 0\}$ and a Brownian filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ and two stochastic processes $\{U_t: t \geqslant 0\}$ and $\{Y_t: t \geqslant 0\}$ such that*

(1) *both U and Y are jointly measurable and adapted to $\{\mathcal{F}_t\}_{t\geqslant 0}$,*
(2) *for all $t \geqslant 0$,*

$$\int_0^t |U_s|\mathrm{d}s < \infty \ \wedge \ \int_0^t Y_s^2 \mathrm{d}s < \infty \quad \text{a.s.} \tag{12.2}$$

(3) *for all $t \geqslant 0$,*

$$X_t = X_0 + \int_0^t U_s \mathrm{d}s + \int_0^t Y_s \, \mathrm{d}B_s \quad \text{a.s.} \tag{12.3}$$

In order to explain the terms mentioned in the above definition, we note the following standard concepts:

Definition 12.2 (Local martingale) *A* local martingale *is a process $\{M_t: t \geqslant 0\}$ for which there is a sequence $\{T_n: n \geqslant 1\}$ of non-negative random variables such that*

(1) *$\forall n \geqslant 1$: T_n is a stopping time such that $\{M_{T_n \wedge t}: t \geqslant 0\}$ is a martingale, and*
(2) *$T_n \to \infty$ a.s. as $n \to \infty$.*

Definition 12.3 (Semimartingale) *A stochastic process $\{X_t: t \geqslant 0\}$ is a* semimartingale *if there exists a local martingale $\{M_t: t \geqslant 0\}$ such that $X_t - M_t$ is of bounded variation on any compact interval a.s. (i.e., $V_t^{(1)}(X - M) < \infty$ a.s. for each $t \geqslant 0$).*

In short, a local martingale is a martingale after localization with the stochastic integral of a locally integrable process playing the role of a salient example. A semimartingale is a sum of a local martingale and an adapted process of bounded variation, just as stipulated in Definition 12.1 above.

We say that a semimartingale is continuous if it has continuous sample paths. An example of a semimartingale that does not have continuous paths is a homogeneous Poisson process $\{N_t: t \geqslant 0\}$. Here the decomposition is $N_t = (N_t - t) + t$ where $N_t - t$ is a martingale, albeit with discontinuous sample paths.

The term "generalized diffusion" arises from a term in the physics literature. There the notion of a diffusion process is reserved to:

Definition 12.4 *We say that a process* $\{X_t\colon t\geqslant 0\}$ *is a* diffusion *(a.k.a. Itô diffusion) if there exist functions* $f,g\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ *such that Definition 12.1 applies with*

$$U_t := f(t,X_t) \ \wedge\ Y_t := g(t,X_t) \tag{12.4}$$

The conditions ensure that the stochastic dynamics governing the process X has no memory and is local. This has been one of the governing principles of physics since the time of Newton's laws.

We remark that, similarly as in (12.1), the identity (12.3) will often be shortened to a differential form as in

$$\mathrm{d}X_t = U_t\mathrm{d}t + Y_t\mathrm{d}B_t \tag{12.5}$$

Here $U_t\mathrm{d}t$ is called the *drift* term while $Y_t\mathrm{d}B_t$ is referred to as the *diffusive* term. Note that writing (12.5) tacitly assumes that all the remaining conditions in Definition 12.1 hold as well. Indeed, these conditions are needed to ensure that the integrals are well defined and yield an adapted jointly measurable process.

As it turns out, the Itô formula extends to semimartingales, but in order to formulate it properly we need an interpretation of the term $(\mathrm{d}X_t)^2$ that will inevitably arise in the Taylor expansion. As for the Brownian motion, this will be interpreted in terms of the quadratic variation. So we first prove:

Theorem 12.5 *Let X be a semimartingale of the form $\mathrm{d}X_t = U_t\mathrm{d}t + Y_t\mathrm{d}B_t$. Then for any $t\geqslant 0$ and any sequence of partitions $\{\Pi_n\}_{n\geqslant 1}$,*

$$\|\Pi_n\|\to 0 \quad\Rightarrow\quad V_t^{(2)}(X,\Pi_n) \ \xrightarrow[n\to\infty]{P}\ \int_0^t Y_s^2\mathrm{d}s \tag{12.6}$$

Proof. Let Π be a partition of $[0,t]$. Throughout we will repeatedly invoke the following identity

$$\Bigl|V_t^{(2)}(X,\Pi)^{1/2} - V_t^{(2)}(\widetilde{X},\Pi)^{1/2}\Bigr| \leqslant V_t^{(2)}(X-\widetilde{X},\Pi)^{1/2} \tag{12.7}$$

which is simply the Minkowski inequality for the Euclidean norm. The fact that

$$V_t^{(2)}\Bigl(\int_0^{\cdot} U_s\mathrm{d}s,\Pi\Bigr) \leqslant V_t^{(1)}\Bigl(\int_0^{\cdot} U_s\mathrm{d}s,\Pi\Bigr)\,\mathrm{osc}_{\int_0^{\cdot} U_s\mathrm{d}s}\bigl([0,t],\|\Pi\|\bigr) \tag{12.8}$$

along with the bound

$$V_t^{(1)}\Bigl(\int_0^{\cdot} U_s\mathrm{d}s,\Pi\Bigr) \leqslant \sup_{\Pi'} V_t^{(1)}\Bigl(\int_0^{\cdot} U_s\mathrm{d}s,\Pi'\Bigr) = \int_0^t |U_s|\mathrm{d}s \tag{12.9}$$

where the integral is finite almost surely by Definition 12.1(2), then show

$$\|\Pi_n\|\to 0 \quad\Rightarrow\quad V_t^{(2)}\Bigl(\int_0^{\cdot} U_s\mathrm{d}s,\Pi_n\Bigr) \ \xrightarrow[n\to\infty]{\text{a.s.}}\ 0 \tag{12.10}$$

Using this in (12.7) we conclude that the drift term has no contribution to the limit and it suffices to prove the claim for X such that $\mathrm{d}X_t = Y_t\mathrm{d}B_t$. We will do this by going from Y simple to $Y\in\mathcal{V}$ and $Y\in\mathcal{V}^{\text{loc}}$.

STEP 1 (Y simple): Let us first assume first that $X_t = \int_0^t Y_s\,\mathrm{d}B_s$ for $Y\in\mathcal{V}_0$ with $Y_0=0$ and, without loss of generality, $Y_u = 0$ for $u\geqslant t$. This permits us to write Y in the form

$Y = \sum_{i=1}^k Z_i 1_{(s_{i-1},s_i]}$ for some $0 = s_0 < s_1 < \cdots < s_k \leqslant t$. Let $t_0 = 0 < t_1 < \cdots < t_m = t$ denote the points defining the partition Π. Assuming henceforth that $\|\Pi\| < \min_{i\leqslant n} |s_i - s_{i-1}|$, each interval $[t_{i-1}, t_i]$ of the partition contains at most one "jump" of Y. Let $\mathcal{I} \subseteq \{1,\dots,m\}$ be the indices denoting the intervals containing such a jump; i.e., those i for which $\{s_0,\dots,s_k\} \cap [t_{i-1}, t_i] \neq \varnothing$. For each $i \in \{1,\dots,m\}$, let

$$j(i) := \max\{j \geqslant 0 \colon s_j \leqslant t_i\} \tag{12.11}$$

We then have

$$\begin{aligned} &\forall i \notin \mathcal{I}\colon \quad s_{j(i)} < t_{i-1} \ \wedge\ s_{j(i)+1} > t_i \\ &\forall i \in \mathcal{I}\colon \quad t_{i-1} \leqslant s_{j(i)} \leqslant t_j \end{aligned} \tag{12.12}$$

This gives

$$\begin{aligned} V_t^{(2)}(X,\Pi) = \sum_{i\notin\mathcal{I}} & Z_{j(i)+1}^2 (B_{t_i} - B_{t_{i-1}})^2 \\ &+ \sum_{i\in\mathcal{I}} \Big(Z_{j(i)}(B_{s_{j(i)}} - B_{t_{i-1}}) + Z_{j(i)+1}(B_{t_i} - B_{s_{j(i)}}) \Big)^2 \end{aligned} \tag{12.13}$$

Since $\max_{i\leqslant n} |Z_i| \leqslant C$ a.s. by the assumption that Y is simple, the expectation of the second term is at most

$$\begin{aligned} 2C^2 \sum_{i\in\mathcal{I}} \Big(E\big((B_{s_{j(i)}} - B_{t_{i-1}})^2\big) + E\big((B_{t_i} - B_{s_{j(i)}})^2\big)\Big) \\ = 2C^2 \sum_{i\in\mathcal{I}} (t_i - t_{i-1}) \leqslant 4C^2(k+1)\|\Pi\| \end{aligned} \tag{12.14}$$

where we first used that $(a+b)^2 \leqslant 2a^2 + 2b^2$ and the noted that $|\mathcal{I}| \leqslant k+1$.

As to the first term on the right of (12.13), we write this as

$$\sum_{i\notin\mathcal{I}} Z_{j(i)+1}^2 (t_i - t_{i-1}) + \sum_{i\notin\mathcal{I}} Z_{j(i)+1}^2 \Big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \Big) \tag{12.15}$$

For the first term in this sum we then get

$$\Big| \sum_{i\notin\mathcal{I}} Z_{j(i)+1}^2 (t_i - t_{i-1}) - \int_0^t Y_s^2 \mathrm{d}s \Big| \leqslant C^2 \sum_{i\in\mathcal{I}} (t_i - t_{i-1}) \leqslant C^2(k+1)\|\Pi\| \tag{12.16}$$

proving that this term tends to the desired integral as $\|\Pi\| \to 0$. Since $Z_{j(i)+1}$ is $\mathcal{F}_{s_{j(i)}}$-measurable and so, by the first line in (12.12), also $\mathcal{F}_{t_{i-1}}$-measurable, the expectation of the second term in (12.15) vanishes thanks to the fact that $(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})$ is a martingale increment. By the same reasoning (and the assumed a.s. bound on the Z_i's) the second moment is at most

$$C^4 \sum_{i\notin\mathcal{I}} E\bigg(\Big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \Big)^2 \bigg) \leqslant C^4 t \operatorname{Var}(\mathcal{N}(0,1)^2)\|\Pi\| \tag{12.17}$$

We conclude that the second term in (12.15) tends to zero in probability as $\|\Pi\| \to 0$ which proves (12.10) for all Y simple.

STEP 2 (Y square integrable): Next assume $Y \in \mathcal{V}$. Then there exists $\{Y^{(n)}\}_{n\geqslant 1} \subseteq \mathcal{V}_0$ such that $Y^{(n)} \to Y$ in $\mathcal{V}$. Denote

$$X_t^{(n)} := \int_0^t Y_s^{(n)}\,\mathrm{d}B_s \tag{12.18}$$

Then (12.7) yields

$$\begin{aligned} E\biggl(\Bigl|\,V_t^{(2)}(X,\Pi)^{1/2} - V_t^{(2)}(X^{(n)},\Pi)^{1/2}\Bigr|^2\biggr) \\ &\leqslant EV_t^{(2)}(X - X^{(n)},\Pi) \\ &= E\bigl(|X_t - X_t^{(n)}|^2\bigr) = E\Bigl(\int_0^t (Y_s - Y_s^{(n)})^2 \mathrm{d}x\Bigr) \end{aligned} \tag{12.19}$$

were we noted that, since $X - X^{(n)}$ is an L^2-martingale, its increments over non-overlapping intervals of the partition are uncorrelated and then invoked the Itô isometry at the end. Similarly we get

$$E\left(\left|\Bigl(\int_0^t Y_s^2 \mathrm{d}s\Bigr)^{1/2} - \Bigl(\int_0^t (Y_s^{(n)})^2 \mathrm{d}s\Bigr)^{1/2}\right|^2\right) \leqslant E\Bigl(\int_0^t (Y_s - Y_s^{(n)})^2 \mathrm{d}x\Bigr) \tag{12.20}$$

Using that $Y^{(n)} \to Y$ in $\mathcal{V}$, the left-hand sides of (12.19–12.20) tend to zero as $n \to \infty$. To get the claim from this note that, for any non-negative random variables $\{Z_n\}_{n\geqslant 1}$ and Z we have

$$|Z_n - Z| \leqslant 2Z^{1/2}\,|Z_n^{1/2} - Z^{1/2}| + (Z_n^{1/2} - Z^{1/2})^2 \tag{12.21}$$

Assuming these random variables are in L^2, the Cauchy-Schwarz inequality shows that $Z_n^{1/2} \to Z^{1/2}$ in L^2 (resp., in probability) implies $Z_n \to Z$ in L^1 (resp., in probability).

STEP 3 (Y locally square integrable): Finally, let us address the case $Y \in \mathcal{V}^{\mathrm{loc}}$ which requires a (by now) routine localization step. Let $\{\Pi_n\}_{n\geqslant 1}$ of partitions of $[0,t]$ with $\|\Pi_n\| \to 0$ and let $T^{(M)}$ be the quantity in (10.2). Then

$$\widetilde{X}_t := X_{T^{(M)}\wedge t} = \int_0^t Y_s 1_{\{T^{(M)}>s\}} \mathrm{d}B_s \tag{12.22}$$

Since $\{Y_s 1_{\{T^{(M)}>s\}} \colon s \geqslant 0\} \in \mathcal{V}$, on $\{T^{(M)} > t\}$ we get

$$V_t^{(2)}(X,\Pi_n) = V_t^{(2)}(\widetilde{X},\Pi_n) \underset{n\to\infty}{\overset{P}{\longrightarrow}} \int_0^t Y_s^2 1_{\{T^{(M)}>s\}} \mathrm{d}s \tag{12.23}$$

by STEP 2 above. Since $Y_s^2 1_{\{T^{(M)}>s\}} = Y_s^2$ for all $s \in [0,t]$ on $\{T^{(M)} > t\}$, Lemma 11.5 shows that the integral on the right equals $\int_0^t Y_s\,\mathrm{d}B_s$ a.s. on $\{T^{(M)} > t\}$. Since $\bigcup_{M\geqslant 1}\{T^{(M)} > t\}$ is a full measure set, the claim is proved. □

We now put forward:

Definition 12.6 *Given a semimartingale X of the form $\mathrm{d}X_t = U_t\mathrm{d}t + Y_t\mathrm{d}B_t$, we call*

$$\langle X\rangle_t := \int_0^t Y_s^2\,\mathrm{d}s \tag{12.24}$$

the quadratic variation process *associated with* X.

With this we observe:

Lemma 12.7 (Doob-Meyer decomposition for stochastic integral squared) *Let X be a semimartingale of the form $\mathrm{d}X_t = Y_t \mathrm{d}B_t$. Then $\{X_t^2 - \langle X\rangle_t \colon t \geqslant 0\}$ is local martingale.*

Proof. Suppose first $Y \in \mathcal{V}$. Denoting by $\{\mathcal{F}_t\}_{t\geqslant 0}$ the underlying Brownian filtration, for all $t \geqslant u \geqslant 0$ the conditional Itô isometry from Lemma 11.3 gives

$$\begin{aligned} E\big(X_t^2 - X_u^2 \,\big|\, \mathcal{F}_u\big) &= E\big((X_t - X_u)^2 \,\big|\, \mathcal{F}_u\big) = E\left(\left(\int_u^t Y_s \,\mathrm{d}B_s\right)^2 \,\middle|\, \mathcal{F}_u\right) \\ &= E\left(\int_u^t Y_s^2 \,\mathrm{d}s \,\middle|\, \mathcal{F}_u\right) = E\big(\langle X\rangle_t - \langle X\rangle_u \,\big|\mathcal{F}_u\big) \end{aligned} \tag{12.25}$$

where we used that $E(X_t - X_u|\mathcal{F}_u) = 0$ in the first step. It follows that $\{X_t^2 - \langle X\rangle_t \colon t \geqslant 0\}$ is a martingale.

For $Y \in \mathcal{V}^{\mathrm{loc}}$ we use that $\widetilde{X}_t := X_{T^{(M)}\wedge t}$ is of the form treated above and so the process $\{\widetilde{X}_t^2 - \langle \widetilde{X}\rangle_t \colon t \geqslant 0\}$ is a martingale. By (12.22) and the fact that the integrand there is in $\mathcal{V}$, the quadratic variation process of $\widetilde{X}$ takes the form

$$\langle \widetilde{X}\rangle_t = \int_0^{T^{(M)}\wedge t} Y_s^2 \mathrm{d}s = \langle X\rangle_{T^{(M)}\wedge t} \tag{12.26}$$

Since $T^{(M)}$ is a stopping with $T^{(M)} \to \infty$ a.s. as $M \to \infty$, we get that $\{X_t^2 - \langle X\rangle_t \colon t \geqslant 0\}$ is a local martingale according to Definition 12.2. $\square$

The conclusion of Lemma 12.7 is an instance of *Doob-Meyer decomposition* of a continuous local submartingale $\{Z_t \colon t \geqslant 0\}$, of which $Z_t := X_t^2$ is an example, as the sum $Z_t = M_t + A_t$ of a local martingale M and a non-decreasing continuous process A. A proof of this is somewhat complicated and can be found in Karatzas and Shreve's book.

An important part of Doob-Meyer decomposition statement is a uniqueness clause: the non-decreasing process (and thus also the martingale) is determined uniquely by its value at time zero. This clause is extracted from the following lemma whose proof we leave to homework.

Lemma 12.8 *Let $\{M_t \colon t \geqslant 0\}$ be a continuous local martingale such that, for some $t \geqslant 0$,*

$$V_t^{(1)}(M) < \infty \quad \text{a.s.} \tag{12.27}$$

Then

$$P\big(\,\forall s \in [0,t] \colon M_t = M_0\big) = 1 \tag{12.28}$$

Indeed, if the submartinagle admits two different decompositions, $X_t = M_t + A_t = \widetilde{M}_t + \widetilde{A}_t$, then the local marginagle $M - \widetilde{M}$ obeys $M_t - \widetilde{M}_t = \widetilde{A}_t - A_t$ for each $t \geqslant 0$. The process on the right is a difference of two non-decreasing functions and so it is of finite variation on any interval. The above lemma then tells us that $M - \widetilde{M}$ is constant a.s. and, if $A_0 = \widetilde{A}_0 = 0$, it vanishes everywhere a.s.

We leave the proof of Lemma 12.8 to homework. As a consequence of the ensuing uniqueness argument, we also get:

Corollary 12.9 *Suppose that* $\{U_t\colon t \geqslant 0\}$, $\{\widetilde{U}_t\colon t \geqslant 0\}$, $\{Y_t\colon t \geqslant 0\}$ *and* $\{\widetilde{Y}_t\colon t \geqslant 0\}$ *obey conditions (1-2) of Definition 12.1 and, for some* $t \geqslant 0$,

$$\forall u \in [0,t]\colon \quad \int_0^u U_s \mathrm{d}s + \int_0^u Y_s \mathrm{d}B_s = \int_0^u \widetilde{U}_s \mathrm{d}s + \int_0^u \widetilde{Y}_t \mathrm{d}B_t \tag{12.29}$$

holds a.s. Then

$$\lambda\Big(s \in [0,t]\colon Y_s \neq \widetilde{Y}_s \ \vee \ U_s \neq \widetilde{U}_s\Big) = 0 \quad \text{a.s.} \tag{12.30}$$

In particular, the representation (12.3) *of a generalized diffusion* X *is unique up to equivalence.*

Proof. Since $u \mapsto \int_0^u (U_s - \widetilde{U}_s)\mathrm{d}s$ is of bounded variation, equality (12.29) forces the continuous local martingale $u \mapsto \int_0^u (Y_s - \widetilde{Y}_s)\mathrm{d}B_s$ to vanish for all $u \in [0,t]$ a.s. Lemma 11.5 then gives that $Y_s \neq \widetilde{Y}_s$ on a Lebesgue null set a.s. But then $\int_0^u (U_s - \widetilde{U}_s)\mathrm{d}s = 0$ for all $u \in [0,t]$ a.s. which by the Lebesgue differentiation theorem shows that also $U_s \neq \widetilde{U}_s$ on Lebesgue null set a.s. □

As a final remark we note that uniqueness of Doob-Meyer decomposition may fail without continuity. Indeed, if $\{B_t\colon t \geqslant 0\}$ is a standard Brownian motion and $\{N_t\colon t \geqslant 0\}$ an independent rate-1 Poisson process, then (assuming the filtration is such that N is Markov) both $\{B_t^2 - t\colon t \geqslant 0\}$ and $\{B_t^2 - N_t\colon t \geqslant 0\}$ are martingales.

Further reading: Section 1.5 of Karatzas-Shreve