Brownian motion and stochastic calculus:
Errata and supplementary material

Martin Larsson

1 Course content and exam instructions

The course covers everything in the script except Sections 1.2–1.3 (Donsker's theorem and applications) and Section 4.9 (Backward stochastic differential equations). We also did not discuss the rather long Example 2.10 on page 66–68. The material in the present document (the errata and clarifications, as well as the proof of Itô's representation theorem and the review of regular conditional distributions) is also included. **Note in particular the points labeled with (∗) below**. Note that the exercises contain material that is important for the course; for instance, components of some proofs appear as exercises.

Instructions for the exam. There is a 30 minute oral exam. Two of the following questions will be randomly selected for the exam. You will have 5 minutes to prepare, after which there will be a 25 minute discussion based on the two questions. While preparing, you will have access to the script. The discussion will involve statements and proofs, as time allows. Note that there may be digressions, for instance about supporting definitions and results.

The numbering below first indicates the chapter and then the numbering in the script. For example, "Theorem 2.3.8" refers to Theorem 3.8 in Chapter 2.

(1) Proposition 1.5.9 and Theorem 1.5.11 (Lévy–Ciesielski construction of Brownian motion)
(2) Proposition 2.1.1, Corollary 2.1.2, Proposition 2.3.4 (properties of Brownian motion)
(3) Theorem 2.2.1 (nowhere differentiability of Brownian paths)
(4) Theorem 2.3.8 (stopping theorem) and Corollary 2.3.10 (maximal inequality)
(5) Theorem 2.1.5 (the law of the iterated logarithm), proof given in Section 2.4.
(6) Definitions 3.1.4 and 3.1.6 (Markov processes) and Proposition 3.1.11 (characterization of Markov property)
(7) Definition 3.2.2 (shift operator) and Proposition 3.2.4 (Markov property for the canonical process), Definition 3.2.13 (strong Markov property)

(8) Definition 3.2.18 (Feller property) and Theorem 3.2.19 (right-continuous Feller processes are strong Markov)
(9) Theorem 3.3.2 (reflection principle) and Corollary 3.3.7.
(10) Theorem 3.3.8 (arcsine law) and Corollary 3.3.5.
(11) Definition 3.4.2 (infinitesimal generator) and Proposition 3.4.7 (a Feller process solves the martingale problem for its infinitesimal generator)
(12) Theorem 3.4.9 (uniquely solvable martingale problem implies strong Markov property)
(13) Theorem 4.1.5 (quadratic variation of continuous local martingales)
(14) Proposition 4.1.10 (Kunita-Watanabe inequality)
(15) Theorem 4.2.5 (stochastic integral for square-integrable integrands)
(16) Properties of stochastic integrals: Proposition 4.2.13 (associativity and stopping) and Theorem 4.2.19 (dominated convergence)
(17) Theorem 4.3.4 (Itô's formula)
(18) Theorem 4.3.13 (Lévy's characterization of Brownian motion)
(19) Theorem 4.3.16 (BDG inequalities)
(20) Theorem 4.4.6 (Girsanov transformation), including Proposition 4.4.4.
(21) Theorem 4.6.3 (Itô's representation theorem); see also Theorem 3.1 below.
(22) Definition 4.7.2 (strong solution to SDE) and Theorem 4.7.4 (existence and uniqueness), comparison with weak solutions.
(23) Theorem 4.8.11 (Feynman-Kac formula)
(24) Definition 5.1.1 (Lévy process) and statement of Theorem 5.1.4.
(25) Theorem 5.2.6 (a restarted Lévy process is Lévy)
(26) Proposition 5.3.6 (sums of jumps of Lévy processes)
(27) Theorem 5.4.1 (Lévy–Itô decomposition), formulation, interpretation, and main steps of the proof.

2 Errata and clarifications as of 09.06.2017

Errors and typos from the script are listed below, along with corrections. Some instances are also included which, while mathematically correct, may merit additional clarification. There is no guarantee that the list below is complete. **Some points are labeled with ($*$) to indicate that they are particularly important.**

Page 5: In (0.9), the statement that the σ-field $\mathcal{B}(\mathbb{R})^{[0,\infty)}$ is "much smaller than the product-σ-field on $\mathbb{R}^{[0,\infty)}$" should be replaced by the statement that it is "much smaller than the Borel σ-field on $\mathbb{R}^{[0,\infty)}$ equipped with the product topology".

Let us provide some further explanation. Given a measurable space $(S, \mathcal{S})$ and an arbitrary set J, define $S^J := \{\text{all maps } f\colon J \to S\}$. The *product σ-field on S^J*, denoted $\mathcal{S}^J$, is defined as the smallest σ-field such that the coordinate map $S^J \to \mathbb{R}$, $f \mapsto f(t)$, becomes measurable for every $t \in J$. In symbols, this can be expressed as $\mathcal{S}^J := \sigma(Y_t \colon t \in J)$, where Y_t, $t \in J$, denote the coordinate maps. In the context of (0.9), we have $S = \mathbb{R}$ and $J = [0, \infty)$, and therefore

$$\mathcal{B}(\mathbb{R})^{[0,\infty)} \text{ is equal to the product } \sigma\text{-field on } \mathbb{R}^{[0,\infty)}.$$

Suppose now that S is a topological space and $\mathcal{S}$ is the Borel σ-field. Then S^J can be equipped with the *product topology*, namely the smallest topology such that the coordinate map $f \mapsto f(t)$ becomes continuous for every $t \in J$. Convergence in this topology is equivalent to *pointwise convergence*: $f_\alpha \to f$ in S^J if and only if $f_\alpha(t) \to f(t)$ in S for every $t \in J$. As usual, let $\mathcal{B}(S^J)$ denote the Borel σ-field on S^J. The key point is now that if J is countable, then $\mathcal{B}(S^J) = \mathcal{B}(S)^J$. But if J is uncountable, as in (0.9), and if S is rich enough ($S = \mathbb{R}$ is enough) then $\mathcal{B}(S^J) \supsetneq \mathcal{B}(S)^J$.

Page 25: In the proof of Proposition 4.3, on page 25, line 4, "$|Y_{n,1}$" should be replaced by "$|Y_{n,1}|$" (an absolute value bar is missing). Moreover, on line 6, the unspecified constant should be inside the exponential. That is, one has

$$P[M_n \le a_n] \le \exp\left(-c_1 \frac{k_n^{1-\alpha}}{\sqrt{\log k_n}}\right) =: f(k_n)$$

for some constant c_1, where now $f(x) := \exp\left(-c_1 \frac{x^{1-\alpha}}{\sqrt{\log x}}\right)$. This leads to minor changes in the subsequent lines, but the argument remains the same: using that $k_n \ge c_2 n$ for some constant c_2, one derives

$$f(k_n) \le \exp\left(-c_3 \frac{k_n^{1-\alpha}}{\sqrt{\log k_n}}\right) \quad \text{for all large } n,$$

where c_3 is a different constant (for example, $c_3 = c_1 c_2^{1-\alpha}/2$ does the job). This still implies that $\sum_{n=1}^{\infty} P[M_n \le a_n] < \infty$, as needed to apply Borel–Cantelli.

Page 27: On the second-to-last line, "$f_0 \equiv 0$" should be replaced by "$f_0 \equiv 1$".

Page 29: On the third-to-last line, "$\varphi_{m,\ell}(t)$" should be replaced by "$\varphi_{n,k}(t)$".

Page 32: On line 10, "$J_{n,k}$" should be replaced by "$J_{k,n}$". Moreover, on line 16, all four occurrences of n should be replaced by N.

Page 36: On the first line, the symbol $\mu|_{C(0,1]}$ denotes the law under μ of the process $Y|_{(0,1]} := (Y_t)_{0<t\le 1}$, where $Y = (Y_t)_{0\le t\le 1}$ is the coordinate process on $C[0,1]$. In particular, $\mu|_{C(0,1]}$ is a probability measure on $C(0,1]$.

Page 41: There is no error here, but on line 6, "$t_i^n \wedge t$" and "$t_{i+1}^n \wedge t$" can be replaced by "t_i^n" and "t_{i+1}^n", respectively, since it has been assumed that $t = 1$. Here, while not explicitly mentioned in the proof, it should be clear that t_i^n and t_{i+1}^n are taken from the given partition Π_n of $[0, 1]$.

Page 44: On line 12, "$e^{-\alpha(W_t - W_s)}$" should be replaced by "$e^{\alpha(W_t - W_s)}$" inside the conditional expectation. Note that, strictly speaking, there is no error here because $-\alpha(W_t - W_s)$ and $\alpha(W_t - W_s)$ both have the same $\mathcal{F}_s$-conditional law, namely $\mathcal{N}(0, \alpha^2(t - s))$.

(∗) **Pages 44–45:** On page 44, line 19, the definition of $\mathcal{F}_\tau$ should read

$$\mathcal{F}_\tau := \{A \in \mathcal{F}_\infty \colon A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\} . \tag{1}$$

The difference is that "$\mathcal{F}$" has been replaced by "$\mathcal{F}_\infty$". This is to ensure that for $\tau \equiv \infty$ we have $\mathcal{F}_\tau = \mathcal{F}_\infty$, and is consistent with the book by Revuz and Yor (see page 42 in the book). Due to this change, an extra assumption is needed in Lemma 3.7 on page 45. Namely, on line 14 we must assume that X_∞ is not only well-defined, but also $\mathcal{F}_\infty$-measurable.

Page 45: This is a follows-up on the previous point. On the third-to-last line on page 45, we should define $M_\infty(\omega) := \liminf_{n\to\infty} M_n(\omega)$, which exists for *every* ω and gives an $\mathcal{F}_\infty$-measurable random variable. The application of Lemma 3.7 on page 46, line 8, is then unproblematic. Note that the martingale convergence theorem still tells us that, in fact, $M_\infty = \lim_{n\to\infty} M_n$ holds P-a.s. and in $\mathcal{L}^1(P)$.

(∗) **Page 46:** Remark 3.9 should be updated to state that "If $(\Omega, \mathcal{F}, \mathbb{F}, P)$ satisfies the usual conditions, then any martingale has a version with RCLL paths." Here, a martingale $M = (M_t)_{t\geq 0}$ having RCCL paths means that $M_\cdot(\omega)$ is an RCLL function for *every* ω. A good reference for this is Revuz and Yor, Theorem II.2.9. If the filtration is right-continuous but not complete, martingales do have right-continuous versions, but left limits might fail to exist on a nullset.

Page 51: On line 12, the "const." in front of $\left(\log \frac{1}{h_{n+1}}\right)^{-(1+\delta)}$ is not needed.

(∗) **Page 57:** In Definition 1.6, $P[X_{t+h} \in A \mid \mathcal{G}_t]$ stands for $E[\mathbf{1}_{\{X_{t+h}\in A\}} \mid \mathcal{G}_t]$, and this notation is used in general throughout the script. See also the review of regular conditional distributions in Section 4 below.

Page 59: In Proposition 1.11, the functions $f_0, f_1, \dots, f_n$ should in addition be bounded. Moreover, the notation $\nu \otimes K_{t_1} \otimes K_{t_2-t_2} \otimes \cdots \otimes K_{t_n-t_{n-1}}$ represents the measure on the product σ-field $\mathcal{S} \otimes \cdots \otimes \mathcal{S}$ ($n+1$ copies) given by

$$(\nu \otimes K_{t_1} \otimes K_{t_2-t_2} \otimes \cdots \otimes K_{t_n-t_{n-1}})(A_0 \times A_1 \times \cdots \times A_n)$$
$$:= \int_{A_0} \nu(dx_0) \int_{A_1} K_{t_1}(x_0, dx_1) \int_{A_2} K_{t_2-t_1}(x_1, dx_2) \cdots \int_{A_n} K_{t_n-t_{n-1}}(x_{n-1}, dx_n).$$

Page 77: On line 10 it is stated that the proof of Theorem 3.8 does not use the strong Markov property. This is not accurate: the proof uses Corollary 3.5, which uses the reflection principle for Brownian motion, which uses the strong Markov property.

Page 79: Lines 4–7 describe why X is a Markov process, but they do not reference the fact that τ is exponential. This fact is however crucial. A calculation shows ($\to$ exercise) that for any $x \geq 0$ and without assumptions on τ,

$$P[X_{t+h} > x \mid \mathcal{F}_t^0] = P[\tau - t < h - x \mid \tau \geq t]\mathbf{1}_{\{X_t=0\}} + \mathbf{1}_{\{X_t>(x-h)^+\}}.$$

But τ is exponential, so the conditional distribution of $\tau - t$ given that $\tau \geq t$ is again exponential. Therefore,

$$P[X_{t+h} > x \mid \mathcal{F}_t^0] = K_h(X_t, (x, \infty)) := \begin{cases} \left(1 - e^{-(h-x)^+}\right)\mathbf{1}_{\{X_t=0\}} + \mathbf{1}_{\{X_t>(x-h)^+\}}, & x \geq 0, \\ 1, & x < 0, \end{cases}$$

which gives the Markov property.

Page 85: On line 1, "Markov process" should be replaced by "Feller process".

Page 85: In Theorem 4.9, "(for $\mathbb{F}$ sufficiently large)" should be deleted; the proof does not depend on it. Nonetheless, the assumption that the martingale problem $MP(\mathcal{A}, F)$ has a unique solution for every $x \in S$ can only be satisfied if the family $F \subseteq C_0(S)$ is rich enough. As an extreme illustration, convince yourself that uniqueness will fail in general if $F = \{0\}$ only contains the zero function!

Page 85: An important part of the proof of Theorem 4.9 is to verify that (4.10) holds. However, not all of (4.10) is discussed in the script. What is not argued is that the starting point under $R(\omega, \cdot)$ is $Y_\tau(\omega)$. More precisely, the statement which is not proved in the script is

$$R(\omega, \{Y_0 = Y_\tau(\omega)\}) = 1 \text{ for } \mathbb{Q}_x\text{-a.e. } \omega. \tag{2}$$

Some details for this were given in class. Specifically, let $H\colon \mathbb{D}(S)\times\mathbb{D}(S)\to\mathbb{R}$ be the measurable function given by

$$H(\omega',y) := \mathbf{1}_{\{y(0)=Y_\tau(\omega')\}} \quad \text{for any } (\omega',y)\in\mathbb{D}(S)\times\mathbb{D}(S).$$

Then, by a property of regular conditional distributions (see Lemma 4.3 below), one has the delicate equality

$$\int_{\mathbb{D}(S)} H(\omega,y)R(\omega,dy) = E_{\mathbb{Q}_x}\left[H(\cdot,Y_{\tau+\cdot}) \mid \mathcal{Y}_\tau\right](\omega) \quad \text{for } \mathbb{Q}_x\text{-a.e. } \omega.$$

The left-hand side is equal to $R(\omega,\{Y_0=Y_\tau(\omega)\})$, while the right-hand side is equal to one, since $H(\omega',Y_{\tau+\cdot}(\omega')) = \mathbf{1}_{\{Y_{\tau+0}(\omega')=Y_\tau(\omega')\}} = 1$. This proves (2).

Page 88: In Remark 1.3, one must additionally assume that Z_0 is bounded, i.e. $|Z_0|\le c$ P-a.s. for some constant c. As an illustration, consider a constant process $Z\equiv Z_0$, where $Z_0\sim N(0,1)$ is a standard normal random variable!

($*$) Pages 92 and 106: Remark 1.15 on page 92 is somewhat inaccurate. Consider a continuous semimartingale $X=X_0+M+A$, where $M\in\mathcal{M}^c_{0,\mathrm{loc}}$ and $A\in cFV_0$. We then *define* the quadratic variation process $\langle X\rangle$ by

$$\langle X\rangle := \langle M\rangle.$$

(This is not clearly stated in the script, but has been mentioned several times in class.) Observe that this is well-defined. Indeed, if $X=X_0+M'+A'$ is another decomposition of X with $M'\in\mathcal{M}^c_{0,\mathrm{loc}}$ and $A'\in cFV_0$, then $M-M'=A'-A\in\mathcal{M}^c_{0,\mathrm{loc}}\cap cFV_0$, so that $M'=M$ by Proposition 1.4 on page 88. In particular, $\langle M'\rangle=\langle M\rangle$. Furthermore, if Y is another continuous semimartingale, we use polarization to define

$$\langle X,Y\rangle := \frac{1}{4}\left(\langle X+Y\rangle-\langle X-Y\rangle\right) = \frac{1}{4}\left(\langle M+N\rangle-\langle M-N\rangle\right) = \langle M,N\rangle.$$

Coming back to Remark 1.15, the statement that "X has quadratic variation $\langle M\rangle$" is now a matter of definition.

Next, on page 106 it is claimed that $\langle X^k,X^\ell\rangle$ satisfies (3.3) for all $t\ge 0$, P-a.s., where X is an $\mathbb{R}^d$-valued continuous semimartingale with components $X^k=X^k_0+M^k+A^k$, $k=1,\dots,d$. This follows from (1.9) on page 91, along with the fact that A^k and A^ℓ are FV. Indeed, writing $\Delta_i Y_t := Y_{t_{i+1}\wedge t}-Y_{t_i}$ for any process Y, observe that

$$\sum_{\substack{t_i\in\Pi_n,\\ t_i\le t}} (\Delta_i X^k_t)(\Delta_i X^\ell_t) = \sum_{\substack{t_i\in\Pi_n,\\ t_i\le t}} (\Delta_i M^k_t)(\Delta_i M^\ell_t) + R^n_t,$$

where

$$R_t^n := \sum_{\substack{t_i \in \Pi_n, \\ t_i \le t}} \Big((\Delta_i X_t^k)(\Delta_i A_t^\ell) + (\Delta_i A_t^k)(\Delta_i M_t^\ell) \Big) .$$

Since almost all trajectories of A^k and A^ℓ are of finite variation, one has $\lim_{n\to\infty} R_t^n = 0$ for all $t \geq 0$, P-a.s. One then obtains (3.3) on page 106 from (1.9) on page 91, using also that $\langle X^k, X^\ell \rangle = \langle M^k, M^\ell \rangle$ by definition.

Page 93: On line 9, "since our filtration $\mathbb{F}$ is right-continuous" should be replaced by "since our filtration $\mathbb{F}$ satisfies the usual conditions". Without completeness, martingales need not have RCLL versions in general. (However, right-continuity of $\mathbb{F}$ *is* enough to prove that any martingale has a *right-continuous* version, and even an *almost surely RCLL* version.)

Page 95: In part 2) of Remark 2.2, one has to assume that H_0 is $\mathcal{F}_0$-measurable to deduce that H is optional.

Page 107: Line 16 (the display) should be replaced by:

$$\sum_{\substack{t_i \in \Pi_n, \\ t_i \le s}} (\Delta_i x)^2 \longrightarrow \langle X \rangle_s(\omega) \quad \text{for all } s \in [0,t] \text{ as } n \to \infty.$$

Furthermore, on line 18, "$\langle x \rangle$" should be replaced by "$\langle X \rangle(\omega)$", and similarly on line 21, "$\int_0^t g(s) d\langle x \rangle(s)$" should be replaced by "$\int_0^t g(s) d\langle X \rangle_s(\omega)$".

Page 115: Some care is needed in the calculation in the proof of Proposition 4.4, since only $Q \ll P$ is assumed but not $Q \approx P$. A detailed calculation is as follows:

$$\begin{aligned}
E_Q[U_\tau \mathbf{1}_A] &= E_Q[U_\tau \mathbf{1}_{A \cap \{Z_\sigma > 0\}}] && (Z_\sigma > 0,\ Q\text{-a.s.}) \\
&= E_P[Z_\tau U_\tau \mathbf{1}_{A \cap \{Z_\sigma > 0\}}] && (\text{Lemma 4.2}) \\
&= E_P\left[E_P[Z_\tau U_\tau \mid \mathcal{F}_\sigma] \mathbf{1}_{A \cap \{Z_\sigma > 0\}} \right] && (\text{tower rule}) \\
&= E_P\left[Z_\sigma \frac{1}{Z_\sigma} E_P[Z_\tau U_\tau \mid \mathcal{F}_\sigma] \mathbf{1}_{A \cap \{Z_\sigma > 0\}} \right] && (Z_\sigma / Z_\sigma = 1 \text{ on } \{Z_\sigma > 0\}) \\
&= E_Q\left[\frac{1}{Z_\sigma} E_P[Z_\tau U_\tau \mid \mathcal{F}_\sigma] \mathbf{1}_{A \cap \{Z_\sigma > 0\}} \right] && (\text{Lemma 4.2}) \\
&= E_Q\left[\frac{1}{Z_\sigma} E_P[Z_\tau U_\tau \mid \mathcal{F}_\sigma] \mathbf{1}_A \right] && (Z_\sigma > 0,\ Q\text{-a.s.})
\end{aligned}$$

The point that is not explicit in the script is that $Z_\sigma > 0$ Q-a.s., which however follows from an obvious calculation: $Q[Z_\sigma = 0] = E_P[Z_\sigma \mathbf{1}_{\{Z_\sigma = 0\}}] = 0$.

Page 116: On line 20, “Z” should be replaced by “$Z - Z_0$”.

Page 118: On line 4, Novikov’s condition is correctly claimed to imply $\langle L\rangle_\infty < \infty$ P-a.s. However, more is true: Due to the inequality $x \leq e^{x/2}$, Novikov’s condition even implies $\langle L\rangle_\infty \in L^1(P)$. Therefore $L \in \mathcal{H}_0^{2,c}$ by Lemma 1.18 on page 94, and in particular L_∞ exists and is finite, so that $Z_\infty := \mathcal{E}(L)_\infty = \exp(L_\infty - \frac{1}{2}\langle L\rangle_\infty) > 0$ P-a.s. We deduce that the measure Q defined by $dQ := Z_\infty dP$ is actually *equivalent to* P.

Page 127: On line 7, “functions a, b” should be replaced by “measurable functions a, b”.

Page 130: On line 16, “every strong solution is also a weak solution” should be replaced by “every strong solution gives rise to a weak solution”. This is an issue of terminology: a weak solution is a tuple $(\Omega, \mathcal{F}, \mathbb{F}, Q, W, X)$, while a strong solution is a process X with certain properties. They can therefore not be the same, but once a strong solution is given (which in particular means that some probability space, filtration, etc. are also given), one can easily build a weak solution.

Page 133: On line 5, “Markov” should be replaced by “Feller”.

Page 152: On line 20, the last equality in the display, $E[X_{t-s}] - (t-s)E[X_1] = 0$, needs an argument. Taking $d = 1$ for simplicity, one can use that $E[e^{iuX_h}] = e^{h\psi(u)}$ to get

$$E[X_h] = -i\frac{d}{du}e^{h\psi(u)}|_{u=0} = -i\psi'(0)h,$$

and hence $E[X_h] = hE[X_1]$. See also Exercise 14.2(e).

Page 155: Before speaking about the strong Markov property of a Lévy process X, we should make sure that X is indeed a Markov process. Fortunately, this easily follows from the defining properties (L1)–(L2), with the transition semigroup given by $K_h(x, A) := P[x + X_h \in A]$. (Exercise: check this! See also Example 1.9 on page 58.)

Page 157: On line 2, “$f\colon \mathbb{R}^d \to \mathbb{R}$” should be replaced by “$f\colon \mathbb{R}^d \to \mathbb{R}^d$”. Indeed, on line 5 we wish to take $f(x) = x$ for $x \in \mathbb{R}^d$.

Page 159: On lines 12–13, “In particular, the measure ν from (3.3) is finite.” should be deleted.

Page 160: On line 2, the inequality $P[|X_t| > 2nC] \le e^t\alpha^n$ is only proved for integer n. Therefore, on line 4, one has to write

$$P\left[|X_t| > \frac{\log m}{\gamma}\right] \le P\left[|X_t| > 2C\left\lfloor\frac{\log m}{2C\gamma}\right\rfloor\right] \le e^t\alpha^{\left\lfloor\frac{\log m}{2C\gamma}\right\rfloor} \le \frac{1}{\alpha}e^t\alpha^{\frac{\log m}{2C\gamma}},$$

where $\lfloor x\rfloor$ denotes the integer part of x. The subsequent calculations work as before, up to an inconsequential factor $1/\alpha$.

3 An alternative proof of Itô's representation theorem

Itô's representation theorem (Theorem 6.3 on page 123) can be stated as follows, where the setting is as described in (6.1) on page 123.

Theorem 3.1. *Every $(P, \mathbb{F}^W)$-local martingale N is of the form*

$$N = N_0 + \int H dW \quad \textit{for some } H \in L^2_{\mathrm{loc}}(W), \tag{3}$$

and in particular admits a continuous version. Consequently, every random variable $F \in L^1(\mathcal{F}^W_\infty, P)$ admits a unique representation

$$F = E[F] + \int_0^\infty H_s dW_s \qquad P\textit{-a.s.} \tag{4}$$

for some $H \in L^2_{\mathrm{loc}}(P)$ such that $\int H dW$ is a uniformly integrable martingale.

The following proof based on the Kunita-Watanabe decomposition was given in class; it differs somewhat from the (sketch of) proof given in the notes.

Proof. (a) The representation (4) follows by applying (3) to $N_t := E[F \mid \mathcal{F}^W_t]$, $t \ge 0$. The uniqueness statement follows because $H_t = d\langle N, W\rangle_t/dt$ up to nullsets.

(b) It is therefore enough to prove (3). We first do this for N locally in $\mathcal{H}^2$. By localization, and after subtracting N_0, we may assume that $N \in \mathcal{H}^2_0$. Then, by Lemma 5.4 (with $M := W$) and Proposition 5.5 (with $\mathcal{A} := \mathcal{I}(M) = \mathcal{I}(W)$), we get

$$N = H \cdot W + L \quad \text{for some } H \in L^2(W) \text{ and } L \perp \mathcal{I}(W),$$

where "$\perp$" as usual denotes strong orthogonality. We must show that $L = 0$ P-a.s.

(c) We claim that for any deterministic $K = (K_s)_{s\ge 0}$ with $\int_0^\infty K_s^2 ds < \infty$, the random variable

$$Z := \exp\left(i\int_0^\infty K_s dW_s\right)$$

satisfies $E[L_t Z] = 0$ for all $t \geq 0$. The proof proceeds by Itô's formula in manner reminiscent of the proof of Lévy's characterization of Brownian motion, Theorem 3.13 on page 111. The independent increments property of Brownian motion yields

$$M_t := E[Z \mid \mathcal{F}_t^W] = \exp\left(i \int_0^t K_s dW_s\right) E\left[\exp\left(i \int_t^\infty K_s dW_s\right)\right],$$

and since $\int_t^\infty K_s dW_s$ is $\mathcal{N}(0, \int_t^\infty K_s^2 ds)$ distributed,

$$M_t = \exp\left(i \int_0^t K_s dW_s - \frac{1}{2} \int_t^\infty K_s^2 ds\right).$$

Itô's formula applied to the real and imaginary parts of M gives

$$M_t = M_0 + i \int_0^t M_s K_s dW_s,$$

so that (the real and imaginary part of) $M - M_0$ lies in $\mathcal{I}(W)$. Thus LM is a martingale, and we get

$$E[L_t Z] = E[L_t M_t] = L_0 M_0 = 0.$$

This proves the claim.

(d) Applying what we just proved with K of the form $K_s := \sum_{j=1}^n \lambda_j \mathbf{1}_{[0,t_j]}(s)$ gives, for all $t \geq 0$,

$$E\left[L_t \exp\left(i \sum_{j=1}^n \lambda_j W_{t_j}\right)\right] = 0 \text{ for all } t_1, \ldots t_n \in [0, \infty),\ \lambda_1, \ldots, \lambda_n \in \mathbb{R}. \tag{5}$$

To prove that this implies $L_t = 0$, we follow the last part of Lemma 6.7 in the script. Define the signed measure μ on $\mathcal{F}_\infty^W$ by $\frac{d\mu}{dP} := L_t$, and observe that the left-hand side of (5) is the Fourier transform of the image measure $\mu \circ (W_{t_1}, \ldots, W_{t_n})^{-1}$ on $\mathcal{B}(\mathbb{R}^n)$. Since this Fourier transform is zero, so is the image measure. Therefore $\mu|_{\sigma(W_{t_1}, \ldots, W_{t_n})} = 0$ (the zero measure) for any $t_1, \ldots, t_n$, which by an application of the π-λ-theorem yields $\mu = \mu|_{\mathcal{F}_\infty^W} = 0$. Therefore

$$E[|L_t|] = E[L_t \text{sign}(L_t)] = \int_\Omega \text{sign}(L_t(\omega)) d\mu(\omega) = 0,$$

so that $L_t = 0$ P-a.s. Since $t \geq 0$ was arbitrary, we have $L = 0$ P-a.s. Consequently, (3) holds for every N locally in $\mathcal{H}^2$.

(e) Let now N be an arbitrary local martingale; from here on, the proof follows points 3) and 4) in the outline on page 124 in the script. By localization, we may assume that N is a uniformly integrable martingale. Then $N_t = E[F \mid \mathcal{F}_t^W]$ for $F := N_\infty \in L^1(P)$. Pick

$F^n \in L^2(P)$ such that $F^n \to F$ in $L^1(P)$ (for instance, $F^n := F\mathbf{1}_{\{|F|\leq n\}}$ does the job). The martingales N^n defined by $N_t^n := E[F^n \mid \mathcal{F}_t^W]$ lie in $\mathcal{H}^2$ by Jensen's inequality and therefore, by what we already proved, have continuous paths P-a.s. Doob's submartingale inequality gives, for any $\varepsilon > 0$,

$$P\left[\sup_{t\geq 0} |N_t - N_t^n| \geq \varepsilon\right] \leq \frac{1}{\varepsilon} E\left[|N_\infty - N_\infty^n|\right] = \frac{1}{\varepsilon}\|F - F^n\|_{L^1(P)} \to 0 \quad (n \to \infty).^1$$

Along a subsequence, the continuous processes N^n therefore converge uniformly to N P-a.s., so N must be continuous as well. In particular, N is then locally in $\mathcal{H}^2$, and thus satisfies (3) by what we already proved. □

4 Brief review of regular conditional distributions

Fix a probability space $(\Omega, \mathcal{F}, P)$ and let $X\colon \Omega \to E$ be a random variable with values in a Polish space E.[2] You can think of $E = \mathbb{R}$, but in the script we also consider the case $E = \mathbb{D}(S)$, the space of RCLL paths on a state space S. For a given sub-σ-field $\mathcal{G} \subset \mathcal{F}$, we can consider the conditional probability

$$P(X \in B \mid \mathcal{G}) := E[\mathbf{1}_{\{X\in B\}} \mid \mathcal{G}] \tag{6}$$

for any $B \in \mathcal{B}(E)$. A natural idea is to view (6) as a function of the set argument B, and think of it as a conditional distribution of X given $\mathcal{G}$. The problem, however, is that (6) is uniquely determined only up to a nullset that *may depend on* B, which gives very little control over the behavior of the map $B \mapsto P(X \in B \mid \mathcal{G})(\omega)$ for fixed ω. The notion of regular conditional distribution lets us to circumvent this difficulty.

Definition 4.1. *A* regular conditional distribution of X given $\mathcal{G}$ *is a stochastic kernel* R *from* $(\Omega, \mathcal{G})$ *into* $(E, \mathcal{B}(E))$ *such that* $R(\cdot, B)$ *is a version of* $P(X \in B \mid \mathcal{G})$ *for every* $B \in \mathcal{B}(E)$*. More precisely,* $R\colon \Omega \times \mathcal{B}(E) \to [0, 1]$ *satisfies*

- $\omega \mapsto R(\omega, B)$ *is* $\mathcal{G}$*-measurable for every* $B \in \mathcal{B}(E)$*,*
- $B \mapsto R(\omega, B)$ *is a probability measure on* $\mathcal{B}(E)$ *for every* $\omega \in \Omega$*,*
- *for every* $B \in \mathcal{B}(E)$*, one has* $P(R(\cdot, B) = P(X \in B \mid \mathcal{G})) = 1$*.*

Existence of a regular conditional distribution turns out to depend crucially on the image space E. Fortunately, if E is a Polish space with its Borel σ-field, regular conditional distributions always exist.

[1] Doob's submartingale inequality states that $\varepsilon P(\sup_{t\geq 0} X_t \geq \varepsilon) \leq \sup_{t\geq 0} E[X_t]$ for any $\varepsilon > 0$ and any nonnegative submartingale X; see Revuz and Yor, Theorem II.1.7. Note that $X := |M|$ is a submartingale if M is a martingale, and that $\sup_{t\geq 0} E[X_t] = E[|M_\infty|]$ if in addition M is uniformly integrable.

[2] That E is *Polish* means that it admits a metric for which it becomes a complete separable metric space.

Theorem 4.2. *Let $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, E, and X be as above. Then there exists a regular conditional distribution of X given $\mathcal{G}$.*

Rather than conditional probabilities (6), one is often interested in conditional expectations of the form $E[f(X) \mid \mathcal{G}]$. As expected, such conditional expectations are obtained by integrating against R.

Lemma 4.3. *Let $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, E, and X be as above, and let R be a regular conditional distribution of X given $\mathcal{G}$. Then, for any $\mathcal{G} \otimes \mathcal{B}(E)$-measurable map $H \colon \Omega \times E \to \mathbb{R}$ that is nonnegative or bounded, one has*

$$E[H(\cdot, X) \mid \mathcal{G}](\omega) = \int_E H(\omega, x) R(\omega, dx) \quad P\text{-a.e. } \omega.^3$$

In particular, if $f \colon E \to \mathbb{R}$ is measurable as well as nonnegative or bounded, then

$$E[f(X) \mid \mathcal{G}](\omega) = \int_E f(x) R(\omega, dx) \quad P\text{-a.e. } \omega.$$

Proof. The result holds for every H of the form $H(\omega, x) = \mathbf{1}_{A \times B}(\omega, x)$ with $A \in \mathcal{G}$ and $B \in \mathcal{B}(E)$, because

$$\begin{aligned} E[H(\cdot, X) \mid \mathcal{G}](\omega) &= \mathbf{1}_A(\omega) P(X \in B \mid \mathcal{G})(\omega) \\ &= \mathbf{1}_A(\omega) R(\omega, B) \\ &= \int_E H(\omega, x) R(\omega, dx), \quad P\text{-a.e. } \omega. \end{aligned}$$

The monotone class theorem now gives the result for all bounded measurable H. This relies on the fact that the maps $\mathbf{1}_{A \times B}$ as above generate $\mathcal{G} \otimes \mathcal{B}(E)$, and that the set of maps H for which the result holds is a real vector space closed under pointwise increasing limits (the latter uses both the conditional monotone convergence theorem and the unconditional monotone convergence theorem). □

[3] Here the left-hand side should be read as $E[Y \mid \mathcal{G}](\omega)$ with Y given by $Y(\omega) := H(\omega, X(\omega))$.