6. Quadratic variation

This section presents the first set of calculations that naturally lead to the discovery of the stochastic integral. A key object of interest here is the quadratic variation.

We have seen that Brownian paths are roughly $1/2 + o(1)$-Hölder regular. This refers to local behavior but applications in ODE (using the Brownian path as a driving "noise") and elsewhere need a way to characterize the regularity of the path as a whole. For this we recall the following concepts:

Definition 6.1 *A* partition Π *of interval* $[0,t]$ *is an ordered collection of points* $0 =: t_0 < t_1 < \cdots < t_n := t$. *The* mesh $\|\Pi\|$ *of the partition* Π *is then given by*

$$\|\Pi\| := \max_{i=1,\dots,n} |t_i - t_{i-1}| \tag{6.1}$$

Given $f\colon [0,\infty) \to \mathbb{R}$ *and* $p \geqslant 0$, *the* p*-variation of* f *associated with partition* Π *of interval* $[0,t]$ *is the quantity*

$$V_t^{(p)}(f,\Pi) := \sum_{i=1}^n \bigl|f(t_i) - f(t_{i-1})\bigr|^p \tag{6.2}$$

Given a function $f\colon [0,t] \to \mathbb{R}$, we are generally interested in the behavior of the p-variation $V_t^{(p)}(f,\Pi)$ as the mesh of Π tends to zero. For $p = 1$ (in fact $p \leqslant 1$), this is aided by the fact that the p-variation increases upon refinements of Π. It thus makes sense to take a supremum over Π which we then refer to (for $p = 1$) as *total variation* which is a concept (introduced by C. Jordan in 1881) closely related to rectifiability of curves in analytic geometry. While the monotonicity under refinements is no-longer true for $p > 1$, the supremum can also be considered there. If the result is finite, we we say that f is of *bounded p-variation*. (For $p = 1$, this is simply referred to as being of bounded variation.) We call $V^{(2)}(f,\Pi)$ *quadratic variation*.

As is readily checked, uniform Hölder continuity with exponent $\alpha \in (0,1]$ implies boundedness of p-variation for all p with $p\alpha \geqslant 1$. This suggests that $p = 2 + o(1)$ is the relevant case for standard Brownian motion. And, indeed, given a Brownian path $\{B_t\colon t \in [0,\infty)\}$, we have

$$\begin{aligned} EV_t^{(p)}(B,\Pi) &= \sum_{i=1}^n E\bigl(|B_{t_i} - B_{t_{i-1}}|^p\bigr) \\ &= E\bigl(|\mathcal{N}(0,1)|^p\bigr) \sum_{i=1}^n |t_i - t_{i-1}|^{p/2} \leqslant E\bigl(|\mathcal{N}(0,1)|^p\bigr) t \|\Pi\|^{\frac{p-2}{2}} \end{aligned} \tag{6.3}$$

which implies $V_t^{(p)}(B,\Pi_n) \to 0$ in probability and L^1 as $\|\Pi_n\| \to 0$ for all $p > 2$. For $p = 2$ we have

$$EV_t^{(2)}(B,\Pi) = t \tag{6.4}$$

and so

$$\Bigl\{V_t^{(2)}(B,\Pi)\colon \Pi = \text{ partition of } [0,t]\Bigr\} \text{ is a tight family} \tag{6.5}$$

A natural question is whether $V_t^{(2)}(B,\Pi)$ converges weakly along sequences of partitions whose mesh tends to zero. This is addressed in:

Proposition 6.2 *Let $\{B_t\colon t\in[0,\infty)\}$ be a standard Brownian motion. For any sequence of partitions $\{\Pi_n\}_{n\geqslant 1}$ of $[0,t]$, we have:*

(1) $\|\Pi_n\|\to 0$ *implies* $V_t^{(2)}(B,\Pi_n)\to t$ *in probability and* L^2,
(2) $\sum_{n\geqslant 1}\|\Pi_n\|<\infty$ *implies* $V_t^{(2)}(B,\Pi_n)\to t$ *a.s.*
(3) $\{\Pi_n\}_{n\geqslant 1}$ *are nested and* $\|\Pi_n\|\to 0$ *imply* $V_t^{(2)}(B,\Pi_n)\to t$ *a.s.*

Here $\{\Pi_n\}_{n\geqslant 1}$ are said to be nested if Π_{n+1} contains all points of Π_n, for each $n\geqslant 1$.

Proof. Using that the Brownian increments are independent Gaussians, we observe

$$\begin{aligned}\operatorname{Var}\bigl(V_t^{(2)}(B,\Pi)\bigr) &= \sum_{i=1}^n \operatorname{Var}\bigl((B_{t_i}-B_{t_{i-1}})^2\bigr)\\ &= \operatorname{Var}\bigl(\mathcal{N}(0,1)^2\bigr)\sum_{i=1}^n |t_i-t_{i-1}|^2 \leqslant 3t\|\Pi\|\end{aligned}\tag{6.6}$$

In conjunction with (6.4) this gives L^2-convergence $V_t^{(2)}(B,\Pi)\to t$. Invoking the Markov inequality

$$P\bigl(|V_t^{(2)}(B,\Pi)-t|>\epsilon\bigr)\leqslant \frac{1}{\epsilon^2}\operatorname{Var}\bigl(V_t^{(2)}(B,\Pi)\bigr)\leqslant \frac{3t}{\epsilon^2}\|\Pi\|\tag{6.7}$$

we also get convergence in probability as $\|\Pi\|\to 0$, thus proving (1). For (2) we note that the summability of the mesh-sequence implies summability of the probabilities on the left. Then a.s. convergence $V_t^{(2)}(B,\Pi)\to t$ then follows via a standard argument based on the Borel-Cantelli lemma.

The convergence in (3) will be inferred from martingale convergence. We need a lemma whose statement is more scary that the proof:

Lemma 6.3 *Let $Z_1,\dots,Z_n\in L^2$ be independent with symmetric law, $\forall i\leqslant n\colon Z_i \overset{\text{law}}{=} -Z_i$, and let $I_1,\dots,I_m$ be non-empty disjoint sets such that $\{1,\dots,n\}=\bigcup_{j=1}^m I_j$. Then for any σ-algebra $\mathcal{G}$ satisfying*

$$\sigma\biggl(\Bigl(\sum_{i\in I_j}Z_i\Bigr)^2\colon j=1,\dots,m\biggr)\subseteq\mathcal{G}\subseteq\sigma\biggl(\Bigl(\sum_{i\in I}Z_i\Bigr)^2\colon\ I\subseteq I_j \text{ for some } j\in\{1,\dots,m\}\biggr)\tag{6.8}$$

we have

$$E\biggl(\Bigl(\sum_{i=1}^n Z_i\Bigr)^2\,\bigg|\,\mathcal{G}\biggr)=\sum_{j=1}^m\Bigl(\sum_{i\in I_j}Z_i\Bigr)^2\tag{6.9}$$

Proof. Left to homework. □

We now pick a sequence of nested partitions $\{\Pi_n\}_{n\geqslant 1}$ where the points of Π_n will be denoted as $\{t_i^n\}_{i=0}^{k(n)}$. For $n \leqslant N$ set

$$\mathcal{G}_{n,N} := \sigma\Big((B_{t_i^r} - B_{t_{i-1}^r})^2 \colon i = 1, \dots, k(r),\, r = n, \dots, N\Big) \tag{6.10}$$

Lemma 6.3 then implies

$$V_t^{(2)}(f, \Pi_n) = E(B_t^2 \,|\, \mathcal{G}_{n,N}) \tag{6.11}$$

Taking $N \to \infty$ with the help of the Levy Forward Theorem gives

$$V_t^{(2)}(f, \Pi_n) = E(B_t^2 \,|\, \mathcal{G}_n) \tag{6.12}$$

where

$$\mathcal{G}_n := \sigma\Big(\bigcup_{N\geqslant n} \mathcal{G}_{n,N}\Big) = \sigma\Big((B_{t_i^r} - B_{t_{i-1}^r})^2 \colon i = 1, \dots, k(r),\, r \geqslant n\Big) \tag{6.13}$$

Lévy Backward Theorem then gives a.s. convergence of $V_t(f, \Pi_n)$ as $n \to \infty$. (Note that none of these require that the mesh of the partition tends to zero.) In light of part (1) of the claim, if $\|\Pi_n\| \to 0$ the a.s. limit must equal t, as desired. □

A natural follow-up question is what happens with the corresponding quadratic variation of functions of the Brownian path. We start by looking at $V^{(2)}(B^2, \Pi)$. A calculation shows

$$\begin{aligned} V_t^{(2)}(B^2, \Pi) &= \sum_{i=1}^n (B_{t_i}^2 - B_{t_{i-1}}^2)^2 \\ &= \sum_{i=1}^n 4B_{t_{i-1}}^2 (B_{t_i} - B_{t_{i-1}})^2 + \sum_{i=1}^n 4B_{t_{i-1}} (B_{t_i} - B_{t_{i-1}})^3 + \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^4 \end{aligned} \tag{6.14}$$

The reason for separating terms this way is that, by the defining properties of standard Brownian motion, $B_{t_{i-1}}$ is independent of $B_{t_i} - B_{t_{i-1}}$, and in fact of $B_{t_k} - B_{t_{k-1}}$ for all $k \geqslant i$. This allows us to estimate the L^1-norm of the second term on the right by

$$E|B_t| \sum_{i=1}^n E\big(|B_{t_i} - B_{t_{i-1}}|^3\big) \leqslant E|B_t|\, t\, E\big(|\mathcal{N}(0,1)|^3\big)\, \|\Pi\|^{1/2} \tag{6.15}$$

The L^1-norm of the last term on the right of (6.14) is in turn bounded by a constant times $\|\Pi\|$. These terms thus tend to zero in probability as $\|\Pi\| \to 0$. For the first term we in turn write

$$\sum_{i=1}^n 4B_{t_{i-1}}^2 (B_{t_i} - B_{t_{i-1}})^2 = \sum_{i=1}^n 4B_{t_{i-1}}^2 |t_i - t_{i-1}| + \sum_{i=1}^n 4B_{t_{i-1}}^2 \Big[(B_{t_i} - B_{t_{i-1}})^2 - |t_i - t_{i-1}|\Big] \tag{6.16}$$

The second term on the right has mean zero and variance equal to

$$\sum_{i=1}^{n} 16E\big(B_{t_{i-1}}^4\big)\text{Var}\big((B_{t_i}-B_{t_{i-1}})^2\big) \leqslant 16t^4\, E\big(|\mathcal{N}(0,1)^4\big)\,\text{Var}\big(\mathcal{N}(0,1)^2\big)\sum_{i=1}^{n}|t_i-t_{i-1}|^2 \tag{6.17}$$

where the bound on the right uses the scaling property of Brownian increments. The quantity on the right is order $\|\Pi\|$ and thus decays to zero as $\|\Pi\| \to 0$. Since the first term on the right of (6.16) is the Riemann sum, we have proved that, for any sequence $\{\Pi_n\}_{n\geqslant 1}$ of partitions of $[0,t]$,

$$\|\Pi_n\| \to 0 \quad \Rightarrow \quad V_t^{(2)}(B^2,\Pi_n) \ \underset{n\to\infty}{\overset{P}{\longrightarrow}} \int_0^t 4B_s^2\,\mathrm{d}s \tag{6.18}$$

More interesting that this conclusion is the mechanics of the underlying calculation. The key point is to arrange terms to isolate Brownian increments over non-overlapping intervals. These increments are then formally handled according to the infinitesimal-calculus "rules"

$$(\mathrm{d}B_t)^2 = \mathrm{d}t \ \wedge\ \mathrm{d}B_t\,\mathrm{d}t = 0 \ \wedge (\mathrm{d}t)^2 = 0 \tag{6.19}$$

Using these, we easily check the formal validity of the following claim:

Lemma 6.4 *Let $\{B_t\colon t\in[0,\infty)\}$ be standard Brownian motion. Then for any $f\in C^1(\mathbb{R})$, any $t\geqslant 0$ and any sequence $\{\Pi_n\}_{n\geqslant 1}$ of partitions of $[0,t]$,*

$$\|\Pi_n\| \to 0 \quad \Rightarrow \quad V_t^{(2)}(f\circ B,\Pi_n) \ \underset{n\to\infty}{\overset{P}{\longrightarrow}} \int_0^t f'(B_s)^2\,\mathrm{d}s \tag{6.20}$$

We leave a detailed proof of this lemma (which has to go beyond just applying the "rules" (6.19)) to a homework assignment. To give a hint, we note that the proof should be based on the following use of the first-order Taylor expansion

$$\sum_{i=1}^{n}\big[f(B_{t_i})-f(B_{t_{i-1}})\big]^2 = \sum_{i=1}^{n} f'(B_{t_{i-1}})^2(B_{t_i}-B_{t_{i-1}})^2 + \text{error} \tag{6.21}$$

Leaving the explicit form and control of the "error" to the reader, inspired by (6.16) we now write the sum on the right as

$$\begin{aligned}\sum_{i=1}^{n} f'(B_{t_{i-1}})^2(B_{t_i}-B_{t_{i-1}})^2 \\ = \sum_{i=1}^{n} f'(B_{t_{i-1}})^2(t_i-t_{i-1}) + \sum_{i=1}^{n} f'(B_{t_{i-1}})^2\Big[(B_{t_i}-B_{t_{i-1}})^2-(t_i-t_{i-1})\Big]\end{aligned} \tag{6.22}$$

The first object on the right is the Riemann sum associated with the integral on the right of (6.20). The second sum is treated using the observation that $f'(B_{t_{i-1}})$ and the quantity $(B_{t_i}-B_{t_{i-1}})^2-(t_i-t_{i-1})$ are independent, with the latter of zero mean. The expression is thus a martingale whose second moment is bounded by sum of the squares of the second moments of the increments; i.e., by the expression

$$E\biggl(\sum_{i=1}^n f'(B_{t_{i-1}})^4\,\mathrm{Var}\bigl((B_{t_i}-B_{t_{i-1}})^2\bigr)\biggr) \tag{6.23}$$

If f' is bounded (by its supremum norm $\|f'\|_\infty$) , this would be at most

$$\mathrm{Var}(\mathcal{N}(0,1)^2)\|f'\|_\infty^4 t\|\Pi\| \tag{6.24}$$

proving that this term does not contribute as $\|\Pi\|\to 0$. An unbounded f' situation has to be handled by a suitable truncation.

A main take-away message of Proposition 6.2 is that the limiting *quadratic-variation process* is generally random. Note, however, that all of the above statements concerning the second variation involve a limit in probability along sequences of partitions and thus offer no information about the supremum of $V_t^{(2)}(f\circ B,\Pi)$ over all partitions Π of $[0,t]$. And, indeed, this supremum is actually badly behaved:

Lemma 6.5 *For a.e. path of the standard Brownian motion and every $t>0$,*

$$\sup_\Pi V_t^{(2)}(B,\Pi)=+\infty \tag{6.25}$$

where the supremum is over the partitions of $[0,t]$.

We again leave the proof of this lemma to homework. The conclusion of this section is thus the following: While Brownian paths (and generally, functions thereof) are almost never of finite second variation, the concept of quadratic variation introduced by the in-probability limits along sequences of partitions above results in an interesting characterization of the sample path regularity.

Further reading: Chapter 3 of Øksendal