STOCHASTIC PROCESSES

Contents

Date: Updated April 23, 2018.

These are notes for a course on Stochastic Calculus, and are meant to supplement the texts Durrett, *Probability: Theory and Examples*; Karatzas-Sheve [KS], *Brownian Motion and Stochastic Calculus*; and Øksendahl, *Stochastic Differential Equations: An Introduction with Applications*. There's almost surely typos in the text, so please use at your own risk!

1. Preliminaries

1.1. **Stochastic process.** Throughout, the probability space $(\Omega, \mathscr{F}, P)$ will be the sample space, on which a collection of random variables, i.e., measurable functions $X = \{X_t; 0 \le t < \infty\}$ will be defined and called a stochastic process, taking values in the state space $(S, \mathscr{S})$. Most of the time, we will take $(S, \mathscr{S}) = (\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$, where for any topological space U the Borel sets $\mathscr{B}(U)$ will denote the σ-algebra generated by the open sets in U.

For a fixed sample point $\omega \in \Omega$, the function $t \mapsto X_t(\omega), t \ge 0$ is the sample path of the process X associated to ω. If X and Y are stochastic processes defined on $(\Omega, \mathscr{F}, P)$ and $(\Omega', \mathscr{F}', P')$ respectively, then we say they have the same finite-dimensional distributions if for any integer $n \ge 1$, real numbers $0 \le t_1 < \cdots < t_n < \infty$, and $A \in \mathscr{B}(\mathbf{R}^d)$, we have

$$P[(X_{t_1}, \ldots, X_{t_n}) \in A] = P'[(Y_{t_1}, \ldots, Y_{t_n}) \in A]. \tag{1.1.1}$$

We say X is measurable if the mapping $(t, \omega) \mapsto X_t(\omega) : ([0, \infty) \times \Omega, \mathscr{B}(\mathbf{R}^d) \otimes \mathscr{F}) \to (\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$ is measurable.

1.1.1. *Filtrations.* Next we equip our sample space with a filtration $\{\mathscr{F}_t; 0 \le t < \infty\}$, a non-decreasing family of σ-subalgebras $\mathscr{F}_s \subset \mathscr{F}_t \subset \mathscr{F}, 0 \le s < t < \infty$. Given a stochastic process X, the simplest filtration is that generated by X itself $\mathscr{F}_t^X = \sigma(X_s; 0 \le s < t)$, the smallest σ-algebra with respect to which X_s is measurable for every $s \in [0, t]$. On the other hand, we say a process X is adapted to a filtration $\{\mathscr{F}_t\}$ if X_t is $\mathscr{F}_t$-measurable for every $t \ge 0$.

Let $\mathscr{F}_t$ be a filtration. Define $\mathscr{F}_{t-} = \sigma(\cup_{s<t} \mathscr{F}_s)$ to be the σ-algebra of past events, and $\mathscr{F}_{t+} = \cap_{s>t} \mathscr{F}_s$ to be the σ-algebra of future events after $t \ge 0$. Define $\mathscr{F}_{0-} := \mathscr{F}_0$. We say the filtration is right-(resp. left-)continuous if $\mathscr{F}_t = \mathscr{F}_{t+}$ (resp. $\mathscr{F}_t = \mathscr{F}_{t-}$) for all $t \ge 0$. When $X = \{X_t, \mathscr{F}_t^X; 0 \le t < \infty\}$ is a process on $(\Omega, \mathscr{F})$, then left-continuity of $\mathscr{F}_t$ at some fixed $t > 0$ can be interpreted to mean that Xt can be discovered by observing $X_s, 0 \le s < t$. Right-continuity means intuitively that if X_s has been observed for $0 \le s < t$, then nothing more can be learned by peeking infinitesimally far into the future. Here $\mathscr{F}_t^X = \sigma(X_s; 0 \le s < t)$.

1.1.2. *Time.* A random time T is an $\mathscr{F}$-measurable random variable on $(\Omega, \mathscr{F})$ taking values in $[0, \infty]$. If X is a stochastic process, define the function X_T on the event $\{T < \infty\}$ by $X_T(\omega) := X_{T(\omega)}(\omega)$. If $X_\infty(\omega)$ is defined for all $\omega \in \Omega$, then X_T can be defined on Ω by setting $X_T(\omega) = X_\infty(\omega)$ on $\{T = \infty\}$. If X is measurable and T is finite, then X_T is a random variable.

Now let $\{\mathscr{F}_t\}$ be a filtration on $(\Omega, \mathscr{F})$. A random time T is called a stopping time of the filtration if $\{T \le t\} \in \mathscr{F}_t$ for all ≥ 0, and an optional time if $\{T < t\} \in \mathscr{F}_t$ for all ≥ 0. Every random time equal to a nonnegative constant is a stopping time. Every stopping time is optional, and the two concepts coincide if the filtration is right-continuous.

Let X be a process with right-continuous paths, adapted to $\{\mathscr{F}_t\}$. Let $A \in \mathscr{S}$ in state space, then define the hitting time $H_A(\omega) = \inf\{t \ge 0; X_t(\omega) \in A\}$, with the convention that the infimum of the empty set is infinity. If A is open, then H_A is an optional time; if A is closed and the sample paths of X are continuous, then H_A is a stopping time.

1.2. **Continuous-time martingales.** An example of a discrete-time martingale is a symmetric simple random walk. An example of a continuous-time martingale is a Brownian motion. Brownian motion also happens to be a continuous function of t, but this is not always true of continuous-time martingales. Consider the **R**-valued process $X = \{X_t; 0 \le t < \infty\}$ on a probability space $(\Omega, \mathscr{F}, P)$, adapted to a filtration $\{\mathscr{F}_t\}$ and such that $E|X_t| < \infty$ holds for all $t \ge 0$. Then we say that $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a submartingale (resp. supermartingale) if for every

$0 \leq s < t < \infty$ we have $E[X_t|\mathscr{F}_s] \geq X_s$ (resp. $E[X_t|\mathscr{F}_s] \leq X_s$). If it is both a super- and sub-martingale, then we call X a martingale.

We shall sometimes consider process X whose sample paths $X_t(\omega)$ are RCLL, i.e., right-continuous on $t \in [0,\infty)$ and with finite left-hand limits on $(0,\infty)$, or some other combination of Rs and Ls. Sometimes this is also called a *cádlág* process (French abbreviation).

1.2.1. *Continuous, square-integrable martingales.* Now, let $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ be a right-continuous martingale. We say that X is square-integrable if $E[X_t^2] < \infty$ for all $t \geq 0$. Let $\mathscr{M}_2$ be the collection of square-integrable, right-continuous martingales X with $X_0 = 0$ a.e., and $\mathscr{M}_2^c$ the subset of $X \in \mathscr{M}_2$ that are continuous. Define a metric by

$$||X|| = \sum_{n=1}^{\infty} 2^{-n}||X||_n \wedge 1, \qquad ||X||_n^2 = E[X_t^2], \tag{1.2.1}$$

under which $\mathscr{M}_2$ forms a complete metric space, and $\mathscr{M}_2^c$ a closed subspace.

Let $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ be a process with $X_0 = 0$ a.e. If there exists a non-decreasing sequence $\{T_n\}$ of stopping times of $\{\mathscr{F}_t\}$ such that $\{X_{t\wedge T_n}, \mathscr{F}_t; 0 \leq t < \infty\}$ is a martingale for each $n \geq 1$ and $P(\lim_{n\to\infty} T_n = \infty) = 1$, then we say that X is a *local martingale.*

For any $X \in \mathscr{M}_2$, X^2 is a nonnegative submartingale, and has a unique Doob-Meyer decomposition

$$X_t^2 = M_t + A_t, \quad 0 \leq t < \infty \tag{1.2.2}$$

where $M = \{M_t, \mathscr{F}_t, 0 \leq t < \infty\}$ is a right-continuous martingale and $A = \{A_t, \mathscr{F}_t, 0 \leq t < \infty\}$ is a natural increasing process, i.e., $t \mapsto A_t(\omega)$ is a non-decreasing, right-continuous function, $E[A_t] < \infty$ for all $0 \leq t < \infty$, and (natural) for every bounded right continuous martingale N we have

$$E[\int_{(0,t]} N_s dA_s] = E[\int_{(0,t]} M_s dA_s], \quad 0 < t < \infty. \tag{1.2.3}$$

We normalise these processes so that $M_0 = A_0 = 0$ P-a.s. If $X \in \mathscr{M}_2^c$, then M and A are also continuous.

For $X \in \mathscr{M}_2$, the *quadratic variation* of X can be defined as $\langle X\rangle_t = A_t$, and one checks that this coincides with the definition given later below. In other words, $\langle X\rangle$ is the unique adapted, natural increasing process for which $\langle X\rangle_0 = 0$ a.s. and $X^2 - \langle X\rangle$ is a martingale.

For any two martingales $X, Y \in \mathscr{M}_2$, define their *cross-variation process* $\langle X, Y\rangle$ by

$$\langle X, Y\rangle_t := \frac{1}{4}(\langle X+Y\rangle_t - \langle X-Y\rangle_t), \quad 0 \leq t < \infty. \tag{1.2.4}$$

And observe that $XY - \langle X, Y\rangle$ is a martingale. Two elements X, Y are called orthogonal if $\langle X, Y\rangle_t = 0$ a.s. for every $0 \leq t < \infty$. Also, $\langle \cdot, \cdot\rangle$ is a bilinear form on $\mathscr{M}_2$.

1.3. **The Markov property.** The Markov chain condition $P(X_{n+1} \in B|\mathcal{F}_n) = p(X_n, B)$ is clear, especially if you consider the discrete case

(1.3.1)
$$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j|X_n = i) := p(i,j).$$

The Markov property $E_\mu(Y \circ \theta_n|\mathcal{F}_n) = E_{X_n}Y$ tells us that shifting by θ_n, we can basically forget the conditioning $\mathcal{F}_n$ (information up to n) and start at X_n.

The strong Markov applies Markov to stopping times. Recall that N is a stopping time if $\{N = n\} \in \mathcal{F}_n$ for all n. Then we restrict ourselves to

$$\mathcal{F}_N = \{A \in \mathcal{F} : A \cap \{N = n\} \in \mathcal{F}_n \forall n\}. \tag{1.3.2}$$

We want to make sense of $E_\mu(Y_N \circ \theta_N | \mathcal{F}_N) = E_{X_N} Y_N$ on $\{N < \infty\}$, which heuristically means that shifting to the stopping time N, we can forget the conditioning up to $\mathcal{F}_N$ and start fresh at X_N. The way to see this is by fixing $A \in \mathcal{F}_N$, and computing

$$E_\mu(Y_N \circ \theta_N; A \cap \{N < \infty\}) = \sum_{n=0}^{\infty} E_\mu(Y_N \circ \theta_N; A \cap \{N = n\}) \tag{1.3.3}$$

The random variable N takes values at integers n. So at each level we understand $E_\mu(Y_N \circ \theta_N; A \cap \{N = n\}) = E_\mu(Y_n \circ \theta_n; A \cap \{N = n\})$. Then applying the Markov property we get

$$\sum_{n=0}^{\infty} E_\mu(E_{X_n} Y_n; A \cap \{N = n\}) = E_\mu(E_{X_N} Y_N; A \cap \{N < \infty\}) \tag{1.3.4}$$

recalling that for conditional expectation $E(E(Y|\mathcal{F})) = E(Y)$, and that $E(X|\mathcal{F}) = E(X; \Omega_i)/P(\Omega_i)$ on Ω_i, where $\mathcal{F} = \sigma(\Omega_1, \Omega_2 \dots)$ where the Ω_i are disjoint. In our case, take $\Omega_i = A \cap \{N < \infty\}$ (the other being $A \cap \{N = \infty\}$).

Example 1.1. In the random walk example, μ was the distribution of the iid random variables $\xi_1, \xi_2, \dots \in \mathbf{Z}^d$, meaning that $\mu(A) = P(\xi_n \in A)$ for any measurable set A, and P a probability measure on $\mathbf{Z}^d$. So if $X_n = i$, the probability that $X_{n+1} = j$, is captured by the transition probability

$$p(i,j) = P(\xi_{n+1} = i - j). \tag{1.3.5}$$

(For a nice example of a probability measure on $\mathbf{Z}$, look up the zeta distribution.)

2. Brownian motion

A *one-dimensional Brownian motion* is a continuous adapted process $B = \{B_t, \mathscr{F}_t; 0 \geq t\}$ defined on a probability space $(\Omega, \mathscr{F}, P)$ such that for $0 \leq s < t < \infty$, (a) the increment $B_t - B_s$ is independent of $\mathscr{F}_s$, and (b) $B_t - B_s$ is normally distributed with mean zero and variance $t - s$. If $B_0 = 0$, we call B a standard Brownian motion.

Proposition 2.1. *B is a square-integrable martingale.*

Proof. By Jensen's inequality, $E|B_t|^2 \leq E[B_t^2] = \text{var}(B_t) = t < \infty$, so it is square-integrable. Also, for $0 \leq s < t < \infty$, we have

$$E[B_t|\mathscr{F}_s] = E[B_t - B_s|\mathscr{F}_s] + E[B_s|\mathscr{F}_s] = E[B_t - B_s] + E[B_s] = B_s, \tag{2.0.1}$$

where we have used the independence of $B_t - B_s$ of $\mathscr{F}_s$ and $B_s \in \mathscr{F}_s$. □

2.1. The invariance principle.

Let $\{\xi_n\}$ be a sequence of i.i.d. random variables with mean 0 and variance $\sigma^2, 0 < \sigma < \infty$. Consider the partial sums $S_n = \sum_{i=1}^n \xi_i, S_0 = 0$. We can define a continuous-time process by linear interpolation:

$$Y_t = S_{[t]} + (t - [t])\xi_{[t]+1}, \quad t \geq 0, \tag{2.1.1}$$

where $[t]$ is the floor function. Scaling by time and space, we obtain a sequence of processes

$$X_t^{(n)} = \frac{1}{\sigma\sqrt{n}} Y_{nt}, \quad t \geq 0. \tag{2.1.2}$$

Observe that for $s = k/n$ and $t = (k+1)/n$, the increment $X(n)_t - X(n)_s = (1/\sigma\sqrt{n})\xi_{k+1}$ is independent of $\sigma(\xi_1, \dots, \xi_k)$, has zero mean and variance $t - s$. This suggests that $\{X^{(n)}\}$ is approximately Brownian motion. Even though the random variables ξ_i are not necessarily normal, the central limit theorem implies that the limiting distributions of the increments are.

Theorem 2.2. *Let $\{X^{(n)}\}$ be defined as above. Then for $0 \le t_1 < \cdots < t_d < \infty$ we have the convergence in distributions*

$$(X^{(n)}_{t_1}, \ldots, X^{(n)}_{t_d}) \xrightarrow{D} (B_{t_1}, \ldots, B_{t_d}) \tag{2.1.3}$$

as $t \to \infty$, and $\{B_t, \mathscr{F}_t; t \ge 0\}$ is standard one-dimensional Brownian motion.

Proof. See [KS, 4.17] for the full proof. Here we'll prove only convergence in probability, and with $d = 2$. That is, for $s = t_1, t = t_2$, we want to show that $(X^{(n)}_s, X^{(n)}_t) \to (B_s, B_t)$ in probability. Now, since

$$\left| X^{(n)}_t - \frac{1}{\sigma\sqrt{n}} S_{[tn]} \right| \le \frac{1}{\sigma\sqrt{n}} |\xi_{[tn]+1}|, \tag{2.1.4}$$

we have by the Čebyšev inequality,

$$P\left[\left| X^{(n)}_t - \frac{1}{\sigma\sqrt{n}} S_{[tn]} \right| > \epsilon \right] \le \frac{1}{\epsilon^2 n} \to 0 \tag{2.1.5}$$

as $n \to \infty$. It is clear then that

$$\left\| (X^{(n)}_s, X^{(n)}_t) - \frac{1}{\sigma\sqrt{n}} (S_{[sn]}, S_{[tn]}) \right\| \to 0 \tag{2.1.6}$$

in probability.

We claim that if in addition

$$\frac{1}{\sigma\sqrt{n}} (S_{[sn]}, S_{[tn]}) \xrightarrow{D} (B_s, B_t) \tag{2.1.7}$$

then $(X^{(n)}_s, X^{(n)}_t) \xrightarrow{D} (B_s, B_t)$ by [KS, 4.16]. Since $B_t - B_s \mapsto B_t$ is a continuous function, this is equivalent to proving that [KS, 4.5]

$$\frac{1}{\sigma\sqrt{n}} (S_{[sn]}, S_{[tn]} - S_{[sn]}) \xrightarrow{D} (B_s, B_t - B_s). \tag{2.1.8}$$

By the independence of the random variables $\{\xi_n\}$, we have

$$\begin{aligned} &\lim_{n\to\infty} E[\exp\Big(\frac{iu}{\sigma\sqrt{n}} \sum_{i=1}^{[sn]} \xi_i + \frac{iv}{\sigma\sqrt{n}} \sum_{i=[sn]+1}^{[tn]} \xi_i\Big)] \\ &= \lim_{n\to\infty} E[\exp\Big(\frac{iu}{\sigma\sqrt{n}} \sum_{i=1}^{[sn]} \xi_i\Big)] \cdot \lim_{n\to\infty} E[\Big(\frac{iv}{\sigma\sqrt{n}} \sum_{i=[sn]+1}^{[tn]} \xi_i\Big)] \end{aligned}$$

provided the limits exist. Since

$$\left| \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{[sn]} \xi_i - \frac{\sqrt{s}}{\sigma\sqrt{[sn]}} \sum_{i=1}^{[sn]} \xi_i \right| \to 0 \tag{2.1.9}$$

in probability, and by the central limit theorem $\sqrt{(s/\sigma^2[sn]} \sum_{i=1}^{[sn]} \xi_i$ converges in distribution to a normal random variable with mean 0 and variance s, we have

$$\lim_{n\to\infty} E[\exp\Big(\frac{iu}{\sigma\sqrt{n}} \sum_{i=1}^{[sn]} \xi_i\Big)] = e^{-u^2 s/2} \tag{2.1.10}$$

and similarly,

$$\lim_{n\to\infty} E[\Big(\frac{iv}{\sigma\sqrt{n}} \sum_{i=[sn]+1}^{[tn]} \xi_i\Big)] = e^{-v^2 t - s/2} \tag{2.1.11}$$

completing the proof of convergence in probability. To extend to convergence in distribution use [KS, Lemma 4.18, 4.19] to prove tightness of the sequence. □

Theorem 2.3 (The invariance principle of Donsker). *Let $\{\xi_i\}$ be a sequence of i.i.d. random variables with zero mean and finite variance $\sigma^2 > 0$, defined on a probability space $(\Omega, \mathscr{F}, P)$. Define $X^{(n)} = \{X_t^{(n)}; t \geq 0\}$ as above, and let P_n be the measure induced by $X^{(n)}$ on $(C[0,\infty), \mathscr{B}(C[0,\infty)))$. Then $\{P_n\}$ converges weakly to a measure P, under which the coordinate mapping process $B_t(\omega) := \omega(t)$ on $C[0,\infty)$ is a standard one-dimensional Brownian motion.*

The resulting probability measure P is called the Wiener measure.

2.2. **The Markov property.** We first need to define d-dimensional Brownian motion and Brownian families.

Definition 2.4. Let d be a positive integer and μ a probability measure on $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$. Let $B = \{B_t, \mathscr{F}_t; t \geq 0\}$ be a continuous, adapted process with values in $\mathbf{R}^d$, defined on a probability space $(\Omega, \mathscr{F}, P)$. We call it a **d-dimensional Brownian motion** with distribution μ if (i) $P(B_0 \in A) = \mu(A)$ for all $A \in \mathscr{B}(\mathbf{R}^d)$, and (ii) for $0 \leq s < t < \infty$, the increment $B_t - B_s$ is independent of $\mathscr{F}_s$ and is normally distributed with mean zero and covariance matrix equal to $(t-s)I_d$ where I_d is the $(d \times d)$-identity matrix. If μ assigns measure 1 to some singleton $\{x\}$, we say that B is a d-dimensional Brownian motion starting at x.

We have given a construction of this using Kolmogorov's extension theorem (see Durrett). Here is a second way: Let $P^0 = P^{(1)} \times \cdots \times P^{(n)}$ be d copies of Wiener measure. Under P^0, the coordinate mapping process $B_t(\omega) := \omega(t)$ together with the filtration $\mathscr{F}_t$ generated by B_t is a d-dimensional Brownian motion starting at the origin. Given $x \in \mathbf{R}^d$, we can also define the probability measure on $(C[0,\infty)^d, \mathscr{B}(C[0,\infty)^d))$ by

$$P^x(A) = P^0(A - x), \quad A \in \mathscr{B}(C[0,\infty)^d) \tag{2.2.1}$$

where $F - x = \{\omega \in C[0,\infty)^d : \omega + x \in A\}$, giving Brownian motion starting at x. Finally, for a probability measure μ on $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$, we define P^μ on $\mathscr{B}(C[0,\infty)^d)$ by

$$P^\mu(A) = \int_{\mathbf{R}^d} P^x(A)\mu(dx). \tag{2.2.2}$$

Now, given a metric space (S,d), we denote by $\overline{\mathscr{B}(S)^\mu}$ the competion of the Borel σ-algebra $\mathscr{B}(S)$ with respect to the finite measure μ on $(S, \mathscr{B}(S))$. We define the universal σ-algebra $\mathscr{U} = \cap_\mu \overline{\mathscr{B}(S)^\mu}$, where the intersection is over all finite/probability measures μ. A $\mathscr{U}(S)/\mathscr{B}(|R)$-measurable, real-valued function is called universally measurable.

Definition 2.5. A **d-dimensional Brownian family** is an adapted, d-dimensional process $B = \{B_t, \mathscr{F}_t; t \geq 0\}$ on a measurable space $(\Omega, \mathscr{F})$ and a family of probability measures $\{P^x\}$ such that

(1) for each $A \in \mathscr{F}$, the mapping $x \mapsto P^x(A)$ is universally measurable,
(2) for each $x \in \mathbf{R}^d$, we have $P^x(B_0 = x) = 1$,
(3) under each P^x, the process B is a d-dimensional Brownian motion starting at x.

In fact, the construction above shows that $x \mapsto P^x(A)$ is Borel-measurable for each $A \in \mathscr{F}$, which implies (i).

Exercise 1. Let X and Y be d-dimensional random vectors on $(\Omega, \mathscr{F}, P)$. If $\mathscr{G} \subset \mathscr{F}$ is a sub-σ-algebra, X is indepdent of $\mathscr{G}$ adn Y is $\mathscr{G}$-measurable, then for every $A \in \mathscr{B}(\mathbf{R}^d)$ we have

(1) $P[X + Y \in A|\mathscr{G}] = P[X + Y \in A|Y]$, P-a.e.
(2) $P[X + Y \in A|Y = y] = P[X + y \in A]$ for PY^{-1}-a.e. $y \in \mathbf{R}^d$.

where $PY^{-1}(B) = P(\omega \in \Omega : X(\omega) \in B)$ for any $B \in \mathscr{B}(\mathbf{R}^d.)$

From this it follows that

$$P^\mu[B_t \in A|\mathscr{F}_s] = P^\mu[B_t \in A|B_s], \quad 0 \le s < t, A \in \mathscr{B}(R^d). \tag{2.2.3}$$

That is, information for B_t up to time s is the same as the information of B_s. Secondly,

$$P^\mu[B_t \in A|B_s = y] = P^y[B_{t-s} \in A]], \quad 0 \le s < t, A \in \mathscr{B}(\mathbf{R}^d). \tag{2.2.4}$$

That is, $B_t = (B_t - B_s) + B_s$ is distributed the same as B_{t-s} under $P^y, B_s = y$.

Definition 2.6. Let μ be a probability measure on $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$. An adapted, d-dimensional process $X = \{X_t, \mathscr{F}_t; t \ge 0\}$ on some probability space $(\Omega, \mathscr{F}, P)$ is called a **Markov process** with initial distribution μ if (i) $P^\mu(X_0 \in \mathcal{A}) = \mu(A)$ for all $A \in \mathscr{B}(\mathbf{R}^d)$, and (ii) for $s, t \ge 0$ and $A \in \mathscr{B}(\mathbf{R}^d)$ we have

$$P^\mu[X_{t+s} \in A|\mathscr{F}_s] = P^\mu[X_{t+s} \in X_s], \quad P^\mu\text{-a.s.} \tag{2.2.5}$$

If $\{P^x\}$ is a family of probability measures on $(\Omega, \mathscr{F})$, then X is a **Markov family** if

(1) for each $A \in \mathscr{F}$, the mapping $x \mapsto P^x(A)$ is universally measurable,
(2) for each $x \in \mathbf{R}^d$, we have $P^x(X_0 = x) = 1$,
(3) for each $x \in \mathbf{R}^d, s, t \ge 0$ and $A \in \mathscr{B}(\mathbf{R}^d)$,

$$P^x[X_{t+s} \in A|\mathscr{F}_s] = P^x[X_{t+s} \in X_s], \quad P^x\text{-a.s.} \tag{2.2.6}$$

(4) for each $x \in \mathbf{R}^d, s, t \ge 0$ and $A \in \mathscr{B}(\mathbf{R}^d)$,

$$P^x[X_{t+s} \in A|X_s = y] = P^y[X_t \in A], \quad P^x X_s^{-1}\text{-a.s. } y \tag{2.2.7}$$

It the follows form that a d-dimensional Brownian motion (resp. family) is a Markov process (resp. family).

2.3. **Brownian sample paths.** An $\mathbf{R}^d$-valued stochastic process $X = \{X_t; 0 \le t < \infty\}$ is called Gaussian if for any integer $k \ge 1$ and real numbers $0 \le t_1 < t_2 < \cdots < t_k < \infty$, the random vector $(X_{t_1}, \ldots, X_{t_k})$ has a joint normal distribution. If the distribution $(X_{t+t_1}, \ldots, X_{t+t_k})$ does not depend on t, then we say that the process is **stationary**.

The finite-dimensional distributions of a Gaussian process X are determined by its expectation vector $\mu(t) := EX_t, t \ge 0$ and its covariance matrix

$$\rho(s,t) := E[(X_s - \mu(s))(X_t - \mu(t))^T], \quad s, t \ge 0. \tag{2.3.1}$$

If $\mu(t) = 0$ for all $t \ge 0$, we say that X is a zero-mean Gaussian process.

A one-dimensional Brownian motion is a zero-mean Gaussian process with covariance $\rho(s,t) = s \wedge t$. Conversely, any zero mean Gaussian process with a.s. continuous paths and covariance function $s \wedge t$ is a one-dimensional Brownian motion.

Remark 2.7 (Equivalence of definitions for Brownian motion). (Refer to [D].) By translation invariance we'll set $B_0 = 0$. By (a) and (b) and the definitions that (a') B_t is a Gaussian process. To get (b'), let $s < t$, check that

$$EB_sB_t = E[B_s^2 + B_s(B_t - B_s)] = E[B_s^2] = s. \tag{2.3.2}$$

To go the other direction, notice that (a') and (b') determines the finite-dimensional distributions of of B_t, and from the above equation you can see that they agree with the ones defined in (a) and (b).

Exercise 2. Let $W = \{W_t, \mathscr{F}_t; 0 \le t < \infty\}$ be a standard Brownian motion. Show that the following processes obtained from 'equivalence transformations' are also standard Brownian motion:

(1) Scaling: $c > 0$

$$X_t = \frac{1}{\sqrt{c}} W_{ct}, \quad 0 \le t < \infty \tag{2.3.3}$$

(2) Time-inversion:

$$Y_t = \begin{cases} tW_{1/t}, & 0 < t < \infty \\ 0, & t = 0 \end{cases} \tag{2.3.4}$$

(3) Time-reversal: for fixed $T > 0$, $Z_t = W_T - W_{T-t}, \quad 0 \le t \le T$

(4) Symmetry: $-W_t$.

2.3.1. *Nowhere differentiability.* Let $f : [0,\infty) \to \mathbf{R}$ be a continuous function. Define the upper and lower (right and left) Dini derivatives at t by

$$D^{\pm} f(t) = \limsup_{h\to 0^{\pm}} \frac{f(t+h) - f(t)}{h}, \quad D^{\pm} f(t) = \liminf_{h\to 0^{\pm}} \frac{f(t+h) - f(t)}{h}. \tag{2.3.5}$$

We say f is differentiable at t from the right (resp. left) if $D^+ f(t) = D_+ f(t)$ exist (resp. $D^- f(t) = D_- f(t)$), and differentiable at $t > 0$ if all Dini derivatives at t exists and are equal. At $t = 0$, differentiability is defined to be differentiability from the right.

Theorem 2.8. *The Brownian sample path $W_t(\omega)$ is nowhere differentiable for a.e. ω. More precisely, the set*

$$\{\omega \in \Omega : \forall t \in [0,\infty), \text{ either } D^+ W_t(\omega) = \infty \text{ or } D_+ W_t(\omega) = -\infty\} \tag{2.3.6}$$

contains an event $A \in \mathscr{F}$ with $P(A) = 1$.

Proof. By scaling, it is enough to consider $t \in [0,1]$. For fixed integers $j, k \ge 1$, consider the set

$$A_{jk} = \bigcup_{t\in[0,1]} \bigcap_{h\in[0,1/k]} \{\omega \in \Omega : |W_{t+h}(\omega) - W_t(\omega)| \le jh\}. \tag{2.3.7}$$

Observe that

$$\bigcup_{j,k=1}^{\infty} A_{jk} = \{\omega \in \Omega : -\infty < D_+ W_t(\omega) \le D^+ W_t(\omega) < \infty \text{ for some } t \in [0,1]\} \tag{2.3.8}$$

Since for any ω in this set, there exists some $t \in [0,1]$ and $j \ge 1$ such that as $h \to 0$ we have

$$\left| \frac{W_{t+h}(\omega) - W_t(\omega)}{h} \right| \le j. \tag{2.3.9}$$

So if for every j, k we can find an event $C \in \mathscr{F}$ such that $P(C) = 0$ and $A_{jk} \subset C$, the complement of C will will prove the theorem.

To that end, fix a sample path $\omega \in A_{jk}$. So there exists a $t \in [0,1]$ such that $|W_{t+h}(\omega) - W_t(\omega)| \le jh$ for every $0 \le h \le 1/k$. Take an integer $n \ge 4k$. Then there exists an integer $1 \le i \le n$, such that $(i-1)/n \le t \le i/n$, and it follows that

$$|W_{(i+1)/n}(\omega) - W_{i/n}(\omega)| \le |W_{(i+1)/n}(\omega) - W_t(\omega)| + |W_{i/n}(\omega) - W_t(\omega)| \le \frac{2j}{n} + \frac{j}{n} = \frac{3j}{n}. \tag{2.3.10}$$

The last inequality follows from the fact that for $\nu = 1, 2, 3$ we have $(i+\nu)/n - t \le (\nu+1)/n \le 1/k$.

Now observe that from this last fact, $\omega \in A_{jk}$ gives information about the size of the Brownian increment not only over the interval $[i/n, (i+1)/n]$, but over the neighboring intervals $[(i+1)/n, (i+2)/n]$ and $[(i+2)/n, (i+3)/n]$. Indeed, by the same argument we have

$$|W_{(i+2)/n}(\omega) - W_{(i+1)/n}(\omega)| \le \frac{3j}{n} + \frac{2j}{n} = \frac{5j}{n}, \tag{2.3.11}$$

$$|W_{(i+3)/n}(\omega) - W_{(i+2)/n}(\omega)| \le \frac{4j}{n} + \frac{3j}{n} = \frac{7j}{n}. \tag{2.3.12}$$

So if we define

$$C_i^{(n)} := \bigcap_{\nu=1,2,3} \Big\{\omega \in \Omega : |W_{(i+\nu)/n}(\omega) - W_{(i+\nu-1)/n}(\omega)| \le \frac{2\nu+1}{n} j\Big\}, \tag{2.3.13}$$

we see that $A_{jk} \subset \cup_{i=1}^n C_i^{(n)}$ for each $n \ge 4k$.

On the other hand, after Brownian scaling

$$Z_\nu := \sqrt{n}(W_{(i+\nu)/n}(\omega) - W_{(i+\nu-1)/n}(\omega)), \quad \nu = 1, 2, 3 \tag{2.3.14}$$

are independent standard normal variables, and one checks that $P(|Z_\nu| \le \epsilon) \le \epsilon$ for any $\epsilon > 0$. So then for $i = 1, 2, \dots, n$ we have

$$P(C_i^{(n)}) \le \Big(\frac{3j}{n^{1/2}}\Big)\Big(\frac{5j}{n^{1/2}}\Big)\Big(\frac{7j}{n^{1/2}}\Big) = \frac{105 j^3}{n^{3/2}}, \tag{2.3.15}$$

and

$$A_{jk} \subset C := \bigcap_{n=4k}^{\infty} \bigcup_{i=1}^{n} C_i^{(n)} \in \mathscr{F} \tag{2.3.16}$$

and $P(C) \le \inf_{n \le 4k} P(\cup_{i=1}^n C_i^{(n)}) = 0$. □

Exercise 3. Modifying the above proof, show that with probability one, the Brownian sample path is not Hölder continuous of exponent $\gamma > 1/2$, hence nowhere differentiable. On the other hand, using the Kolmogorov-Čentsov lemma it is easy to show that Brownian motion is Hölder continuous with exponent $\gamma < 1/2$.

2.4. **On the construction of Brownian motion.** (Refer to [D], [KS] for the full proof.) The initial construction of the measure ν_x satisfies (a) and (b) of Brownian motion but not continuity (c). The idea for fixing this is to construct a similar function, also denoted ν_x such that

$$\nu_x(\omega : \omega(0) = x\} = 1 \tag{2.4.1}$$

and

$$\nu_x(\{\omega : \omega(t_i) \in A_i, i = 1, \dots, n) = \mu_{x,t_1,\dots,t_n}(A_1 \times \cdots \times A_n) \tag{2.4.2}$$

with $\mu_{x,t_1,\dots,t_n}$ defined in class. From the construction of $\mu_{x,t_1,\dots,t_n}$ we see that properties (a) and (b) again hold, and it remains to check for (c). To do this, we have to first show **Claim 1**:

$$\nu_x(\omega : \mathbf{Q}_2 \to \mathbf{R}, \text{ uniformly continuous on } \mathbf{Q}_2 \cap [0,T]) = 1 \tag{2.4.3}$$

Any uniformly continuous function ω on $\mathbf{Q}_2 \cap [0,T]$ has a unique continuous extension to $[0,T]$ by

$$\omega(t) := \lim_{s \in \mathbf{Q}_2 \to t} \omega(s) \tag{2.4.4}$$

for any $t \in [0,T]$. Let ψ be this function, so

$$\psi : \Omega_2 \to C \tag{2.4.5}$$

where C is the set of continuous functions $\omega : [0,\infty) \to \mathbf{R}$. It is a measurable function, so we can define a probability measure on $(C, \mathscr{C})$ where $\mathscr{C}$ is the σ-algebra generated by the sets $\{\omega : \omega(t) \in At, \ge 0\}$ for any Borel set $A \subset \mathbf{R}$ (the smallest σ-algebra that makes each ω measurable).

$$P_x := \nu_x \circ \psi^{-1} \tag{2.4.6}$$

which will gives us (c).

As usual, take $B_0(\omega) = 0$ and $T = 1$. Then to prove uniform continuity, we want to show **Claim 2** (Kolmogorov-Čentsov), which gives

$$|B_q - B_r| \le C_\omega |q - r|^\gamma \tag{2.4.7}$$

for all $q, r \in \mathbf{Q}_2 \cap [0,1]$ and $\gamma < \alpha/\beta$, where $\alpha, \beta > 0$. The proof looks at the dyadic subdivisions G_n, for which Chebyshev's inequality gives

$$P(G_n^c) \ll 2^{-n\lambda} \tag{2.4.8}$$

where $\lambda = \alpha - \beta\gamma > 0$.

For the bound on G_n, we start with **Claim 3** on the set H_N for a fixed $N > 0$:

$$|B_q - B_r| \leq \frac{3}{1 - 2^{-\gamma}} |q - r|^\gamma \tag{2.4.9}$$

for $q, r \in \mathbf{Q}_2 \cap [0,1]$ with $|q - r| < 2^{-N}$. This was proved in class. To apply this, we use the Borel-Cantelli lemma on the fact that $\sum P(G_N^c) < \infty$, to get that for any fixed $A > 0$, there exists δ_ω such that

$$|B_q - B_r| \leq A|q - r|^\gamma \tag{2.4.10}$$

for any $q, r \in \mathbf{Q}_2$ with $|q - r| < \delta_\omega$.

Now to prove Claim 2, we extend the latter to all q, r in $\mathbf{Q}_2 \cap [0,1]$ by subdividing $[q, r]$ into intervals of length less than δ_ω and applying the triangle inequality, and we are done.

3. Stochastic calculus

3.1. **Quadratic variation.** We have already seen Brownian motion is not differentiable in the usual sense. This discussion is to explain why stochastic integration also cannot be defined in the old way, the first variation being unbounded, and how to fix it using bounded quadratic variation.

Definition 3.1. Here is the definition for a p-variation, $p > 0$, with $p = 2$ being the quadratic variation. Let $X = \{X_t; 0 \leq t < \infty\}$ be a process. Fix $t > 0$, let $\Pi = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[0, t]$ with $0 = t_0 \leq t_1 \leq \cdots \leq t_n = t$. Define the p-th variation of X over the partition Π to be

$$V_t^{(p)}(\Pi) = \sum_{k=1}^{n} |X_{t_k} - X_{t_{k-1}}|^p. \tag{3.1.1}$$

Now define the mesh of the partition Π to be $||\Pi|| = \max_{1 \leq k \leq n} |t_k - t_{k-1}|$. If $V_t^{(2)}(\Pi)$ converges (in some sense) as $||\Pi|| \to 0$, then the limit can be called the quadratic variation $\langle X \rangle_t$ of X on $[0, t]$.

Example 3.2. Let $X = \{X_t, \mathscr{F}_t, 0 \leq t < \infty\}$ be a continuous martingale such that $X_0 = 0$ a.s., and $EX_t^2 < \infty$ for every $t \geq 0$ (i.e., X is square-integrable). Then $X^2 = \{X_t^2, \mathscr{F}_t; 0 \leq t < \infty\}$ is a nonnegative submartingale, and hence has a unique Doob-Meyer decomposition

$$X_t^2 = M_t + A_t, \qquad 0 \leq t < \infty \tag{3.1.2}$$

where $M = \{M_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a continuous martingale and $A = \{A_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is an increasing process. Then for partitions Π of $[0, t]$, we have

$$\lim_{||\Pi|| \to 0} V_t^{(2)}(\Pi) = A_t \tag{3.1.3}$$

in probability, and we may *define* $\langle X \rangle_t = A_t$ in this case. In other words, for every $\epsilon, \eta > 0$, there exists a $\delta > 0$ such that $||\Pi|| < \delta$ implies

$$P(|V_t^{(2)}(\Pi) - \langle X \rangle_t| > \epsilon) < \eta. \tag{3.1.4}$$

Exercise 4. Let $X = \{X_t, \mathscr{F}_t; 0 \leq t < \infty\}$ be a continuous process such that for every fixed $t > 0$ and some $p > 0$

$$\lim_{||\Pi|| \to 0} V_t^{(p)}(\Pi) = L_t \quad \text{(in probability)} \tag{3.1.5}$$

where L_t is some random variable taking values in $[0, \infty)$ a.s. Show that for all $q > p$ $\lim_{||\Pi|| \to 0} V_t^{(q)}(\Pi) = 0$ in probability, and for $0 < q < p$, $\lim_{||\Pi|| \to 0} V_t^{(q)}(\Pi) = \infty$ in probability on the event $\{L_t > 0\}$.

The conclusion is that the unbounded first variation of continuous square-integrable martingales M means they cannot be differentiable, and it is impossible to define integrals of the form $\int_0^t X_s(\omega)\, dM_s(\omega)$ for (almost) every $\omega \in \Omega$ in the Riemann-Stieltjes sense.

3.2. **Construction of the stochastic integral.** Let $M = \{M_t, \mathscr{F}_t; 0 \le t \le \infty\}$ be a square integrable martingale on $(\Omega, \mathscr{F}, P)$, such that the filtration $\{\mathscr{F}_t\}$ satisfies $\mathscr{F}_t = \cap_{\epsilon > 0} \mathscr{F}_{t+\epsilon}$ for all t, and that $\mathscr{F}_0$ contains all sets with P-measure zero. Define a measure on $([0, \infty) \times \Omega, \mathscr{B}([0, \infty)) \otimes \mathscr{F})$ by

$$\mu_M(A) = E[\int_0^\infty 1_A(t, \omega) d\langle M \rangle_t(\omega)]. \tag{3.2.1}$$

Call two $\mathscr{F}_t$-adapted processes $X = \{X_t, \mathscr{F}_t; 0 \le t \le \infty\}, Y = \{Y_t, \mathscr{F}_t; 0 \le t \le \infty\}$ equivalent if $X_t(\omega) = Y_t(\omega)$ μ_M-a.e. (t, ω), meaning almost everywhere on $[0, \infty) \times \Omega$ with respect to the measure μ_M. This gives an equivalence relation.

3.2.1. *Spaces of processes.* Define the L^2-norm for X as a function of (t, ω) restricted to $[0, T] \times \Omega$ under the measure μ_M,

$$[X]_T^2 := E[\int_0^T X_t^2 d\langle M \rangle_t], \tag{3.2.2}$$

when it exists. We have $[X - Y]_T = 0$ for all $T > 0$ if and only if X and Y are equivalent. Now let $\mathscr{L}$ be the set of equivalence classes of all measurable, $\{\mathscr{F}_t\}$-adapted processes X such that $[X]_T < \infty$ for all $T > 0$. Define a metric on $\mathscr{L}$ by

$$[X] := \sum_{n=1}^\infty 2^{-n} \min(1, [X]_n). \tag{3.2.3}$$

If the function $t \mapsto \langle M \rangle_t(\omega)$ is absolutely continuous for P-a.e. ω, we would be able to construct the integral $\int_0^T X_t dM_t$ for all $X \in \mathscr{L}$ and $T \ge 0$. But without this condition, we have to restrict ourselves to a smaller subspace.

Let $\mathscr{L}^{**}$ be the subspace of processes $X \in \mathscr{L}$ that are progressively measurable with respect to the filtration $\{\mathscr{F}_t\}$, i.e., the mapping $(t, \omega) \mapsto X_t(\omega)$ with $(t, \omega) \in ([0, t] \times \Omega, \mathscr{B}([0, t]) \otimes \mathscr{F})$ is measurable for all $t \ge 0$.

Let $\mathscr{L}^*$ be the subspace of processes $X \in \mathscr{L}$ that are predictable with respect to the filtration $\{\mathscr{F}_t\}$, i.e., the mapping $(t, \omega) \mapsto X_t(\omega)$ is measurable with respect to the predictable σ-algebra, which is the σ-algebra on $[0, \infty) \times \Omega$ generated by the sets $\{0\} \times A$ with $A \in \mathscr{F}_0$ and $\{(s, t]\} \times A'$ with $A' \in \mathscr{F}_s, s < t$.

Let $\mathscr{L}^0$ be the class of all simple processes. A process $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty)$ is called simple if there exists a strictly increasing sequence of real numbers $\{t_n\}$ with $t_0 = 0$ and $\lim_{n \to \infty} t_n = \infty$, and a sequence of random variables $\{\xi_n\}$ with $\sup_{n \ge 0} |\xi(\omega)| \le C < \infty$ for every ω, such that ξ_n is $\mathscr{F}_{t_n}$-measurable for every $n \ge 0$ and

$$X_t(\omega) = \xi_0(\omega) 1_{\{0\}}(t) + \sum_{i=0}^\infty \xi_i(\omega) 1_{(t_i, t_{i+1}]}(t) \tag{3.2.4}$$

for $0 \le t < \infty, \omega \in \Omega$. We have the chain of inclusions $\mathscr{L}^0 \subset \mathscr{L}^* \subset \mathscr{L}^{**} \subset \mathscr{L}$.

Definition 3.3. Let $X \in \mathscr{L}^0$. The stochastic integral is defined to be the martingale transform with respect to M,

$$I_t(X) := \sum_{i=0}^{n-1} \xi_i (M_{t_{i+1}} - M_{t_i}) + \xi_n (M_t - M_{t_n}) = \sum_{i=0}^\infty \xi_i (M_{t \wedge t_{i+1}} - M_{t \wedge t_i}), \tag{3.2.5}$$

for $0 \le t < \infty$. Here n is the unique integer for which $t_n \le t < t_{n+1}$. The definition shall be extended to integrands $X \in \mathscr{L}$ by successive approximations using simple processes.

Proposition 3.4 (KS 3.2.6)**.** *If $t \mapsto \langle M\rangle_t(\omega)$ is absolutely continuous with respect to Lebesgue measure for P-a.e. $\omega \in \Omega$, then $\mathscr{L}^0$ is dense in $\mathscr{L}$ with respect to the metric $[\cdot]$ defined above.*

Proposition 3.5 (KS 3.2.8)**.** *$\mathscr{L}^0$ is dense in $\mathscr{L}^{**}$ with respect to the metric $[\cdot]$.*

The following properties of simple processes and their integrals can be found in [KS] pp.137–138. Refer there for the proof.

Lemma 3.6. *Let $X, Y \in L^0$, and $0 \le s < t < \infty$. Then*

1. $I_0(X) = 0$ *a.s.* P
2. $I(aX + bY) = aI(X) + bI(Y))$, $a, b \in \mathbf{R}$
3. $E[I_t(X)|\mathscr{F}_s] = I_s(X)$ *a.s.* P
4. $E[I_t(X)]^2 = E[\int_0^t X_u^2 d\langle M\rangle_u]$
5. $||I(X)|| = [X]$, *where* $||X|| := \sum_{n=1}^{\infty} 2^{-n}\min(1, ||X||_n)$ *and* $||X||_n^2 := E[X_n^2]$.
6. $E[(I_t(X) - I_s(X))^2|\mathscr{F}_s] = E[\int_s^t X_u^2 d\langle M\rangle_u|\mathscr{F}_s]$ *a.s.* P

Proof. First, note that (1) and (2) are clear. (6) implies (4) and (5) by setting $s = 0$. (3) implies that $I(X) = \{I_t(X), \mathscr{F}_t; 0 \le t < \infty\}$ is a continuous martingale. (6) implies that it is square-integrable.

Now to prove (3), observe that for any $i \ge 1$,

$$E[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathscr{F}_s] = \xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i}), \quad P\text{-a.s.}. \tag{3.2.6}$$

we have to check the three cases $s \le t_i, t_i < s \le t_{i+1}, t_{i+1} < s$. In the first case, we condition twice using the fact that $\mathscr{F}_{t_i} \supset \mathscr{F}_s$,

$$E[E[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge t_i})|\mathscr{F}_{t_i}]|\mathscr{F}_s] = E[\xi_i E[M_{t\wedge t_{i+1}} - M_{t\wedge t_i}|\mathscr{F}_{t_i}]|\mathscr{F}_s] = 0 \tag{3.2.7}$$

by the stopping time. In the second case, use the same idea on the first summand below

$$\begin{aligned}&E[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge s}) - \xi_i(M_{t\wedge s} - M_{t\wedge t_i})|\mathscr{F}_s]\\ =&E[E[\xi_i(M_{t\wedge t_{i+1}} - M_{t\wedge s})|\mathscr{F}_{t_{i+1}}]|\mathscr{F}_s] + \xi_i E[M_{t\wedge s} - M_{t\wedge t_i}|\mathscr{F}_s] = \xi_i(M_s - M_{s\wedge t_i})\end{aligned}$$

and the third case is straightforward since $\mathscr{F}_{t_i} \subset \mathscr{F}_s$.

For (6), choose m, n such that $t_{m-1} \le s < t_m$ and $t_n \le t < t_{n+1}$, and

$$\begin{aligned}&E[(I_t(X) - I_s(X))^2|\mathscr{F}_s]\\ &= E[\{\xi_{m-1}(M_{t_m} - M_s) + \sum_{i=m}^{n-1}\xi_i(M_{t_{i+1}} - M_{t_i}) + \xi_n(M_t - M_{t_n})\}^2|\mathscr{F}_s]\\ &= E[\xi_{m-1}^2(M_{t_m} - M_s)^2 + \sum_{i=m}^{n-1}\xi_i^2(M_{t_{i+1}} - M_{t_i})^2 + \xi_n^2(M_t - M_{t_n})^2|\mathscr{F}_s]\\ &= E[\xi_{m-1}^2(\langle M\rangle_{t_m} - \langle M\rangle_s) + \sum_{i=m}^{n-1}\xi_i^2(\langle M\rangle_{t_{i+1}} - \langle M\rangle_{t_i}) + \xi_n^2(\langle M\rangle_t - \langle M\rangle_{t_n})|\mathscr{F}_s]\\ &= E[\int_s^t X_u^2 d\langle M\rangle_u|\mathscr{F}_s]\end{aligned}$$

where in the second equality we use the vanishing of cross-terms, below in Example 3.9., and for the third equality, $E[M_t^2] = E[\langle M\rangle_t]$ for any square integrable martingale M with $M_t = 0$. □

The square-integrable continuous martingales forms a complete metric space $\mathscr{M}_2^c$ under the metric $||\cdot||$ above. Given $X \in \mathscr{L}^*$, by the density of $\mathscr{L}^0$ in $\mathscr{L}$ there exists a sequence $\{X^{(n)}\} \subset \mathscr{L}^0$ such that $[X^{(n)} - X] \to 0$ and by (5) we can show that $\{I(X^{(n)})\}$ is a Cauchy sequence in $\mathscr{M}_2^c$. One checks that the limit $I(X)$ again lies in $\mathscr{L}^*$, and that it is well-defined. This should also be done for $X \in \mathscr{L}^{**}$.

Definition 3.7. Let $X \in \mathscr{L}^*$. Then the stochastic integral of X with respect to the martingale $M \in \mathscr{M}_2^c$ is the unique, square-integrable martingale $I(X) = \{I_t(X), \mathscr{F}_t; 0 \le t < \infty\}$ such that $\lim_{n\to\infty} ||I(X^{(n)} - X|| = 0$ for every sequence $\{X^{(n)}\} \subset \mathscr{L}^0$ such that $\lim_{n\to\infty}[X^{(n)} - X] = 0$. We write

$$I_t(X) = \int_0^t X_s dM_s, \quad 0 \le t < \infty. \tag{3.2.8}$$

Example 3.8. Let X be a constant process with $X_t = c$ for all t, and take M to be Brownian motion B. Then for any partition $\Pi = \{t_0, \dots, t_n\}$ of $[0, t]$, we get straightaway the integral

$$I_t(X) = \sum_{i=0}^{\infty} c(B_{t\wedge t_{i+1}} - B_{t\wedge t_i}) = c(B_t - B_0). \tag{3.2.9}$$

Taking standard Brownian motion so that $B_0 = 0$, and $c = 1$ we have

$$\int_0^t dB_s = B_t. \tag{3.2.10}$$

Example 3.9. Now let $X = M = B$ be standard Brownian motion. Let $\Pi_n = \{t_0, \dots, t_n\}$ be a partition of $[0, t]$ with $0 = t_0 < t_1 < \cdots < t_n = t$. Approximate the stochastic integral $\int_0^t B_s\ dB_s$ by the sum

$$V(\Pi_n) = \sum_{n=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i}) = \sum_{i=0}^{n-1} \frac{1}{2}(B_{t_i+1}^2 - B_{t_i}^2 - (B_{t_{i+1}} - B_{t_i})^2). \tag{3.2.11}$$

Cancelling terms, the sum reduces to

$$\frac{1}{2}B_{t_n}^2 - \frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2, \tag{3.2.12}$$

and the last sum converges in L^2 to t. To see this, show (exercise) that

$$E\left[\frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2 - t\right]^2 = \frac{1}{2}\sum_{i=0}^{n-1} E[B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)], \tag{3.2.13}$$

and that this last quantity is bounded by $C\frac{1}{2}\sum_{i=0}^{n-1}(t_{i+1} - t_i)] \le Ct||\Pi||$. This uses the properties

1. If $B_s - B_t, 0 \le s < t$ is normally distributed with mean 0 and variance $t - s$, then for each positive integer n there is a constant C_n such that $E[B_t - B_s]^{2n} \le C_n|t - s|^n$.
2. If X is a square-integrable martingale and $0 \le s < t \le u < v$ then we have $E[(X_v - X_u)(X_t - X_s)] = E\{E[X_v - X_u|\mathscr{F}_s](X_t - X_s)\} = 0$, i.e., the cross product terms vanish.

Thus we get

$$\int_0^t B_s\ dB_s = \frac{1}{2}B_t^2 + \frac{1}{2}t. \tag{3.2.14}$$

Exercise 5. Fill in the gaps in the last example.

3.3. **A characterization of the integral.** Suppose $M = \{M_t, \mathscr{F}_t; 0 \le t < \infty\}$ and $N = \{N_t, \mathscr{F}_t; 0 \le t < \infty\}$ belong to $\mathscr{M}_2^c$, and take $X \in \mathscr{L}^*(M), Y \in \mathscr{L}^*(N)$. Then $I_t^M(X) = \int_0^t X_s dM_s$ and $I_t^N(Y) = \int_0^t Y_s dN_s$ are also in $\mathscr{M}_2^c$, and from the proof of Lemma 3.6(iii), we know that

$$\langle I_t^M(X)\rangle = \int_0^t X_u^2 d\langle M\rangle_u, \quad \langle I_t^N(Y)\rangle = \int_0^t Y_u^2 d\langle N\rangle_u. \tag{3.3.1}$$

We want to establish the cross variation formula

$$\langle I_t^M(X), I_t^N(Y)\rangle = \int_0^t X_u Y_u d\langle M, N\rangle_u, \quad t \ge 0, P\text{-a.s.}. \tag{3.3.2}$$

If X, Y are simple, this follows the computation as in the proof of Lemma 3.6(iii) that for $0 \le s < t < \infty$,

$$E[(I_s^M(X) - I_s^N(Y))(I_t^M(X) - I_t^N(Y))|\mathscr{F}_s] = E\left[\int_s^t X_u Y_u d\langle M, N\rangle_u\right] \tag{3.3.3}$$

P-a.s., which is equivalent to (3.3.2). We extend to the general case in several steps. The first is the following:

Proposition 3.10 (Kunita-Wanabe)**.** *Let $M, N \in \mathscr{M}_2^c, X \in \mathscr{L}^*(M), Y \in \mathscr{L}^*(N)$. Then a.s. we have*

$$\int_0^t |X_u Y_u| d\check{\xi}_u \le \left(\int_0^t X_u^2 d\langle M\rangle_u\right)^{1/2} \left(\int_0^t Y_u^2 d\langle N\rangle_u\right)^{1/2}, \quad 0 \le t < \infty, \tag{3.3.4}$$

where $\check{\xi}_s$ denote the total variation of the process $\xi = \langle M, N\rangle$ on $[0, u]$.

Proof sketch. $\check{\xi}$ is absolutely continuous with respect to $\varphi(\omega) := \frac{1}{2}(\langle M\rangle + \langle N\rangle)(\omega)$ for every $\omega \in \hat{\Omega}$ with $P(\hat{\Omega}) = 1$, and for every such ω the Radon-Nikodym theorem implies the existence of functions $f_i(\cdot, \omega) : [0, \infty) \to \mathbf{R}, i = 1, 2, 3$, such that

$$\langle M\rangle_t(\omega) = \int_0^t f_1(s, \omega) d\varphi_s(\omega), \quad \langle N\rangle_t(\omega) = \int_0^t f_2(s, \omega) d\varphi_s(\omega) \tag{3.3.5}$$

$$\xi_t(\omega) = \langle M, N\rangle_t(\omega) = \int_0^t f_3(s, \omega) d\varphi_s(\omega), \quad 0 \le t < \infty. \tag{3.3.6}$$

Consequently, for $\alpha, \beta \in \mathbf{R}$ and $\omega \in \tilde{\Omega}_{\alpha\beta} \subset \hat{\Omega}$ such that $P(\tilde{\Omega}_{\alpha\beta}) = 1$, we have

$$0 \le \langle \alpha M + \beta N\rangle_t(\omega) - \langle \alpha M + \beta N\rangle_s(\omega) \tag{3.3.7}$$

$$= \int_0^t (\alpha^2 f_1(s, \omega) + 2\alpha\beta f_3(s, \omega) + \beta^2 f_s(s, \omega)) d\varphi_s(\omega), \quad 0 \le s < t < \infty. \tag{3.3.8}$$

This can happen only if for every $\omega \in \tilde{\Omega}_{\alpha\beta}$, there exists a set $T_{\alpha\beta}(\omega) \in \mathscr{B}([0, \infty))$ with $\int_{T_{\alpha\beta}(\omega)} d\varphi_t(\omega) = 0$ and such that

$$\alpha^2 f_1(s, \omega) + 2\alpha\beta f_3(s, \omega) + \beta^2 f_s(s, \omega) \ge 0 \tag{3.3.9}$$

holds for every $t \notin T_{\alpha\beta}(\omega)$. Now let $\tilde{\Omega} = \cap_{\alpha,\beta \in \mathbf{Q}} \tilde{\Omega}_{\alpha\beta}, T(\omega) = \cup_{\alpha,\beta \in \mathbf{Q}} T_{\alpha\beta}(\omega)$ so that

$$P(\tilde{\Omega}) = 1, \quad \int_{T(\omega)} d\varphi_t(\omega) = 0, \quad \omega \in \tilde{\Omega}. \tag{3.3.10}$$

Fix $\omega \in \tilde{\Omega}$, then the last inequality holds for every $t \notin T(\omega)$ and every $\alpha, \beta \in \mathbf{Q}$, and thus also for every $\alpha, \beta \in \mathbf{R}$. In particular,

$$\alpha^2 |X_t(\omega)|^2 f_1(s, \omega) + 2\alpha |X_t(\omega) Y_t(\omega)| f_3(s, \omega) + |Y_t(\omega)|^2 |\beta^2 f_s(s, \omega) \ge 0, \quad t \notin T(\omega). \tag{3.3.11}$$

Integrating with respect to $d\varphi_t$, we obtain a.s.

$$\alpha^2 \int_0^t |X_u|^2 \langle M\rangle_u + 2\alpha \int_0^t |X_s Y_s| d\check{\xi} + \int_0^t |Y_t|^2 |\langle N\rangle_u \ge 0, \quad 0 \le t < \infty, \tag{3.3.12}$$

and the desired result follows by minimization over α. □

Lemma 3.11. *If $M, N \in \mathscr{M}_2^c, X \in \mathscr{L}^*(M)$, and $\{X^{(n)}\}_{n=1}^\infty \subset \mathscr{L}^*(M)$ is such that for some $T > 0$,*

$$\lim_{n\to\infty} \int_0^T |X_u^{(n)} - X_u| d\langle M\rangle_u = 0, \quad P\text{-a.s.} \tag{3.3.13}$$

then

$$\lim_{n\to\infty} \langle I(X^{(n)}), N\rangle_t = \langle I(X), N\rangle_t, \quad P\text{-a.s.}, 0 \le t \le T. \tag{3.3.14}$$

Proof. Using the property that $|\langle M, N\rangle|^2 \le \langle M\rangle\langle N\rangle$, we have for $0 \le t \le T$,
(3.3.15)

$$|\langle I(X^{(n)}) - I(X), N\rangle_t|^2 \le \langle I(X^{(n)}) - I(X)\rangle_t \langle N\rangle_t \le \int_0^T |X_u^{(n)} - X_u|^2 d\langle M\rangle_u \cdot \langle N\rangle_T.$$

□

Lemma 3.12. *If $M, N \in \mathscr{M}_2^c, X \in \mathscr{L}^*(M)$, then*

$$\langle I^M(X), N\rangle_t = \int_0^t X_u d\langle M, N\rangle_u, \quad P\text{-a.s.}, 0 \le t < \infty. \tag{3.3.16}$$

Proof. From [KS, Lemma 3.2.7] there exists a sequence $\{X^{(n)}\}_{n=1}^\infty$ of simple processes such that

$$\sup_{T>0} \lim_{n\to\infty} E\int_0^T |X_u^{(n)} - X_u|^2 d\langle M\rangle_u = 0. \tag{3.3.17}$$

Consequently for each $T > 0$ there is a subsequence $\{\tilde{X}^{(n)}\}_{n=1}^\infty$ such that

$$\lim_{n\to\infty} E\int_0^T |\tilde{X}_u^{(n)} - X_u|^2 d\langle M\rangle_u = 0 \text{ a.s..} \tag{3.3.18}$$

But then for simple processes we can conclude that

$$\langle I^M(\tilde{X}^{(n)}), N\rangle_t = \int_0^t \tilde{X}_u^{(n)} d\langle M, N\rangle_u, \quad P\text{-a.s.}, 0 \le t \le T. \tag{3.3.19}$$

Then letting $n \to \infty$ we obtain the result from the previous lemma and the Kunita-Watanabe inequality. □

Proposition 3.13. *Let $M, N \in \mathscr{M}_2^c, X \in \mathscr{L}^*(M), Y \in \mathscr{L}^*(N)$. Then (3.3.2) and (3.3.3) hold.*

Proof. The previous lemma tells us that $d\langle M, I^N(Y)\rangle_u = Y_u d\langle M, N\rangle_u$. Replacing N by $I^N(Y)$, in the previous lemma, we have
(3.3.20)

$$\langle I^M(X), I^N(Y)\rangle_t = \int_0^t X_u d\langle M, I^N(Y)\rangle_u = \int_0^t X_u Y_u d\langle M, N\rangle_u \quad P\text{-a.s.}, 0 \le t < \infty.$$

□

Exercise 6. Suppose $M = \{M_t, \mathscr{F}_t; 0 \le t < \infty\}$ and $N = \{N_t, \mathscr{F}_t; 0 \le t < \infty\}$ belong to $\mathscr{M}_2^c$, and take $X \in \mathscr{L}_\infty^*(M), Y \in \mathscr{L}_\infty^*(N)$. Then the martingales $I^M(X), I^N(Y)$ are uniformly integrable and have last elements $I_\infty^M(X), I_\infty^N(Y)$, the cross variation $\langle I^M(X), I^N(Y)\rangle_t$ converges a.s. as $t \to \infty$, and

$$E[I_\infty^M(X) I_\infty^N(Y)] = E\langle I^M(X), I^N(Y)\rangle_\infty = E\int_0^\infty X_t Y_t d\langle M, N\rangle_t. \tag{3.3.21}$$

In particular,

$$E\left(\int_0^\infty X_t dM_t\right)^2 = E\int_0^\infty X_t^2 d\langle M\rangle_t. \tag{3.3.22}$$

3.4. **Itô's formula.** A continuous semimartingale $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ is an adapted process which has the decomposition P-a.s.

$$X_t = X_0 + M_t + C_t = X_0 + M_t + (A_t^+ - A_t^-), \qquad 0 \le t < \infty \tag{3.4.1}$$

where $M = \{M_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a local martingale and $A^\pm = \{A_t^\pm, \mathscr{F}_t; 0 \le t < \infty\}$ are continuous, nondecreasing, adapted processes with $A_0^\pm = 0$ P-a.e. One can show that this decomposition is unique.

Theorem 3.14. *Let f be a real-valued function in $C^2(\mathbf{R})$, and $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ a continuous semimartingale. Then P-a.s.,*
(3.4.2)

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dM_s + \int_0^t f'(X_s)dC_s + \frac{1}{2}\int_0^t f''(X_s)d\langle M\rangle_s, \quad 0 \le t < \infty.$$

The first integral is the stochastic integral, whereas the second two are the usual Lebesgue-Stieltjes integrals.

The proof of this is in [KS, 3.3]. We will give a proof of the special case where X is Brownian motion, following [D, 7.6]. The claim then is that P-a.s.,

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds, \quad 0 \le t < \infty. \tag{3.4.3}$$

Proof. (Sketch) We first make a reduction: define the stopping time $T_n = \inf\{t \ge 0 : |B_t| \ge n \text{ or } \langle B\rangle_t \ge n\}$ for any $n \ge 1$. Also set $T_n = \infty$ if the set is empty. Certainly T_n is nondecreasing and tends to infinity as n grows large. Thus if we can establish Itô's formula for the stopped process $B_{t\wedge T_n}, t \ge 0$, then we obtain the desired result by letting $n \to \infty$. So we may assume that $B_t(\omega)$ and $\langle B\langle_t(\omega)$ are bounded on $[0,\infty) \times \Omega$, and hence we may assume also that $f, f', f""$ are bounded.

Let Π_n be a partition $0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_k^{(n)} = t$ of the interval $[0,t]$, such that $\max_{1 \le i \le k(n)} t_i^{(n)} - t_{i-1}^{(n)} \to 0$ as $n \to \infty$. An application of the mean value theorem and Taylors theorem implies that for any real $a < b$ there is a $c \in (a,b)$ such that

$$f(b) - f(a) = (b-a)f'(a) + \frac{1}{2}(b-a)^2 f''(c). \tag{3.4.4}$$

(Can you see why?) Applying the partitions, we get

$$\begin{aligned} f(B_t) - f(B_0) &= \sum_{i=0}^{k-1} f(B_{t_i^{(n)}}) - f(B_{t_{i-1}^{(n)}}) \\ &= \sum_{i=0}^{k-1} f'(B_{t_{i-1}^{(n)}})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) + \frac{1}{2}\sum_{i=0}^{k-1} f''(c_{t_i^{(n)},t_{i+1}^{(n)}})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2. \end{aligned}$$

So we want to show that the two sums converge a.s. to the two integrals respectively. The process

$$Y_s(\omega) := f'(B_s(\omega)), \quad 0 \le s < t, \omega \in \Omega \tag{3.4.5}$$

lies in $\mathscr{L}^*$, i.e., it is an adapted, continuous and bounded process, so we shall approximate it by the simple process

$$Y_s^{\Pi_n}(\omega) := f'(X_0(\omega))1_{\{0\}}(s) + \sum_{i=0}^{k-1} f'(X_{t_i^{(n)}}(\omega))1_{(t_i^{(n)},t_{i+1}^{(n)}]}(s). \tag{3.4.6}$$

Indeed, since B_t is uniformly continuous, we have by the bounded convergence theorem

$$E[I_t^2(Y^{\Pi_n} - Y)] = E\Big[\int_0^t |Y^{\Pi_n} - Y|^2 d\langle M\rangle_s\Big] \to 0 \tag{3.4.7}$$

as $||\Pi_n|| \to 0$. It follows then that

$$\sum_{i=0}^{k-1} f'(B_{t_i^{(n)}})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}}) = \int_0^t Y_s^{\Pi} dM_s \to \int_0^t Y_s dM_s \tag{3.4.8}$$

in quadratic mean.

For the second term, let $c_i^{(n)} = c_{t_i^{(n)}, t_{i+1}^{(n)}}$. We want to show that

$$\sum_{i=0}^{k-1} f''(c_i^{(n)})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 \to \int_0^t f''(B_s) d\langle B\rangle_s \tag{3.4.9}$$

in $L^1(\Omega, \mathscr{F}, P)$ as $||\Pi_n|| \to 0$. Define the function $g_s^{(n)}$ to be the function taking values $c_i^{(n)}$ on the interval $(t_i^{(n)}, t_{i+1}^{(n)}]$, and also $A_s = \sum_{t_{i+1}^{(n)} \leq s} (B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2$. Then we can write

$$\sum_{i=0}^{k-1} f''(c_i^{(n)})(B_{t_{i+1}^{(n)}} - B_{t_i^{(n)}})^2 = \int_0^t g_s^{(n)} dA_s. \tag{3.4.10}$$

We know that A_s converges a.e. to the quadratic variation $\langle B\rangle_s$, and by the continuity of f'' we have $g_{s_n}^{(n)} \to f''(B_s)$ as $n \to \infty$, for any $s_n \to s$.

Then the convergence now follows from [D, Lemma 7.6.2]. Namely, If there exists (i) finite measure μ_n on $[0, t$ converging weakly to a finite measure μ, and (ii) a sequence of functions g_n with $|g_n| \leq M$ and such that $g_n(s_n) \to g(s)$ for any sequence s_n in $[0, t]$ converging to s, then as $n \to \infty$ one has

$$\int_0^t g_n d\mu_n \to \int_0^t g d\mu. \tag{3.4.11}$$

□

We also need the formula for functions $f(t, x)$. We might as well state the formula for d-dimensional Brownian motion:

Theorem 3.15. *Let $f(t,x) : [0, \infty) \times \mathbf{R}^d \to \mathbf{R}$ be a $C^{1,2}$ function. Then for all $t \geq 0$ we have P-a.s.,*

(3.4.12)

$$f(t, B_t) - f(0, B_0) = \int_0^t \frac{\partial f}{\partial t}(s, B_s) ds + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(B_s) dB_s^i + \frac{1}{2} \sum_{i=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i^2}(B_s) ds.$$

Proof. See [D, Theorem 7.6.7], and [KS, 3.6] for the statement for general continuous local martingales. □

Corollary 3.16. *Let $\Delta = \sum_{i=1}^d \partial^2/\partial x_i^2$ be the Laplace operator. Then*

$$f(B_t) = f(B_0) + \int_0^t \nabla f(B_s) dB_s + \frac{1}{2} \int_0^t \Delta f(B_s) ds. \tag{3.4.13}$$

Now let's look at some examples.

Example 3.17. Taking $f = x$ recovers $B_t - B_0 = \int_0^t dB_s$. Taking $f = x^2$, we have

$$B_t^2 - B_0^2 = 2 \int_0^t B_s \, dB_s + t. \tag{3.4.14}$$

as before.

Example 3.18. Let's integrate

$$\int_0^t s \, dB_s. \tag{3.4.15}$$

From calculus one might guess the answer tB_t should be involved. So let's define $f(t, B_t) = tB_t$. Then using Itô's formula for $f(t, B_t)$,

$$tB_t = \int_0^t B_s ds + \int_0^t s dB_s, \tag{3.4.16}$$

or

$$\int_0^t s dB_s = tB_t - \int_0^t B_s ds. \tag{3.4.17}$$

Exercise 7 (Integration by parts)**.** Let $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ and $Y = \{Y_t, \mathscr{F}_t; 0 \le t < \infty\}$ be continuous (square-integrable) martingales. Prove that

$$\int_0^t X_s dY_s = X_t Y_t - X_0 Y_0 - \int_0^t Y_s dX_s - \langle X, Y \rangle_t, \quad 0 \le t < \infty \tag{3.4.18}$$

where $\langle X, Y \rangle_t$ is the cross-variation $\frac{1}{4}\langle X+Y \rangle_t - \frac{1}{4}\langle X-Y \rangle_t$.

This shows one difference from the familiar integration formula, where the cross-variation produces a correction term. The correction term can be compensated for by accounting for it in the definition of the stochastic integral, as is with the Fisk-Stratanovich integral. It is defined for a smaller-class of integrands than the Itô integral, but is a useful tool in modeling because it is more robust under perturbations. Let X, Y be continuous semimartingales. The Fisk-Stratanovich integral of X with respect to Y can be defined as:

$$\int_0^t Y_s \circ dX_s := \int_0^t Y_s dM_s + \int_0^t Y_s d(A^+ - A^-)_s + \frac{1}{2}\langle M, N \rangle_t, \tag{3.4.19}$$

for $0 \le t < \infty$. The first integral on the right-hand side is an Itô integral.

Exercise 8 (Exponential martingales)**.** Let $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ be a progressively measurable process such that for all $T \in [0, \infty)$, we have

$$P\left[\int_0^T X_t^2 d\langle M \rangle_t < \infty\right] = 1. \tag{3.4.20}$$

Then define for $0 \le s < t < \infty$,

$$\zeta_t^s(X) := \int_s^t X_u dB_u - \frac{1}{2}\int_s^t X_u^2 du \tag{3.4.21}$$

and $\zeta_t := \zeta_t^0(X)$. Let $Z_t = \exp(\zeta_t)$.

(1) Show that the process $Z = \{Z_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a supermartingale, and a martingale if X is also a simple process.
(2) Show that Z satisfies the stochastic integral equation

$$Z_t = 1 + \int_0^t Z_s X_s dB_s, \quad 0 \le t < \infty \tag{3.4.22}$$

by applying $f(x) = e^x$ to the semimartingale ζ_t. (Solution: First write

$$f(\zeta_t) = f(\zeta_0) + \int_0^t f'(\zeta_s) dM_s + \int_0^t f'(\zeta_s) dB_s + \frac{1}{2}\int_0^t f''(\zeta_s) d\langle M \rangle_s. \tag{3.4.23}$$

Now we have to use some facts from [KS]. The Z is a semimartingale with (local) martingale part $M_t = \int_0^t X_s dB_s$ and bounded variation part $C_t = -\frac{1}{2}X_s^2 ds$. Also, $dM_s = X_s dB_s$. Then the result follows.)

Note that it is easier to prove the last formula using the differential form of Itô's formula.

3.5. **Martingale characterisation of Brownian motion.** Recall that if B is a d-dimensional standard Brownian motion, then $\langle B^{(k)}, B^{(j)}\rangle_t = \delta_{kj}t$ for $1 \le k, j \le d, 0 \le t < \infty$. It turns out that this property characterises Brownian motion among continuous local martingales. The compensated Poisson process with intensity $\lambda = 1$ provides an example of a discontinuous square-integrable martingale with $\langle M\rangle_t = t$.

Theorem 3.19 (P. Lévy, 1948)**.** *Let $X = \{X_t = (X_t^{(1)}, \dots, X_t^{(d)}\}, \mathscr{F}_t, 0 \le t < \infty\}$ be a continuous adapted proces in $\mathbf{R}^d$ such that for every component $1 \le k \le d$ the process $M_t := X_t^{(k)} - X_0^{(k)}, 0 \le t < \infty$ is a continuous local martingale relative to $\{\mathscr{F}_t\}$ and $\langle M^{(k)}, M^{(j)}\rangle_t = \delta_{kj}t$ for $1 \le k, j \le d, 0 \le t < \infty$.*

Then X is a d-dimensional Brownian motion.

Proof. We want to show that for $0 \le s < t$, the random vector $X_t - X_s$ is independent of $\mathscr{F}_s$ and has the d-variate normal distribution with mean zero and covariance matrix equal to $(t-s)$ times the $(d \times d)$ identity. Using [KS, Lemma 2.6.13], it suffices to prove that for each $u \in \mathbf{R}^d$,

$$E[e^{i(u\cdot(X_t - X_s))}|\mathscr{F}_s] = e^{-1/2||u||^2(t-s)}, \quad P\text{-a.s.} \tag{3.5.1}$$

For a fixed u, the function $f(x) = e^{i(u\cdot x)}$ satisfies

$$\frac{\partial}{\partial x_j} f(x) = iu_j f(x), \quad \frac{\partial^2}{\partial x_j \partial x_k} f(x) = -u_j u_k f(x). \tag{3.5.2}$$

Applying Itô's formula to the real and imaginary parts of $f(x)$, we obtain

$$e^{i(u\cdot X_t)} = e^{i(u\cdot X_s)} + i\sum_{j=1}^d u_j \int_s^t e^{i(u\cdot X_u)} dM_u^{(j)} - \frac{1}{2}\sum_{j=1}^d u_j^2 \int_s^t e^{i(u, X_u)} du. \tag{3.5.3}$$

Since $|f(x)| \le 1$ for all x and $\langle M^{(j)}\rangle_t = t$, we have that $M^{(j)} \in \mathscr{M}_2^c$. Thus the real and imaginary parts of

$$\{\int_0^t e^{i(u\cdot X_s)} dM_s^{(j)}, \mathscr{F}_t; 0 \le t < \infty\}$$

belong to $\mathscr{M}_s^c$. Consequently, we have

$$E\left[\int_0^t e^{i(u\cdot X_s)} dM_s^{(j)} \middle| \mathscr{F}_s\right] = 0, \quad P\text{-a.s.}. \tag{3.5.4}$$

For any $A \in \mathscr{F}_s$, we may multiply (3.5.3) by $e^{-i(u\cdot X_s)}1_A$ and take expectations to obtain

$$E[e^{-i(u\cdot(X_t - X_s))}1_A] = P(A) - \frac{1}{2}||u||^2 \int_s^t E[e^{-i(u\cdot(X_t - X_s))}1_A] du. \tag{3.5.5}$$

This integral equation for the deterministic function $t \mapsto E[e^{-i(u\cdot(X_t - X_s))}1_A]$ is readily solved to yield

$$E[e^{-i(u\cdot(X_t - X_s))}1_A] = P(A)e^{-1/2||u||^2(t-s)}. \tag{3.5.6}$$

□

3.6. **The Girsanov theorem.** Here is the setup: Let W be a standard Brownian motion, and X a predictable process. We define for $0 \le s < t < \infty$,

$$\zeta_t^s(X) := \int_s^t X_u dW_u - \frac{1}{2}\int_s^t X_u^2 du, \tag{3.6.1}$$

and $\zeta_t(X) := \zeta_t^0(X)$. The process $\{\exp(\zeta_t(X)), \mathscr{F}_t, 0 \le t < \infty\}$ is a supermartingale. It is a martingale if X is a simple process. The Girsanov theorem shall give more general conditions for which it is a martingale.

In differential notation,

$$d\zeta_t = X_t dW_t - \frac{1}{2} X_t^2 dt, \tag{3.6.2}$$

and computing formally we have $(d\zeta_t)^2 = X_t^2 dt$. Then Itô's rule can be written as

$$df(\zeta_t) = f'(\zeta_t) d\zeta_t + \frac{1}{2} f''(\zeta_t)(d\zeta_t)^2, \tag{3.6.3}$$

so that with $f(x) = e^x$ and $Z_t := \exp(\zeta_t(X))$, we obtain

$$dZ_t = Z_t X_t dW_t - \frac{1}{2} Z_t x_t^2 dt + \frac{1}{2} Z_t x_t^2 dt = Z_t X_t dW_t, \tag{3.6.4}$$

and taking into account the initial condition $Z_0 = 1$, we then have the stochastic integral equation

$$Z_t = 1 + \int_0^t Z_s X_s dW_s, \quad 0 \le t < \infty. \tag{3.6.5}$$

3.6.1. Fix a probability space $(\Omega, \mathscr{F}, P)$ and a d-dimensional Brownian motion $W = \{W_t = (W_t^{(1)}, \ldots, W_t^{(d)}), \mathscr{F}_t; 0 \le t < \infty\}$ on it with $P(W_0 = 0) = 1$. Assume that the filtration $\{\mathscr{F}_t\}$ satisfies the usual conditions. Let $X = \{X_t = (X_t^{(1)}, \ldots, X_t^{(d)}), \mathscr{F}_t; 0 \le t < \infty\}$. be a vector of measurable, adapted processes satisfying

$$P\left[\int_0^T (X_t^{(i)})^2 dt < \infty\right] = 1, \quad 1 \le i \le d, 0 \le T < \infty. \tag{3.6.6}$$

Then for each i the stochastic integral $I^{W^{(i)}}(X^{(i)})$ is defined, and is a continuous local martingale. Define

$$Z_t(X) := \exp\left[\sum_{i=1}^d \int_0^t X_s^{(i)} dW_s^{(i)} - \frac{1}{2}\int_0^t ||X_s||^2 ds\right] \tag{3.6.7}$$

We have

$$Z_t(X) = 1 + \sum_{i=1}^d \int_0^t Z_s(X) X_s^{(i)} dW_s^{(i)} \tag{3.6.8}$$

which shows that $Z(X)$ is a continuous local martingale with $Z_0(X) = 1$.

If $Z(X)$ is a martingale, then $EZ_t(X) = 1, t \ge 0$, and for each $T \ge 0$ we can defined a probability measure $\tilde{P}_T$ on $\mathscr{F}_T$ by

$$\tilde{P}_T(A) := E[1_A Z_T(X)], \quad A \in \mathscr{F}_T. \tag{3.6.9}$$

The martingale property shows that the family of probability measures $\{\tilde{P}_T : 0 \le T < \infty\}$ satisfies the consistency condition $\tilde{P}_T(A) = \tilde{P}_t(A), A \in \mathscr{F}_t, 0 \le t \le T$.

Theorem 3.20 (Girsanov)**.** *Assume that $Z(X)$ above is a martingale. Define a process $\tilde{W} = \{\tilde{W}_t = (\tilde{W}_t^{(1)}, \ldots, \tilde{W}_t^{(d)}), \mathscr{F}_t; 0 \le t < \infty\}$ by*

$$\tilde{W}_t^{(i)} := W_t^{(i)} - \int_0^t X_t^{(i)} ds, \quad 1 \le i \le d, 0 \le t < \infty. \tag{3.6.10}$$

For each fixed $T \in [0, \infty)$, the process W is a d-dimensional Brownian motion $(\Omega, \mathscr{F}_T, \tilde{P}_T)$.

Corollary 3.21. *Let $W = \{W_t, \mathscr{F}_t; 0 \le t < \infty\}$ be the coordinate mapping process on $\Omega := C[0,\infty)^d$ so that $\mathscr{F}_\infty^W = \mathscr{B}(C[0,\infty)^d)$. Let P be the Weiner measure on $(\Omega, \mathscr{F}_\infty^W)$. Let $X = \{X_t, \mathscr{F}_t^W : 0 \le t < \infty\}$ be a d-dimensional process satisfying (3.6.6). If $Z(X)$ is a martingale, then there is a unique probability measure $\tilde{P}$ satisfying (3.6.9) and $\tilde{W}$ is a d-dimensional Brownian motion on $(\Omega, \mathscr{F}_\infty^W, \tilde{P})$.*

Proof. To see that $\tilde{W}$ is a Brownian motion on $(\Omega, \mathscr{F}_\infty^W, \tilde{P})$, let $0 \le t_1 < \cdots < t_n \le t$ be given. We have then

$$\tilde{P}[(\tilde{W}_t^{(1)}, \ldots, \tilde{W}_t^{(d)}) \in A] = \tilde{P}_t[(\tilde{W}_t^{(1)}, \ldots, \tilde{W}_t^{(d)}) \in A], \quad \mathcal{A} \in \mathscr{B}(\mathbf{R}^d). \tag{3.6.11}$$

The result then follows from the theorem. □

We will denote by $\tilde{E}_T$ (resp. $\tilde{E)}$ the expectation operator with respect to $\tilde{P}_T$ (resp. $\tilde{P}$).

Lemma 3.22. *Fix $0 \le T < \infty$ and assume that $Z(X)$ is a martingale. If $0 \le s \le t \le T$and Y is an $\mathscr{F}_t$-measurable random variable satisfying $\tilde{E}_T|Y| \le \infty$, then we have the Bayes' rule:*

$$\tilde{E}_T[Y|\mathscr{F}_s] = \frac{1}{Z_s(X)} E[Y Z_t(X)|\mathscr{F}_s], \quad P\text{- and } \tilde{P}\text{-a.s.}. \tag{3.6.12}$$

Proof. Using the definition of $\tilde{E}_T$, the definition of conditional expectation, and the martingale property, we have for any $A \in \mathscr{F}_s$:

$$\tilde{E}_T\left[1_A \frac{1}{Z_s(X)} E[Y Z_t(X)|\mathscr{F}_s]\right] = E[1_A E[Y Z_t(X)|\mathscr{F}_s]] \tag{3.6.13}$$

$$= E[1_A Y Z_t(X)] = \tilde{E}_T[1_A Y] \tag{3.6.14}$$

□

Proposition 3.23. *Fix $0 \le T < \infty$ and assume that $Z(X)$ is a martingale. If $M \in \mathscr{M}_T^{c,loc}$, then the process*

$$\tilde{M}_t := M_t - \sum_{i=1}^d \int_0^t X_s^{(i)} d\langle M, W^{(i)}\rangle_s, \quad 0 \le t \le T \tag{3.6.15}$$

is in $\tilde{\mathscr{M}}_T^{c,loc}$. If $N \in \mathscr{M}_T^{c,loc}$ and

$$\tilde{N}_t := N_t - \sum_{i=1}^d \int_0^t X_s^{(i)} d\langle N, W^{(i)}\rangle_s, \quad 0 \le t \le T, \tag{3.6.16}$$

then $\langle \tilde{M}, \tilde{N}\rangle_t = \langle M, N\rangle_t, 0 \le t \le T$ a.s. P and $\tilde{P}_T$.

Proof. We consider only the case where M and N are bounded martingales with bounded quadratic variations, and assume also that $Z_t(X)$ and $\sum_{i=1}^d \int_0^t (X_s^{(j)})^2 ds$ are bounded in t and ω. The general case can be reduced to this one by localization. From the Kunita-Watanabe inequality,

$$\left|\int_0^t X_s^{(i)} d\langle M, W^{(i)}\rangle_s\right|^2 \le \langle M\rangle_t \int_0^t (X_s^{(i)})^2 ds, \tag{3.6.17}$$

so $\tilde{M}$ is also bounded. The integration by parts formula gives

$$Z_t(X)\tilde{M}_t = \int_0^t Z_u(X) dM_u + \sum_{i=1}^d \int_0^t \tilde{M}_u X_u^{(i)} Z_u(X) dW_u^{(i)} \tag{3.6.18}$$

which is martingale under P. Therefore, we have from the previous lemma

$$\tilde{E}_T[\tilde{M}|\mathscr{F}_s] = \frac{1}{Z_s(X)} E[\tilde{M} Z_t(X)|\mathscr{F}_s], \quad P\text{- and } \tilde{P}\text{-a.s.}. \tag{3.6.19}$$

for $0 \le s \le t \le T$. It follows that $\tilde{M} \in \tilde{\mathscr{M}}^{c,\mathrm{loc}}$. The change of variable formula also implies

$$\tilde{M}_t\tilde{N}_t - \langle M, N\rangle_t = \int_0^t \tilde{M}_u dN_u + \int_0^t \tilde{N}_u dM_u \tag{3.6.20}$$

$$- \sum_{i=1}^d \left[\int_0^t [\tilde{M}_u\tilde{N}_u - \langle M, N\rangle_u] X_u^{(i)} Z_u(X) dW^{(i)}\right] \tag{3.6.21}$$

and

(3.6.22)

$$Z_t(X)[\tilde{M}_t\tilde{N}_t - \langle M, N\rangle_t] = \int_0^t Z_u(X)\tilde{M}_u dN_u + \int_0^t Z_u(X)\tilde{N}_u dM_u$$

$$- \sum_{i=1}^d \left[\int_0^t \tilde{M}_u X_u^{(i)} d\langle N, W^{(i)}\rangle_u + \int_0^t \tilde{N}_u X_u^{(i)} d\langle M, W^{(i)}\rangle_u\right] \tag{3.6.23}$$

This last process is consequently a martingale under P, and so the lemma implies that

$$\tilde{E}_T[\tilde{M}_t\tilde{N}_t - \langle M, N\rangle_t|\mathscr{F}_s] = \tilde{M}_s\tilde{N}_s - \langle M, N\rangle_s \tag{3.6.24}$$

This proves that $\langle \tilde{M}, \tilde{N}\rangle_t = \langle M, N\rangle_t, 0 \le t \le T$ a.s. $\tilde{P}_T$ and P. □

Proof of Theorem 3.20. We show that the continuous process $\tilde{W}$ on $(\omega, \mathscr{F}_T, \tilde{P}_T)$ satisfies the hypotheses Lévy's theorem. Setting $M = W^{(j)}$ in the proposition above, we obtain $\tilde{M} = \tilde{W}^{(j)}$, so $\tilde{W}^{(j)} \in \tilde{M}_T^{c,\mathrm{loc}}$. Setting $N = W^{(k)}$, we obtain $\tilde{P}_T$ and P-a.s.,

$$\langle \tilde{W}^{(j)}, \tilde{W}^{(k)}\rangle_t = \langle W^{(j)}, W^{(k)}\rangle_t = \delta_{jk}t, \quad 0 \le t \le T. \tag{3.6.25}$$

□

3.7. **The Novikov condition.** To use the Girsanov theorem, we need some conditions under which the process $Z(X)$ becomes a martingale. Define

$$T_n := \inf\left\{t \ge 0 : \max_{1\le i\le d} \int_0^t (Z_s(X)X_s^{(i)})^2 ds = n\right\}, \tag{3.7.1}$$

then the stopped processes $Z_{t\wedge T_n}(X)$ are martingales. Consequently, we have

$$E[Z_{t\wedge T_n}|\mathscr{F}_s] = Z_{s\wedge T_n}, \quad 0 \le s \le t, n \ge 1 \tag{3.7.2}$$

and using Fatou's lemma we have $E[Z_t(X)|\mathscr{F}_s] \le Z_s$, so $Z(X)$ is always a supermartingale and is a martingale if and only if

$$E[Z_t(X)] = 1, \quad 0 \le t < \infty \tag{3.7.3}$$

by [KS, 1.3.25].

Proposition 3.24. *Let $M = \{M_t, \mathscr{F}_t; 0 \le t < \infty\}$ be in $\mathscr{M}^{c,loc}$. and define $Z_t = \exp(M_t - \frac{1}{2}\langle M\rangle_t), 0 \le t < \infty$. If $E[\exp(\frac{1}{2}\langle M\rangle_t] < \infty$ for all t then $EZ_t = 1, 0 \le t < \infty$.*

Proof. See [KS 3.5.12]. □

Corollary 3.25 (Novikov)**.** *Let $W = \{W_t = (W_t^{(1)}, \ldots, W_t^{(d)}), \mathscr{F}_t, 0 \le t < \infty\}$ be a d-dimensional Brownian motion, and let $X = \{X_t = (X_t^{(1)}, \ldots, X_t^{(d)}), \mathscr{F}_t, 0 \le t < \infty\}$ be a vector of measurable, adapted processes satisfying (3.6.6). If*

$$E\left[\exp\left(\frac{1}{2}\int_0^T ||X_s||^2 ds\right)\right] < \infty, \quad 0 \le T < \infty, \tag{3.7.4}$$

then $Z(X)$ defined by (3.6.7) is a martingale.

Corollary 3.26. *The above corollary holds if (3.7.4) is replaced by the following assumption: There exists a sequence of real numbers $\{t_n\}$ with $0 = t_0 < t_1 < \cdots < t_n$, and $t_n \to \infty$ such that*

$$E\left[\exp\left(\frac{1}{2}\int_{t_{n-1}}^{t_n} ||X_s||^2 ds\right)\right] < \infty, \quad \forall n \geq 1. \tag{3.7.5}$$

Proof. Let $X_t(n) = (X_t^{(1)} 1_{[t_{n-1},t_n)}(t), \ldots, X_t^{(d)} 1_{[t_{n-1},t_n)}(t))$, so that $Z(X(n))$ is a martingale by Corollary 3.25. In particular,

$$E[Z_{t_n}(X(n))|\mathscr{F}_{t_{n-1}}] = E[Z_{t_n-1}(X(n))] = 1 \tag{3.7.6}$$

for all $n \geq 1$. But then

$$E[Z_{t_n}(X)] = E[Z_{t_{n-1}}(X)E[Z_{t_n}(X(n))|\mathscr{F}_{t_{n-1}}]] = E[Z_{t_{n-1}}(X)], \tag{3.7.7}$$

and by induction on n we can show that $E[Z_{t_n}(X)] = 1$ holds for all $n \geq 1$. Since $E[Z_t(X)]$ is nonincreasing in t and $t_n \to \infty$, we get $EZ_t(X) = 1$. $\square$

Definition 3.27. Let $C[0,\infty)^d$ be the space of continuous functions $x : [0,\infty) \to \mathbf{R}^d$. For $0 \leq t < \infty$, define $\mathscr{G}_t := \sigma(x(s) : 0 \leq s \leq t)$ and set $\mathscr{G} = \mathscr{G}_\infty$. A progressively measurable functional on $C[0,\infty)^d$ is a mapping $\mu : [0,\infty) \times C[0,\infty)^d \to \mathbf{R}$ which has the property that for each fixed $0 \leq t < \infty$, μ restricted to $[0,t] \times C[0,\infty)^d$ is $\mathscr{B}([0,t]) \otimes \mathscr{G}_t/\mathscr{B}(\mathbf{R})$-measurable.

If $\mu = (\mu^{(1)}, \ldots, \mu^{(d)})$ is a vector of progressively measurable functionals on $C[0,\infty)^d$ and W a d-dimensional Brownian motion on some $(\Omega, \mathscr{F}, P)$, then the processes

$$X_t^{(i)}(\omega) := \mu^{(i)}(t, W(\omega)), \quad 0 \leq t < \infty, 1 \leq i \leq d, \tag{3.7.8}$$

are progressively measurable relative to $\{\mathscr{F}_t\}$.

Corollary 3.28 (Beneš)**.** *Let the vector μ of progressively measurable functionals on $C[0,\infty)^d$ satisfy, for each $0 \leq T < \infty$ and for some $K_T > 0$, the condition*

$$||\mu(t,x)|| \leq K_T(1 + x^*(t)), \quad 0 \leq t \leq T \tag{3.7.9}$$

where $x^*(t) := \max_{0 \leq s \leq t} ||x(s)||$. Then with $X_t = (X_t^{(1)}, \ldots, X_t^{(d)})$ defined by (3.7.8), $Z(X)$ of (3.6.7) is a martingale.

Proof. If, for arbitrary $T > 0$, we can find $\{t_0, \ldots, t_n\}, n = n(T)$ such that $0 = t_0 < \cdots < t_n = T$ and (3.7.5) holds for $1 \leq n \leq n(T)$, then we can construct a sequence $\{t_n\}$ satisfying the previous corollary. Thus, fix $T > 0$. We have from (3.7.8) and (3.7.9) that whenever $0 \leq t_{n-1} < t_n \leq T$,

$$\int_{t_{n-1}}^{t_n} ||X_s^2|| ds \leq (t_n - t_{n-1})K_T^2(1 + W_T^*)^2 \tag{3.7.10}$$

where $W_T^* := \max_{0 \leq t \leq T} ||W_t||$. We claim that the process

$$Y_t := \exp(\frac{1}{4}(t_n - t_{n-1})K_T^2(1 + W_T^*)^2) \tag{3.7.11}$$

is a submartingale, and by Doob's maximal inequality we have

$$E[Y_t] = E\left[\max_{0 \leq t \leq T} T_t^2\right] \leq 4E[Y_t^2], \tag{3.7.12}$$

which is finite provided that $t_n - t_{n-1} \leq 1/TK_T^2$. This allows us to construct $t_0, \ldots, t_{n(T)}$ as described previously. $\square$

4. Stochastic differential equations

4.1. **Diffusion processes.** The study of stochastic differential equations, the existence and uniqueness of solutions, is really the study of diffusion processes.[1] The term diffusion is loosely attributed to a Markov process with continuous sample paths and that can be characterized in terms of an infinitesimal generator.

Consider a d-dimensional Markov family $X = \{X_t, \mathscr{F}_t, 0 \le t < \infty\}$ on $(\Omega, \mathscr{F})$ with the family of measures $\{P^x\}, x \in \mathbf{R}^d$, and assume that X has continuous paths. Suppose also that

$$\lim_{t\downarrow 0} \frac{1}{t} E^x[f(X_t) - f(X)] = (\mathscr{A} f)(x), \quad x \in \mathbf{R}^d \tag{4.1.1}$$

holds for all f in a suitable subclass of $C^2(\mathbf{R}^d)$. The limit is called the infinitesimal generator of the Markov family, applied to the test function f. The operator $\mathscr{A}$ is given by

$$(\mathscr{A} f)(x) := \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x) \frac{\partial f(x)}{\partial x_i} \tag{4.1.2}$$

for suitable Borel-measurable functions $b_i, a_{ij} : \mathbf{R}^d \to \mathbf{R}$, $1 \le i, j \le d$. This operator is called the second-order differential operator associated with the drift vector $(b_1, \dots, b_d)$ and diffusion matrix $a = (a_{ij})$ which is assumed to symmetric and nonnegative-definite for all $x \in \mathbf{R}^d$.

Heuristically, we can interpret the coefficients as follows: fix $x \in \mathbf{R}^d$, and let $f_i(y) = y_i$ and $f_{ij}(y) = (y_i - x_i)(y_j - x_j)$, for $y \in \mathbf{R}^d$. Assuming the limit exists, we obtain

$$E^x[X_t^{(i)} - x_i] = t b_i(x) + o(t) \tag{4.1.3}$$

$$E^x[(X_t^{(i)} - x_i)(X_t^{(j)} - x_j)] = t a_{ij}(x) + o(t) \tag{4.1.4}$$

as $t \downarrow 0$ for $1 \le i, j \le d$. In other words, the drift vector $b(x)$ measures locally the mean velocity of the random motion modeled by X, and $a(x)$ approximates the rate of change in the covariance matrix of the vector $X_t - x$ for small values of $t > 0$.

Definition 4.1. Let X be a d-dimensional Markov family such that

(1) X has continuous sample paths,
(2) (4.1.1) holds for every $f \in C^2(\mathbf{R}^d)$ which itself, its first- and second-order derivatives are bounded,
(3) (4.1.3) and (4.1.4) hold for every $x \in \mathbf{R}^d$,
(4) (a) for each $A \in \mathscr{F}$, the mapping $x \mapsto P^x(A)$ is universally measurable,
 (b) for each $x \in \mathbf{R}^d$, we have $P^x(X_0 = x) = 1$,
 (c) for each $x \in \mathbf{R}^d, A \in \mathscr{B}(\mathbf{R}^d)$, and any stopping time S of $\{\mathscr{F}_t\}$,
$$P^x[X_{S+t} \in A | \mathscr{F}_{S^+}] = P^x[X_{S^+} \in X_{S^+}], \quad P^x\text{-a.s.on } \{S < \infty\}$$
 (d) for each $x \in \mathbf{R}^d, A \in \mathscr{B}(\mathbf{R}^d)$, and any stopping time S of $\{\mathscr{F}_t\}$,
$$P^x[X_{S+t} \in A | X_S = y] = P^y[X_t \in A], \quad P^x X_S^{-1}\text{-a.s. } y$$

Then X is called a Kolmogorov-Feller diffusion process.

There are several approaches to the study of diffusions, ranging from the purely analytical to the purely probabilistic. In order to illustrate the traditional analytical approach, let us suppose that the Markov family above has a transition probability density function

$$P^x[X_t \in dy] = \Gamma(t; x, y) dy, \quad \forall x \in \mathbf{R}^d, t > 0. \tag{4.1.5}$$

[1]Another important branch of study are *jump processes*, which we shall not cover here.

Various heuristic arguments, with (4.1.1) as their starting point, can then be employed to suggest that $\Gamma(t;x,y)$ should satisfy the *forward Kolmogorov equation* or the *Fokker-Planck equation*, for every fixed $x \in \mathbf{R}^d$,

$$\frac{\partial}{\partial t}\Gamma(t;x,y) = \mathscr{A}^*\Gamma(t;x,y), \quad (t,y) \in (0,\infty) \times \mathbf{R}^d, \tag{4.1.6}$$

and the *backward Kolmogorov equation*, for every fixed $x \in \mathbf{R}^d$,

$$\frac{\partial}{\partial t}\Gamma(t;x,y) = \mathscr{A}\Gamma(t;x,y), \quad (t,y) \in (0,\infty) \times \mathbf{R}^d. \tag{4.1.7}$$

The operator $\mathscr{A}^*$ is the formal adjoint operator of $\mathscr{A}$, given by

$$(\mathscr{A}^* f)(y) := \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial y_i \partial y_j}[a_{ij}(y)f(y)] + \sum_{i=1}^{d} \frac{\partial}{\partial x_i}[b_i(y)f(y)] \tag{4.1.8}$$

provided of course that the coefficients a_{ij}, b_i posses the smoothness properties required. The early work of Kolmogorov (1931) and Feller (1936) used tools from the theory of partial differential equations to establish, under suitable and rather restrictive conditions, the existence of a solution $\Gamma(t;x,y)$ to the forward and backward Kolmogorov equations.

The methodology of stochastic differential equations was suggested by Lévy as an alternative, probabilistic approach to diffusions and was carried out in a masterly way by Itô. Suppose we have a continuous, adapted d-dimensional process $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ which satisfies for every $x \in \mathbf{R}^d$, the stochastic integral equation

$$X_t^{(i)} = x_i + \int_0^t b_i(X_s)ds + \sum_{i=1}^{r} \int_0^t \sigma_{ij}(X_s)dW_s^{(j)}, \quad 0 \le t < \infty, 1 \le i \le d, \tag{4.1.9}$$

on a probability space $(\Omega, \mathscr{F}, P^x)$ where W is a Brownian motion in $\mathbf{R}^r$ and the coefficients $b_i, \sigma_{ij} : \mathbf{R}^d \to \mathbf{R}$ are Borel-measurable. Then it is reasonable to expect that under certain conditions, (4.1.1), (4.1.3), and (4.1.4) will be valid, with

$$a_{ij}(x) = \sum_{k=1}^{r} \sigma_{ik}(x)\sigma_{kj}(x). \tag{4.1.10}$$

Thus the diffusion processes will be solutions to stochastic differential equations, which we now turn to.

4.2. **Strong solutions.** Let $b_i(t,x), \sigma_{ij}(t,x), 1 \le i \le d, 1 \le j \le r$ be Borel-measurable functions from $[0,\infty) \times \mathbf{R}^d$ to $\mathbf{R}$. Define the drift vector $b(t,x) = (b_i(t,x))$ and the dispersion matrix $\sigma(t,x) = (\sigma_{ij}(t,x))$. We want to assign meaning to the stochastic differential equation

$$dX_t = b(x, X_t)dt + \sigma(t, X_t)dW_t, \tag{4.2.1}$$

written component wise as

$$dX_t^{(i)} = b_i(x, X_t)dt + \sum_{j=1}^{r} \sigma_{ij}(t, X_t)dW_t^{(j)}, \quad 1 \le i \le d, \tag{4.2.2}$$

where $W = \{W_t, \mathscr{F}_t, 0 \le t < \infty\}$ is an r-dimensional Brownian motion and $X = \{X_t; 0 \le t < \infty\}$ is a suitable $\mathbf{R}^d$-valued stochastic process with continuous sample paths, which will be the 'solution' of the equation. The matrix $a(t,x) = \sigma(t,x)\sigma(t,x)^T$ with entries

$$a_{ik}(t,x) := \sum_{j=1}^{r} \sigma_{ij}(t,x)\sigma_{ik}(t,x), 1 \le i,k \le d \tag{4.2.3}$$

is called the diffusion matrix.

Fix a probability space $(\Omega, \mathscr{F}, P)$. Assume that this space is rich enough to accommodate an $\mathbf{R}^d$-valued random vector ξ, independent of $\mathscr{F}_\infty^W$ and with the given distribution

$$\mu(A) = P(\xi \in A), \quad A \in \mathscr{B}(\mathbf{R}^d). \tag{4.2.4}$$

Consider the left-continuous filtration

$$\mathscr{G}_t := \sigma(\xi, W_s; 0 \le s \le t, \quad 0 \le t < \infty, \tag{4.2.5}$$

and the collection of null sets

$$\mathscr{N} := \{N \subset \Omega : \exists G \in \mathscr{G}_\infty \text{ s.t. } N \subset G, P(G) = 0\}, \tag{4.2.6}$$

then define the augmented filtration

$$\mathscr{F}_t := \sigma(\mathscr{G}_t \cup \mathscr{N}), \quad 0 \le t < \infty. \tag{4.2.7}$$

Definition 4.2. A strong solution of the stochastic differential equation (4.2.1) on $(\Omega, \mathscr{F}, P)$ with respect to the fixed Brownian motion W and initial condition ξ, is a process $X = \{X_t; 0 \le t < \infty\}$ with continuous sample paths and satisfying

(1) X is adapted to the filtration $\{\mathscr{F}_t\}$ above,
(2) $P(X_0 = \xi) = 1$,
(3)

$$P(\int_0^t b_i(s, X_s) + \sigma_{ij}^2(s, X_s) ds < \infty) = 1 \tag{4.2.8}$$

for every $1 \le i \le d, 1 \le j \le r$, and $0 \le t < \infty$,

(4) the integral version of (4.2.1),

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad 0 \le t < \infty, \tag{4.2.9}$$

or equivalently

$$X_t^{(i)} = X_0^{(i)} + \int_0^t b_i(s, X_s) ds + \sum_{j=1}^r \int_0^t \sigma_{ij}(s, X_s) dW_s^{(j)}, \quad 1 \le i \le d, 0 \le t < \infty, \tag{4.2.10}$$

holds almost surely.

We note that property (1) is the most important part of the definition. It corresponds to X as the 'output' of the dynamical system described by b and σ, and whose input is W and ξ. The principle of causality in dynamics requires that X_t at time t only depend on ξ and the values of $W_s, 0 \le s \le t$. Moreover, the latter two should determine the output $\{X_t; 0 \le t < \infty\}$ in an unambiguous way. So we want to talk about uniqueness.

Definition 4.3. Let $b(t, x)$ and $\sigma(t, x)$ be given. Suppose that whenever W is an r-dimensional Brownian motion on some $(\Omega, \mathscr{F}, P)$, ξ is an independent d-dimensional random vector, $\{\mathscr{F}_t\}$ is given as above, and $X, \tilde{X}$ are two strong solutions relative to W with initial condition ξ, then $P(X_t = \tilde{X}_t; 0 \le t < \infty) = 1$. We then say that strong uniqueness holds for the pair b and σ.

Example 4.4. Let $d = 1$. Consider the equation $dX_t = b(t, X_t)dt + dW_t$, where b is a bounded, Borel-measurable,, and nonincreasing in the space variable, i.e., $b(t, x) \le b(t, y)$ for all $y \le x, 0 \le t < \infty$. Strong uniqueness holds for this equation. Indeed, for any two processes $X^{(1)}, X^{(2)}$ satisfying P-a.s.,

$$X_t^{(i)} = X_0 + \int_0^t b(s, X_s^{(i)}) ds + W_t, \quad i = 1, 2, 0 \le t < \infty. \tag{4.2.11}$$

Define the continuous process $\Delta = X_t^{(1)} - X_t^{(2)}$, and observe that P-a.s.,

$$\Delta_t^2 = 2 \int_0^t (X_s^{(1)} - X_s^{(2)})(b(s, X_s^{(1)} - b(s, X_s^{(2)})) ds \le 0, 0 \le t < \infty. \tag{4.2.12}$$

If the dispersion matrix $\sigma(t,x)=0$, then the stochastic integral equation reduces to an ordinary integral equation

$$X_t = X_0 + \int_0^t b(s,X_s)ds. \tag{4.2.13}$$

In the theory of such equations it is common to impose the assumption that the vector field $b(t,x)$ satisfies a local Lipschitz condition in the space variable x, and is bonded on compact subsets of $[0,\infty)\times\mathbf{R}^d$. These conditions ensure that for sufficiently small $t>0$, the so-called Picard-Lindelöf iterations $X_t^{(n)}$ converge to a solution. Here then is an existence result. For any $(d\times r)$ matrix σ, we define $||\sigma||^2=\sum_{I=1}^d\sum_{j=1}^r\sigma_{ij}^2$.

Theorem 4.5. *Suppose that the global Lipchitz and linear growth conditions are satisfied, namely,*

$$||b(t,x)-b(t,y)||+||\sigma(t,x)-\sigma(t,y)||\le K||x-y|| \tag{4.2.14}$$

$$||b(t,x)||^2+||\sigma(t,x)||^2\le K^2(1+||x||^2) \tag{4.2.15}$$

for every $0\le t<\infty$, and $x,y\in\mathbf{R}^d$, where K is a positive constant. Also, on some probability space $(\Omega,\mathscr{F},P)$ let ξ be an $\mathbf{R}^d$-valued random vector, independent of r-dimensional Brownian motion W, with finite second moment $E||\xi||^2<\infty$. Let $\mathscr{F}_t$ be as above.

Then there exists a continuous, adapted process $X=\{X_t,\mathscr{F}_t;0\le t<\infty\}$ which is a strong solution of (4.2.1) relative to W with the initial condition ξ. Moreover, X is square-integrable, i.e., for every $T>0$ there is a constant C depending only on K and T, such that

$$E||X_t||^2\le C(1+E||\xi||^2)e^{Ct},\quad 0\le t<T. \tag{4.2.16}$$

Proof. The idea of the proof is to construct a sequence of successive approximations by $X_t^{(0)}=\xi$, and

$$X_t^{(k+1)}:=\xi+\int_0^t b(s,X_s^{(k)})ds+\int_0^t\sigma(s,X_s^{(k)})dW_s, 0\le t<\infty, \tag{4.2.17}$$

for $k\ge 0$. Such processes are continuous and adapted to $\{\mathscr{F}_t\}$. We would like to show that it converges to the solution of (4.2.1).

Write $X_t^{(k+1)}-X_t^{(k)}=B_t+M_t$ where

(4.2.18)
$$B_t:=\int_0^t b(s,X_s^{(k)})-b(s,X_s^{(k-1)})ds,\quad M_t:=\int_0^t\sigma(s,X_s^{(k)})-\sigma(s,X_s^{(k-1)})dW_s.$$

By the Lipschitz and linear growth conditions, the process $M=\{M_t=(M_t^{(1)},\dots,M_t^{(d)}),\mathscr{F}_t;0\le t<\infty\}$ is seen to be a vector of square-integrable martingales. We claim that

$$E[\max_{0\le s\le t}||M_s||^2]\le\Lambda_1 E\left[\int_0^t||\sigma(s,X_s^{(k)})-\sigma(s,X_s^{(k-1)})||^2ds\right] \tag{4.2.19}$$

$$\le\Lambda_1K^2E\left[\int_0^t||X_s^{(k)}-X_s^{(k-1)}||^2ds\right]. \tag{4.2.20}$$

To prove the claim we show that for an vector M of continuous local martingales, notice that there exists a positive constant Λ such that

$$E[\max_{0\le s\le t}||M_s||]^2\le d\sum_{i=1}^d E[\max_{0\le s\le t}M_s]^2\le d\sum_{i=1}^d K_1E[\langle M^{(i)}\rangle_t], \tag{4.2.21}$$

and refer to [KS, Remark 3.3.30]

On the other hand, we have by Jensen's inequality and the Lipschitz condition,

$$E||B_t||^2\le K^2t\int_0^t E||X_s^{(k)}-X_s^{(k-1)}||^2ds, \tag{4.2.22}$$

so with $L = 4K^2(\Lambda_1 + T)$,

$$(4.2.23)\quad E\left[\max_{0\le s\le t} ||X_s^{(k)} - X_s^{(k-1)}||^2\right] \le L\int_0^t E||X_s^{(k)} - X_s^{(k-1)}||^2 ds, \quad 0\le t\le T.$$

The last inequality can be iterated to yield successive upper bounds

$$(4.2.24)\qquad E\left[\max_{0\le s\le t} ||X_s^{(k)} - X_s^{(k-1)}||^2\right] \le 4C\frac{(Lt)^k}{k!}, \quad 0\le t\le T.$$

where $C = \max_{0\le t\le T} E||X_t^{(1)} - \xi||^2$, which is finite because of (4.2.16). Applying the Čebyšev inequality, we get

$$(4.2.25)\qquad P\left(\max_{0\le s\le t} ||X_s^{(k)} - X_s^{(k-1)}|| > \frac{1}{2^{k+1}}\right) \le 4C\frac{(4LT)^k}{k!}, \quad k = 1, 2, \dots$$

and this upper bound is the general term in a convergent series. From the Borel-Cantelli lemma, we conclude that there exists an event $\Omega^* \in \mathscr{F}$ with $P(\Omega^*) = 1$ and an integer-valued random variable $N(\omega)$ such that for every $\omega \in \Omega^*$,

$$(4.2.26)\qquad \max_{0\le t\le T} ||X_s^{(k)}(\omega) - X_s^{(k-1)}(\omega)|| \le \frac{1}{2^{k+1}}, \quad k > N(\omega),$$

and consequently,

$$(4.2.27)\qquad \max_{0\le t\le T} ||X_s^{(k+m)}(\omega) - X_s^{(k)}(\omega)|| \le \frac{1}{2^{k}}, \quad m\ge 1, k > N(\omega).$$

We see then that the sequence of sample paths $\{X_t^{(k)}(\omega); 0\le t\le T\}_{k=1}^\infty$ is convergent in the sup norm on continuous functions, from which follows the existence of a continuous limit $\{X_t(\omega); 0\le t\le T\}$ for all $\omega\in\Omega^*$. Since T is arbitrary, we have the existence of a continuous process $X = \{X_t; 0\le t<\infty\}$ with the property that for P-a.e. ω, the sample paths $X^{(k)}(\omega)$ converge to $X(\omega)$ uniformly on compact subsets of $[0,\infty)$.

The inequality (4.2.16) is a consequence of the fact that for every $T > 0$ there is a constant C depending only on K and T, such that

$$(4.2.28)\qquad E||X_t^{(k)}||^2 \le C(1 + E||\xi||^2)e^{Ct}, \quad k\ge 0, 0\le t<T.$$

and Fatou's lemma. We first show that $X_t^{(k)}$ is defined for all $t\ge 0$. It will be enough to show that

$$(4.2.29)\qquad \int_0^t ||b(s, X_s^{(k)})||^2 + ||\sigma(s, X_s)^{(k)}||^2 ds < \infty, k\ge 0, \quad 0\le t<\infty, \text{ a.s.}.$$

By the linear growth condition, this will follow from

$$(4.2.30)\qquad \sup_{0\le t\le T} E||X_t^{(k)}|| < \infty.$$

We will prove this by induction. For $k = 0$, this is clear. Now assume it (4.2.30) true for some $k > 0$. Then following the proof of Theorem 4.8, we obtain the bound

$$(4.2.31)\qquad E||X^{(k+1)}||^2 \le 9E||\xi||^2 + 9(T+1)K^2\int_0^t (1 + E||X_s^{(k)}||^2) ds,$$

giving then (4.2.30) for $k+1$. Using the last inequality, we also have

$$(4.2.32)\qquad E||X^{(k+1)}||^2 \le C(1 + E||X_s^{(k)}||^2) + C\int_0^t E||X_s^{(k)}||^2 ds, \quad 0\le t\le T,$$

where C depends only on K and T. Then iterating the inequality gives

$$(4.2.33)\qquad E||X^{(k+1)}||^2 \le C(1 + E||X_s^{(k)}||^2)(1 + Ct + \frac{(Ct)^2}{2!} + \dots + \frac{(Ct)^{k+1}}{(k+1)!}),$$

and thus (4.2.16) follows.

This and the linearity condition gives condition (3) of Definition 4.2. Conditions (i) and (ii) are clearly satisfied. We leave as an exercise the proof that (4) is also satisfied, i.e., argue that

$$\left|\left|\int_0^t b(s, X_s^{(k)})ds - \int_0^t b(s, X_s)ds\right|\right|^2 \tag{4.2.34}$$

and

$$E\left|\left|\int_0^t \sigma(s, X_s^{(k)})dW_s - \int_0^t \sigma(s, X_s)dW_s\right|\right|^2 \tag{4.2.35}$$

converge to 0 a.s. for $0 \le t \le T$ as $k \to \infty$. Note that $\{X_t^{(k)}\}$ is a Cauchy sequence and $X_t^{(k)} \to X_t$ a.s. in $L^2(\Omega, \mathscr{F}, P)$.

□

Remark 4.6. The equation

$$\frac{dX_t}{dt} = X_t^2, \quad X_0 = 1 \tag{4.2.36}$$

corresponding to $b(x) = x^2$, which is does not satisfy the linear growth condition, has the unique solution

$$X_t = \frac{1}{1-t}, \quad 0 \le t < 1, \tag{4.2.37}$$

thus is is impossible to find a global solution i.e., one that is defined for all t in this case. More generally, the condition ensures that the solution $X_t(\omega)$ does not explode, i.e., $|X_t(\omega)|$ does not tend to infinity in a finite amount of time.

Remark 4.7. The equation

$$\frac{dX_t}{dt} = 3X_t^{\frac{2}{3}}, \quad X_0 = 0 \tag{4.2.38}$$

has more than one solution, namely, for any $a > 0$,

$$X_t = \begin{cases} 0 & t \le a \\ (t-a)^3 & t > a \end{cases}. \tag{4.2.39}$$

is a solution to the differential equation. In this case $b(x) = 3x^{\frac{2}{3}}$ does not satisfy the Lipschtiz condition at $x = 0$. Indeed, the condition guarantees uniqueness of the solution.

Indeed, the uniqueness is captured by the following theorem:

Theorem 4.8. *Suppose that $b(t, x)$ and $\sigma(t, x)$ are locally Lipschitz-continuous in x, i.e., for every integer $n \ge 1$ there exists a constant $K_n > 0$ such that for every $t \ge 0$, and $||x||, ||y|| \le n$,*

$$||b(t, x) - b(t, y)|| + ||\sigma(t, x) - \sigma(t, y)|| \le K_n||x - y||. \tag{4.2.40}$$

Then strong uniqueness holds for (4.2.1).

Proof. Suppose that X, X' are strong solutions defined for all $t \ge 0$ of (4.2.1) relatie to the same W and ξ on some $(\Omega, \mathscr{F}, P)$. Define the stopping times $\tau_n = \inf\{t \ge 0 : ||X_t|| \ge n\}$ for $n \ge 1$, and similarly τ'. Set $S_n := \tau_n \wedge \tau_n'$. Clearly $\lim_{n\to\infty} S_n = \infty$, P-a.s., and

$$X_{t\wedge S_n} - X'_{t\wedge S_n} = \int_0^{t\wedge S_n} (b(u, X_u) - b(u, X'_u))du + \int_0^{t\wedge S_n} (\sigma(u, X_u) - \sigma(u, X'_u))dW_u. \tag{4.2.41}$$

Using the vector inequality $||v_1 + \dots v_k||^2 \leq k^2(||v_1||^2 + \cdots + ||v_k||^2)$, the Hölder inequality for Lebesgue integrals, and the local Lipschitz condition, we have

$$\begin{aligned} &E||X_{t\wedge S_n} - X'_{t\wedge S_n}||^2 \\ &\leq 4E\left[\int_0^{t\wedge S_n} ||b(u,X_u) - b(u,X'_u)||du\right]^2 \\ &\quad + 4E\sum_{i=1}^d \left[\sum_{j=1}^r \int_0^{t\wedge S_n} (\sigma_{ij}(u,X_u) - \sigma_{ij}(u,X'_u))dW^{(j)}\right]^2. \end{aligned}$$

Then to the stochastic integral we apply the d-dimensional analogue of Itô's isometry, i.e., the 1-dimensional formula $E[I_t(X)]^2 = E\int_0^t X_u^2 du$, the property that

$$\begin{aligned} &E[\int_0^{t\wedge S_n} (\sigma_{ij}(u,X_u) - \sigma_{ij}(u,X'_u))dW_u^{(j)} \int_0^{t\wedge S_n} (\sigma_{ij}(u,X_u) - \sigma_{ij}(u,X'_u))dW_u^{(k)}] \\ &= E\int_0^{t\wedge S_n} (\sigma_{ij}(u,X_u) - \sigma_{ij}(u,X'_u))d\langle W^{(j)}, W^{(k)}\rangle_u \end{aligned}$$

and that for Brownian motion, $\langle W^{(j)}, W^{(k)}\rangle_t = \delta_{jk}t$.

We then have the upper bound

$$\begin{aligned} &4tE\left[\int_0^{t\wedge S_n} ||b(u,X_u) - b(u,X'_u)||^2 du\right] + 4E\left[\int_0^{t\wedge S_n} ||\sigma_{ij}(u,X_u) - \sigma_{ij}(u,X'_u)||^2\right] du \\ &\leq 4(T+1)K_n^2 \int_0^t E||X_{u\wedge S_n} - X'_{u\wedge S_n}||^2 du. \end{aligned}$$

We now need the Gronwall inequality: given a continuous function $a(t)$ such that

$$0 \leq g(t) \leq \alpha(t) + \beta\int_0^t g(s)ds, \quad 0 \leq t \leq T, \tag{4.2.42}$$

with $\beta \geq 0$ and $\alpha : [0,T] \to \mathbf{R}$ integrable. Then

$$g(t) \leq \alpha(t) + \beta\int_0^t \alpha(s)e^{\beta(t-s)}ds, \quad 0 \leq t \leq T. \tag{4.2.43}$$

We leave the proof of this as an exercise.

Apply the Gronwall inequality with $g(t) = E||X_{u\wedge S_n} - X'_{u\wedge S_n}||^2$ to conclude that $g(t) = 0$. Hence $\{X_{t\wedge S_n}; 0 \leq t < \infty\}$ and $\{X'_{t\wedge S_n}; 0 \leq t < \infty\}$ are modifications of one another, and thus indistinguishable, i.e.,

$$P(X_{t\wedge S_n} = X'_{t\wedge S_n}; 0 \leq t < \infty) = 1. \tag{4.2.44}$$

Letting $n \to \infty$, we see that the same is true for $\{X_t; 0 \leq t < \infty\}$ and $\{X_t; 0 \leq t < \infty\}$. □

4.3. **Examples.** If we allow for some randomness in some of the coefficients of a differential equation, we often obtain a more realistic mathematical model of the situation. Consider the simple population growth model

$$\frac{dN}{dt} = a(t)N(t), \qquad N(0) = N_0, \tag{4.3.1}$$

where $N(t)$ is the population at time t, and $a(t)$ is the relative growth rate at time t. It might happen that $a(t)$ is not completely known, but subject to some random environmental effects. So we have $a(t) = r(t) +$ 'noise', where we do not know the exact behaviour of the noise term, only its probability distribution. Or, more generally, we consider equations of the form

$$\frac{dX}{dt} = b(t,X_t) + \sigma(t,X_t)(\text{noise}) \tag{4.3.2}$$

where b, σ are given real-valued functions. The noise will be described by some stochastic process W_t. Considering first a discrete version of this,

$$X_{k+1} - X_k = b(t_k, X_k)\Delta t_k + \sigma(t_k, X_k)W_{t_k}\Delta t_k \tag{4.3.3}$$

for t_k, t_{k+1} in some partition $0 = t_0 < t_1 < \cdots < t_n = t$, and $\Delta t_k = t_{k+1} - t_k$. So we want to know that the limit of the right-hand side exists, in some sense, as $\Delta t_k \to 0$. If it does, then we can understand this to mean that X_t satisfies the stochastic integral equation

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \tag{4.3.4}$$

or in differential form, $dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$. So the work we have to put in is to give meaning to the stochastic integral $\int_0^t f(s,\omega)dB_s(\omega)$.

Remark 4.9. The differeintal form of Itô's rule can be written as

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \tag{4.3.5}$$

with the rule that $dX_t \wedge dX_t = dt, dt \wedge dt = 0, dX_t \wedge dt = 0$.

Example 4.10. Let's try to solve our population model

$$\frac{dN_t}{dt} = (r_t + \alpha W_t)N_t, \tag{4.3.6}$$

where α is a constant. Let's also assume that r_t is constant. Then we write this as

$$dN_t = rN_t dt + \alpha N_t dW_t, \tag{4.3.7}$$

hence

$$\int_0^t \frac{dN_s}{N_s} = rt + \alpha W_t, \quad W_0 = 0. \tag{4.3.8}$$

To evaluate the integral we use the differential form of Itô's formula on $g(t,x) = \ln x, x > 0$ to get in differential form,

$$d(\ln N_t) = \frac{1}{N_t}dN_t + \frac{1}{2}\Big(-\frac{1}{N_t^2}\Big)(dN_t)^2 = \frac{dN_t}{N_t} - \frac{1}{2}\alpha^2 dt, \tag{4.3.9}$$

since

$$(dN_t)^2 = (rN_t dt + \alpha N_t dW_t)^2 = \alpha^2 N_t^2 dt. \tag{4.3.10}$$

Hence

$$\ln\frac{N_t}{N_0} = (r - \frac{1}{2}\alpha^2)t + \alpha W_t, \tag{4.3.11}$$

or

$$N_t = N_0 \exp((r - \frac{1}{2}\alpha^2)t + \alpha W_t). \tag{4.3.12}$$

This is our solution. If W_t is Brownian motion, then the solution N_t is an example of geometric Brownian motion, which looks like $X_t = X_0 \exp(\mu t + \alpha B_t)$, for constants μ, α.

Remark 4.11. It seems reasonable that if B_t is independent of N_0 we should have $E[N_t] = E[N_0]e^{rt}$, so it is the sam as when there is no noise in $a(t)$. Let $Y^t = e^{\alpha B_t}$ and apply Itô's formula to get

$$Y_t = Y_0 + \alpha\int_0^t e^{\alpha B_s}dB_s + \frac{1}{2}\alpha^2\int_0^t e^{\alpha B_s}ds. \tag{4.3.13}$$

Since $E[\int_0^t f(s, X_s)dX_s] = 0$ is true for simple processes, it follows that $E[\int_0^t e^{\alpha B_s}dB_s] = 0$ also. Hence

$$E[Y_t] = E[Y_0] + \frac{1}{2}\alpha^2\int_0^t e[Y_s]ds, \tag{4.3.14}$$

or,

$$\frac{d}{dt}E[Y_t] = \frac{1}{2}\alpha^2 E[Y_t], \quad E[Y_0] = 1. \tag{4.3.15}$$

So it follows that $E[Y_t] = e^{\alpha^2 t/2}$ and hence $E[N_t] = E[N_0]e^{rt}$.

Example 4.12. The charge $Q(t)$ at time t at a fixed point in an electric circuit satisfies the differential equation

$$LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = F(t) \tag{4.3.16}$$

where L is inductance, R is resistance, C is capacitance, and $F(t)$ the potential source at time t. Again if we have a situation where some of the coefficients are not deterministic, say $F(t) = G(t) + \alpha W(t)$, we first introduce the vector $X(t,\omega) = (Q_t Q_t')^T$, to get

$$X_1' = X_2 \tag{4.3.17}$$

$$LX_2 = -RX_2 - \frac{1}{C}X_1 + G_t + \alpha W_t, \tag{4.3.18}$$

or, in matrix notation,

$$dX(t) = AX_t dt + H_t dt + K dB_t \tag{4.3.19}$$

where

$$dX = \begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix}, A = \begin{pmatrix} 0 & 1 \\ -\frac{1}{CL} & -\frac{R}{L} \end{pmatrix}, H_t = \begin{pmatrix} 0 \\ \frac{1}{L}G_t \end{pmatrix}, K = \begin{pmatrix} 0 \\ \frac{\alpha}{L} \end{pmatrix}, \tag{4.3.20}$$

and B_t is a one-dimensional Brownian motion.

Thus we are led to a 2-dimensional stochastic differential equation. Write

$$e^{-At}dX_t - e^{-At}AX_t dt = e^{-At}(H_t dt + K dB_t). \tag{4.3.21}$$

First apply Itô's formula to the function

$$f(t, x_1, x_2) = e^{-At}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{4.3.22}$$

to get

$$d(e^{-At}X_t) = -Ae^{-At}X_t dt + e^{-At}dX_t, \tag{4.3.23}$$

and substitute in the previous equation to get

$$e^{-At}X_t - X_0 = \int_0^t e^{-As}H_s ds + \int_0^t e^{-As}K dB_s, \tag{4.3.24}$$

or,

$$X_t = e^{At}\Big[X_0 + e^{-At}KB_t + \int_0^t e^{-At}(H_s + AKB_s)ds\Big] \tag{4.3.25}$$

by integration by parts.

Example 4.13 (Brownian motion on the unit circle)**.** Let $W = B = \{B_t, \mathscr{F}_t; 0 \le t < \infty\}$ be a 1-dimensional Brownian motion, and $g(t,x) = (\cos x, \sin x)$ for $x \in \mathbf{R}$. Then $X(t) = g(t, X_t)$ satisfies

$$dX_1(t) = -\sin(B_t)dB_t - \frac{1}{2}\cos(B_t)dt \qquad = -X_2 dBt - \frac{1}{2}X_1 dt \tag{4.3.26}$$

$$dX_2(t) = \cos(B_t)dB_t - \frac{1}{2}\sin(B_t)dt \qquad = X_1 dB_t - \frac{1}{2}X_2 dt. \tag{4.3.27}$$

Or, in matrix notation,

$$dX(t) = -\frac{1}{2}X(t)dt + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} Y(t)dB_t. \tag{4.3.28}$$

Exercise 9. Show that

(1) $X_t = e^{B_t}$ solves $dX_t = \frac{1}{2}X_t dt + X_t dB_t$.
(2) $X_t = B_t/(1+t), B_0$ solves $dX_t = -\frac{1}{1+t}X_t dt + \frac{1}{1+t}X_t dB_t$.
(3) $(X_1(t), X_2(t)) = (t, e^t B_t)$ solves $\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = \begin{pmatrix} 1 \\ X_2 \end{pmatrix} dt + \begin{pmatrix} 0 \\ e^{X_1} \end{pmatrix} dB_t$

Exercise 10 (Ornstein-Uhlenbeck process)**.** Solve the stochastic differential equation $dX_t = \mu X_t dt + \sigma dB_t$ where $\mu, \sigma \in R$. (Hint: multiply with the integrating factor e^{-t} and compare with $d(e^{-t}X_t)$.) Then find $E[X_t]$ and $\text{Var}[X_t] := E[(X_t - E[X_t])^2]$.

4.4. **Weak solutions.** In the case of strong solutions, the probability space $(\Omega, \mathscr{F}, P)$ is given, together with the data of a Brownian motion W and initial condition ξ on it. In the case of a weak solution, we only ask for the solution to be defined on *some* probability space, and for *some* Brownian motion and filtration.

Definition 4.14. A weak solution of the stochastic differential equation (4.2.1) is consists of

(1) $(\Omega, \mathscr{F}, P)$ a probability space
(2) $\{\mathscr{F}_t\}$ a filtration of sub-σ-algebras satisfying the usual conditions
(3) $W = \{W_t, \mathscr{F}_t; 0 \le t < \infty\}$ an r-dimensional Brownian motion
(4) $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ a continuous, adapted $\mathbf{R}^d$-valued process, such that conditions (3) and (4) of Definition 4.2 are satisfied.

The probability measure $\mu(A) := P(X_0 \in A), A \in \mathscr{B}(\mathbf{R}^d)$ is called the initial distribution of the solution.

Note that the filtration $\{\mathscr{F}_t\}$ is not necessarily the augmented filtration as in the strong solution, thus the value of the solution $X_t(\omega)$ at time t is not necessarily given by a measurable functional of the Brownian path $\{W_s(\omega); 0 \le s \le t\}$ and the initial condition $\xi(\omega) = X_0(\omega)$. On the other hand, since W is a Brownian motion relative to $\{\mathscr{F}_t\}$, the solution $X_t(\omega)$ cannot anticipate the future of the Brownian motion, besides $\{W_s(\omega); 0 \le s \le t\}$ and $\xi(\omega)$, whatever extra information required to compute $X_t(\omega)$ must be independent of $\{W_\theta(\omega) - W_t(\omega); t \le \theta < \infty\}$.

One consequence of this arrangement is that the existence of a weak solution does not guarantee, for a given Brownian motion $\tilde{W}$ on a (possibly different) probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ the existence of a process $\tilde{X}$ such that the tuple $(\tilde{X}, \tilde{W}), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{F}_t\}$ is again a weak solution. It is clear, however, that a strong solution is a weak solution.

Definition 4.15. Suppose that whenever $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ and $(\tilde{X}, \tilde{W}), (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}), \{\tilde{F}_t\}$ are weak solutions to (4.2.1) with a common Brownian motion (but possibly different filtrations) on a common probability space $(\Omega, \mathscr{F}, P)$ and with common initial value, i.e., $P(X_0 = \tilde{X}_0) = 1$, we have then the two $X, \tilde{X}$ are indistinguishable $P(X_t = \tilde{X}_t; 0 \le t < \infty) = 1$. We say then that pathwise uniqueness holds for (4.2.1).

Definition 4.16. We say that uniqueness in the sense of probability law holds for (4.2.1) if for any two weak solutions with the same initial distribution $P(X_0 \in A) = \tilde{P}(\tilde{X}_0 \in A)$ for all $A \in \mathscr{B}(\mathbf{R}^d)$, the two processes have the same law $PX^{-1} = \tilde{P}\tilde{X}^{-1}$.

Example 4.17. Consider the one dimensional equation

$$X_t = \int_0^t \text{sign}(X_s) dW_s, \quad 0 \le t < \infty. \tag{4.4.1}$$

If $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is a weak solution, then the process $X = \{X_t, \mathscr{F}_t; 0 \le t < \infty\}$ is a continuous square-integrable martingale with quadratic variation process $\langle X \rangle_t = \int_0^t \text{sign}(X_s)^2 ds = t$. Therefore, X is a Brownian motion and uniqueness in the sense of probability law holds. On the other hand, $(-X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$

is also a weak solution, so once we establish the existence of a weak solution, we have also shown that pathwise uniqueness cannot hold for X_t.

So start with a probability space $(\Omega, \mathscr{F}, P)$ and a one-dimensional Brownian motion $X = \{X_t, \dot{\mathscr{F}}_t^X; 0 \leq t < \infty\}$ on it with $P(X_0 = 0) = 1$ and $\{\dot{\mathscr{F}}_t^X\}$ the augmentation of the filtration $\{\mathscr{F}_t^X\}$ under P. The same argument as before shows that

$$W_t := \int_0^t \operatorname{sign}(X_s) dX_s, \quad 0 \leq t < \infty \tag{4.4.2}$$

is a Brownian motion adapted to $\{\dot{\mathscr{F}}_t^X\}$. Then one can show that this is a weak solution to (4.4.1) above.

Now we can produce weak solutions to stochastic differential equations using transformation of drift, via the Girsanov theorem.

Proposition 4.18. *Fix $T > 0$, and let W be a d-dimensional Brownian motion, and $b(t, x)$ a Borel measurable, $\mathbf{R}^d$-valued function on $[0, T] \times \mathbf{R}^d$ such that*

$$||b(t, x)|| \leq K(1 + ||x||) \quad 0 \leq t \leq T, x \in \mathbf{R}^d, \tag{4.4.3}$$

for some $K > 0$. Then for any probability measure μ on $(\mathbf{R}^d, \mathscr{B}(\mathbf{R}^d))$, the stochastic differential equation

$$dX_t = b(t, X_t)dt + dW_t, \quad 0 \leq t \leq T \tag{4.4.4}$$

has a weak solution with initial distribution μ.

Proof. Let $X = \{X_t, \mathscr{F}_t; 0 \leq t \leq T\}, (\Omega, \mathscr{F}), \{P^x\}_{x \in \mathbf{R}^d}$ be a Brownian family. According to Corollary 3.7.9

$$Z_t := \exp\left(\sum_{i=1}^d \int_0^t b_i(s, X_s) dX_s^{(i)} - \frac{1}{2}\int_0^t ||b(s, X_s)||^2 ds\right) \tag{4.4.5}$$

is a martingale under each measure P^x, so the Girsanov Theorem 3.20 implies that under Q^x given by the Radon-Nikodym derivative $\frac{dQ^x}{dP^x} = Z_T$, the process

$$W_t = X_t - X_0 - \int_0^t b(s, X_s)ds, \quad 0 \leq t \leq T, \tag{4.4.6}$$

is a Brownian motion with $Q^x(W_0 = 0) = 1$ for all $x \in \mathbf{R}^d$. Rewriting this as

$$X_t = X_0 + \int_0^t b(s, X_s)ds + W_t, \quad 0 \leq t \leq T, \tag{4.4.7}$$

we see that with $Q^\mu := \int_{\mathbf{R}^d} Q^x(A)\mu(dx)$, the triple $(X, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_t\}$ is a weak solution to (4.4.4). ☐

The Girsanov theorem can also be used to study the uniqueness in law of weak solutions.

Proposition 4.19. *Assume that $(X^{(i)}, W^{(i)}), (\Omega^{(i)}, \mathscr{F}^{(i)}, P^{(i)}), \{\mathscr{F}_t^{(i)}\}, i = 1, 2$ are weak solutions to (4.4.4) with the same initial distribution. If*

$$P^{(i)}\left[\int_0^T ||b(t, X_t^{(i)})||^2 dt < \infty\right] = 1, \quad i = 1, 2, \tag{4.4.8}$$

then $(X^{(i)}, W^{(i)}), i = 1, 2$ have the same law under their respective probability measures.

Proof. For each $k \geq 1$, let

$$\tau_k^{(i)} := T \wedge \inf\left\{0 \leq t \leq T : \int_0^T ||b(t, X_t^{(i)})||^2 dt = k\right\}. \tag{4.4.9}$$

Then by Novikov's condition, Corollary 3.25,

$$\xi_t^{(k)}(X^{(i)}) := \exp\left(-\int_0^{t\wedge\tau_k^{(i)}} b(s,X_s^{(i)})dX_s - \frac{1}{2}\int_0^{t\wedge\tau_k^{(i)}} ||b(s,X_s^{(i)})||^2 ds\right) \tag{4.4.10}$$

is a martingale, so we may define probability measures $\tilde{P}_k^{(i)}$ on $\mathscr{F}_T^{(i)}, i = 1,2$ according to $d\tilde{P}_k^{(i)}/dP^{(i)} = \xi_T^{(k)}(X^{(i)})$. The Girsanov Theorem 3.20 then implies that under $\tilde{P}_k^{(i)}$ the process

$$X^{(i)}_{t\wedge\tau_k^{(i)}} = X_0^{(i)} + \int_0^{t\wedge\tau_k^{(i)}} b(s,X^{(i)})ds + W^{(i)}_{t\wedge\tau_k^{(i)}}, \quad 0\le t\le T \tag{4.4.11}$$

is a d-dimensional Brownian motion with initial distribution μ, stopped at time $t\wedge\tau_k^{(i)}$.

But $\tau_k^{(i)}, \{W_t^{(i)} : 0\le t\le t\wedge\tau_k^{(i)}\}$, and $\xi_T^{(k)}(X^{(i)})$ can all be defined in terms of this latter process. Therefore, for $0 = t_0 < \cdots < t_n = T$ and $A\in\mathscr{B}(\mathbf{R}^{2d(n+1)})$ we have

$$P^{(1)}[(X_{t_0}^{(1)}, W_{t_0}^{(1)},\dots,X_{t_n}^{(1)},W_{t_n}^{(1)})\in A : \tau_k^{(1)} = T] \tag{4.4.12}$$

$$= \int_{\Omega^{(1)}} \frac{1}{\xi_T^{(k)}(X^{(1)})} 1_{\{(X_{t_0}^{(1)}, W_{t_0}^{(1)},\dots,X_{t_n}^{(1)},W_{t_n}^{(1)})\in A:\tau_k^{(1)}=T\}} d\tilde{P}_k^{(1)} \tag{4.4.13}$$

$$= \int_{\Omega^{(2)}} \frac{1}{\xi_T^{(k)}(X^{(2)})} 1_{\{(X_{t_0}^{(2)}, W_{t_0}^{(2)},\dots,X_{t_n}^{(2)},W_{t_n}^{(2)})\in A:\tau_k^{(2)}=T\}} d\tilde{P}_k^{(2)} \tag{4.4.14}$$

$$= P^{(2)}[(X_{t_0}^{(2)}, W_{t_0}^{(2)},\dots,X_{t_n}^{(2)},W_{t_n}^{(2)})\in A : \tau_k^{(2)} = T]. \tag{4.4.15}$$

By assumption (4.4.8),

$$\lim_{k\to\infty} P^{(i)}(\tau_k^{(i)} = T) = 1, \quad i = 1,2, \tag{4.4.16}$$

so passage to the limit in the last computation gives the desired conclusion. □

5. Applications

5.1. Basics on PDEs. Let's go back to analytic geometry. We have the familiar conic sections: the circle, ellipse, parabola, and hyperbola:

$$x^2+y^2=a^2, \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \quad y^2=4ax, \quad \frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{5.1.1}$$

A partial differential equation, (PDE) is a differential equation that contains unknown multivariable functions and their partial derivatives. This is in contrast to ordinary differential equations (ODE), which deal with functions of a single variable and their derivatives.

Consider a real-valued function $u(x,y)$. A second-order, linear, constant coefficient PDE for u is of the form

$$Au_{xx} + 2Bu_{xy} + Du_x + Eu_y + F = 0. \tag{5.1.2}$$

If we are able to replace the partials $\partial x, \partial y$ by say, X, Y, (which can be done formally by a Fourier transform), then we may convert the PDE into a polynomial of the same degree, with the top degree (a homogeneous polynomial) being most significant for the classification.

Indeed, as with the conic sections, we classify the PDE based on the discriminant, namely,

$$B^2 - 4AC \begin{cases} <0 & \text{elliptic,} \\ =0, & \text{parabolic,} \\ >0, & \text{hyperbolic.} \end{cases} \tag{5.1.3}$$

Important examples of such PDEs are the heat equation (parabolic) and the wave equation (hyperbolic). More generally, if we have a function u of $x_1, \dots, x_d$ variables, then it is classified according to the signature of the eigenvalues of the coefficient matrix (a_{ij}), where a_{ij} is the coefficient of $\partial^2 u/\partial x_i \partial x_j$.

Remark 5.1. We point out here that what we shall be studying in this section are deterministic PDEs. There are stochastic PDEs the same way there are stochastic (ordinary) differential equations like we have studied in the last section. In this section, we are taking a different approach, namely, studying deterministic PDEs using the stochastic methods which we have developed.

5.2. **The Dirichlet problem.** Recall that a function $u : D \subset \mathbf{R}^d \to \mathbf{R}$, where U is an open set, is called harmonic in D if u is of class C^2 and $\Delta u := \sum_{i=1}^d (\partial^2 u/\partial x_i^2) = 0$ in D. Let $W = \{W_t, \mathscr{F}_t; 0 \leq t < \infty\}, (\Omega, \mathscr{F}), \{P^x\}_{x \in \mathbf{R}^d}$ be a d-dimensional Brownian family and $\{\mathscr{F}_t\}$ satisfies the usual conditions. Define the stopping time

$$\tau_D = \inf\{t \geq 0 : W_t \in D^c\} \tag{5.2.1}$$

the first time of exit from D. Since W_t is almost surely unbounded (the Law of the Iterated Logarithm states that

$$\limsup_{t\to\infty} \frac{W_t(\omega)}{\sqrt{2t \log\log t}} = 1, \quad \liminf_{t\to\infty} \frac{W_t(\omega)}{\sqrt{2t \log\log t}} = -1 \tag{5.2.2}$$

), so

$$P^x(\tau_D < \infty) = 1, \quad x \in D \tag{5.2.3}$$

Let $B_r := \{x \in \mathbf{R}^d : ||x|| < r\}$ be the open ball of radius r centered at the origin. Its volume is

$$V_r := \frac{2r^d \pi^{d/2}}{d\Gamma(d/2)} \tag{5.2.4}$$

and its surface is $S - r := \frac{d}{r} V_r$. Define a probability measure μ_r on ∂B_r by

$$\mu_r(dx) = P^0[W_{\tau_{B_r}} \in dx], \quad r > 0. \tag{5.2.5}$$

Now, d-dimensional Brownian motion is rotationally invariant in the sense that for any $d \times d$ orthogonal matrix Q, i.e., $Q^T = Q^{-1}$, one checks that QW_t is also a d-dimensional Brownian motion. Therefore the measure μ_r is also rotationally invariant and thus proportional to the surface measure on ∂B_r. In particular, the Lebesgue integral of a function f over B can be written in the iterated form as

$$\int_{B_r} f(x)dx = \int_0^r S_\rho \int_{\partial B_\rho} f(x)\mu_\rho(dx)d\rho. \tag{5.2.6}$$

we say that the function u has the mean-value property if for every $a \in D$ and $0 < r < \infty$ such that $a + \bar{B}_r \subset D$, we have

$$u(a) = \int_{\partial B_r} u(a + x)dx. \tag{5.2.7}$$

One can derive using the previous identity

$$u(a) = \frac{1}{V_r} \int_{B_r} u(a + x)dx, \tag{5.2.8}$$

which says that the mean integral value of u over a ball is equal to the value at the center.

Proposition 5.2. *If u is harmonic in D, then it has the mean-value property there.*

Proof. For $a \in D$ and $0 < r < \infty$ such that $a + \bar{B}_r \subset D$, we have from Itô's rule (5.2.9)

$$u(W_{t\wedge\tau_{a+B_r}}) = u(W_0) + \sum_{i=1}^d \int_0^{t\wedge\tau_{a+B_r}} \frac{\partial u}{\partial x_i}(W_s)dW_s^{(i)} + \frac{1}{2}\int_0^{t\wedge\tau_{a+B_r}} \Delta u(W_s)ds,$$

for $0 \le t < \infty$. Since u is harmonic, the last integral vanishes, and since $\frac{\partial u}{\partial x_i}, 1 \le i \le d$ are bounded functions on $a + B$, the expectations under P^a of the stochastic integrals are all equal to 0. Taking expectation on both sides and letting $t \to \infty$, we use the fact that $P^x(\tau_D < \infty) = 1, \quad x \in D$ to obtain

$$u(a) = E^a u(W_{\tau_{a+B_r}}) = \int_{\partial B_r} u(a + x)\mu_r(dx) \tag{5.2.10}$$

□

Corollary 5.3 (Maximum principle)**.** *Suppose that u is harmonic in an open, connected domain D. If u achieves its supremum over D at some point in D, then it is identically constant.*

Proof. Let $M = \sup_{x \in D} u(x)$ and $D_M = \{x \in D : u(x) = M\}$. Assume that D_M is nonempty. We want to show that $D_M = D$. Since u is continuous, D_M is closed relative to D. But for $a \in D_M$ and $0 < r < \infty$ such that $a + \bar{B}_r \subset D$, we have the mean value property

$$M = u(a) = \frac{1}{V_r} \int_{B_r} u(a + x)dx, \tag{5.2.11}$$

which shows that $u = M$ on $a + B_r$. Therefore D_M is open. Since D is connected, then either D_M or $D \backslash D_M$ must be empty. □

We also state a converse:

Proposition 5.4 (KS 4.2.5)**.** *If $u : D \to \mathbf{R}$ has the mean value property, then u is smooth and harmonic.*

We can now describe the Dirichlet problem: Let D be an open subset of $\mathbf{R}^d$, and $f : \partial D \to \mathbf{R}$ be a continuous function. Find a continuous function $u : \bar{D} \to \mathbf{R}$ such that $u \in C^2(D)$, and

$$\Delta u = 0 \text{ in } D, \tag{5.2.12}$$

$$u = f \text{ on } \partial D. \tag{5.2.13}$$

If such a solution exists, then we call it a solution to the Dirichlet problem (D, f). One may interpret the solution $u(x)$ as the steady-state temperature at $x \in D$ when the boundary temperatures are specified by f.

In fact, using probabilistic methods we can immediately write down a likely solution.

Proposition 5.5. *Let*

$$u(x) = E^x[f(W_{\tau_D})], \quad x \in \bar{D}. \tag{5.2.14}$$

If $E^x|f(W_{\tau_D})| < \infty$ for all $x \in D$, then u is harmonic in D.

Proof. By definition of τ_D, u satisfies (5.2.13). Furthermore, for $a \in D$ and B_r chosen so that $a + \bar{B}_r \subset D$, we have from the strong Markov property

$$u(a) = E^a[f(W_{\tau_D})] = E^a\{E^a[f(W_{\tau_D})|\mathscr{F}_{\tau_{a+B_r}}]\} \tag{5.2.15}$$

$$= E^a[u(W_{\tau_{a+B_r}})] = \int_{\partial B_r} u(a + x)\mu_r(dx), \tag{5.2.16}$$

therefore u has the mean-value property, and so it must satisfy (5.2.12). The only unresolved issue is whether u is continuous up to and including ∂D. It turns out that this depends on the regularity of ∂D. □

5.2.1. *Uniqueness.* We can also establish a uniqueness result for the solution just obtained.

Proposition 5.6. *If f is bounded and*

$$P^a(\tau_D < \infty) = 1, \quad \forall a \in D, \tag{5.2.17}$$

then any bounded solution to (D, f) has the representation (5.2.14).

Proof. Let u be any bounded solution to (D, f), and let

$$D_n := \{x \in D : \inf_{y\in\partial D} ||x - y|| > \frac{1}{n}\}. \tag{5.2.18}$$

From Itô's rule, we have

(5.2.19)

$$u(W_{t\wedge\tau_{B_n}\wedge\tau_{D_n}}) = u(W_0) + \sum_{i=1}^{d}\int_0^{t\wedge\tau_{B_n}\wedge\tau_{D_n}} \frac{\partial u}{\partial x_i}(W_s)dW_s^{(i)}, \quad 0 \le t < \infty, n \ge 1.$$

Since $\partial u/\partial x_i$ is bounded in $B_n \cap D_n$, we may take expectations and conclude that

$$u(a) = E^a[u(W_{t\wedge\tau_{B_n}\wedge\tau_{D_n}})]; \quad 0 \le t < \infty, n \ge 1, a \in D_n. \tag{5.2.20}$$

As $t \to \infty$ and $n \to \infty$, (5.2.17) implies that $u(W_{t\wedge\tau_{B_n}\wedge\tau_{D_n}})$ converges to $f(W_{\tau_D})$, P^a-a.s. The representation (5.2.14) then follows from the bounded convergence theorem. □

5.2.2. *Regularity.* It remains to characterise points $a \in \partial D$ for which

$$\lim_{\substack{x\to a\\ x\in D}} E^x[f(W_{\tau_D})] = f(a) \tag{5.2.21}$$

holds for every bounded, measurable function $f : \partial D \to \mathbf{R}$ which is continuous at the point a.

Definition 5.7. Consider the stopping time of the right-continuous filtration $\{\mathscr{F}_t\}$ given by $\sigma_D := \inf\{t > 0 : W_t \in D^c\}$, in contrast to the definition of τ_D. we call a point $a \in \partial D$ regular for D if $P^a(\sigma_D = 0) = 1$, i.e., a Brownian path started at a does not immediately return to D and remain there for a nonempty time interval.

Also we call a irregular if $P^a(\sigma_D = 0) < 1$. However, the event $\{\sigma_D = 0\}$ belongs to $\mathscr{F}_{0+}^W$ so the Blumenthal zero-one law gives for an irregular point $P^a(\sigma_D = 0) = 0$.

Clearly regularity is a local condition, i.e., $a \in \partial D$ is regular for D if and only if a is regular for $(a + B_r) \cap D$ for some $r > 0$.

IN the one-dimensional case, every point of ∂D is regular [KS, 2.7.18], and the Dirichlet problem is always solvable, the solution being piecewise linear. When $d \ge 2$, more interesting behaviour can occur. In particular, if $D = \{x \in \mathbf{R}^d : 0 < ||x|| < 1\}$ is a punctured ball, then for any $x \in D$ the Brownian motion starting at x exits from D on its outer boundary, not at the origin [KS, 3.3.22]. This means that u defined by (5.2.14) is determined solely by the values of f along the outer boundary of D, and, except at the origin, this u will agree with the harmonic function

$$\tilde{u}(x) := E^x[f(W_{\tau_{B_1}})] = E^x[f(W_{\sigma_D})], \quad x \in B_1. \tag{5.2.22}$$

Now $u(0) := f(0)$, so u is continuous at the origin if and only if $f(0) = \tilde{u}(0)$. When $d \ge 3$, it is even possible for ∂D to be connected but contain irregular points.

Theorem 5.8. *Assume that $d \ge 2$ and fix $a \in \partial D$. The following are equivalent:*

1. (5.2.21) *holds for every bounded, measurable function $f : \partial D \to \mathbf{R}$ which is continuous at a,*
2. *a is regular for D,*
3. *for all $\epsilon > 0$, we have*

$$\lim_{\substack{x\to a\\ x\in D}} P^x(\tau_D > \epsilon) = 0. \tag{5.2.23}$$

Proof. Assume without loss of generality that $a = 0$. We first prove (i) $\Rightarrow$ (ii) by contradiction. If the origin is irregular, then $P^0(\sigma_D = 0) = 0$. Since a Brownian motion of dimension $d \geq 2$ never returns to its origin [KS 3.3.22], we have

$$\lim_{r \downarrow 0} P^0(W_{\sigma_D} \in B_r) = P^0(W_{\sigma_D} = 0) = 0. \tag{5.2.24}$$

Fix $r > 0$ for which $P^0(W_{\sigma_D} \in B_r) < \frac{1}{4}$, and choose a sequence $\{\delta_n\}_{n=1}^{\infty}$ such that $0 < \delta_n < r$ for all n and $\delta_n \downarrow 0$. With $\tau_n := \inf\{t \geq 0 : ||W_t|| \geq \delta_n\}$, we have $P^0(\tau_n \downarrow 0) = 1$, and thus

$$\lim_{n \to \infty} P^0(\tau_n < \sigma_D) = 1. \tag{5.2.25}$$

Furthermore, on the event $\{\tau_n < \sigma_D\}$, we have $W_{\tau_n} \in D$. For n large enough so that $P^0(\tau_n < \sigma_D) \geq \frac{1}{2}$, we may write

$$\frac{1}{4} > P^0(W_{\sigma_D} \in B_r) \geq P^0(W_{\sigma_D} \in B_r, \tau_n < \sigma_D) \tag{5.2.26}$$

$$= E^0[1_{\{\tau_n < \sigma_D\}} P^0(W_{\sigma_D} \in B_r | \mathscr{F}_{\tau_n})] \tag{5.2.27}$$

$$= \int_{D \cap B_{\delta_n}} P^x(W_{\tau_D} \in B_r) P^0(\tau_n < \sigma_D, W_{\tau_n} \in dx) \tag{5.2.28}$$

$$\geq \frac{1}{2} \inf_{x \in D \cap B_{\delta_n}} P^x(W_{\tau_D} \in B_r). \tag{5.2.29}$$

Hence we conclude that $P^{x_n}(W_{\tau_D} \in B_r) \leq \frac{1}{2}$ for some $x_n \in D \cap B_{\delta_n}$. Now choose a bounded continuous function $f : \partial D \to \mathbf{R}$ such that $f = 0$ outside B_r, $f \leq 1$ inside B_r and $f(0) = 1$. For such a function we have

$$\limsup_n E^{x_n}[f(W_{\tau_D})] \leq \limsup_n P^{x_n}(W_{\tau_D \in B_r}) \leq \frac{1}{2} < f(0), \tag{5.2.30}$$

and (i) fails.

(ii) $\Rightarrow$ (iii). Observe first of all that for $0 < \delta < \epsilon$, the function

$$g_\delta(x) := P^x(W_s \in D : \delta \leq s \leq \epsilon) = E^x[P^{W_\delta}(\tau_D > \epsilon - \delta)] \tag{5.2.31}$$

$$= \int_{\mathbf{R}^d} P^y(\tau_D > \epsilon - \delta) P^x(W_\delta \in dy) \tag{5.2.32}$$

is continuous in x. But

$$g_\delta(x) \searrow g(x) P^x(W_s \in D : 0 < s \leq \epsilon) = P^x(\sigma_D > \epsilon) \tag{5.2.33}$$

as $\delta \downarrow 0$, so g is upper semicontinuous. From this fact and the inequality $\tau_D \leq \sigma_D$, we conclude that

$$\limsup_{x \in D \to 0} P^x(\tau_D > \epsilon) \leq \limsup_{x \to 0} g(x) \leq g(0) = 0 \tag{5.2.34}$$

by (ii).

(iIi) $\Rightarrow$ (i). We know that for each $r > 0$,

$$P^x(\max_{0 \leq t \leq \epsilon} ||W_t - W_0|| < r) \tag{5.2.35}$$

does not depend on x and approaches 1 as $\epsilon \downarrow 0$. But then

$$P^x(||W_{\tau_D} - W_0|| < r) \geq P^x \left[\left\{ \max_{0 \leq t \leq \epsilon} ||W_t - W_0|| < r \right\} \cap \{\tau_D \leq \epsilon\} \right] \tag{5.2.36}$$

$$\geq P^0 \left[\max_{0 \leq t \leq \epsilon} ||W_t|| < r \right] - P^x[\tau_D > \epsilon]. \tag{5.2.37}$$

Letting $x \in D \to 0$ and $\epsilon \to 0$ successively, we obtain from (iii)

$$\lim_{x \in D \to 0} P^x(||W_{\tau_D} - x|| < r) = 1, \quad 0 < r < \infty. \tag{5.2.38}$$

The continuity of f at the origin and its boundedness on ∂D then give us (i). $\square$

5.3. **The one-dimensional heat equation.** In this section, we shall establish stochastic representations for the temperatures in a rod. Consider an infinite rod, insulated and extended along the x-axis of the (t, x)-plane, and let $f(x)$ denote the temperature of the rod at time $t = 0$ and location x. If $u(t, x)$ is the temperature of the rod at time $t \geq 0$ and position $x \in \mathbf{R}$, then wit hthe appropriate choice of units, u will satisfy the *heat equation*,

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} \tag{5.3.1}$$

with initial condition $u(0, x) = f(x), x \in \mathbf{R}$. The starting point of our probablistic treatment is furnished by the observation that the transition density

$$p(t; x, y) := \frac{1}{y}P^x[W_t \in dy] = \frac{1}{\sqrt{2\pi t}}e^{-(x-y)^2/2t}, \quad t > 0, x, y \in \mathbf{R} \tag{5.3.2}$$

of the one-dimesional Brownian family satisfies the partial differential equation

$$\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial x^2} \tag{5.3.3}$$

(prove this). Suppose then that $f : \mathbf{R} \to \mathbf{R}$ is a Borel-measurable function satisfying the condition

$$\int_{-\infty}^{\infty} e^{-ax^2}|f(x)|dx < \infty, \tag{5.3.4}$$

for some $a > 0$. It is well known that

$$u(t, x) := E^x[f(W_t)] = \int_{-\infty}^{\infty} f(y)p(t; x, y)dy \tag{5.3.5}$$

is defined for $0 < t < 1/2a, x \in \mathbf{R}$, has derivatives of all orders, and satisfies the heat equation (5.3.1).

Exercise 11. Show that for any nonnegative integers n, m, under the assumption (5.3.4), we have

$$\frac{\partial^{n+m}}{\partial t^n \partial x^m}u(t, x) = \int_{-\infty}^{\infty} f(y)\frac{\partial^{n+m}}{\partial t^n \partial x^m}p(t; x, y)dy, \quad 0 < t < \frac{1}{2a}. \tag{5.3.6}$$

If f is bounded and continuous, then rewriting (5.3.5) as $u(t, x) = E^0[f(x+W_t)]$, we can use the bouned convergence theorem to conclude

$$f(x) = \lim_{t\downarrow 0, y\to x} u(t, y) \tag{5.3.7}$$

for all $x \in \mathbf{R}$.

5.3.1. *Tychonoff uniqueness.* We shall call $p(t; x, y)$ a fundamental solution to the problem of finding a function u which satisfies (5.3.1) and agrees with the specified function f at time $t = 0$.

We shall say that a function $u : \mathbf{R}^m \to \mathbf{R}$ has continuous derivatives up to a certain order on a set G, if these derivatives exist and are continuous in the interior of G, and have continuous extensions to that part of the boundary ∂G which is included in G. With this convention, we can state the following uniqueness theorem.

Theorem 5.9 (Tychonoff, 1935)**.** *Suppose that the function u is $C^{1,2}$ on $(0, T] \times \mathbf{R}$ an satisfies (5.3.1) there, as well as the conditions*

$$\lim_{t\downarrow 0, y\to x} u(t, y) = 0, \quad x \in \mathbf{R}, \tag{5.3.8}$$

$$\sup_{0<t\leq T} |u(t, x)| \leq Ke^{ax^2}, \quad x \in \mathbf{R} \tag{5.3.9}$$

for constants $K, a > 0$. Then $u = 0$ on $(0, T] \times \mathbf{R}$.

Proof. See [KS, 4.3.3]. □

Remark 5.10. In particular, applying this to u_1, u_2 satisfying

$$\lim_{t\downarrow 0, y\to x} u_1(t,y) = \lim_{t\downarrow 0, y\to x} u_2(t,y), \quad x \in \mathbf{R} \tag{5.3.10}$$

and the other conditions of the theorem, then applying the theorem to $u_1 - u_2$ implies that $u_1 = u_2$ on $(0,T)\times\mathbf{R}$.

5.3.2. *Nonnegative solutions.* If the initial temperature f is nonnegative, as it always is if measured on the absolute scale, then the temperature should remain nonnegative for all $t > 0$; this is evident from the representation (5.3.1). Is it possible to characterize the nonnegative solutions of the heat equation? This was done by Widder (1944), who showed that such functions u have a representation

$$u(t,x) = \int_{-\infty}^{\infty} p(t;x,y)dF(y), \quad x \in \mathbf{R}, \tag{5.3.11}$$

where $F : \mathbf{R} \to \mathbf{R}$ is nondecreasing.

Theorem 5.11. *Let $v(t,x)$ be a nonnegative function on $(0,T)\times\mathbf{R}$. The following are equivalent*

(1) *For some nondecreasing function $F : \mathbf{R} \to \mathbf{R}$, we have*

$$v(t,x) = \int_{-\infty}^{\infty} p(T-t;x,y)dF(y), \quad 0 < t < T, x \in \mathbf{R}, \tag{5.3.12}$$

(2) *$v \in C^{1,2}((0,T)\times\mathbf{R})$ and satisfies the 'backward' heat equation*

$$\frac{\partial v}{\partial t} + \frac{1}{2}\frac{\partial^2 v}{\partial x^2} = 0 \tag{5.3.13}$$

on this strip,

(3) *For a Brownian family $\{W_s, \mathscr{F}_s; 0 \le s < \infty\}, (\Omega, \mathscr{F}), \{P^x\}_{x\in\mathbf{R}}$, and each fixed $t \in (0,T), x \in \mathbf{R}$, the process $\{v(t+s, W_s), \mathscr{F}_s, 0 \le s < T-t\}$ is a martingale on $(\Omega, \mathscr{F}, P^x)$*

(4) *For a Brownian family $\{W_s, \mathscr{F}_s; 0 \le s < \infty\}, (\Omega, \mathscr{F}), \{P^x\}_{x\in\mathbf{R}}$, we have*

$$v(t,x) = E^x[v(t+s, W_s))], \quad 0 < t \le t+s < T, x \in \mathbf{R}. \tag{5.3.14}$$

Proof. See [KS, 4.3.6]. □

5.4. **The Feynman-Kac formula.** The Feynman-Kac formula is a representation for the solution of the *parabolic equation*

$$\frac{\partial u}{\partial t} + ku = \frac{1}{2}\Delta u + g, \quad (t,x) \in (0,\infty)\times\mathbf{R}^d, \tag{5.4.1}$$

subject to the initial condition

$$u(0,x) = f(x), x \in \mathbf{R}^d \tag{5.4.2}$$

for suitable functions $k : \mathbf{R}^d \to [0,\infty), g : (0,\infty)\times\mathbf{R}^d \to \mathbf{R}$, and $f : \mathbf{R}^d \to \mathbf{R}$.

In the special case where $g = 0$, we may defined the Laplace transform

$$z_\alpha(x) := \int_0^\infty e^{-\alpha t} u(t,x)dt, \quad x \in \mathbf{R}^d, \tag{5.4.3}$$

and using (5.4.1), (5.4.2), integration by pats, and the assumption that

$$\lim_{t\o\infty} e^{-\alpha t} u(t,x) = 0, \quad \alpha > 0, x \in \mathbf{R}^d, \tag{5.4.4}$$

we may compute formally

$$\frac{1}{2}\Delta z_\alpha = \frac{1}{2}\int_0^\infty e^{-\alpha t}\Delta u dt = (\alpha + k)z_\alpha - f. \tag{5.4.5}$$

The stochastic representation for the solution z_α of the *elliptic equation* (5.4.5) is known as the Kac formula.

Definition 5.12. Consider the continuous functions $k : \mathbf{R}^d \to [0, \infty), g : [0, T] \times \mathbf{R}^d \to \mathbf{R}$, and $f : \mathbf{R}^d \to \mathbf{R}$. Suppose that $v \in C^{1,2}([0, T) \times \mathbf{R}^d)$, and is continuous on $[0, T] \times \mathbf{R}^d$, and satisfies

$$-\frac{\partial v}{\partial t} + kv = \frac{1}{2}\Delta v + g, \quad \text{on } [0, T) \times \mathbf{R}^d, \tag{5.4.6}$$

$$v(T, x) = f(x), \quad x \in \mathbf{R}^d. \tag{5.4.7}$$

Then the function is said to be a solution of the *Cauchy problem* for the backward heat equation (5.3.13) with potential k and Lagrangian g, subject to the terminal condition (5.4.7).

Theorem 5.13 (Feynman 1948, Kac 1949)**.** *Let v be as above, and assume that*

$$\max_{0 \le t \le T} |v(t, x)| + \max_{0 \le t \le T} |g(t, x)| \le K e^{a||x||^2}, \quad x \in \mathbf{R}^d \tag{5.4.8}$$

for constants $K > 0$ and $0 < a < 1/2Td$. Then v admits the stochastic representation

$$\begin{aligned} v(t, x) =& E^x \Big[f(W_{T-t} \exp \left(- \int_0^{T-t} k(W_s) ds \right) \\ & + \int_0^{T-t} g(t + \theta, W_\theta) \exp \left(- \int_0^{\theta} k(W_s) ds \right) \Big] \end{aligned} \tag{5.4.9}$$

for $0 \le t \le T, x \in \mathbf{R}^d$. In particular, such a solution is unique.

Proof. We obtain from Itô's rule, with (5.4.6)

$$d \left[v(t + \theta, W_\theta) \exp \left(- \int_0^{\theta} k(W_s) ds \right) \right] \tag{5.4.10}$$

$$= \exp \left(- \int_0^{\theta} k(W_s) ds \right) \left[-g(t + \theta, W_\theta) d\theta + \sum_{i=1}^{d} \frac{\partial}{\partial x_i} v(t + \theta, W_\theta) dW_\theta^{(i)} \right]. \tag{5.4.11}$$

Let $S_n = \inf\{t \ge 0 : ||W_t|| \ge n\sqrt{d}\}, n \ge 1$. Choose $0 < r < T - t$ and integrate on $[0, r \wedge S_n]$, the resulting stochastic integrals have expectation zero,, so

$$\begin{aligned} v(t, x) =& E^x \int_0^{r \wedge S_n} g(t + \theta, W_\theta) \exp \left(- \int_0^{\theta} k(W_s) ds \right) d\theta \\ & + E^x \left[v(t + S_n, W_{S_n}) \exp \left(- \int_0^{S_n} k(W_s) ds \right) 1_{\{S_n \le r\}} \right] \\ & + E^x \left[v(t + r, W_r) \exp \left(- \int_0^{r} k(W_s) ds \right) 1_{\{S_n \, r\}} \right] \end{aligned}$$

The first term on the right converges to

$$E^x \int_0^{T-t} g(t + \theta, W_\theta) \exp \left(- \int_0^{\theta} k(W_s) ds \right) d\theta \tag{5.4.12}$$

as $n \to \infty$ and $r \to T - t$, either by monotone convergence (if $g \ge 0$) or by dominated convergence (it is bounded in absolute value by $\int_0^{T-t} |g(t + \theta, W_\theta) d\theta$, which has finite expectation by (5.4.8)). The second term is dominated by

(5.4.13)

$$E^x[|v(t+S_n, W_{S_n})|1_{\{S_n\le T-t\}}] \le Ke^{adn^2}P^x[S_n \le T]$$

$$\le Ke^{adn^2}\sum_{j=1}^d P^x\Big[\max_{0\le t\le T}|W_t^{(j)}|\ge n\Big] \tag{5.4.14}$$

$$\le 2Ke^{adn^2}\sum_{j=1}^d P^x[W_T^{(j)}\ge n] + P^x[-W_T^{(j)}\ge n], \tag{5.4.15}$$

where we have used [KS,2.6.2], which tells us that for the *passage time* $T_b(\omega) = \inf\{t\ge 0; B_t(\omega) = b\}$, the reflection principle gives

$$P^0[T_b < t] = 2P^0[B_t > b] = \sqrt{\frac{2}{\pi}}\int_{b\sqrt{t}}^\infty e^{-x^2/2}dx. \tag{5.4.16}$$

But since for every $x > 0$, (prove this)

$$\frac{x}{1+x^2}e^{-x^2/2} \le \int_x^\infty e^{-u^2/2}du \le \frac{1}{x}e^{-x^2/2}, \tag{5.4.17}$$

it follows that

$$e^{adn^2}P^x[\pm W_T^{(j)}\ge n] \le e^{adn^2}\sqrt{\frac{T}{2\pi}}\frac{1}{n\mp x^{(j)}}e^{-(n\mp x^{(j)})^2/2T} \tag{5.4.18}$$

which converges to zero as $n\to\infty$, because $0 < a < 1/2Td$. Again by the dominated convergence theorem, the third term is seen to converge to

$$E^x\left[v(T, W_{T-t})\exp\left(-\int_0^{T-t}k(W_s)ds\right)\right] \tag{5.4.19}$$

as $n\to\infty, r\uparrow T-t$. The Feynman-Kac formula (5.4.9) follows. □

Corollary 5.14. *Assume that $f:\mathbf{R}^d\to\mathbf{R}, k:\mathbf{R}^d\to[0,\infty)$ and $g:[0,\infty)\times\mathbf{R}^d\to\mathbf{R}$ are continuous, and that the continuous function $u:[0,\infty)\times\mathbf{R}^d\to\mathbf{R}$ is $C^{1,2}$ on $(0,\infty)\times\mathbf{R}^d$ and satisfies (5.4.6), (5.4.7). If for each $T>0$ there exists constants $K>0$ and $0<a<1/2Td$ such that (5.4.8) holds for u in place of v, then u admits the stochastic representation*

$$\begin{aligned} u(t,x) = &E^x\Big[f(W_t\exp\left(-\int_0^t k(W_s)ds\right) \\ &+\int_0^t g(t+\theta, W_\theta)\exp\left(-\int_0^\theta k(W_s)ds\right)\Big] \end{aligned} \tag{5.4.20}$$

for $0\le t\le T, x\in\mathbf{R}^d$.

5.5. **SDEs and PDEs.** Consider the solution to the stochastic integral equation

$$X_s^{(t,x)} = x + \int_t^s b(\theta, X_\theta^{(t,x)})d\theta + \int_t^s \sigma(\theta, X_\theta^{(t,x)})dW_\theta, \quad t\le s<\infty \tag{5.5.1}$$

together with the assumptions that

(1) $b_i(t,x)$ and $\sigma_{ij}(t,x) : [0,\infty)\to\mathbf{R}^d\to\mathbf{R}$ are continuous and satisfy the linear growth condition (4.2.15)
(2) (5.5.1) has a weak solution $(X^{(t,x)}, W), (\Omega, \mathscr{F}, P), \{\mathscr{F}_s\}$ for every pair (t,x), and
(3) the solution is unique in the sense of probability law.

Associated with (5.5.1) is the second-order differential operator
(5.5.2)
$$(\mathscr{A}_t f)(x) := \frac{1}{2}\sum_{i,k}^{d} a_{ik}(t,x)\frac{\partial^2 f(x)}{\partial x_i \partial x_k} + \sum_{i=1}^{d} b_i(t,x)\frac{\partial f(x)}{\partial x_i}, \quad f \in C^2(\mathbf{R}^d), t \geq 0,$$

where $a_{ik}(t,x)$ are the components of the diffusion matrix. If f is a function of $t \in [0,\infty)$ and $x \in \mathbf{R}^d$, then $(\mathscr{A}_t)(t,x)$ is obtained by applying $\mathscr{A}_t$ to $f(t,\cdot)$.

Proposition 5.15. *Let $f(t,x) : [0,\infty) \times \mathbf{R}^d \to \mathbf{R}$ belong to $C^{1,2}((0,\infty)\times\mathbf{R}^d)$. Then the process $M^f = M^f_t, \mathscr{F}_t; 0 \leq t < \infty\}$ given by*

$$M^f_t = f(t,X_t) - f(0,X_0) - \int_0^t \Big(\frac{\partial f}{\partial s} + \mathscr{A}_s f\Big)(s,X_s)ds \tag{5.5.3}$$

is a continuous local martingale. If g another such function, then

$$\langle M^f, M^g\rangle_t = \sum_{i,k=1}^{d}\int_0^t a_{ik}(s,X_s)\frac{\partial}{\partial x_i}f(s,X_s)\frac{\partial}{\partial x_k}g(s,X_s)ds. \tag{5.5.4}$$

Furthermore, if $f \in C_0([0,\infty)\times\mathbf{R}^d)$ and the coefficients $\sigma_{ij}, 1 \leq i \leq d, 1 \leq j \leq r$ are of bounded on the support of f, then $M^f \in \mathscr{M}_2^c$.

Proof. Itô's rule expresses M^f as a sum of stochastic integrals

$$M^f_t = \sum_{i=1}^{d}\sum_{j=1}^{r} M^{(i,j)}_t, \quad M^{(i,j)}_t := \int_0^t \sigma_{ij}(s,X_s)\frac{\partial}{\partial x_i}f(s,X_s)dW^{(i)}_s. \tag{5.5.5}$$

Introducing the stopping times

$$S_n = \inf\{t \geq 0 : ||X_t|| \geq n \text{ or } \int_0^t \sigma^2_{ij}(s,X_s) \geq n \text{ for some } i,j\}, \tag{5.5.6}$$

and recalling that a weak solution must satisfy condition (iii) of Definition 4.2, we see that $\lim_{n\to\infty} S_n = \infty$ a.s. The processes $M^f_t(n) := M^f_{t\wedge S_n}$ are continuous martingales, so $M^f \in \mathscr{M}^{c,\mathrm{loc}}$. The cross variation formula holds readily from $M^f_t(n)$. If f has compact support on which each σ_{ij} is bounded, then the integrand in the expression for $M^{(ij)}$ is bounded, and so $M^f \in \mathscr{M}_2^c$. □

Example 5.16. The simplest case is that of a d-dimensional Brownian motion, which corresponds to $b_i(t,x) \equiv 0$ and $\sigma_{ij}(t,x) \equiv \delta_{ij}, 1 \leq i,j \leq d$, then we have

$$\mathscr{A} f = \frac{1}{2}\Delta f = \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2 f}{\partial x_i^2}, \quad f \in C^2(\mathbf{R}^d). \tag{5.5.7}$$

Exercise 12. Show that a continuous adapted process $W = \{W_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a d-dimensional Brownian motion if and only if

$$f(W_t) - f(W_0) - \frac{1}{2}\int_0^t \Delta f(W_s)ds, \tag{5.5.8}$$

is in $\mathscr{M}^{c,\mathrm{loc}}$ for every $f \in C^2(\mathbf{R}^d)$. This provides a martingale characterisation of Brownian motion.

5.5.1. *The Dirichlet problem.* Let D be an open subset of $\mathbf{R}^d$, and assume that b, σ are independent of t.

Definition 5.17. The operator $\mathscr{A}$ is called elliptic at the point $x \in \mathbf{R}^d$ if

$$\sum_{i,k=1}^{d} a_{ik}(x)\xi_i\xi_k > 0, \quad \forall \xi \in \mathbf{R}^d\backslash\{0\}. \tag{5.5.9}$$

If $\mathscr{A}$ is elliptic at every point of D, we say that $\mathscr{A}$ is elliptic in D. If there exists a number $\delta > 0$ such that

$$\sum_{i,k=1}^{d} a_{ik}(x)\xi_i\xi_k \geq \delta||\xi||^2, \quad \forall x \in D, \xi \in \mathbf{R}^d\backslash\{0\}, \tag{5.5.10}$$

then we say that $\mathscr{A}$ is uniformly elliptic in D.

Let $\mathscr{A}$ be elliptic in the open, bounded domain D, and consider continuous functions $k : \overline{D} \to [0,\infty), g : \overline{D} \to \mathbf{R}$, and $f : \partial D \to \mathbf{R}$. The Dirichlet problem is then to find a continuous function $u : \overline{D} \to \mathbf{R}$ such that u is $C^{1,2}(D)$ and satisfies the elliptic equation

$$\mathscr{A}u - ku = -g, \text{ in } D \tag{5.5.11}$$

and the boundary condition

$$u = f, \text{ on } \partial D. \tag{5.5.12}$$

Proposition 5.18. *Let u be a solution of the Dirichlet problem above, in the bounded open domain D, and let $\tau_D := \inf\{t \geq 0 : X_t \notin D\}$. If for all $x \in D$,*

$$E^x\tau_D < \infty, \tag{5.5.13}$$

then under the assumptions (1)–(3) of (5.5.1), we have

(5.5.14)

$$u(x) = E^x\left[f(X_{\tau_D})\exp\left(-\int_0^{\tau_D} k(X_s)ds\right) + \int_0^{\tau_D} g(X_t)\exp\left(-\int_0^t k(X_s)ds\right)dt\right]$$

for every $x \in \overline{D}$.

Proof. We show that for $t \geq 0$,

(5.5.15)

$$M_t := u(X_{t\wedge\tau_D})\exp\Big(-\int_0^{t\wedge\tau_D} k(X_s)ds\Big) + \int_0^{t\wedge\tau_D} g(X_s)\exp\Big(-\int_0^s k(X_\theta)d\theta\Big)ds$$

is a uniformly martingale under P^x. To that end, consider an increasing sequence $\{D_n\}_{n=1}^\infty$ of open sets with $\overline{D}_n \subset D$ for all $n \geq 1$, and $\cup_{n=1}^\infty D_n = D$, so that the stopping times $\tau_D = \inf\{t \geq 0 : X_t \notin D_n\}$ satisfy $\lim_{n\to\infty}\tau_n = \tau_D$ a.s. P^x. Using Itô's formula, argue that

$$M_t^{(n)} := u(X_{t\wedge\tau_n})\exp\Big(-\int_0^{t\wedge\tau_n} k(X_s)ds\Big) + \int_0^{t\wedge\tau_n} g(X_s)\exp\Big(-\int_0^s k(X_\theta)d\theta\Big)ds$$

is a P^x-martingale for every $n \geq 1, x \in D$.

Also, $|M_t(\omega)|$ and $|M_t^{(n)}(\omega)$ are bounded above by

$$\max_{x\in\overline{D}}|u(x)| + (t \wedge \tau_D(\omega))\max_{x\in\overline{D}}|g(x)|$$

for P^x-a.e. $\omega \in \Omega$. Then letting $n \to \infty$ and using the bounded convergence theorem, it follows that the process $M = \{M_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a martingale, and moreover by (5.5.13), it is uniformly integrable.

Then $M_\infty = \lim_{t\to\infty} M_t$ is martingale (see [KS, 1.3.20]), and the identity $E^xM_0 = E^xM_\infty$ then gives the representation. □

5.5.2. *The Cauchy problem and a Feynman-Kac representation.* Fix $T > 0$ and appropriate constants $L > 0, \lambda \geq 1$. Consider $f : \mathbf{R}^d \to \mathbf{R}, g : [0,T] \times \mathbf{R}^d \to \mathbf{R}$, and $k : [0,T] \times \mathbf{R}^d \to [0,\infty)$ which are continuous and satisfy

$$\text{(i) } |f(x)| \leq L(1 + ||x||^{2\lambda}) \text{ or (ii) } f(x) \geq 0, \quad x \in \mathbf{R}^d \tag{5.5.16}$$

$$\text{(i) } |g(t,x)| \leq L(1 + ||x||^{2\lambda}) \text{ or (ii) } g(t,x) \geq 0, \quad 0 \leq t \leq T, x \in \mathbf{R}^d. \tag{5.5.17}$$

We now formulate the analogue of the Feynman-Kac theorem 5.13.

Theorem 5.19. *Under the preceding assumptions, and those for (5.5.1), suppose thatv* $: [0,T] \times \mathbf{R}^d \to \mathbf{R}^d$ *is continuous, of class* $C^{1,2}([0,T] \times \mathbf{R}^d)$*, and satisfied the Cauchy problem*

$$-\frac{\partial v}{\partial t} + kv = \mathscr{A}_t v + g, \text{ in } [0,T) \times \mathbf{R}^d, \tag{5.5.18}$$

$$v(T,x) = f(x), \quad x \in \mathbf{R}^d, \tag{5.5.19}$$

as well as the polynomial growth condition

$$\max_{0 \le t \le T} |v(t,x)| \le M(1 + ||x||^{2\mu}), \quad x \in \mathbf{R}^d, \tag{5.5.20}$$

for some $M > 0, \mu \ge 1$*. Then* $v(t,x)$ *admits the stochastic representation*

$$v(t,x) = E^{(t,x)}\Big[f(X_T)\exp\left(-\int_t^T k(\theta, X_\theta)d\theta\right) \tag{5.5.21}$$

$$+ \int_t^T g(s, X_s)\exp\left(-\int_t^s k(\theta, X_\theta)d\theta\right) ds\Big] \tag{5.5.22}$$

on $[0,T] \times \mathbf{R}^d$*. In particular, such a solution is unique.*

Proof. Proceeding as in the proof of Theorem 5.13, we apply the Itô formula to the process $v(s, X_s)\exp(-\int_t^s l(\theta, X_\theta)d\theta), s \in [t,T]$, and obtain with $\tau_n := \inf\{s \ge t : ||X_s|| \ge n\}$,

$$\begin{aligned} v(t,x) =& E^{(t,x)}\left[\int_t^{T\wedge\tau_n} g(s, X_s)\exp\left(-\int_t^s k(\theta, X_\theta)d\theta\right) ds\right] \\ &+ E^{(t,x)}\left[v(\tau_n, X_{\tau_n})\exp\left(-\int_t^{\tau_n} k(\theta, X_\theta)d\theta\right) 1_{\{\tau_n \le T\}}\right] \\ &+ E^{(t,x)}\left[f(X_T)\exp\left(-\int_t^T k(\theta, X_\theta)d\theta\right) 1_{\{\tau_n\ T\}}\right] \end{aligned}$$

We shall use the estimate (see [KS,(5.3.17)]

$$E^{(t,x)}\left[\max_{t \le \theta \le s} ||X_\theta||^{2m}\right] \le C(1 + ||x||^{2m})e^{C(s-t)}, \quad t \le s \le T \tag{5.5.23}$$

which is valid for every $m \ge 1$ and some $C = C(m, K, T, d) > 0$. Now the first term on the right of (**??**) converges as $n \to \infty$ to

$$E^{(t,x)}\left[\int_t^T g(s, X_s)\exp\left(-\int_t^s k(\theta, X_\theta)d\theta\right) ds\right] \tag{5.5.24}$$

either by the dominated convergence theorem by (5.5.17)(i) and (5.5.23), or by the monotone convergence if (5.5.17)(ii) prevails. The second term is bounded in absolute value by

$$E^{(t,x)}[|v(\tau_n, X_{\tau_n})|1_{\{\tau_n \le T\}}] \le M(1 + n^{2\mu})P^{(t,x)}[\tau_n \le T], \tag{5.5.25}$$

and this last probability can be written using Markov's inequality as
(5.5.26)

$$P^{(t,x)}\left[\max_{t \le \theta \le T} ||X_\theta||| \ge n\right] \le n^{-2m}P^{(t,x)}\left[\max_{t \le \theta \le T} ||X_\theta|||\right] \le Cn^{-2m}(1 + ||x||^{2m})e^{CT},$$

by virtue of (5.5.23) and Chebyshev's inequality. Choosing $m > \mu$, we see that the right-hand side of (5.5.25) converges to 0 as $n \to \infty$. Finally, the last term in (**??**) converges to

$$E^{(t,x)}\left[f(X_T)\exp\left(-\int_t^T k(\theta, X_\theta)d\theta\right)\right] \tag{5.5.27}$$

either by the dominated convergence theorem or the monotone convergence theorem. □

5.6. **Portfolio and Consumption Processes.** Let us consider a market in which $d+1$ assets (or 'securities') are traded continuously. We assume throughout this setion that there is a fixed time horizon $0 \le T < \infty$. One of the assets, called the *bond*, has a price $P_0(t)$ which evolves according to the differential equation

(5.6.1) $$dP_0(t) = r(t)P_0(t)dt, \quad P_0(0) = p_0, 0 \le t \le T.$$

The remaining d assets, called *stocks*, are 'risky'; their prices are modelled by the linear stochastic differential equation for $i = 1, \dots, d$,

(5.6.2)
$$dP_i(t) = b_i(t)P_i(t)dt + P_i(t)\sum +j = 1^d\sigma_{ij}(t)dW^{(j)}, \quad P_i(0) = p_i, 0 \le t \le T.$$

The process $W = \{W_t = (W^{(1)}, \dots, w^{(d)}, \mathscr{F}_t; 0 \le t \le T\}$ is a d-dimensional Brownian motion on a probability space $(\Omega, \mathscr{F}, P)$, and the filtration $\{\mathscr{F}_t\}$ is the augementation under P of the filtration $\{\mathscr{F}_t^W\}$ generated by W. The *interest rate* process $\{r(t), \mathscr{F}_t; 0 \le t \le T\}$, as well as the vector of *mean rates of return* $\{b(t) = (b_1, \dots, b_d(t))^T, \mathscr{F}_t; 0 \le t \le T\}$ and the *dispersion matrix* $\{\sigma(t) = \{(\sigma_{ij}(t))_{1 \le i,j \le d}, \mathscr{F}_t, 0 \le t \le T\}$ are assumed to be measurable, adpated, and bounded uniformly in $(t, \omega) \in [0,T] \times \Omega$. We set $a(t) := \sigma(t)\sigma^T(t)$ and assume that for some number $\epsilon > 0$,

(5.6.3) $$\xi^T a(t)\xi \ge \epsilon||\xi||^2, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s..}$$

Exercise 13. Under the (5.6.3), $\sigma^T(t)$ has an inverse, and

(5.6.4) $$||(\sigma^T(t))^{-1}\xi|| \le \epsilon - 1/2||\xi||, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s..}$$

Moreover, with $\hat{a}(t) := \sigma^T(t)\sigma(t)$, we have

(5.6.5) $$\xi^T \hat{a}(t)\xi \ge \epsilon||\xi||^2,, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s.,}$$

so $\sigma(t)$ also has an inverse and

(5.6.6) $$||(\sigma(t))^{-1}\xi|| \le \epsilon^{-1/2}||\xi||, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s..}$$

We imagine now an investor who starts with some initial endowment $x \ge 0$ and invests it in the $d+1$ assets described previously. Let $N_i(t)$ denote the number of shares of asset i owned by the investor at time t. Then $X_0 \equiv x = \sum_{i=0}^d N_i(0)p_i$, and the investor's *wealth* at time t is

(5.6.7) $$X_t = \sum_{i=0}^d N_i(t)P_i(t).$$

If trading of shares takes place at discrete time points, say at t and $t+h$, and there is no infusion or withdrawal of funds, then

(5.6.8) $$X_{t+h} - X_t = \sum_{i=0}^d N_i(t)[P_i(t+h) - P_i(t)].$$

If, on the other hand, the investor chooses at time $t+h$ to consime an amount hC_{t+h}, and reduce the wealth accordingly, then the last equation should be replaced by

(5.6.9) $$X_{t+h} - X_t = \sum_{i=0}^d N_i(t)[P_i(t+h) - P_i(t)] - hC_{t+h}.$$

The continuous time analogue of this is

(5.6.10) $$dX_t = \sum_{i=0} N_i(t)dP_t(t) - C_t dt.$$

Taking (5.6.1), (5.6.2), and (5.6.7) into account, and denoting by $\pi_i(t) := \mathbf{N_i(t)P_i(t)}$ the amount invested in the i-th stock, we may write this as

$$dX_t = (r(t)X_t - C_t)dt + \sum_{i=1}^{d}(b_i(t) - r(t))\pi_i(t)dt + \sum_{i,j=1}^{d} \pi_i(t)\sigma_{ij}(t)dW_t^{(j)}. \tag{5.6.11}$$

Definition 5.20. A *portfolio process* $\pi = \{\pi(t) = (\pi_1(t), \ldots, \pi_d(t))^T, \mathscr{F}_t; 0 \le t \le T\}$ is a measurable, adapted process for which

$$\sum_{i=1}^{d} \int_0^T \pi_i^2(t)dt < \infty, \text{a.s.} \tag{5.6.12}$$

A *consumption process* $C = \{C_t, \mathscr{F}_t; 0 \le t \le T\}$ is a measurable, adapted process with values in $[0, \infty)$ and

$$\int_0^T C_t dt < \infty, \text{a.s.} \tag{5.6.13}$$

Remark 5.21. Note that any component of $\pi(t)$ may become negative, which is to be interpreted as short-selling a stock. The amount invested in the bond,

$$\pi_0(t) := X_t - \sum_{i=1}^{d} \pi_t(t) \tag{5.6.14}$$

may also be negative, and this amounts to borrowing at the interest rate $r(t)$.

The conditions (5.6.12) and (5.6.13) ensure that the stochastic differential equation (5.6.11) has a unique strong solution. Indeed, formal applications of [KS, Problem 5.6.15] leads to the formula

$$\begin{aligned} X_t = e^{\int_0^t r(s)ds}\Big[x &+ \int_0^t e^{-\int_0^s r(u)du}\{\pi(s)^T(b(s) - r(s)\vec{1}) - C_s\}ds \\ &+ \int_0^t e^{-\int_0^s r(u)du}\pi^T(s)\sigma(s)dW_s\Big], \quad 0 \le t \le T. \end{aligned} \tag{5.6.15}$$

Definition 5.22. A pair (π, C) or portfolio and consumption processes is said to be *admissible for initial endowment* $x \ge 0$ if the wealth process (5.6.15) satisfies $X_t \ge 0, 0 \le t \le T$ a.s.

If $b(t) = r(t)\vec{1}$ for $0 \le t \le T$ then the discount factor $e^{-\int_0^t r(s)ds}$ exactly offsets the rate of growth of all assets and (5.6.15) shows that

$$M_t := X_t e^{-\int_0^t r(s)ds} - x + \int_0^t e^{-\int_0^s r(u)du} C_s ds \tag{5.6.16}$$

is a stochastic integral. In other words, the process consisting of current wealth plus cumulative consumption, both properly discounted, is a local martingale. When $b(t) \ne r(t)\vec{1}$, then M_t is no longer a local martingale under P, but becomes one under a new measure $\tilde{P}$ which removes the drift term $\pi(t)^T(b(t) - r(t)\vec{1})$ in (5.6.11). More specifically, from Exercise 13 we know that the process

$$\theta(t) := ((\sigma(t))^{-1}(b(t) - r(t)\vec{1}) \tag{5.6.17}$$

is bounded, and set

$$Z_t = \exp\left[-\sum_{i=1}^{d} \int_0^t \theta_j(s)dW_s^{(j)} - \frac{1}{2}\int_0^t ||\theta(s)||^2 ds\right]. \tag{5.6.18}$$

Then $Z = \{Z_t, \mathscr{F}_t; 0 \le t \le T\}$ is a martingale by the Novikov condition, Theorem 3.25. The new probability measure

$$\tilde{P}(A) := E[Z_T 1_A], \quad A \in \mathscr{F}_T \tag{5.6.19}$$

is such that P and $\tilde{P}$ are mutually absolutely continuous on $\mathscr{F}_T$, and the process

$$\tilde{W}_t := W_t + \int_0^t \theta(s)ds, \quad 0 \le t \le T \tag{5.6.20}$$

is a d-dimensional Brownian motion under $\tilde{P}$ by the Girsanov Theorem 3.20. In terms of $\tilde{W}$, (5.6.11) may be written as

$$X_t e^{-\int_0^t r(s)ds} + \int_0^t e^{-\int_0^s r(u)du} C_s ds = x + \int_0^t e^{-\int_0^s r(u)du} \pi^T(s)\sigma(s)d\tilde{W}_s \tag{5.6.21}$$

for $0 \le t \le T$ a.s.

For an admissible pair (π, C), the left-hand side of (8c) is nonnegative and the right-hand side is a $\tilde{P}$-local martingale. It follows that the left-hand side, and hence also $X_t e^{-\int_0^t r(s)ds}$ is a nonnegative supermartingale under $\tilde{P}$ by [KS, Problem 1.5.19]. Let

$$\tau_0 = T \wedge \inf\{0 \le t \le T; X_t = 0\}. \tag{5.6.22}$$

According to [KS, Problem 1.3.29],

$$X_t = 0, \tau_0 \le t \le T \tag{5.6.23}$$

holds a.s. on $\{\tau_0 < T\}$. If $\tau_0 < T$, we say that *bankruptcy* occurs at time τ_0.

From the supermartingale property in (8c), we obtain

$$\tilde{E}\left[X_T e^{-\int_0^T r(s)ds} + \int_0^T e^{-\int_0^s r(u)du} C_s ds\right] \le x, \tag{5.6.24}$$

whence the following *necessary condition for admissibility*,

$$\tilde{E}\int_0^T e^{-\int_0^s r(u)du} C_s ds \le x. \tag{5.6.25}$$

This condition is also sufficient for in the sense of the following proposition.

Proposition 5.23. *Suppose $x \ge 0$ and a consumption process C are given so that (5.6.25) is satisfied. Then there exists a portfolio process π such that the pair (π, C) is admissible for the endowment x.*

Proof. Let $D := \int_0^T C_t e^{-\int_0^t r(s)ds} dt$, and define the nonnegative process

$$\xi_t := \tilde{E}\left[\int_t^T C_s e^{-\int_t^s r(u)du} \Big| \mathscr{F}_t\right] + (x - \tilde{E}D)e^{\int_0^t r(s)ds}, \tag{5.6.26}$$

so that

$$\xi_t = e^{\int_0^t r(s)ds}\left\{x + m_t - \int_0^t C_s e^{-\int_0^s r(u)du} ds\right\}, \tag{5.6.27}$$

where

$$m_t := \tilde{E}[D|\mathscr{F}_t] - \tilde{E}D = \frac{E[DZ_T|\mathscr{F}_t]}{Z_t} - E(DZ_T) \tag{5.6.28}$$

from Bayes rule of Lemma 3.22. From [KS, Theorem 1.3.13], we may assume that the P-a.e. path of the martingale

$$N_t := E(DZ_T|\mathscr{F}_t), \quad 0 \le t \le T, \tag{5.6.29}$$

is right continuous with left limits (RCLL), so by [KS, Problem 3.4.16] there exists a measurable $\{\mathscr{F}_t\}$-adapted $\mathbf{R}^d$-valued process Y with

$$\int_0^T ||Y(t)||^2 dt < \infty, \tag{5.6.30}$$

and

$$N_t = E(DZ_T) + \sum_{j=1}^{d} \int_0^t Y_j(s) dW_s^{(j)}, \quad 0 \le t \le T, \tag{5.6.31}$$

valid P-a.s. Now $m_t = u(N_t, Z_t) - E(DZ_T)$, where $u(x, y) = (x/y)$, and from Itô's rule we obtain with $\varphi(t) := (Y(t) + N_t\theta(t))/Z_t$,

$$m_t = \sum_{j=1}^{d} \int_0^t \varphi_j(s) d\tilde{W}_s^{(j)}, \quad 0 \le t \le T. \tag{5.6.32}$$

we have used the relations $dZ_t = -Z_t\theta^T(t)dW_t$ and (5.6.20). Now define

$$\pi(t) := e^{\int_0^t r(s)ds}(\sigma^T(t))^{-1}\varphi(t), \tag{5.6.33}$$

so that ξ_t in (5.6.26) becomes (8c) when we make the identification $\xi = X$. Condition (5.6.12) follows from (5.6.4) and (5.6.30), the boundedness of θ, and the path continuity of Z and N, the latter being a consequence of (5.6.31). □

Remark 5.24. The representation (5.6.31) cannot be obtained from a direct application of [KS, Problem 3.4.16] to the $\tilde{P}$ martingale $\{m_t, \mathscr{F}_t; 0 \le t \le T\}$. The reason is that the filtration $\{\mathscr{F}_t\}$ is the augmentation (under P or $\tilde{P}$) of $\{\mathscr{F}_t^W\}$, *not* of $\{\mathscr{F}^{\tilde{W}}\}$.

5.7. **Option pricing and the Black-Scholes.** In the context of the previous subsection, suppose that at time $t = 0$ we sign a contract which gives us the option to buy, at a specified time T (called *maturity* or *expiration date*), one share of stock 1 at a specified price q, called the *exercise price*. At maturity, if the price $P_T^{(1)}$ of stock 1 is below the exercise price, the contract is worthless to us; on the other hand, if $P_T^{(1)} > q$, we can exercise our option (i.e., buy one share at the preassigned price q) and then sell the share immediately in the market for $P_T^{(1)}$. This contract, which is called an option, is thus equivalent to a payment of $(P_T^{(1)} - q)^+ := \max(0, P_T^{(1)} - q)$ dollars at maturity. Sometimes the term *European option* is used to describe this financial instrument, in contrast to an *American option*, which can be exercised at any time between $t = 0$ and maturity.

The following definition provides a generalization of the concept of option.

Definition 5.25. A *contingent claim* is a financial instrument consisting of

(1) a *payoff rate* $g = \{g_t, \mathscr{F}_t; 0 \le t \le T\}$, and
(2) a *terminal payoff* f_T at maturity.

Here g is a nonnegative, measurable, and adapted process, and f_T is a nonnegative, $\mathscr{F}_T$-measurable random variable, and for some $\mu > 1$, we have

$$E\left[f_T + \int_0^T g_t dt\right]^{\mu} < \infty. \tag{5.7.1}$$

Remark 5.26. An option is a special case of a contingent claim with $g \equiv 0$ and $f_T = (P_T^{(1)} - q)^+$.

Definition 5.27. Let $x \ge 0$ be given, and let (π, C) be a portfolio/consumption process pair which is admissible for the initial endowment x. The pair (π, C) is called a *hedging strategy against the contingent claim* (g, f_T), provided

(1) $C_t = g_t, 0 \le t \le T$, and
(2) $X_T = f_T$.

holds a.s., where X is the wealth process associated with the pair (π, C) and with the initial condition $X_0 = x$.

The concept of hedging strategy is introduced in order to allow the solution of the *contingent claim valuation problem*: What is a fair price to pay at time $t = 0$ for a contingent claim? If there exists a hedging strategy which is admissible for an initial endowment $X_0 = x$, then an agent who buys at time $t = 0$ the contingent claim (g, f_T) for the price x could instead have invested the wealth in such a way as to duplicate the payoff of the contingent claim. Consequently, the price of the claim should not be greater than x. Could one begin with an initial wealth strictly smaller than x and again duplicate the payoff ofthe contingent claim? The answer to this question may be affirmative, as shown by the following exercise.

Exercise 14. Consider the case $r \equiv 0, d = 1, b_1 \equiv 0$, and $\sigma \equiv 1$. Let the contingent claim $g \equiv 0, f_T \equiv 0$ be given, so obviously there exists a hedging strategy $x = 0, C \equiv 0$, and $\pi \equiv 0$. Show that for each $x > 0$, there is a hedging strategy with $X_0 = x$.

The *fair price* for a contingent claim is the smallest number $x \geq 0$ which allows the construction of a hedging strategy with initial wealth x. We shall show that under condition (5.6.3) and the assumptions preceding it, every contingent claim has a fair price; we shall also derive the explicit Black-Scholes formula for the fair price of an option.

Lemma 5.28. *Let the contingent claim* (g, f_T) *be given, and define*

$$Q = e^{-\int_0^T r(u)du} f_T + \int_0^T e^{-\int_0^s r(u)du} g_s ds. \tag{5.7.2}$$

Then $\tilde{E}Q$ *is finite and is a lower bound on the fair price of* (g, f_T).

Proof. Recalling that r is uniformly bounded in t and ω, we may write

$$Q \leq L\Big(f_T + \int_0^T g_s ds\Big), \tag{5.7.3}$$

where L is some nonrandom constant. From (5.6.18), we have for every $\nu \geq 1$,

$$Z_T^\nu = \exp\left(-\sum_{i=1}^d \int_0^T \nu\theta_j(s)dW_s^{(j)} - \frac{1}{2}\int_0^T ||\nu\theta(s)||^2 ds\right) \tag{5.7.4}$$

$$\times \exp\left(\frac{\nu(\nu-1)}{2}\int_0^T ||\theta(s)||^2 ds\right), \tag{5.7.5}$$

and because $||\theta||$ is bounded by some constant K, it follows that

$$EZ_T^\nu \leq \exp\left(\frac{\nu(\nu-1)}{2}K^2 T\right). \tag{5.7.6}$$

With μ as in (5.7.1), and ν given by $(1/\nu)+(1/\mu) = 1$, the Hölder inequality implies that

$$\tilde{E}Q \leq LE\left[Z_T\Big(f_T + \int_0^T g_s ds\Big)\right] \tag{5.7.7}$$

$$\leq L(EZ_T^\nu)^{1/\nu}\left[E\Big(f_T + \int_0^T g_s ds\Big)^\mu\right]^{1/\mu} < \infty. \tag{5.7.8}$$

Now suppose that (π, C) is a hedging strategy against the contingent claim (g, f_T), and the corresponding wealth process is X with initial condition $X_0 = x$. Recalling Definition 5.27 and (5.7.2), we may rewrite (5.6.24) as $\tilde{E}Q \leq x$. □

Theorem 5.29. *Under condition* (5.6.3) *and the assumptions preceding it, the fair price of a contingent claim* (g, f_T) *is* $\tilde{E}Q$. *Moreover, there exists a hedging strategy with initial wealth* $x = \tilde{E}Q$.

Proof. Define

$$\xi_T := e^{\int_0^t r(s)ds}\left[\tilde{E}Q + m_t - \int_0^t e^{-\int_0^s r(u)du} g_s ds\right], \tag{5.7.9}$$

where $m_t = \tilde{E}[Q|\mathscr{F}_t] = \tilde{E}Q$. Proceeding exactly in the proof of Proposition 5.23 with D replaced by Q, we define π by (5.6.33) and $C \equiv g$, so that (5.7.9) becomes (8c) with the identifications $X = \xi, x = \tilde{E}Q$. But then (5.7.9) can also be written as

$$X_t = \tilde{E}\left[e^{-\int_0^T r(u)du} f_T + \int_0^T e^{-\int_s^t r(u)du} g_s ds\right], \quad 0 \le t \le T, \tag{5.7.10}$$

whence $X_t \ge 0, 0 \le t \le T$, and $X_T = f_T$ are seen to hold almost surely. □

Exercise 15. Show that the hedging strategy constructed in the proof of Theorem 5.29 is essentially (in the sense of meas $\times$ P-a.e. equivalence) the only hedging strategy corresponding to initial wealth $x = \tilde{E}Q$. In particular, the process X of (5.7.10) gives the unique wealth process corresponding to the fair price; it is called the *valuation process* of the contingent claim.

Example 5.30 (Black-Scholes option valuation formula)**.** In the setting of Remark 5.26 with $d = 1$ and constant coefficients $r(t) \equiv r > 0$, $\sigma_{11}(t) \equiv \sigma > 0$, the price of the bond is

$$P_0(t) = p_0 e^{rt}, \quad 0 \le t \le T, \tag{5.7.11}$$

and the price of the stock obeys

$$dP_1(t) = b_1(t)P_1(t)dt + \sigma P_1(t)dW_1(t) = (t)P_1(t)dt + \sigma P_1(t)d\tilde{W}_1(t). \tag{5.7.12}$$

For the option to buy one share of stock at time T and price q, we have from (5.7.10) the valuation process

$$X_t = \tilde{E}[e^{-r(T-t)}(P_1(T) - q)^+|\mathscr{F}_t], \quad 0 \le t \le T. \tag{5.7.13}$$

In order to write X_t in a more explicit form, let us observe that the function

$$v(t,x) := \begin{cases} x\Phi(\rho_+(T-t,x)) - qe^{-r(T-t)}\Phi(\rho_-(T-t,x)), & 0 \le t < T, x \ge 0 \\ (x-q)^+; & t = T, x \ge 0, \end{cases} \tag{5.7.14}$$

with

$$\rho_\pm(t,x) = \frac{1}{\sigma\sqrt{x}}\left[\log\frac{x}{q} + t\Big(r \pm \frac{\sigma^2}{2}\Big)\right], \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-x^2/2}dx \tag{5.7.15}$$

satisfies the Cauchy problem

$$-\frac{\partial v}{\partial t} + rv = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 v}{\partial x^2} + rx\frac{\partial v}{\partial x}, \text{ on } [0,T) \times (0,\infty) \tag{5.7.16}$$

$$v(T,x) = (x-q)^+, \quad x \ge 0, \tag{5.7.17}$$

as well as the conditions of Theorem 5.19. We conclude from that theorem and the Markov property applied to X_t that

$$X_t = v(t, P_1(t)), \quad 0 \le t \le T, \text{ a.s.} \tag{5.7.18}$$

We thus have an explicit formula for the value of the option at time t in terms of the current stock price $P_1(t)$, the time to maturity $T - t$, and the exercise price q.

Exercise 16. In the setting of the example above, but with $f_T = h(P_1(T))$ where $h : [0,\infty) \to [0,\infty)$ is a convex piecewise C^2 function with $h(0) = h'(0) = 0$, show that the valuation process for the contingent claim $(0, f_T)$ is given by

$$X_t = \tilde{E}[e^{-r(T-t)}h(P_t(T))|\mathscr{F}_t] = \int_0^\infty h''(q)v_{q,T}(t, P_1(t))dq. \tag{5.7.19}$$

We denote here by $v_{q,t}(t,x)$ the function of (5.7.14).

5.8. **Optimal consumption and investment.** In this section we pose and solve a stochastic *optimal control problem* for the economics model in Section 5.6. Suppose that, in addition to the data given there, we have a measurable, adapted, uniformly bounded *discount process* $\beta = \{\beta(s), \mathscr{F}_s; 0 \le s \le T\}$ and a strictly increasing, strictly concave, continuously differentiable *utility function* $U : [0,\infty) \to [0,\infty)$ for which $U(0) = 0$ and $U'(\infty) := \lim_{c\to\infty} U'(c) = 0$. We allow the possibility that $U'(0) := \lim_{c\downarrow 0} U'(c) = \infty$. Given an initial endowment $x \ge 0$, an investor wishes to choose an admissible pair (π, C) of portfolio and consumption processes, so as to maximise

$$V_{\pi,C}(x) := E\int_0^T e^{-\int_0^s \beta(u)du} U(C_s)ds. \tag{5.8.1}$$

We define the *value function* for this problem to be

$$V(x) = \sup_{(\pi,C)} V_{\pi,C}(x), \tag{5.8.2}$$

where the supremum is over all pairs (π, C) admissible for x. From the necessary condition for admissibility (5.6.25), it is clear that $V(0) = 0$.

Recall from Proposition 5.23 that for a given consumption process C, (5.6.25) is satisfied if and only if there exists a portfolio π such that (π, C) is admissible for x. Let us define $\mathscr{D}(x)$ be the class of consumption processes C for which

$$\tilde{E}\int_0^T e^{-\int_0^t r(s)ds} C_t dt = x. \tag{5.8.3}$$

It turns out that in the maximisation indicated in (5.8.2) we may ignore the portfolio process π, and we need only consider $C \in \mathscr{D}(x)$.

Proposition 5.31. *For every $x \ge 0$, we have*

$$V(x) = \sup_{C\in\mathscr{D}(x)} E\int_0^T e^{-\int_0^s \beta(u)du} U(C_s)ds. \tag{5.8.4}$$

Proof. Suppose that (π, C) is admissible for $x > 0$, ad set

$$y := \tilde{E}\int_0^T e^{-\int_0^s \beta(u)du} U(C_s)ds. \le x. \tag{5.8.5}$$

If $y > 0$, we may defined $\hat{C}_t := (x/y)C_t$ so that $\hat{C} \in \mathscr{D}(x)$. There exists then a portfolio process $\hat{pi}$ such that $(\hat{\pi}, \hat{C})$ is admissible for x, and

$$V_{\pi,C}(x) \le V_{\hat{\pi},\hat{C}}(x). \tag{5.8.6}$$

If $y = 0$, then $C_t = 0$ a.e., $t \in [0,T]$ almost surely and we can find a constant $c > 0$ such that $\hat{C} \equiv c$ satisfies (5.8.3). Again, (5.8.6) holds for some $\hat{\pi}$ chosen so that $(\hat{\pi}, \hat{C})$ is admissible for x. □

Since $U' : [0,\infty] \to [0, U'(0)]$ is surjective and strictly decreasing, it has a strictly deceasing and surjective inverse function $I : [0, U'(0)] \to [0,\infty]$. We extend I by setting $I(y) = 0$ for $y > U'(0)$. Note that $I(0) = \infty$ abd $I(\infty) = 0$. It is easily verified that

$$U(I(y)) - yI(y) \ge U(c) - yc, \quad 0 \le c < \infty, 0 < y < \infty. \tag{5.8.7}$$

Define a function $\mathscr{X} : [0,\infty] \to [0,\infty]$ by

$$\mathscr{X}(y) = \tilde{E}\int_0^T e^{-\int_0^s r(u)du} I(yZ_s e^{\int_0^s (\beta(u)-r(u))du})ds, \tag{5.8.8}$$

and assume that

$$\mathscr{X}(y) < \infty, \quad 0 < y < \infty. \tag{5.8.9}$$

Also define $\bar{y} := \sup\{y \geq 0 : \mathscr{X} \text{ is strictly decreasing on } [0, y]\}$.

Exercise 17. Under condition (5.8.9), show that $\mathscr{X}$ is continuous and strictly decreasing on $[0, \bar{y}]$ with $\mathscr{X}(0) = \infty$ and $\mathscr{X}(\bar{y}) = 0$.

Let $\mathscr{Y} : [0, \infty] \to [0, \bar{y}]$ be the inverse of $\mathscr{X}$. For a given initial endowment $x \geq 0$, define the processes

$$\eta^*_s := \mathscr{Y}(x) Z_s e^{\int_0^s (\beta(u) - r(u))du} \tag{5.8.10}$$

$$C^*_s := I(\eta^*_s). \tag{5.8.11}$$

The definition of $\mathscr{Y}$ implies that $C^* \in \mathscr{D}(x)$. We show now that C^* is an *optimal consumption process.*

Theorem 5.32. *Let $x \geq 0$ be a given and assume that (5.8.9) holds. Then the consumption process C^* is optimal, i.e.,*

$$V(x) = E \int_0^\infty e^{-\int_0^t \beta(s)ds} U(C^*_t)dt \tag{5.8.12}$$

Proof. It suffices to compare C^* to an arbitrary $C \in \mathscr{D}(x)$. For such a C, we have

$$E \int_0^\infty e^{-\int_0^t \beta(s)ds} (U(C^*_t) - U(C_t))dt \tag{5.8.13}$$

$$= E \int_0^\infty e^{-\int_0^t \beta(s)ds} (U(I(\eta^*_t))) - \eta^*_t I(\eta^*_t)) - (U(C_t) - \eta^*_t C_t)dt \tag{5.8.14}$$

$$+ \mathscr{Y}(x) \tilde{E} \int_0^T e^{-\int_0^t r(s)ds} (C^*_t - C_t)dt. \tag{5.8.15}$$

The first expectation on the right-hand side is nonnegative because of (5.8.7), while the second vanishes because both C^* and C are in $\mathscr{D}(x)$. □

Having thus determined the value function and the optimal consumption process, we appeal to the construction in the proof of Proposition 5.23 for the determination of a corresponding portfolio process π^*. This does not provide us with a very useful representation of π^*, but one can specialise the model in various ways so as to obtain V, C^*, π^* more explicitly.

Appendix A. Measure theory background

A.1. **Probability space.** A probability space is a triple $(\Omega, \mathscr{F}, P)$. Here Ω is a set, $\mathscr{F}$ a σ-algebra, i.e., a nonempty collection of subsets of Ω such that (i) if $A \in \mathscr{F}$ then the complement $A^c \in \mathscr{F}$ and (ii) if A_i is a countable sequence of sets in $\mathscr{F}$ then the union $\cup_i A_i \in \mathscr{F}$. As a consequence $\mathscr{F}$ is closed under countable intersection.

A measure $\mu : \mathscr{F} \to \mathbf{R}$ is a nonnegative countably additive function, i.e., $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathscr{F}$ and $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ for any countable sequence $A_i \in \mathscr{F}$. If $\mu(\Omega) = 1$, then we call μ a probability measure and denote it P.

Example A.1. Let Ω be a countable set, $\mathscr{F}$ the power set of Ω, and

$$P(A) = \sum_{x \in A} p(x), \qquad \sum_{x \in \Omega} p(x) = 1. \tag{A.1.1}$$

Example A.2. Let $\Omega = \mathbf{R}$, $\mathscr{F}$ the Borel subgalgebra, i.e., the smallest σ-algebra containing the open sets. Define a Stieltjes measure function F nondecreasing and right-continuous on $\mathbf{R}$, i.e.,

$$\lim_{y \to x^+} F(y) = F(x). \tag{A.1.2}$$

Associated to each such F is a unique measure such that $\mu((a, b]) = F(b) - F(a)$. The special case $F(x) = x$ gives the Lebesgue measure. To get Lebesgue measure on $\mathbf{R}^d$ is slightly more complicated, and requires an extra condition.

Now a random variable on the probability space Ω is a function $X : \Omega \to \mathbf{R}$ such that for every Borel set B in $\mathbf{R}$ we have $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathscr{F}$.

Example A.3. (1) If Ω is a discrete probability space then any X is a random variable. (2) The indicator (or characteristic) function 1_A of a set $A \in \mathscr{F}$ is also a random variable.

A.2. **Distributions.** X induces a probability measure on $\mathbf{R}$, called its distribution by setting $\mu(A) = P(X \in A) := P(X^{-1}(A))$ for all Borel sets $A \subset \mathbf{R}$. The distribution is described by its distribution function

$$F(x) = P(X \leq x) \tag{A.2.1}$$

Now let F be a distribution function. Then it is nondecreasing, right-continuous, and

$$\lim_{x \to \infty} F(x) = 1, \qquad \lim_{x \to -\infty} F(x) = 0. \tag{A.2.2}$$

Conversely, any function F satisfying these properties is the distribution function of some random variable. Also, if random variables X and Y induce the same distribution μ on $\mathbf{R}$, then we say they are equal in distribution. From the above characterisation, we that this holds if and only if X and Y have the same distribution function, namely $P(X \leq x) = P(Y \leq x)$ for all x. We denote this by

$$X \stackrel{d}{=} Y. \tag{A.2.3}$$

We say X has a density function f_X if a distribution function F has the form

$$F(x) = \int_{-\infty}^{x} f_X(y) dy. \tag{A.2.4}$$

On the other hand, we can start with f and use the above to define an F. All we need is for $f(x) \geq 0$ and $\int_{\mathbf{R}} f(x) dx = 1$.

Example A.4. Exercise: determine $F(x)$ for the first two examples below:

1. The uniform distribution on $(0, 1)$ given by $f(x) = 1_{(0,1)}$
2. The exponential distribution with rate λ given by $f(x) = \lambda e^{-\lambda(x)}$ for $x \geq 0$ and 0 otherwise.
3. The standard normal distribution given by $f(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

Call a distribution function on $\mathbf{R}$ absolutely continuous if it has a density, and singular if the corresponding measure is singular with respect to Lebesgue measure. Also, call a distribution function (or the induced probability measure) discrete if there is a countable set A with $P(A^c) = 0$. The simplest example is taking $F(x) = 1$ for $x \geq 0$ and $F(x) = 0$ for $x < 0$.

A.3. **Random variables.** Let's generalise a little. Let $(S, \mathcal{S})$ be a measure space with σ-algebra $\mathcal{S}$. We call a function $X : \Omega \to S$ measurable if $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathscr{F}$ for all $B \in \mathcal{S}$. Note that if $\{\omega : X(\omega) \in A\} \in \mathscr{F}$ for all $A \in \mathcal{A}$ and $\mathcal{A}$ generates $\mathcal{S}$, then X is measurable.

Example A.5. Conisder $\mathbf{R}^d$ with the Lebesgue measure. Then the sets $(a_1, b_1) \times \cdots \times (a_d, b_d)$ for all $a_i < b_i$ form such an $\mathcal{A}$.

If $X : \Omega \to \mathcal{S}$ and $f : \mathcal{S} \to \mathcal{T}$ are measurable, then $f(X)$ is measurable. (Prove this.) Moreover, if $X_1, X_2, \dots$ are random variables then so are (also prove this)

(A.3.1) $$\inf_n X_n, \quad \sup_n X_n, \quad \liminf_n X_n, \quad \limsup_n X_n.$$

From this we see that

(A.3.2) $$\Omega_0 := \{\omega : \lim_{n\to\infty} X \text{ exists}\} = \{\omega : \limsup_{n\to\infty} X - \liminf_n X_n = 0\}$$

is a measurable set. If $P(\Omega_0) = 1$, we say that X_n converges almost surely or almost everywhere. Note that $X_\infty = \limsup_n X_n$ may be infinite, so it is sometimes useful to extend the definition of random variable to the extended real line $[-\infty, \infty]$.

A.4. **Integration.** We want to define integration. We will do it quickly, but step by step over a measure space $(\Omega, \mathscr{F})$ with a σ-finite measure μ. First, call f a simple function if $f = \sum_{i=1}^n a_i 1_{A_i}$ for disjoint sets A_i with $\mu(A_i) < \infty$. Then define

(A.4.1) $$\int f d\mu = \sum_{i=1}^n \mu(A_i).$$

We say $f \geq g$ almost everywhere $\mu(\{\omega : f(\omega) < g(\omega)\}) = 0$, and write a.s. for short.

Lemma A.6. *Let f, g be simple functions. Then*

1. *If $f \geq 0$ a.s. then $\int f \; d\mu \geq 0$.*
2. *$\int af \; d\mu = a \int f \; d\mu$ for all $a \in \mathbf{R}$.*
3. *$\int f + g \; d\mu = \int f \; d\mu + \int g \; d\mu$.*

If $g \leq f$ a.s. then $\int g \; d\mu \leq \int f \; d\mu$.
$|\int f \; d\mu| \leq \int |f| d\mu$.

On to the next step. Let A be a set with $\mu(A) < \infty$ and let h be a bounded function that vanishes on E^c. If f, g are simple functions such that $f \leq h \leq g$, then we want $\int f \; d\mu \leq \int h \; d\mu \leq \int g \; d\mu$. so define

(A.4.2) $$\int h \; d\mu = \sup_{f \leq h} \int f \; d\mu = \inf_{g \leq h} \int g \; d\mu.$$

The last equality is statement to be proven. Then we can prove the above lemma for bounded functions.

On to nonnegative functions. Let $f \geq 0$. Then define

(A.4.3) $$\int f \; d\mu = \sup\{\int h d\mu : 0 \leq h \leq f, h \text{ is bounded and } \mu(\{x : h(x) > 0\}) < \infty\}$$

Finally, for a general function f, we say f is integrable if $\int |f| d\mu < \infty$. Define $f^+ = \max(f, 0)$ and $f^- = \max(-x, 0)$, then define the integral of f by

(A.4.4) $$\int f \; d\mu = \int f^+ d\mu + \int f^- d\mu.$$

Note the right hand side is well defined since the integrands are bounded by $|f|$. Then all the properties of the above lemma hold true.

Exercise 18. Prove the Riemann-Lebesgue lemma: If f is integrable then

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} g(x)\cos(nx)dx = 0 \tag{A.4.5}$$

Theorem A.7 (Jensen's inequality)**.** *Let f be a convex function on $\mathbf{R}$, that is, $cf(x)+(1-c)f(y)\geq f(cx+(1-c)y)$ for all $c\in(0,1)$ and $x,y\in\mathbf{R}$. If μ is a probability measure and $g, f(g)$ are integrable then*

$$f\Big(\int g\ d\mu\Big) \leq \int f(g)d\mu. \tag{A.4.6}$$

Define $||f||_p = (\int |f|^p d\mu)^{1/p}$ for any $1\leq p\leq\infty$.

Theorem A.8 (Hölder's inequality)**.** *Given $p,q\in(1,\infty)$ such that $\frac{1}{p}+\frac{1}{q}=1$, then*

$$\int |fg|d\mu \leq ||f||_p||g||_q. \tag{A.4.7}$$

Note that the case $p=q=2$ is known as the Cauchy-Schwartz inequality.

We say that $f_n\to f$ in measure if for any $\epsilon>0$, we have $\mu(\{x:|f_n(x)-f(x)|>\epsilon\})=0$ as $n\to\infty$. The next theorem is used to prove the one following it, and so on.

Theorem A.9. *The following theorems hold:*

(1) *(Bounded convergence theorem) Let A be a set with $\mu(A)<\infty$, and suppose that (i) $f_n=0$ on A^c, (ii) $|f_n(x)|\leq M$ for some $M>0$, and (iii) $f_n\to f$ in measure. Then*

$$\int f\ d\mu = \lim_{n\to\infty}\int f_n d\mu \tag{A.4.8}$$

(2) *(Fatou's lemma) If $f_n\geq 0$ for all n, then*

$$\liminf_{n\to\infty}\int f_n d\mu \geq \int(\liminf_{n\to\infty} f_n)d\mu \tag{A.4.9}$$

(3) *(Monotone convergence theorem) If $f_n\geq 0$ for all n, and $f_n\uparrow f$ then*

$$\int f_n d\mu \uparrow \int f\ d\mu. \tag{A.4.10}$$

(4) *(Dominated convergence theorem) If $f_n\to f$ almost everywhere, $|f_n|\leq g$ for all n, and g is integrable, then*

$$\int f_n d\mu \to \int f\ d\mu. \tag{A.4.11}$$

Exercise 19. Prove Minkowski's inequality: $||f+g||_p \leq ||f||_p+||g||_p$ for any $p\in[1,\infty]$.

A.5. **Expected value.** Now we specialise back to a probability measure $\mu=P$. If $X\geq 0$ is a random variable on $(\Omega,\mathscr{F},P)$ then define its expected value or mean to be $EX=\int X\ dP$. Also set $EX=EX^+-EX^-$. We observe the following basic properties:

(1) $E(X+Y)=EX+EY$
(2) $E(aX+b)=aEX+b$ for any $a,b\in\mathbf{R}$
(3) If $X\geq Y$, then $EX\geq EY$.

Since EX is defined by integration, we have the following results from the previous theorems:

Theorem A.10. *The following theorems hold:*

(1) *(Jensen) If f is convex, then $E(f(X)) \geq f(EX)$ as long as both expectations exist, that is, $E|X|, E|f(X)| < \infty$.*
(2) *(Hölder) If $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$, then $E|XY| \leq ||X||_p ||Y||_q$.*
(3) *(Fatou) If $X_n \geq 0$ for all n, then $\liminf_{n\to\infty} EX_n \geq E(\liminf_{n\to\infty} X_n)$.*
(4) *(Monotone convergence) If $X_n \geq 0$ for all n and $X_n \uparrow X$, then $EX_n \uparrow EX$.*
(5) *(Dominated convergence) If $X_n \to X$ almost surely, $|X_n| \leq Y$ for all n, and $|EY| \leq \infty$ then $EX_n \to EX$.*

Note that if Y is constant, then the last statement is the bounded convergence theorem. To state the next theorem, we define $E(X; A) = \int_A X dP$.

Theorem A.11 (Chebyshev-Markov inequality)**.** *Let $f \geq 0$, and $i_A = \inf\{f(a) : a \in A\}$ for a Borel set $A \subset \mathbf{R}$. Then*

$$i_A P(X \in A) \leq E(f(X); X \in A) \leq Ef(X). \tag{A.5.1}$$

Some call Markov's inequality the special case $f(a) = a^2$ and $A = \{a : |a| \geq b\}$, in which case $b^2 P(|X| \geq b) \leq EX^2$.

Proposition A.12. *Suppose $X_n \to X$ almost surely, and let f, g be continuous functions such that (i) $f \geq 0$ and $f(x) \to \infty$ as $|x| \to \infty$, (ii) $|g(x)|/f(x) \to 0$ as $|x| \to \infty$, and (iii) for all n, we have $Ef(X_n) \leq M$ for a fixed M. Then $Eg(X_n) \to Eg(X)$.*

The most important special case of the above result is when $f(x) = |x|^p, p > 1$ and $g(x) = x$.

Theorem A.13 (Change of variables)**.** *Let X be a random variable on $(S, \mathcal{S})$ with distribution μ, that is, $\mu(A) = P(X \in A)$. If f is a measurable function from $S \to \mathbf{R}$ with either $f \geq 0$ or $E|f(X)| \leq \infty$, then*

$$Ef(X) = \int_S f(y)\mu(dy) \tag{A.5.2}$$

If we write h for X and $P \circ h^{-1}$ for μ we have

$$\int_\Omega f(h(\omega))dP = \int_S f(y)d(P \circ h^{-1}). \tag{A.5.3}$$

Using this theorem we can compute expected values by integrating on $\mathbf{R}$. To set up our examples, we define the k-th moment of X to be EX^k, for some positive integer k. If $k = 1$, EX is called the mean and denoted μ. If $EX^2 < \infty$, the variance of X is defined to be

$$\text{var}(X) := E(X - \mu)^2 = EX^2 - 2\mu EX + \mu^2 = EX^2 - \mu^2. \tag{A.5.4}$$

Exercise 20. Show that if X is an exponential distribution of rate 1 then

$$EX = \int_0^\infty x^k e^{-x} dx = k! \tag{A.5.5}$$

Moreover, if we scale $\frac{1}{\lambda}X$, it has density $\lambda e^{-\lambda y}$ for $y \geq 0$, and has mean λ^{-1} and variance λ^{-2}.

Exercise 21. Show that if X is the standard normal distribution, then $EX = 0$ and $EX = 1$. If $Y = \mu X + \sigma$ with $\mu > 0, \sigma \in \mathbf{R}$, then $EY = \mu$ and $\text{var}(Y) = \sigma^2$, and Y has density

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}. \tag{A.5.6}$$

It is the normal distribution with mean μ and variance σ^2.

Exercise 22. We say that X has Bernoulli distribution with parameter p if $P(X = 1) = p$ and $P(X = 0) = 1 - p$. Show that $EX = p$ and $\text{var}(X) = p(1 - p)$.

Exercise 23. We say that X has Poisson distribution with parameter λ if $P(X = k) = e^{-\lambda}\lambda^k/k!$ for $k = 0, 1, \dots$. Show that

$$E(X(X-1)\dots(X-k+1)) = \lambda^k, \tag{A.5.7}$$

and deduce that $EX = \lambda$ and $\mathrm{var}(X) = \lambda$.

Exercise 24. (Inclusion-exclusion) Let $A_1, A_2, \dots$ be events with $A = \cup_{i=1}^n A_i$. Prove that $1_A = 1 - \prod_{i=1}^n (1 - 1_{A_i})$. Expand the right hand side and take expected value to conclude

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots \tag{A.5.8}$$

$$\cdots (-1)^{-1} P(\cap_{i=1}^n A_i) \tag{A.5.9}$$

Exercise 25. (Bonferroni inequalities) Let $A_1, A_2, \dots$ be events with $A = \cup_{i=1}^n A_i$. Prove that $1_A \le \sum_{i=1}^n 1_{A_i}$. Then take expected values to conclude

$$P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i) \tag{A.5.10}$$

$$P(\cup_{i=1}^n A_i) \ge \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) \tag{A.5.11}$$

$$P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \tag{A.5.12}$$

In general, if we stop the inclusion-exclusion formula after an even (odd) number of sums, we get an lower (upper) bound.

A.6. **Product measures.** Let $(X_1, \mathcal{A}_1, \mu_1), (X_2, \mathcal{A}_2, \mu_2)$ be two σ-finite measure spaces. Define $\Omega = X_1 \times X_2$ and $\mathscr{F} = \mathcal{A}_1 \times \mathcal{A}_2$, the σ-algebra generated by $A_1 \times A_2$ for $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$.

Theorem A.14. *There is a unique measure μ on $\mathscr{F}$ with $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$.*

Applying the theorem with induction, we obtain Lebesgue measure on the Borel subsets in $\mathbf{R}^n$.

Theorem A.15 (Fubini's theorem)**.** *If $f \ge 0$ and $\int |f| d\mu < \infty$, then*

$$\int_{X_1} \int_{X_2} f(x_1, x_2) \mu_2(dx_2) \mu_1(dx_1) = \int_{X_1 \times X_2} f d\mu = \int_{X_2} \int_{X_1} f(x_1, x_2) \mu_1(dx_1) \mu_2(dx_2) \tag{A.6.1}$$

Example A.16. Let $(X_1, \mathcal{A}_1, \mu_1) = (X_2, \mathcal{A}_2, \mu_2)$ with $X_1 = \mathbf{N}$, $\mathcal{A}_1$ all subsets of $\mathbf{N}$, and μ_1 the counting measure. Define $f(n,n) = 1, f(n+1, n) = -1$ for $n \ge 1$ and 0 otherwise. But then

$$\sum_m \sum_n f(m,n) = 1, \qquad \sum_n \sum_m f(m,n) = 0 \tag{A.6.2}$$

Example A.17. Let $X_1 = (0,1), X_2 = (1, \infty)$ with Borel sets and Lebesgue measure. Let $f(x,y) = e^{-xy} - 2e^{-2xy}$. Then

$$\int_0^1 \int_1^\infty f(x_1, x_2) dx_2 dx_1 = \int_0^1 x_1^{-1}(e^{-x_1} - e^{-2x_1}) dx_1 > 0 \tag{A.6.3}$$

$$\int_1^\infty \int_0^1 f(x_1, x_2) dx_2 dx_1 = \int_1^\infty x_2^{-1}(e^{-x_2} - e^{-2x_2}) dx_2 < 0 \tag{A.6.4}$$

Appendix B. Additional exercises

B.1. **Discrete Markov chains and martingales.**

(1) (The Pólya Urn scheme) Consider two urns containing a total of N balls. Pick one of the N balls at random and move it to the other urn. Let X_n be the number of balls in one of the urns, call it U, after the n-th draw. This forms a Markov chain over the state space $S = \{1, \dots, N\}$, with transition probability given by

$$p(k, k+1) = \frac{N-k}{N}, \quad p(k, k-1) = \frac{k}{N}, \tag{B.1.1}$$

for $1 \le k \le N$, and $p(i, j) = 0$ otherwise.

(a) Show that all states are recurrent.
(b) Show that $\mu(x) = 2^{-N}\binom{N}{x}$ is a stationary distribution.
(c) Show that $p^n(x, x) = 0$ if n is odd.
(d) Show that $E_x X_{n+1} = 1 + (1 - \frac{2}{N}) E_x X_n$
(e) Using the last result and induction conclude that

$$E_x X_n = \frac{N}{2} - \left(1 - \frac{2}{N}\right)^n \left(x - \frac{N}{2}\right) \tag{B.1.2}$$

i.e., the mean $E_x X_n$ converges exponentially to the equilibrium $N/2$.

(2) (Brother-sister mating) Two animals are mated, and among their direct descendants two of opposite sex are selected at random, are mated and the process continues. Suppose each individual can be one of three genotypes AA, Aa, aa, (denote it by 2,1,0) and suppose that the type of the offspring is determined by selecting a letter from each parent. With these rules, the pair of genotypes in the n-th generation is a Markov chain with six states:

$$22, 21, 20, 11, 10, 00 \tag{B.1.3}$$

(a) Compute its transition probability.
(b) Show that the number of A's in the pair is a martingale.
(c) A state a is called absorbing if $P_a(X_1 = a) = 1$. Notice that 22 and 00 are absorbing states for the chain. Show that the probability of absorption in 22 is equal to the fraction of A's in the state. (You may use (b) if you like.)
(d) Let $T = \min\{n \ge 0 : X_n = 22 \text{ or } 00\}$ be the absorption time. Find $E_x T$ for all states x.

(3) (Ehrenfest chain) Consider an urn that contains red and green balls. At each time n, choose a ball at random, then put it back and add one more ball of the same colour. Let X_n be the fraction of red balls at time n. Since $X_n \ge 0$ for all n and X_n is a martingale, it follows that X_n converges, say to X_∞.

(a) Show that X_n is a martingale.
(b) Suppose that at time 0 there is one ball of each colour. Show that the probability that red balls are drawn on the first j draws and then green balls are drawn on the next $n - j$ is $j!(n-j)!/(n+1)!$.
(c) Using (b) to conclude that for any $1 \le j \le n+1$,

$$P(X_n = \frac{j}{n+2}) = \frac{1}{n+1}. \tag{B.1.4}$$

and therefore the distribution of the limit X_∞ is uniform.
(d) Suppose that at time 0 there are r red balls and g green balls. Show that at X_∞ has the distribution

$$\frac{(g+r-1)!}{(g-1)!(r-1)!} x^{g-1}(1-x)^{r-1}. \tag{B.1.5}$$

(4) (Galton-Watson process) Recall the branching process in class: Given i.i.d nonnegative integer-valued random variables $\{\xi_i^n\}$, define Z_n by $Z_0 = 1$, and

$$Z_{n+1} = \begin{cases} \xi_1^n + \cdots + \xi_{Z_n}^n, & \text{if } Z_n > \\ 0 & \text{if } Z_n = 0. \end{cases} \tag{B.1.6}$$

Also let $p_k = P(\xi_i^n = k)$, called the offspring distribution.

(a) (Subcritical) Show that if $\mu < 1$, then $Z_n = 0$ for all n sufficiently large.

(b) (Critical) Show that if $\mu = 1$ and $p_1 < 1$, then $Z_n = 0$ for all n sufficiently large.

(c) (Supercitical) This process was invented to study the survival of family names. Suppose each family has exactly 3 children, which are male or female with equal probability. If only female children keep the family name, this leads to a branching process with $p_0 = 1/8, p_1 = 3/8, p_2 = 3/8$, and $p_3 = 1/8$. Show that $\mu > 1$, and compute the probability ρ that the family name will die out.

(5) (Gambler's ruin chain) Let $\{X_n\}$ be independent random variables with

$$P(X_i = 1) = p, \qquad P(X_i = -1) = 1 - p = q \tag{B.1.7}$$

with $0 < p < 1$. Let $S_n = X_0 + X_1 + \cdots + X_n$, and $h(x) = (q/p)^x$.

(a) Show that X_n is a Markov chain.

(b) Show that $M_n = h(S_n)$ is a martingale with respect to X_n.

(c) (Unfair games) Let $\frac{1}{2} < p < 1$. Given $a, b > 0$, define the stopping time $N = \min\{n : S_n \notin (a, b)\}$. Show that

$$P_x(S_N = a) = \frac{(\frac{q}{p})^b - (\frac{q}{p})^x}{(\frac{q}{p})^b - (\frac{q}{p})^a}. \tag{B.1.8}$$

(Hint: First argue that $h(x) = E_x M_{N \wedge n}$, and letting n tend to infinity it is equal to

$$E_x M_N = (q/p)^a P(S_N = a) + (q/p)^b P(S_N = b) \tag{B.1.9}$$

and solve.)

(d) (Fair games) Let $p = q = \frac{1}{2}$, and $a < 0 < b$. We have seen that S_n and $S_n^2 - n$ are martingales. Show that

$$E_0 N = -ab. \tag{B.1.10}$$

(Hint: Show that as n tends to infinity, the right-hand side of $E_x(S_{N\wedge n}^2 - N \wedge n) = 0$ tends to

$$a^2 P_0(S_N = a) + b^2 P_0(S_N = b) - E_0 N.) \tag{B.1.11}$$

B.2. **Brownian motion.**

(1) Prove that Brownian motion B_t is not differentiable, with probability one. (Hint: See Durrett, Theorem 7.1.6)

(2) (The invariance principle) Let's come back to the fair game, $S_n = X_1 + \cdots + X_n$ where the X_i are independent random variables with $P(X_i = 1) = P(X_i = -1) = \frac{1}{2}$, and $S_0 = 0$. Define

$$B_n(t) = \frac{S_{[nt]}}{\sqrt{n}} \tag{B.2.1}$$

for $t \geq 0$, and $[t]$ denotes the largest integer less than or equal to t. As n tends to infinity, $B_n(t)$ behaves like a standard Brownian motion. This convergence of partials sums of i.i.d's with zero mean and unit variance to a

Brownian motion is called the *invariance principle*, though we do not prove this here.

(a) Show that for $a < 0 < b$ and $N = \min\{n : S_n \notin (a,b)\}$, we have

$$P_0(B_n(N) = a) = P_0(S_N = a) = \frac{b}{b-a}. \tag{B.2.2}$$

Then conclude that $E_0 T = E_0 N = -ab$, where $T = \min\{n : B_n(t) \notin (a,b)\}$.

(b) Let's stop the random walk once it drops a units below the historical maximum. Set

$$M_n := \max_{0\le k\le n} S_k, \quad Y_n = M_n - S_n, \quad \tau = \min\{n \ge 0 : Y_n = a\}. \tag{B.2.3}$$

Show that $P(M_\tau = 0) = \dfrac{1}{1+a}$.

(c) Show that $P(M_\tau \ge k) = \left(\dfrac{a}{1+a}\right)^k$. What is the distribution of M_τ?

(d) Let $B(t)$ be standard Brownian motion. Also let

$$M(t) = \max_{0\le s\le t} B(s), \quad Y(t) = M(t) - B(t), \quad \tau = \min\{t \ge 0 : Y(t) = a\}. \tag{B.2.4}$$

Assuming the invariance principle, argue that $M(\tau)$ has an exponential distribution with mean a.

(Note that τ is a popular strategy for the sale of a stock, i.e., keep the stock as long as it is going up, but sell it once it drops a units past its historical best. Since $E[B(\tau)] = E[M(\tau)] - a = 0$, in the Brownian motion model of the stock market this strategy does not profit on average.)

(3) Let $\{X_n\}$ be i.i.d. random variables on $(\Omega, \mathscr{F}, P)$ with $EX_i = 0$ and $\mathrm{var}(X_i) \in (0,\infty)$ for all i. Let $S_n = X_1 + \cdots + X_n$.

(a) Show using the Central Limit Theorem and Kolmogorov's 0-1 law to conclude that $\limsup S_n/\sqrt{n} = \infty$ a.e.

(b) Show that $S_n/\sqrt{n}$ does not converge in probability. Hint: Consider $n = m!$ and argue by contradiction.

(4) (Quadratic variation of Brownian motion) Let $\{\Pi_n\}$ be a sequence of partitions of $[0,t]$ such that $\lim_{n\to\infty} ||\Pi_n|| = 0$.

(a) Show that the quadratic variation

$$V_t^{(2)}(\Pi_n) = \sum_{k=1}^{m_n} |B_{t_k^{(n)}} - B_{t_{k-1}^{(n)}}|^2 \tag{B.2.5}$$

converges in L^2 to t as $n \to \infty$. (Hint: Write $V_t^{(2)}(\Pi_n) - t$ as a sum of independent, mean-zero random variables, and show that

$$E(V_t^{(2)}(\Pi_n))^2 \le tE(Z^2-1)^2 ||\Pi||, \tag{B.2.6}$$

where Z is a standard normal random variable. See the notes also for a different proof sketch.)

(b) If moreover $\sum_{n=1}^\infty ||\Pi_n|| < \infty$, then the convergence takes place with probability one. (Hint: Show that for all $\epsilon > 0$,

$$\sum_{n=1}^\infty P(|V_t^{(2)}(\Pi_n) - t| > \epsilon) \le \frac{K}{\epsilon^2} \sum_{n=1}^\infty ||\Pi_n|| \tag{B.2.7}$$

for some constant K, and apply the Borel-Cantelli lemma.)

(5) A Wiener process W_t is a stochastic process adapted to a filtration $\mathscr{F}_t$ such that (a) $W_0 = 0$, (b) W_t is a martingale with $E[W_t^2] < \infty$ for all $t \geq 0$, such that

$$E[(W_t - W_s)^2] = t - s, \qquad s \leq t, \tag{B.2.8}$$

and (c) W_t is continuous in t. By a theorem of Lévy, a Wiener process is a Brownian motion process. Prove that a Brownian motion process is a Wiener process.

(6) (Brownian motion is Markov.) Let X and Y be d-dimensional random vectors on $(\Omega, \mathscr{F}, P)$. If $\mathscr{G} \subset \mathscr{F}$ is a sub-σ-algebra, X is indepdent of $\mathscr{G}$ adn Y is $\mathscr{G}$-measurable, then for every $A \in \mathscr{B}(\mathbf{R}^d)$ we have
(a) $P[X + Y \in A|\mathscr{G}] = P[X + Y \in A|Y]$, P-a.e.
(b) $P[X + Y \in A|Y = y] = P[X + y \in A]$ for PY^{-1}-a.e. $y \in \mathbf{R}^d$.
where $PY^{-1}(B) = P(\omega \in \Omega : X(\omega) \in B)$ for any $B \in \mathscr{B}(\mathbf{R}^d.)$

(7) Prove directly from the definition of the Itô integral that

$$\int_0^t s\ dB_s = tB_s - \int_0^t B_s ds. \tag{B.2.9}$$

(Hint: Note that $\sum_i \Delta(s_i B_i) = \sum_i s_i \Delta B_i + \sum_i B_{i+1} \Delta s_i$, where $\Delta a_i := a_{t_{i+1}} - a_{t_i}$.)

(8) Use Itô's formula to prove that

$$\int_0^t B_s^2 dB_s = \frac{1}{3} B_t^3 - \int_0^t B_s ds \tag{B.2.10}$$

(9) (The Fisk-Stratanovich integral) Let B_t be a standard Brownian motion, and $\epsilon \in [0, 1]$. Given a partition $\Pi = \{t_0, t_1, \dots, t_n\}$ of $[0, t]$ by $0 = t_0 < t_1 < \cdots < t_n = t$, and consider the approximating sum

$$S_\epsilon(\Pi) := \sum_{i=0}^{n-1} [(1-\epsilon) B_{t_i} + \epsilon B_{t_{i+1}}](B_{t_{i+1}} - B_{t_i}) \tag{B.2.11}$$

for the stochastic integral $\int_0^t B_s\ dB_s$. Show that the supremum over partitions converges to

$$\frac{1}{2} B_t^2 + \Big(\epsilon - \frac{1}{2}\Big) t, \tag{B.2.12}$$

in L^2. This expression is a martingale if and only if $\epsilon = 0$, and gives Itô's integral. If $\epsilon = \frac{1}{2}$ we get the Fisk-Stratonovich integral, which has the usual calculus rule $\int_0^t B_s\ dB_s = \frac{1}{2} B_t^2$. Finally, $\epsilon = 1$ leads to the so-called backwards Itô integral. The sensitivity of the limit to the value of ϵ is a consequence of the unbounded variation of the Brownian path.

(Hint: First show that $S_\epsilon(\Pi) = \frac{1}{2} B_t^2 + (\epsilon - \frac{1}{2}) \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2$. Then argue as in Itô's integral.)

B.3. **Stochastic differential equations and PDEs.**

(1) (Gronwall inequality): Given a function $g(t) \geq 0$ such that

$$g(t) \leq \alpha(t) + \beta \int_0^t g(s) ds, \quad 0 \leq t \leq T,$$

with $\beta \geq 0$ and $\alpha : [0,T] \to \mathbf{R}$ integrable, then

$$g(t) \leq \alpha(t) + \beta \int_0^t \alpha(s) e^{\beta(t-s)} ds, \quad 0 \leq t \leq T.$$

Hint: Argue that $d/dt(e^{-\beta t} \int_0^t g(s)ds) \leq \alpha(t) e^{-\beta t}$.

(2) Let $\{B_t = (B_t^{(1)}, \dots, B_t^{(d)}, \mathscr{F}_t; 0 \leq t < \infty\}$ be a d-dimensional Brownian motion. Show that the cross-variation $\langle B^{(i)}, B^{(j)} \rangle_t = \delta_{ij} t$ for $1 \leq i,j \leq d$.

Hint: Show that for $X, Y \in \mathscr{M}_2^c$ and a partition $\Pi = \{t_0, \dot{;} t_n\}$ of $[0,t]$, we have in probability,

$$\lim_{||\Pi|| \to 0} \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}}) = \langle X, Y \rangle_t.$$

(3) Show that the stochastic process X constructed from the proof of the existence of strong solutions [KS, 5.2.9] satisfies the expected stochastic integral equation, i.e., argue that

$$\left|\left| \int_0^t b(s, X_s^{(k)}) ds - \int_0^t b(s, X_s) ds \right|\right|^2$$

and

$$E \left|\left| \int_0^t \sigma(s, X_s^{(k)}) dW_s - \int_0^t \sigma(s, X_s) dW_s \right|\right|^2$$

converge to 0 a.s. for $0 \leq t \leq T$ as $k \to \infty$. Note that $\{X_t^{(k)}\}$ is a Cauchy sequence and $X_t^{(k)} \to X_t$ a.s. in $L^2(\Omega, \mathscr{F}, P)$.

(4) Let $B = \{B_t, \mathscr{F}_t; 0 \leq t < \infty\}$ be a 1-dimensional standard Brownian motion.

(a) Solve the stochastic differential equation:

$$\begin{pmatrix} dX_1 \\ dX_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & X_1 \end{pmatrix} \begin{pmatrix} dB_1 \\ dB_2 \end{pmatrix}$$

(b) Solve

$$dX_t = X_t dt + dB_t.$$

(Hint: multiply the equation by e^{-t}, and compare with $d(e^{-t} X_t)$.)

(c) Solve the Ornstein-Uhlenbeck, or Langevin equation

$$dX_t = \mu X_t dt + \sigma dB_t$$

where $\mu, \sigma \in \mathbf{R}$.

(5) Let D be an open subset of $\mathbf{R}^d$, and $g : D \to \mathbf{R}, f : \partial D \to \mathbf{R}$ continuous, bounded functions. Assume that $u : \overline{D} \to \mathbf{R}$ is continuous, of class $C^2(D)$, and solves the Poisson equation

$$\frac{1}{2} \Delta u = -g$$

in D, subject to the boundary condition

$$u = f$$

on ∂D. Let's show that $u(x)$ can be represented by

$$u(x) = E^x[f(W_{\tau_D})] + \int_0^{\tau_D} g(W_t) dt, \qquad x \in \overline{D}. \tag{B.3.1}$$

(a) Consider an increasing sequence $\{D_n\}_{n=1}^{\infty}$ of open sets with $\overline{D}_n \subset D$ for all $n \geq 1$, and $\cup_{n=1}^{\infty} D_n = D$, so that the stopping times $\tau_D = \inf\{t \geq 0 : W_t \notin D_n\}$ satisfy $\lim_{n\to\infty} \tau_n = \tau_D$ a.s. P^x. Using Itô's formula, argue that

$$M_t^{(n)} := u(W_{t\wedge\tau_n}) + \int_0^{t\wedge\tau_n} g(W_s)ds$$

is a P^x-martingale for every $n \geq 1, x \in D$.

(b) Let

$$M_t := u(W_{t\wedge\tau_D}) + \int_0^{t\wedge\tau_D} g(W_s)ds.$$

Argue that $|M_t(\omega)|$ and $|M_t^{(n)}(\omega)$ are bounded above by

$$\max_{x\in\overline{D}} |u(x)| + (t \wedge \tau_D(\omega)) \max_{x\in\overline{D}} |g(x)|$$

for P^x-a.e. $\omega \in \Omega$.

(c) Letting $n \to \infty$ and using the bounded convergence theorem, we know that the process $M = \{M_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a martingale. Show that if M is uniformly integrable, then $M_\infty = \lim_{t\to\infty} M_t$

$$M_\infty = u(W_{\tau_D}) + \int_0^{\tau_D} g(W_s)ds, \qquad P^x\text{-a.s..}$$

is a martingale, and from $E^x M_0 = E^x M_\infty$, deduce the representation (B.3.1). (Hint: The general case is Problem 1.3.20 in Karatzas-Shreve.)

(d) (Optional) Prove that M is uniformly integrable, e.g., by showing that $E^x\tau_D < \infty, \forall x \in D$.

(6) (Optional) Let W be a Brownian motion. Define

$$T = \inf\{0 \leq t \leq 1 : t + W_t^2 = 1\},$$

$$X_t = -\frac{2}{(1-t)^2} W_t 1_{t\leq T}, \quad 0 \leq t < 1,$$

and $X_1 = 0$.

(a) Prove that $P(T < 1) = 1$, and therefore $\int_0^1 X_t^2 dt < \infty$ a.s.

(b) Apply Ito's formula to the process $\{(W_t/(1-t))^2; 0 \leq t < 1$ to conclude that

$$\int_0^1 X_t dW_t - \frac{1}{2}\int_0^1 X_t^2 dt = -1 - 2\int_0^T ((1-t)^{-4} - (1-t)^{-3})W_t^2 dt \leq -1.$$

(7) Recall that d-dimensional Brownian motion corresponds to $b_i(t,x) \equiv 0$ and $\sigma_{ij}(t,x) \equiv \delta_{ij}, 1 \leq i,j \leq d$, and we have

$$\mathscr{A} f = \frac{1}{2}\Delta f = \frac{1}{2}\sum_{i=1}^{d} \frac{\partial^2 f}{\partial x_i^2}, \quad f \in C^2(\mathbf{R}^d).$$

Show that a continuous adapted process $W = \{W_t, \mathscr{F}_t; 0 \leq t < \infty\}$ is a d-dimensional Brownian motion if and only if

$$f(W_t) - f(W_0) - \frac{1}{2}\int_0^t \Delta f(W_s)ds,$$

is in $\mathscr{M}^{c,\mathrm{loc}}$ for every $f \in C^2(\mathbf{R}^d)$. This provides a martingale characterisation of Brownian motion.

(8) (a) Recall the dispersion matrix $\{\sigma(t) = \{(\sigma_{ij}(t))_{1\le i,j\le d}, \mathscr{F}_t, 0 \le t \le T\}$ for the stocks $i = 1, \dots, d$, assumed to be measurable, adapted, and bounded uniformly in $(t, \omega) \in [0, T] \times \Omega$. We set $a(t) := \sigma(t)\sigma^T(t)$ and assume that for some number $\epsilon > 0$,

$$\xi^T a(t)\xi \ge \epsilon||\xi||^2, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s..}$$

Show that $\sigma^T(t)$ has an inverse, and

$$||(\sigma^T(t))^{-1}\xi|| \le \epsilon{-}1/2||\xi||, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s..}$$

Moreover, with $\hat{a}(t) := \sigma^T(t)\sigma(t)$, we have

$$\xi^T \hat{a}(t)\xi \ge \epsilon||\xi||^2, , \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s.,}$$

so $\sigma(t)$ also has an inverse and

$$||(\sigma(t))^{-1}\xi|| \le \epsilon^{-1/2}||\xi||, \quad \forall \xi \in \mathbf{R}^d, 0 \le t \le T, \text{ a.s.}$$

(b) Derive the stochastic differential equation (5.6.11) in the notes.

$$dX_t = (r(t)X_t - C_t)dt + \sum_{i=1}^{d}(b_i(t) - r(t))\pi_i(t)dt + \sum_{i,j=1}^{d} \pi_i(t)\sigma_{ij}(t)dW_t^{(j)}.$$

Then derive the strong solution (5.6.11):

$$X_t = e^{\int_0^t r(s)ds}\Big[x + \int_0^t e^{-\int_0^s r(u)du}\{\pi(s)^T(b(s) - r(s)\vec{1}) - C_s\}ds + \int_0^t e^{-\int_0^s r(u)du}\pi^T(s)\sigma(s)dW_s\Big], \quad 0 \le t \le T.$$

(c) In your own words, explain how the Novikov condition and Girsanov theorem are applied to the strong solution to obtain the expression

$$X_t e^{-\int_0^t r(s)ds} + \int_0^t e^{-\int_0^s r(u)du}C_s ds = x + \int_0^t e^{-\int_0^s r(u)du}\pi^T(s)\sigma(s)d\tilde{W}_s.$$

Prove that $X_t e^{-\int_0^t r(s)ds}$ is a nonnegative supermartingale under $\tilde{P}$ (hint: see [KS, Problem 1.5.19]).

(d) Complete the proof of the sufficient condition for admissibility, Proposition 5.23.

(i) Show that martingale $N_t := E(DZ_T|\mathscr{F}_t), \quad 0 \le t \le T$, is right continuous with left limits (RCLL) P-a.e.

(ii) Show that there exists a measurable $\{\mathscr{F}_t\}$-adapted $\mathbf{R}^d$-valued process Y with $\int_0^T ||Y(t)||^2 dt < \infty$, and

$$N_t = E(DZ_T) + \sum_{j=1}^{d}\int_0^t Y_j(s)dW_s^{(j)}, \quad 0 \le t \le T, \ P\text{-a.s.}$$

(iii) Show using Itô's formula that for $m_t := \tilde{E}[D|\mathscr{F}_t] - \tilde{E}D$ and $\varphi(t) := (Y(t) + N_t\theta(t))/Z_t$,

$$m_t = \sum_{j=1}^{d}\int_0^t \varphi_j(s)d\tilde{W}_s^{(j)}, \quad 0 \le t \le T.$$

(iv) Conclude that ξ_t is a strong solution as desired.

(9) (a) To complete the proof of the Black-Scholes formula, show that the the function

$$v(t,x) := \begin{cases} x\Phi(\rho_+(T-t,x)) - qe^{-r(T-t)}\Phi(\rho_-(T-t,x)), & 0 \le t < T, x \ge 0 \\ (x-q)^+; & t = T, x \ge 0, \end{cases}$$

with

$$\rho_\pm(t,x) = \frac{1}{\sigma\sqrt{x}}\left[\log\frac{x}{q} + t\Big(r \pm \frac{\sigma^2}{2}\Big)\right], \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-x^2/2}dx$$

satisfies the Cauchy problem

$$-\frac{\partial v}{\partial t} + rv = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 v}{\partial x^2} + rx\frac{\partial v}{\partial x}, \text{ on } [0,T)\times(0,\infty)$$

$$v(T,x) = (x-q)^+, \quad x \ge 0,$$

as well as the conditions of Theorem 5.19 (the Feynman-Kac formula).

(b) In the setting of the Black-Scholes formula, but with $f_T = h(P_1(T))$ where $h : [0,\infty) \to [0,\infty)$ is a convex piecewise C^2 function with $h(0) = h'(0) = 0$, show that the valuation process for the contingent claim $(0, f_T)$ is given by

$$X_t = \tilde{E}[e^{-r(T-t)}h(P_t(T))|\mathscr{F}_t] = \int_0^\infty h''(q)v_{q,T}(t, P_1(t))dq.$$

We denote here by $v_{q,t}(t,x)$ the function of (5.7.14).

(10) For a fixed $a, b \in \mathbf{R}$ consider the 1-dimensional equation

$$dY_t = \frac{b - Y_t}{1-t}dt + dB_t, \quad 0 \le t < 1, \ Y_0 = a.$$

Verify that

$$Y_t = a(1-t) + bt + (1-t)\int_0^t \frac{dB_s}{1-s}, \quad 0 \le t < 1$$

is a solution, and that $\lim_{t\to 1} Y)_t = b$ a.s. The process Y_t is called the Brownian bridge.

(11) Find the generator of the following diffusions:
(a) $dX_t = \mu X_t dt + \sigma dB_t$, where μ, σ are constants. (Ornstein-Uhlenbeck process)
(b) $dX_t = rX_t dt + \alpha X_t dB_t$, where r, α are constants. (Geometric Brownian motion)
(c) $\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & X_1 \end{bmatrix}\begin{bmatrix} dB_1 \\ dB_2 \end{bmatrix}$, where B_i are Brownian motion.

(12) Let $C_0^2(\mathbf{R})$ denote functions on $\mathbf{R}$ that are compactly supported and twice differentiable. Find the diffusions whose generator is given by the following:
(a) $\mathscr{A}f(x) = f'(x) + f''(x)$, where $f \in C_0^2(\mathbf{R})$
(b) $\mathscr{A}f(x) = \frac{\partial f}{\partial t} + cx\frac{\partial f}{\partial x} + \frac{1}{2}\alpha^2 x^2 \frac{\partial^2 f}{\partial x^2}$, where $f \in C_0^2(\mathbf{R})$.

(13) Let $\alpha(t) = \frac{1}{2}\ln(1 + \frac{2}{3}t^3)$. If W_t is a Brownian motion, prove that there exists another Brownian motion $\tilde{W}_t$ such that

$$\int_0^{\alpha(t)} e^s dW_s = \int_0^t s \, d\tilde{W}_s.$$

(14) Let W_t be a Brownian motion in $\mathbf{R}$. Show that $X_t := W_t^2$ is a weak solution to the stochastic differential equation

$$dX_t = dt + 2|X_t|^{\frac{1}{2}} d\tilde{W}_t.$$

(Hint: Use Itô's formula to express X_t as a stochastic integral.)