5 Stochastic integration

From the discrete-time theory developed in Chapters 1–3, we know that the trading gains or losses from a self-financing strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ are described by the *stochastic integral*

$$G(\vartheta) = \vartheta \bullet S = \int \vartheta \, \mathrm{d}S = \sum_j \vartheta_j^{\mathrm{tr}} \Delta S_j = \sum_j \vartheta_j^{\mathrm{tr}} (S_j - S_{j-1}).$$

To be able to develop an analogous theory in continuous time, we therefore need to understand how to define, and how to work with, a continuous-time stochastic integral process $\int \vartheta \, \mathrm{d}S$. From classical integration theory, the obvious idea is to start with approximating Riemann sums of the form $\sum \vartheta_{\tilde{t}_i}^{\mathrm{tr}} (S_{t_{i+1}} - S_{t_i})$, with $\tilde{t}_i$ lying between t_i and t_{i+1}, and then pass to the limit in a suitable sense. The simplest idea for that would be to fix ω, look at the trajectories $t \mapsto S_t(\omega)$ and $t \mapsto \vartheta_t(\omega)$ and take limits of

$$\sum \vartheta_{\tilde{t}_i}(\omega) \big(S_{t_{i+1}}(\omega) - S_{t_i}(\omega) \big)$$

like in courses on measure and integration theory. But unfortunately, this works well (i.e., for many integrands ϑ) only if the function $t \mapsto S_t(\omega)$ is of *finite variation* — and this would immediately exclude as integrator a process like Brownian motion which does not have this property. So one must use a different approach, and this will be explained in this chapter. For an amplification (and proof) of the above point that "naive stochastic integration is impossible", we refer to Protter [13, Section I.8]; the idea originally goes back to C. Stricker.

Remarks. 1) To avoid misunderstandings later, let us clarify that defining stochastic integrals as above in a pathwise manner (i.e. ω by ω) may well be possible if the integrator S and the integrand ϑ match up nicely enough, even if $t \mapsto S_t(\omega)$ is not of finite variation. We shall see this later in the context of Itô's formula, where ϑ has the form $\vartheta_t = g(S_{t-})$ for some C^1-function g. But if we want to fix S and allow many ϑ without imposing undue restrictions, an ω-wise approach leads to problems.

2) In classical integration theory, it does not matter in which point $\tilde{t}_i \in [t_i, t_{i+1}]$ one evaluates the integrand when defining the Riemann approximation. For stochastic integrals, this is different — choosing the left endpoint $\tilde{t}_i = t_i$ leads to the *Itô integral*, the right endpoint $\tilde{t}_i = t_{i+1}$ yields the *forward integral*, and the midpoint choice $\tilde{t}_i = \frac{1}{2}(t_i + t_{i+1})$ produces the *Stratonovich integral*. However, for applications in finance, it is clear that one must choose $\tilde{t}_i = t_i$ (and hence the Itô integral) because the strategy must be decided before the price move. $\diamond$

5.1 The basic construction

Our goal in this section is to construct a stochastic integral process $H \bullet M = \int H \, \mathrm{d}M$ when M is a (real-valued) local martingale null at 0 and H is a (real-valued) predictable process with a suitable integrability property (relative to M). In Section 5.3 below, we also explain how to extend this from local martingales to semimartingales; but the key step and the main work happen in the martingale case.

Remark. For simplicity, we take both M and H to be real-valued. It is reasonably straightforward, although somewhat technical, to extend the theory from this section to M and H that are both $\mathbb{R}^d$-valued, and we comment on the necessary changes a bit later. We then also point out some of the pitfalls that one has to avoid in that context. $\diamond$

Throughout this chapter, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ satisfying the usual conditions of right-continuity and P-completeness. If needed, we define $\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t$. We also fix a (real-valued) *local martingale* $M = (M_t)_{t \geq 0}$ null at 0 (as defined before Proposition 4.2.3) and having RCLL (right-continuous with left limits) trajectories. (The latter property, as pointed out earlier in Remark 4.2.1, is not a restriction; we can always find an RCLL version of M thanks to the usual conditions on $\mathbb{F}$.) Because we want to define stochastic integrals $\int_a^b H \, \mathrm{d}M$ and these are always over half-open intervals of the form $(a, b]$ with $0 \leq a < b \leq \infty$, the value of M at 0 is irrelevant and it is enough to look at processes $H = (H_t)$ defined for $t > 0$. This will simplify some definitions. For any process $Y = (Y_t)_{t \geq 0}$ with RCLL trajectories, we denote by $\Delta Y_t := Y_t - Y_{t-} := Y_t - \lim_{s \to t, s < t} Y_s$ the jump of Y at time $t > 0$.

The *simplest example* to be kept in mind is when $M = W$ is a Brownian motion. From Proposition 4.2.3, we know that both W itself and $(W_t^2 - t)_{t\geq 0}$ are then martingales, and by Theorem 4.1.4, the quantity t we subtract from W_t^2 is the quadratic variation of W, which can be obtained as a pathwise limit of sums of squared increments of W. As already mentioned in Remark 4.1.5, a similar result is true for a general local martingale M, and this is the key for constructing stochastic integrals.

Theorem 1.1. *For any local martingale $M = (M_t)_{t\geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t\geq 0}$ null at 0 with $\Delta[M] = (\Delta M)^2$ and having the property that $M^2 - [M]$ is also a local martingale. This process $[M]$ can be obtained as the quadratic variation of M in the following sense: There exists a sequence $(\Pi_n)_{n\in I\!N}$ of partitions of $[0,\infty)$ with $|\Pi_n| \to 0$ as $n \to \infty$ such that*

$$P\left[[M]_t(\omega) = \lim_{n\to\infty} \sum_{t_i\in\Pi_n} \big(M_{t_i\wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega)\big)^2 \text{ for all } t \geq 0\right] = 1.$$

We call $[M]$ the optional quadratic variation *or* square bracket *process of M.*

If M satisfies $\sup_{0\leq s\leq T} |M_s| \in L^2$ for some $T > 0$ (and hence is in particular a square-integrable martingale on $[0,T]$), then $[M]$ is integrable on $[0,T]$ (i.e. $[M]_T \in L^1$) and $M^2 - [M]$ is a martingale on $[0,T]$.

Proof. See Protter [13, Section II.6] or Dellacherie/Meyer [5, Theorem VII.42] or Jacod/Shiryaev [11, Section I.4c].

Remarks. 1) Recall from Theorem 1.4 that Brownian motion W has $[W]_t = t$.

2) Note that $[M]$ has paths of finite variation. So one can easily define integrals $\int \ldots \, \mathrm{d}[M]$ in a pathwise manner as usual Lebesgue–Stieltjes integrals. This does not need any new theory.

3) The sequence $(\Pi_n)_{n\in I\!N}$ of partitions in Theorem 1.4 of course depends on M. ◇

For two local martingales M, N null at 0, we define the (optional) *covariation process*

$[M,N]$ by polarisation, i.e.

$$[M,N] := \frac{1}{4}([M+N]-[M-N]).$$

From the characterisation of $[M]$ in Theorem 1.1, it follows easily that the operation $[\,\cdot\,,\,\cdot\,]$ is bilinear, and also that $[M,N]$ is the unique adapted RCLL process B null at 0, of finite variation with $\Delta B = \Delta M \Delta N$ and such that the difference $MN - B$ is again a local martingale.

Remark 1.2. **1)** If M is a square-integrable martingale, then $[M]$ is integrable and therefore, by the general theory of stochastic processes, admits a so-called *(predictable) compensator* or *dual predictable projection*: There exists a unique increasing predictable integrable process $\langle M\rangle = (\langle M\rangle_t)_{t\geq 0}$ null at 0 such that $[M]-\langle M\rangle$, and therefore also $M^2 - \langle M\rangle = M^2 - [M] + [M] - \langle M\rangle$, is a martingale. The process $\langle M\rangle$ is called the *sharp bracket* (or sometimes the *predictable variance*) process of M. See Dellacherie/Meyer [5, Theorem VI.65 and Definition VI.77] or Jacod/Shiryaev [11, Theorem I.3.17]. Note that we still need to define what "predictable" means in continuous time.

2) Once we know what localisation means (see the end of this section for more details), we can easily extend the results in 1). It is enough if M is a locally square-integrable local martingale; then $\langle M\rangle$ is also locally integrable, and then both $[M]-\langle M\rangle$ and $M^2-\langle M\rangle$ are local martingales.

3) We already point out here that any adapted process which is continuous is automatically locally bounded (see later for the definition) and therefore also locally square-integrable. Again, we refer to the end of this section for more details.

4) If M is *continuous*, then so is $[M]$, because $\Delta[M] = (\Delta M)^2 = 0$. This implies then also that $[M] = \langle M\rangle$. In particular, for a Brownian motion W, we have $[W]_t = \langle W\rangle_t = t$ for all $t \geq 0$.

5) If both M and N are locally square-integrable (e.g. if they are continuous), we also get $\langle M,N\rangle$ via polarisation.

6) If M is $\mathbb{R}^d$-valued, then $[M]$ becomes a $d\times d$-matrix-valued process with entries $[M]^{ik} = [M^i, M^k]$. To work with that, one needs to establish more properties. The same applies to $\langle M\rangle$, if it exists.

7) The key difference between $[M]$ and $\langle M\rangle$ is that $[M]$ exists for any local martingale M null at 0, whereas the existence of $\langle M\rangle$ requires some extra local integrability of M. $\diamond$

Definition. We denote by $b\mathcal{E}$ the set of all *bounded elementary processes* of the form

$$H=\sum_{i=0}^{n-1} h_i I_{(t_i,t_{i+1}]}$$

with $n\in I\!N$, $0\le t_0<t_1<\cdots<t_n<\infty$ and each h_i a bounded (real-valued) $\mathcal{F}_{t_i}$-measurable random variable. For any stochastic process $X=(X_t)_{t\ge 0}$, we then define the *stochastic integral* $\int H\,\mathrm{d}X$ of $H\in b\mathcal{E}$ by

$$\int_0^t H_s\,\mathrm{d}X_s := H\bullet X_t := \sum_{i=0}^{n-1} h_i(X_{t_{i+1}\wedge t}-X_{t_i\wedge t})\qquad \text{for } t\ge 0.$$

Note that if X is RCLL, then so is $\int H\,\mathrm{d}X = H\bullet X$.

If X and H are both $I\!R^d$-valued, the integral is still real-valued, and we simply replace products by scalar products everywhere. But then Lemma 1.3 below looks more complicated.

Lemma 1.3. *Suppose that M is a square-integrable martingale (i.e., M is a martingale with $M_t\in L^2$ for all $t\ge 0$, or equivalently with $\sup_{0\le s\le T}|M_s|\in L^2$ for all $T>0$). For every $H\in b\mathcal{E}$, the stochastic integral process $H\bullet M=\int H\,\mathrm{d}M$ is then also a square-integrable martingale, and we have $[H\bullet M]=\int H^2\,\mathrm{d}[M]$ and the* isometry property

$$E[(H\bullet M_\infty)^2]=E\bigg[\bigg(\int_0^\infty H_s\,\mathrm{d}M_s\bigg)^2\bigg] \tag{1.1}$$
$$=E\bigg[\sum_{i=0}^{n-1} h_i^2([M]_{t_{i+1}}-[M]_{t_i})\bigg]$$
$$=E\bigg[\int_0^\infty H_s^2\,\mathrm{d}[M]_s\bigg].$$

Note that the last $\mathrm{d}[M]$-integral can be defined ω by ω via classical measure and integration theory, because $t \mapsto [M]_t(\omega)$ is increasing and hence of finite variation. But of course it is here also just a finite sum, because H has such a simple form.

Proof of Lemma 1.3. Adaptedness of $H \bullet M$ is clear, and so is square-integrability because H is bounded and each $H \bullet M_t$ is just a finite sum. Moreover, H is identically 0 after t_n so that both infinite integrals actually end at t_n. We first argue the martingale property, for simplicity only for $s = t_i$, $t = t_{i+1}$. [$\rightarrow$ *Exercise:* Prove this in detail for arbitrary $s \le t$.] Indeed, by first using that h_i is $\mathcal{F}_{t_i}$-measurable and bounded, and then that M is a martingale, we get

$$E[H \bullet M_t - H \bullet M_s \,|\, \mathcal{F}_s] = E[h_i(M_{t_{i+1}} - M_{t_i}) \,|\, \mathcal{F}_{t_i}] = h_i E[M_{t_{i+1}} - M_{t_i} \,|\, \mathcal{F}_{t_i}] = 0.$$

Next, it is easy to check [$\rightarrow$ *exercise*] for any square-integrable martingale N that

$$E[N_t^2 - N_s^2 \,|\, \mathcal{F}_s] = E[(N_t - N_s)^2 \,|\, \mathcal{F}_s] \qquad \text{for } s \le t.$$

Applying this once to $H \bullet M$ and once to M yields

$$\begin{aligned} E[(H \bullet M_{t_{i+1}})^2 - (H \bullet M_{t_i})^2 \,|\, \mathcal{F}_{t_i}] &= E[(H \bullet M_{t_{i+1}} - H \bullet M_{t_i})^2 \,|\, \mathcal{F}_{t_i}] \\ &= E[h_i^2 (M_{t_{i+1}} - M_{t_i})^2 \,|\, \mathcal{F}_{t_i}] \\ &= h_i^2 E[M_{t_{i+1}}^2 - M_{t_i}^2 \,|\, \mathcal{F}_{t_i}] \\ &= h_i^2 E\big[[M]_{t_{i+1}} - [M]_{t_i} \,\big|\, \mathcal{F}_{t_i}\big] \\ &= E\big[h_i^2([M]_{t_{i+1}} - [M]_{t_i}) \,\big|\, \mathcal{F}_{t_i}\big] \\ &= E\big[H^2 \bullet [M]_{t_{i+1}} - H^2 \bullet [M]_{t_i} \,\big|\, \mathcal{F}_{t_i}\big], \end{aligned}$$

where we have used twice that h_i is $\mathcal{F}_{t_i}$-measurable and bounded, and in the fourth step also that $M^2 - [M]$ is a martingale. Summing up and taking expectations then gives (1.1). Moreover, it is not very difficult to argue that

$$\Delta\bigg(\int H^2 \,\mathrm{d}[M]\bigg) = H^2 \Delta[M] = H^2 (\Delta M)^2 = \big(\Delta(H \bullet M)\big)^2$$

for $H \in b\mathcal{E}$, by exploiting that H is piecewise constant and $\Delta[M] = (\Delta M)^2$. In view of Theorem 1.1 and the uniqueness there, the combination of these two properties can also be formulated as saying that

$$[H \bullet M] = \left[\int H \, \mathrm{d}M \right] = \int H^2 \, \mathrm{d}[M] = H^2 \bullet [M] \qquad \text{for } H \in b\mathcal{E}.$$

This completes the proof. **q.e.d.**

Remark. The argument in the proof of Lemma 1.3 actually shows that the process $(H \bullet M)^2 - \int H^2 \, \mathrm{d}[M]$ is a martingale. [$\to$ *Exercise:* Prove this in detail.] See also Remark 1.2. $\diamond$

Our goal is now to extend the above results from $H \in b\mathcal{E}$ to a larger class of integrands. To that end, it is useful to view stochastic processes as random variables on the *product space* $\overline{\Omega} := \Omega \times (0, \infty)$. (Recall that the values at 0 are irrelevant for stochastic integrals.) We define the *predictable σ-field* $\mathcal{P}$ on $\overline{\Omega}$ as the σ-field generated by all adapted left-continuous processes, and we call a stochastic process $H = (H_t)_{t>0}$ *predictable* if it is $\mathcal{P}$-measurable when viewed as a mapping $H : \overline{\Omega} \to I\!R$. As a consequence, every $H \in b\mathcal{E}$ is then predictable as it is adapted and left-continuous. We also define the (possibly infinite) measure $P_M := P \otimes [M]$ on $(\overline{\Omega}, \mathcal{P})$ by setting

$$\int_{\overline{\Omega}} Y \, \mathrm{d}P_M := E_M[Y] := E\left[\int_0^\infty Y_s(\omega) \, \mathrm{d}[M]_s(\omega) \right]$$

for $Y \geq 0$ predictable; the inner integral is defined ω-wise as a Lebesgue–Stieltjes integral because $t \mapsto [M]_t(\omega)$ is increasing, null at 0 and RCLL and so can be viewed as the distribution function of a (possibly infinite) ω-dependent measure on $(0, \infty)$. (Actually, one could even allow Y to be product-measurable here.) Note that $P_M = P \otimes [M]$ is not a product measure in general because unlike $\langle W \rangle_t = t$, the quadratic variation $[M]$ of a general local martingale M depends on both t and ω. Finally, we introduce the space

$$\begin{aligned} L^2(M) &:= L^2(M, P) := L^2(\overline{\Omega}, \mathcal{P}, P_M) \\ &= \Bigg\{ \text{all (equivalence classes of) predictable } H = (H_t)_{t>0} \text{ such that} \\ &\qquad \|H\|_{L^2(M)} := (E_M[H^2])^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 \, \mathrm{d}[M]_s \right] \right)^{\frac{1}{2}} < \infty \Bigg\}. \end{aligned}$$

(As usual, taking equivalence classes means that we identify H and H' if they agree P_M-a.e. on $\overline{\Omega}$ or, equivalently, if $E[\int_0^\infty (H_s - H'_s)^2 \,\mathrm{d}[M]_s] = 0$.)

With the above notations, we can restate the first half of Lemma 1.3 as follows:

For a fixed square-integrable martingale M, the mapping $H \mapsto H \bullet M$ is linear and goes from $b\mathcal{E}$ to the space $\mathcal{M}_0^2$ of all RCLL martingales $N = (N_t)_{t \geq 0}$ null at 0 which satisfy $\sup_{t \geq 0} E[N_t^2] < \infty$.

The last assertion is true because each $H \bullet M$ remains constant after some t_n given by $H \in b\mathcal{E}$, and because Doob's inequality gives for any martingale N and any $t \geq 0$ that

$$E\Big[\sup_{0 \leq s \leq t} |N_s|^2\Big] \leq 4E[N_t^2].$$

Now the martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t \to \infty} N_t$ P-a.s., and we have $N_\infty \in L^2$ by Fatou's lemma, and the process $(N_t)_{0 \leq t \leq \infty}$ defined up to ∞, i.e. on the *closed* interval $[0, \infty]$, is still a martingale. Moreover, Doob's maximal inequality implies that two martingales N and N' which have the same final value, i.e. $N_\infty = N'_\infty$ P-a.s., must coincide. Therefore we can identify $N \in \mathcal{M}_0^2$ with its limit $N_\infty \in L^2(\mathcal{F}_\infty, P)$, and so $\mathcal{M}_0^2$ becomes a *Hilbert space* with the norm

$$\|N\|_{\mathcal{M}_0^2} = \|N_\infty\|_{L^2} = (E[N_\infty^2])^{\frac{1}{2}}$$

and the scalar product

$$(N, N')_{\mathcal{M}_0^2} = (N_\infty, N'_\infty)_{L^2} = E[N_\infty N'_\infty].$$

Rephrasing Lemma 1.3 once again, we see that

the mapping $H \mapsto H \bullet M$ from $b\mathcal{E}$ to $\mathcal{M}_0^2$ is linear *and an* isometry because (1.1) says that for $H \in b\mathcal{E}$,

$$\|H \bullet M\|_{\mathcal{M}_0^2} = \big(E[(H \bullet M_\infty)^2]\big)^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 \,\mathrm{d}[M]_s\right]\right)^{\frac{1}{2}} = \|H\|_{L^2(M)}. \tag{1.2}$$

By general principles, this mapping can therefore be uniquely extended to the *closure of $b\mathcal{E}$ in $L^2(M)$*; in other words, we can define a stochastic integral process $H \bullet M$ for every H that can be approximated, with respect to the norm $\|\cdot\|_{L^2(M)}$, by processes from $b\mathcal{E}$, and the resulting $H \bullet M$ is again a martingale in $\mathcal{M}_0^2$ and still satisfies the isometry property (1.2).

(The argument behind these general principles is quite standard. If $(H^n)_{n\in I\!N}$ is a sequence of predictable processes converging to H with respect to $\|\cdot\|_{L^2(M)}$, then (H^n) is also a Cauchy sequence with respect to $\|\cdot\|_{L^2(M)}$. If all the H^n are in $b\mathcal{E}$, then the stochastic integral process $H^n \bullet M$ is well defined and in $\mathcal{M}_0^2$ for each n by Lemma 1.3. Moreover, by the isometry property in Lemma 1.3 for integrands in $b\mathcal{E}$, the sequence $(H^n \bullet M)_{n\in I\!N}$ is then also a Cauchy sequence in $\mathcal{M}_0^2$, and because $\mathcal{M}_0^2$ is a Hilbert space, hence complete, that Cauchy sequence must have a limit which is again in $\mathcal{M}_0^2$. This limit is then defined to be the stochastic integral $H \bullet M$ of H with respect to M. That the isometry property extends to the limit is also standard.)

The crucial question now is of course how we can describe the closure of $b\mathcal{E}$ and especially how big it is — the bigger the better, because we then have many integrands.

Proposition 1.4. *Suppose that M is in $\mathcal{M}_0^2$. Then:*

1) *$b\mathcal{E}$ is dense in $L^2(M)$, i.e. the closure of $b\mathcal{E}$ in $L^2(M)$ is $L^2(M)$. In other words, every $H \in L^2(M)$ can be written as a limit, with respect to the norm $\|\cdot\|_{L^2(M)}$, of a sequence $(H^n)_{n\in I\!N}$ in $b\mathcal{E}$.*

2) *For every $H \in L^2(M)$, the stochastic integral process $H \bullet M = \int H \,\mathrm{d}M$ is well defined, in $\mathcal{M}_0^2$ and satisfies* (1.2).

Proof. Assertion 1) uses a martingale approximation argument on $\overline{\Omega}$ which we do not give here. However, we point out that the assumption $M \in \mathcal{M}_0^2$ is used to ensure that P_M is a finite measure. Assertion 2) is then clear from the discussion above. **q.e.d.**

By definition, saying that M is in $\mathcal{M}_0^2$ means that M is an RCLL martingale null at 0 with $\sup_{t\geq 0} E[M_t^2] < \infty$. In particular, we then have $E[M_t^2] < \infty$ for every $t \geq 0$ so that every $M \in \mathcal{M}_0^2$ is also a square-integrable martingale. However, the converse is not true; Brownian motion W for example is a martingale and has $E[W_t^2] = t$ so that $\sup_{t\geq 0} E[W_t^2] = +\infty$, which means that BM is not in $\mathcal{M}_0^2$. This makes it clear that we need to extend our approach to stochastic integration further. This can be done via *localisation*.

Definition. We call a local martingale M null at 0 *locally square-integrable* and write $M \in \mathcal{M}_{0,\mathrm{loc}}^2$ if there is a sequence of stopping times $\tau_n \nearrow \infty$ P-a.s. such that $M^{\tau_n} \in \mathcal{M}_0^2$ for each n. We say for a predictable process H that $H \in L_{\mathrm{loc}}^2(M)$ if there exists a sequence of stopping times $\tau_n \nearrow \infty$ P-a.s. such that $HI_{]\!]0,\tau_n]\!]} \in L^2(M)$ for each n. Here we use the *stochastic interval* notation $]\!]0,\tau_n]\!] := \{(\omega,t) \in \overline{\Omega} : 0 < t \leq \tau_n(\omega)\}$.

More generally, if we have a class $\mathcal{C}$ of stochastic processes, we define the *localised class* $\mathcal{C}_{\mathrm{loc}}$ by saying that a process X *is in* $\mathcal{C}_{\mathrm{loc}}$ or that X *is locally in* $\mathcal{C}$ if there exists a sequence of stopping times $\tau_n \nearrow \infty$ P-a.s. such that X^{τ_n} is in $\mathcal{C}$ for each n. If the process we consider is an integrand H, then we have to require instead that $HI_{]\!]0,\tau_n]\!]}$ is in $\mathcal{C}$ for each n.

For $M \in \mathcal{M}_{0,\mathrm{loc}}^2$ and $H \in L_{\mathrm{loc}}^2(M)$, defining the stochastic integral is straightforward; we simply set

$$H \bullet M := (HI_{]\!]0,\tau_n]\!]}) \bullet M^{\tau_n} \qquad \text{on }]\!]0,\tau_n]\!]$$

which gives a definition on all of $\overline{\Omega}$, because $\tau_n \nearrow \infty$ so that $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$. The only point we need to check is that this definition is *consistent*, i.e. that the definition on $]\!]0,\tau_{n+1}]\!] \supseteq]\!]0,\tau_n]\!]$ does not clash with the definition on $]\!]0,\tau_n]\!]$. This can be done by using the (subsequently listed) properties of stochastic integrals, but we do not go into details here. Of course, $H \bullet M$ is then in $\mathcal{M}_{0,\mathrm{loc}}^2$.

Remarks. 1) A closer look at the developments so far shows that the *definitions* (but not the preceding results and arguments) for P_M and $L^2(M)$ only need $[M]$; hence one can introduce and use them for any local martingale M, due to Theorem 1.1.

2) One can also define a stochastic integral process $H \bullet M$ for $H \in L^2_{\mathrm{loc}}(M)$ when M is a general local martingale, but this requires substantially more theory. For more details, see Dellacherie/Meyer [5, Theorem VIII.37].

3) If M is $\mathbb{R}^d$-valued with components M^i that all are local martingales null at 0, one can also define the so-called *vector stochastic integral* $H \bullet M$ for $\mathbb{R}^d$-valued predictable processes in a suitable space $L^2_{\mathrm{loc}}(M)$; the result is then a real-valued process. Details can be found in Jacod/Shiryaev [11, Sections III.4a and III.6a]. However, one *warning* is indicated: $L^2_{\mathrm{loc}}(M)$ is not obtained by just asking that each component H^i should be in $L^2_{\mathrm{loc}}(M^i)$ and then setting $H \bullet M = \sum_i H^i \bullet M^i$. In fact, it can happen that $H \bullet M$ is well defined whereas the individual $H^i \bullet M^i$ are not. So the intuition for the multidimensional case is that

$$\text{“}\int H \,\mathrm{d}M = \int \sum_i H^i \,\mathrm{d}M^i \neq \sum_i \int H^i \,\mathrm{d}M^i\text{”},$$

as we have already pointed out in Remark 1.2.2.

4) One can extend the stochastic integral even further to more general integrands in a space called $L(M)$, but this becomes technical and also has a nontrivial *pitfall*: There are (real-valued) local martingales M and predictable integrands H such that the stochastic integral process $\int H \,\mathrm{d}M$ is well defined, but not a local martingale (!). This is in marked contrast to discrete time; see Theorem 1.3.1. We remark, however, that this can only happen if M has jumps. ◇

To end this section on a positive note, let us consider the case where M is a *continuous* local martingale null at 0, briefly written as $M \in \mathcal{M}^c_{0,\mathrm{loc}}$. This includes in particular the case of a Brownian motion W. Then M is in $\mathcal{M}^2_{0,\mathrm{loc}}$ because it is even *locally bounded*: For the stopping times

$$\tau_n := \inf\{t \geq 0 : |M_t| > n\} \nearrow \infty \qquad P\text{-a.s.},$$

we have by continuity that $|M^{\tau_n}| \leq n$ for each n, because

$$\left|M_t^{\tau_n}\right| = |M_{t\wedge\tau_n}| = \begin{cases} |M_t| \leq n & \text{if } t < \tau_n, \\ |M_{\tau_n}| = n & \text{if } t \geq \tau_n. \end{cases}$$

(Note that continuity of M is only used to obtain the equality $|M_{\tau_n}| = n$; everything else works just as well if M is only assumed to be adapted and RCLL.) The set $L^2_{\mathrm{loc}}(M)$ of nice integrands for M can here be explicitly described as

$$L^2_{\mathrm{loc}}(M) = \Big\{ \text{all predictable processes } H = (H_t)_{t>0} \text{ such that}$$
$$\int_0^t H_s^2 \,\mathrm{d}[M]_s = \int_0^t H_s^2 \,\mathrm{d}\langle M\rangle_s < \infty \ P\text{-a.s. for each } t \geq 0 \Big\}.$$

Finally, the resulting stochastic integral $H \bullet M = \int H \,\mathrm{d}M$ is then (as we shall see from the properties in Section 5.2 below) also a continuous local martingale, and of course null at 0.

5.2 Properties

As with usual integrals, one very rarely computes a stochastic integral by passing to the limit from some approximation. One works with stochastic integrals by using a set of *rules* and *properties*. These are listed in this section, without proofs.

- *(Local) Martingale properties:*
  - If M is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, then $\int H\,\mathrm{d}M$ is a local martingale in $\mathcal{M}^2_{0,\mathrm{loc}}$. If $H \in L^2(M)$, then $\int H\,\mathrm{d}M$ is even a martingale in $\mathcal{M}^2_0$.
  - If M is a local martingale and H is predictable and locally bounded (which means that there are stopping times $\tau_n \nearrow \infty$ P-a.s. such that $HI_{]\!]0,\tau_n]\!]}$ is bounded by a constant c_n, say, for each $n \in I\!N$), then $\int H\,\mathrm{d}M$ is a local martingale.
  - If M is a martingale in $\mathcal{M}^2_0$ and H is predictable and bounded, then $\int H\,\mathrm{d}M$ is again a martingale in $\mathcal{M}^2_0$.
  - *Warning*: If M is a martingale and H is predictable and bounded, then $\int H\,\mathrm{d}M$ need not be a martingale; this is in striking contrast to the situation in discrete time.

- *Linearity:*
  - If M is a local martingale and H, H' are in $L^2_{\mathrm{loc}}(M)$ and $a, b \in I\!R$, then also $aH+bH'$ is in $L^2_{\mathrm{loc}}(M)$ and

$$(aH + bH')\bullet M = (aH)\bullet M + (bH')\bullet M = a(H\bullet M) + b(H'\bullet M).$$

- *Associativity:*
  - If M is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$, we already know that $H\bullet M$ is again a local martingale. Then a predictable process K is in $L^2_{\mathrm{loc}}(H\bullet M)$ if and only if the product KH is in $L^2_{\mathrm{loc}}(M)$, and then

$$K\bullet(H\bullet M) = (KH)\bullet M,$$

i.e.

$$\int K \,\mathrm{d}\Big(\int H \,\mathrm{d}M \Big) = \int KH \,\mathrm{d}M.$$

- *Behaviour under stopping:*
  - Suppose that M is a local martingale, $H \in L^2_{\mathrm{loc}}(M)$ and τ is a stopping time. Then M^τ is a local martingale by the stopping theorem, H is in $L^2_{\mathrm{loc}}(M^\tau)$, $HI_{]\!]0,\tau]\!]}$ is in $L^2_{\mathrm{loc}}(M)$, and we have

    $$(H \bullet M)^\tau = H \bullet (M^\tau) = (HI_{]\!]0,\tau]\!]}) \bullet M = (HI_{]\!]0,\tau]\!]}) \bullet (M^\tau).$$

    In words: A stopped stochastic integral is computed by either first stopping the integrator and then integrating, or setting the integrand equal to 0 after the stopping time and then integrating, or combining the two.

- *Quadratic variation and covariation:*
  - Suppose that M, N are local martingales, $H \in L^2_{\mathrm{loc}}(M)$ and $K \in L^2_{\mathrm{loc}}(N)$. Then

    $$\Big[\int H \,\mathrm{d}M, N \Big] = \int H \,\mathrm{d}[M, N]$$

    and

    $$\Big[\int H \,\mathrm{d}M, \int K \,\mathrm{d}N \Big] = \int HK \,\mathrm{d}[M, N].$$

    In words: The covariation process of two stochastic integrals is obtained by integrating the product of the integrands with respect to the covariation process of the integrators.
  - In particular, $[\int H \,\mathrm{d}M] = \int H^2 \,\mathrm{d}[M]$. (We have seen this already for $H \in b\mathcal{E}$ in Lemma 1.3.)

- *Jumps:*
  - Suppose M is a local martingale and $H \in L^2_{\mathrm{loc}}(M)$. Then we already know that $H \bullet M$ is in $\mathcal{M}^2_{0,\mathrm{loc}}$ and therefore RCLL. Its jumps are given by

$$\Delta\left(\int H \,\mathrm{d}M\right)_t = H_t \,\Delta M_t \qquad \text{for } t > 0,$$

    where $\Delta Y_t := Y_t - Y_{t-}$ again denotes the jump at time t of a process Y with trajectories which are RCLL (right-continuous and having left limits).

Example. To illustrate why the direct use of the definitions is complicated, let us compute the stochastic integral $\int W \,\mathrm{d}W$ for a Brownian motion W. This is well defined because $M := W$ is in $\mathcal{M}^2_{0,\mathrm{loc}}$ (it is even continuous) and $H := W$ is predictable and locally bounded, because it is adapted and continuous.

Because

$$2W_{t_i}(W_{t_{i+1}} - W_{t_i}) = W^2_{t_{i+1}} - W^2_{t_i} - (W_{t_{i+1}} - W_{t_i})^2$$

by elementary algebra, we obtain by summing up that

$$\sum_{t_i \in \Pi_n} W_{t_i \wedge t}(W_{t_{i+1} \wedge t} - W_{t_i \wedge t}) = \frac{1}{2}(W^2_t - W^2_0) - \frac{1}{2} \sum_{t_i \in \Pi_n} (W_{t_{i+1} \wedge t} - W_{t_i \wedge t})^2.$$

If the mesh size $|\Pi_n|$ of the partition sequence (Π_n) goes to 0, then the sum on the right-hand side converges P-a.s. to t by Theorem 4.1.4, if the partitions are also refining. We therefore expect to obtain

$$\int_0^t W_s \,\mathrm{d}W_s = \frac{1}{2} W^2_t - \frac{1}{2} t,$$

and we shall see later from Itô's formula that this is indeed correct. Note that we should expect the first term $\frac{1}{2}W^2_t$ from classical calculus (where we have $\int_0^x y \,\mathrm{d}y = \frac{1}{2}x^2$); the second-order correction term $\frac{1}{2}t$ appears due to the quadratic variation of Brownian trajectories.

Exercise: Prove directly (without using the above result) that the stochastic integral process $\int W \,\mathrm{d}W$ is a martingale, but not in $\mathcal{M}^2_0$.

Exercise: Compute the Stratonovich integral and the backward integral for $\int W \, \mathrm{d}W$, and analyse their properties.

Exercise: Prove that if H is predictable and bounded, then $\int H \, \mathrm{d}W$ is a square-integrable martingale.

Exercise: For any local martingale M null at 0 and any stopping time τ, prove that we have $[M]^\tau = [M^\tau]$.

5.3 Extension to semimartingales

So far, we have seen two ideas for constructing stochastic integrals $\int H \, \mathrm{d}X$ of some process H with respect to another process X:

a) In Section 5.1, we have taken for $X = M$ a *local martingale* null at 0 and for H a process in $L^2_{\mathrm{loc}}(M)$; this means that H must be predictable and possess some integrability.

b) If $X = A$ has trajectories $t \mapsto A_t(\omega)$ that are of *finite variation*, we can classically define $\int H_s(\omega) \, \mathrm{d}A_s(\omega)$ for each ω (pathwise) as a Lebesgue–Stieltjes integral. This requires some measurability and integrability for $s \mapsto H_s(\omega)$.

Because integration is a linear operation, the obvious and easy idea for an extension is therefore to look at processes that are sums of the above two types, because we can then define an integral with respect to the sum as the sum of the two integrals.

Definition. A *semimartingale* is a stochastic process $X = (X_t)_{t \geq 0}$ that can be decomposed as $X = X_0 + M + A$, where M is a local martingale null at 0 and A is an adapted RCLL process null at 0 and having trajectories of finite variation. A semimartingale X is called *special* if there exists such a decomposition where A is in addition predictable.

Remark 3.1. 1) If X is a special semimartingale, the decomposition with A predictable is *unique* and called the *canonical decomposition*. The uniqueness result is based on the useful fact that *any local martingale which is predictable and of finite variation must be constant*.

2) If X is a *continuous* semimartingale, both M and A can be chosen continuous as well. Therefore X is special because A is then predictable, as it is adapted and continuous.

3) If X is a semimartingale, we define its *optional quadratic variation* or *square bracket* process $[X] = ([X]_t)_{t \geq 0}$ via

$$[X] := [M] + 2[M, A] + [A] := [M] + 2 \sum \Delta M \Delta A + \sum (\Delta A)^2.$$

One can show that this is well defined and does not depend on the chosen decomposition of X. Moreover, $[X]$ can also be obtained as a quadratic variation similarly as in Theorem 1.1; see Section 6.1 below for more details. However, $X^2 - [X]$ is no longer a local martingale, but only a semimartingale in general. $\diamond$

If X is a semimartingale, we can define a stochastic integral $H \bullet X = \int H \, \mathrm{d}X$ at least for any process H which is predictable and locally bounded. We simply set

$$H \bullet X := H \bullet M + H \bullet A,$$

where $H \bullet M$ is as in Section 5.1 and $H \bullet A$ is defined ω-wise as a Lebesgue–Stieltjes integral. Of course one still needs to check that this is well defined (e.g. without ambiguity if X has several decompositions), but this can be done; see for instance Dellacherie/Meyer [5, Section VIII.1] or Jacod/Shiryaev [11, Section I.4d].

The resulting stochastic integral then has all the *properties* from Section 5.2 except those that rest in an essential way on the (local) martingale property; so the isometry property for example is of course lost. But we still have, for H predictable and locally bounded:

- $H \bullet X$ is a *semimartingale*.
- If X is special with canonical decomposition $X = X_0 + M + A$, then $H \bullet X$ is also special, with canonical decomposition $H \bullet X = H \bullet M + H \bullet A$.
  [This uses the non-obvious fact that if A is predictable and of finite variation and H is predictable and locally bounded, the pathwise defined integral $H \bullet A$ can be chosen to be predictable again.]
- *linearity:* same formula as before.
- *associativity:* same formula as before.
- *behaviour under stopping:* same formula as before.
- *quadratic variation and covariation:* same formula as before.

– *jumps:* same formula as before.

– If X is *continuous*, then so is $H \bullet X$; this is clear from $\Delta(H \bullet X) = H\Delta X = 0$.

In addition, there is also a sort of *dominated convergence theorem*: If H^n, $n \in I\!N$, are predictable processes with $H^n \to 0$ pointwise on $\overline{\Omega}$ and $|H^n| \le |H|$ for some locally bounded H, then $H^n \bullet X \to 0$ uniformly on compacts in probability, which means that

$$\sup_{0 \le s \le t} |H^n \bullet X_s| \longrightarrow 0 \qquad \text{in probability as } n \to \infty, \text{ for every } t \ge 0. \tag{3.1}$$

This can also be viewed as a *continuity property* of the stochastic integral operator $H \mapsto H \bullet X$, because (pointwise and locally bounded) convergence of (H^n) implies convergence of $(H^n \bullet X)$, in the ucp sense of (3.1).

From the whole approach above, the definition of a semimartingale looks completely ad hoc and rather artificial. But it turns out that this concept is in fact very natural and has a number of very good properties:

1) If X is a semimartingale and f is a C^2-function, then $f(X)$ is again a semimartingale. This will follow from *Itô's formula*, which even gives an explicit expression for $f(X)$.

2) If X is a semimartingale with respect to P and R is a probability measure equivalent to P, then X is still a semimartingale with respect to R. This will follow from *Girsanov's theorem*, which even gives a decomposition of X under R.

3) If X is any adapted process with RC trajectories, we can always define the (elementary) stochastic integral $H \bullet X$ for processes H in $b\mathcal{E}$. If X is such that this mapping on $b\mathcal{E}$ also has the continuity property (3.1) for any sequence $(H^n)_{n \in I\!N}$ in $b\mathcal{E}$ converging pointwise to 0 and with $|H^n| \le 1$ for all n, then X must in fact be a semimartingale. This deep result is due to *Bichteler and Dellacherie* and shows that semimartingales are a *natural class of integrators*.

One direct consequence of 2) for finance is that semimartingales are the natural processes to model discounted asset prices in financial markets. In fact, the fundamental theorem of asset pricing (in a suitably general version for continuous-time models) essentially says that a suitably arbitrage-free model should be such that S is a local martingale (or more generally a σ-martingale) under some $Q \approx P$. But then S is a Q-semimartingale and thus by 2) also a P-semimartingale.

Put differently, the above result implies that if we start with any model where S is *not a semimartingale*, there will be *arbitrage* of some kind. Things become different if one includes transaction costs; but in frictionless markets, one must be careful about this issue.

Remark. We have explained so far how to obtain a stochastic integral $H \bullet X$ for semimartingales X and locally bounded predictable H. The Bichteler–Dellacherie result shows that one cannot go beyond semimartingales without a serious loss; but because not every predictable process is locally bounded, one can ask if, for a given semimartingale X, there are more possible integrands H for X. This leads to the notion and definition of the class $L(X)$ of X-*integrable* processes; but the development of this requires rather advanced results and techniques from stochastic calculus, and so we cannot go into details here. See Dellacherie/Meyer [5, Section VIII.3] or Jacod/Shiryaev [11, Section III.6]. Alternatively, this is usually presented in the course “Mathematical Finance”. $\diamond$

6 Stochastic calculus

Our goal in this chapter is to provide the basic tools, results and techniques for working with stochastic processes and especially stochastic integrals in continuous time. This will be used in the next chapter when we discuss continuous-time option pricing and in particular the famous Black–Scholes formula.

Throughout this chapter, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)$ satisfying the usual conditions of right-continuity and P-completeness. For all local martingales, we then can and tacitly do choose a version with RCLL trajectories. For the time parameter t, we have either $t \in [0, T]$ with a fixed time horizon $T \in (0, \infty)$ or $t \geq 0$. In the latter case, we set

$$\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t := \sigma\bigg(\bigcup_{t \geq 0} \mathcal{F}_t\bigg).$$

6.1 Itô's formula

The question to be addressed in this section is very simple. If X is a semimartingale and f is some (suitable) function, what can we say about the stochastic process $f(X)$? What kind of process is it, and what does it look like in more detail?

In the simplest case, let $x : [0, \infty) \to \mathbb{R}$ be a function $t \mapsto x(t)$ and think of x as a typical trajectory $t \mapsto X_t(\omega)$ of X. The *classical chain rule from analysis* then says that if x is in C^1 (i.e. continuously differentiable) and $f : \mathbb{R} \to \mathbb{R}$ is in C^1, the composition $f \circ x : [0, \infty) \to \mathbb{R}$, $t \mapsto f(x(t))$ is again in C^1 and its derivative is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}(f \circ x)(t) = \frac{\mathrm{d}f}{\mathrm{d}x}\big(x(t)\big)\,\frac{\mathrm{d}x}{\mathrm{d}t}(t),$$

or more compactly

$$(f \circ x)^{\boldsymbol{\cdot}}\,(t) = f'\big(x(t)\big)\,\dot{x}(t),$$

where the dot $^{\boldsymbol{\cdot}}$ denotes the derivative with respect to t and the prime $'$ is the derivative with respect to x. In formal differential notation, we can rewrite this as

$$\mathrm{d}(f \circ x)(t) = f'\big(x(t)\big)\,\mathrm{d}x(t), \tag{1.1}$$