2

Integrators and Martingales

Now that the basic notions of filtration, process, and stopping time are at our disposal, it is time to develop the stochastic integral $\int X \, dZ$, as per Itô's ideas explained on page 5. We shall call X the ***integrand*** and Z the ***integrator***. Both are now processes.

For a guide let us review the construction of the ordinary Lebesgue–Stieltjes integral $\int x \, dz$ on the half-line; the stochastic integral $\int X \, dZ$ that we are aiming for is but a straightforward generalization of it. The Lebesgue–Stieltjes integral is constructed in two steps. First, it is defined on step functions x. This can be done whatever the integrator z. If, however, the Dominated Convergence Theorem is to hold, even on as small a class as the step functions themselves, restrictions must be placed on the integrator: z must be right-continuous and must have finite variation. This chapter discusses the stochastic analog of these restrictions, identifying the processes that have a chance of being useful stochastic integrators.

Given that a distribution function z on the line is right-continuous and has finite variation, the second step is one of a variety of procedures that extend the integral from step functions to a much larger class of integrands. The most efficient extension procedure is that of Daniell; it is also the only one that has a straightforward generalization to the stochastic case. This is discussed in chapter 3.

Step Functions and Lebesgue–Stieltjes Integrators on the Line

By way of motivation for this chapter let us go through the arguments in the second paragraph above in "abbreviated detail." A function $x : s \mapsto x_s$ on $[0, \infty)$ is a *step function* if there are a partition

$$\mathcal{P} = \{0 = t_1 < t_2 < \ldots < t_{N+1} < \infty\}$$

and constants $r_n \in \mathbb{R}$, $n = 0, 1, \ldots, N$, such that

$$x_s = \begin{cases} r_0 & \text{if} \quad s = 0 \\ r_n & \text{for} \quad t_n < s \le t_{n+1} \ , \ n = 1, 2, \ldots, N, \\ 0 & \text{for} \quad s > t_{N+1} . \end{cases} \tag{2.1}$$

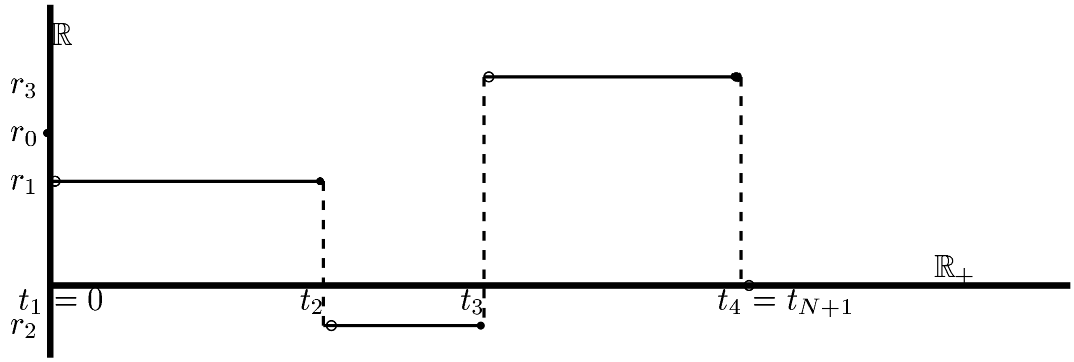

Figure 2.7 A step function on the half-line

The point $t = 0$ receives special treatment inasmuch as the measure $\mu = dz$ might charge the singleton $\{0\}$. The integral of such an *elementary integrand* x against a distribution function or *integrator* $z : [0, \infty) \to \mathbb{R}$ is

$$\int x\, dz = \int x_s\, dz_s \stackrel{\text{def}}{=} r_0 \cdot z_0 + \sum_{n=1}^{N} r_n \cdot \Big(z_{t_{n+1}} - z_{t_n} \Big) \ . \tag{2.2}$$

The collection $\mathfrak{e}$ of step functions is a vector space, and the map $x \mapsto \int x\, dz$ is a linear functional on it. It is called the *elementary integral*.

If z is just any function, nothing more of use can be said. We are after an extension satisfying the Dominated Convergence Theorem, though. If there is to be one, then z must be right-continuous; for if (t_n) is any sequence decreasing to t, then

$$z_{t_n} - z_t = \int 1_{(t,t_n]}\, dz \xrightarrow[n\to\infty]{} 0 \ ,$$

because the sequence $\big(1_{(t,t_n]}\big)$ of elementary integrands decreases pointwise to zero. Also, for every t the set

$$\int \mathfrak{e}_1\, dz^t \stackrel{\text{def}}{=} \left\{ \int x\, dz^t \ : \ x \in \mathfrak{e}\,,\ |x| \le 1 \right\}$$

must be bounded.[1] For if it were not, then there would exist elementary integrands $x^{(n)}$ with $|x^{(n)}| \le 1_{[0,t]}$ and $\int x^{(n)}\, dz > n$; the functions $x^{(n)}/n \in \mathfrak{e}$ would converge pointwise to zero, being dominated by $1_{[0,t]} \in \mathfrak{e}$, and yet their integrals would all exceed 1. The condition can be rewritten quantitatively

as $$\lceil z \rceil_t \stackrel{\text{def}}{=} \sup \left\{ \left| \int x\, dz \right| \ : \ |x| \le 1_{[0,t]} \right\} < \infty \quad \forall\, t < \infty \ , \tag{2.3}$$

or as $$\|y\|_z \stackrel{\text{def}}{=} \sup \left\{ \left| \int x\, dz \right| \ : \ |x| \le y \right\} < \infty \quad \forall\, y \in \mathfrak{e}_+ \ ,$$

[1] Recall from page 23 that z^t is z stopped at t.

or again thus: the image under $\int \cdot\, dz$ of any ***order interval***

$$[-y, y] \stackrel{\text{def}}{=} \{x \in \mathfrak{e} : -y \le x \le y\}$$

is a bounded subset of the range $\mathbb{R}$, $y \in \mathfrak{e}_+$. If (2.3) is satisfied, we say that z has *finite variation.*

In summary, if there is to exist an extension satisfying the Dominated Convergence Theorem, then z must be right-continuous and have finite variation. As is well known, these two conditions are also sufficient for the existence of such an extension.

The present chapter defines and analyzes the stochastic analogs of these notions and conditions; the elementary integrands are certain step functions on the half-line that depend on chance $\omega \in \Omega$; z is replaced by a process Z that plays the role of a "random distribution function"; and the conditions of right-continuity and finite variation have their straightforward analogs in the stochastic case. Discussing these and drawing first conclusions occupies the present chapter; the next one contains the extension theory via Daniell's procedure, which works just as simply and efficiently here as it does on the half-line.

Exercise 2.1 According to most textbooks, a distribution function $z : [0, \infty) \to \mathbb{R}$ has finite variation if for all $t < \infty$ the number

$$\text{¦}z\text{¦}_t = \sup\Big\{ |z_0| + \sum_i |z_{t_{i+1}} - z_{t_i}| : \ 0 = t_1 \le t_2 \le \ldots \le t_{I+1} = t \Big\} ,$$

called the variation of z on $[0, t]$, is finite. The supremum is taken over all finite partitions $0 = t_1 \le t_2 \le \ldots \le t_{I+1} = t$ of $[0, t]$. To reconcile this with the definition given above, observe that the sum is nothing but the integral of a step function, to wit, the function that takes the value $\operatorname{sgn}(z_0)$ on $\{0\}$ and $\operatorname{sgn}(z_{t_{i+1}} - z_{t_i})$ on the interval $(z_{t_i}, z_{t_{i+1}}]$. Show that

$$\text{¦}z\text{¦}_t = \sup\Big\{ \Big| \int x_s \, dz_s \Big| : |x| \le [0, t] \Big\} = \big\| [0, t] \big\|_z .$$

Exercise 2.2 The map $y \mapsto \|y\|_z$ is additive and extends to a positive measure on step functions. The latter is called the *variation measure* $\text{¦}\mu\text{¦} = d\text{¦}z\text{¦} = |dz|$ of $\mu = dz$. Suppose that z has finite variation. Then z is right-continuous if and only if $\mu = dz$ is σ-additive. If z is right-continuous, then so is $\text{¦}z\text{¦}$. $\text{¦}z\text{¦}$ is increasing and its limit at ∞ equals

$$\text{¦}z\text{¦}_\infty = \sup\Big\{ \Big| \int x_s \, dz_s \Big| : |x| \le 1 \Big\} .$$

If this number is finite, then z is said to have *bounded* or *totally finite variation.*

Exercise 2.3 A function on the half-line is a step function if and only if it is left-continuous, takes only finite many values, and vanishes after some instant. Their collection $\mathfrak{e}$ forms both an algebra and a vector lattice closed under chopping. The uniform closure of $\mathfrak{e}$ contains all continuous functions that vanish at infinity. The confined uniform closure of $\mathfrak{e}$ contains all continuous functions of compact support.

2.1 The Elementary Stochastic Integral

Elementary Stochastic Integrands

The first task is to identify the stochastic analog of the step functions in equation (2.1). The simplest thing coming to mind is this: a process X is an ***elementary stochastic integrand*** if there are a finite partition

$$\mathcal{P} = \{0 = t_0 = t_1 < t_2 \ldots < t_{N+1} < \infty\}$$

of the half-line and simple random variables $f_0 \in \mathcal{F}_0, f_n \in \mathcal{F}_{t_n}$, $n = 1, 2, \ldots, N$ such that

$$X_s(\omega) = \begin{cases} f_0(\omega) & \text{for} \quad s = 0 \\ f_n(\omega) & \text{for} \quad t_n < s \le t_{n+1}\ ,\ n = 1, 2, \ldots, N, \\ 0 & \text{for} \quad s > t_{N+1}\ . \end{cases}$$

In other words, for $t_n < s \le t \le t_{n+1}$, the random variables $X_s = X_t$ are simple and measurable on the σ-algebra $\mathcal{F}_{t_n}$ that goes with the left endpoint t_n of this interval. If we fix $\omega \in \Omega$ and consider the path $t \mapsto X_t(\omega)$, then we see an ordinary step function as in figure 2.7 on page 44. If we fix t and let ω vary, we see a simple random variable measurable on a σ-algebra strictly prior to t. Convention A.1.5 on page 364 produces this compact notation for X:

$$X = f_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f_n \cdot (\!(t_n, t_{n+1}]\!]\ . \tag{2.1.1}$$

The collection of elementary integrands will be denoted by $\mathcal{E}$, or by $\mathcal{E}[\mathcal{F}_{\bullet}]$ if we want to stress the fact that the notion depends – through the measurability assumption on the f_n – on the filtration.

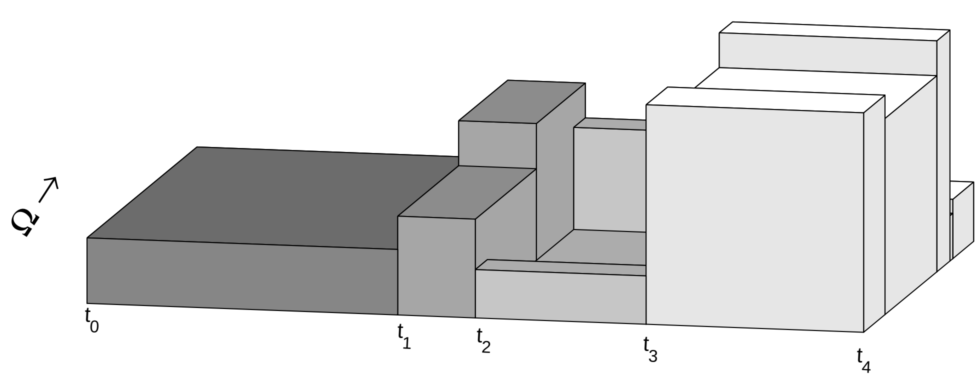

Figure 2.8 An elementary stochastic integrand

Exercise 2.1.1 An elementary integrand is an adapted left-continuous process.

Exercise 2.1.2 If X, Y are elementary integrands, then so are any linear combination, their product, their pointwise infimum $X \wedge Y$, their pointwise maximum $X \vee Y$, and the "chopped function" $X \wedge 1$. In other words, $\mathcal{E}$ is an algebra and vector lattice of bounded functions on $\boldsymbol{B}$ closed under chopping. (For the proof of proposition 3.3.2 it is worth noting that this is the sole information about $\mathcal{E}$ used in the extension theory of the next chapter.)

Exercise 2.1.3 Let $\mathcal{A}$ denote the collection of idempotent functions, i.e., sets,[2] in $\mathcal{E}$. Then $\mathcal{A}$ is a ring of subsets of $\boldsymbol{B}$ and $\mathcal{E}$ is the linear span of $\mathcal{A}$. $\mathcal{A}$ is the ring generated by the collection $\big\{\{0\} \times A :\ A \in \mathcal{F}_0\big\} \cup \big\{(s,t] \times A :\ s < t,\ A \in \mathcal{F}_s\big\}$ of rectangles, and $\mathcal{E}$ is the linear span of these rectangles.

The Elementary Stochastic Integral

Let Z be an adapted process. The integral against dZ of an elementary integrand $X \in \mathcal{E}$ as in (2.1.1) is, in complete analogy with the deterministic case (2.2), defined by

$$\int X\, dZ = f_0 \cdot Z_0 + \sum_{n=1}^{N} f_n \cdot (Z_{t_{n+1}} - Z_{t_n})\ . \tag{2.1.2}$$

This is a random variable: for $\omega \in \Omega$

$$\Big(\int X\, dZ\Big)(\omega) = f_0(\omega) \cdot Z_0(\omega) + \sum_{n=1}^{N} f_n(\omega) \cdot \big(Z_{t_{n+1}}(\omega) - Z_{t_n}(\omega)\big)\ .$$

However, although stochastic analysis is about dependence on chance ω, it is considered babyish to mention the ω; so mostly we shan't after this. The path of X is an ordinary step function as in (2.1). The present definition agrees ω-for-ω with the classical definition (2.2). The linear map $X \mapsto \int X\, dZ$ of (2.1.2) is called the ***elementary stochastic integral***.

Exercise 2.1.4 $\int X\, dZ$ does not depend on the representation (2.1.1) of X and is linear in both X and Z.

The Elementary Integral and Stopping Times

A description in terms of stopping times and stochastic intervals of both the elementary integrands and their integrals is natural and most useful. Let us call a stopping time ***elementary*** if it takes only finitely many values, all of them finite.

Let $S \le T$ be two elementary stopping times. The elementary stochastic interval $(\!(S, T]\!]$ is then an elementary integrand.[2] To see this let

$$\{0 \le t_1 < t_2 < \ldots < t_{N+1} < \infty\}$$

[2] See convention A.1.5 and figure A.14 on page 365.

be the values that S and T take, written in order. If $s \in (t_n, t_{n+1}]$, then the random variable $(\!(S,T]\!]_s$ takes only the values 0 or 1; in fact, $(\!(S,T]\!]_s(\omega) = 1$ precisely if $S(\omega) \le t_n$ and $T(\omega) \ge t_{n+1}$. In other words, for $t_n < s \le t_{n+1}$

$$\begin{aligned}(\!(S,T]\!]_s &= [S \le t_n] \cap [T \ge t_{n+1}] \\ &= [S \le t_n] \setminus [T \le t_n] \in \mathcal{F}_{t_n} \ ,\end{aligned}$$

so that $$(\!(S,T]\!] = \sum_{n=1}^{N} \ (t_n, t_{n+1}] \times \Big([S \le t_n] \cap [T \ge t_{n+1}] \Big) \ :$$

$(\!(S,T]\!]$ is a set in $\mathcal{E}$. Let us compute its integral against the integrator Z:

$$\begin{aligned}\int (\!(S,T]\!] \, dZ &= \sum_{n=1}^{N} ([S \le t_n][T \ge t_{n+1}])(Z_{t_{n+1}} - Z_{t_n}) \\ &= \sum_{1 \le m < n \le N+1} ([S = t_m][T = t_n])(Z_{t_n} - Z_{t_m}) \\ &= \sum_{1 \le m < n \le N+1} ([S = t_m][T = t_n])(Z_T - Z_S) \\ &= Z_T - Z_S \ . \end{aligned} \tag{2.1.3}$$

This is just as it should be.

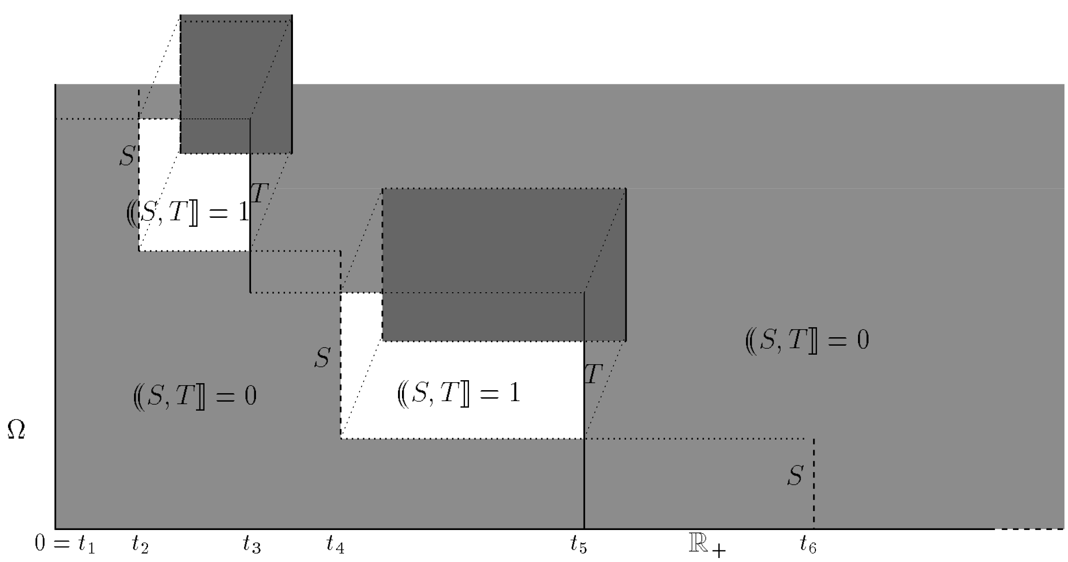

Figure 2.9 The indicator function of the stochastic interval $(\!(S,T]\!]$

Next let $A \in \mathcal{F}_0$. The stopping time 0 can be reduced by A to produce the stopping time 0_A (see exercise 1.3.18 on page 31). Its graph $[\![0_A]\!] = \{0\} \times A$ is evidently an elementary integrand with integral $A \cdot Z_0$. Finally, let $0 = T_1 < \ldots < T_{N+1}$ be elementary stopping times and $r_1, \ldots, r_N$ real numbers, and let f_0 be a simple random variable measurable on $\mathcal{F}_0$. Since f_0 can be written as $f_0 = \sum_k \rho_k \cdot A_k$, $A_k \in \mathcal{F}_0$, the process $f_0 \cdot [\![0]\!]$ is again an elementary integrand with integral $f_0 \cdot Z_0$. The linear combination

$$X = f_0 \cdot [\![0]\!] + \textstyle\sum_{n=1}^{N} r_n \cdot (\!(T_n, T_{n+1}]\!] \tag{2.1.4}$$

is then also an elementary integrand and its integral against dZ is

$$\int X \, dZ = f_0 \cdot Z_0 + \textstyle\sum_{n=1}^{N} r_n \cdot \left(Z_{T_{n+1}} - Z_{T_n}\right) .$$

Exercise 2.1.5 Let $0 = T_1 \le T_2 \le \ldots \le T_{N+1}$ be elementary stopping times and let $f_0 \in \mathcal{F}_0, f_1 \in \mathcal{F}_{T_1}, \ldots, f_N \in \mathcal{F}_{T_N}$ be simple functions. Then

$$X = f_0 \cdot [\![0]\!] + \textstyle\sum_{n=1}^{N} f_n \cdot (\!(T_n, T_{n+1}]\!]$$

is an elementary integrand, and its integral is

$$\int X \, dZ = f_0 \cdot Z_0 + \textstyle\sum_{n=1}^{N} f_n \cdot \left(Z_{T_{n+1}} - Z_{T_n}\right) .$$

Exercise 2.1.6 Every elementary integrand is of the form (2.1.4).

L^p-Integrators

Formula (2.1.2) associates with every elementary integrand $X : \boldsymbol{B} \to \mathbb{R}$ a random variable $\int X \, dZ$. The linear map $X \mapsto \int X \, dZ$ from $\mathcal{E}$ to L^0 is just like a signed measure, except that its values are random variables instead of numbers – the technical term is that the elementary integral defined by (2.1.2) is a ***vector measure***. Measures with values in topological vector spaces like L^p, $0 \le p < \infty$, turn out to have just as simple an extension theory as do measures with real values, provided they satisfy some simple conditions. Recall from the introduction to this chapter that a distribution function z on the half-line must be right-continuous, and its associated elementary integral must map order-bounded sets of step functions to bounded sets of reals, if there is to be a satisfactory extension.

Precisely this is required of our random distribution function Z, too:

Definition 2.1.7 (Integrators) *Let Z be a numerical process adapted to $\mathcal{F}_.$, $\mathbb{P}$ a probability on $\mathcal{F}_\infty$, and $0 \le p < \infty$.*

*(i) Let T be any stopping time, possibly $T = \infty$. We say that Z is **$\mathcal{I}^p$-bounded on the stochastic interval** $[\![0, T]\!]$ if the family of random variables*

$$\int \mathcal{E}_1 \, dZ^T = \left\{ \int X \, dZ^T : X \in \mathcal{E}, |X| \le 1 \right\}$$

$$\text{if } T \text{ is elementary:} \qquad = \left\{ \int X \, dZ : X \in \mathcal{E}, |X| \le [\![0, T]\!] \right\}$$

is a bounded subset of L^p.

(ii) Z is an ***L^p-integrator*** *if it satisfies the following two conditions:*

$$Z \text{ is right-continuous in probability;} \tag{RC-0}$$

$$Z \text{ is } \mathcal{I}^p\text{-bounded on every bounded interval } [\![0,t]\!]. \tag{B-p}$$

(B-p) simply says that the image under $\int \cdot\, dZ$ of any ***order interval***

$$[-Y,Y] \stackrel{\text{def}}{=} \{X \in \mathcal{E} : -Y \le X \le Y\}\,, \qquad Y \in \mathcal{E}_+\,,$$

is a bounded subset of the range L^p, or again that $\int \cdot\, dZ$ is continuous in the topology of confined uniform convergence (see item A.2.5 on page 370).

(iii) Z is a ***global L^p-integrator*** *if it is right-continuous in probability and $\mathcal{I}^p$-bounded on $[\![0,\infty)\!)$.*

If there is a need to specify the probability, then we talk about $\mathcal{I}^p[\mathbb{P}]$-boundedness and (global) $L^p(\mathbb{P})$-integrators.

The reader might have wondered why in (2.1.1) the values f_n that X takes on the interval $(t_n, t_{n+1}]$ were chosen to be measurable on the smallest possible σ-algebra, the one attached to the left endpoint t_n. The way the question is phrased points to the answer: had f_n been allowed to be measurable on the σ-algebra that goes with the right endpoint, or the midpoint, of that interval, then we would have ended up with a larger space $\mathcal{E}$ of elementary integrands. A process Z would have a harder time satisfying the boundedness condition (B-p), and the class of L^p-integrators would be smaller. We shall see soon (theorem 2.5.24) that it is precisely the choice made in equation (2.1.1) that permits martingales to be integrators.

The reader might also be intimidated by the parameter p. Why consider all exponents $0 \le p < \infty$ instead of picking one, say $p = 2$, to compute in Hilbert space, and be done with it? There are several reasons. First, a given integrator Z might not be an L^2-integrator but merely an L^1-integrator or an L^0-integrator. One could argue here that every integrator is an L^0-integrator, so that it would suffice to consider only these. In fact, L^0-integrators are very flexible (see proposition 2.1.9 and proposition 3.7.4); almost every reasonable process can be integrated in the sense L^0 (theorem 3.7.17); neither the feature of being an integrator nor the integral change when $\mathbb{P}$ is replaced by an equivalent measure (proposition 2.1.9 and proposition 3.6.20), which is of principal interest for statistical analysis; and finally L^0 is an algebra. On the other hand, the topological vector space L^0 is not locally convex, and the absence of a single homogeneous gauge measuring the size of its functions makes for cumbersome arguments – this problem can be overcome by replacing in a controlled way the given probability $\mathbb{P}$ by an equivalent one for which the driving term is an L^2-integrator or better – see theorem 4.1.2 on page 191. Second and more importantly, in the stability theory of stochastic differential equations Kolmogoroff's lemma A.2.37 will be used. The exponent p in inequality (A.2.4) will generally have to be strictly greater than the dimension of some parameter space (theorem 5.3.10) or of the state space (example 5.6.2).

The notion of an L^∞-integrator could be defined along the lines of definition 2.1.7, but this would be useless; there is no satisfactory extension theory for L^∞-valued vector measures. Replacing L^p with an Orlicz space whose defining Young function satisfies a so-called Δ_2-condition leads to a satisfactory integration theory, as does replacing it with a Lorentz space $L^{p,\infty}$, $p < \infty$. The most reasonable generalization is touched upon in exercise 3.6.19. We shall not pursue these possibilities.

Local Properties

A word about global *versus* "plain" L^p-integrators. The former are evidently the analogs of distribution functions with *totally finite* or *bounded* variation $\left\bracevert z \right\bracevert_\infty$, while the latter are the analogs of distribution functions z on $\mathbb{R}_+$ with just plain finite variation: $\left\bracevert z \right\bracevert_t < \infty \quad \forall\, t < \infty$. $\left\bracevert z \right\bracevert_t$ may well tend to ∞ as $t \to \infty$, as witness the distribution function $z_t = t$ of Lebesgue measure.

Note that a global integrator is defined in terms of the sup-norm on $\mathcal{E}$: the image of the ***unit ball***

$$\mathcal{E}_1 \stackrel{\text{def}}{=} \big\{X \in \mathcal{E} : \ |X| \le 1\big\} = \big\{X \in \mathcal{E} : \ -1 \le X \le 1\big\}$$

under the elementary integral must be a bounded subset of L^p. It is not good enough to consider only global integrators – a Wiener process, for instance, is not one. Yet it is frequently sufficient to prove a general result for them; given a "plain" integrator Z, the result in question will apply to every one of the stopped processes Z^t, $0 \le t < \infty$, these being evidently global L^p-integrators. In fact, in the stochastic case it is natural to consider an even more local notion:

Definition 2.1.8 *Let P be a property of processes – P might be the property of being a (global) L^p-integrator or of having continuous paths, for example. A stopping time T is said to **reduce** Z to a process having the property P if the stopped process Z^T has P.*

*The process Z is said to have the property P **locally** if there are **arbitrarily large stopping times** that reduce Z to processes having P, that is to say, if for every $\epsilon > 0$ and $t \in (0, \infty)$ there is a stopping time T with $\mathbb{P}[T < t] < \epsilon$ such that the stopped process Z^T has the property P.*

A local L^p-integrator is generally not an L^p-integrator. If $p = 0$, though, it is; this is a first indication of the flexibility of L^0-integrators. A second indication is the fact that being an L^0-integrator depends on the probability only up to local equivalence:

Proposition 2.1.9 *(i) A local L^0-integrator is an L^0-integrator; in fact, it is $\mathcal{I}^0$-bounded on every finite stochastic interval.*

(ii) If Z is a global $L^0(\mathbb{P})$-integrator, then it is a global $L^0(\mathbb{P}')$-integrator for any measure $\mathbb{P}'$ absolutely continuous with respect to $\mathbb{P}$.

(iii) Suppose that Z is an $L^0(\mathbb{P})$-integrator and $\mathbb{P}'$ is a probability on $\mathcal{F}_\infty$ locally absolutely continuous with respect to $\mathbb{P}$. Then Z is an $L^0(\mathbb{P}')$-integrator.

Proof. (i) To say that Z is a local $L^0(\mathbb{P})$-integrator means that, given an instant t and an $\epsilon > 0$, we can find a stopping time T with $\mathbb{P}[T \le t] < \epsilon$ such that the set of classes

$$\mathcal{B}^\epsilon \stackrel{\text{def}}{=} \left\{ \int X \, dZ^{T \wedge t} : X \in \mathcal{E} \,, \, |X| \le 1 \right\}$$

is bounded in $L^0(\mathbb{P})$. Every random variable $\int X \, dZ$ in the set

$$\mathcal{B} \stackrel{\text{def}}{=} \left\{ \int X \, dZ^t : \; X \in \mathcal{E} \,, \, |X| \le 1 \right\}$$

differs from the random variable $\int X \, dZ^{T \wedge t} \in \mathcal{B}^\epsilon$ only in the set $[T \le t]$. That is, the distance of these two random variables is less than ϵ if measured with $\lceil\!\lceil \; \rceil\!\rceil_0$ [3]. Thus $\mathcal{B} \subset L^0$ is a set with the property that for every $\epsilon > 0$ there exists a bounded set $\mathcal{B}^\epsilon \subset L^0$ with $\sup_{f \in \mathcal{B}} \inf_{f' \in \mathcal{B}'} \lceil\!\lceil f - f' \rceil\!\rceil_0 \le \epsilon$. Such a set is itself bounded in L^0. The second half of the statement follows from the observation that the instant t above can be replaced by an almost surely finite stopping time without damaging the argument. For the right-continuity in probability see exercise 2.1.11.

(iii) If the set $\{\int X \, dZ : \; X \in \mathcal{E}, |X| \le [\![0, t]\!]\}$ is bounded in $L^0(\mathbb{P})$, then it is bounded in $L^0(\mathcal{F}_t, \mathbb{P})$. Since the injection of $L^0(\mathcal{F}_t, \mathbb{P})$ into $L^0(\mathcal{F}_t, \mathbb{P}')$ is continuous (exercise A.8.19), this set is also bounded in the latter space. Since it is known that $t_n \downarrow t$ implies $Z_{t_n} \to Z_t$ in $L^0(\mathcal{F}_{t_1}, \mathbb{P})$, it also implies $Z_{t_n} \to Z_t$ in $L^0(\mathcal{F}_{t_1}, \mathbb{P}')$ and then in $L^0(\mathcal{F}_\infty, \mathbb{P}')$. (ii) is even simpler. ∎

Exercise 2.1.10 (i) If Z is an L^p-integrator, then for any stopping time T so is the stopped process Z^T. A local L^p-integrator is locally a global L^p-integrator.
(ii) If the stopped processes Z^S and Z^T are plain or global L^p-integrators, then so is the stopped process $Z^{S \vee T}$. If Z is a local L^p-integrator, then there is a sequence of stopping times reducing Z to global L^p-integrators and increasing a.s. to ∞.

Exercise 2.1.11 A locally nearly (almost surely) right-continuous process is nearly (respectively almost surely) right-continuous. An adapted process that has locally nearly finite variation nearly has finite variation.

Exercise 2.1.12 The sum of two (local, plain, global) L^p-integrators is a (local, plain, global) L^p-integrator. If Z is a (local, plain, global) L^q-integrator and $0 \le p \le q < \infty$, then Z is a (local, plain, global) L^p-integrator.

Exercise 2.1.13 Argue along the lines on page 43 that both conditions (B-p) and (RC-0) are necessary for the existence of an extension that satisfies the Dominated Convergence Theorem.

---

[3] The topology of L^0 is discussed briefly on page 33 ff., and in detail in section A.8.

Exercise 2.1.14 The map $X \mapsto \int X\, dZ$ is evidently a measure (that is to say a linear map on a space of functions) that has values in a vector space (L^p). Not every vector measure $\mathcal{I} : \mathcal{E} \to L^p$ is of the form $\mathcal{I}[X] = \int X\, dZ$. In fact, the stochastic integrals are exactly the vector measures $\mathcal{I} : \mathcal{E} \to L^0$ that satisfy $\mathcal{I}[f \cdot [\![0,0]\!]] \in \mathcal{F}_0$ for $f \in \mathcal{F}_0$ and

$$\mathcal{I}[f \cdot (\!(s,t]\!]] = f \cdot \mathcal{I}[(\!(s,t]\!]] \in \mathcal{F}_t$$

for $0 \le s \le t$ and simple functions $f \in L^\infty(\mathcal{F}_s)$.

2.2 The Semivariations

Numerical expressions for the boundedness condition (B-p) of definition 2.1.7 are desirable, in fact are necessary, to do the estimates we should expect, for instance, in Picard's scheme sketched on page 5. Now, the only difference with the classical situation discussed on page 44 is that the range $\mathbb{R}$ of the measure has been replaced by $L^p(\mathbb{P})$. It is tempting to emulate the definition (2.3) of the ordinary variation on page 44. To do that we have to agree on a substitute for the absolute value, which measures the size of elements of $\mathbb{R}$, by some device that measures the size of the elements of $L^p(\mathbb{P})$.

The obvious choice is one of the subadditive p-means, $0 \le p < \infty$, of equation (1.3.1) on page 33. With it the analog of inequality (2.3) becomes

$$\lceil\!\lceil Y \rceil\!\rceil_{Z\text{–}p} \stackrel{\text{def}}{=} \sup\left\{ \left\lceil\!\left\lceil \int X\, dZ \right\rceil\!\right\rceil_p : X \in \mathcal{E}\,,\ |X| \le Y \right\}, \tag{2.2.1}$$

The functional $\mathcal{E}_+ \ni Y \mapsto \lceil\!\lceil Y \rceil\!\rceil_{Z\text{–}p}$ is called ***the*** $\lceil\!\lceil\ \rceil\!\rceil_p$***-semivariation*** of Z. Recall our little mnemonic device: functionals with "straight sides" like $\|\ \|$ are homogeneous, and those with a little "crossbar" like $\lceil\!\lceil\ \rceil\!\rceil$ are subadditive. Of course, for $1 \le p < \infty$, $\|\ \|_p = \lceil\!\lceil\ \rceil\!\rceil_p$ is both; we then also write $\|Y\|_{Z\text{–}p}$ for $\lceil\!\lceil Y \rceil\!\rceil_{Z\text{–}p}$. In the case $p = 0$, the homogeneous gauges $\|\ \|_{[\alpha]}$ occasionally come in handy; the corresponding semivariation is

$$\|Y\|_{Z\text{–}[\alpha]} \stackrel{\text{def}}{=} \sup\left\{ \left\| \int X\, dZ \right\|_{[\alpha]} : X \in \mathcal{E},\ |X| \le Y \right\}, \qquad p = 0\,,\ \alpha \in \mathbb{R}\,.$$

If there is need to mention the measure $\mathbb{P}$, we shall write $\|\ \|_{Z\text{–}p;\mathbb{P}}$, $\lceil\!\lceil\ \rceil\!\rceil_{Z\text{–}p;\mathbb{P}}$, and $\|\ \|_{Z\text{–}[\alpha;\mathbb{P}]}$. It is clear that we could define a Z-semivariation for any other functional on measurable functions that strikes the fancy. We shall refrain from that.

In view of exercise A.8.18 on page 451 the boundedness condition (B-p) can be rewritten in terms of the semivariation as

$$\lceil\!\lceil \lambda \cdot Y \rceil\!\rceil_{Z\text{–}p} \xrightarrow[\lambda \to 0]{} 0 \qquad \forall\, Y \in \mathcal{E}_+\,. \tag{B-p}$$

When $0 < p < \infty$, this reads simply: $\lceil\!\lceil Y \rceil\!\rceil_{Z\text{–}p} < \infty \quad \forall\, Y \in \mathcal{E}_+$.

Proposition 2.2.1 *The semivariation* $\lceil\!\lceil\ \rceil\!\rceil_{Z-p}$ *is subadditive.*

Proof. Let $Y_1, Y_2 \in \mathcal{E}_+$ and let $r < \lceil\!\lceil Y_1 + Y_2 \rceil\!\rceil_{Z-p}$. There exists an integrand $X \in \mathcal{E}$ with $|X| \le Y_1 + Y_2$ and $r < \lceil\!\lceil \int X\, dZ \rceil\!\rceil_p$. Set $Y_1' \stackrel{\text{def}}{=} |X| \wedge Y_1$, $Y_2' \stackrel{\text{def}}{=} |X| - |X| \wedge Y_1 \le Y_2$, and

$$\begin{aligned} X_{1+} &\stackrel{\text{def}}{=} Y_1' \wedge X_+ \,, & X_{2+} &\stackrel{\text{def}}{=} X_+ - Y_1' \wedge X_+ \,, \\ X_{1-} &\stackrel{\text{def}}{=} Y_1' - Y_1' \wedge X_+ \,, & X_{2-} &\stackrel{\text{def}}{=} X_- + Y_1' \wedge X_+ - Y_1' \,. \end{aligned}$$

The columns of this matrix add up to Y_1' and Y_2', the rows to X_+ and X_-. The entries are positive elementary integrands. This is evident, except possibly for the positivity of X_{2-}. But on $[X_- = 0]$ we have $Y_1' = X_+ \wedge Y_1$ and with it $X_{2-} = 0$, and on $[X_+ = 0]$ we have instead $Y_1' = X_- \wedge Y_1$ and therefore $X_{2-} = X_- - X_- \wedge Y_1 \ge 0$. We estimate

$$\begin{aligned} r < \left\lceil\!\!\left\lceil \int X\, dZ \right\rceil\!\!\right\rceil_p &= \left\lceil\!\!\left\lceil \int (X_{1+} - X_{1-})\, dZ + \int (X_{2+} - X_{2-})\, dZ \right\rceil\!\!\right\rceil_p \\ &\le \left\lceil\!\!\left\lceil \int (X_{1+} - X_{1-})\, dZ \right\rceil\!\!\right\rceil_p + \left\lceil\!\!\left\lceil \int (X_{2+} - X_{2-})\, dZ \right\rceil\!\!\right\rceil_p \qquad (*) \\ &\le \lceil\!\lceil X_{1+} + X_{1-} \rceil\!\rceil_{Z-p} + \lceil\!\lceil X_{2+} + X_{2-} \rceil\!\rceil_{Z-p} \\ \text{as } Y_i' \le Y_i: \quad &= \lceil\!\lceil Y_1' \rceil\!\rceil_{Z-p} + \lceil\!\lceil Y_2' \rceil\!\rceil_{Z-p} \le \lceil\!\lceil Y_1 \rceil\!\rceil_{Z-p} + \lceil\!\lceil Y_2 \rceil\!\rceil_{Z-p} \,. \end{aligned}$$

The subadditivity of $\lceil\!\lceil\ \rceil\!\rceil_{Z-p}$ is established. Note that the subadditivity of $\lceil\!\lceil\ \rceil\!\rceil_p$ was used at $(*)$. ∎

At this stage the case $p = 0$ seems complicated, what with the boundedness condition (B-p) looking so clumsy. As the story unfolds we shall see that L^0-integrators are actually rather flexible and easy to handle. Proposition 2.1.9 gave a first indication of this; in theorem 3.7.17 it is shown in addition that every halfway decent process is integrable in the sense L^0 on every almost surely finite stochastic interval.

Exercise 2.2.2 The semivariations $\lceil\!\lceil\ \rceil\!\rceil_{Z-p}$, $\|\ \|_{Z-p}$, and $\|\ \|_{Z-[\alpha]}$ are *solid*; that is to say, $Y \le Y' \implies \|Y\|_{\bullet} \le \|Y'\|_{\bullet}$. The last two are absolute-homogeneous.

Exercise 2.2.3 Suppose that V is an adapted increasing process. Then for $X \in \mathcal{E}$ and $0 \le p < \infty$, $\lceil\!\lceil X \rceil\!\rceil_{V-p}$ equals the p-mean of the Lebesgue–Stieltjes integral $\int |X|\, dV$.

The Size of an Integrator

Saying that Z is a global L^p-integrator simply means that the elementary stochastic integral with respect to it is a continuous linear operator from one topological vector space, $\mathcal{E}$, to another, L^p; the size of such is customarily measured by its operator norm. In the case of the Lebesgue–Stieltjes integral this was the total variation $\|z\|_\infty$ (see exercise 2.2). By analogy we are led to set

$$\|Z\|_{\mathcal{I}^p} \stackrel{\text{def}}{=} \sup\left\{\left\|\int X\,dZ\right\|_p : X \in \mathcal{E},\ |X| \le 1\right\}, \qquad 0 < p < \infty\,,$$

$$\lceil Z \rceil_{\mathcal{I}^p} \stackrel{\text{def}}{=} \sup\left\{\left\lceil\!\left\lceil\int X\,dZ\right\rceil\!\right\rceil_p : X \in \mathcal{E},\ |X| \le 1\right\}, \qquad 0 \le p < \infty\,,$$

$$\|Z\|_{[\alpha]} \stackrel{\text{def}}{=} \sup\left\{\left\|\int X\,dZ\right\|_{[\alpha]} : X \in \mathcal{E},\ |X| \le 1\right\}, \qquad p = 0,\ \alpha \in \mathbb{R}\,,$$

depending on our current predilection for the size-measuring functional. If Z is merely an L^p-integrator, not a global one, then these numbers are generally infinite, and the quantities of interest are their finite-time versions

$$\lceil Z^t \rceil_{\mathcal{I}^p} \stackrel{\text{def}}{=} \sup\left\{\left\lceil\!\left\lceil\int X\,dZ\right\rceil\!\right\rceil_p : X \in \mathcal{E}\,,\ |X| \le [\![0,t]\!]\right\},$$

$0 \le t < \infty,\ 0 < p < \infty$, etc.

Exercise 2.2.4

(i)
$$\lceil Z + Z' \rceil_{\mathcal{I}^p} \le \lceil Z \rceil_{\mathcal{I}^p} + \lceil Z' \rceil_{\mathcal{I}^p}\,,$$
$$\lceil Z \rceil_{\mathcal{I}^p} = \|Z\|_{\mathcal{I}^p}^{p\wedge 1}\,, \quad \text{and} \quad \lceil \lambda Z \rceil_{\mathcal{I}^p} = |\lambda|^{1\wedge p} \lceil Z \rceil_{\mathcal{I}^p}$$

for $p > 0$. (ii) $\mathcal{I}^p$ forms a vector space on which $Z \mapsto \lceil Z \rceil_{\mathcal{I}^p}$ is subbadditive. (iii) If $0 < p < \infty$, then (B-p) is equivalent with $\lceil Z \rceil_{\mathcal{I}^p} < \infty$ or $\|Z\|_{\mathcal{I}^p} < \infty$. (iv) If $p = 0$, then (B-p) is equivalent with $\|Z\|_{[\alpha]} < \infty \quad \forall\,\alpha > 0$.

Exercise 2.2.5 If Z is an L^p-integrator and T is an elementary stopping time, then the stopped process Z^T is a global L^p-integrator and

$$\lceil \lambda \cdot Z^T \rceil_{\mathcal{I}^p} = \lceil\!\lceil \lambda \cdot [\![0,T]\!] \rceil\!\rceil_{Z\text{–}p} \qquad \forall\,\lambda \in \mathbb{R}\,.$$

Also,
$$\lceil\!\lceil [\![0,T]\!] \rceil\!\rceil_{Z\text{–}p} = \lceil Z^T \rceil_{\mathcal{I}^p}\,, \quad \|[\![0,T]\!]\|_{Z\text{–}p} = \|Z^T\|_{\mathcal{I}^p}\,,$$

and
$$\|[\![0,T]\!]\|_{Z\text{–}[\alpha]} = \|Z^T\|_{[\alpha]}\,.$$

Exercise 2.2.6 Let $0 \le p \le q < \infty$. An L^q-integrator is an L^p-integrator. Give an inequality between $\lceil\!\lceil Z \rceil\!\rceil_{\mathcal{I}^p}$ and $\lceil\!\lceil Z \rceil\!\rceil_{\mathcal{I}^q}$ and between $\|Z^T\|_{\mathcal{I}^p}$ and $\|Z^T\|_{\mathcal{I}^q}$ in case p is strictly positive.

Exercise 2.2.7 If Z is an L^p-integrator and $X \in \mathcal{E}$, then $Y_t \stackrel{\text{def}}{=} \int X\,dZ^t = \int_0^t X\,dZ$ defines a global L^p-integrator Y. For any $X' \in \mathcal{E}$,

$$\int X'\,dY = \int X'X\,dZ\,.$$

Exercise 2.2.8 $1/p \mapsto \log \|Z\|_{\mathcal{I}^p}$ is convex for $0 < p < \infty$.

Vectors of Integrators

A stochastic differential equation frequently is driven by not one or two but a whole slew $\boldsymbol{Z} = (Z^1, Z^2, \ldots, Z^d)$ of integrators, even an infinity of them.[4] It eases the notation to set,[5] for $\boldsymbol{X} = (X_1, \ldots, X_d) \in \mathcal{E}^d$,

$$\int \boldsymbol{X}\, d\boldsymbol{Z} \stackrel{\text{def}}{=} \int X_\eta \, dZ^\eta \tag{2.2.2}$$

and to define the integrator size of the d-tuple $\boldsymbol{Z}$ by

$$\lceil\!\lceil \boldsymbol{Z} \rceil\!\rceil_{\mathcal{I}^p} = \sup \left\{ \left\| \int \boldsymbol{X}\, d\boldsymbol{Z} \right\|_{L^p} : \ \boldsymbol{X} \in \mathcal{E}^d_1 \right\}, \qquad p > 0\ ;$$

$$\lceil\!\lceil \boldsymbol{Z} \rceil\!\rceil_{[\alpha]} = \sup \left\{ \left\| \int \boldsymbol{X}\, d\boldsymbol{Z} \right\|_{[\alpha]} : \ \boldsymbol{X} \in \mathcal{E}^d_1 \right\}, \qquad p = 0, \alpha > 0\ ,$$

and so on. These definitions take advantage of possible cancellations among the Z^η. For instance, if $\boldsymbol{W} = (W^1, W^2, \ldots, W^d)$ are independent standard Wiener processes stopped at the instant t, then $\lceil\!\lceil \boldsymbol{W} \rceil\!\rceil_{\mathcal{I}^2}$ equals $\sqrt{d \cdot t}$ rather than the first-impulse estimate $d\sqrt{t}$. Lest the gentle reader think us too nitpicking, let us point out that this definition of the integrator size is instrumental in establishing previsible control of random measures in theorem 4.5.25 on page 251, control which in turn greatly facilitates the solution of differential equations driven by random measures (page 296).

Definition 2.2.9 *A vector $\boldsymbol{Z}$ of adapted processes is an L^p-integrator if its components are right-continuous in probability and its $\mathcal{I}^p$-size $\lceil\!\lceil \boldsymbol{Z}^t \rceil\!\rceil_{\mathcal{I}^p}$ is finite for all $t < \infty$.*

Exercise 2.2.10 $\mathcal{E}^d$ is a self-confined algebra and vector lattice closed under chopping of bounded functions, and the vector $\boldsymbol{Z}$ of càdlàg adapted processes is an L^p-integrator if and only if the map $\boldsymbol{X} \mapsto \int \boldsymbol{X}\, d\boldsymbol{Z}$ is continuous from $\mathcal{E}^d$ equipped with the topology of confined uniform convergence (see item A.2.5) to L^p.

The Natural Conditions

The notion of an L^p-integrator depends on the filtration. If Z is an L^p-integrator with respect to the given filtration $\mathcal{F}_.$ and we change every $\mathcal{F}_t$ to a larger σ-algebra $\mathcal{G}_t$, then Z will still be adapted and right-continuous in probability – these features do not mention the filtration. But doing so will generally increase the supply of elementary integrands, so that now Z

[4] See equation (1.1.9) on page 8 or equation (5.1.3) on page 271 and section 3.10 on page 171.

[5] We shall use the ***Einstein convention*** throughout: summation over repeated indices *in opposite positions* (the η in (2.2.2)) is implied.

has a harder time satisfying the boundedness condition (B-p). Namely, since $\mathcal{E}[\mathcal{F}_{\bullet}] \subset \mathcal{E}[\mathcal{G}_{\bullet}]$,

the collection $$\left\{\int X\, dZ :\ X \in \mathcal{E}[\mathcal{G}_{\bullet}],\ |X| \le 1\right\}$$

is larger than $$\left\{\int X\, dZ :\ X \in \mathcal{E}[\mathcal{F}_{\bullet}],\ |X| \le 1\right\} ;$$

and while the latter is bounded in L^p, the former need not be. However, a slight enlargement is innocuous:

Proposition 2.2.11 *Suppose that Z is an $L^p(\mathbb{P})$-integrator on $\mathcal{F}_{\bullet}$ for some $p \in [0,\infty)$. Then Z is an $L^p(\mathbb{P})$-integrator on the natural enlargement $\mathcal{F}^{\mathbb{P}}_{\bullet+}$, and the sizes $\lceil Z^t \rceil_{\mathcal{I}^p}$ computed on $\mathcal{F}^{\mathbb{P}}_{\bullet+}$ are at most twice what they are computed on $\mathcal{F}_{\bullet}$ – if $Z_0 = 0$, they are the same.*

Proof. Let $\mathcal{E}^{\mathbb{P}} = \mathcal{E}[\mathcal{F}^{\mathbb{P}}_{\bullet+}]$ denote the elementary integrands for the natural enlargement and set

$$\mathcal{B} \stackrel{\text{def}}{=} \left\{\int X\, dZ^t :\ X \in \mathcal{E}_1\right\} \quad \text{and} \quad \mathcal{B}^{\mathbb{P}} \stackrel{\text{def}}{=} \left\{\int X\, dZ^t :\ X \in \mathcal{E}^{\mathbb{P}}_1\right\} .$$

$\mathcal{B}$ is a bounded subset of L^p, and so is its “solid closure”

$$\mathcal{B}^{\diamond} \stackrel{\text{def}}{=} \{f \in L^p :\ |f| \le |g| \ \text{ for some } \ g \in \mathcal{B}\} .$$

We shall show that $\mathcal{B}^{\mathbb{P}}$ is contained in $\overline{\mathcal{B}^{\diamond}} + \overline{\mathcal{B}^{\diamond}}$, where $\overline{\mathcal{B}^{\diamond}}$ is the closure of $\mathcal{B}^{\diamond}$ in the topology of convergence in measure; the claim is then immediate from this consequence of solidity and Fatou's lemma A.8.7:

$$\sup\left\{\lceil\!\lceil f \rceil\!\rceil_p :\ f \in \overline{\mathcal{B}^{\diamond}} + \overline{\mathcal{B}^{\diamond}}\right\} \le 2 \sup\left\{\lceil\!\lceil f \rceil\!\rceil_p :\ f \in \mathcal{B}\right\} .$$

Let then $X \in \mathcal{E}^{\mathbb{P}}_1$, writing it as in equation (2.1.1):

$$X = f_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f_n \cdot (\!(t_n, t_{n+1}]\!] , \qquad f_n \in \mathcal{F}^{\mathbb{P}}_{t_n+} .$$

For every $n \in \mathbb{N}$ there is a simple random variable $f'_n \in \mathcal{F}_{t_n+}$ that differs negligibly from f_n. Let k be so large that $t_n + 1/k < t_{n+1}$ for all n and set

$$X^{(k)} \stackrel{\text{def}}{=} f'_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f'_n \cdot (\!(t_n + 1/k, t_{n+1}]\!] , \qquad k \in \mathbb{N} .$$

The sum on the right clearly belongs to $\mathcal{E}_1$, so its stochastic integral

$$f'_0 \cdot Z_0 + \sum f'_n \cdot \left(Z_{t_{n+1}} - Z_{t_n+1/k}\right)$$

belongs to $\mathcal{B}$. The first random variable $f'_0 Z_0$ is majorized in absolute value by $|Z_0| = \left|\int [\![0]\!]\, dZ\right|$ and thus belongs to $\mathcal{B}^{\diamond}$. Therefore $\int X^{(k)}\, dZ$ lies in the sum $\mathcal{B}^\diamond + \mathcal{B} \subset \overline{\mathcal{B}^\diamond} + \overline{\mathcal{B}^\diamond}$. As $k \to \infty$ these stochastic integrals on $\mathcal{E}$ converge in probability to

$$f_0' \cdot Z_0 + \sum f_n' \cdot \left(Z_{t_{n+1}} - Z_{t_n}\right) = \int X \, dZ \ ,$$

which therefore belongs to $\overline{\mathcal{B}^\diamond} + \overline{\mathcal{B}^\diamond}$. ▮

Recall from exercise 1.3.30 that a regular and right-continuous filtration has more stopping times than just a plain filtration. We shall therefore make our life easy and replace the given measured filtration by its natural enlargement:

Assumption 2.2.12 ***The given measured filtration*** $(\mathcal{F}_{\bullet}, \mathbb{P})$ ***is henceforth assumed to be both right-continuous and regular.***

Exercise 2.2.13 On Wiener space $(\mathscr{C}, \mathcal{B}^\bullet(\mathscr{C}), \mathbb{W})$ consider the canonical Wiener process $w_{\bullet}$ (w_t takes a path $w \in \mathscr{C}$ to its value at t). The $\mathbb{W}$-regularization of the basic filtration $\mathcal{F}^0_{\bullet}[w]$ is right-continuous (see exercise 1.3.47): it is the natural filtration $\mathcal{F}_{\bullet}[w]$ of w. Then the triple $(\mathscr{C}, \mathcal{F}_{\bullet}[w], \mathbb{W})$ is an instance of a measured filtration that is right-continuous and regular. $w : (t, w) \mapsto w_t$ is a continuous process, adapted to $\mathcal{F}_{\bullet}[w]$ and p-integrable for all $p \ge 0$, but not L^p-bounded for any $p \ge 0$.

Exercise 2.2.14 Let $\mathcal{A}'$ denote the ring on $\boldsymbol{B}$ generated by $\{[\![0_A]\!] : A \in \mathcal{F}_0\}$ and the collection $\{(\!(S, T]\!] : \ S, T$ bounded stopping times$\}$ of stochastic intervals, and let $\mathcal{E}'$ denote the step functions over $\mathcal{A}'$. Clearly $\mathcal{A} \subset \mathcal{A}'$ and $\mathcal{E} \subset \mathcal{E}'$. Every $X \in \mathcal{E}'$ can be written in the form

$$X = f_0 \cdot [\![0]\!] + \textstyle\sum_{n=0}^N f_n \cdot (\!(T_n, T_{n+1}]\!] \ ,$$

where $0 = T_0 \le T_1 \le \ldots \le T_{N+1}$ are bounded stopping times and $f_n \in \mathcal{F}_{T_n}$ are simple. If Z is a global L^p-integrator, then the definition

$$\int X \, dZ \stackrel{\text{def}}{=} f_0 \cdot Z_0 + \textstyle\sum_n f_n \cdot (Z_{T_{n+1}} - Z_{T_n}) \tag{$*$}$$

provides an extension of the elementary integral that has the same modulus of continuity. Any extension of the elementary integral that satisfies the Dominated Convergence Theorem must have a domain containing $\mathcal{E}'$ and coincide there with $(*)$.

2.3 Path Regularity of Integrators

Suppose Z, Z' are modifications of each other, that is to say, $Z_t = Z'_t$ almost surely at every instant t. An inspection of (2.1.2) then shows that for every elementary integrand X the random variables $\int X \, dZ$ and $\int X \, dZ'$ nearly coincide: as integrators, Z and Z' are the same and should be identified. It is shown in this section that from all of the modifications one can be chosen that has rather regular paths, namely, càdlàg ones.

Right-Continuity and Left Limits

Lemma 2.3.1 *Suppose Z is a process adapted to $\mathcal{F}_{\bullet}$ that is $\mathcal{I}^0[\mathbb{P}]$-bounded on bounded intervals. Then the paths whose restrictions to the positive rationals have an oscillatory discontinuity occur in a $\mathbb{P}$-nearly empty set.*

Proof. Fix two rationals a, b with $a < b$, an instant $u < \infty$, and a finite set $\mathcal{S} = \{s_0 < s_1 < \ldots < s_N\}$ of rationals in $[0, u)$. Next set $T_0 \stackrel{\text{def}}{=} \min\{s \in \mathcal{S} : Z_s < a\} \wedge u$ and continue by induction:

$$T_{2k+1} = \inf\{s \in \mathcal{S} : \ s > T_{2k} \, , \ Z_s > b\} \wedge u$$
$$T_{2k} = \inf\{s \in \mathcal{S} : \ s > T_{2k-1} \, , \ Z_s < a\} \wedge u \, .$$

It was shown in proposition 1.3.13 that these are stopping times, evidently elementary. ($T_n(\omega)$ will be equal to u for some index $n(\omega)$ and all higher ones, but let that not bother us.) Let us now estimate the number $U_{\mathcal{S}}^{[a,b]}$ of upcrossings of the interval $[a, b]$ that the path of Z performs on $\mathcal{S}$. (We say that $\mathcal{S} \ni s \mapsto Z_s(\omega)$ *upcrosses* the interval $[a, b]$ on $\mathcal{S}$ if there are points $s < t$ in $\mathcal{S}$ with $Z_s(\omega) < a$ and $Z_t(\omega) > b$. To say that this path has n ***upcrossings*** means that there are n pairs: $s_1 < t_1 < s_2 < t_2 < \ldots < s_n < t_n$ in $\mathcal{S}$ with $Z_{s_\nu} < a$ and $Z_{t_\nu} > b$.) If $\mathcal{S} \ni s \mapsto Z_s(\omega)$ upcrosses the interval $[a, b]$ n times or more on $\mathcal{S}$, then $T_{2n-1}(\omega)$ is strictly less than u, and vice versa:

$$\Big[U_{\mathcal{S}}^{[a,b]} \geq n\Big] = [T_{2n-1} < u] \in \mathcal{F}_u \, . \tag{2.3.1}$$

This observation produces the inequality[2]

$$\Big[U_{\mathcal{S}}^{[a,b]} \geq n\Big] \leq \frac{1}{n(b-a)} \left(\sum_{k=0}^{\infty} (Z_{T_{2k+1}} - Z_{T_{2k}}) \ + \ |Z_u - a| \right) , \tag{2.3.2}$$

for if $U_{\mathcal{S}}^{[a,b]} \geq n$, then the (finite!) sum on the right contributes more than n times a number greater than $b - a$. The last term of the sum might be negative, however. This occurs when $T_{2k}(\omega) < s_N$ and thus $Z_{T_{2k}}(\omega) < a$, and $T_{2k+1}(\omega) = u$ because there is no more $s \in \mathcal{S}$ exceeding $T_{2k}(\omega)$ with $Z_s(\omega) > b$. The last term of the sum is then $Z_u(\omega) - Z_{T_{2k}}(\omega)$. This number might well be negative. However, it will not be less than $Z_u(\omega) - a$: the last term $\big|Z_u(\omega) - a\big|$ of (2.3.2) added to the last non-zero term of the sum will always be positive.

The stochastic intervals $(\!(T_{2k}, T_{2k+1}]\!]$ are elementary integrands, and their integrals are $Z_{T_{2k+1}} - Z_{T_{2k}}$. This observation permits us to rewrite (2.3.2) as

$$\Big[U_{\mathcal{S}}^{[a,b]} \geq n\Big] \leq \frac{1}{n(b-a)} \left(\int \sum_{k=0}^{\infty} (\!(T_{2k}, T_{2k+1}]\!] \, dZ^u \ + \ |Z_u - a| \right) . \tag{2.3.3}$$

This inequality holds for any adapted process Z. To continue the estimate observe now that the integrand $\sum_{k=1}^{\infty} (\!(T_{2k}, T_{2k+1}]\!]$ is majorized in absolute value by 1. Measuring both sides of (2.3.3) with $\lceil\!\lceil \ \rceil\!\rceil_0$ yields the inequality

$$\begin{aligned} \mathbb{P}\Big[U_{\mathcal{S}}^{[a,b]} \geq n\Big] &= \Big\lceil\!\Big\lceil \Big[U_{\mathcal{S}}^{[a,b]} \geq n\Big] \Big\rceil\!\Big\rceil_{L^0(\mathbb{P})} \\ &\leq \lceil\!\lceil 1/n(b-a) \rceil\!\rceil_{Z^u\text{–}0;\mathbb{P}} + \Big\lceil\!\Big\lceil \big(a - Z_u\big)/n(b-a) \Big\rceil\!\Big\rceil_{L^0(\mathbb{P})} \\ &\leq 2 \cdot \lceil\!\lceil 1/n(b-a) \rceil\!\rceil_{Z^u\text{–}0;\mathbb{P}} + |a|/n(b-a) \, . \end{aligned}$$

Now let $\mathbb{Q}_+^{u-}$ denote the set of positive rationals less than u. The right-hand side of the previous inequality does not depend on $\mathcal{S} \subset \mathbb{Q}_+^{u-}$. Taking the supremum over all finite subsets $\mathcal{S}$ of $\mathbb{Q}_+^{u-}$ results, in obvious notation, in the inequality

$$\mathbb{P}\Big[U_{\mathbb{Q}_+^{u-}}^{[a,b]} \geq n\Big] \leq 2 \cdot \lceil\!\lceil 1/n(b-a) \rceil\!\rceil_{Z^u\text{–}0;\mathbb{P}} + a/n(b-a) \ .$$

Note that the set on the left belongs to $\mathcal{F}_u$ (equation (2.3.1)). Since Z is assumed $\mathcal{I}^0$-bounded on $[0,u]$, taking the limit as $n \to \infty$ gives

$$\mathbb{P}\Big[U_{\mathbb{Q}_+^{u-}}^{[a,b]} = \infty\Big] = 0 \ .$$

That is to say, the restriction to $\mathbb{Q}_+^{u-}$ of nearly no path upcrosses the interval $[a,b]$ infinitely often. The set

$$Osc = \bigcup_{u \in \mathbb{N}} \ \bigcup_{a,b \in \mathbb{Q},\, a<b} \Big[U_{\mathbb{Q}_+^{u}}^{[a,b]} = \infty\Big]$$

therefore belongs to $\mathcal{A}_{\infty\sigma}$ and is $\mathbb{P}$-negligible: it is a $\mathbb{P}$-nearly empty set. If ω is in its complement, then the path $t \mapsto Z_t(\omega)$ restricted to the rationals has no oscillatory discontinuity. ∎

The upcrossing argument above is due to Doob, who used it to show the regularity of martingales (see proposition 2.5.13).

Let Ω_0 be the complement of Osc. By our standing regularity assumption 2.2.12 on the filtration, Ω_0 belongs to $\mathcal{F}_0$. The path $Z_{\boldsymbol{.}}(\omega)$ of the process Z of lemma 2.3.1 has for every $\omega \in \Omega_0$ left and right limits through the rationals at any time t. We may define

$$Z'_t(\omega) = \begin{cases} \lim\limits_{\mathbb{Q} \ni q \downarrow t} Z_q(\omega) & \text{for } \omega \in \Omega_0, \\ \qquad 0 & \text{for } \omega \in Osc. \end{cases}$$

The limit is understood in the extended reals $\overline{\mathbb{R}}$, since nothing said so far prevents it from being $\pm\infty$. The process Z' is right-continuous and has left limits at any finite instant.

Assume now that Z is an L^0-integrator. Since then Z is right-continuous in probability, Z and Z' are modifications of one another. Indeed, for fixed t let q_n be a sequence of rationals decreasing to t. Then $Z_t = \lim_n Z_{q_n}$ in measure, in fact nearly, since $\big[Z_t \neq \lim_n Z_{q_n}\big] \in \mathcal{F}_{q_1}$. On the other hand, $Z'_t = \lim_n Z_{q_n}$ nearly, by definition. Thus $Z_t = Z'_t$ nearly for all $t < \infty$. Since the filtration satisfies the natural conditions, Z' is adapted. Any other right-continuous modification of Z is indistinguishable from Z' (exercise 1.3.28).

Note that these arguments apply to any probability with respect to which Z is an L^p-integrator: Osc is nearly empty for every one of them. Let $\mathfrak{P}[Z]$

denote their collection. The version Z' that we found is thus "universally regular" in the sense that it is adapted to the "small enlargement"

$$\mathcal{F}_{\bullet+}^{\mathfrak{P}[Z]} \stackrel{\text{def}}{=} \bigcap \big\{ \mathcal{F}_{\bullet+}^{\mathbb{P}} : \ \mathbb{P} \in \mathfrak{P}[Z] \big\} \ .$$

Denote by $\mathfrak{P}_0[Z]$ the class of probabilities under which Z is actually a global $L^0(\mathbb{P})$-integrator. If $\mathbb{P} \in \mathfrak{P}_0[Z]$, then we may take ∞ for the time u of the proof and see that the paths of Z that have an oscillatory discontinuity anywhere, including at ∞, are negligible. In other words, then $Z'_\infty \stackrel{\text{def}}{=} \lim_{t\uparrow\infty} Z'_t$ exists, except possibly in a set that is $\mathbb{P}$-negligible simultaneously for all $\mathbb{P} \in \mathfrak{P}_0[Z]$.

Boundedness of the Paths

For the remainder of the section we shall assume that a modification of the L^0-integrator Z has been chosen that is right-continuous and has left limits at all finite times and is adapted to $\mathcal{F}_{\bullet+}^{\mathfrak{P}[Z]}$. So far it is still possible that this modification takes the values $\pm\infty$ frequently. The following ***maximal inequality*** of weak type rules out this contingency, however:

Lemma 2.3.2 *Let T be any stopping time and $\lambda > 0$. The maximal process $Z^\star$ of Z satisfies, for every $\mathbb{P} \in \mathfrak{P}[Z]$,*

$$\mathbb{P}[Z_T^\star \geq \lambda] \leq \lceil Z^T/\lambda \rceil_{\mathcal{I}^0[\mathbb{P}]} \ , \qquad p = 0;$$

$$\|Z_T^\star\|_{[\alpha]} \leq \vert Z^T \vert_{[\alpha;\mathbb{P}]} \ , \qquad p = 0 \, , \ \alpha \in \mathbb{R};$$

$$\mathbb{P}[Z_T^\star \geq \lambda] \leq \lambda^{-p} \cdot \vert Z^T \vert_{\mathcal{I}^p[\mathbb{P}]}^p \ , \qquad 0 < p < \infty.$$

Proof. We resurrect our finite set $\mathcal{S} = \{s_0 < s_1 < \ldots < s_N\}$ of positive rationals strictly less than u and define

$$U = \inf\{s \in \mathcal{S} : \ |Z_s^T| > \lambda\} \wedge u \ .$$

This is an elementary stopping time (proposition 1.3.13). Now

$$\Big[\sup_{s\in\mathcal{S}} |Z_s^T| > \lambda\Big] = [U < u] \in \mathcal{F}_u \ ,$$

on which set $$|Z_U^T| = \Big|\int [\![0, U]\!] \, dZ^T\Big| > \lambda \ .$$

Applying $\lceil\!\lceil \ \rceil\!\rceil_p$ to the resulting inequality

$$\Big[\sup_{s\in\mathcal{S}} |Z_s^T| > \lambda\Big] \leq \lambda^{-1} \Big| \int [\![0, U]\!] \, dZ^T \Big|$$

gives $$\Big\lceil\!\!\Big\lceil \Big[\sup_{s\in\mathcal{S}} |Z_s^T| > \lambda\Big] \Big\rceil\!\!\Big\rceil_p \leq \Big\lceil\!\!\Big\lceil \lambda^{-1} \int [\![0, U]\!] \, dZ^T \Big\rceil\!\!\Big\rceil_p \leq \lceil \lambda^{-1} Z^T \rceil_{\mathcal{I}^p} \ .$$

We observe that the ultimate right does not depend on $\mathcal{S} \subset \mathbb{Q}_+ \cap [0,u)$. Taking the supremum over $\mathcal{S} \subset \mathbb{Q}_+ \cap [0,u)$ therefore gives

$$\left\|\!\left| \left[\sup_{s<u} |Z^T_s| > \lambda\right] \right|\!\right\|_p = \left\|\!\left| \left[\sup_{s\in\mathbb{Q}_+, s<u} |Z^T_s| > \lambda\right] \right|\!\right\|_p \le \lceil \lambda^{-1} Z^T \rceil_{\mathcal{I}^p} . \tag{2.3.4}$$

Letting $u \to \infty$ yields the stated inequalities (see exercise A.8.3). ∎

Exercise 2.3.3 Let $\boldsymbol{Z} = (Z^1, \ldots, Z^d)$ be a vector of L^0-integrators. The maximal process of its euclidean length

$$|\boldsymbol{Z}|_t \stackrel{\mathrm{def}}{=} \Bigl(\sum_{1\le\eta\le d} (Z^\eta_t)^2\Bigr)^{1/2} \quad \text{satisfies} \quad \bigl\|\, |\boldsymbol{Z}|^\star_t \bigr\|_{[\alpha]} \le K_0^{(\mathrm{A.8.6})} \cdot \lceil \boldsymbol{Z} \rceil_{[\alpha\kappa_0]} , \quad 0 < \alpha < 1 .$$

(See theorem 2.3.6 on page 63 for the case $p > 0$ and a hint.)[6]

Redefinition of Integrators

Note that the set $\left[Z^{\star T}_u > \lambda\right]$ on the left in inequality (2.3.4) belongs to $\mathcal{F}_u \in \mathcal{A}_\infty$. Therefore

$$N \stackrel{\mathrm{def}}{=} \left[Z^\star_T = \infty\right] \cap [T < \infty] = \bigcup_{u\in\mathbb{N}} \left[Z^{\star T}_u = \infty\right]$$

is a $\mathbb{P}$-negligible set of $\mathcal{A}_{\infty\sigma}$; it is $\mathbb{P}$-nearly empty. This is true for all $\mathbb{P} \in \mathfrak{P}[Z]$. We now alter Z by setting it equal to zero on N. Since $\mathcal{F}_{\bullet}$ is assumed to be right-continuous and regular, we obtain an adapted right-continuous modification of Z whose paths are real-valued, in fact bounded on bounded intervals. The upshot:

Theorem 2.3.4 *Every L^0-integrator Z has a modification all of whose paths are right-continuous, have left limits at every finite instant, and are bounded on every finite interval. Any two such modifications are indistinguishable. Furthermore, this modification can be chosen adapted to $\mathcal{F}^{\mathfrak{P}[Z]}_{\bullet+}$. Its limit at infinity exists and is $\mathbb{P}$-almost surely finite for all $\mathbb{P}$ under which Z is a global L^0-integrator.*

Convention 2.3.5 *Whenever an L^p-integrator Z on a regular right-continuous filtration appears it will henceforth be understood that a right-continuous real-valued modification with left limits has been chosen, adapted to $\mathcal{F}^{\mathfrak{P}[Z]}_{\bullet+}$ as it can be.*

Since a local L^p-integrator Z is an L^0-integrator (proposition 2.1.9), it is also understood to have càdlàg paths and to be adapted to $\mathcal{F}^{\mathfrak{P}[Z]}_{\bullet+}$.

In remark 3.8.5 we shall meet a further regularity property of the paths of an integrator Z; namely, while the sums $\sum_k |Z_{T_k} - Z_{T_{k-1}}|$ may diverge as the random partition $\{0 = T_1 \le T_2 \le \ldots \le T_K = t\}$ of $[\![0,t]\!]$ is refined, the sums $\sum_k |Z_{T_k} - Z_{T_{k-1}}|^2$ of squares stay bounded, even converge.

---

[6] The superscript (A.8.6) on K_0 means that this is the constant K_0 of inequality (A.8.6).

The Maximal Inequality

The last "weak type" inequality in lemma 2.3.2 can be replaced by one of "strong type," which holds even for a whole vector

$$\boldsymbol{Z} = (Z^1, \ldots, Z^d)$$

of L^p-integrators and extends the result of exercise 2.3.3 for $p = 0$ to strictly positive p. The ***maximal process*** of $\boldsymbol{Z}$ is the d-tuple of increasing processes

$$\boldsymbol{Z}^\star = (Z^{\eta\star})_{\eta=1}^d \stackrel{\text{def}}{=} \big(|Z^\eta|^\star\big)_{\eta=1}^d .$$

Theorem 2.3.6 *Let $0 < p < \infty$ and let $\boldsymbol{Z}$ be an L^p-integrator. The euclidean length $|\boldsymbol{Z}^\star|$ of its maximal process satisfies $|\boldsymbol{Z}|^\star \le |\boldsymbol{Z}^\star|$ and*

$$\big\| |\boldsymbol{Z}^\star|_t \big\|_{L^p} = \big\lceil |\boldsymbol{Z}^\star|^t \big\rceil_{\mathcal{I}^p} \le C_p^\star \cdot \big\lceil \boldsymbol{Z}^t \big\rceil_{\mathcal{I}^p} , \tag{2.3.5}$$

with universal constant[6]
$$C_p^\star \le \frac{10}{3} \cdot K_p^{(\text{A.8.5})} \le 3.35 \cdot 2^{\frac{2-p}{2p} \vee 0} . \tag{2.3.6}$$

Proof. Let $\mathcal{S} = \{0 = s_0 < s_1 < \ldots < t\}$ be a finite partition of $[0,t]$ and pick a $q > 1$. For $\eta = 1 \ldots d$, set $T_0^\eta = -1$, $Z_{-1}^\eta = 0$, and define inductively $T_1^\eta = 0$ and

$$T_{n+1}^\eta = \inf\{s \in \mathcal{S} : s > T_n^\eta \text{ and } |Z_s^\eta| > q\Big|Z_{T_n^\eta}^\eta\Big|\} \wedge t .$$

These are elementary stopping times, only finitely many distinct. Let N^η be the *last* index n such that $|Z_{T_n^\eta}^\eta| > |Z_{T_{n-1}^\eta}^\eta|$. Clearly $\sup_{s\in\mathcal{S}} |Z_s^\eta| \le q\,|Z_{T_{N^\eta}^\eta}^\eta|$. Now $\omega \mapsto T_{N^\eta(\omega)}^\eta$ is not a stopping time, inasmuch as one has to check $\boldsymbol{Z}_t$ at instants t later than $T_{N^\eta}^\eta$ in order to determine whether $T_{N^\eta}^\eta$ has arrived. This unfortunate fact necessitates a slightly circuitous argument.

Set $\zeta_0^\eta = 0$ and $\zeta_n^\eta = \big|Z_{T_n^\eta}^\eta\big|$ for $n = 1, \ldots, N^\eta$. Since $\zeta_n^\eta \ge q\zeta_{n-1}^\eta$ for $1 \le n \le N^\eta$,

$$\zeta_{N^\eta} \le L_q \left(\sum_{n=1}^{N^\eta} (\zeta_n^\eta - \zeta_{n-1}^\eta)^2\right)^{1/2} ;$$

we leave it to the reader to show by induction that this holds when the choice $L_q^2 \stackrel{\text{def}}{=} (q+1)/(q-1)$ is made.

Since
$$\sup_{s\in\mathcal{S}} |Z_s^\eta| \le q\Big|Z_{T_{N^\eta}^\eta}^\eta\Big| = q\,\zeta_{N^\eta}^\eta \le qL_q \left(\sum_{n=1}^{N^\eta} (\zeta_n^\eta - \zeta_{n-1}^\eta)^2\right)^{1/2} \tag{2.3.7}$$
$$\le qL_q \left(\sum_{n=1}^{\infty} \big(Z_{T_n^\eta}^\eta - Z_{T_{n-1}^\eta}^\eta\big)^2\right)^{1/2} ,$$

the quantity
$$\zeta^{\mathcal{S}} \stackrel{\text{def}}{=} \left\| \left(\sum_{\eta=1}^d \big(\sup_{s\in\mathcal{S}} |Z_s^\eta|\big)^2\right)^{1/2} \right\|_{L^p}$$

satisfies, thanks to the Khintchine inequality of theorem A.8.26,

$$\zeta^{\mathcal{S}} \le qL_q \left\| \left(\sum_{\eta=1}^{d} \sum_{n=1}^{\infty} \left| Z^\eta_{T^\eta_n} - Z^\eta_{T^\eta_{n-1}} \right|^2 \right)^{1/2} \right\|_{L^p(\mathbb{P})}$$

$$\le qL_q K_p^{(\mathrm{A.8.5})} \left\| \left\| \sum_{n,\eta} \big(Z^\eta_{T^\eta_n} - Z^\eta_{T^\eta_{n-1}} \big) \epsilon_{n,\eta}(\tau) \right\|_{L^p(d\tau)} \right\|_{L^p(\mathbb{P})}$$

by Fubini:
$$= qL_q K_p \left\| \left\| \sum_{n,\eta} \big(Z^\eta_{T^\eta_n} - Z^\eta_{T^\eta_{n-1}} \big) \epsilon_{n,\eta}(\tau) \right\|_{L^p(\mathbb{P})} \right\|_{L^p(d\tau)}$$

$$= qL_q K_p \left\| \left\| \int \sum_{\eta} \Big(\sum_{n} (\!(T^\eta_{n-1}, T^\eta_n]\!] \, \epsilon_{n,\eta}(\tau) \Big) \, dZ^\eta \right\|_{L^p(\mathbb{P})} \right\|_{L^p(d\tau)}$$

$$\le qL_q K_p \left\| \lceil \boldsymbol{Z}^t \rceil_{\mathcal{I}^p} \right\|_{L^p(d\tau)} = qL_q K_p \, \lceil \boldsymbol{Z}^t \rceil_{\mathcal{I}^p} \ .$$

In the penultimate line $(\!(T^\eta_0, T^\eta_1]\!]$ stands for $[\![0]\!]$. The sums are of course really finite, since no more summands can be non-zero than $\mathcal{S}$ has members. Taking now the supremum over all finite partitions of $[0,t]$ results, in view of the right-continuity of Z, in $\| \, |\boldsymbol{Z}^\star|_t \|_{L^p} \le qL_q K_p \lceil \boldsymbol{Z}^t \rceil_{\mathcal{I}^p}$. The constant qL_q is minimal for the choice $q = (1+\sqrt{5})/2$, where it equals $(q+1)/\sqrt{q-1} \le 10/3$. Lastly, observe that for a positive increasing process I, $I = |\boldsymbol{Z}^\star|$ in this case, the supremum in the definition of $\lceil I^t \rceil_{\mathcal{I}^p}$ on page 55 is assumed at the elementary integrand $[\![0,t]\!]$, where it equals $\| I^t \|_{L^p}$. This proves the equality in (2.3.5); since $|\boldsymbol{Z}^\star|$ is plainly right-continuous, it is an L^p-integrator. ❑

Exercise 2.3.7 The absolute value $|Z|$ of an L^p-integrator Z is an L^p-integrator, and

$$\lceil |Z|^t \rceil_{\mathcal{I}^p} \le 3 \lceil Z^t \rceil_{\mathcal{I}^p} \ , \qquad 0 \le p < \infty, 0 \le t \le \infty \ .$$

Consequently, $\mathcal{I}^p$ forms a vector lattice under pointwise operations.

Law and Canonical Representation

2.3.8 Adapted Maps between Filtered Spaces Let $(\Omega, \mathcal{F}_.)$ and $(\overline{\Omega}, \overline{\mathcal{F}}_.)$ be filtered probability spaces. We shall say that a map $\underline{R} : \Omega \to \overline{\Omega}$ is ***adapted to $\mathcal{F}_.$ and $\overline{\mathcal{F}}_.$*** if $\underline{R}$ is $\mathcal{F}_t/\overline{\mathcal{F}}_t$-measurable at all instants t. This amounts to saying that for all t

$$\underline{\mathcal{F}}_t \stackrel{\text{def}}{=} \underline{R}^{-1}(\overline{\mathcal{F}}_t) = \overline{\mathcal{F}}_t \circ \underline{R}^{\ 2} \tag{2.3.8}$$

is a sub-σ-algebra of $\mathcal{F}_t$. Occasionally we call such a map $\underline{R}$ a ***morphism of filtered spaces*** or a ***representation*** of $(\Omega, \mathcal{F}_.)$ on $(\overline{\Omega}, \overline{\mathcal{F}}_.)$, the idea being that it forgets unwanted information and leaves only the "aspect of interest" $(\overline{\Omega}, \overline{\mathcal{F}}_.)$. With such $\underline{R}$ comes naturally the map $(t,\omega) \mapsto \big(t, \underline{R}(\omega)\big)$ of the base space of Ω to the base space of $\overline{\Omega}$. We shall denote this map by $\underline{R}$ as well; this won't lead to confusion.

The following facts are obvious or provide easy exercises:

(i) If the process $\overline{X}$ on $\overline{\Omega}$ is left-continuous (right-continuous, càdlàg, continuous, of finite variation), then $X \stackrel{\text{def}}{=} \overline{X} \circ \underline{R}$ has the same property on Ω.

(ii) If $\overline{T}$ is an $\overline{\mathcal{F}}_{\bullet}$-stopping time, then $T \stackrel{\text{def}}{=} \overline{T} \circ \underline{R}$ is an $\underline{\mathcal{F}}_{\bullet}$-stopping time. If the process $\overline{X}$ is adapted (progressively measurable, an elementary integrand) on $(\overline{\Omega}, \overline{\mathcal{F}}_{\bullet})$, then $X \stackrel{\text{def}}{=} \overline{X} \circ \underline{R}$ is adapted (progressively measurable, an elementary integrand) on $(\Omega, \underline{\mathcal{F}}_{\bullet})$ and on $(\Omega, \mathcal{F}_{\bullet})$. $\overline{X}$ is predictable[7] on $(\overline{\Omega}, \overline{\mathcal{F}}_{\bullet})$ if and only if X is predictable on $(\Omega, \underline{\mathcal{F}}_{\bullet})$; it is then predictable on $(\Omega, \mathcal{F}_{\bullet})$.

(iii) If a probability $\mathbb{P}$ on $\underline{\mathcal{F}}_{\infty} \subset \mathcal{F}_{\infty}$ is given, then the image of $\mathbb{P}$ under $\underline{R}$ provides a probability $\overline{\mathbb{P}}$ on $\overline{\mathcal{F}}_{\infty}$. In this way the whole slew $\mathfrak{P}$ of pertinent probabilities gives rise to the pertinent probabilities $\overline{\mathfrak{P}}$ on $(\overline{\Omega}, \overline{\mathcal{F}}_{\infty})$.

Suppose $\overline{Z}_{\bullet}$ is a càdlàg process on $\overline{\Omega}$. Then $\overline{Z}$ is an $L^p(\overline{\mathbb{P}})$-integrator on $(\overline{\Omega}, \overline{\mathcal{F}}_{\bullet})$ if and only if $X \stackrel{\text{def}}{=} \overline{Z} \circ \underline{R}$ is an $L^p(\mathbb{P})$-integrator on $(\Omega, \underline{\mathcal{F}}_{\bullet})$.[8] To see this let $\underline{\mathcal{E}}$ denote the elementary integrands for the filtration $\underline{\mathcal{F}}_{\bullet} \stackrel{\text{def}}{=} \underline{R}^{-1}(\overline{\mathcal{F}}_{\bullet})$. It is easily seen that $\underline{\mathcal{E}} = \overline{\mathcal{E}} \circ \underline{R}$, in obvious notation, and that the collections of random variables

$$\left\{ \int \overline{X}\, d\overline{Z} : \overline{X} \in \overline{\mathcal{E}}, |\overline{X}| \leq \overline{Y} \right\} \quad \text{and} \quad \left\{ \int X\, dZ : X \in \underline{\mathcal{E}}, |X| \leq Y \right\}$$

upon being measured with $\lceil\!\lceil\ \rceil\!\rceil^*_{L^p(\overline{\mathbb{P}})}$ and $\lceil\!\lceil\ \rceil\!\rceil^*_{L^p(\mathbb{P})}$, respectively, produce the same sets of numbers when $\mathcal{E}_+ \ni Y = \overline{Y} \circ R$. The equality of the suprema reads

$$\left\lceil\!\!\left\lceil \overline{Y} \right\rceil\!\!\right\rceil_{\overline{Z}\text{–}p;\overline{\mathbb{P}}} = \lceil\!\lceil Y \rceil\!\rceil_{Z\text{–}p;\mathbb{P}} \tag{2.3.9}$$

for $Y = \overline{Y} \circ \underline{R}$, $\overline{\mathbb{P}} = \underline{R}[\mathbb{P}]$, and $Z = \overline{Z} \circ \underline{R}$ considered as an integrator on $(\Omega, \underline{\mathcal{F}}_{\bullet})$.[8]

Let us then henceforth forget information that may be present in $\mathcal{F}_{\bullet}$ but not in $\underline{\mathcal{F}}_{\bullet}$, by replacing the former filtration with the latter. That is to say,

$$\mathcal{F}_t = \underline{R}^{-1}(\overline{\mathcal{F}}_t) = \overline{\mathcal{F}}_t \circ \underline{R} \quad \forall\, t \geq 0\,, \quad \text{and then} \quad \mathcal{E} = \overline{\mathcal{E}} \circ \underline{R}\,.$$

Once the integration theory of Z and $\overline{Z}$ is established in chapter 3, the following further facts concerning a process X of the form $X = \overline{X} \circ \underline{R}$ will be obvious:

(iv) X is previsible with $\mathbb{P}$ if and only if $\overline{X}$ is previsible with $\overline{\mathbb{P}}$.

(v) X is Z–$p;\mathbb{P}$-integrable if and only if $\overline{X}$ is $\overline{Z}$–$p;\overline{\mathbb{P}}$-integrable, and then

$$(\overline{\boldsymbol{X}} {*} \overline{Z})_{\bullet} \circ \underline{R} = (\boldsymbol{X} {*} \boldsymbol{Z})_{\bullet}\,. \tag{2.3.10}$$

(vi) X is Z-measurable if and only if $\overline{X}$ is $\overline{Z}$-measurable. Any Z-measurable process differs Z-negligibly from a process of this form.

---

[7] A process is predictable if it belongs to the sequential closure of the elementary integrands – see section 3.5.

[8] Note the underscore! One cannot expect in general that Z be an $L^p(\mathbb{P})$-integrator, i.e., be bounded on the potentially much larger space $\mathcal{E}$ of elementary integrands for $\mathcal{F}_{\bullet}$.

2.3.9 Canonical Path Space In algebra one tries to get insight into the structure of an object by representing it with morphisms on objects of the same category that have additional structure. For example, groups get represented on matrices or linear operators, which one can also add, multiply with scalars, and measure by size. In a similar vein[9] the typical target space of a representation is a space of paths, which usually carries a topology and may even have a linear structure:

Let (E,ρ) be some polish space. $\mathscr{D}_E$ denotes the set of all càdlàg paths $x_. : [0,\infty) \to E$. If $E = \mathbb{R}$, we simply write $\mathscr{D}$; if $E = \mathbb{R}^d$, we write $\mathscr{D}^d$. A path in $\mathscr{D}^d$ is identified with a path on $(-\infty,\infty)$ that vanishes on $(-\infty,0)$.

A natural topology on $\mathscr{D}_E$ is the topology τ of uniform convergence on bounded time-intervals; it is given by the complete metric

$$d(x_., y_.) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{N}} 2^{-n} \wedge \rho(x_., y_.)^\star_n \qquad x_., y_. \in \mathscr{D}_E\ ,$$

where

$$\rho(x_., y_.)^\star_t \stackrel{\text{def}}{=} \sup_{0\le s\le t} \rho(x_s, y_s)\ .$$

The maximal theorem 2.3.6 shows that this topology is pertinent. Yet its Borel σ-algebra is rarely useful; it is too fine. Rather, it is the ***basic filtration*** $\mathcal{F}^0_.[\mathscr{D}_E]$, generated by the right-continuous ***evaluation process***

$$\overline{R}_s : x_. \mapsto x_s\ , \qquad 0 \le s < \infty\ , x_. \in \mathscr{D}_E\ ,$$

and its right-continuous version $\mathcal{F}^0_{.+}[\mathscr{D}_E]$ that play a major role. The final σ-algebra $\mathcal{F}^0_\infty[\mathscr{D}_E]$ of the basic filtration coincides with the Baire σ-algebra of the topology σ of pointwise convergence on $\mathscr{D}_E$. On the space $\mathscr{C}_E$ of continuous paths the σ-algebras generated by σ and τ coincide (generalize equation (1.2.5)).

The right-continuous version $\mathcal{F}^0_{.+}[\mathscr{D}_E]$ of the basic filtration will also be called the ***canonical filtration***. The space $\mathscr{D}_E$ equipped with the topology τ[10] and its canonical filtration $\mathcal{F}^0_{.+}[\mathscr{D}_E]$ is ***canonical path space.***[11]

Consider now a càdlàg adapted E-valued process R on $(\Omega, \mathcal{F}_.)$. Just as a Wiener process was considered as a random variable with values in canonical path space $\mathscr{C}$ (page 14), so can now our process R be regarded as a map $\underline{R}$ from Ω to path space $\mathscr{D}_E$, the image of an $\omega \in \Omega$ under $\underline{R}$ being the path $R_.(\omega) : t \mapsto R_t(\omega)$. Since R is assumed adapted, $\underline{R}$ represents $(\Omega, \mathcal{F}_.)$ on path space $(\mathscr{D}_E, \mathcal{F}^0_.[\mathscr{D}_E])$ in the sense of item 2.3.8. If $\mathcal{F}_.$ is right-continuous, then $\underline{R}$ represents $(\Omega, \mathcal{F}_.)$ on canonical path space $(\mathscr{D}_E, \mathcal{F}^0_{.+}[\mathscr{D}_E])$. We call $\underline{R}$ the ***canonical representation*** of R on path space.

---

[9] I hope that the reader will find a little farfetchedness more amusing than offensive.

[10] A glance at theorems 2.3.6, 4.5.1, and A.4.9 will convince the reader that τ is most pertinent, despite the fact that it is not polish and that its Borels *properly* contain the pertinent σ-algebra $\mathcal{F}_\infty$.

[11] "Path space", like "frequency space" or "outer space," may be used without an article.

If $(\Omega, \mathcal{F}_.)$ carries a distinguished probability $\mathbb{P}$, then the ***law*** of the process R is of course nothing but the image $\overline{\mathbb{P}} \stackrel{\text{def}}{=} \underline{R}[\mathbb{P}]$ of $\mathbb{P}$ under $\underline{R}$. The triple $(\mathscr{D}_E, \mathcal{F}^0_{.+}[\mathscr{D}_E], \overline{\mathbb{P}})$ carries all statistical information about the process R – which now "is" the evaluation process $\overline{R}_.$ – and has forgotten all other information that might have been available on $(\Omega, \mathcal{F}_., \mathbb{P})$.

2.3.10 Integrators on Canonical Path Space Suppose that E comes equipped with a distinguished slew $\boldsymbol{z} = (z^1, \ldots, z^d)$ of continuous functions. Then $t \mapsto \overline{Z}_t \stackrel{\text{def}}{=} \boldsymbol{z} \circ \overline{R}_t$ is a distinguished adapted $\mathbb{R}^d$-valued process on the path space $(\mathscr{D}_E, \mathcal{F}^0_.[\mathscr{D}_E], \overline{\mathbb{P}})$. These data give rise to the collection $\mathfrak{P}[\overline{Z}]$ of all probabilities on path space for which $\overline{Z}$ is an integrator. We may then define the ***natural filtration*** on $\mathscr{D}_E$: it is the regularization of $\mathcal{F}^0_{.+}[\mathscr{D}_E]$, taken for the collection $\mathfrak{P}[\overline{Z}]$, and it is denoted by $\mathcal{F}_.[\mathscr{D}_E]$ or $\mathcal{F}_.[\mathscr{D}_E; \boldsymbol{z}]$.

2.3.11 Canonical Representation of an Integrator Suppose that we face an integrator $\boldsymbol{Z} = (Z^1, \ldots, Z^d)$ on $(\Omega, \mathcal{F}_., \mathfrak{P})$ and a collection $C = (C^1, C^2, \ldots)$ of real-valued processes, certain functions f_η of which we might wish to integrate with $\boldsymbol{Z}$, say. We glob the data together in the obvious way into a process $R_t \stackrel{\text{def}}{=} (C_t, \boldsymbol{Z}_t) : \Omega \to E \stackrel{\text{def}}{=} \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^d$, which we identify with a map $\underline{R} : \Omega \to \mathscr{D}_E$. "$\underline{R}$ forgets all information except the aspect of interest $(C, \boldsymbol{Z})$." Let us write $\overline{\omega}_. = (c^\nu_., \boldsymbol{z}^\eta_.)$ for the generic point of $\overline{\Omega} = \mathscr{D}_E$. On E there are the distiguished last d coordinate functions $z^1, \ldots, z^d$. They give rise to the distinguished process $\overline{Z} : t \mapsto \big(z^1(\overline{\omega}_t), \ldots, z^d(\overline{\omega}_t)\big)$. Clearly the image under $\underline{R}$ of any probability in $\mathfrak{P} \subset \mathfrak{P}[\boldsymbol{Z}]$ makes $\overline{Z}$ into an integrator on path space.[11] The integral $\int_0^t f_\eta[\overline{\omega}_.]_s \, d\overline{Z}^\eta_s(\overline{\omega}_.)$, which is frequently and with intuitive appeal written as

$$\int_0^t f_\eta[c_., \boldsymbol{z}_.]_s \, dz^\eta \, , \tag{2.3.11}$$

then equals $\int_0^t f_\eta[C_., \boldsymbol{Z}_.]_s \, dZ^\eta_s$, after composition with $\underline{R}$, that is, and after information beyond $\underline{\mathcal{F}}_.$ has been discarded.[8] In other words,

$$\boldsymbol{X} * \boldsymbol{Z} = (\overline{\boldsymbol{X}} * \overline{\boldsymbol{Z}}) \circ \underline{R} \, . \tag{2.3.12}$$

In this way we arrive at the ***canonical representation*** $\underline{R}$ of $(C, \boldsymbol{Z})$ on $(\mathscr{D}_E, \mathcal{F}_.[\mathscr{D}_E])$ with pertinent probabilities $\overline{\mathfrak{P}} \stackrel{\text{def}}{=} \underline{R}[\mathfrak{P}]$. For an application see page 316.

2.4 Processes of Finite Variation

Recall that a process V has ***bounded variation*** if its paths are functions of bounded variation on the half-line, i.e., if the number

$$\lvert V \rvert_\infty(\omega) = |V_0| + \sup \Big\{ \sum_{i=1}^{I} |V_{t_{i+1}}(\omega) - V_{t_i}(\omega)| \Big\}$$

is finite for every $\omega \in \Omega$. Here the supremum is taken over all finite partitions $\mathcal{T} = \{t_1 < t_2 < \ldots < t_{I+1}\}$ of $\mathbb{R}_+$. V has ***finite variation*** if the stopped processes V^t have bounded variation, at every instant t. In this case the variation process $¦V¦$ of V is defined by

$$¦V¦_t(\omega) = |V_0(\omega)| + \sup_{\mathcal{T}} \left\{ \sum_{i=1}^{I} |V_{t\wedge t_{i+1}}(\omega) - V_{t\wedge t_i}(\omega)| \right\} . \tag{2.4.1}$$

The integration theory of processes of finite variation can of course be handled path-by-path. Yet it is well to see how they fit in the general framework.

Proposition 2.4.1 *Suppose V is an adapted right-continuous process of finite variation. Then $¦V¦$ is adapted, increasing, and right-continuous with left limits. Both V and $¦V¦$ are L^0-integrators.*

If $¦V¦_t \in L^p$ at all instants t, then V is an L^p-integrator. In fact, for $0 \le p < \infty$ and $0 \le t \le \infty$

$$\lceil\!\lceil [\![0,t]\!] \rceil\!\rceil_{V\text{-}p} = \lceil V^t \rceil_{\mathcal{I}^p} \le \Big\lceil\!\!\Big\lceil ¦V¦_t \Big\rceil\!\!\Big\rceil_p . \tag{2.4.2}$$

Proof. Due to the right-continuity of V, taking the partition points t_i of equation (2.4.1) in the set $\mathbb{Q}^t = (\mathbb{Q} \cap [0,t]) \cup \{t\}$ will result in the same path $t \mapsto ¦V¦_t(\omega)$; and since the collection of finite subsets of $\mathbb{Q}^t$ is countable, the process $¦V¦$ is adapted. For every $\omega \in \Omega$, $t \mapsto ¦V¦_t(\omega)$ is the cumulative distribution function of the variation $¦dV¦(\omega)$ of the scalar measure $dV_{\boldsymbol{.}}(\omega)$ on the half-line. It is therefore right-continuous (exercise 2.2). Next, for $X \in \mathcal{E}_1$, f_n, and t_n as in equation (2.1.1) we have

$$\begin{aligned} \left| \int X \, dV^t \right| &= \left| f_0 \cdot V_0 + \sum\nolimits_n f_n \cdot (V^t_{t_{n+1}} - V^t_{t_n}) \right| \\ &\le |V_0| + \sum\nolimits_n \left| V^t_{t_{n+1}} - V^t_{t_n} \right| \le ¦V¦_t . \end{aligned}$$

We apply $\lceil\!\lceil \ \rceil\!\rceil_p$ to this and obtain inequality (2.4.2). ∎

Our adapted right-continuous process of finite variation therefore can be written as the difference of two adapted increasing right-continuous processes $V^{\pm}$ of finite variation: $V = V^+ - V^-$ with

$$V^+_t = 1/2 \left(¦V¦_t + V_t \right) \ , \ V^-_t = 1/2 \left(¦V¦_t - V_t \right) .$$

It suffices to analyze *increasing* adapted right-continuous processes I.

Remark 2.4.2 The reverse of inequality (2.4.2) is not true in general, nor is it even true that $¦V¦_t \in L^p$ if V is an L^p-integrator, except if $p = 0$. The reason is that the collection $\mathcal{E}$ is too small; testing V against its members is not enough to determine the variation of V, which can be written as

$$¦V¦_t = |V_0| + \sup \int_0^t \operatorname{sgn}(V_{t_{i+1}} - V_{t_i}) \, dV .$$

Note that the integrand here is not elementary inasmuch as $(V_{t_{i+1}} - V_{t_i}) \notin \mathcal{F}_{t_i}$. However, in (2.4.2) equality holds if V is previsible (exercise 4.3.13) or increasing. Example 2.5.26 on page 79 exhibits a sequence of processes whose variation grows beyond all bounds yet whose $\mathcal{I}^2$-norms stay bounded.

Exercise 2.4.3 Prove the right-continuity of $¦V¦$ directly.

Decomposition into Continuous and Jump Parts

A measure μ on $[0,\infty)$ is the sum of a measure ${}^c\mu$ that does not charge points and an atomic measure ${}^j\mu$ that is carried by a countable collection $\{t_1, t_2, \ldots\}$ of points. The cumulative distribution function[12] of ${}^c\mu$ is continuous and that of ${}^j\mu$ constant, except for jumps at the times t_n, and the cumulative distribution function of μ is the sum of these two. All of this is classical, and every path of an increasing right-continuous process can be decomposed in this way. In the stochastic case we hope that the continuous and jump components are again adapted, and this is indeed so; also, the times of the jumps of the discontinuous part are not too wildly scattered:

Theorem 2.4.4 *A positive increasing adapted right-continuous process I can be written uniquely as the sum of a continuous increasing adapted process ${}^cI$ that vanishes at 0 and a right-continuous increasing adapted process ${}^jI$ of the following form: there exist a countable collection $\{T_n\}$ of stopping times with bounded disjoint graphs,[13] and bounded positive $\mathcal{F}_{T_n}$-measurable functions f_n, such that*

$$ {}^jI = \sum_n f_n \cdot [\![T_n, \infty)\!) \ . $$

Proof. For every $i \in \mathbb{N}$ define inductively $T^{i,0} = 0$ and

$$ T^{i,j+1} = \inf \left\{ t > T^{i,j} : \ \Delta I_t \geq 1/i \right\} \ . $$

From proposition 1.3.14 we know that the $T^{i,j}$ are stopping times. They increase a.s. strictly to ∞ as $j \to \infty$; for if $T = \sup_j T^{i,j} < \infty$, then $I_t = \infty$ after T. Next let $T^{i,j}_k$ denote the reduction of $T^{i,j}$ to the set

$$ [\Delta I_{T^{i,j}} \leq k+1] \cap [T^{i,j} \leq k] \in \mathcal{F}_{T^{i,j}} \ . $$

(See exercises 1.3.18 and 1.3.16.) Every one of the $T^{i,j}_k$ is a stopping time with a bounded graph. The jump of I at time $T^{i,j}_k$ is bounded, and the set $[\Delta I \neq 0]$ is contained in the union of the graphs of the $T^{i,j}_k$. Moreover, the collection $\{T^{i,j}_k\}$ is countable; so let us count it: $\{T^{i,j}_k\} = \{T'_1, T'_2, \ldots\}$. The T'_n do not have disjoint graphs, of course. We force the issue by letting T_n be the reduction of T'_n to the set $\bigcup_{m<n} [T'_n \neq T'_m] \in \mathcal{F}_{T'_n}$ (exercise 1.3.16). It is plain upon inspection that with $f_n = \Delta I_{T_n}$ and ${}^cI = I - {}^jI$ the statement is met. ∎

Exercise 2.4.5 ${}^jI_t = \sum_{s \leq t} \Delta I_s = \sum_n \sum_{s \leq t} f_n \cdot [T_n = s]$.

Exercise 2.4.6 Call a subset S of the base space ***sparse*** if it is contained in the union of the graphs of countably many stopping times. Such stopping times can be chosen to have disjoint graphs; and if S is measurable, then it actually equals the union of the disjoint graphs of countably many stopping times (use theorem A.5.10).

---

[12] See page 406.
[13] We say that T has a ***bounded graph*** if $[\![T]\!] \subset [\![0, t]\!]$ for some finite instant t.

Next let V be an adapted càdlàg process of finite variation. Then V is the sum $V = {}^{c}V + {}^{j}V$ of two adapted càdlàg processes of finite variation, of which ${}^{c}V$ has continuous paths and $d\,{}^{j}V = S \cdot d\,{}^{j}V$ with $S \overset{\text{def}}{=} [\Delta V \neq 0] = [\Delta\,{}^{j}V \neq 0]$ sparse. For more see exercise 4.3.4.

The Change-of-Variable Formula

Theorem 2.4.7 *Let I be an adapted positive increasing right-continuous process and $\Phi : [0, \infty) \to \mathbb{R}_+$ a continuously differentiable function. Set*

$$T^\lambda = \inf\{t : I_t \geq \lambda\} \quad \textit{and} \quad T^{\lambda+} = \inf\{t : I_t > \lambda\} \,, \qquad \lambda \in \mathbb{R} \,.$$

Both $\{T^\lambda\}$ and $\{T^{\lambda+}\}$ form increasing families of stopping times, $\{T^\lambda\}$ left-continuous and $\{T^{\lambda+}\}$ right-continuous. For every bounded measurable process X [14]

$$\int_{[0}^{\infty} X_s \, d\Phi(I_s) = \int_0^\infty X_{T^\lambda} \cdot \Phi'(\lambda) \cdot [T^\lambda < \infty] \, d\lambda \tag{2.4.3}$$

$$= \int_0^\infty X_{T^{\lambda+}} \cdot \Phi'(\lambda) \cdot [T^{\lambda+} < \infty] \, d\lambda \,. \tag{2.4.4}$$

Proof. Thanks to proposition 1.3.11 the T^λ are stopping times and are increasing and left-continuous in λ. Exercise 1.3.30 yields the corresponding claims for $T^{\lambda+}$. $T^\lambda < T^{\lambda+}$ signifies that $I = \lambda$ on an interval of strictly positive length. This can happen only for countably many different λ. Therefore the right-hand sides of (2.4.3) and (2.4.4) coincide.

To prove (2.4.3), say, consider the family $\mathcal{M}$ of bounded measurable processes X such that for all finite instants u

$$\int [\![0, u]\!] \cdot X \, d\Phi(I) = \int_0^\infty X_{T^\lambda} \cdot \Phi'(\lambda) \cdot [T^\lambda \leq u] \, d\lambda \,. \tag{?}$$

$\mathcal{M}$ is clearly a vector space closed under pointwise limits of bounded sequences. For processes X of the special form

$$X = f \cdot [0, t] \,, \qquad f \in L^\infty(\mathcal{F}_\infty) \,, \tag{$*$}$$

the left-hand side of (?) is simply

$$f \cdot \big(\Phi(I_{t\wedge u}) - \Phi(I_{0-})\big) = f \cdot \big(\Phi(I_{t\wedge u}) - \Phi(0)\big)$$

---

[14] Recall from convention A.1.5 that $[T^\lambda < \infty]$ equals 1 if $T^\lambda < \infty$ and 0 otherwise. Indicator function aficionados read these integrals as $\int_0^\infty X_{T^\lambda} \cdot \Phi'(\lambda) \cdot 1_{[T^\cdot < \infty]}(\lambda) \, d\lambda$, etc.

and the right-hand side is[14]

$$
\begin{aligned}
&f \cdot \int_0^\infty [0,t](T^\lambda) \cdot \Phi'(\lambda) \cdot [T^\lambda \le u]\, d\lambda \\
=&f \cdot \int_0^\infty [T^\lambda \le t] \cdot \Phi'(\lambda) \cdot [T^\lambda \le u]\, d\lambda \\
=&f \cdot \int_0^\infty [T^\lambda \le t \wedge u] \cdot \Phi'(\lambda)\, d\lambda \\
=&f \cdot \int_0^\infty [\lambda \le I_{t\wedge u}] \cdot \Phi'(\lambda)\, d\lambda = f \cdot \big(\Phi(I_{t\wedge u}) - \Phi(0)\big)
\end{aligned}
$$

as well. That is to say, $\mathcal{M}$ contains the processes of the form $(*)$, and also the constant process 1 (choose $f \equiv 1$ and $t \ge u$). The processes of the form $(*)$ generate the measurable processes and so, in view of theorem A.3.4 on page 393, $(*)$ holds for all bounded measurable processes. Equation (2.4.3) follows upon taking u to ∞. ∎

Exercise 2.4.8 $I_t = \inf\{\lambda : T^\lambda > t\} = \int [T^\lambda, t]\, d\lambda = \int [\![T^\lambda, \infty)\!)_t\, d\lambda$ (see convention A.1.5). A stochastic interval $[\![T, \infty)\!)$ is an increasing adapted process (ibidem). Equation (2.4.3) can thus be read as saying that $\Phi(I)$ is a "continuous superposition" of such simple processes:

$$\Phi(I) = \int_0^\infty \Phi'(\lambda)[\![T^\lambda, \infty)\!)\, d\lambda\,.$$

Exercise 2.4.9 (i) If the right-continuous adapted process I is strictly increasing, then $T^\lambda = T^{\lambda+}$ for every $\lambda \ge 0$; in general, $\{\lambda : T^\lambda < T^{\lambda+}\}$ is countable.

(ii) Suppose that $T^{\lambda+}$ is nearly finite for all λ and $\mathcal{F}_{\bullet}$ meets the natural conditions. Then $(\mathcal{F}_{T^{\lambda+}})_{\lambda\ge 0}$ inherits the natural conditions; if Λ is an $\mathcal{F}_{T^{\cdot+}}$-stopping time, then $T^{\Lambda+}$ is an $\mathcal{F}_{\bullet}$-stopping time.

Exercise 2.4.10 Equations (2.4.3) and (2.4.4) hold for measurable processes X whenever one or the other side is finite.

Exercise 2.4.11 If $T^{\lambda+} < \infty$ almost surely for all λ, then the filtration $\big(\mathcal{F}_{T^{\lambda+}}\big)_\lambda$ inherits the natural conditions from $\mathcal{F}_{\bullet}$.

2.5 Martingales

Definition 2.5.1 *An integrable process M is an $(\mathcal{F}_{\bullet}, \mathbb{P})$-**martingale** if*[15]

$$\mathbb{E}^{\mathbb{P}}[M_t | \mathcal{F}_s] = M_s$$

*for $0 \le s < t < \infty$. We also say that M is a $\mathbb{P}$-**martingale on** $\mathcal{F}_{\bullet}$, or simply a **martingale** if the filtration $\mathcal{F}_{\bullet}$ and probability $\mathbb{P}$ meant are clear from the context.*

Since the conditional expectation above is unique only up to $\mathbb{P}$-negligible and $\mathcal{F}_s$-measurable functions, the equation should be read "M_s is a (one of very many) conditional expectation of M_t given $\mathcal{F}_s$."

[15] $\mathbb{E}^{\mathbb{P}}[M_t|\mathcal{F}_s]$ is the conditional expectation of M_t given $\mathcal{F}_s$ – see theorem A.3.24 on page 407.

A martingale on $\mathcal{F}_{\bullet}$ is clearly adapted to $\mathcal{F}_{\bullet}$. The martingales form a class of integrators that is complementary to the class of finite variation processes – in a sense that will become clearer as the story unfolds – and that is much more challenging. The name "martingale" seems to derive from the part of a horse's harness that keeps the beast from throwing up its head and thus from rearing up; the term has also been used in gambling for centuries. The defining equality for a martingale says this: given the whole history $\mathcal{F}_s$ of the game up to time s, the gambler's fortune at time $t > s$, $\$M_t$, is expected to be just what she has at time s, namely, $\$M_s$; in other words, she is engaged in a fair game. Roughly, martingales are processes that show, on the average, no drift (see the discussion on page 4).

The class of L^0-integrators is rather stable under changes of the probability (proposition 2.1.9), but the class of martingales is not. It is rare that a process that is a martingale with respect to one probability is a martingale with respect to an equivalent or otherwise pertinent measure. For instance, if the dice in a fair game are replaced by loaded ones, the game will most likely cease to be fair, that being no doubt the object of the replacement. Therefore we will fix a probability $\mathbb{P}$ on $\mathcal{F}_\infty$ throughout this section. $\mathbb{E}$ is understood to be the expectation $\mathbb{E}^{\mathbb{P}}$ with respect to $\mathbb{P}$.

Example 2.5.2 Here is a frequent construction of martingales. Let g be an integrable random variable, and set $M^g_t = \mathbb{E}[g|\mathcal{F}_t]$, the conditional expectation of g given $\mathcal{F}_t$. Then M^g is a uniformly integrable martingale – it is shown in exercise 2.5.14 that all uniformly integrable martingales are of this form. It is an easy exercise to establish that the collection

$$\{\mathbb{E}[g|\mathcal{G}] : \mathcal{G} \text{ a sub-}\sigma\text{-algebra of } \mathcal{F}_\infty\}$$

of random variables is uniformly integrable.

Exercise 2.5.3 Suppose M is a martingale. Then $\mathbb{E}[f \cdot (M_t - M_s)] = 0$ for $s < t$ and any $f \in L^\infty(\mathcal{F}_s)$. Next assume M is ***square integrable***: $M_t \in L^2(\mathcal{F}_t, \mathbb{P})\quad \forall t$. Then

$$\mathbb{E}[(M_t - M_s)^2|\mathcal{F}_s] = \mathbb{E}[M_t^2 - M_s^2|\mathcal{F}_s]\,, \qquad 0 \le s < t < \infty\,.$$

Exercise 2.5.4 If W is a Wiener process on the filtration $\mathcal{F}_{\bullet}$, then it is a martingale on $\mathcal{F}_{\bullet}$ and on the natural enlargement of $\mathcal{F}_{\bullet}$, and so are $W_t^2 - t$, and $e^{zW_t - z^2 t/2}$ for any $z \in \mathbb{C}$.

Exercise 2.5.5 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathfrak{F}$ a collection of sub-σ-algebras of $\mathcal{F}$ that is increasingly directed. That is to say, for any two $\mathcal{F}_1, \mathcal{F}_2 \in \mathfrak{F}$ there is a σ-algebra $\mathcal{G} \in \mathfrak{F}$ containing both $\mathcal{F}_1$ and $\mathcal{F}_2$. Let $g \in L^1(\mathcal{F}, \mathbb{P})$ and for $\mathcal{G} \in \mathfrak{F}$ set $g^{\mathcal{G}} \overset{\text{def}}{=} \mathbb{E}[g|\mathcal{G}]$. The collection $\{g^{\mathcal{G}} : \mathcal{G} \in \mathfrak{F}\}$ is uniformly integrable and converges in L^1-mean to the conditional expectation of g with respect to the σ-algebra $\bigvee \mathfrak{F} \subset \mathcal{F}$ generated by $\mathfrak{F}$.

Exercise 2.5.6 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $f, f' : \Omega \to \overline{\mathbb{R}}_+$ two $\mathcal{F}$-measurable functions such that $[f \le q] = [f' \le q]$ $\mathbb{P}$-almost surely for all rationals q. Then $f = f'$ $\mathbb{P}$-almost surely.

Submartingales and Supermartingales

The martingales are the processes of primary interest in the sequel. It eases their analysis though to introduce the following generalizations. An integrable process Z adapted to $\mathcal{F}_{\bullet}$ is a ***submartingale*** (***supermartingale***) if

$$\mathbb{E}\left[Z_t|\mathcal{F}_s\right] \;\geq Z_s \text{ a.s. } (\leq Z_s \text{ a.s., respectively}), \qquad 0 \leq s \leq t < \infty\,.$$

Exercise 2.5.7 The fortune of a gambler in Las Vegas is a supermartingale, that of the casino is a submartingale.

Since the absolute-value function $|\cdot|$ is convex, it follows immediately from Jensen's inequality in A.3.24 that the absolute value of a martingale M is a submartingale:

$$\big|M_s\big| = \Big|\mathbb{E}[M_t|\mathcal{F}_s]\Big| \leq \mathbb{E}\Big[\big|M_t\big|\,|\mathcal{F}_s\Big] \text{ a.s.,} \quad 0 \leq s < t < \infty\,.$$

Taking expectations, $\quad \mathbb{E}\big[\big|M_s\big|\big] \leq \mathbb{E}\big[\big|M_t\big|\big] \qquad 0 \leq s < t < \infty$

follows. This argument lends itself to some generalizations:

Exercise 2.5.8 Let $\boldsymbol{M} = (M^1, \ldots, M^d)$ be a vector of martingales and $\Phi : \mathbb{R}^d \to \mathbb{R}$ convex and so that $\Phi(\boldsymbol{M}_t)$ is integrable for all t. Then $\Phi(\boldsymbol{M})$ is a submartingale. Apply this with $\Phi(\boldsymbol{x}) = |\boldsymbol{x}|_p$ to conclude that $t \mapsto |\boldsymbol{M}_t|_p$ is a submartingale for $1 \leq p \leq \infty$ and that $t \mapsto \|M^1_t\|_{L^p}$ is increasing if M^1 is p-integrable (bounded if $p = \infty$).

Exercise 2.5.9 A martingale M on $\mathcal{F}_{\bullet}$ is a martingale on its own basic filtration $\mathcal{F}^0_{\bullet}[M]$. Similar statements hold for sub– and supermartingales.

Here is a characterization of martingales that gives a first indication of the special role they play in stochastic integration:

Proposition 2.5.10 *The adapted integrable process M is a martingale (submartingale, supermartingale) on $\mathcal{F}_{\bullet}$ if and only if*

$$\mathbb{E}\Big[\int X\,dM\Big] = 0 \qquad (\geq 0\,,\ \leq 0\,,\ \textit{respectively})$$

for every positive elementary integrand X that vanishes on $[\![0]\!]$.

Proof. $\Longrightarrow$: A glance at equation (2.1.1) shows that we may take X to be of the form $X = f \cdot (\!(s,t]\!]$, with $0 \leq s < t < \infty$ and $f \geq 0$ in $L^\infty(\mathcal{F}_s)$. Then

$$\begin{aligned}\mathbb{E}\Big[\int X\,dM\Big] &= \mathbb{E}\big[f\cdot(M_t - M_s)\big] = \mathbb{E}\big[f\cdot M_t\big] - \mathbb{E}\big[f\cdot M_s\big]\\ &= \mathbb{E}\Big[f\cdot\Big(\mathbb{E}\big[M_t|\mathcal{F}_s\big] - M_s\Big)\Big]\\ &= (\geq,\leq)\;\; \mathbb{E}\Big[f\cdot(M_s - M_s)\Big] = 0\,.\end{aligned}$$

$\Longleftarrow$: If, on the other hand,

$$\mathbb{E}\Big[\int X\,dM\Big] = \mathbb{E}\Big[f\cdot\Big(\mathbb{E}\big[M_t|\mathcal{F}_s\big] - M_s\Big)\Big] = \;(\geq,\leq)\;0$$

for all $f \geq 0$ in $L^\infty(\mathcal{F}_s)$, then $\mathbb{E}\big[M_t|\mathcal{F}_s\big] - M_s = \;(\geq,\leq)\;0$ almost surely, and M is a martingale (submartingale, supermartingale). ▮

Corollary 2.5.11 *Let M be a martingale (submartingale, supermartingale). Then for any two elementary stopping times $S \le T$ we have nearly*

$$\mathbb{E}\big[M_T - M_S|\mathcal{F}_S\big] = 0 \ \ (\ge 0, \ \le 0) \ .$$

Proof. We show this for submartingales. Let $A \in \mathcal{F}_S$ and consider the reduced stopping times S_A, T_A. From equation (2.1.3) and proposition 2.5.10

$$\begin{aligned} 0 \le \mathbb{E}\Big[\int (\!(S_A \wedge T, T_A \wedge T]\!]\, dM\Big] &= \mathbb{E}\Big[M_{T_A \wedge T} - M_{S_A \wedge T}\Big] \\ = \mathbb{E}\Big[\big(M_T - M_S\big) \cdot 1_A\Big] &= \mathbb{E}\Big[\Big(\mathbb{E}\big[M_T|\mathcal{F}_S\big] - M_S\Big) \cdot 1_A\Big] \ . \end{aligned}$$

As $A \in \mathcal{F}_S$ was arbitrary, this shows that $\mathbb{E}\big[M_T|\mathcal{F}_S\big] \ge M_S$, except in a negligible set of $\mathcal{F}_T \subset \mathcal{F}_{\max T}$. ∎

Exercise 2.5.12 (i) An adapted integrable process M is a martingale (submartingale, supermartingale) if and only if $\mathbb{E}\,[M_T - M_S] = 0$ $(\ge 0, \ \le 0)$ for any two elementary stopping times $S \le T$.

(ii) The infimum of two supermartingales is a supermartingale.

Regularity of the Paths: Right-Continuity and Left Limits

Consider the estimate (2.3.3) of the number of upcrossings of the interval $[a, b]$ that the path of our martingale M performs on the finite subset

$$S = \{s_0 < s_1 < \ldots < s_N\} \ \text{ of } \ \mathbb{Q}_+^{u-} \stackrel{\text{def}}{=} \mathbb{Q}_+ \cap [0, u) :$$

$$\Big[U_S^{[a,b]} \ge n\Big] \le \frac{1}{n(b-a)} \left(\int \sum_{k=0}^{\infty} (\!(T_{2k}, T_{2k+1}]\!]\, dM \ + \ \big|M_u - a\big|\right) \ .$$

Applying the expectation and proposition 2.5.10 we obtain an estimate of the probability that this number exceeds $n \in \mathbb{N}$:

$$\mathbb{P}\left[U_S^{[a,b]} \ge n\right] \le \frac{1}{n(b-a)} \cdot \mathbb{E}\big[\big|M_u\big| + |a|\big] \ .$$

Taking the supremum over all finite subsets $S \subset \mathbb{Q}_+^{u-}$ and then over all $n \in \mathbb{N}$ and all pairs of rationals $a < b$ shows as on page 60 that the set

$$Osc \stackrel{\text{def}}{=} \bigcup_n \ \bigcup_{a,b \in \mathbb{Q}, a<b} \Big[U_{\mathbb{Q}_+^{n-}}^{[a,b]} = \infty\Big]$$

belongs to $\mathcal{A}_{\infty\sigma}$ and is $\mathbb{P}$-nearly empty. Let Ω_0 be its complement and define

$$M'_t(\omega) = \begin{cases} \lim\limits_{\mathbb{Q} \ni q \downarrow t} M_q(\omega) & \text{for } \omega \in \Omega_0, \\ \qquad 0 & \text{for } \omega \in Osc. \end{cases}$$

The limit is understood in the extended reals $\overline{\mathbb{R}}$ as nothing said so far prevents it from being $\pm\infty$. The process M' is right-continuous and has left limits at any finite instant. If M is right-continuous in probability, then clearly $M_t = M'_t$ nearly for all t. The same is true when the filtration is right-continuous. To see this notice that $M_q \to M'_t$ in $\| \|_1$-mean as $\mathbb{Q} \ni q \downarrow t$, since the collection $\{M_q : q \in \mathbb{Q}, t < q < t+1\}$ is uniformly integrable (example 2.5.2 and theorem A.8.6); both M_t and M'_t are measurable on $\mathcal{F}_t = \bigcap_{\mathbb{Q} \ni q > t} \mathcal{F}_q$ and have the same integral over every set $A \in \mathcal{F}_t$:

$$\int_A M_t \, d\mathbb{P} = \int_A M_q \, d\mathbb{P} \xrightarrow[t<q\to t]{} \int_A M'_t \, d\mathbb{P} \, .$$

That is to say, M' is a modification of M.

Consider the case that M is L^1-bounded. Then we can take ∞ for the time u of the proof and see that the paths of M that have an oscillatory discontinuity anywhere, including at ∞, are negligible. In other words, then

$$M'_\infty \stackrel{\text{def}}{=} \lim_{t\uparrow\infty} M'_t$$

exists almost surely.

A ***local martingale*** is, of course, a process that is locally a martingale. Localizing with a sequence T_n of stopping times that reduce M to uniformly integrable martingales, we arrive at the following conclusion:

Proposition 2.5.13 *Let M be a local $\mathbb{P}$-martingale on the filtration $\mathcal{F}_.$. If M is right-continuous in probability or $\mathcal{F}_.$ is right-continuous, then M has a modification adapted to the $\mathbb{P}$-regularization $\mathcal{F}^{\mathbb{P}}_.$, one all of whose paths are right-continuous and have left limits. If M is L^1-bounded, then this modification has almost surely a limit $M_\infty \in \overline{\mathbb{R}}$ at infinity.*

Exercise 2.5.14 M also has a modification that is adapted to $\mathcal{F}_.$ and whose paths are nearly right-continuous. If M is uniformly integrable, then it is L^1-bounded and is a modification of the martingale M^g of example 2.5.2, where $g = M_\infty$.

Example 2.5.15 The positive martingale M_t of example 1.3.45 on page 41 converges at every point, the limit being $+\infty$ at zero and zero elsewhere. It is L^1-bounded but not uniformly integrable, and its value at time t is not $\mathbb{E}[M_\infty|\mathcal{F}_t]$.

Exercise 2.5.16 Doob considered martingales first in discrete time: let $\{\mathcal{F}_n : n \in \mathbb{N}\}$ be an increasing collection of sub-σ-algebras of $\mathcal{F}$. The random variables M_n, $n \in \mathbb{N}$ form a ***martingale*** on $\{\mathcal{F}_n\}$ if $\mathbb{E}[F_{n+1}|\mathcal{F}_n] = F_n$ almost surely for all $n \geq 1$. He developed the upcrossing argument to show that an L^1-bounded martingale converges almost surely to a limit in the extended reals $\overline{\mathbb{R}}$ as $n \to \infty$, in $\mathbb{R}$ if $\{M_n\}$ is uniformly integrable.

Exercise 2.5.17 (A Strong Law of Large Numbers) The previous exercise allows a simple proof of an uncommonly strong version of the strong law of large numbers. Namely, let $F_1, F_2, \ldots$ be a sequence of square integrable random variables that all have expectation p and whose variances all are bounded by some σ^2. Assume that the conditional expectation of F_{n+1} given $F_1, F_2, \ldots, F_n$ equals p as well, for $n = 1, 2, 3, \ldots$. [To paraphrase: knowledge of previous executions of the experiment may influence the law of its current replica only to the extent that the expectation does not change and the variance does not increase overly much.] Then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{\nu=1}^{n} F_\nu = p$$

almost surely. See exercise 4.2.14 for a generalization to the case that the F_n merely have bounded moments of order q for some $q > 1$ and [81] for the case that the random variables F_n are merely orthogonal in L^2.

Boundedness of the Paths

Lemma 2.5.18 (Doob's Maximal Lemma) *Let M be a right-continuous martingale. Then at any instant t and for any $\lambda > 0$*

$$\mathbb{P}\Big[M_t^\star > \lambda\Big] \le \frac{1}{\lambda}\int_{[M_t^\star>\lambda]} |M_t| \, d\mathbb{P} \le \frac{1}{\lambda}\cdot \mathbb{E}[|M_t|] \, .$$

Proof. Let $S = \{s_0 < s_1 < \ldots < s_N\}$ be a finite set of rationals contained in the interval $[0,t]$, let $u > t$, and set

$$M^S = \sup_{s\in S} |M_s| \quad \text{and} \quad U = \inf\big\{s \in S : \ |M_s| > \lambda\big\} \wedge u \, .$$

Clearly U is an elementary stopping time and $|M_U| = |M_{U\wedge t}| > \lambda$ on

$$[U < u] = [U \le t] = \Big[M^S > \lambda\Big] \subset \big[M_t^\star > \lambda\big] \, .$$

Therefore
$$1_{\left[M^S>\lambda\right]} \le \frac{|M_U|}{\lambda}\cdot 1_{\left[U\le t\right]} \in \mathcal{F}_t \, .$$

We apply the expectation; since $|M|$ is a submartingale,

$$\mathbb{P}\Big[M^S > \lambda\Big] \le \lambda^{-1}\cdot \int_{[U\le t]} |M_U| \, d\mathbb{P} = \lambda^{-1}\cdot \int_{[U\le t]} |M_{U\wedge t}| \, d\mathbb{P}$$

by corollary 2.5.11:
$$\le \lambda^{-1}\cdot \int_{[U\le t]} \mathbb{E}\big[|M_t| \ \big| \ \mathcal{F}_{U\wedge t}\big] \, d\mathbb{P}$$

$$= \lambda^{-1}\cdot \int_{[M^S>\lambda]} |M_t| \, d\mathbb{P} \le \lambda^{-1}\cdot \int_{[M_t^\star>\lambda]} |M_t| \, d\mathbb{P} \, .$$

We take the supremum over all finite subsets S of $\{t\} \cup (\mathbb{Q} \cap [0,t])$ and use the right-continuity of M: Doob's inequality follows. ∎

Theorem 2.5.19 (Doob's Maximal Theorem) *Let M be a right-continuous martingale on $(\Omega, \mathcal{F}_., \mathbb{P})$ and p, p' conjugate exponents, that is to say, $1 \le p, p' \le \infty$ and $1/p + 1/p' = 1$. Then*

$$\| M_\infty^\star \|_{L^p(\mathbb{P})} \le p' \cdot \sup_t \| M_t \|_{L^p(\mathbb{P})} \, .$$

Proof. If $p = 1$, then $p' = \infty$ and the inequality is trivial; if $p = \infty$, then it is obvious. In the other cases consider an instant $t \in (0,\infty)$ and resurrect the finite set $S \subset [0,t]$ and the random variable M^S from the previous proof. From equation (A.3.9) and lemma 2.5.18,

$$\begin{aligned}\int (M^S)^p \, d\mathbb{P} &= p\int_0^\infty \lambda^{p-1}\mathbb{P}[M^S > \lambda]\, d\lambda \\ &\le p\int_0^\infty \int \lambda^{p-2}|M_t| \cdot [M^S > \lambda]\, d\mathbb{P}\, d\lambda \\ &= \frac{p}{p-1}\int |M_t|(M^S)^{p-1}\, d\mathbb{P}\,.\end{aligned}$$

by A.8.4: $$\int (M^S)^p \, d\mathbb{P} \le \frac{p}{p-1} \cdot \left(\int |M_t|^p \, d\mathbb{P}\right)^{\frac{1}{p}} \cdot \left(\int (M^S)^p \, d\mathbb{P}\right)^{\frac{p-1}{p}} .$$

Now $\int (M^S)^p \, d\mathbb{P}$ is finite if $\int |M_t|^p \, d\mathbb{P}$ is, and we may divide by the second factor on the right to obtain

$$\left\| M^S \right\|_{L^p(\mathbb{P})} \le p' \cdot \left\| M_t \right\|_{L^p(\mathbb{P})} .$$

Taking the supremum over all finite subsets S of $\{t\} \cup (\mathbb{Q} \cap [0,t])$ and using the right-continuity of M produces $\|M^\star_t\|_{L^p(\mathbb{P})} \le p' \cdot \|M_t\|_{L^p(\mathbb{P})}$. Now let $t \to \infty$. ∎

Exercise 2.5.20 For a vector $\boldsymbol{M} = (M^1, \ldots, M^d)$ of right-continuous martingales set

$$|\boldsymbol{M}_t|_\infty = \|\boldsymbol{M}_t\|_{\ell^\infty} \stackrel{\text{def}}{=} \sup_\eta |M^\eta_t| \quad \text{and} \quad \boldsymbol{M}^\star_t \stackrel{\text{def}}{=} \sup_{s \le t} |\boldsymbol{M}_s|_\infty .$$

Using exercise 2.5.8 and the observation that the proofs above use only the property of $|M|$ of being a positive submartingale, show that

$$\|\boldsymbol{M}^\star_\infty\|_{L^p(\mathbb{P})} \le p' \cdot \sup_{t<\infty} \big\| |\boldsymbol{M}_t|_\infty \big\|_{L^p(\mathbb{P})} .$$

Exercise 2.5.21 (i) For a standard Wiener process W and $\alpha, \beta \in \mathbb{R}_+$,

$$\mathbb{P}\big[\sup_t (W_t - \alpha t/2) > \beta\big] \le e^{-\alpha\beta} .$$

(ii) $\lim_{t\to\infty} W_t/t = 0$.

Doob's Optional Stopping Theorem

In support of the vague principle "what holds at instants t holds at stopping times T," which the reader might be intuiting by now, we offer this generalization of the martingale property:

Theorem 2.5.22 (Doob) *Let M be a right-continuous uniformly integrable martingale. Then $\mathbb{E}[M_\infty|\mathcal{F}_T] = M_T$ almost surely at any stopping time T.*

Proof. We know from exercise 2.5.14 that $M_t = \mathbb{E}\Big[M_\infty \Big| \mathcal{F}_t\Big]$ for all t. To start with, assume that T takes only countably many values $0 \le t_0 \le t_1 \le \ldots$, among them possibly the value $t_\infty = \infty$. Then for any $A \in \mathcal{F}_T$

$$\int_A M_\infty \, d\mathbb{P} = \sum_{0\le k\le\infty} \int_{A\cap[T=t_k]} M_\infty \, d\mathbb{P} = \sum_{0\le k\le\infty} \int_{A\cap[T=t_k]} M_{t_k} \, d\mathbb{P}$$
$$= \sum_{0\le k\le\infty} \int_{A\cap[T=t_k]} M_T \, d\mathbb{P} = \int_A M_T \, d\mathbb{P} \, .$$

The claim is thus true for such T. Given an arbitrary stopping time T we apply this to the discrete-valued stopping times $T^{(n)}$ of exercise 1.3.20. The right-continuity of M implies that

$$M_T = \lim_n \mathbb{E}[\, M_\infty | \mathcal{F}_{T^{(n)}} \,] \, .$$

This limit exists in mean, and the integral of it over any set $A \in \mathcal{F}_T$ is the same as the integral of M_∞ over A (exercise A.3.27). ∎

Exercise 2.5.23 (i) Let M be a right-continuous uniformly integrable martingale and $S \le T$ any two stopping times. Then $M_S = \mathbb{E}[M_T|\mathcal{F}_S]$ almost surely.
(ii) If M is a right-continuous martingale and T a stopping time, then the stopped process M^T is a martingale; if T is bounded, then M^T is uniformly integrable.
(iii) A local martingale is locally uniformly integrable. (iv) A positive local martingale M is a supermartingale; if $\mathbb{E}[M_t]$ is constant, M is a martingale. In any case, if $\mathbb{E}[M_S] = \mathbb{E}[M_T] = \mathbb{E}[M_0]$, then $\mathbb{E}[M_{S\vee T}] = \mathbb{E}[M_0]$.

Martingales Are Integrators

A simple but pivotal result is this:

Theorem 2.5.24 *A right-continuous square integrable martingale M is an L^2-integrator whose size at any instant t is given by*

$$\lceil\!\lceil M^t \rceil\!\rceil_{\mathcal{I}^2} = \| M_t \|_2 \, . \tag{2.5.1}$$

Proof. Let X be an elementary integrand as in (2.1.1):

$$X = f_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f_n \cdot (\!(t_n, t_{n+1}]\!] \, , \qquad 0 = t_1 < \ldots, \; f_n \in \mathcal{F}_{t_n} \, ,$$

that vanishes past time t. Then

$$\Big(\int X \, dM\Big)^2 = \Big(f_0 M_0 + \sum_{n=1}^{N} f_n \cdot \big(M_{t_{n+1}} - M_{t_n}\big)\Big)^2$$
$$= f_0^2 M_0^2 + 2 f_0 M_0 \cdot \sum_{n=1}^{N} f_n \cdot \big(M_{t_{n+1}} - M_{t_n}\big)$$
$$+ \sum_{m,n=1}^{N} f_m (M_{t_{m+1}} - M_{t_m}) \cdot f_n (M_{t_{n+1}} - M_{t_n}) \, . \tag{$*$}$$

If $m \neq n$, say $m < n$, then $f_m(M_{t_{m+1}} - M_{t_m}) \cdot f_n$ is measurable on $\mathcal{F}_{t_n}$. Upon taking the expectation in $(*)$, terms with $m \neq n$ will vanish. At this point our particular choice of the elementary integrands pays off: had we allowed the steps to be measurable on a σ-algebra larger than the one attached to the left endpoint of the interval of constancy, then $f_n = X_{t_n}$ would not be measurable on $\mathcal{F}_{t_n}$, and the cancellation of terms would not occur. As it is we get

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int X\,dM\Big)^2\Big] &= \mathbb{E}\Big[f_0^2 M_0^2 + \sum_{n=1}^{N} f_n^2 \cdot (M_{t_{n+1}} - M_{t_n})^2\Big] \\
&\le \mathbb{E}\Big[M_0^2 + \textstyle\sum_n \big(M_{t_{n+1}} - M_{t_n}\big)^2\Big] \\
&= \mathbb{E}\Big[M_0^2 + \textstyle\sum_n \big(M_{t_{n+1}}^2 - 2M_{t_{n+1}}M_{t_n} + M_{t_n}^2\big)\Big] \\
\text{by exercise 2.5.3:} \quad &= \mathbb{E}\Big[M_0^2 + \textstyle\sum_{n=1}^{N} \big(M_{t_{n+1}}^2 - M_{t_n}^2\big)\Big] = \mathbb{E}\big[M_{t_{N+1}}^2\big] \le \mathbb{E}[M_t^2]\,.
\end{aligned}$$

Taking the square root and the supremum over elementary integrands X that do not exceed[2] $[\![0,t]\!]$ results in equation (2.5.1). ∎

Exercise 2.5.25 If W is a standard Wiener process on the filtration $\mathcal{F}_.$, then it is an L^2-integrator on $\mathcal{F}_.$ and on its natural enlargement, and for every elementary integrand $X \ge 0$

$$[\![X]\!]_{W\text{–}2} = \|X\|_{W\text{–}2} = \Big(\int\!\!\int X_s^2\,ds\,d\mathbb{P}\Big)^{1/2}\,.$$

In particular $\quad \lceil W^t \rceil_{\mathcal{I}^2} = \sqrt{t}\,.$ (For more see 4.2.20.)

Example 2.5.26 Let $X_1, X_2, \ldots$ be independent identically distributed Bernoulli random variables with $\mathbb{P}[X_k = \pm 1] = 1/2$. Fix a (large) natural number n and set

$$Z_t = \frac{1}{\sqrt{n}} \sum_{k \le tn} X_k \qquad 0 \le t \le 1\,.$$

This process is right-continuous and constant on the intervals $[k/n, (k+1)/n)$, as is its basic filtration. Z is a process of finite variation. In fact, its variation process clearly is

$$\lceil Z \rceil_t = \frac{1}{\sqrt{n}} \cdot \lfloor tn \rfloor \quad \approx t \cdot \sqrt{n}\,.$$

Here $\lfloor r \rfloor$ denotes the largest integer less than or equal to r. Thus if we estimate the size of Z as an L^2-integrator through its variation, using proposition 2.4.1 on page 68, we get the following estimate:

$$\lceil Z^t \rceil_{\mathcal{I}^2} \le t\sqrt{n}\,. \tag{v}$$

Z is also evidently a martingale. Also, the L^2-mean of Z_t is easily seen to be $\sqrt{\lfloor tn \rfloor / n} \le \sqrt{t}$. Theorem 2.5.24 yields the much superior estimate

$$\lceil Z^t \rceil_{\mathcal{I}^2} \le \sqrt{t}\,, \tag{m}$$

which is, in particular, independent of n.

Let us use this example to continue the discussion of remark 1.2.9 on page 18 concerning the driver of Brownian motion. Consider a point mass on the line that receives at the instants k/n a kick of momentum $p_0 X_k$, i.e., either to the right or to the left with probability $1/2$ each. Let us scale the units so that the total energy transfer up to time 1 equals 1. An easy calculation shows that then $p_0 = 1/\sqrt{n}$. Assume that the point mass moves through a viscous medium. Then we are led to the stochastic differential equation

$$\begin{pmatrix} dx_t \\ dp_t \end{pmatrix} = \begin{pmatrix} p_t/m\; dt \\ -\alpha\, p_t\; dt \;+\; dZ_t \end{pmatrix}, \tag{2.5.2}$$

just as in equation (1.2.1). If we are interested in the solution at time 1, then the pertinent probability space is finite. It has 2^n elements. So the problem is to solve finitely many ordinary differential equations and to assemble their statistics. Imagine that n is on the order of 10^{23}, the number of molecules per mole. Then 2^n far exceeds the number of elementary particles in the universe! This makes it impossible to do the computations, and the estimates toward any procedure to solve the equation become useless if inequality (v) is used. Inequality (m) offers much better prospects in this regard but necessitates the development of stochastic integration theory.

An aside: if dt is large as compared with $1/n$, then $dZ_t = Z_{t+dt} - Z_t$ is the superposition of a large number of independent Bernoulli random variables and thus is distributed approximately $N(0, dt)$. It can be shown that Z tends to a Wiener process in law as $n \to \infty$ (theorem A.4.9) and that the solution of equation (2.5.2) accordingly tends in law to the solution of our idealized equation (1.2.1) for physical Brownian motion (see exercise A.4.14).

Martingales in L^p

The question arises whether perhaps a p-integrable martingale M is an L^p-integrator for exponents p other than 2. This is true in the range $1 < p < \infty$ (theorem 2.5.30) but not in general at $p = 1$, where M can only be shown to be a *local* L^1-integrator. For the proof of these claims some estimates are needed:

Lemma 2.5.27 *(i) Let Z be a bounded adapted process and set*

$$\lambda = \sup |Z| \quad \textit{and} \quad \mu = \sup \Big\{ \mathbb{E}\Big[\int X\, dZ \Big] : \; X \in \mathcal{E}_1 \Big\} .$$

Then for all X in the unit ball $\mathcal{E}_1$ of $\mathcal{E}$

$$\mathbb{E}\Big[\Big| \int X\, dZ \Big| \Big] \le \sqrt{2} \cdot (\lambda + \mu) . \tag{2.5.3}$$

*In other words, Z has global L^1-integrator size $\lceil Z \rceil_{\mathcal{I}^1} \le \sqrt{2} \cdot (\lambda + \mu)$. Inequality (2.5.3) holds if $\mathbb{P}$ is merely a **subprobability:** $0 \le \mathbb{P}[\Omega] \le 1$.*

(ii) Suppose Z is a positive bounded supermartingale. Then for all $X \in \mathcal{E}_1$

$$\mathbb{E}\Big[\Big|\int X\, dZ\Big|^2\Big] \le 8 \cdot \sup Z \cdot \mathbb{E}[Z_0] \,. \tag{2.5.4}$$

That is to say, Z has global L^2-integrator size $\lceil Z \rceil_{\mathcal{I}^2} \le 2\sqrt{2 \sup Z \cdot \mathbb{E}[Z_0]}$.

Proof. It is easiest to argue if the elementary integrand $X \in \mathcal{E}_1$ of the claims (2.5.3) and (2.5.4) is written in the form (2.1.1) on page 46:

$$X = f_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f_n \cdot (\!(t_n, t_{n+1}]\!] \,, \qquad 0 = t_1 < \ldots < t_{N+1},\ f_n \in \mathcal{F}_{t_n} \,.$$

Since X is in the unit ball $\mathcal{E}_1 \stackrel{\text{def}}{=} \{X \in \mathcal{E} : |X| \le 1\}$ of $\mathcal{E}$, the f_n all have absolute value less than 1. For $n = 1, \ldots, N$ let

$$\zeta_n \stackrel{\text{def}}{=} Z_{t_{n+1}} - Z_{t_n} \quad \text{and} \quad Z'_n \stackrel{\text{def}}{=} \mathbb{E}\Big[Z_{t_{n+1}} \Big| \mathcal{F}_{t_n}\Big] \,;$$

$$\widehat{\zeta}_n \stackrel{\text{def}}{=} \mathbb{E}\Big[\zeta_n \Big| \mathcal{F}_{t_n}\Big] = Z'_n - Z_{t_n} \quad \text{and} \quad \widetilde{\zeta}_n \stackrel{\text{def}}{=} \zeta_n - \widehat{\zeta}_n = Z_{t_{n+1}} - Z'_n \,.$$

Then
$$\begin{aligned}
\int X\, dZ &= f_0 \cdot Z_0 + \sum_{n=1}^{N} f_n \cdot \big(Z_{t_{n+1}} - Z_{t_n}\big) = f_0 \cdot Z_0 + \sum_{n=1}^{N} f_n \cdot \zeta_n \\
&= \Big(f_0 \cdot Z_0 + \sum_{n=1}^{N} f_n \cdot \widetilde{\zeta}_n\Big) + \Big(\sum_{n=1}^{N} f_n \cdot \widehat{\zeta}_n\Big) \\
&= M + V \,.
\end{aligned}$$

The L^1-means of the two terms can be estimated separately. We start on M. Note that $\mathbb{E}\big[f_m \widetilde{\zeta}_m \cdot f_n \widetilde{\zeta}_n\big] = 0$ if $m \ne n$ and compute

$$\begin{aligned}
\mathbb{E}[M^2] &= \mathbb{E}\Big[f_0^2 \cdot Z_0^2 + \sum_{n=1}^{N} f_n^2 \cdot \widetilde{\zeta}_n^2\Big] \le \mathbb{E}\Big[Z_0^2 + \sum_{n=1}^{N} \widetilde{\zeta}_n^2\Big] \\
&= \mathbb{E}\Big[Z_{t_1}^2 + \textstyle\sum_n \big(Z_{t_{n+1}} - Z'_n\big)^2\Big] \\
&= \mathbb{E}\Big[Z_{t_1}^2 + \textstyle\sum_n \Big(Z_{t_{n+1}}^2 - 2 Z_{t_{n+1}} Z'_n + Z'^2_n\Big)\Big] \\
&= \mathbb{E}\Big[Z_{t_1}^2 + \textstyle\sum_n \Big(Z_{t_{n+1}}^2 - Z'^2_n\Big)\Big] \\
&= \mathbb{E}\Big[Z_{t_1}^2 + \textstyle\sum_n \Big(Z_{t_{n+1}}^2 - Z_{t_n}^2\Big) + \textstyle\sum_n \big(Z_{t_n}^2 - Z'^2_n\big)\Big] \\
&= \mathbb{E}\Big[Z_{t_{N+1}}^2 + \textstyle\sum_n \big(Z_{t_n} + Z'_n\big) \cdot \big(Z_{t_n} - Z'_n\big)\Big] \\
&= \mathbb{E}\Big[Z_{t_{N+1}}^2 + \textstyle\sum_n \big(Z_{t_n} + Z'_n\big) \cdot \big(Z_{t_n} - Z_{t_{n+1}}\big)\Big] \\
&= \mathbb{E}\Big[Z_{t_{N+1}}^2\Big] - \mathbb{E}\Big[\int \Big(\textstyle\sum_{n=1}^{N} \big(Z_{t_n} + Z'_n\big) \cdot (\!(t_n, t_{n+1}]\!]\Big)\, dZ\Big] \\
&\le \lambda^2 + 2\lambda\mu \,.
\end{aligned} \tag{*}$$

After this preparation let us prove (i). Since $\mathbb{P}$ has mass less than 1, (*) results in

$$\mathbb{E}[|M|] \le \big(\mathbb{E}[M^2]\big)^{1/2} \le \sqrt{\lambda^2 + 2\lambda\mu}\ .$$

We add the estimate of the expectation of

$$|V| \le \sum_n \big|f_n \widehat{\zeta}_n\big| = \sum_n |f_n| \operatorname{sgn}\big(\widehat{\zeta}_n\big) \cdot \widehat{\zeta}_n :$$

$$\begin{aligned}
\mathbb{E}|V| &\le \mathbb{E}\Big[\sum_{n=1}^N |f_n| \operatorname{sgn}\big(\widehat{\zeta}_n\big) \cdot \widehat{\zeta}_n\Big] = \mathbb{E}\Big[\sum_{n=1}^N |f_n| \operatorname{sgn}\big(\widehat{\zeta}_n\big) \cdot \zeta_n\Big] \\
&= \mathbb{E}\Big[\int \sum_{n=1}^N |f_n| \operatorname{sgn}\big(\widehat{\zeta}_n\big) \cdot (\!(t_n, t_{n+1}]\!]\, dZ\Big] \le \mu
\end{aligned}$$

to get $\quad \mathbb{E}\Big[\Big|\int X\ dZ\Big|\Big] \le \sqrt{\lambda^2 + 2\lambda\mu} + \mu \le \sqrt{2} \cdot (\lambda + \mu)\ .$

We turn to claim (ii). Pick a $u > t_{N+1}$ and replace Z by $Z \cdot [\![0, u)\!)$. This is still a positive bounded supermartingale, and the left-hand side of inequality (2.5.4) has not changed. Since $X = 0$ on $(\!(t_{N+1}, u]\!]$, renaming the t_n so that $t_{N+1} = u$ does not change it either, so we may for convenience assume that $Z_{t_{N+1}} = 0$. Continuing (*) we find, using proposition 2.5.10, that

$$\begin{aligned}
\mathbb{E}[M^2] &\le -\mathbb{E}\Big[\int_{(\!(0, t_{N+1}]\!]} 2\lambda\, dZ\Big] \\
&= 2\lambda \cdot \mathbb{E}\Big[Z_0 - Z_{t_{N+1}}\Big] = 2 \sup Z \cdot \mathbb{E}\Big[Z_0\Big]\ . \qquad (**)
\end{aligned}$$

To estimate $\mathbb{E}[V^2]$ note that the $\widehat{\zeta}_n$ are all negative: $\Big|\sum_n f_n \cdot \widehat{\zeta}_n\Big|$ is largest when all the f_n have the same sign. Thus, since $-1 \le f_n \le 1$,

$$\begin{aligned}
\mathbb{E}[V^2] &= \mathbb{E}\Big[\Big(\sum_{n=1}^N f_n \cdot \widehat{\zeta}_n\Big)^2\Big] \le \mathbb{E}\Big[\Big(\sum_{n=1}^N \widehat{\zeta}_n\Big)^2\Big] \\
&\le 2 \sum_{1 \le m \le n \le N} \mathbb{E}\big[\widehat{\zeta}_m \cdot \widehat{\zeta}_n\big] = 2 \sum_{1 \le m \le n \le N} \mathbb{E}\big[\widehat{\zeta}_m \cdot \zeta_n\big] \\
&= 2 \sum_{1 \le m \le N} \mathbb{E}\Big[\widehat{\zeta}_m \cdot \big(Z_{t_{N+1}} - Z_{t_m}\big)\Big] = 2 \sum_{1 \le m \le N} \mathbb{E}\big[-\widehat{\zeta}_m \cdot Z_{t_m}\big] \\
&\le 2 \sup Z \sum_{1 \le m \le N} \mathbb{E}\big[-\widehat{\zeta}_m\big] = -2 \sup Z \sum_{1 \le m \le N} \mathbb{E}\big[\zeta_m\big] \\
&= -2 \sup Z \cdot \big(Z_{t_{N+1}} - Z_0\big) = 2 \sup Z \cdot \mathbb{E}\big[Z_0\big]\ .
\end{aligned}$$

Adding this to inequality (**) we find

$$\mathbb{E}\Big[\Big|\int X\ dZ\Big|^2\Big] \le 2\mathbb{E}[M^2] + 2\mathbb{E}[V^2] \le 8 \cdot \sup Z \cdot \mathbb{E}[Z_0]\ . \qquad \blacksquare$$

The following consequence of lemma 2.5.27 is the first step in showing that p-integrable martingales are L^p-integrators in the range $1 < p < \infty$ (theorem 2.5.30). It is a "weak-type" version of this result at $p = 1$:

Proposition 2.5.28 *An L^1-bounded right-continuous martingale M is a global L^0-integrator. In fact, for every elementary integrand X with $|X| \le 1$ and every $\lambda > 0$,*

$$\mathbb{P}\Big[\Big|\int X\, dM\Big| > \lambda\Big] \le \frac{2}{\lambda} \cdot \sup_t \|M_t\|_{L^1(\mathbb{P})} . \tag{2.5.5}$$

Proof. This inequality clearly implies that the linear map $X \mapsto \int X\, dM$ is bounded from $\mathcal{E}$ to L^0, in fact to the Lorentz space $L^{1,\infty}$. The argument is again easiest if X is written in the form (2.1.1):

$$X = f_0 \cdot [\![0]\!] + \sum_{n=1}^{N} f_n \cdot (\!(t_n, t_{n+1}]\!] , \qquad 0 = t_1 < \ldots,\ f_n \in \mathcal{F}_{t_n} .$$

Let U be a bounded stopping time strictly past t_{N+1}, and let us assume to start with that M is positive at and before time U. Set

$$T = \inf\{t_n : M_{t_n} \ge \lambda\} \wedge U .$$

This is an elementary stopping time (proposition 1.3.13). Let us estimate the probabilities of the disjoint events

$$B_1 = \Big[\Big|\int X\, dM\Big| > \lambda,\, T < U\Big] \quad\text{and}\quad B_2 = \Big[\Big|\int X\, dM\Big| > \lambda,\, T = U\Big]$$

separately. B_1 is contained in the set $M^\star_U \ge \lambda$, and Doob's maximal lemma 2.5.18 gives the estimate

$$\mathbb{P}[B_1] \le \lambda^{-1} \cdot \mathbb{E}\big[|M_U|\big] . \tag{$*$}$$

To estimate the probability of B_2 consider the right-continuous process

$$Z = M \cdot [\![0, T)\!) .$$

This is a positive supermartingale bounded by λ; indeed, using A.1.5,

$$\begin{aligned}\mathbb{E}\big[Z_t|\mathcal{F}_s\big] &= \mathbb{E}\Big[M_t \cdot [T > t]\Big|\mathcal{F}_s\Big] \\ &\le \mathbb{E}\Big[M_t \cdot [T > s]\Big|\mathcal{F}_s\Big] = \mathbb{E}\Big[M_s \cdot [T > s]\Big|\mathcal{F}_s\Big] = Z_s .\end{aligned}$$

On B_2 the paths of M and Z coincide. Therefore $\int X\, dZ = \int X\, dM$ on B_2, and B_2 is contained in the set

$$\Big[\Big|\int X\, dZ\Big| > \lambda\Big] .$$

Due to Chebyschev's inequality and lemma 2.5.27, the probability of this set is less than

$$\lambda^{-2} \cdot \mathbb{E}\Big[\Big(\int X \, dZ\Big)^2\Big] \le \frac{8\lambda \cdot \mathbb{E}[Z_0]}{\lambda^2} = \frac{8\mathbb{E}[M_0]}{\lambda} \le \frac{8}{\lambda} \cdot \mathbb{E}[|M_U|] \ .$$

Together with (*) this produces

$$\mathbb{P}\Big[\Big|\int X \, dM\Big| > \lambda\Big] \le \frac{9}{\lambda} \cdot \mathbb{E}[M_U] \ .$$

In the general case we split M_U into its positive and negative parts $M_U^{\pm}$ and set $M_t^{\pm} = \mathbb{E}\big[M_U^{\pm}|\mathcal{F}_t\big]$, obtaining two positive martingales with difference M^U. We estimate

$$\begin{aligned}\mathbb{P}\Big[\Big|\int X \, dM\Big| \ge \lambda\Big] &\le \mathbb{P}\Big[\Big|\int X \, dM^+\Big| \ge \lambda/2\Big] + \mathbb{P}\Big[\Big|\int X \, dM^-\Big| \ge \lambda/2\Big] \\ &\le \frac{9}{\lambda/2} \cdot \Big(\mathbb{E}[M_U^+] + \mathbb{E}[M_U^-]\Big) = \frac{18}{\lambda} \cdot \mathbb{E}[|M_U|] \\ &\le \frac{18}{\lambda} \cdot \sup_t \mathbb{E}[|M_t|] \ .\end{aligned}$$

This is inequality (2.5.5), except for the factor of $1/\lambda$, which is 18 rather than 2, as claimed. We borrow the latter value from Burkholder [14], who showed that the following inequality holds and is best possible: for $|X| \le 1$

$$\mathbb{P}\Big[\sup_t \Big|\int_0^t X \, dM\Big| > \lambda\Big] \le \frac{2}{\lambda} \cdot \sup_t \| M_t \|_{L^1} \ . \qquad \blacksquare$$

The proof above can be used to get additional information about local martingales:

Corollary 2.5.29 *A right-continuous local martingale M is a local L^1-integrator. In fact, it can locally be written as the sum of an L^2-integrator and a process of integrable total variation.* (According to exercise 4.3.14, M can actually be written as the sum of a finite variation process and a locally square integrable local ***martingale***.)

Proof. There is an arbitrarily large bounded stopping time U such that M^U is a uniformly integrable martingale and can be written as the difference of two positive martingales $M^{\pm}$. Both can be chosen right-continuous (proposition 2.5.13). The stopping time $T = \inf\{t : M_t^{\pm} \ge \lambda\} \wedge U$ can be made arbitrarily large by the choice of λ. Write

$$(M^{\pm})^T = M^{\pm} \cdot [\![0, T)\!) + M_T^{\pm} \cdot [\![T, \infty)\!) \ .$$

The first summand is a positive bounded supermartingale and thus is a global $L^2(\mathbb{P})$-integrator; the last summand evidently has integrable total variation $|M_T^{\pm}|$. Thus M^T is the sum of two global $L^2(\mathbb{P})$-integrators and two processes of integrable total variation. ∎

Theorem 2.5.30 *Let $1 < p < \infty$. A right-continuous L^p-integrable martingale M is an L^p-integrator. Moreover, there are universal constants A_p independent of M such that for all stopping times T*

$$\lceil\!\lceil M^T \rceil\!\rceil_{\mathcal{I}^p} \le A_p \cdot \| M_T \|_p \,. \tag{2.5.6}$$

Proof. Let X be an elementary integrand with $|X| \le 1$ and consider the following linear map U from $L^\infty(\mathcal{F}_\infty, \mathbb{P})$ to itself:

$$U(g) = \int X \, dM^g \,.$$

Here M^g is the right-continuous martingale $M_t^g = \mathbb{E}[g|\mathcal{F}_t]$ of example 2.5.2. We shall apply Marcinkiewicz interpolation to this map (see proposition A.8.24). By (2.5.5) U is of weak type 1–1:

$$\mathbb{P}\big[|U(g)| > \lambda\big] \le \frac{2}{\lambda} \cdot \| g \|_1 \,.$$

By (2.5.1), U is also of strong type 2–2:

$$\| U(g) \|_2 \le \| g \|_2 \,.$$

Also, U is self-adjoint: for $h \in L^\infty$ and $X \in \mathcal{E}_1$ written as in (2.1.1)

$$\begin{aligned}
\mathbb{E}[U(g) \cdot h] &= \mathbb{E}\left[\left(f_0 M_0^g + \sum_n f_n \left(M_{t_{n+1}}^g - M_{t_n}^g\right)\right) M_\infty^h\right] \\
&= \mathbb{E}\left[f_0 M_0^g M_0^h + \sum_n f_n \left(M_{t_{n+1}}^g M_{t_{n+1}}^h - M_{t_n}^g M_{t_n}^h\right)\right] \\
&= \mathbb{E}\left[\left(f_0 M_0^h + \sum_n f_n \left(M_{t_{n+1}}^h - M_{t_n}^h\right)\right) M_\infty^g\right] \\
&= \mathbb{E}[U(h) \cdot g] \,.
\end{aligned}$$

A little result from Marcinkiewicz interpolation, proved as corollary A.8.25, shows that U is of strong type p–p for all $p \in (1, \infty)$. That is to say, there are constants A_p with $\| \int X \, dM \|_p \le A_p \cdot \| M_\infty \|_p$ for all elementary integrands X with $|X| \le 1$. Now apply this to the stopped martingale M^T to obtain (2.5.6). ∎

Exercise 2.5.31 Provide an estimate for A_p from this proof.

Exercise 2.5.32 Let S_t be a positive $\mathbb{P}$-supermartingale on the filtration $\mathcal{F}_{.}$ and assume that S is almost surely strictly positive; that is to say, $\mathbb{P}[S_t = 0] = 0 \quad \forall\, t$. Then there exists a $\mathbb{P}$-nearly empty set N outside which the restriction of every path of S to the positive rationals is bounded away from zero on every bounded time-interval.

Exercise 2.5.33 A right continuous local martingale M that is an L^1-integrator is actually a martingale. M is a global L^1-integrator if and only if $M^\star_\infty \in L^1$; then M is a uniformly integrable martingale.

Repeated Footnotes: 47^{2} 56^{5} 62^{6} 65^{8} 66^{11} 70^{14}