Contents

14 Martingale transforms, variation, integration **300**
14.1 Rademacher transform and the quadratic function 301
14.2 Quadratic function in continuous time 304
14.3 Local martingales 306
14.4 BDG inequalities revisited 307

15 Optional times **310**
15.1 Progressive stochastic processes 311
15.2 Optional sampling theorem 312
15.3 Exercises 315
15.4 References 316

16 Martingale transforms, variation, integration **317**
16.1 Rademacher transform and the quadratic function 317
16.2 Quadratic function in continuous time 320
16.3 Local martingales 322
16.4 BDG inequalities revisited 323

17 Introduction to semimartingales **326**
17.1 Preliminary definitions 326
17.2 Spaces and topologies 328
17.3 Exercises 331

18 Chaos representation of random functionals **332**
18.1 Tensor products 333
18.2 Tensor operations on tensor products 334
18.3 The problem of diagonals 335
18.4 Diagonal integrals 335
18.5 Recursion 337
18.6 Chaos polynomials 338
18.7 Generating functions 338
18.8 Examples of chaos polynomials 339

14 Martingale transforms, variation, integration

Consider a stochastic process X_t adapted to a filtration $\mathcal{F}_t$, $t \in T$. Let τ be an optimal time. Recall the stopped process

$$X_t^\tau \overset{\mathsf{df}}{=} X_{\tau \wedge t}. \tag{14.1}$$

Let $T = \mathbb{N}$ and denote the increments by $D_k = \Delta X_k = X_k - X_{k-1}$, and predictable weights by $v_k = 1\!\!1_{\{\tau \geq k\}}$. Then, we can rewrite the stopped process as a "*discrete-time integral*"

$$(v \cdot X)_t = \int_0^t v\, dX = \sum_{k=1}^t v_k \Delta X_k. \tag{14.2}$$

Clearly, not only for the specific 0-1 valued sequence v_k, but for an arbitrary integrable predictable sequence v_k, the formula preserves the martingale property. That is, if X is a martingale, so is $v \cdot X$. It is a fundamental question to ask which further properties of a martingale still appear in the transform, properties such as L^p-boundedness, uniform integrability, convergence, etc. In other words, this the question of continuity of the transform.

Turning the question around, is the transform continuous for non-martingales? In fact, a specific mode of continuity defines a new class of processes, sometimes similar to yet sometimes well beyond martingales. For example, consider a class of processes (somewhat resembling martingales) such that

$$X \text{ is bounded} \quad \Rightarrow \quad v \cdot X \text{ is bounded} \tag{14.3}$$

for some class of predictable sequences v_k (e.g., such that $|v_k| \leq C$), where boundedness could be defined with the help of the L^p-metric. Notice that the Orlicz Summation Theorem may be seen through the angle of above transform. That is, if a sequence (D_k) is summable (i.e., the series $\sum_k D_k$ converges unconditionally, i.e., for every permutation) in L^p, $0 \leq p \leq \infty$, then (16.3) holds for every $\{0, 1\}$ valued v_k, or, equivalently (when $p \geq 1$), whenever v_k are uniformly bounded non-random numbers. In other words, such sequence D_k defines a random vector measure on subsets of $\mathbb{N}$, and convergence of the transform with bounded coefficients corresponds to the integrability of a bounded function with respect to that random measure.

No surprise that procesess satisfying (16.3) with uniformly bounded predictable coefficients are called **semi-martingales**.

The next crucial question is how to extend the concept of transform (16.2) to a time-continuous process. In other words, this the question of a construction of stochastic integrals.

14.1 Rademacher transform and the quadratic function

Although the simplest case seems to be that of bounded deterministic coefficients, even $\{0, 1\}$- or $\{\pm 1\}$-valued, this is not so. Rather, we will begin with $v_k = r_k$, for a Rademacher sequence independent of the martingale $X_t, t \in T = \mathbb{N}$. Recall that all L^p norms or F-norms of a Rademacher series $\sum_k r_k d_k$ are comparable, where $0 \le p < \infty$. For two quantities A and B, depending on some parameter, let's write $A \asymp B$, if there is a constant $c > 0$ such that

$$\frac{1}{c} A \le B \le cA, \quad \text{uniformly with respect to the said parameter}$$

Similarly, we introduce the one-sided relation "$\prec$". So, by Fubini's Theorem, for any $p \in (0, \infty)$,

$$\Big\| \sum_n r_n D_n \Big\|_p \asymp \Big\| \Big(\mathsf{E}_r | \sum_n r_n D_n |^2 \Big)^{1/2} \Big\| = \Big\| \Big(\sum_n |D_n|^2 \Big)^{1/2} \Big\|$$

The latter series $Q(X)$ is called the **quadratic variation** of X, and can be seen as a nondecreasing process

$$Q_n = Q_n(X) = \Big(\sum_{k=1}^n |D_k|^2 \Big)^{1/2}$$

We will prove the main part ($p > 1$) of

Theorem 14.1 (Burkholder-Davis-Gundy: $p > 1$, Davis: $p = 1$) *For any martingale X_n and $p \in [1, \infty)$,*

$$\|Q(X)\|_p \asymp \|X^*\|_p \, .$$

Hence, every martingale with p-integrable maximum function is a semi-martingale in regard to the L^p-norm.

For the proof we need several lemmas. First, denote $\|X\| = \sup_n \mathsf{E}\,|X_n|$, and let

$$\tau = \tau_\lambda = \inf\{\, n \ge 1 : |X|_n > \lambda \,\} \, .$$

We often skip the subscript λ below, for typographic reasons.

Lemma 14.2 *Let X_n be a martingale or a nonnegative submartingale, and $\tau = \tau_\lambda$ for $\lambda > 0$. Then,*

$$\begin{aligned} &(1) \quad \mathsf{E}\, Q_{\tau-1}^2 + \mathsf{E}\, X_{\tau-1}^2 \le 2\mathsf{E}\, X_\tau X_{\tau-1} \le 2\lambda \|X\|; \\ &(2) \quad \lambda \mathsf{P}(Q > \lambda,\ X^* \le \lambda) \le 2\|X\|; \\ &(3) \quad \lambda \mathsf{P}(Q > \lambda) \le 3\|X\|. \end{aligned}$$

Proof. The second inequality in (1) is obvious. For any index k,

$$Q_{k-1}^2 + X_{k-1}^2 = 2X_k X_{k-1} - 2\sum_{j=1}^k X_{j-1} D_j.$$

If k is randomized, replaced by τ, the expectation will vanish the second term.

(2): Notice that $\{ X^* \le \lambda \} = \{ \tau_\lambda = \infty \}$. Then, by (1),

$$\lambda \mathsf{P}(Q > \lambda,\, X^* \le \lambda) \le \lambda \mathsf{P}(Q_{\tau-1} > \lambda) \le \frac{1}{\lambda} \mathsf{E}\, Q^2_{\tau-1} \le 2\|X\|.$$

(3) follows from (2) and from Doob's maximal tail inequality. ■

Lemma 14.3 *Let X_n be a martingale and $\lambda, \theta > 0$. Denote $Y_n = Q_n(\theta X) \vee X_n^*$. Then there is $\beta = \beta(\theta) > 1$ such that*

$$\lambda \mathsf{P}(Y_n > \beta\lambda) \le 3\mathsf{E}\,[X_n; Y_n > \lambda]. \tag{14.4}$$

Proof. The choice $\beta^2 = 1 + 2\theta^2$ will be sufficient. Define another stopping time

$$\sigma = \inf \{ n \ge 1 : Q_n(\theta X) > \lambda \}.$$

The following sequence is a positive submartingale, since the multiplier defined by the indicator increases with n:

$$Z_n = X_n 1\!\!1_{\{Q_n(\theta X) > \lambda\}}.$$

Note the inclusions

$$A \stackrel{\mathsf{df}}{=} \{ Q_n(\theta X) > \beta\lambda,\, X_n^* \le \lambda \} \subset \{ \sigma \le n,\, Z_n^* \le \lambda,\, |D_\sigma| \le \lambda \}$$

Hence, on A

$$\begin{aligned}
\lambda(1+2\theta^2) Q_n^2(\theta X) = &\ Q^2_{\sigma-1}(\theta X) + \theta^2 D_\sigma^2 + \theta^2 \sum_{j=\sigma+1}^{n} D_j^2 \\
\le &\ \lambda^2 + \theta^2\lambda^2 + \theta^2 \sum_{j=\sigma+1}^{n} (Z_j - Z_{j-1})^2 \\
\le &\ (1+\theta^2)\lambda^2 + \theta^2 Q_n^2(Z)
\end{aligned}$$

Since $A \subset Q_n(\theta X) > \beta\lambda$, we infer that $A \subset \{ Q_n(Z) > \lambda \}$. Now, applying (2) of the preceding lemma and Doob's maximal tail inequality to the submartingale Z, we obtain

$$\begin{aligned}
\lambda \mathsf{P}(Q_n(\theta X) \vee X_n^* > \beta\lambda) \le &\ \lambda \mathsf{P}(Q_n(\theta X) > \beta\lambda) + \lambda \mathsf{P}(X_n^* > \beta\lambda) \\
\le &\ \lambda \mathsf{P}(Q_n(Z) > \lambda) + \lambda \mathsf{P}(X_n^* > \lambda) \\
\le &\ 3\mathsf{E}\,[|X_n|; Y_n > \lambda]
\end{aligned}$$

■

We repeat the derivation of Doob's moment maximal inequality from Doob's tail maximal inequality. ■

Proof of BDG Theorem.

The proof of the left inequality $\|Q(X)\|_p \le C\|X\|_p$ is essentially contained in Exercise 16.1, since $Q(X) \le Y$. The right inequality follows by duality (cf. Exercise 16.2) ■

14.2 Quadratic function in continuous time

Let $\mathcal{F}_t$ be a filtration on $T = [0, \infty)$, and $\tau = (\tau_n)$ be a nondecreasing sequence of optional times diverging to ∞, with $\tau_0 = 0$. The first extension of the quadratic function of a martingale $X = (X_t)$ is defined as a nondecreasing process

$$B_t = B_t(X, \tau) = \sum_k \left| X_{t \wedge \tau_k} - X_{t \wedge \tau_{k-1}} \right|^2$$

To continue the construction, we will refine the random partitions induced by the times τ_n while assumming continuous paths of the martingale. The limit process can be intepreted as the quadratic variation. In fact, no continuous non-constant martingale will have paths of bounded variation.

Lemma 14.4 *Let X_t be a continuous martingale on $T = [0, \infty)$. If*

$$V(t) = \sup \left\{ \sum_{j=1}^{n} |X_{t_j} - X_{t_{j-1}}| : 0 = t_0 < t_1 < \ldots < t_n \leq t \right\} < \infty \ a.s.,$$

Then $X_t = X_0$ a.s.

Proof. First, if the variation is finite, we may assume that is bounded, by considering the stoppped martingale X^τ, where τ is the first hitting time of a level a by the process V_t.

The modulus of continuity

$$\rho_t(\epsilon) = \sup \{ |X_u - X_v| : |u - v| < \epsilon, \, u \vee v \leq t \} \to 0 \quad \text{a.s.}$$

when $\epsilon \to 0$. Then, choosing $\epsilon = n^{-1}$

$$\sum_{k=1}^{n} |X_{kt/n} - X_{(k-1)t/n}|^2 \leq V(t) \rho(n^{-1}) \leq a \, \rho(n^{-1})$$

By the BDG inequality (or using just the second moment)

$$\|X_t - X_0\|_p \leq C a \mathsf{E} \, \rho(n^{-1}) \to 0$$

Hence $X_t = X_0$. ■

For a fixed t, denoting

$$D_k = X_{t \wedge \tau_k} - X_{t \wedge \tau_{k-1}},$$

we derive the identity

$$X_t^2 = M_t + B_t, \tag{14.5}$$

with a certain martingale M_t, by twice telescoping the suitable sums:

$$
\begin{aligned}
X_t^2 &= \left(\sum_k D_k\right)^2 \\
&= \sum_j \sum_k D_j D_k \\
&= 2\sum_k \sum_{j=1}^{k-1} D_j\, D_k + \sum_k |D_k|^2 \\
&= 2\sum_k X_{t\wedge\tau_{k-1}} D_k + B_t \\
&= 2\sum_k X_{\tau_{k-1}} D_k + B_t,
\end{aligned}
$$

In the sum we recognize a martingale transform $M_t = (v \cdot Y)_t$ of a discrete time martingale Y with differences D_k and predictable coefficients $v_k = X_{\tau_{k-1}}$. Thus, we may apply the discrete-time BDG inequalities to it. Still, these coefficients define a predictable process

$$v_t = \sum_k v_k \mathbb{1}_{(\tau_{k-1},\tau_k]}(t).$$

Theorem 14.5 *Let X_t be a continuous martingale with $X^* \in L^p$, for some $p \geq 1$. Let $r \in \mathbb{N}$, and let $\tau_n = \tau_n^r$ be optional times defined by induction:*

$$\tau_{n+1} = \inf t > \tau_n : |X_t - X_{\tau_n}| = 2^{-r}.$$

Then the continuous martingales M_t^r defined above converge and in $\mathbb{M}_p$ to some continuous martingale M_t, as $r \to \infty$, and the decomposition formula (16.5) *holds, where B_t is a continuous nondecreasing process. Further, the decomposition is unique if $B_0 = 0$.*

Proof. In virtue of Exercise 16.3, it suffices to show that M_t^r is Cauchy. Let v^r denote the suitable predictable process introduced above. Then,

$$\sup_t \|v_t^r - X_t| \leq 2^{-r}$$

Therefore, for an integer $s \geq r$,

$$\sup_t \|v_t^r - v_t^s| \leq 2^{-r} + 2^{-s} \leq 2^{-r+1}.$$

Whence, applying the discrete-time BDG inequalities with constant C_p (or, just plain orthogonality with $C_2 = 1$ when $p = 2$),

$$\|(M^r - M^s)^*\|_p \leq C_p 2^{-r+1} \|X^*\|_p \to 0, \quad \text{as } r, s \to \infty.$$

The underlying process B_t^r will automatically converge, also a.s. uniformly, so to a continuous limit. The increase will be preserved in the limit as well.

Should also $X_t^2 = M_t' + B_t'$, we would have a martingale with paths of bounded variation, as a difference of two increasing processes:

$$M_t - M_t' = B_t' - B_t$$

Then the left hand side would be constant. The convention that B_t starts at 0 implies that the constant is 0. ■

Corollary 14.6 *The assumption that $X^* \in L^1$ can be relaxed.*

Proof. Even for a non-bounded martingale, the stopped martingale X^{σ_a} is bounded where σ_a is the first hitting time of the level $a > 0$ by the $|X_t|$. So, (16.5) holds:

$$|X_t^{\sigma_a}|^2 = M_t^a + B_t^a.$$

However, by uniqueness, $M_t^a = M_t^b$ and $B_t^a = B_t^b$ on $[0, \sigma_a \wedge \sigma_b]$. Since X_t is unbounded, $\sigma_a \to \infty$ when $a \to \infty$. Thus both processes admit the unique extensions over $[0, \infty)$. ■

From the polarization formula

$$XY = \frac{1}{4}\Big((X+Y)^2 - (X-Y)^2\Big),$$

we obtain the decomposition

$$XY = \frac{1}{4}\Big(M_{X+Y} - M_{X-Y}\Big)) + \frac{1}{4}\Big(B_{X+Y} - B_{X-Y}\Big) = M + A$$

into a martingale and a process of bounded variation. Like before, the decomposition is unique, provided the latter process starts at 0. We shall use the square-bracket notation

$$XY = M + [X, Y], \qquad X^2 = M + [X]$$

14.3 Local martingales

In the proof of Theorem **??**, we used a sequence of stopping times T_m, which strengthened the L^1-boundedness of a martingale X_t to the uniform integrability, or even to the integrability of the supremum function, of stopped martingales X^{T_m}. That suggests a new class of processes. Given a filtration $\mathcal{F}_t$, an adapted process X_t is called a **local martingale**, if for some sequence of stopping times $\tau_n \to \infty$, called a **localizing sequence**, X^{τ_n} is a martingale for every n. By the same method, called **localization**, we can define other **local properties**, such as the local boundedness or L^p-boundedness, local continuity, etc.

For example, X_t is a locally bounded martingale, if for some sequence $\tau_n \to \infty$ of optional times, each stopped process X^{τ_n} is bounded, i.e., there exists a constant $C = C_n$ such that $\sup_t |X_{t \wedge \tau_n}| < C$ a.s.

A "*local local*" property is just a local property. For a local martingale X_t, any any escaping nondecreasing sequence of stopping time can be used as a localizing sequence. Indeed, if τ_n was the original seqeunce, and σ is another optional time, then X_n^τ is a martingale, so is $(X_n^\tau)^\sigma$, but the latter equals $(X^\sigma)^{\tau_n}$, so by definition X^σ is a local martingale. The same is true for a sequence of σ_m.

The results in the previous section carry over to the case of local martingales. Let's list the main statements.

Proposition 14.7 *Let X_t be a continuous local martingale.*

1. *If X_t has finite variation, then $X_t = X_0$ a.s.*
2. *There exists a unique local continuous martingale M_t and a unique adapted nondecreasing continuous process $B_t = [X]_t$ with $[X]_0 = 0$ such that*

$$X_t^2 = M_t + [X]_t$$

14.4 BDG inequalities revisited

Recall that BDG inequalities established the equivalence of p-norms of the quadratic function $Q(X)$ and the maximum function X^* of a discrete-time martingale. Since the variation processes stems from the quadratic function, the inequalities can be carried over to the case of continuous time. However, we will prove even more powerful inequalities anew in an easier way

Lemma 14.8 *Let X_t be a non-constant (i.e., $\mathsf{P}(X^* > 0) = 1$) continuous local martingale with $X_0 = 0$. For $a \in \mathbb{R}$, consider the hitting time $\tau_a = \inf t > 0 : X_t = a$. Then*

$$\mathsf{P}(\tau_a < \tau_b) \leq \frac{b}{b-a}, \qquad a < 0 < b$$

Proof. Since $\mathsf{E}\, X_{\tau_a \wedge \tau_b \wedge t} = \mathsf{E}\, X_0 = 0$ for every $t > 0$, and the integrand is bounded, so, letting $t \to \infty$, $0 = \mathsf{E}\, X_{\tau_a \wedge \tau_b}$. Also,

$$\mathsf{E}\,[X_\infty; \tau_a \wedge \tau_b = \infty] = \lim_{s \to \infty} \mathsf{E}\,[X_\infty; \tau_a \wedge \tau_b > s] = 0.$$

Hence,

$$\begin{aligned} 0 = \mathsf{E}\, X_{\tau_a \wedge \tau_b} &= a\mathsf{P}(\tau_a < \tau_b) + b\mathsf{P}(\tau_b \leq \tau_a) + \mathsf{E}\,[X_\infty; \tau_a = \tau_b = \infty] \\ &= a\mathsf{P}(\tau_a < \tau_b) + b(1 - \mathsf{P}(\tau_a < \tau_b)) \end{aligned}$$

A little algebra yields the inequality. ■

Lemma 14.9 (Good λ inequality) *Given a continuous local martingale X_t, denote $U = X^*$ and $V = [X]_\infty^{1/2}$, or switch U, V. Then*

$$\mathsf{P}(U > \beta\lambda,\, V \le \delta\lambda) \le \frac{\delta^2}{(\beta-1)^2}\mathsf{P}(U > \lambda).$$

Proof. Let $\beta > 1$, $\lambda > 0$. Consider $\tau = \inf\{\, t > 0 : |X_t| > \lambda \,\}$. Pick $\delta \in (0, \beta - 1)$. Then

$$M_t = \Big(X_{\tau+t} - X_t\Big)^2 - \Big([X]_{\tau+t} - [X]_t\Big)$$

is a continuous local martingale w.r.t. $\mathcal{F}_{\tau+t}$. On the event

$$A = \big\{\, X^* > \beta\lambda,\, [X]_\infty^{1/2} \le \delta\lambda \,\big\}$$

M_t hits $b = (\beta-1)^2\lambda^2 - \delta^2\lambda^2 > 0$ before it hits $a = -\delta^2\lambda^2$. Hence, by the previous lemma,

$$\mathsf{P}(A) = \mathsf{P}(A, \tau < \infty) = \mathsf{E}\Big(\mathsf{P}[A|\mathcal{F}_\tau]\, 1\!\!1_{\{\tau<\infty\}}\Big) \le \mathsf{P}(\tau < \infty)\frac{\delta^2}{(\beta-1)^2}.$$

The switch between U and V yields the same inequality. ■

Lemma 14.10 *Let Φ be a continuous moderately growing function, $\Phi(x) = 0$ iff $x = 0$. That is, for some (and then for every) $\alpha > 1$*

$$\sup_x \frac{\Phi(ax)}{\Phi(x)} < \infty$$

Let a good λ inequality hold, i.e., for some $\beta > 1$, for all $\delta, \lambda > 0$,

$$\mathsf{P}(U > \beta\lambda,\, V \le \delta\lambda) \le \phi(\delta)\,\mathsf{P}(U > \lambda),$$

for some continuous monotonic function ϕ with $\phi(0) = 0$. Then there is a constant C depending only on $\beta.\phi, \Phi$, such that

$$\mathsf{E}\,\Phi(V) \le C\mathsf{E}\,\Phi(U)$$

Proof. First, we apply integration $\int_0^\infty ...d\lambda$ to both sides of the good λ inequality, and use Fubini's theorem. After some elementary algebra, the moderate growth condition of the function Φ will yield the desired inequality. ■

As a corollary, we obtain

Theorem 14.11 (BDG) *For any moderate Φ,*

$$C_1\Phi([X]_\infty^{1/2}) \le \mathsf{E}\,\Phi(X^*) \le C_2\Phi([X]_\infty^{1/2}).$$

In particular, we may choose $\Phi(t) = t^p$, $p \in (0,\infty)$, or $\Phi(t) = x \wedge 1$, thus establishing the equivalence of convergence in probability for seqeunces X_n^* and $[X_n]_\infty$. ■

Corollary 14.12 *For a sequence of continuous local martingales $X_t^{(n)}$, $(X^{(n)})^* \to 0$ iff $[X^{(n)}]_\infty \to 0$ in L^p, $p \in [0,\infty]$.*

Exercise 14.1 From the inequality

$$\lambda \mathsf{P}(Y > \beta\lambda) \leq \mathsf{E}\,[X; Y > \lambda]$$

held for some $\beta > 1$ and nonnegative r.vs. X, Y, deduce the inequality

$$\|Y\|_p \leq C\|X\|_p,$$

where $p > 1$ and $C = C_{p,\beta}$ (similarly to the proof of Doob's maximal moment inequality; in fact that inequality will be obtained again, as a by-product).

Exercise 14.2 Fox $p \geq 1$.

1. Show that the following spaces of martingales are Banach spaces:

$$\begin{aligned} \mathbb{M}_p &= \left\{ X = (X_n) : \text{ martingales with } \|X\|_p \stackrel{\mathsf{df}}{=} \sup_n \|X_n\|_p < \infty \right\} \\ \mathbb{Q}_p &= \left\{ X = (X_n) : \text{ martingales with } |||X|||_p \stackrel{\mathsf{df}}{=} \|Q(X)\|_p < \infty \right\} \end{aligned}$$

2. Let $1 < p < \infty$, and $1/q + 1/p = 1$. Show the dualities $\mathbb{M}'_p = \mathbb{M}_q$ and $\mathbb{Q}'_p = \mathbb{Q}_q$. Hint: show that, given a martingale $X = (X_n) \in \mathbb{M}_p$,

$$Y_n = \frac{\operatorname{sign}(X_n)|X_n|^{p-1}}{\|X_n\|_p^{p-1}}$$

   is a unit norm martingale in $\mathbb{M}_q$, with $\mathsf{E}\, X_n Y_n = \|X_n\|_p$.

3. Thus, given a continuous mapping $T : \mathbb{M}_p \to \mathbb{Q}_p$, the adjoint $T' : \mathbb{Q}_q \to \mathbb{M}_q$ is also continuous.

Exercise 14.3 Extend the previous properties to the case of continuous time $T = [0, \infty)$. Also, show that the subspaces $\mathbb{M}_p$ consisting of continuous martingales are Banach.

Exercise 14.4 Show that every bounded local martingale is a martingale.

Exercise 14.5 Show an example of a local martingale that is not a martingale.

15 Optional times

In Subsection **??** we introduced the notion of a stopping, or optional, time associated with a discrete-time filtration $\{ \mathcal{F}_n \}$ on a probability space $(\mathcal{F}, \mathfrak{F}, \mathsf{P})$. The definition extends easily to any index set T, but here we will consider $T \subset \overline{\mathbb{R}} = [-\infty, \infty]$ (the extended real line). Again, a measurable mapping $\tau : \Omega \to \overline{\mathbb{R}}$ is dubbed a **stopping** (or **optional**) time with respect to a filtration $\mathcal{F}_t$, if

$$\forall \ t \in T \quad \{ \tau \leq t \} \in \mathcal{F}_t$$

Note that the definition means precisely that the stochastic process

$$I_t = \mathbb{1}_{\{\tau \leq t\}}$$

is adapted to the filtration. A stopping time τ induces its "own" σ-field

$$\mathcal{F}_\tau = \{ A \in \mathcal{F} : A \cap \{ \tau \leq t \} \in \mathcal{F}_t, \text{ for every } t \}.$$

The property of being optional is preserved by the lattice operations of maximum and minimum. Namely, if σ, τ are optional times, so are $\sigma \vee \tau$ and $\sigma \wedge \tau$.

A continuous index set raises many continuity issues, of course. Let's say $T = [0, \infty)$. Any filtration $\mathcal{F}_t$ is covered by a larger filtration

$$\mathcal{F}_t^+ = \bigcap_{s>t} \mathcal{F}_s.$$

It is natural to call $\mathcal{F}_t$ **right-continuous** if it coincides with the above enlargement. Since the property of being optional is defined by the relation "$\in$", the larger the filtration, the easier it is to be optional. So, there might be random times that are optional with respect to $\mathcal{F}_t^+$, and we call them **weakly optional**, but not optional with respect to $\mathcal{F}_t$.

Lemma 15.1 *Let $T = \mathbb{R}$. A random time τ is weakly optional iff $\{ \tau < t \} \in \mathcal{F}_t$ for every t.*

Proof. Below, the subscripts r are rational numbers. Let τ be weakly optional. Then

$$\{ \tau < t \} = \bigcup_{r<t} \{ \tau \leq r \} \in \bigcup_{r<t} \mathcal{F}_r \subset \mathcal{F}_t.$$

Conversely,

$$\{ \tau \leq t \} = \bigcup_{r>t} \{ \tau < r \} \in \bigcup_{r>t} \mathcal{F}_r = \mathcal{F}_t^+$$

■

Similarly we prove that

$$\bigcap_{\epsilon>0} \mathcal{F}_{\tau+\epsilon} = \{ A : A \cap \{ \tau < t \} \in \mathcal{F}_t, \text{ for all } t \}.$$

Now, the countable lattice operation of infimum might not preserve the optionality, only it will preserve it in the weak sense. The weakly optional times can be monotonically approximated from above by optional times. Indeed, the identity $id(x) = x$ is the pointwise infimum of the step-functions

$$f_n(x) = 2^{-n} f(2^n x), \quad \text{where } f(x) = [x+1].$$

So, given a weakly optional time τ, discrete-valued optional times $\tau_n = f_n(\tau)$ decrease monotonically down to τ.

15.1 Progressive stochastic processes

For a discrete-time adapted process X_t, and an optional time τ, X_τ is $\mathcal{F}_\tau$-measurable. Whence we extend the concept of the adaptedness to the process with random (optional) time. For the continuous-time process, this property may fail.

A stochastic process $X = X(t, \omega)$, $t \in T, \omega \in \Omega$ may be looked at from three angles of perspective.

1. X - as a collection of random variables $\{ X_t \}$;
2. X - as a bivariate function, measurable as a mapping
$$\Big(\Omega, \mathcal{F}) \otimes \mathcal{B}[0, \infty) \mapsto (\mathbb{R}, \mathcal{B});$$
3. X - as a random element from Ω into some space of trajectories x_t (e.g., continuous, differentiable, or integrable, etc.)

When the parameter t is seen as time, one may consider the ever growing information concerning the process, encoded as a filtration $\mathcal{F}_t$. In the third approach the connection to that filtration is essentially lost. In the first approach, we talk about adapted processes with each single X_t being measurable with respect to $\mathcal{F}_t$, for every t.

The progression can be also embedded into the second approach. So we call a process $X = X(t, \omega)$ on $[0, \infty)$ **progressive** if the "bivariate measurability" is strengthened to the requirement of measurability of the mappings

$$X : (\Omega, \mathcal{F}_t) \otimes \mathcal{B}[0, t) \mapsto (\mathbb{R}, \mathcal{B}), \quad \text{for every } t > 0$$

(see formula (15.1) below for another equivalent condition). If a process is progressive, then it is adapted. It follows from the fact that sections of a measurable bivariate function are measurable along the variable component. A trace of continuity, e.g., the right-continuity makes adapted processes progressive. That strengthened property makes the process "adapted" in regard to random (optional) time.

Lemma 15.2 *Let $T = \mathbb{R}_+$, $\mathcal{F}_t$ be a filtration, and τ be an optional time. If a process X_t is progressive, then X_τ is $\mathcal{F}_\tau$-measurable.*

Proof. We must show that

$$\{ X_\tau \in B \} \cap \{ \tau \le t \} \in \mathcal{F}_t, \qquad \text{for all Borel sets } B \text{ and } t > 0 \tag{15.1}$$

Since, obviously,

$$\{ X_t \in B \} \cap \{ \tau > t \} \in \mathcal{F}_t,$$

both relations put together mean exactly that $X_{\tau \wedge t}$ is $\mathcal{F}_t$ measurable, for every $t > 0$. Renaming $\sigma = \tau \wedge t$, it suffices to show that, if an optional time $\sigma \le t$, then X_σ is $\mathcal{F}_t$-measurable. However, the latter random variable $X(\omega, \sigma(\omega))$ is a composition

$$\begin{array}{ccccc} \Omega & \overset{\psi}{\mapsto} & \Omega \times T & \overset{X}{\mapsto} & \mathbb{R} \\ \omega & \mapsto & (\omega, \sigma(\omega)) & \mapsto & X(\omega, \sigma(\omega)) \end{array}$$

Considering the σ-fields $\mathcal{F}_t$ on Ω and $\mathcal{B}[0, t]$ on $[0, t] \subset T$, we see that ψ is appropriately measurable, and the progressiveness assumptions makes the composition measurable. ■

15.2 Optional sampling theorem

Let $(X_t, \mathcal{F}_t)$ be a martingale over a continuous subset $T \subset \mathbb{R}$, and $\tau \le \sigma$ be optional times. We wonder if the martingale defining projection carries over to the randomized time martingale. First, we need $\mathcal{X}_\tau$ to be $\mathcal{F}_\tau$ measurable, which will be true, if X_t is right-continuous. Then, we want

$$\mathsf{E}\,[\, X_\tau \,|\, \mathcal{F}_\sigma \,] = X_\sigma \quad a.e.$$

However, optional times are partially not linearly ordered. The more natural yet equivalent property, for given optional times, is

$$\mathsf{E}\,[\, X_\tau \,|\, \mathcal{F}_\sigma \,] = X_{\tau \wedge \sigma} \quad a.e. \tag{15.2}$$

The similar question arises for submartingales and other types of processes. However, (15.2) does not hold always, and additional assumptions are needed.

Theorem 15.3 (Doob Optional Sampling) *Let $(X_t, \mathcal{F}_t)$ be a martingale on $T = [0, \infty)$ that is right-continuous both in regard to paths and the filtration, and σ, τ be optional times. Then* (15.2) *holds true, if one of the following conditions is satisfied:*

1. *τ is bounded;*
2. *X_t is uniformly integrable.*

Proof.

Step 1: a linguistic convention.

Given a measurable set A, and r.vs. ξ, η, we say that "$\xi = \eta$ *on* A if $\xi \mathbb{1}_A = \eta \mathbb{1}_A$ a.s. Given two σ-fields $\mathcal{F}$ and $\mathcal{G}$, we say that "$\mathcal{F} = \mathcal{G}$ *on* A", if $A \in \mathcal{F} \cap \mathcal{G}$ and $A \cap \mathcal{F} = A \cap \mathcal{G}$. To illustrate the usage of the phrase, let's prove the following statemt

Lemma 15.4 *Let σ-fields $\mathcal{F}$ and $\mathcal{G}$ and integrable random variables ξ, η equal on A. Then* $\mathsf{E}[\xi \,|\, \mathcal{F}] = \mathsf{E}[\eta \,|\, \mathcal{G}]$ *on A.*

Proof. Exercise 15.3. ■

Hence (trivially), given a stopping time τ w.r.t. filtration $\mathcal{F}_t$,

$$\mathcal{F}_\tau = \mathcal{F}_t \text{ on } \{\tau = t\}. \tag{15.3}$$

Let T be discrete, τ be bounded, say, $\tau \leq u$, and X_t be a martingale. Then

$$\mathsf{E}[X_u \,|\, \mathcal{F}_\tau] = X_\tau. \tag{15.4}$$

Indeed, by (15.3),

$$\mathsf{E}[X_u \,|\, \mathcal{F}_\tau] = \mathsf{E}[X_u \,|\, \mathcal{F}_t] = X_t = X_\tau \text{ on } \{\tau = t\}.$$

Step 2: Discrete time.

Consider a countable subset $T_0 \subset T$, or, w.l.o.g., suppose that T is countable. Then, for a bounded τ and any σ, (15.4) enables us to replace X_τ by the conditional expectation, and hence

$$\mathsf{E}[X_\tau \,|\, \mathcal{F}\sigma] = \mathsf{E}[X_\tau \,|\, \mathcal{F}_{\sigma\wedge\tau}] = X_{\sigma\wedge\tau} \text{ on } \{\sigma \leq \tau\} \tag{15.5}$$

Again by Step 1,

$$\mathsf{E}[X_\tau \,|\, \mathcal{F}\sigma] = \mathsf{E}[X_{\sigma\wedge\tau} \,|\, \mathcal{F}_\sigma] = X_{\sigma\wedge\tau} \text{ on } \{\sigma > \tau\}. \tag{15.6}$$

That is, when τ is bounded, (15.5) and (15.6) combined yield (15.2).

Now, even for an unbounded τ,

$$\mathsf{E}[X_{u\wedge\tau} \,|\, \mathcal{F}_\sigma] = X_{u\wedge u\tau\wedge\sigma},$$

for every $u \in T$. The assumption of uniform integrability of X_t ensures its a.s. and L^1-convergence. Hence, letting $u \to \infty$, we again obtain (15.2).

Step 3: Continuous time

We approximate optimal times τ and σ monotonically from above by $\tau_n = f_n(\tau)$ and $\sigma_n = f_n(\sigma_n)$. From the discrete-time case, where T_0 is the aritmethic progression of increment 2^{-n}, we infer that

$$\mathsf{E}\left[X_{\tau_n} \,|\, \mathcal{F}_{\sigma_m}\right] = X_{\sigma_m \wedge \tau_n}$$

Letting $m \to \infty$, we obtain in the limit

$$\mathsf{E}\left[X_{\tau_n} \,|\, \mathcal{F}_{\sigma}\right] = X_{\sigma \wedge \tau_n},$$

where we utilized the right-continuity of X_t on the right hand side, and the convergence theorem for closed martingales on the left hand side. X_{τ_n} is uniformly integrable, either by assumption, or as a closed martingale when τ is bounded. The right-continuity on the right, and the L^1- and a.s.-convergence of X_{τ_n} on the left, finish the proof, as $n \to \infty$. ■

If both τ and σ are bounded, or X_t is uniformly integrable, then, taking the mathematical expectation in (15.2), we infer that

$$\mathsf{E}\, X_\tau = \mathsf{E}\, X_{\sigma \wedge \tau} = \mathsf{E}\, X_\sigma. \tag{15.7}$$

Recall a gambler's or investor's concept of the fair game or investment modeled by a martingale. In average the profit remains constant in time. Now, regardless of any strategy as long as it has an end, no 'system' (understood as an implementation of the time-randomized values) will beat the cassino or the market, but fortunately, neither the opposite will happen. Of course many non-martingales have constant mean. Interestingly, if time-limited 'strategies' have no effect on the average profit, the process must be a martingale. More rigorously, the following implication is true:

> If (15.7) holds, i.e., $\mathsf{E}\, X_\tau = \mathsf{E}\, X_\sigma$ for any optimal times σ and τ, then X_t must be a martingale. It even suffices to assume that $\mathsf{E}\, X_t = \mathsf{E}\, X_\sigma$, where σ takes only two values.

Indeed, fix $s < t$. For any $A \in \mathcal{F}_s$, define the optional time

$$\sigma_A = s\mathbb{1}_A + t\mathbb{1}_{A^c}.$$

The assumption reads $\mathsf{E}\, X_t = \mathsf{E}\, X_{\sigma_A}$, i.e.,

$$\mathsf{E}\, X_t = \mathsf{E}\,[X_s; A] + \mathsf{E}\,[X_t; A^c] = \mathsf{E}\,[X_s; A] + \Big(\mathsf{E}\, X_t - \mathsf{E}\,[X_t; A]\Big).$$

That is, for every $A \in \mathcal{F}_s$,

$$\mathsf{E}\,[X_s; A] = \mathsf{E}\,[X_t; A] \quad \Rightarrow \quad X_s = \mathsf{E}\left[X_t \,|\, \mathcal{F}_s\right].$$

15.3 Exercises

All objects below refer to some filtration $\mathcal{F}_t$, where $t \in T \subset \overline{\mathbb{R}}$. The Greek letters $\sigma, \tau, ...$ denote optional times. The objective is to verify each of the following statement, i.e., to prove it or find an example to the contrary

A. *If σ, τ are optional times , then $\sigma + \tau$ is also an optional time.*

B. *τ is $\mathcal{F}_\tau$-measurable.*

C. $\mathcal{F}_{\sigma \wedge \tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$

D. $\mathcal{F}_{\sigma \vee \tau} = \mathcal{F}_\sigma \vee \mathcal{F}\tau$ *(the symbol $\mathcal{A} \vee \mathcal{B}$ applied to σ-fields means $\sigma\{\mathcal{A} \cup \mathcal{B}\}$, for the plain union might not be even a field).*

E. Consider three cases: $T = \mathbb{N}$, $T = [0, \infty)$, or T being a countable subset of $\mathbb{R}$. *Then, the meanings of "optional" and "weakly optional" times are the same.*

F. *Let τ_n be a sequence of optional times. Then*

a. $\sup_n \tau_n$ *is an optional time.*

b. $\tau = \inf_n \tau_n$ *is an*

i. optional time, and $\mathcal{F}_\tau = \bigcap_n \mathcal{F}_{\tau_n}$

ii. weakly optional time, and $\mathcal{F}^+_\tau = \bigcap_n \mathcal{F}^+_{\tau_n}$

G. *An adapted process is progressive* (Hint: false)

H. *An adapted and right-continuous (or left-continuous) process is progressive* (Hint: true).

I. Prove the tautology in Lemma 15.4 from the definition of the conditional expectation. That is, given $A \in \mathcal{F} \cap \mathcal{G}$ and $G \in \mathcal{G}$, show the chain of identities

$$\mathsf{E}\left[\ \mathsf{E}\left[\xi \mathbb{1}_A \,|\, \mathcal{F}\right]\ \mathbb{1}_G\right] = \cdots = \mathsf{E}\left[\eta \mathbb{1}_A \mathbb{1}_G\right].$$

J. The equality of randomized means (15.7) for bounded optional times continues to hold for unbounded optional times (if untrue, that would mean that within infinite time horizon, one can beat or be beaten by the market or the cassino even if the game is fair).

15.4 References

Besides using Chapters 6 and 7 in Kallenberg's book, we borrowed heavily from other sources.

Standard proofs of Doob's inequalities and theorems came from the original Doob's works in 1950s, gathered in the fundamental

J.L. Doob, *Stochastic Processes*, Wiley 1953

The special product martingale $\xi_n(a)$ (**??**) was investigated in

J. Szulga, *Introduction to Random Chaos*, Chapman-Hall 1998.

The discussion on the RNP and counter-examples in c_0 and L^1 came from

J. Diestel & J.J. Uhl, *Vector Measures*, AMS 1977.

The "positive" MCT valid in a separable dual came from

J. Neveu, *Discrete-parameter Martingales*, Elsevier 1972.

The c_0 in the role of a "bad space", e.g., as described in Remark **??** was taken from

J. Hoffman-Jørgensen, *Sums of independent Banach space valued random variables*, Studia Math. 52, 159-189, 1974.

16 Martingale transforms, variation, integration

Consider a stochastic process X_t adapted to a filtration $\mathcal{F}_t$, $t \in T$. Let τ be an optimal time. Recall the stopped process

$$X_t^\tau \overset{\mathsf{df}}{=} X_{\tau\wedge t}. \tag{16.1}$$

Let $T = \mathbb{N}$ and denote the increments by $D_k = \Delta X_k = X_k - X_{k-1}$, and predictable weights by $v_k = 1\!\!1_{\{\tau \geq k\}}$. Then, we can rewrite the stopped process as a "*discrete-time integral*"

$$(v \cdot X)_t = \int_0^t v\, dX = \sum_{k=1}^t v_k \Delta X_k. \tag{16.2}$$

Clearly, not only for the specific 0-1 valued sequence v_k, but for an arbitrary integrable predictable sequence v_k, the formula preserves the martingale property. That is, if X is a martingale, so is $v \cdot X$. It is a fundamental question to ask which further properties of a martingale still appear in the transform, properties such as L^p-boundedness, uniform integrability, convergence, etc. In other words, this the question of continuity of the transform.

Turning the question around, is the transform continuous for non-martingales? In fact, a specific mode of continuity defines a new class of processes, sometimes similar to yet sometimes well beyond martingales. For example, consider a class of processes (somewhat resembling martingales) such that

$$X \text{ is bounded} \quad \Rightarrow \quad v \cdot X \text{ is bounded} \tag{16.3}$$

for some class of predictable sequences v_k (e.g., such that $|v_k| \leq C$), where boundedness could be defined with the help of the L^p-metric. Notice that the Orlicz Summation Theorem may be seen through the angle of above transform. That is, if a sequence (D_k) is summable (i.e., the series $\sum_k D_k$ converges unconditionally, i.e., for every permutation) in L^p, $0 \leq p \leq \infty$, then (16.3) holds for every $\{0, 1\}$ valued v_k, or, equivalently (when $p \geq 1$), whenever v_k are uniformly bounded non-random numbers. In other words, such sequence D_k defines a random vector measure on subsets of $\mathbb{N}$, and convergence of the transform with bounded coefficients corresponds to the integrability of a bounded function with respect to that random measure.

No surprise that procesess satisfying (16.3) with uniformly bounded predictable coefficients are called **semi-martingales**.

The next crucial question is how to extend the concept of transform (16.2) to a time-continuous process. In other words, this the question of a construction of stochastic integrals.

16.1 Rademacher transform and the quadratic function

Although the simplest case seems to be that of bounded deterministic coefficients, even $\{0, 1\}$- or $\{\pm 1\}$-valued, this is not so. Rather, we will begin with $v_k = r_k$, for a Rademacher sequence independent of the martingale $X_t, t \in T = \mathbb{N}$. Recall that all L^p norms or F-norms of a Rademacher series $\sum_k r_k d_k$ are comparable, where $0 \le p < \infty$. For two quantities A and B, depending on some parameter, let's write $A \asymp B$, if there is a constant $c > 0$ such that

$$\frac{1}{c} A \le B \le cA, \quad \text{uniformly with respect to the said parameter}$$

Similarly, we introduce the one-sided relation "$\prec$". So, by Fubini's Theorem, for any $p \in (0, \infty)$,

$$\Big\| \sum_n r_n D_n \Big\|_p \asymp \Big\| \Big(\mathsf{E}_r |\sum_n r_n D_n|^2 \Big)^{1/2} \Big\| = \Big\| \Big(\sum_n |D_n|^2 \Big)^{1/2} \Big\|$$

The latter series $Q(X)$ is called the **quadratic variation** of X, and can be seen as a nondecreasing process

$$Q_n = Q_n(X) = \Big(\sum_{k=1}^n |D_k|^2 \Big)^{1/2}$$

We will prove the main part ($p > 1$) of

Theorem 16.1 (Burkholder-Davis-Gundy: $p > 1$, Davis: $p = 1$) *For any martingale X_n and $p \in [1, \infty)$,*

$$\|Q(X)\|_p \asymp \|X^*\|_p .$$

Hence, every martingale with p-integrable maximum function is a semi-martingale in regard to the L^p-norm.

For the proof we need several lemmas. First, denote $\|X\| = \sup_n \mathsf{E} |X_n|$, and let

$$\tau = \tau_\lambda = \inf \{ n \ge 1 : |X|_n > \lambda \} .$$

We often skip the subscript λ below, for typographic reasons.

Lemma 16.2 *Let X_n be a martingale or a nonnegative submartingale, and $\tau = \tau_\lambda$ for $\lambda > 0$. Then,*

$$\begin{aligned}
&(1) \quad \mathsf{E}\, Q_{\tau-1}^2 + \mathsf{E}\, X_{\tau-1}^2 \le 2\mathsf{E}\, X_\tau X_{\tau-1} \le 2\lambda \|X\|; \\
&(2) \quad \lambda \mathsf{P}(Q > \lambda,\ X^* \le \lambda) \le 2\|X\|; \\
&(3) \quad \lambda \mathsf{P}(Q > \lambda) \le 3\|X\|.
\end{aligned}$$

Proof. The second inequality in (1) is obvious. For any index k,

$$Q_{k-1}^2 + X_{k-1}^2 = 2X_k X_{k-1} - 2\sum_{j=1}^k X_{j-1} D_j.$$

If k is randomized, replaced by τ, the expectation will vanish the second term.

(2): Notice that $\{ X^* \le \lambda \} = \{ \tau_\lambda = \infty \}$. Then, by (1),

$$\lambda \mathsf{P}(Q > \lambda,\, X^* \le \lambda) \le \lambda \mathsf{P}(Q_{\tau-1} > \lambda) \le \frac{1}{\lambda} \mathsf{E}\, Q_{\tau-1}^2 \le 2\|X\|.$$

(3) follows from (2) and from Doob's maximal tail inequality. ■

Lemma 16.3 *Let X_n be a martingale and $\lambda, \theta > 0$. Denote $Y_n = Q_n(\theta X) \vee X_n^*$. Then there is $\beta = \beta(\theta) > 1$ such that*

$$\lambda \mathsf{P}(Y_n > \beta\lambda) \le 3\mathsf{E}\,[X_n; Y_n > \lambda]. \tag{16.4}$$

Proof. The choice $\beta^2 = 1 + 2\theta^2$ will be sufficient. Define another stopping time

$$\sigma = \inf \{ n \ge 1 : Q_n(\theta X) > \lambda \}.$$

The following sequence is a positive submartingale, since the multiplier defined by the indicator increases with n:

$$Z_n = X_n 1\!\!1_{\{Q_n(\theta X) > \lambda\}}.$$

Note the inclusions

$$A \stackrel{\mathsf{df}}{=} \{ Q_n(\theta X) > \beta\lambda,\, X_n^* \le \lambda \} \subset \{ \sigma \le n,\, Z_n^* \le \lambda,\, |D_\sigma| \le \lambda \}$$

Hence, on A

$$\begin{aligned} \lambda(1+2\theta^2) Q_n^2(\theta X) = \;& Q_{\sigma-1}^2(\theta X) + \theta^2 D_\sigma^2 + \theta^2 \sum_{j=\sigma+1}^{n} D_j^2 \\ \le \;& \lambda^2 + \theta^2\lambda^2 + \theta^2 \sum_{j=\sigma+1}^{n} (Z_j - Z_{j-1})^2 \\ \le \;& (1+\theta^2)\lambda^2 + \theta^2 Q_n^2(Z) \end{aligned}$$

Since $A \subset Q_n(\theta X) > \beta\lambda$, we infer that $A \subset \{ Q_n(Z) > \lambda \}$. Now, applying (2) of the preceding lemma and Doob's maximal tail inequality to the submartingale Z, we obtain

$$\begin{aligned} \lambda \mathsf{P}(Q_n(\theta X) \vee X_n^* > \beta\lambda) \le \;& \lambda \mathsf{P}(Q_n(\theta X) > \beta\lambda) + \lambda \mathsf{P}(X_n^* > \beta\lambda) \\ \le \;& \lambda \mathsf{P}(Q_n(Z) > \lambda) + \lambda \mathsf{P}(X_n^* > \lambda) \\ \le \;& 3\mathsf{E}\,[|X_n|; Y_n > \lambda] \end{aligned}$$

■

We repeat the derivation of Doob's moment maximal inequality from Doob's tail maximal inequality. ■

Proof of BDG Theorem.

The proof of the left inequality $\|Q(X)\|_p \le C\|X\|_p$ is essentially contained in Exercise 16.1, since $Q(X) \le Y$. The right inequality follows by duality (cf. Exercise 16.2) ■

16.2 Quadratic function in continuous time

Let $\mathcal{F}_t$ be a filtration on $T = [0, \infty)$, and $\tau = (\tau_n)$ be a nondecreasing sequence of optional times diverging to ∞, with $\tau_0 = 0$. The first extension of the quadratic function of a martingale $X = (X_t)$ is defined as a nondecreasing process

$$B_t = B_t(X, \tau) = \sum_k \left| X_{t\wedge\tau_k} - X_{t\wedge\tau_{k-1}} \right|^2$$

To continue the construction, we will refine the random partitions induced by the times τ_n while assumming continuous paths of the martingale. The limit process can be intepreted as the quadratic variation. In fact, no continuous non-constant martingale will have paths of bounded variation.

Lemma 16.4 *Let X_t be a continuous martingale on $T = [0, \infty)$. If*

$$V(t) = \sup \left\{ \sum_{j=1}^{n} |X_{t_j} - X_{t_{j-1}}| : 0 = t_0 < t_1 < \ldots < t_n \leq t \right\} < \infty \ \textit{a.s.},$$

Then $X_t = X_0$ a.s.

Proof. First, if the variation is finite, we may assume that is bounded, by considering the stoppped martingale X^τ, where τ is the first hitting time of a level a by the process V_t.

The modulus of continuity

$$\rho_t(\epsilon) = \sup \{ |X_u - X_v| : |u - v| < \epsilon, \, u \vee v \leq t \} \to 0 \quad \text{a.s.}$$

when $\epsilon \to 0$. Then, choosing $\epsilon = n^{-1}$

$$\sum_{k=1}^{n} |X_{kt/n} - X_{(k-1)t/n}|^2 \leq V(t)\rho(n^{-1}) \leq a\,\rho(n^{-1})$$

By the BDG inequality (or using just the second moment)

$$\|X_t - X_0\|_p \leq Ca\mathsf{E}\,\rho(n^{-1}) \to 0$$

Hence $X_t = X_0$. ∎

For a fixed t, denoting

$$D_k = X_{t\wedge\tau_k} - X_{t\wedge\tau_{k-1}},$$

we derive the identity

$$X_t^2 = M_t + B_t, \tag{16.5}$$

with a certain martingale M_t, by twice telescoping the suitable sums:

$$\begin{aligned} X_t^2 = & \left(\sum_k D_k\right)^2 \\ = & \sum_j \sum_k D_j D_k \\ = & 2\sum_k \sum_{j=1}^{k-1} D_j\, D_k + \sum_k |D_k|^2 \\ = & 2\sum_k X_{t\wedge\tau_{k-1}} D_k + B_t \\ = & 2\sum_k X_{\tau_{k-1}} D_k + B_t, \end{aligned}$$

In the sum we recognize a martingale transform $M_t = (v \cdot Y)_t$ of a discrete time martingale Y with differences D_k and predictable coefficients $v_k = X_{\tau_{k-1}}$. Thus, we may apply the discrete-time BDG inequalities to it. Still, these coefficients define a predictable process

$$v_t = \sum_k v_k \mathbb{I}_{(\tau_{k-1},\tau_k]}(t).$$

Theorem 16.5 *Let X_t be a continuous martingale with $X^* \in L^p$, for some $p \geq 1$. Let $r \in \mathbb{N}$, and let $\tau_n = \tau_n^r$ be optional times defined by induction:*

$$\tau_{n+1} = \inf t > \tau_n : |X_t - X_{\tau_n}| = 2^{-r}.$$

Then the continuous martingales M_t^r defined above converge and in $\mathbb{M}_p$ to some continuous martingale M_t, as $r \to \infty$, and the decomposition formula (16.5) *holds, where B_t is a continuous nondecreasing process. Further, the decomposition is unique if $B_0 = 0$.*

Proof. In virtue of Exercise 16.3, it suffices to show that M_t^r is Cauchy. Let v^r denote the suitable predictable process introduced above. Then,

$$\sup_t \| v_t^r - X_t| \leq 2^{-r}$$

Therefore, for an integer $s \geq r$,

$$\sup_t \| v_t^r - v_t^s| \leq 2^{-r} + 2^{-s} \leq 2^{-r+1}.$$

Whence, applying the discrete-time BDG inequalities with constant C_p (or, just plain orthogonality with $C_2 = 1$ when $p = 2$),

$$\|(M^r - M^s)^*\|_p \leq C_p 2^{-r+1} \|X^*\|_p \to 0, \quad \text{as } r, s \to \infty.$$

The underlying process B_t^r will automatically converge, also a.s. uniformly, so to a continuous limit. The increase will be preserved in the limit as well.

Should also $X_t^2 = M_t' + B_t'$, we would have a martingale with paths of bounded variation, as a difference of two increasing processes:

$$M_t - M_t' = B_t' - B_t$$

Then the left hand side would be constant. The convention that B_t starts at 0 implies that the constant is 0. ■

Corollary 16.6 *The assumption that $X^* \in L^1$ can be relaxed.*

Proof. Even for a non-bounded martingale, the stopped martingale X^{σ_a} is bounded where σ_a is the first hitting time of the level $a > 0$ by the $|X_t|$. So, (16.5) holds:

$$|X_t^{\sigma_a}|^2 = M_t^a + B_t^a.$$

However, by uniqueness, $M_t^a = M_t^b$ and $B_t^a = B_t^b$ on $[0, \sigma_a \wedge \sigma_b]$. Since X_t is unbounded, $\sigma_a \to \infty$ when $a \to \infty$. Thus both processes admit the unique extensions over $[0, \infty)$. ■

From the polarization formula

$$XY = \frac{1}{4}\Big((X+Y)^2 - (X-Y)^2\Big),$$

we obtain the decomposition

$$XY = \frac{1}{4}\Big(M_{X+Y} - M_{X-Y}\Big)\Big) + \frac{1}{4}\Big(B_{X+Y} - B_{X-Y}\Big) = M + A$$

into a martingale and a process of bounded variation. Like before, the decomposition is unique, provided the latter process starts at 0. We shall use the square-bracket notation

$$XY = M + [X, Y], \qquad X^2 = M + [X]$$

16.3 Local martingales

In the proof of Theorem **??**, we used a sequence of stopping times T_m, which strengthened the L^1-boundedness of a martingale X_t to the uniform integrability, or even to the integrability of the supremum function, of stopped martingales X^{T_m}. That suggests a new class of processes. Given a filtration $\mathcal{F}_t$, an adapted process X_t is called a **local martingale**, if for some sequence of stopping times $\tau_n \to \infty$, called a **localizing sequence**, X^{τ_n} is a martingale for every n. By the same method, called **localization**, we can define other **local properties**, such as the local boundedness or L^p-boundedness, local continuity, etc.

For example, X_t is a locally bounded martingale, if for some sequence $\tau_n \to \infty$ of optional times, each stopped process X^{τ_n} is bounded, i.e., there exists a constant $C = C_n$ such that $\sup_t |X_{t \wedge \tau_n}| < C$ a.s.

A "*local local*" property is just a local property. For a local martingale X_t, any any escaping nondecreasing sequence of stopping time can be used as a localizing sequence. Indeed, if τ_n was the original seqeunce, and σ is another optional time, then X_n^τ is a martingale, so is $(X_n^\tau)^\sigma$, but the latter equals $(X^\sigma)^{\tau_n}$, so by definition X^σ is a local martingale. The same is true for a sequence of σ_m.

The results in the previous section carry over to the case of local martingales. Let's list the main statements.

Proposition 16.7 *Let X_t be a continuous local martingale.*

1. *If X_t has finite variation, then $X_t = X_0$ a.s.*
2. *There exists a unique local continuous martingale M_t and a unique adapted nondecreasing continuous process $B_t = [X]_t$ with $[X]_0 = 0$ such that*

$$X_t^2 = M_t + [X]_t$$

16.4 BDG inequalities revisited

Recall that BDG inequalities established the equivalence of p-norms of the quadratic function $Q(X)$ and the maximum function X^* of a discrete-time martingale. Since the variation processes stems from the quadratic function, the inequalities can be carried over to the case of continuous time. However, we will prove even more powerful inequalities anew in an easier way

Lemma 16.8 *Let X_t be a non-constant (i.e., $\mathsf{P}(X^* > 0) = 1$) continuous local martingale with $X_0 = 0$. For $a \in \mathbb{R}$, consider the hitting time $\tau_a = \inf t > 0 : X_t = a$. Then*

$$\mathsf{P}(\tau_a < \tau_b) \leq \frac{b}{b-a}, \qquad a < 0 < b$$

Proof. Since $\mathsf{E}\, X_{\tau_a \wedge \tau_b \wedge t} = \mathsf{E}\, X_0 = 0$ for every $t > 0$, and the integrand is bounded, so, letting $t \to \infty$, $0 = \mathsf{E}\, X_{\tau_a \wedge \tau_b}$. Also,

$$\mathsf{E}\,[X_\infty; \tau_a \wedge \tau_b = \infty] = \lim_{s \to \infty} \mathsf{E}\,[X_\infty; \tau_a \wedge \tau_b > s] = 0.$$

Hence,

$$\begin{aligned} 0 = \mathsf{E}\, X_{\tau_a \wedge \tau_b} = &\ a\mathsf{P}(\tau_a < \tau_b) + b\mathsf{P}(\tau_b \leq \tau_a) + \mathsf{E}\,[X_\infty; \tau_a = \tau_b = \infty] \\ = &\ a\mathsf{P}(\tau_a < \tau_b) + b(1 - \mathsf{P}(\tau_a < \tau_b)) \end{aligned}$$

A little algebra yields the inequality. ∎

Lemma 16.9 (Good λ inequality) *Given a continuous local martingale X_t, denote $U = X^*$ and $V = [X]_\infty^{1/2}$, or switch U, V. Then*

$$\mathsf{P}(U > \beta\lambda,\, V \le \delta\lambda) \le \frac{\delta^2}{(\beta-1)^2}\mathsf{P}(U > \lambda).$$

Proof. Let $\beta > 1$, $\lambda > 0$. Consider $\tau = \inf\{\, t > 0 : |X_t| > \lambda \,\}$. Pick $\delta \in (0, \beta - 1)$. Then

$$M_t = \Big(X_{\tau+t} - X_t\Big)^2 - \Big([X]_{\tau+t} - [X]_t\Big)$$

is a continuous local martingale w.r.t. $\mathcal{F}_{\tau+t}$. On the event

$$A = \big\{\, X^* > \beta\lambda,\, [X]_\infty^{1/2} \le \delta\lambda \,\big\}$$

M_t hits $b = (\beta-1)^2\lambda^2 - \delta^2\lambda^2 > 0$ before it hits $a = -\delta^2\lambda^2$. Hence, by the previous lemma,

$$\mathsf{P}(A) = \mathsf{P}(A, \tau < \infty) = \mathsf{E}\Big(\,\mathsf{P}[A|\mathcal{F}_\tau]\,\mathbb{1}_{\{\tau<\infty\}}\Big) \le \mathsf{P}(\tau < \infty)\frac{\delta^2}{(\beta-1)^2}.$$

The switch between U and V yields the same inequality. ∎

Lemma 16.10 *Let Φ be a continuous moderately growing function, $\Phi(x) = 0$ iff $x = 0$. That is, for some (and then for every) $\alpha > 1$*

$$\sup_x \frac{\Phi(ax)}{\Phi(x)} < \infty$$

Let a good λ inequality hold, i.e., for some $\beta > 1$, for all $\delta, \lambda > 0$,

$$\mathsf{P}(U > \beta\lambda,\, V \le \delta\lambda) \le \phi(\delta)\,\mathsf{P}(U > \lambda),$$

for some continuous monotonic function ϕ with $\phi(0) = 0$. Then there is a constant C depending only on $\beta.\phi, \Phi$, such that

$$\mathsf{E}\,\Phi(V) \le C\mathsf{E}\,\Phi(U)$$

Proof. First, we apply integration $\int_0^\infty ...d\lambda$ to both sides of the good λ inequality, and use Fubini's theorem. After some elementary algebra, the moderate growth condition of the function Φ will yield the desired inequality. ∎

As a corollary, we obtain

Theorem 16.11 (BDG) *For any moderate Φ,*

$$C_1\Phi([X]_\infty^{1/2}) \le \mathsf{E}\,\Phi(X^*) \le C_2\Phi([X]_\infty^{1/2}).$$

In particular, we may choose $\Phi(t) = t^p$, $p \in (0,\infty)$, or $\Phi(t) = x \wedge 1$, thus establishing the equivalence of convergence in probability for seqeunces X_n^* and $[X_n]_\infty$. ∎

Corollary 16.12 *For a sequence of continuous local martingales $X_t^{(n)}$, $(X^{(n)})^* \to 0$ iff $[X^{(n)}]_\infty \to 0$ in L^p, $p \in [0,\infty]$.*

Exercise 16.1 From the inequality

$$\lambda \mathsf{P}(Y > \beta\lambda) \le \mathsf{E}\,[X; Y > \lambda]$$

held for some $\beta > 1$ and nonnegative r.vs. X, Y, deduce the inequality

$$\|Y\|_p \le C\|X\|_p,$$

where $p > 1$ and $C = C_{p,\beta}$ (similarly to the proof of Doob's maximal moment inequality; in fact that inequality will be obtained again, as a by-product).

Exercise 16.2 Fox $p \ge 1$.

1. Show that the following spaces of martingales are Banach spaces:

$$\begin{aligned} \mathbb{M}_p &= \left\{ X = (X_n):\ \text{martingales with } \|X\|_p \stackrel{\mathsf{df}}{=} \sup_n \|X_n\|_p < \infty \right\} \\ \mathbb{Q}_p &= \left\{ X = (X_n):\ \text{martingales with } |||X|||_p \stackrel{\mathsf{df}}{=} \|Q(X)\|_p < \infty \right\} \end{aligned}$$

2. Let $1 < p < \infty$, and $1/q + 1/p = 1$. Show the dualities $\mathbb{M}_p' = \mathbb{M}_q$ and $\mathbb{Q}_p' = \mathbb{Q}_q$. Hint: show that, given a martingale $X = (X_n) \in \mathbb{M}_p$,

$$Y_n = \frac{\operatorname{sign}(X_n)|X_n|^{p-1}}{\|X_n\|_p^{p-1}}$$

   is a unit norm martingale in $\mathbb{M}_q$, with $\mathsf{E}\,X_nY_n = \|X_n\|_p$.

3. Thus, given a continuous mapping $T : \mathbb{M}_p \to \mathbb{Q}_p$, the adjoint $T' : \mathbb{Q}_q \to \mathbb{M}_q$ is also continuous.

Exercise 16.3 Extend the previous properties to the case of continuous time $T = [0, \infty)$. Also, show that the subspaces $\mathbb{M}_p$ consisting of continuous martingales are Banach.

Exercise 16.4 Show that every bounded local martingale is a martingale.

Exercise 16.5 Show an example of a local martingale that is not a martingale.

17 Introduction to semimartingales

Like before, $\mathcal{F} = (\mathcal{F}_t)$ is a filtration on $T = [0, \infty)$. Also, we shall also assume that an underlying probability space is complete. Recall that, for a simple predictable left continuous process

$$V_t = \sum_{i=1}^{n} V_{i-1} \mathbb{1}_{(\tau_{i-1}, \tau_i]},$$

where τ_i is a nondecreasing sequence of stopping times and each V_i is $\mathcal{F}_{\tau_i}$-measurable, and for a local martingale X_t, the martingale transform

$$(V \cdot X)_t = \int_0^t V\, dX = \sum_{i=1}^{n} V_{i-1} \Big(X_{\tau_i \wedge t} - X_{\tau_i \wedge t} \Big)$$

is also a local martingale. The set of all such processes is a vector space, and we can turn it into a normed space under the (t, ω)-supremum norm (i.e., with the topology of uniform convergence). In particular, the **stochastic integral** is well defined.

$$I(V) = X_X(V) = \int_T V\, dX = \sum_{i=1}^{n} V_{i-1} \Big(X_{\tau_i} - X_{\tau_i} \Big) \tag{17.1}$$

From BDG inequalities[1] it follows that, whenever $|V| \leq 1$, then the mapping

$$V \mapsto I_X(V)$$

is continuous (even in every L^p, $p \in [0, \infty)$). The same pair of essential properties - linearity and continuity - is obviously well known in the classical calculus, that is, for a process with finite local variation. One of important consequences is the the possibility of an extension of the process I_X to the completion of $\mathbb{S}$ in the uniform norm.

From the decomposition theorem, $X^2 = M + [X]$ into the sum of a martingale and a process of bounded variation, we see that the square process shares the two essential properties. Since the decomposition holds for a product, thus any integer power X_t^k, and thus, any polynomial $p(X_t)$ again will stay within the discussed class. Obviously, one needs to use some approximation tools to extend the class even further, to incorporate functionals $f(X_t)$ for analytic functions, then, even for some smooth functions.

17.1 Preliminary definitions

Let $\mathbb{S}$ be the vector space of simple predictable left-continuous processes of the form

$$V_t = V_0 \mathbb{1}_{\{0\}} + \sum_{i=1}^{n} V_{i-1} \mathbb{1}_{(\tau_{i-1}, \tau_i]}, \tag{17.2}$$

[1] even more trivially, when X_t is an L^2-martingale

where $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n$, $|V_i| \leq C < \infty$ and V_i are $\mathcal{F}_{\tau_i}$-measurable. $\mathbb{S}$ becomes a normed space under the norm

$$\|V\| = \sup_t \|V_t\|_\infty.$$

Given an adapted process X_t that is right continuous and have left limits (**cadlag** or **rcll**), we define the linear mappings on $\mathbb{S}$ with values in L^0

$$[I_X(V)]_t = \int_0^t V\,dX = V_0X_0 + \sum_{i=1}^n V_{i-1}\Big(X_{t\wedge\tau_i} - X_{t\wedge\tau_i}\Big).$$

Such process X_t is called a **semimartingale** if the mappings $[I_X]_t : \mathbb{S} \mapsto L^0$ are continuous for every $t \geq 0$. Obviously, the set of all semimartingales is a vector space.

The definition depends on the probability, since the continuity involves the convergence in probability. We may underline the dependence calling a semimartingale more precisely a P-semimartingale. Thus, if P' is another probability that is absolutely continuous with respect to the original probability P, a P-semimartingale is still a P'-semimartingale. By the same token, if X_t is a P_k-semimartingale, then X_t is a $\mathsf{P} = \sum_k c_k\mathsf{P}_k$-semimartingale for a convex combination of probabilities.

The definition also depends on the filtration, for the continuity uses convergent predictable processes. The fewer such processes, the more likely X_t is a semimartingale. That is, the smaller the filtration, the more likely a process is a semimartingale. Thus, if X_t is an $\mathcal{F}$-semimartingale $\mathcal{G} \subset \mathcal{F}$, but X_t is adapted to $\mathcal{G}$, then X_t is a $\mathcal{G}$-semimartingale.

It is interesting to observe that the property of being a semimartingale is already local.

Lemma 17.1 *Every local semimartingale is a semimartingale.*

Proof. Let $\sigma_n \to \infty$ be an increasing sequence of stopping times and $X_t^{\sigma_n}$ be a semimartingale, for each n. Fix a $t \geq 0$. To prove the continuity of I_X, let $\mathbb{S} \ni V_m \to 0$ uniformly. Fix $c > 0$. We need to show that

$$\mathsf{P}(I_{X^t}(V_m)| > c) \to 0.$$

Let $\epsilon > 0$. Choose n large enough so that $\mathsf{P}(\sigma_n \leq t) < \epsilon/2$. Given that n, choose m large enough so that

$$\mathsf{P}(I_{X^{\sigma_n \wedge t}}(V_m)| > c) < \epsilon/2.$$

That is, given $\epsilon > 0$, for m large enough

$$\begin{aligned}\mathsf{P}(I_{X^t}(V_m)| > c) = &\ \mathsf{P}(I_{X^t}(V_m)| > c, \sigma_n \leq t) + \mathsf{P}(I_{X^t}(V_m)| > c, \sigma_n > t)\\ \leq &\ \mathsf{P}(\sigma_n \leq t) + \mathsf{P}(I_{X^{\sigma_n \wedge t}}(V_m)| > c) \leq \epsilon/2 + \epsilon/2 = \epsilon\end{aligned}$$

So, the integral is continuous, i.e., X_t is a semimartingale. ∎

17.2 Spaces and topologies

In addition to the space $\mathbb{S}$ of simple predictable processes, consider also the vector space $\mathbb{D}$ of right-continuous processes with left limits (**rcll**, or French **cadlag**). Switching "left" with "right" yields the twin vector space $\mathbb{L}$ (**lcrl**, or French **caglad**). Note the assignment

$$\mathbb{L} \ni Y \quad \mapsto \quad Y^+ \in \mathbb{D}, \qquad Y_t^+ \stackrel{\mathsf{df}}{=} Y_{t+} = \lim_{s \to t+} Y_s$$

Recall the notation

$$X_t^* = \sup_{0 \le s \le t} |X_s|.$$

Let us endow $\mathbb{L}$ with the topology through a family of semi-F-norms[2]

$$\|V\|_t = \mathsf{E}\, 1 \wedge X_t^*.$$

Obviously, a countable family, e.g., $(\|V\|_n, n \in \mathbb{N})$, is sufficient, and thus this topology is metrizable via the F-norm

$$\|V\| = \sum_n \frac{\|V\|_n}{2^n}.$$

So, the metric space induces the convergence in probability, uniform on compact sets in T. We shall use the acronym **ucp** to denote this topology. It follows (Exercise 17.1) that, e.g., $\mathbb{D}_{ucp}$ is a complete metric space.

Recall that a semimartingale X is a process for which the mappings

$$\mathbb{S}_u \ni V \mapsto V \cdot X^t \in L^0$$

are continuous for every $t > 0$, where the subscript "u" marks the uniform convergence. Now, we shall strengthen the mode of continuity.

Theorem 17.2 *For a semimartingale X, the mapping*

$$\mathbb{S}_{ucp} \ni V \mapsto V \cdot X \in \mathbb{D}_{ucp}$$

is continuous.

Proof. That the integral $V \cdot X$ is cadlag is obvious. First, we will show the continuity of

$$\mathbb{S}_u \ni V \mapsto V \cdot X \in \mathbb{D}_{ucp}.$$

Suppose that $\|V_k\|_\infty \to 0$, and w.l.o.g, we may assume that these processes are uniformly bounded. We need to show that for every $c > 0$,

$$\mathsf{P}((V_k \cdot X)_t^* > c) \to 0, \qquad \text{as } k \to \infty. \tag{17.3}$$

[2]the F-norm $\mathsf{E}\, 1 \wedge |X|$ induces the convergence in probability.

If σ_k are the stopping times

$$\sigma_k = \inf\{\, t > 0 : |(V_k \cdot X)_t| \geq c\,\},$$

then $V_k 1\!\!1_{[0,\sigma_k]} \in \mathbb{S}$ and converge uniformly to 0, as $k \to \infty$. Also

$$(V_k \cdot X)_{\sigma_k \wedge t} = (V_k 1\!\!1_{[0,\sigma_k]} \cdot X)_t$$

Thus, by the definition of semimartingale,

$$\mathsf{P}((V_k \cdot X)^*_t > c) \leq \mathsf{P}((V_k \cdot X)_{\sigma_k \wedge t} \geq c) = \mathsf{P}(V_k 1\!\!1_{[0,\sigma_k]} \cdot X)_t \geq c) \to 0.$$

Let now simple predictable $V_k \overset{ucp}{\to} 0$. We will construct processes V'_k that are "close" to V_k and converge to 0 uniformly. So, for an arbitrary $c > 0$ we must prove (17.3). In the first part of the proof, we established the implication: for every t and c,

$$\forall\, \epsilon > 0 \;\; \exists\, \eta > 0 \quad \|V\|_\infty < \eta \quad \Rightarrow \quad \mathsf{P}((V \cdot X)^*_t > c) < \epsilon.$$

Now, let $\rho_k = \inf\{\, t > 0 : |(V_k)_t| > \eta\,\}$ and define a simple predictable caglad process

$$V'_k \overset{\mathsf{df}}{=} V_k \, 1\!\!1_{[0,\rho_k]} \, 1\!\!1_{\{\rho_k > 0\}}.$$

Further, $\|V'_k\|_\infty \leq \eta$. Also

$$t \leq \rho_k \;\Rightarrow\; (V_k \cdot X)^*_t = (V'_k \cdot X)^*_t.$$

So, (17.3) is obtained as follows:

$$\begin{aligned}
\mathsf{P}((V_k \cdot X)^*_t > c) \leq \;& P((V_k \cdot X)^*_t > c,\, \rho_k \geq t) + P((V_k \cdot X)^*_t > c, \rho_k > t) \\
\leq \;& P((V'_k \cdot X)^*_t > c,\, \rho_k \geq t) + P(\rho_k > t) \\
\leq \;& \epsilon + \mathsf{P}((V_k)^*_t > \eta)
\end{aligned}$$

Apply the limes superior with respect to k, so the latter probability goes to 0, since we assumed that $V_k \to 0$ ucp. Now, let $\epsilon \to 0$. The ucp-convergence of the integrals $V_k \cdot X$ has been established. ■

Now we can extend the **stochastic integral** to integrands from $\mathbb{L}$ if only we can show that $\mathbb{S}$ is dense in $\mathbb{L}$ in **ucp**.

Proposition 17.3 *The space* $\mathbb{S}$ *is* **ucp**-*dense in* $\mathbb{L}$.

Proof. Let us write the letter "b" to denote the subspace of a given space of processes that consists of bounded processes, e.g., $b\mathbb{L}$. Since $b\mathbb{L}$ is dense (Exercise 17.2) in $\mathbb{L}$, it suffices to approximate a given bounded caglad process U by simple predictable processes. Given $\epsilon > 0$, define the increasing sequence of stoppimg times τ_n by induction

$$\tau_0 = 0, \quad \tau_{n+1} = \inf\left\{\, t > \tau_n : |U^+_t - U^+_{\tau_n}| > \epsilon \,\right\}.$$

Notice that we needed to switch to the cadlag version U^+ because the original process U would yield only weak stopping times. This produces a bounded cadlag simple process

$$\sum_n U_{\tau_n} \mathbb{1}_{[\tau_n, \tau_{n+1})}$$

that uniformly converges to U as $\epsilon \to 0$. Now, turn it to a bounded caglad (so, predictable) process

$$U_0 \mathbb{1}_{\{0\}} + \sum_n U_{\tau_n} \mathbb{1}_{(\tau_n, \tau_{n+1}]}.$$

It is not yet simple, but it can be made so:

$$V^{N,\epsilon} \stackrel{\mathsf{df}}{=} U_0 \mathbb{1}_{\{0\}} + \sum_{n \leq N} U_{\tau_n} \mathbb{1}_{(\tau_n \wedge N, \tau_{n+1} \wedge N]}.$$

The former process is uniformly approximated when $N \to \infty$. ■

17.3 Exercises

Exercise 17.1 Show that $\boldsymbol{D}_{ucp}$ (or $\boldsymbol{L}_{ucp}$) is complete.

Exercise 17.2 Show that $b\mathbb{L}$ is **ucp**-dense in $\mathbb{L}$ (Hint: stop the process when it is about to leave the interval $(-n, n)$)

18 Chaos representation of random functionals

As before, $T \subset [0, \infty)$. Let $X = (X_t,\, t \in T)$ be a real stochastic process on some probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Consider the σ-field $\sigma(X) = \sigma\{\, X_t : t \in T\,\}$ induced by all values of the process, and denote by $L^0(X)$ (respectively, $L^p(X)$) the space of all random variables (respectively, p-integrable random variables) measurable with respect to the σ-field $\sigma(X)$. The members of $L^0(X)$ will be called **random functionals of the process** X. Our goal is to give a characterization of random functionals for some basic processes in terms of so called **random chaos**.

For example, for n moments $t_1, \ldots, t_n$ and a Borel function $\phi : \mathbb{R}^n \to \mathbb{R}$, $Y = \phi(X_{t_1}, \ldots, X_{t_n})$ is such functional. The notion of a **statistic**, i.e., a functional expression based on a random sample $X_1, \ldots, X_n$ is then an example of a random functional of a discrete-time process $X = (X_n)$. That, in particular, includes the sample mean and sample variance. Notice that these popular statistics involve linear and quadratic forms. These forms and forms of higher order,

$$\begin{aligned} &\sum_i a_i X_i, \\ &\sum_{i,j} a_{ij} X_i X_j, \\ &\sum_{i,j} a_{ijk} X_i X_j X_k, \\ &\ldots \sum_{i_1,\ldots,i_k} a_{i_1 \ldots i_k} X_{i_1} \cdots X_{i_k}, \end{aligned}$$

and their sums are examples of random functionals again. Observe that the forms of order greater than 1 involve various powers X_i^m. So, assuming for simplicity that the coefficients are symmetric, i.e., take the same values for any permutation of its indices, we have:

$$\begin{aligned} \sum_{i,j} a_{ij} X_i X_j &= \sum_i a_{ii} X_i^2 + 2 \sum_{i<j} a_{ij} X_i X_j \\ \sum_{i,j} a_{ijk} X_i X_j X_k &= \sum_i a_{iii} X_i^3 + 3 \sum_{i<j} a_{iij} X_i^2 X_j + 6 \sum_{i<j<k} a_{ijk} X_i X_j X_k, \text{ etc.} \end{aligned}$$

It is natural to call the last multiple sums in each line above **tetrahedral forms**, and the sums involving powers greater than 1 **diagonal sums**, because of the geometric shapes in the index set,

18.1 Tensor products

Let T be a set and $f, f_1, \ldots, f_n : T \to \mathbb{R}$ be real[3] functions. The **tensor product** or the **tensor power** is the n-variate function

$$f_1 \otimes \cdots \otimes f_n\,(t_1, \ldots, t_n) \overset{\mathsf{df}}{=} f_1(t_1) \cdots f_n(t_n), \qquad f^{\otimes n}\,(t_1, \ldots, t_n) \overset{\mathsf{df}}{=} f(t_1) \cdots f(t_n).$$

By symmetrizing or restricting the order of indices we obtain other tensor products. First, introduce the symmetrization of an n-variate function $f(t_1, \ldots, t_n)$:

$$\tilde{f}(t_1, \ldots, t_n) = \frac{1}{n!} \sum_{\pi} f(t_{\pi(1)}, \ldots, t_{\pi(n)}),$$

where the sum is taken over all permutations of the set $\{1, \ldots, n\}$. This yields the **symmetric tensor product**:

$$f_1 \tilde{\otimes} \cdots \tilde{\otimes} f_n \overset{\mathsf{df}}{=} (f_1 \otimes \cdots \otimes f_n)\tilde{}.$$

Of course, the symmetrization has no effect on tensor powers, i.e.,

$$f^{\tilde{\otimes} n} = f^{\otimes n}$$

Let $\boldsymbol{F}, \boldsymbol{F}_1, \ldots, \boldsymbol{F}_n$ be some vector spaces of functions $f : T \to \mathbb{R}$. Let us introduce the tensor products or powers of the spaces:

$$\boldsymbol{F}_1 \otimes \cdots \otimes \boldsymbol{F}_n = \operatorname{lin}\{f_1 \otimes \cdots \otimes f_n : f_1 \in \boldsymbol{F}_1, \cdots, f_n \in \boldsymbol{F}_n\}, \qquad \boldsymbol{F}^{\otimes n} = \operatorname{lin}\left\{f^{\otimes n} : f \in \boldsymbol{F}\right\}.$$

Warning. Even when $\boldsymbol{F}_1 = \ldots = \boldsymbol{F}_n = \boldsymbol{F}$,

$$\boldsymbol{F} \otimes \cdots \otimes F \neq \boldsymbol{F}^{\otimes n},$$

because the space on the right consists of symmetric functions while the space on the left typically contains also non-symmetric functions.

In contrast, when we introduce the symmetric tensor products, using "$\tilde{\otimes}$" instead of "$\otimes$".

$$\boldsymbol{F} \tilde{\otimes} \cdots \tilde{\otimes} F = \boldsymbol{F}^{\tilde{\otimes} n} = \boldsymbol{F}^{\otimes n} \tag{18.1}$$

It is obvious when $n = 2$, for we may use the polarization formula

$$xy = \frac{1}{2}\Big((x+y)^2 - x^2 - y^2\Big).$$

[3] as follows, the required feature of the objects whose tensor products are considered is the ability to be multiplied. Hence, instead of real function, complex functions may be considered, as well as functions whose values are real or complex functions again, e.g., random variables.

Its extension[4] to n-factors, known as **Mazur-Orlicz polarization formula** presents as follows:

$$\begin{aligned} f_1\tilde{\otimes}\cdots\tilde{\otimes}f_n\,(t_1,\ldots,t_n) &= \frac{1}{n!}\sum_\sigma f_1(t_{\sigma_1})\cdots f_n(t_{\sigma_n}) \\ &= \frac{1}{n!}\sum_{\delta\in\{\,0,1\,\}^n}(-1)^{n-|\delta|}(\sum_i \delta_i f_i(t_i))^{\otimes n}. \end{aligned}$$

Here, $\delta = (\delta_1, \ldots, \delta_n)$ is a 0-1 multindex (vector), and $|\delta| = \sum_i |\delta_i|$ counts the number of 1's in it. The formula says exactly what (18.1) does.

18.2 Tensor operations on tensor products

Suppose that we have operators (linear mappings) $X, X_1, \ldots, X_n : \boldsymbol{F} \to \boldsymbol{L}$, where $\boldsymbol{L}$ is another vector space that allows multiplication. In the context of stochastic integration, we may think of stochastic integrals or series:

Examples.

1. X is a Brownian motion on $T = [0,1]$, $\boldsymbol{F} = L^2[0,1]$, and $\boldsymbol{L} = L^2(\Omega)$.

2. X is a Poisson process on $T = [0,\infty)$, $\boldsymbol{F} = \big\{\, f : \int(|f|\wedge 1)dx < \infty \,\big\}$, and $\boldsymbol{L} \subset L^0(\Omega)$[5].

3. $X = (r_i)$ is a Rademacher sequence, $T = \mathbb{N}$, $\boldsymbol{F} = \ell^2$, $\boldsymbol{L} \subset \bigcap_p L^p(\Omega)$ (integrals, i.e., Rademacher series, posses all moments).

Let us introduce the tensor product (then, tensor powers) of operators:

$$\begin{aligned} &X_1 \otimes \cdots \otimes X_n : \boldsymbol{F}_1 \otimes \cdots \otimes \boldsymbol{F}_n \to \mathbb{R} \\ &X_1 \otimes \cdots \otimes X_n f_1 \otimes \cdots \otimes f_n \stackrel{\mathsf{df}}{=} X_1 f_1 \cdots X_n f_n \end{aligned}$$

Following Example 3, where $Xa = \sum_i a_i r_i$, we would have

$$X^{\otimes n} a^{\otimes n} = (\sum_i a_i r_i)^n$$

---

[4]another polarization may use ± 1 instead, extending $xy = [(x+y)^2 - (x-y)^2]/4$,

$$f_1\tilde{\otimes}\cdots\tilde{\otimes}f_n\,(t_1,\ldots,t_n) = \frac{1}{2^n}\sum_{\sigma\in\{\,-1,1\,\}^n}(-1)^{???}(\sum_i \sigma_i f_i(t_i))^{\otimes n}.$$

Exercise: establish the correct exponent of the sign -1

[5]Exercise: Is the range $\boldsymbol{L}$ of the Poisson integral a proper subset of $L^0(\Omega)$?

Following Example 2, where $Xf = \sum_i f(\tau_i)$ and τ_i are Poisson arrival times,

$$X^{\otimes n} f_1 \otimes \cdots \otimes f_n = \sum_{i_1,\dots,i_n} f_1(\tau_{i1}) \cdots f_n(\tau_{in}).$$

In Example 1, taking for simplicity $n = 2$,

$$X^{\otimes 2} f \tilde{\otimes} f = (Xf)^2 = 1 + 2 \int_0^1 f(t) \int_0^t f(s)\, X(ds)\, X(dt)$$

By linearity, the operator defined for tensor products extends to their linear combinations, which establishes the operator acting on, e.g., $\boldsymbol{F}^{\otimes n}$.

18.3 The problem of diagonals

Diagonals in the n-dimensional cube are uniquely determined by equivalence relations on, or the partitions of, the set $[1, n]$. Let $\kappa = \{ K_1, \dots, K_r \}$, where $K_i \subset [1, n]$, $K_i \cap K_j = \emptyset$, if $i \neq j$, be such an equivalence relation $\sim$ (more precisely, the set of the induced equivalence classes). That is, $i \sim j$, if i, j belong to the same subset. The corresponding diagonal is of the form

$$D(\kappa) = \{ (t_1, \dots, t_n) \in T^n : t_i = t_j \text{ if } i \sim j \}.$$

What is the "diagonal" based on the partition into singletons, i.e., $D(\{1\}, \dots, \{n\})$? This is the union of all diagonal-free tetrahedra, and it is also called a "diagonal", although it may seem strange:

$$T_n = \{ (t_1, \dots, t_n) : t_1 < \cdots < t_n \} = D(\{1\}, \dots, \{n\}).$$

As the dimension $n = 1, 2, 3, 4, \dots$ grows, so do the numbers[6] of diagonals: $1, 2, 5, 15, \dots$ and quite rapidly too, at a rate slightly slower than n^n.

18.4 Diagonal integrals

Similarly to the definition of the quadratic variation of a stochastic process X_t we might consider (signed) variations of order k. However, instead of using the partitions and the limits we will proceed algebraically. The obtained relation will be valid for tensor powers of operators beyond stochastic integrals. Yet we shall stick to the terminology and insight of stochastic integration. Let us list the assumptions

Assumption 1 Suppose that a stochastic integral Xf is well defined for some space $\boldsymbol{F}$ of functions on T.

[6] a.k.a., Bell's numbers

Assumption 2 Supppose that tetrahedral integrals

$$X^{\otimes n}\Big(f^{\otimes n}\mathbb{1}_{T_n}\Big).$$

are well defined for every n; **Assumption 3** Supppose that diagonal integrals

$$X^{\otimes n}\Big(f^{\otimes n}\mathbb{1}_{D_{k,n}}\Big)$$

are well defined for every n and k, where

$$D_{k,n} = D(\kappa_k), \qquad \kappa_k = \{\ \{1,\dots,k\}, k+1,\dots,\{n\}\ \}$$

Next, we introduce the tensor power $X^{\otimes n}$ on $\boldsymbol{F}^{\otimes n}$. In general, the existence of

$$X^{\otimes n}\Big(f^{\otimes n}\mathbb{1}_D\Big)$$

is not guaranteeed, where $D \subset T^n$ (e.g., a tetrahedron or a "true" diagonal). But, without specifying any construction, let us assume . In particular, choosing the partition

$$\kappa_k = \{\ \{1,\dots,k\}, k+1,\dots,\{n\}\ \}$$

and the corresponding diagonal

$$D_k = D(\kappa_k) = \{\,(t_1,\dots,t_n) : t_1 = \dots = t_k,\ \text{and all } t_j \neq t_k\text{'s are distinct for } j > k\,\},$$

we assume that there exist

$$V_k(f) \stackrel{\mathsf{df}}{=} X^{\otimes n}(f^{\otimes n}\mathbb{1}_{D_k}), \quad k = 1,\dots,n \tag{18.2}$$

Denote the main tetrahedron by

$$T_n = \{\,(t_1,\dots,t_n) : t_1 < \dots < t_n\,\}$$

and the off-diagonal integral

$$J_n(f) \stackrel{\mathsf{df}}{=} \frac{1}{n!}X^{\otimes n}(f^{\otimes n}\mathbb{1}_{T_n}).$$

By convention, $J_0 \stackrel{\mathsf{df}}{=} 1$. Also, $J_1 = V_1$.

We will see that two systems of integrals, $(J_1,\dots,J_n)$ and $(V_1,\dots,V_n)$, are equivalent in the sense that one is computable from the other. So, if one type, e.g., of tetrahedral integrals J_k can be constructed, so can be the integrals of the other type.

18.5 Recursion

Denote $D_k^n = D(\{1\},\cdots\{k-1\},\{k,k+1,\ldots,n\})$ for $1 \le k \le n$. Notice that $D_1^n = D(\{1,k+1,\ldots,n\})$ is the union of all tetrahedra. Let the corresponding diagonal integral be V_k^n.

Lemma 18.1 *If $1 \le k \le n$, then*

$$V_k^n = V_{k-1}^{k-1} \cdot V_1^{n-k+1} - (k-1)V_{k-1}^n.$$

Proof. We have

$$\mathbb{1}_{D_k^n} = \prod_{1 \le i < j \le k} \mathbb{1}\{t_i \ne t_j\}\mathbb{1}\{t_k = \cdots = t_n\}.$$

De Moivre's logical law says that

$$\prod_{m=1}^{k-1} \mathbb{1}\{t_m \ne t_k\} = 1 - \max_{1 \le m \le k-1} \mathbb{1}\{t_m = t_k\}.$$

Hence

$$\begin{aligned}
\mathbb{1}\{D_k^n\} &= \prod_{1 \le i < j \le k-1} \mathbb{1}\{t_i \ne t_j\}\Big(1 - \max_{1 \le m \le k-1} \mathbb{1}\{t_m = t_k\}\Big)\mathbb{1}\{t_k = \cdots = t_n\} \\
&= \prod_{1 \le i < j \le k-1} \mathbb{1}\{t_i \ne t_j\}\mathbb{1}\{t_k = \cdots = t_n\} \\
&\qquad - \sum_{m=1}^{k-1} \prod_{1 \le i < j \le k-1} \mathbb{1}\{t_i \ne t_j\}\mathbb{1}\{t_m = t_k = \cdots = t_n\}.
\end{aligned}$$

Since the multiple integral $X^{\otimes n}$ takes the same values for functions with permuted variables, the lemma follows. ■

Theorem 18.2 *Under Assumptions 1, 2, 3, writing $J_k = J_k(f)$, $V_k = V_k(f)$,*

$$nJ_n = \sum_{k=1}^{n} (-1)^{k-1} J_{n-k} V_k. \tag{18.3}$$

Proof. From the lemma:

$$\begin{aligned}
V_n^n &= V_{n-1}^{n-1} V_1^1 - (n-1)V_{n-1}^n \\
&= V_{n-1}^{n-1} V_1^1 - (n-1)V_{n-2}^{n-2} V_1^2 \\
&\qquad + (n-1)(n-2)V_{n-2}^n.
\end{aligned}$$

Therefore, by induction,

$$V_n^n = \sum_{k=1}^{n} (n-1)_{(k-1)} (-1)^{i-1} V_{n-k}^{n-k} V_1^k, \tag{18.4}$$

where $(m)_r \overset{\mathsf{df}}{=} m(m-1)\cdots(m-r+1)$. Substituting

$$V_k^k = k!J_k = X^{\otimes k}(f^{\otimes k}\mathbb{1}_{T_k}),$$

where T_k is the main tetrahedron in T^k, and

$$V_1^k = V_k = X^{\otimes k}(f^{\otimes k}\mathbb{1}_{D_k}),$$

where $D_k = \{\, t_1 = \ldots = t_k \,\}$ is the one-dimensional diagonal in T^k, we obtain equality (18.3) from (18.4). ■

18.6 Chaos polynomials

18.7 Generating functions

For coefficients V_k, J_k define

$$\psi(t) = \sum_{k=1}^{\infty}(-1)^{k-1}V_k t^{k-1}, \quad \text{and} \quad \phi(t) = \sum_{k=0}^{\infty} J_k t^k.$$

Corollary 18.3 *Equality* (18.3) *is equivalent to the formal differential equation,*

$$\phi'(t) = \phi(t)\psi(t),$$

whose solution can be written formally as

$$\phi(t) = \ \exp\left\{ \int_0^t \psi(s)\, ds \right\} \ .$$

The coefficients can be computed by formal differentiation $J_k = \phi^{(k)}(0)/k!$.

The formal procedures are rigorous in the examples below.

Corollary 18.4 *There exists a sequence of polynomials* $p_n(v_1, \ldots, v_n)$ *with the properties*

(i) $p_n(cv_1, \ldots, cv_n) = c^n p_n(v_1, \ldots, v_n)$;

(ii) p_n *is a polynomial of degree* k *of the variable* v_k, $k = 1, \ldots, n$;

such that

$$J_n(f) = p_n(V_1, \ldots, V_n), \quad J_1(f) = V_1.$$

Conversely, there is a polynomial $V_n = V_n(J_1, \ldots, J_n)$, *where*

$$V_n = (-1)^{n-1}\left(nJ_n - \sum_{i=1}^{n-1} J_{n-i}V_i \right).$$

Proof. Use the mathematical induction and formula (18.2). ■

18.8 Examples of chaos polynomials

Hermite polynomials are coefficients, dependent on x, in the series expansion of the function

$$e(t,x) = \exp\left\{ tx - t^2/2 \right\} = \sum_{k=0}^{\infty} t^k H_k(x).$$

The direct computation shows that

$$H_k(x) = \frac{\partial^k e}{\partial t^k}|_{t=0} \quad k \geq 0; \qquad nH_n = H_{n-1}x - H_{n-2}, \quad n \geq 2.$$

For example, $H_0 = 1, H_1(x) = x, H_2(x) = (x^2-1)/2$, etc. If γ is a standard normal random variable, then

$$1 = \mathsf{E} \exp\left\{ t\gamma - t^2/2 \right\} = \sum_{k=0}^{\infty} t^k \mathsf{E}\, H_k(\gamma),$$

hence, $H_0 = 1$ and $\mathsf{E}\, H_k(\gamma) = 0$. Also,

$$e^{st} = \mathsf{E}\, e^{t\gamma - t^2/2 + s\gamma - s^2/2} = \sum_{k,n} t^k s^n \mathsf{E}\, H_k(\gamma) H_n(\gamma).$$

Whence

$$\mathsf{E}\, H_k(\gamma)H_n(\gamma) = 0, \quad \text{if } k \neq n, \text{ and } \quad \mathsf{E}\, H_k^2(\gamma) = 1/k!.$$

The orthonormal sequence $(\sqrt{k!}H_k)$ is complete, i.e., it is an orthonormal basis in $L^2(\phi)$, where

$$\phi(x) = (2\pi)^{-1/2} \exp\left\{ -x^2/2 \right\}\, dx$$

is the Gaussian density.

Example 18.5 *Let $(X_t, t \in T)$ be a square integrable continuous martingale (e.g., the Wiener process) on a finite or infinite interval T. Like before, we assume that $X(0) = 0$, which ensures $\mathsf{E}\, X_t = 0$ and put $\mu(A) = \mathsf{E}\, |X(A)|^2$. Then, the tetrahedral integral of $h^{\otimes n}$ is the n-th Hermite's polynomial of $x = J_1(f)$,*

$$J_n(f) = H_n(I_1(f)) = H_n(Xf)$$

($I_n(f) = H_n(Xf)/\sqrt{n!}$, using the symmetric integral). Indeed, for $f \in \boldsymbol{F} = L^2(T)$ with $\int_T |f|^2\, d\mu = 1$, $x = J_1(f) = V_1(f)$, $V_2 = 1$, $V_k = 0$, for $k \geq 3$. Then $\psi(t) = x - t$ is well defined, so is $\phi(t) = \exp\{xt - t^2/2\}$. Also, the recurrence formula (18.3) takes the form $nJ_n = J_{n-1}J_1 - J_{n-2}V_2$. ■

Example 18.6 *Let $X_t = \xi_t$ be a Poisson process with intensity λ. Then for every set $A \subset [0,\infty)$ of finite measure,*

$$J_k(\mathbb{1}_A) = \xi(A)(\xi(A)-1)\cdots(\xi(A)-k+1).$$

It suffices to take $A = [0, t]$. Let $f = 1\!\!1_{[0,t]}$. Then, $x = J_1 = V_i$, for $i \geq 1$. Hence

$$\psi(s) = \frac{x}{1+s}; \qquad \phi(v) = (1+v)^x.$$

Thus $J_k(1\!\!1_{[0,t]}) = \xi_t(\xi_t - 1) \cdots (\xi_t - k + 1)$. ■

Example 18.7 *Let $X_t = \xi_t - \lambda t$ be a compensated Poisson process with intensity $\lambda > 0$ and $f = 1\!\!1_{[0,t]}$. Consider*

$$x = J_1 = V_1 = X_t, \qquad V_2 = V_3 = \ldots \xi_t = x + \lambda t.$$

Then $\phi(v) = (1+v)^{x+\lambda t} e^{-\lambda t v}$.

Indeed,

$$\begin{aligned} \psi(u) &= \tfrac{1}{u} \left(xu + (x+\lambda t) \sum_{i=2}^{\infty} (-1)^{i-1} u^i \right) \\ &= x - (x+\lambda t) \frac{u}{1+u}. \end{aligned}$$

$\phi^{(n)}(0)/n!$ are called **Poisson-Charlier** or **Charlier polynomials**. ■

Example 18.8 *Let ξ_t be a Poisson process with intensity λ. Denote the atoms of ξ_t by τ_i. Let $J_1 = \xi f$. Then*

$$\phi(v) = \prod_j (1 + v f(\tau_j)).$$

If $J_1 = \xi f - \int f dt$, then

$$\phi(v) = e^{-v \int_0^\infty f(s)\, ds} \prod_j (1 + v f(\tau_j)).$$

For $X_t = \xi_t$,

$$J_1 = V_1 = \xi_1(f), \qquad V_i = \sum_j f^i(\tau_j).$$

We infer that $\phi(v) = \exp\{\, \xi(\ln(1+vf)\, \}$, because

$$\psi(u) = \frac{1}{u} \sum_j \sum_{i=1}^{\infty} (-1)^{i-1} (u f(\tau_j)^i = \xi \left(\frac{f}{1+uf} \right).$$

Functions $\psi(u)$ and $\psi(v)$ are finite, since the Poisson integrals

$$\xi(f/(1+uf)) \quad \text{and} \quad \xi(\ln(1+vf))$$

are finite iff $\xi f < \infty$. ■

Example 18.9 *Consider a symmetrized Poisson process*

$$\widetilde{\xi}_t = \sum_{j=1}^{\xi_t} \varepsilon_j, \quad or \quad \widetilde{\xi}(f) = \sum_j \varepsilon_j f(\tau_j), \ \xi f^2 \ < infty.$$

where (ε_j) *is a Rademacher sequence independent of* ξ_t. *Then for* $J_1 = \widetilde{\xi} f$,

$$\phi(v) = \prod_j (1 + v f(\tau_j).$$

Indeed, for $\xi(f^2) < \infty$,

$$J_1 = V_1 = \widetilde{\xi}(f) \quad \text{and} \quad V_i = \widetilde{\xi}_i(f^i) = \sum_j f^i(\tau_j)\varepsilon_j^i.$$

In other words, $V_{2i} = \xi$ and $V_{2i-1} = \widetilde{\xi}$, $i = 1, 2 \ldots$. Therefore,

$$\begin{aligned} t\psi(t) &= \widetilde{\xi}\left(\sum_{i=1}^{\infty}(-1)^{i-1}(tf)^i\right) \\ &= \sum_j \sum_{i=1}^{\infty}(-1)^{i-1}t^i f^i(\tau_j)\varepsilon_j^i = \psi_\varepsilon(t) - \psi_0(t), \end{aligned}$$

where

$$\begin{aligned} t\psi_\varepsilon(t) &= \sum_j (\sum_{i=0}^{\infty}(tf(\tau_j))^{2i+1})\varepsilon_j = \widetilde{\xi}(tf/(1-(tf)^2)), \\ t\psi_0(t) &= \sum_j \sum_{i=1}^{\infty}(tf(\tau_j))^{2i} = \xi(t^2 f^2/(1-(tf)^2)). \end{aligned}$$

Since $\xi f^2 < \infty$, both series $\psi_\varepsilon(t)$, $\psi_0(t)$ converge in probability. The integral $\int_0^t \psi(u)\, du$ exists for almost everywhere on the set

$$\left\{ \sup_j (tf(\tau_j))^2 < 1 \right\} \supset \left\{ \sum_j (tf(\tau_j))^2 < 1 \right\} = \left\{ \xi((tf)^2) < 1 \right\}.$$

Hence

$$\begin{aligned} \ln\phi(t) = \int_0^t \psi(u)\, du &= \int_0^t \psi_\varepsilon(u)\, du - \int_0^t \psi_0(u)\, du \\ &= \tfrac{1}{2}\left(\widetilde{\xi}\left(\ln\tfrac{1+ft}{1-ft}\right) + \xi(\ln(1-(tf)^2)\right) \\ &= \ln\prod_j (1 + \varepsilon_j f(\tau_j)t). \end{aligned}$$

The last equality follows from the formula

$$\ln(1 + a\varepsilon) = \frac{1}{2}\left(\varepsilon \ln\frac{1+a}{1-a} + \ln(1-a^2)\right), \tag{18.5}$$

where ε is a Rademacher variable. We must confine initially to the set $\{\, \xi(tf)^2 < 1 \,\}$. Next, we let $t \to 0$, and this set will increase to Ω. ∎

Example 18.10 *Let* (θ_j) *be a sequence of either symmetric or nonnegative i.i.d. random variables, independent of a Poisson process* ξ_t. *Consider the compound Poisson integral* $Xf = \sum_j \theta_j f(\tau_j)$. *If* $|v| \le 1/\sup_j |\theta_j f(\tau_j)|$ *then*

$$\phi(v) = \exp\left\{ \sum_j \ln(1 + v\theta_j f(\tau_j)) \right\} = \prod_j (1 + v\theta_j f(\tau_j)),$$

Indeed, if the integral Xf exists, then $\sup_j |\theta_j f(\tau_j)| < \infty$ a.s. Then, we mimic the argument from the preceding example. ■

Let us replace ξ by a locally finite jump process $Qf = \sum_n f(\tau_n)$ or $\widetilde{Q}f = \sum_n \varepsilon_n f(\tau_n)$. Then, formulae for the generating function $\phi(v)$ in Examples 18.6–18.9 hold verbatim if f is a locally bounded function.

Exercise 18.1 Replace the latter constraint by integrability conditions, appropriately for each formula.