11

Brownian Motion and Potential Theory

Introduction

Diffusion can be understood on several levels. The study of diffusion on a macroscopic level, of a substance such as heat, involves the notion of the flux of the quantity. If $u(t,x)$ measures the intensity of the quantity that is diffusing, the flux J across the boundary of a region $\mathcal{O}$ in x-space satisfies the identity

$$\frac{\partial}{\partial t} \int_{\mathcal{O}} u(t,x)\, dV(x) = -\int_{\partial\mathcal{O}} \nu \cdot J\, dS(x), \tag{0.1}$$

as long as the substance is being neither created nor destroyed. By the divergence theorem, this implies

$$\frac{\partial u}{\partial t} = -\operatorname{div} J. \tag{0.2}$$

The mechanism of diffusion creates a flux in the direction from greater concentration to lesser concentration. In the simplest model, the quantitative relation specified is that the flux is proportional to the x-gradient of u:

$$J = -D \operatorname{grad} u, \tag{0.3}$$

with $D > 0$. Applying (0.2), we obtain for u the PDE

$$\frac{\partial u}{\partial t} = D\, \Delta u, \tag{0.4}$$

in case D is constant. In such a case we can make $D = 1$, by rescaling, and this PDE is the one usually called “the heat equation.”

Many real diffusions result from jitterings of microscopic or submicroscopic particles, in a fashion that appears random. This motivates a probabilistic attack on diffusion, including creating probabilistic tools to analyze the heat equation. This is the topic of the present chapter.

In §1 we give a construction of Wiener measure on the space of paths in $\mathbb{R}^n$, governed by the hypothesis that a particle located at $x \in \mathbb{R}^n$ at time t_1 will have the probability $P(t, x, U)$ of being in an open set $U \subset \mathbb{R}^n$ at time $t_1 + t$, where

$$P(t, x, U) = \int\limits_U p(t, x, y)\, dy, \tag{0.5}$$

and $p(t, x, y)$ is the fundamental solution to the heat equation. We prove that, with respect to Wiener measure, almost every path is continuous, and we establish a modulus of continuity. Our choice of $e^{t\Delta}$ rather than $e^{t\Delta/2}$ to define such probabilities differs from the most popular convention and leads to minor differences in various formulas. Of course, translation between the two conventions is quite easy.

In §2 we establish the Feynman-Kac formula, for the solution to

$$\frac{\partial u}{\partial t} = \Delta u + V(x)u, \tag{0.6}$$

in terms of an integral over path space. A limiting argument made in §3 gives us formulas for the solution to (0.4) on a bounded domain Ω, with Dirichlet boundary conditions. This also leads to formulas for solutions to

$$\Delta u = f \text{ on } \Omega, \quad u = 0 \text{ on } \partial\Omega, \tag{0.7}$$

and

$$\Delta u = 0 \text{ on } \Omega, \quad u = g \text{ on } \partial\Omega. \tag{0.8}$$

A different, and more natural, formula for the solution to (0.8) is derived in §5, after the development in §4 of a tool known as the "strong Markov property." In §6 we present a study of the Newtonian capacity of a compact set $K \subset \mathbb{R}^n$, in the case $n \geq 3$, which is related to the probability that a Brownian path starting outside K will hit K. We give Wiener's criterion for a point y in $\partial\Omega$ to be regular for the Dirichlet problem (0.8), in terms of the capacity of $K_r = \{z \in \partial\Omega : |z - y| \leq r\}$, as $r \to 0$, which has a natural probabilistic proof.

In §7 we introduce the notion of the stochastic integral, such as

$$\int_0^t f\big(s, \omega(s)\big)\, d\omega(s), \tag{0.9}$$

which is not straightforward since almost all Brownian paths fail to have locally bounded variation. We show how the solution to

$$\frac{\partial u}{\partial t} = \Delta u + Xu \tag{0.10}$$

can be given in terms of an integral over path space, whose integrand involves a stochastic integral, in case X is a first-order differential operator. The derivation of this formula, like the derivation of the Feynman-Kac formula in §2, uses a tool from functional analysis known as the Trotter product formula, which we establish in Appendix A at the end of this chapter.

In §8 we consider a more general sort of stochastic integral, needed to solve stochastic differential equations:

$$d\mathfrak{X} = b(t, \mathfrak{X})\, dt + \sigma(t, \mathfrak{X})\, d\omega, \tag{0.11}$$

which we study in §9. Via Ito's formulas, stochastic differential equations can be used to treat diffusion equations of the form

$$\frac{\partial u}{\partial t} = \sum A_{jk}(x)\, \partial_j \partial_k u + \sum b_j(x)\, \partial_j u + V(x)u, \tag{0.12}$$

in terms of path space integrals. We look at this in §10. Results there, specialized to (0.10), yield a formula with a different appearance than that derived in §7. The identity of these two formulas leads to a formula of Cameron-Martin-Girsanov, representing the "Jacobian determinant" of a certain nonlinear transformation of path space.

An important topic that we do not treat here is Malliavin's stochastic calculus of variations, introduced in [Mal], which has had numerous interesting applications to PDE. We refer the reader to [Stk2] and [B] for material on this, and further references.

1. Brownian motion and Wiener measure

One way to state the probabilistic connection with the heat equation

$$\frac{\partial u}{\partial t} = \Delta u \tag{1.1}$$

is in terms of the heat kernel, $p(t, x, y)$, satisfying

$$e^{t\Delta} f(x) = \int p(t, x, y) f(y)\ dV(y). \tag{1.2}$$

If Δ in (1.1) is the Friedrichs extension of the Laplacian on any Riemannian manifold M, the maximum principle implies

$$p(t, x, y) \geq 0. \tag{1.3}$$

In many cases, including all compact M and $M = \mathbb{R}^n$, we also have

$$\int p(t, x, y)\ dV(y) = 1. \tag{1.4}$$

Consequently, for each $x \in M$, $p(t, x, y)\ dV(y)$ defines a probability distribution, which we can interpret as giving the probability that a particle starting at the point x at time 0 will be in a given region in M at time t.

Restricting our attention to the case $M = \mathbb{R}^n$, we proceed to construct a probability measure, known as "Wiener measure," on the set of paths

$\omega : [0, \infty) \to \mathbb{R}^n$, undergoing a random motion, sometimes called Brownian motion, described as follows. Given $t_1 < t_2$ and that $\omega(t_1) = x_1$, the probability density for the location of $\omega(t_2)$ is

$$e^{t\Delta}\delta_{x_1}(x) = p(t, x - x_1) = (4\pi t)^{-n/2} e^{-|x-x_1|^2/4t}, \quad t = t_2 - t_1. \tag{1.5}$$

The motion of a random path for $t_1 \le t \le t_2$ is supposed to be independent of its past history. Thus, given $0 < t_1 < t_2 < \cdots < t_k$, and given Borel sets $E_j \subset \mathbb{R}^n$, the probability that a path, starting at $x = 0$ at $t = 0$, lies in E_j at time t_j for each $j \in [1, k]$ is

$$\int_{E_1} \cdots \int_{E_k} p(t_k - t_{k-1}, x_k - x_{k-1}) \cdots p(t_1, x_1)\, dx_k \cdots dx_1. \tag{1.6}$$

It is not obvious that there is a countably additive measure characterized by these properties, and Wiener's result was a great achievement. The construction we give here is a slight modification of one in Appendix A of [Nel2].

Anticipating that Wiener measure is supported on the set of continuous paths, we will take a path to be characterized by its locations at all positive *rational* t. Thus, we consider the set of "paths"

$$\mathfrak{P} = \prod_{t \in \mathbb{Q}^+} \dot{\mathbb{R}}^n. \tag{1.7}$$

Here, $\dot{\mathbb{R}}^n$ is the one-point compactification of $\mathbb{R}^n$ (i.e., $\dot{\mathbb{R}}^n = \mathbb{R}^n \cup \{\infty\}$). Thus $\mathfrak{P}$ is a compact, metrizable space. We construct Wiener measure W as a positive Borel measure on $\mathfrak{P}$.

By the Riesz theorem, it suffices to construct a positive linear functional $E : C(\mathfrak{P}) \to \mathbb{R}$, on the space $C(\mathfrak{P})$ of real-valued, continuous functions on $\mathfrak{P}$, satisfying $E(1) = 1$. We first define E on the subspace $\mathcal{C}^\#$, consisting of continuous functions that depend on only finitely many of the factors in (1.7); that is, functions on $\mathfrak{P}$ of the form

$$\varphi(\omega) = F\big(\omega(t_1), \dots, \omega(t_k)\big), \quad t_1 < \cdots < t_k, \tag{1.8}$$

where F is continuous on $\prod_1^k \dot{\mathbb{R}}^n$, and $t_j \in \mathbb{Q}^+$. To be consistent with (1.6), we take

$$\begin{aligned} E(\varphi) = \int \cdots \int & p(t_1, x_1) p(t_2 - t_1, x_2 - x_1) \\ & \cdots p(t_k - t_{k-1}, x_k - x_{k-1}) \\ & F(x_1, \dots, x_k)\, dx_k \cdots dx_1. \end{aligned} \tag{1.9}$$

If $\varphi(\omega)$ in (1.8) actually depends only on $\omega(t_\nu)$ for some proper subset $\{t_\nu\}$ of $\{t_1, \dots, t_k\}$, there arises a formula for $E(\varphi)$ with a different appearance from (1.9). The fact that these two expressions are equal follows from the semigroup property of $e^{t\Delta}$. From this it follows that $E : \mathcal{C}^\# \to \mathbb{R}$ is well defined. It is also a positive linear functional, satisfying $E(1) = 1$.

Now, by the Stone-Weierstrass theorem, $\mathcal{C}^\#$ is dense in $C(\mathfrak{P})$. Since $E : \mathcal{C}^\# \to \mathbb{R}$ is a positive linear functional and $E(1) = 1$, it follows that E has a unique continuous extension to $C(\mathfrak{P})$, possessing these properties. Thus there is a unique probability measure W on $\mathfrak{P}$ such that

$$E(\varphi) = \int_{\mathfrak{P}} \varphi(\omega)\, dW(\omega). \tag{1.10}$$

This is the Wiener measure.

Proposition 1.1. *The set $\mathfrak{P}_0$ of paths from $\mathbb{Q}^+$ to $\mathbb{R}^n$, which are uniformly continuous on bounded subsets of $\mathbb{Q}^+$ (and which thus extend uniquely to continuous paths from $[0,\infty)$ to $\mathbb{R}^n$), is a Borel subset of $\mathfrak{P}$ with Wiener measure* 1.

For a set S, let $\operatorname{osc}_S(\omega)$ denote $\sup_{s,t\in S} |\omega(s) - \omega(t)|$. Set

$$E(a,b,\varepsilon) = \big\{\omega \in \mathfrak{P} : \operatorname{osc}_{[a,b]}(\omega) > 2\varepsilon\big\}; \tag{1.11}$$

here $[a,b]$ denotes $\{s \in \mathbb{Q}^+ : a \le s \le b\}$. Its complement is

$$E^c(a,b,\varepsilon) = \bigcap_{t,s\in[a,b]} \big\{\omega \in \mathfrak{P} : |\omega(s) - \omega(t)| \le 2\varepsilon\big\}, \tag{1.12}$$

which is closed in $\mathfrak{P}$. Below we will demonstrate the following estimate on the Wiener measure of $E(a,b,\varepsilon)$:

$$W\big(E(a,b,\varepsilon)\big) \le 2\rho\big(\frac{\varepsilon}{2}, |b-a|\big), \tag{1.13}$$

where

$$\rho(\varepsilon,\delta) = \sup_{t\le\delta} \int_{|x|>\varepsilon} p(t,x)\, dx, \tag{1.14}$$

with $p(t,x) = e^{t\Delta}\delta(x)$, as in (1.5). In fact, the sup is assumed at $t = \delta$, so

$$\rho(\varepsilon,\delta) = \int_{|y|>\varepsilon/\sqrt{\delta}} p(1,y)\, dy = \psi_n\big(\frac{\varepsilon}{\sqrt{\delta}}\big), \tag{1.15}$$

where

$$\psi_n(r) = (4\pi)^{-n/2} \int_{|y|>r} e^{-|y|^2/4}\, dy \le \alpha_n r^{n-1} e^{-r^2/4}, \tag{1.16}$$

as $r \to \infty$.

The relevance of the analysis of $E(a,b,\varepsilon)$ is that if we set

$$F(k,\varepsilon,\delta) = \big\{\omega \in \mathfrak{P} : \exists\ J \subset [0,k] \cap \mathbb{Q}^+, \ell(J) \le \delta,\ \operatorname{osc}_J(\omega) > 4\varepsilon\big\}, \tag{1.17}$$

where $\ell(J)$ is the length of the interval J, then

$$F(k,\varepsilon,\delta) = \bigcup \big\{E(a,b,2\varepsilon) : [a,b] \subset [0,k], |b-a| \le \delta\big\} \tag{1.18}$$

is an open set, and, via (1.13), we have

$$W\big(F(k,\varepsilon,\delta)\big) \le 2k\frac{\rho(\varepsilon,\delta)}{\delta}. \tag{1.19}$$

Furthermore, with $F^c(k,\varepsilon,\delta) = \mathfrak{P}\setminus F(k,\varepsilon,\delta)$,

$$\begin{aligned} \mathfrak{P}_0 &= \big\{\omega : \forall k<\infty, \forall \varepsilon>0, \exists \delta>0 \text{ such that } \omega\in F^c(k,\varepsilon,\delta)\big\} \\ &= \bigcap_k \bigcap_{\varepsilon=1/\nu} \bigcup_{\delta=1/\mu} F^c(k,\varepsilon,\delta) \end{aligned} \tag{1.20}$$

is a Borel set (in fact, an $\mathcal{F}_{\sigma\delta}$ set), and we can conclude that $W(\mathfrak{P}_0)=1$ from (1.19), given the observation that, for any $\varepsilon>0$,

$$\frac{\rho(\varepsilon,\delta)}{\delta} \longrightarrow 0, \quad \text{as } \delta\to 0, \tag{1.21}$$

which follows immediately from (1.15) and (1.16). Thus, to complete the proof of Proposition 1.1, it remains to establish the estimate (1.13).

Lemma 1.2. *Given $\varepsilon,\delta>0$, take ν numbers $t_j\in\mathbb{Q}^+$, $0\le t_1<\cdots<t_\nu$, such that $t_\nu - t_1\le\delta$. Let*

$$A = \big\{\omega\in\mathfrak{P} : |\omega(t_1)-\omega(t_j)|>\varepsilon, \ \text{for some } j=1,\dots,\nu\big\}. \tag{1.22}$$

Then

$$W(A) \le 2\rho\Big(\frac{\varepsilon}{2},\delta\Big). \tag{1.23}$$

Proof. Let

$$\begin{aligned} B &= \Big\{\omega : |\omega(t_1)-\omega(t_\nu)| > \frac{\varepsilon}{2}\Big\}, \\ C_j &= \Big\{\omega : |\omega(t_j)-\omega(t_\nu)| > \frac{\varepsilon}{2}\Big\}, \\ D_j &= \{\omega : |\omega(t_1)-\omega(t_j)|>\varepsilon \text{ and} \\ &\qquad |\omega(t_1)-\omega(t_k)|\le\varepsilon, \forall\, k\le j-1\}. \end{aligned} \tag{1.24}$$

Then $A\subset B\cup\bigcup_{j=1}^{\nu}\big(C_j\cap D_j\big)$, so

$$W(A) \le W(B) + \sum_{j=1}^{\nu} W\big(C_j\cap D_j\big). \tag{1.25}$$

Clearly, $W(B)\le\rho(\varepsilon/2,\delta)$. Furthermore, via (1.8)–(1.9), if we set

$$\begin{aligned} D\big(\omega(t_1),\dots,\omega(t_j)\big) &= 1, \quad \text{if } \omega\in D_j, \ 0 \text{ otherwise}, \\ C\big(\omega(t_j),\omega(t_\nu)\big) &= 1, \quad \text{if } \omega\in C_j, \ 0 \text{ otherwise}, \end{aligned}$$

we have $C(x_j, x_\nu) = C_1(x_j - x_\nu)$ and

$$\begin{aligned} &W(C_j \cap D_j) \\ &= \int \cdots \int D(x_1, \dots, x_j) C(x_j, x_\nu) p(t_1, x_1) p(t_2 - t_1, x_2 - x_1) \cdots \\ &\qquad p(t_j - t_{j-1}, x_j - x_{j-1}) p(t_\nu - t_j, x_\nu - x_j)\, dx_\nu dx_j \cdots dx_1 \\ &\le \rho\big(\frac{\varepsilon}{2}, \delta\big) \int \cdots \int D(x_1, \dots, x_j) p(t_1, x_1) \cdots p(t_j - t_{j-1}, x_j - x_{j-1}) \\ &\qquad \cdot\, dx_j \cdots dx_1 \\ &\le \rho\big(\frac{\varepsilon}{2}, \delta\big) W(D_j), \end{aligned} \tag{1.26}$$

so

$$\sum_j W\big(C_j \cap D_j\big) \le \rho\big(\frac{\varepsilon}{2}, \delta\big), \tag{1.27}$$

since the D_j are mutually disjoint. This proves (1.23).

Let us note an intuitive approach to (1.26). Since D_j describes properties of $\omega(t)$ for $t \in [t_1, t_j]$ and C_j describes a property of $\omega(t_\nu) - \omega(t_j)$, these sets describe *independent* events, so $W(C_j \cap D_j) = W(C_j) W(D_j)$; meanwhile $W(C_j) \le \rho(\varepsilon/2, \delta)$.

We continue the demonstration of (1.13). Now, given such t_j as in the statement of Lemma 1.2, if we set

$$E = \big\{\omega : |\omega(t_j) - \omega(t_k)| > 2\varepsilon, \text{ for some } j, k \in [1, \nu]\big\}, \tag{1.28}$$

it follows that

$$W(E) \le 2\rho\big(\frac{\varepsilon}{2}, \delta\big), \tag{1.29}$$

since E is a subset of A, given by (1.22). Now, $E(a, b, \varepsilon)$, given by (1.11), is a countable increasing union of sets of the form (1.28), obtained, say, by letting $\{t_1, \dots, t_\nu\}$ consist of all $t \in [a, b]$ that are rational with denominator $\le K$, and taking $K \nearrow +\infty$. Thus we have (1.13), and the proof of Proposition 1.1 is complete.

We make the natural identification of paths $\omega \in \mathfrak{P}_0$ with continuous paths $\omega : [0, \infty) \to \mathbb{R}^n$. Note that a function φ on $\mathfrak{P}_0$ of the form (1.8), with $t_j \in \mathbb{R}^+$, not necessarily rational, is a pointwise limit on $\mathfrak{P}_0$ of functions in $\mathcal{C}^\#$, as long as F is continuous on $\prod_1^k \dot{\mathbb{R}}^n$, and consequently such φ is measurable. Furthermore, (1.9) continues to hold, by the dominated convergence theorem.

An alternative approach to the construction of W would be to replace (1.7) by $\widetilde{\mathfrak{P}} = \prod\big\{\dot{\mathbb{R}}^n : t \in \mathbb{R}^+\big\}$. With the product topology, this is compact but not metrizable. The set of continuous paths is a Borel subset of $\widetilde{\mathfrak{P}}$, but not a Baire set, so some extra measure-theoretic considerations arise if one takes this route.

Looking more closely at the estimate (1.19) of the measure of the set $F(k,\varepsilon,\delta)$, defined by (1.17), we note that you can take $\varepsilon = K\sqrt{\delta\log 1/\delta}$, in which case

$$\rho(\varepsilon,\delta) = \psi_n\Big(K\sqrt{\log\frac{1}{\delta}}\Big) \le C_n\Big(\log\frac{1}{\delta}\Big)^{n/2-1}\delta^{K^2/4}. \tag{1.30}$$

Then we obtain the following refinement of Proposition 1.1.

Proposition 1.3. *For almost all $\omega \in \mathfrak{P}$, we have the modulus of continuity $8\sqrt{\delta\log 1/\delta}$, that is, given $0 \le s,t \le k < \infty$,*

$$\limsup_{|s-t|=\delta\to 0}\ \Big(\big|\omega(s)-\omega(t)\big| - 8\sqrt{\delta\log\frac{1}{\delta}}\ \Big) \le 0. \tag{1.31}$$

In fact, (1.30) gives $W(S_k) = 1$, where S_k is the set of paths satisfying (1.31), with 8 replaced by $8+1/k$, and then $\bigcap_k S_k$ is precisely the set of paths satisfying (1.31).

This result is not quite sharp; P. Levy showed that, for almost all $\omega \in \mathfrak{P}$, with $\mu(\delta) = 2\sqrt{\delta\log 1/\delta}$, $0 \le s,t \le k < \infty$,

$$\limsup_{|s-t|\to 0}\ \frac{|\omega(s)-\omega(t)|}{\mu(|s-t|)} = 1. \tag{1.32}$$

See [McK] for a proof. We also refer to [McK] for a proof of the result, due to Wiener, that almost all paths ω are nowhere differentiable.

By comparison with (1.31), note that if we define functions X_t on $\mathfrak{P}$, taking values in $\mathbb{R}^n$, by

$$X_t(\omega) = \omega(t), \tag{1.33}$$

then a simple application of (1.8)–(1.10) yields

$$\|X_t\|^2_{L^2(\mathfrak{P})} = \int |x|^2 p(t,x)\ dx = 2nt, \tag{1.34}$$

and more generally

$$\|X_t - X_s\|_{L^2(\mathfrak{P})} = \sqrt{2n}\ |s-t|^{1/2}. \tag{1.35}$$

Note that (1.35) depends on n, while (1.32) does not.

Via a simple translation of coordinates, we have a similar construction for the set of Brownian paths ω starting at a general point $x \in \mathbb{R}^\ell$, yielding the positive functional $E_x : C(\mathfrak{P}) \to \mathbb{R}$, and Wiener measure W_x, such that

$$E_x(\varphi) = \int_{\mathfrak{P}} \varphi(\omega)\ dW_x(\omega). \tag{1.36}$$

When $\varphi(\omega)$ is given by (1.8), $E_x(\varphi)$ has the form (1.9), with the function $p(t_1, x_1)$ replaced by $p(t_1, x_1 - x)$. To put it another way, $E_x(\varphi)$ has the form (1.9) with $F(x_1, \dots, x_k)$ replaced by $F(x_1 + x, \dots, x_k + x)$.

We will often use such notation as

$$E_x\big(f(\omega(t)\big)$$

instead of $\int_{\mathfrak{P}} f\big(X_t(\omega)\big)\,dW_x(\omega)$ or $E_x\big(f(X_t(\omega))\big)$.

The following simple observation is useful.

Proposition 1.4. *If $\varphi \in C(\mathfrak{P})$, then $E_x(\varphi)$ is continuous in x.*

Proof. Continuity for $\varphi \in \mathcal{C}^{\#}$, the set of functions of the form (1.8), is clear from (1.9) and its extension to $x \neq 0$ discussed above. Since $\mathcal{C}^{\#}$ is dense in $C(\mathfrak{P})$, the result follows easily.

Exercises

1. Given $a > 0$, define a transformation $D_a : \mathfrak{P}_0 \to \mathfrak{P}_0$ by

$$(D_a\omega)(t) = a\omega(a^{-2}t).$$

   Show that D_a preserves the Wiener measure W. This transformation is called Brownian scaling.
2. Let $\widetilde{\mathfrak{P}}_0 = \{\omega \in \mathfrak{P}_0 : \lim_{s\to\infty} s^{-1}\omega(s) = 0\}$. Show that $W(\widetilde{\mathfrak{P}}_0) = 1$. Define a transformation $\rho : \widetilde{\mathfrak{P}}_0 \to \mathfrak{P}_0$ by

$$(\rho\omega)(t) = t\omega(t^{-1}),$$

   for $t > 0$. Show that ρ preserves the Wiener measure W.
3. Given $a > 0$, define a transformation $R_a : \mathfrak{P}_0 \to \mathfrak{P}_0$ by

$$\begin{aligned}(R_a\omega)(t) = {} & \omega(t), && \text{for } 0 \le t \le a,\\ & 2\omega(a) - \omega(t), && \text{for } t \ge a.\end{aligned}$$

   Show that R_a preserves the Wiener measure W.
4. Show that $L^p(\mathfrak{P}_0, dW_0)$ is separable, for $1 \le p < \infty$. (*Hint*: $\mathfrak{P}$ is a compact metric space. Show that $C(\mathfrak{P})$ is separable.)
5. If $0 \le a_1 < b_1 \le a_2 < b_2$, show that $X_{b_1} - X_{a_1}$ is orthogonal to $X_{b_2} - X_{a_2}$ in $L^2(\mathfrak{P}, dW_x, \mathbb{R}^n)$, where $X_t(\omega) = \omega(t)$, as in (1.33).
6. Verify the following identities (when $n = 1$):

$$E_x\Big(e^{\lambda(\omega(t)-\omega(s))}\Big) = e^{|t-s|\lambda^2}, \tag{1.37}$$

$$E_x\Big(\big[\omega(t) - \omega(s)\big]^{2k}\Big) = \frac{(2k)!}{k!}|t-s|^k, \tag{1.38}$$

$$E\Big(\omega(s)\omega(t)\Big) = 2\,\min(s,t). \tag{1.39}$$

7. Show that $e^{\lambda|\omega(t)|^2} \in L^2(\mathfrak{P}_0, dW_0)$ if and only if $\lambda < 1/8t$.

2. The Feynman-Kac formula

To illustrate the application of Wiener measure to PDE, we now derive a formula, known as the Feynman-Kac formula, for the solution operator $e^{t(\Delta-V)}$ to

$$\frac{\partial u}{\partial t} = \Delta u - Vu, \quad u(0) = f, \tag{2.1}$$

given f in an appropriate Banach space, such as $L^p(\mathbb{R}^n),\ 1 \le p < \infty$, or $f \in C_o(\mathbb{R}^n)$, the space of continuous functions on $\mathbb{R}^n$ vanishing at infinity. To start, we will assume V is bounded and continuous on $\mathbb{R}^n$. Following [Nel2], we will use the Trotter product formula

$$e^{t(\Delta-V)}f = \lim_{k\to\infty} \Big(e^{(t/k)\Delta}e^{-(t/k)V}\Big)^k f. \tag{2.2}$$

For any k, $\Big(e^{(t/k)\Delta}e^{-(t/k)V}\Big)^k f$ is expressed as a k-fold integral:

$$\begin{aligned}
&\Big(e^{(t/k)\Delta}e^{-(t/k)V}\Big)^k f(x) \\
&= \int\cdots\int f(x_k)e^{-(t/k)V(x_k)}\, p\big(\frac{t}{k}, x_k - x_{k-1}\big)\, e^{(t/k)V(x_{k-1})}\cdots \\
&\qquad\cdot\, e^{-(t/k)V(x_1)}\, p\big(\frac{t}{k}, x - x_1\big)\, dx_1\cdots dx_k.
\end{aligned} \tag{2.3}$$

Comparison with (1.36) gives

$$\Big(e^{(t/k)\Delta}e^{-(t/k)V}\Big)^k f(x) = E_x(\varphi_k), \tag{2.4}$$

where

$$\varphi_k(\omega) = f\big(\omega(t)\big)\, e^{-S_k(\omega)}, \quad S_k(\omega) = \frac{t}{k}\sum_{j=1}^{k} V\Big(\omega\big(\frac{jt}{k}\big)\Big). \tag{2.5}$$

We are ready to prove the Feynman-Kac formula.

Proposition 2.1. *If V is bounded and continuous on $\mathbb{R}^n$, and $f \in C(\mathbb{R}^n)$ vanishes at infinity, then, for all $x \in \mathbb{R}^n$,*

$$e^{t(\Delta-V)}f(x) = E_x\Big(f\big(\omega(t)\big)e^{-\int_0^t V(\omega(\tau))\,d\tau}\Big). \tag{2.6}$$

Proof. We know that $e^{t(\Delta-V)}f$ is equal to the limit of (2.4) as $k \to \infty$, in the sup norm. Meanwhile, since almost all $\omega \in \mathfrak{P}$ are continuous paths, $S_k(\omega) \to \int_0^t V(\omega(\tau))d\tau$ boundedly and a.e. on $\mathfrak{P}$. Hence, for each $x \in \mathbb{R}^n$, the right side of (2.4) converges to the right side of (2.6). This finishes the proof.

Note that if V is real-valued and in $L^\infty(\mathbb{R}^n)$, then $e^{t(\Delta-V)}$ is defined on $L^\infty(\mathbb{R}^n)$, by duality from its action on $L^1(\mathbb{R}^n)$, and

$$f_\nu \in C_0^\infty(\mathbb{R}^n),\ f_\nu \nearrow 1 \Longrightarrow e^{t(\Delta-V)} f_\nu \nearrow e^{t(\Delta-V)} 1. \tag{2.7}$$

Thus, if V is real-valued, bounded, and continuous, then, for all $x \in \mathbb{R}^n$,

$$e^{t(\Delta-V)} 1(x) = E_x\left(e^{-\int_0^t V(\omega(\tau))d\tau}\right). \tag{2.8}$$

We can extend these identities to some larger classes of V. First we consider the nature of the right side of (2.6) for more general V.

Lemma 2.2. *Fix $t \in [0,\infty)$. If $V \in L^\infty(\mathbb{R}^n)$, then*

$$I_V(\omega) = \int_0^t V(\omega(\tau))\ d\tau \tag{2.9}$$

is well defined in $L^\infty(\mathfrak{P})$. If V_ν is a bounded sequence in $L^\infty(\mathbb{R}^n)$ and $V_\nu \to V$ in measure, then $I_{V_\nu} \to I_V$ boundedly and in measure on $\mathfrak{P}$. This is true for each measure $W_x,\ x \in \mathbb{R}^n$.

Proof. Here, L^∞ is the set of equivalence classes (mod a.e. equality) of bounded measurable functions, that is, elements of $\mathcal{L}^\infty(\mathbb{R}^n)$. Suppose $W \in \mathcal{L}^\infty(\mathbb{R}^n)$ is a pre-image of V. Then $\int_0^t W(\omega(\tau))\ d\tau = \iota_W(\omega)$ is defined and measurable, and $\|\iota_W\|_{\mathcal{L}^\infty(\mathfrak{P})} \le \|W\|_{\mathcal{L}^\infty(\mathbb{R}^n)} t$. If $W^\#$ is also a pre-image of V, then $W = W^\#$ almost everywhere on $\mathbb{R}^n$. Look at U, defined on $\mathfrak{P} \times \mathbb{R}^+$ by

$$U(\omega, s) = W(\omega(s)) - W^\#(\omega(s)).$$

This is measurable. Let $K \subset \mathbb{R}^n$ be the set where $W(x) \neq W^\#(x)$; this has measure 0. Now, for fixed s, the set of $\omega \in \mathfrak{P}$ such that $\omega(s) \in K$ has Wiener measure 0. By Fubini's theorem it follows that $U = 0$ a.e. on $\mathfrak{P} \times \mathbb{R}^+$, and hence, for almost all $\omega \in \mathfrak{P}$, $U(\omega, \cdot) = 0$ a.e. on $\mathbb{R}^+$. Thus $\int_0^t W^\#(\omega(\tau))\ d\tau = \int_0^t W(\omega(\tau))\ d\tau$ for a.e. $\omega \in \mathfrak{P}$, so I_V is well defined in $L^\infty(\mathfrak{P})$ for each $V \in L^\infty(\mathbb{R}^n)$. Clearly, $\|I_V\|_{L^\infty} \le \|V\|_{L^\infty} t$.

If $V_\nu \to V$ boundedly and in measure, in view of the previous argument we can assume without loss of generality that, upon passing to a subsequence, $V_\nu(x) \to V(x)$ for all x. Consider

$$U_\nu(\omega, s) = V(\omega(s)) - V_\nu(\omega(s)),$$

which is bounded in $L^\infty(\mathfrak{P} \times \mathbb{R}^+)$. This converges to 0 for each $(\omega, s) \in \mathfrak{P} \times \mathbb{R}^+$, so by Fubini's theorem again, $\int_0^t U_\nu(\omega, s)\ ds \to 0$ for a.e. ω. This completes the proof.

A similar argument yields the following.

Lemma 2.3. *If $V \in L^1_{\text{loc}}(\mathbb{R}^n)$ is bounded from below, then*

$$e_V(\omega) = e^{-\int_0^t V(\omega(\tau))\ d\tau} \tag{2.10}$$

is well defined in $L^\infty(\mathfrak{P})$*. If* $V_\nu \in L^1_{\text{loc}}(\mathbb{R}^n)$ *are uniformly bounded below and* $V_\nu \to V$ *in* L^1_{loc}*, then* $e_{V_\nu} \to e_V$ *boundedly and in measure on* $\mathfrak{P}$*.*

Thus, if $V \in L^1_{\text{loc}}(\mathbb{R}^n)$, $V \geq -K > -\infty$, take bounded, continuous V_ν such that $V_\nu \geq -K$ and $V_\nu \to V$ in L^1_{loc}. We have $\|e^{t(\Delta - V_\nu)}\| \leq e^{Kt}$ for all ν, where $\|\cdot\|$ can be the operator norm on $L^p(\mathbb{R}^n)$ or on $C_o(\mathbb{R}^n)$. Now, if we replace V by V_ν in (2.6), then Lemma 2.3 implies that, for any $f \in C_0^\infty(\mathbb{R}^n)$, the right side converges, for each x, namely,

$$E_x\Big(f\big(\omega(t)\big)e^{-\int_0^t V_\nu(\omega(\tau))\,d\tau}\Big) \longrightarrow P(t)f(x), \quad \text{as } \nu \to \infty. \tag{2.11}$$

Clearly $|P(t)f(x)| \leq e^{Kt}E_x(|f|) \leq e^{Kt}\|f\|_{L^\infty}$. Consequently, for each $x \in \mathbb{R}^n$, if $f \in C_0^\infty(\mathbb{R}^n)$,

$$e^{t(\Delta - V_\nu)}f(x) \longrightarrow P(t)f(x) = E_x\Big(f\big(\omega(t)\big)e^{-\int_0^t V(\omega(\tau))\,d\tau}\Big). \tag{2.12}$$

It follows that $P(t) : C_0^\infty(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$. Since

$$\big|e^{t(\Delta - V_\nu)}f(x)\big| \leq e^{Kt}e^{t\Delta}|f|(x), \tag{2.13}$$

we also have $P(t) : C_0^\infty(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$. Furthermore, we can pass to the limit in the PDE $\partial u_\nu/\partial t = \Delta u_\nu - V_\nu u_\nu$ for $u_\nu = e^{t(\Delta - V_\nu)}f$, to obtain for $u(t) = P(t)f$ the PDE

$$\frac{\partial u}{\partial t} = \Delta u - Vu, \quad u(0) = f. \tag{2.14}$$

If $\Delta - V$, with domain $\mathcal{D} = \mathcal{D}(\Delta) \cap \mathcal{D}(V)$, is self-adjoint, or has self-adjoint closure A, the uniqueness result of Proposition 9.11 in Appendix A, Functional Analysis, guarantees that $P(t)f = e^{tA}f$. For examples of such self-adjointness results on $\Delta - V$, see Chapter 8, §2, and the exercises following that section. Thus the identity (2.6) extends to such V, for example, to $V \in L^\infty(\mathbb{R}^n)$; so does the identity (2.8).

We can derive a similar formula for the solution operator $S(t,0)$ to

$$\frac{\partial u}{\partial t} = \Delta u - V(t,x)u, \quad u(0) = f, \tag{2.15}$$

using the time-dependent Trotter product formula, Proposition A.5, and its consequence, Proposition A.6. Thus, we obtain

$$S(t,0)f(x) = E_x\Big(f\big(\omega(t)\big)e^{-\int_0^t V(\tau,\omega(\tau))\,d\tau}\Big) \tag{2.16}$$

when $V(t) \in C\big([0,\infty), BC(\mathbb{R}^n)\big)$, $BC(\mathbb{R}^n)$ denoting the space of bounded continuous functions on $\mathbb{R}^n$. By arguments such as those used above, we can extend this identity to larger classes of functions $V(t)$.

Exercises

1. Given $\varepsilon > 0, \lambda \in \mathbb{R}$, compute the integral operator giving

$$e^{t(\partial_x^2 - \varepsilon x^2 - \lambda x)} f(x). \tag{2.17}$$

(*Hint*: Use $\varepsilon x^2 + \lambda x = \varepsilon(x + \lambda/2\varepsilon)^2 - \lambda^2/4\varepsilon$ to reduce this to the problem of computing the integral operator giving

$$e^{t(\partial_x^2 - \varepsilon x^2)} g(x). \tag{2.18}$$

For this, see the material on the harmonic oscillator in §6 of Chapter 8, in particular, Mehler's formula.)

2. Obtain a formula for

$$E_x\left(e^{-\varepsilon \int_0^t \omega(s)^2\, ds - \lambda \int_0^t \omega(s)\, ds}\right) = e^{t(\partial_x^2 - \varepsilon x^2 - \lambda x)} 1(x), \tag{2.19}$$

in the case of one-dimensional Brownian motion. (*Hint*: Use the formula

$$\begin{gathered} e^{t(\partial_x^2 - \varepsilon x^2)} 1(x) = a(t) e^{-b(t)x^2}, \\ a(t) = \left(\cosh 2\sqrt{\varepsilon} t\right)^{-1/2}, \quad b(t) = \frac{1}{2}\sqrt{\varepsilon} \tanh 2\sqrt{\varepsilon} t, \end{gathered} \tag{2.20}$$

which follows from the formula for (2.18). Alternatively, verify (2.20) directly, examining the system of ODE

$$a'(t) = -2a(t)b(t), \quad b'(t) = \varepsilon - 4b(t)^2.)$$

3. Pass to the limit $\varepsilon \searrow 0$ in (2.19), to evaluate

$$E_x\left(e^{-\lambda \int_0^t \omega(s) ds}\right). \tag{2.21}$$

Note that the monotone convergence theorem applies.

Exercises 4 and 5 will investigate

$$\psi(\varepsilon) = W_0\left(\left\{\omega \in \mathfrak{P} : \int_0^a \omega(s)^2\, ds < \varepsilon\right\}\right) = P\left(\int_0^a \omega(s)^2\, ds < \varepsilon\right). \tag{2.22}$$

4. Using Exercise 2, show that, for all $\lambda > 0$,

$$\begin{aligned} \int_0^\infty \psi'(s) e^{-\lambda s}\, ds &= E_0\left(e^{-\lambda \int_0^a \omega(s)^2\, ds}\right) \\ &= \left(\cosh 2a\sqrt{\lambda}\right)^{-1/2} = \sqrt{2} e^{-a\sqrt{\lambda}} \left(1 + e^{-4a\sqrt{\lambda}}\right)^{-1/2}. \end{aligned} \tag{2.23}$$

Other derivations of (2.23) can be found in [CM] and [Lev].

5. The subordination identity, given as (5.22) in Chapter 3, implies

$$\int_0^\infty \varphi_a(s) e^{-\lambda s}\, ds = \sqrt{2} e^{-a\sqrt{\lambda}} \quad \text{if} \quad \varphi_a(s) = \frac{a}{\sqrt{2\pi}} s^{-3/2} e^{-a^2/4s}.$$

Deduce that

$$\psi'(s) = \varphi_a(s) - \frac{1}{2}\varphi_{5a}(s) + \frac{3}{8}\varphi_{9a}(s) - \cdots,$$

hence that

$$
\begin{aligned}
&\frac{d}{d\varepsilon} P\left(\int_0^a \omega(s)^2\, ds < \varepsilon\right) \\
(2.24) \qquad &= \frac{a}{\sqrt{2\pi}} \varepsilon^{-3/2} \Bigl[e^{-a^2/4\varepsilon} - \frac{1}{2}\cdot 5e^{-25a^2/4\varepsilon} + \frac{3}{8}\cdot 9e^{-81a^2/4\varepsilon} - \cdots \Bigr].
\end{aligned}
$$

Show that the terms in this alternating series have progressively decreasing magnitude provided $\varepsilon/a^2 \le 1/2$. (*Hint*: Use the power series

$$
(1+y)^{-1/2} = 1 - \frac{1}{2} y + \frac{3}{8} y^2 - \cdots
$$

with $y = e^{-4a\sqrt{\lambda}}$.)

6. Suppose now that $\omega(t)$ is Brownian motion in $\mathbb{R}^n$. Show that

$$
E_0\left(e^{-\lambda \int_0^a |\omega(s)|^2\, ds}\right) = \left(\cosh 2a\sqrt{\lambda}\right)^{-n/2}.
$$

Deduce that in the case $n = 2$,

$$
\frac{d}{d\varepsilon} P\left(\int_0^a |\omega(s)|^2\, ds < \varepsilon\right) = \frac{2a}{\sqrt{\pi}} \varepsilon^{-3/2} \Bigl[e^{-a^2/\varepsilon} - 3e^{-9a^2/\varepsilon} + 5e^{-25a^2/\varepsilon} - \cdots \Bigr].
$$

Show that the terms in this alternating series have progressively decreasing magnitude provided $\varepsilon \le 2a^2$.

3. The Dirichlet problem and diffusion on domains with boundary

We can use results of §2 to provide connections between Brownian motion and the Dirichlet boundary problem for the Laplace operator. We begin by extending Lemma 2.3 to situations where $V_\nu \nearrow V$, with $V(x)$ possibly equal to $+\infty$ on a big set. We have the following analogue of Lemma 2.3.

Lemma 3.1. *Let $V_\nu \in L^1_{\text{loc}}(\mathbb{R}^n)$, $-K \le V_\nu \nearrow V$, with possibly $V(x) = +\infty$ on a set of positive measure. Then $e_V(\omega)$, given by (2.10), is well defined in $L^\infty(\mathfrak{P})$, provided we set $e^{-\infty} = 0$, and $e_{V_\nu} \to e_V$ boundedly and in measure on Ω, for each t.*

Proof. This follows from the monotone convergence theorem.

Thus we again have convergence with bounds in (2.11)–(2.13). We will look at a special class of such sequences. Let $\Omega \subset \mathbb{R}^n$ be open, with smooth boundary (in fact, Lipschitz boundary will more than suffice), and set $E = \mathbb{R}^n \setminus \Omega$. Let $V_\nu \ge 0$ be continuous and bounded on $\mathbb{R}^n$ and satisfy

$$
(3.1) \qquad V_\nu = 0 \text{ on } \overline{\Omega}, \quad V_\nu \ge \nu \text{ on } E_\nu, \quad V_\nu \nearrow,
$$

where E_ν is the set of points of distance $\ge 1/\nu$ from $\overline{\Omega}$. Given $f \in L^2(\mathbb{R}^n)$, $g \in L^2(\Omega)$, set $P_\Omega f = f|_\Omega \in L^2(\Omega)$, and define $E_\Omega g \in L^2(\mathbb{R}^n)$ to be $g(x)$ for $x \in \Omega$, 0 for $x \in E = \mathbb{R}^n \setminus \Omega$.

Proposition 3.2. *Under the hypotheses above, if* $f \in L^2(\mathbb{R}^n)$, *then*

$$e^{t(\Delta - V_\nu)} f \longrightarrow E_\Omega e^{t\Delta_\Omega}\big(P_\Omega f\big), \tag{3.2}$$

as $\nu \to \infty$, *where* Δ_Ω *is the Laplace operator with Dirichlet boundary condition on* Ω.

Proof. We will first show that, for any $\lambda > 0$,

$$\big(\lambda - \Delta + V_\nu\big)^{-1} f \to E_\Omega\big(\lambda - \Delta_\Omega\big)^{-1} P_\Omega f. \tag{3.3}$$

Indeed, denote the left side of (3.3) by u_ν, so $(\lambda - \Delta + V_\nu)u_\nu = f$. Taking the inner product with u_ν, we have

$$\lambda \|u_\nu\|^2_{L^2} + \|\nabla u_\nu\|^2_{L^2} + \int V_\nu |u_\nu|^2\, dx = (f, u_\nu) \le \frac{\lambda}{2}\|u_\nu\|^2_{L^2} + \frac{1}{2\lambda}\|f\|^2_{L^2}, \tag{3.4}$$

so

$$\frac{\lambda}{2}\|u_\nu\|^2_{L^2} + \|\nabla u_\nu\|^2_{L^2} + \int V_\nu |u_\nu|^2\, dx \le \frac{1}{2\lambda}\|f\|^2_{L^2}. \tag{3.5}$$

Thus, for fixed $\lambda > 0$, $\{u_\nu : \nu \in \mathbb{Z}^+\}$ is bounded in $H^1(\mathbb{R}^n)$, while $\int_{E_\nu} |u_\nu|^2\, dx \le C/\nu$. Thus $\{u_\nu\}$ has a weak limit point $u \in H^1(\mathbb{R}^n)$, and $u = 0$ on $\cup E_\nu$. The regularity hypothesized for $\partial\Omega$ implies $u \in H^1_0(\Omega)$. Clearly, $(\lambda - \Delta)u = f$ on Ω, so (3.3) follows, with weak convergence in $H^1(\mathbb{R}^n)$. But note that, parallel to (3.4),

$$\lambda \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2} = (f, u) = \lim_{\nu\to\infty} (f, u_\nu),$$

so

$$\lambda \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2} \ge \limsup_{\nu\to\infty}\ \lambda \|u_\nu\|^2_{L^2} + \|\nabla u_\nu\|^2_{L^2}. \tag{3.6}$$

Hence, in fact, we have H^1-norm convergence in (3.3), and a fortiori L^2-norm convergence.

Now consider the set $\mathcal{F}$ of real-valued $\varphi \in C_o([0, \infty))$ such that, for all $f \in L^2(\mathbb{R}^n)$,

$$\varphi(-\Delta + V_\nu) f \longrightarrow E_\Omega \varphi(-\Delta_\Omega) P_\Omega f, \quad \text{in } L^2(\mathbb{R}^n)\text{-norm}, \tag{3.7}$$

where $\varphi(H)$ is defined via the spectral theorem for a self-adjoint operator H. (Material on this functional calculus can be found in §1 of Chapter 8.) The analysis above shows that, for each $\lambda > 0$, $r_\lambda(s) = (\lambda + s)^{-1}$ belongs to $\mathcal{F}$. Since $P_\Omega E_\Omega$ is the identity on $L^2(\Omega)$, it is clear that $\mathcal{F}$ is an algebra; it is also easily seen to be a closed subset of $C_o([0, \infty))$. Since it contains r_λ for $\lambda > 0$, it separates points, so by the Stone-Weierstrass theorem all real-valued $\varphi \in C_o([0, \infty))$ belong to $\mathcal{F}$. This proves (3.2).

The version of (2.12) we have this time is the following.

Proposition 3.3. *Let $\Omega \subset \mathbb{R}^n$ be open, with smooth boundary, or more generally with the property that*

$$\{u \in H^1(\mathbb{R}^n) : \text{supp } u \subset \overline{\Omega}\} = H_0^1(\Omega).$$

Let $F \in C_0^\infty(\mathbb{R}^n)$, $f = F|_\Omega$. Then, for all $x \in \overline{\Omega}$, $t \geq 0$,

$$e^{t\Delta} f(x) = E_x \Bigl(f\bigl(\omega(t)\bigr) e^{-\int_0^t \ell_\Omega(\omega(\tau))\, d\tau} \Bigr). \tag{3.8}$$

On the left, $e^{t\Delta}$ is the solution operator to the heat equation on $\mathbb{R}^+ \times \Omega$ with Dirichlet boundary condition on $\partial\Omega$, and in the expression on the right

$$\ell_\Omega(x) = 0 \ \text{ on } \overline{\Omega}, \quad +\infty \ \text{ on } \mathbb{R}^n \setminus \overline{\Omega} = \overset{\circ}{E}. \tag{3.9}$$

Note that, for ω continuous,

$$e^{-\int_0^t \ell_\Omega(\omega(\tau))\, d\tau} = \psi_{\overline{\Omega}}(\omega, t) = \begin{matrix} 1 & \text{if } \omega([0,t]) \subset \overline{\Omega}, \\ 0 & \text{otherwise.} \end{matrix} \tag{3.10}$$

The second identity defines $\psi_{\overline{\Omega}}(\omega, t)$. Of course, for ω continuous, $\omega([0,t]) \subset \overline{\Omega}$ if and only if $\omega([0,t] \cap \mathbb{Q}) \subset \overline{\Omega}$.

We now extend Proposition 3.3 to the case where $\Omega \subset \mathbb{R}^n$ is open, with no regularity hypothesis on $\partial\Omega$. Choose a sequence Ω_j of open regions with smooth boundary, such that $\Omega_j \subset\subset \Omega_{j+1} \subset\subset \cdots$, $\bigcup_j \Omega_j = \Omega$. Let Δ_j denote the Laplace operator on Ω_j, with Dirichlet boundary condition, and let Δ denote that of Ω, also with Dirichlet boundary condition.

Lemma 3.4. *Given $f \in L^2(\Omega)$, $t \geq 0$,*

$$e^{t\Delta} f = \lim_{j\to\infty} E_j e^{t\Delta_j} P_j f, \tag{3.11}$$

where $P_j f = f|_{\Omega_j}$ and, for $g \in L^2(\Omega_j)$, $E_j g(x) = g(x)$ for $x \in \Omega_j$, 0 for $x \in \Omega \setminus \Omega_j$.

Proof. Methods of Chapter 5, §5, show that, for $\lambda > 0$,

$$E_j(\lambda - \Delta_j)^{-1} P_j f \to (\lambda - \Delta)^{-1} f \tag{3.12}$$

in L^2-norm, and then (3.11) follows from this, by reasoning used in the proof of Proposition 3.2.

Suppose $f \in C_0^\infty(\Omega_L)$. Then, for $j \geq L$, $E_j e^{t\Delta_j} f \to e^{t\Delta} f$ in L^2-norm, as we have just seen. Furthermore, local regularity implies

$$E_j e^{t\Delta_j} f \longrightarrow e^{t\Delta} f \quad \text{locally uniformly on } \Omega. \tag{3.13}$$

Thus, given such f, and any $x \in \Omega$ (hence $x \in \Omega_j$ for j large),

$$e^{t\Delta} f(x) = \lim_{j\to\infty} E_x \Big(f\big(\omega(t)\big)\psi_{\overline{\Omega}_j}(\omega,t)\Big). \tag{3.14}$$

Now, as $j \to \infty$,

$$\psi_{\overline{\Omega}_j}(\omega,t) \nearrow \psi_\Omega(\omega,t), \tag{3.15}$$

where we define

$$\begin{aligned} \psi_\Omega(\omega,t) = 1 \quad & \text{if } \omega([0,t]) \subset \Omega, \\ 0 \quad & \text{otherwise.} \end{aligned} \tag{3.16}$$

This yields the following:

Proposition 3.5. *For any open $\Omega \subset \mathbb{R}^n$, given $f \in C_0^\infty(\Omega),\ x \in \Omega$,*

$$e^{t\Delta} f(x) = E_x \big(f\big(\omega(t)\big)\psi_\Omega(\omega,t)\big). \tag{3.17}$$

In particular, if Ω has smooth boundary, one can use either $\psi_\Omega(\omega,t)$ or $\psi_{\overline{\Omega}}(\omega,t)$ in the formula for $e^{t\Delta}f(x)$. However, if $\partial\Omega$ is not smooth, it is $\psi_\Omega(\omega,t)$ that one must use.

It is useful to extend this result to more general f. Suppose $f_j \in C_0^\infty(\Omega)$, $f \in L^2(\Omega)$, and $f_j(x) \searrow f(x)$ for each $x \in \Omega$. Then, for any $t > 0$, $e^{t\Delta} f_j \to e^{t\Delta} f$ in $L^2(\Omega) \cap C^\infty(\Omega)$, while, for each $x \in \Omega$, $E_x\big(f_j(\omega(t))\psi_\Omega(\omega,t)\big)$ converges $\searrow$ to the right side of (3.17), by the monotone convergence theorem. Hence (3.17) holds for all such f; denote this class by $\mathcal{L}(\Omega)$. Clearly, the characteristic function $\chi_K \in \mathcal{L}(\Omega)$ for each compact $K \subset \Omega$.

By the same reasoning, the class of functions in $L^2(\Omega)$ for which (3.17) holds is closed under forming monotone limits, either $f_j \nearrow f$ or $f_j \searrow f$, of sequences bounded in $L^2(\Omega)$. An argument used in Lemma 2.2 shows that modifying $f \in L^2(\Omega)$ on a set of measure zero does not change the right side of (3.17). If $S \subset \Omega$ is measurable, then

$$\chi_S(x) = \lim_{j\to\infty} \chi_{K_j}(x), \text{ a.e.,}$$

for an increasing sequence of compact sets $K_j \subset S$, so (3.17) holds for $f = \chi_S$. Thus it holds for finite linear combinations of such characteristic functions, and an easy limiting argument gives the following:

Proposition 3.6. *The identity (3.17) holds for all $f \in L^2(\Omega)$ when $t > 0,\ x \in \Omega$.*

Suppose now that Ω is bounded. Then, for $f \in L^p(\Omega),\ 1 \le p \le \infty$,

$$-\Delta^{-1} f = \int_0^\infty e^{t\Delta} f\, dt, \tag{3.18}$$

the integral being absolutely convergent in L^p-norm. If $f \in C_0^\infty(\Omega)$, we hence have, for each $x \in \Omega$,

$$-\Delta^{-1} f(x) = E_x \left(\int_0^\infty f\big(\omega(t)\big) \psi_\Omega(\omega, t)\, dt \right). \tag{3.19}$$

Furthermore, by an argument such as used to prove Proposition 3.6, this identity holds for almost every $x \in \Omega$, given $f \in L^2(\Omega)$, and for *every* x if $f_j \in C_0^\infty(\Omega)$ and $f_j(x) \nearrow f(x)$ for all x. In particular, for Ω bounded,

$$-\Delta^{-1} 1(x) = E_x\big(\vartheta_\Omega(\omega)\big), \quad x \in \Omega, \tag{3.20}$$

where, if ω is a continuous path starting inside Ω, we define

$$\begin{aligned} \vartheta_\Omega(\omega) = \int_0^\infty \psi_\Omega(\omega, t) dt &= \sup\{t : \omega([0,t]) \subset \Omega\} \\ &= \min\{t : \omega(t) \in \partial\Omega\}. \end{aligned} \tag{3.21}$$

In other words, $\vartheta_\Omega(\omega)$ is the first time $\omega(t)$ hits $\partial\Omega$; it is called the "first exit time." Since $\Delta^{-1} 1 \in C^\infty(\Omega)$, it is clear that the first exit time for a path starting at any $x \in \Omega$ is finite for W_x-almost every ω when Ω is bounded. (If ω starts at a point in $\partial\Omega$ or in $\mathbb{R}^n \setminus \overline{\Omega}$, set $\vartheta_\Omega(\omega) = 0$.) Note that we can write

$$-\Delta^{-1} f(x) = E_x \left(\int_0^{\vartheta_\Omega(\omega)} f\big(\omega(t)\big)\, dt \right). \tag{3.22}$$

If $\partial\Omega$ is smooth enough for Proposition 3.3 to hold, we have the formula (3.19), with $\psi_\Omega(\omega, t)$ replaced by $\psi_{\overline{\Omega}}(\omega, t)$, valid for all $x \in \overline{\Omega}$. In particular, for smooth bounded Ω,

$$-\Delta^{-1} 1(x) = E_x\big(\vartheta_{\overline{\Omega}}(\omega)\big), \quad x \in \overline{\Omega}, \tag{3.23}$$

where we define

$$\vartheta_{\overline{\Omega}}(\omega) = \inf\big\{t : \omega(t) \in \mathbb{R}^n \setminus \overline{\Omega}\big\} = \max\big\{t : \omega\big([0,t]\big) \subset \overline{\Omega}\big\}. \tag{3.24}$$

(If $\omega(0) \in \mathbb{R}^n \setminus \overline{\Omega}$, set $\vartheta_{\overline{\Omega}}(\omega) = 0$.) Comparing this with (3.20), noting that $\vartheta_{\overline{\Omega}}(\omega) \geq \vartheta_\Omega(\omega)$, we have the next result.

Proposition 3.7. *If Ω is bounded and $\partial\Omega$ is smooth enough for Proposition 3.3 to hold, then*

$$x \in \Omega \Longrightarrow \vartheta_\Omega(\omega) = \vartheta_{\overline{\Omega}}(\omega), \quad \textit{for } W_x\textit{- almost every } \omega, \tag{3.25}$$

and

$$x \in \partial\Omega \Longrightarrow \vartheta_{\overline{\Omega}}(\omega) = 0, \quad \textit{for } W_x\textit{- almost every } \omega. \tag{3.26}$$

The probabilistic interpretation of this result is that, for any $x \in \overline{\Omega}$, once a Brownian path ω starting at x hits $\partial\Omega$, it penetrates into the interior of

$\mathbb{R}^n \setminus \overline{\Omega}$ within an arbitrarily short time, for W_x-almost all ω. From here one can show that, given $x \in \partial\Omega$, W_x-a.e. path ω spends a positive amount of time in both Ω and $\mathbb{R}^n \setminus \overline{\Omega}$, on any time interval $[0, s_0]$, for any $s_0 > 0$, however small. This is one manifestation of how wiggly Brownian paths are.

Note that taking $f = 1$ in (3.17) gives, for all $x \in \Omega$, any open set in $\mathbb{R}^n$,

$$e^{t\Delta}1(x) = W_x\big(\{\omega : \vartheta_\Omega(\omega) > t\}\big), \quad x \in \Omega, \tag{3.27}$$

the right side being the probability that a path starting in Ω at x has first exit time $> t$. Meanwhile, if $\partial\Omega$ is regular enough for Proposition 3.3 to hold, then

$$e^{t\Delta}1(x) = W_x\left(\{\omega : \vartheta_{\overline{\Omega}}(\omega) > t\}\right). \tag{3.28}$$

Comparing these identities extends Proposition 3.7 to unbounded Ω.

The following is an interesting consequence of (3.28).

Proposition 3.8. *For one-dimensional Brownian motion, starting at the origin, given $t > 0,\ \lambda > 0$,*

$$W\big(\{\omega : \sup_{0\le s\le t} \omega(s) \ge \lambda\}\big) = 2W\big(\{\omega : \omega(t) \ge \lambda\}\big). \tag{3.29}$$

Proof. The right side is $\int_\lambda^\infty p(t,x)\ dx$, with $p(t,x) = e^{td^2/dx^2}\delta(x) = (4\pi t)^{-1/2}e^{-x^2/4t}$, the $n = 1$ case of (1.5). The left side of (3.29) is the same as $W\big(\{\omega : \vartheta_{(-\infty,\lambda)}(\omega) < t\}\big)$, which by (3.28) is equal to $1 - e^{tL}1(0)$ if $L = d^2/dx^2$ on $(-\infty, \lambda)$, with Dirichlet boundary condition at $x = \lambda$. By the method of images we have, for $x < \lambda$,

$$e^{tL}1(x) = \int p(t,y)H(\lambda - x + y)\ dy,$$

where $H(s) = 1$ for $s > 0,\ -1$ for $s < 0$. From this, the identity (3.29) readily follows.

We next derive an expression for the Poisson integral formula, for the solution $\mathrm{PI}\, f = u$ to

$$\Delta u = 0 \text{ on } \Omega, \quad u|_{\partial\Omega} = f. \tag{3.30}$$

This can be expressed in terms of the integral kernel $G(x,y)$ of Δ^{-1} if $\partial\Omega$ is smooth. In fact, an application of Green's formula gives

$$\mathrm{PI}\, f(x) = \int_{\partial\Omega} f(y)\ \frac{\partial}{\partial\nu_y}G(x,y)\ dS(y), \tag{3.31}$$

where ν_y is the outward normal to $\partial\Omega$ at y. A closely related result is the following. Let f be defined and continuous on a neighborhood of $\partial\Omega$.

Given small $\delta > 0$, set

$$S_\delta = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) < \delta\}, \tag{3.32}$$

and define u_δ by

$$\begin{gathered} \Delta u_\delta = \delta^{-2} f_\delta \text{ on } \Omega, \quad u_\delta = 0 \text{ on } \partial\Omega, \\ f_\delta = f \text{ on } S_\delta, \ 0 \text{ on } \Omega \setminus S_\delta. \end{gathered} \tag{3.33}$$

Lemma 3.9. *If $\partial\Omega$ is smooth, then, locally uniformly on Ω,*

$$\lim_{\delta\to 0} u_\delta = -\frac{1}{2} PI f. \tag{3.34}$$

Proof. If ν is the outward normal, we have

$$\begin{aligned} u_\delta(x) &= \delta^{-2} \int\limits_{\partial\Omega} \int_0^\delta G(x, y - s\nu) f(y) \, ds \, dS(y) + o(1) \\ &= -\delta^{-2} \int\limits_{\partial\Omega} f(y) \frac{\partial}{\partial\nu_y} G(x,y) \Bigl(\int_0^\delta s \, ds \Bigr) \, dS(y) + o(1) \\ &= -\frac{1}{2} \int\limits_{\partial\Omega} f(y) \frac{\partial}{\partial\nu_y} G(x,y) \, dS(y) + o(1), \end{aligned} \tag{3.35}$$

so the result follows from (3.31).

Comparing this with (3.22), we conclude that when $\partial\Omega$ is smooth,

$$\operatorname{PI} f(x) = \lim_{\delta \searrow 0} \frac{2}{\delta^2} \, E_x \left(\int_0^{\vartheta_\Omega(\omega)} f\bigl(\omega(t)\bigr) \iota_{S_\delta}(\omega, t) \, dt \right), \tag{3.36}$$

where S_δ is as in (3.32), and, for $S \subset \Omega$,

$$\begin{aligned} \iota_S(\omega, t) = 1 &\quad \text{if } \omega(t) \in S, \\ 0 &\quad \text{otherwise.} \end{aligned} \tag{3.37}$$

We will discuss further formulas for PI f in §5.

Exercises

1. Looking at the definitions, check that $\psi_{\overline{\Omega}}(\omega, t)$ and $\vartheta_{\overline{\Omega}}(\omega)$ are measurable when $\Omega \subset \mathbb{R}^n$ is open with smooth boundary and that $\psi_\Omega(\omega, t)$ and $\vartheta_\Omega(\omega)$ are measurable, for general open $\Omega \subset \mathbb{R}^n$.
2. Show that if $x \in \mathcal{O}$, then

$$\{\omega \in \mathfrak{P}_0 : \vartheta_{\overline{\mathcal{O}}}(\omega) < t_0\} = \bigcup_{s \in [0, t_0) \cap \mathbb{Q}} \{\omega \in \mathfrak{P}_0 : \omega(s) \in \mathbb{R}^n \setminus \overline{\mathcal{O}}\}. \tag{3.38}$$

3. For any finite set $\mathcal{S} = \{s_1, \dots, s_K\} \subset \mathbb{Q}^+,\ N \in \mathbb{Z}^+$, set

$$F_{N,\mathcal{S}}(\omega) = \Phi_{N,\mathcal{S}}\Bigl(\omega(s_1)\dots\omega(s_K)\Bigr),$$

$$\Phi_{N,\mathcal{S}}(x_1,\dots,x_K) = \min\Bigl(N, \min\{s_\nu : x_\nu \in \mathbb{R}^n \setminus \overline{\mathcal{O}}\}\Bigr).$$

Show that, for any continuous path ω,

$$\vartheta_{\overline{\mathcal{O}}}(\omega) = \sup_N \inf_{\mathcal{S}} F_{N,\mathcal{S}}(\omega). \tag{3.39}$$

Note that the collection of such sets $\mathcal{S}$ is countable.

4. If $\mathfrak{P}_{\overline{\mathcal{O}},N} = \{\omega \in \mathfrak{P}_0 : \vartheta_{\overline{\mathcal{O}}}(\omega) \le N\}$ and $\mathcal{O}$ is bounded, show that

$$W_x\Bigl(\mathfrak{P}_0 \setminus \mathfrak{P}_{\overline{\mathcal{O}},N}\Bigr) \le CN^{-1}. \tag{3.40}$$

(*Hint*: Use (3.23).)

5. If $\omega \in \mathfrak{P}_{\overline{\mathcal{O}},N}$, show that

$$\vartheta_{\overline{\mathcal{O}}}(\omega) = \lim_{\nu\to\infty} \vartheta_{\nu,N}(\omega), \tag{3.41}$$

where

$$\vartheta_{\nu,N}(\omega) = \min\Bigl(N, \inf\Bigl\{s \in 2^{-\nu}\mathbb{Z}^+ : \omega(s) \notin \overline{\mathcal{O}}\Bigr\}\Bigr).$$

Write $\vartheta_{\nu,N}(\omega) = \Phi_{\nu,N}\Bigl(\omega(s_1),\dots,\omega(s_L)\Bigr)$, where $\Phi_{\nu,N}$ has a form similar to $\Phi_{N,\mathcal{S}}$ in Exercise 3.

6. For one-dimensional Brownian motion, establish the following, known as Kolmogorov's inequality:

$$W\Bigl(\{\omega : \sup_{0\le s\le t} |\omega(s)| \ge \varepsilon\}\Bigr) \le \frac{2t}{\varepsilon^2}, \quad \varepsilon > 0. \tag{3.42}$$

(*Hint*: Write the left side of (3.42) as $W\Bigl(\{\omega : \vartheta_{(-\varepsilon,\varepsilon)}(\omega) < t\}\Bigr)$, and relate this to the heat equation on $\Omega = [-\varepsilon, \varepsilon]$, with Dirichlet boundary condition, in a fashion parallel to the proof of Proposition 3.8.)
Note that this estimate is nontrivial only for $t < \varepsilon^2/2$. By Brownian scaling, it suffices to consider the case $\varepsilon = 1$. Compare the estimate

$$W\Bigl(\{\omega : \sup_{0\le s\le t} |\omega(s)| \ge \varepsilon\}\Bigr) \le 4\int_\varepsilon^\infty p(t,x)\, dx,$$

which follows from (3.29).

7. Given $\Omega \subset \mathbb{R}^n$ open, with complement K, and Δ with Dirichlet boundary condition on $\partial\Omega$, show that, for $x \in \Omega$,

$$W_x\Bigl(\{\omega : \vartheta_\Omega(\omega) = \infty\}\Bigr) = H_K(x), \tag{3.43}$$

where

$$H_K(t,x) = e^{t\Delta}1(x) \searrow H_K(x), \quad \text{as } t \nearrow \infty. \tag{3.44}$$

8. Suppose that $K = \mathbb{R}^n \setminus \Omega$ is compact, and suppose there exists $\widetilde{H}_K(x) \in C(\overline{\Omega})$, harmonic on Ω, such that $\widetilde{H}_K = 0$ on ∂K and $\widetilde{H}_K(x) \to 1$, as $|x| \to \infty$. Show

that

$$H_K(t,x) \geq \widetilde{H}_K(x), \quad \text{for all } t < \infty.$$

(*Hint*: Show that $\Delta H_K(t,x) \leq 0$ and that $H_K(t,x) \to 1$ as $|x| \to \infty$, and use the maximum principle.)

Deduce that if such $\widetilde{H}_K(x)$ exists, then $W_x\Big(\{\omega : \vartheta_\Omega(\omega) = \infty\}\Big) > 0$.

9. In the context of Exercise 8, show that if such $\widetilde{H}_K$ exists, then in fact

$$H_K(x) = \widetilde{H}_K(x), \quad \text{for all } x \in \overline{\Omega}. \tag{3.45}$$

(*Hint*: Show that H_K must be harmonic in Ω and that $\limsup_{|x|\to\infty} H_K(x) \leq 1$.)

By explicit construction, produce such a function on $\mathbb{R}^n \setminus B$ when B is a ball of radius $a > 0$, provided $n \geq 3$.

10. Using Exercises 7–9, show that when $n \geq 3$,

$$W_x\Big(\{\omega : |\omega(t)| \to \infty \text{ as } t \to \infty\}\Big) = 1. \tag{3.46}$$

(*Hint*: Given $R > 0$, the probability that $|\omega(t)| \geq R$ for some t is 1. If $R >> a$, and $|\omega(t_0)| \geq R$, show that the probability that $|\omega(t_0 + s)| \leq a$ for some $s > 0$ is small, using (3.43) for $K = B_a = \{x : |x| \leq a\}$.) To restate (3.46), one says that Brownian motion in $\mathbb{R}^n$ is "non-recurrent," for $n \geq 3$.

11. If $n \leq 2$ and $K = B_a$, show that $H_K(t,x) = 0$ in (3.44), and hence the probability defined in (3.46) is zero. Deduce that if $n \leq 2$ and $U \subset \mathbb{R}^n$ is a nonempty open set, almost every Brownian path ω visits U at an infinite sequence of times $t_\nu \to \infty$.

One says that Brownian motion in $\mathbb{R}^n$ is "recurrent," for $n \leq 2$.

12. Relate the formula (3.34) for PI f to representations of PI f by double-layer potentials, discussed in §11 of Chapter 7. Where is the second layer coming from?

13. If Ω is a bounded domain with smooth boundary, show that (3.36) remains true with S_δ replaced by

$$\widetilde{S}_\delta = \{x \in \mathbb{R}^n \setminus \overline{\Omega} : \text{dist}(x, \partial\Omega) < \delta\}$$

and with $\vartheta_\Omega(\omega)$ replaced by $\vartheta_{\Omega_\delta}(\omega)$, where $\Omega_\delta = \overline{\Omega} \cup \widetilde{S}_\delta$. (*Hint*: Start by showing that $\widetilde{u}_\delta(x) \to -(1/2)\text{PI}\, f(x)$, for $x \in \Omega$, where, in place of (3.33),

$$\Delta \widetilde{u}_\delta = \delta^{-2} f_\delta \text{ on } \Omega_\delta, \quad \widetilde{u}_\delta = 0 \text{ on } \partial\Omega_\delta,$$

with $f_\delta = f$ on $\widetilde{S}_\delta$, 0 on $\overline{\Omega}$.

4. Martingales, stopping times, and the strong Markov property

Given $t \in [0,\infty)$, let $\mathfrak{B}_t$ be the σ-field of subsets of $\mathfrak{P}_0$ generated by sets of the form

$$\{\omega \in \mathfrak{P}_0 : \omega(s) \in E\}, \tag{4.1}$$

where $s \in [0,t]$ and E is a Borel subset of $\mathbb{R}^n$. One easily sees that each element of $\mathfrak{B}_t$ is a Borel set in $\mathfrak{P}$. As t increases, $\mathfrak{B}_t$ is an increasing family of σ-fields, each consisting of sets which are W_x-measurable, for all $x \in \mathbb{R}^n$. Set $\mathfrak{B}_\infty = \sigma\big(\bigcup_{t<\infty} \mathfrak{B}_t\big)$.

Given $f \in L^1(\mathfrak{P}_0, \mathfrak{B}_\infty, dW_x)$, we can define the conditional expectation

$$E_x\big(f\big|\mathfrak{B}_t\big), \tag{4.2}$$

a function measurable with respect to $\mathfrak{B}_t$, as follows. Denote by $W_{x,t}$ the restriction of the Wiener measure W_x to the σ-field $\mathfrak{B}_t$. Then

$$\lambda(S) = \int_S f(\omega)\; dW_x(\omega) = E_x(f\chi_S) \tag{4.3}$$

defines a countably additive set function on $\mathfrak{B}_t$, which is absolutely continuous with respect to $W_{x,t}$, so by the Radon-Nikodym theorem there exists a $\mathfrak{B}_t$-measurable function Φ_t, uniquely defined $W_{x,t}$-almost everywhere, such that (4.3) is equal to $\int_S \Phi_t(\omega)\; dW_{x,t}(\omega)$, for all $S \in \mathfrak{B}_t$. This function is $E_x(f|\mathfrak{B}_t)$. Clearly,

$$f \in L^1(\mathfrak{P}_0, \mathfrak{B}_\infty, dW_x) \Longrightarrow E_x\big(f\big|\mathfrak{B}_t\big) \in L^1(\mathfrak{P}_0, \mathfrak{B}_t, dW_{x,t}). \tag{4.4}$$

This construction of conditional expectation generalizes in the obvious way to any situation where f is measurable with respect to some σ-field $\mathfrak{F}$, and is L^1 with respect to a given probability measure on $\mathfrak{F}$, and one wants to define the conditional expectation $E(f|\mathfrak{F}_0)$ with respect to some sub-σ-field $\mathfrak{F}_0$ of $\mathfrak{F}$.

Note that we can regard $L^1(\mathfrak{P}_0, \mathfrak{B}_t, dW_{x,t})$ naturally as a closed linear subspace of $L^1(\mathfrak{P}_0, \mathfrak{B}_\infty, dW_x)$. Then the map $f \mapsto E_x(f|\mathfrak{B}_t)$ is a projection. Similarly, we have

$$f \in L^2(\mathfrak{P}_0, \mathfrak{B}_\infty, dW_x) \Longrightarrow E_x\big(f\big|\mathfrak{B}_t\big) \in L^2(\mathfrak{P}_0, \mathfrak{B}_t, dW_{x,t}),$$

and in this case $E_x(f|\mathfrak{B}_t)$ is simply the orthogonal projection of f onto $L^2(\mathfrak{P}_0, \mathfrak{B}_t, dW_{x,t})$, regarded as a linear subspace of $L^2(\mathfrak{P}_0, \mathfrak{B}_\infty, dW_x)$. The reader might think of this in light of von Neumann's proof of the Radon-Nikodym theorem, which is sketched in the exercises for §2 of Appendix A.

The following is a statement that Brownian motion possesses the *Markov property.*

Proposition 4.1. *Given* $s, t > 0,\ f \in C(\dot{\mathbb{R}}^n)$,

$$E_x\big(f(\omega(t+s))\big|\mathfrak{B}_s\big) = E_{\omega(s)}\big(f(\omega(t))\big), \quad \textit{for } W_x\textit{-almost all } \omega. \tag{4.5}$$

Proof. The right side of (4.5) is $\mathfrak{B}_s$-measurable, so the identity is equivalent to the statement that

$$\int_S f(\omega(t+s))\,dW_x(\omega) = \int_S \Bigl(\int f(\widetilde{\omega}(t))\,dW_{\omega(s)}(\widetilde{\omega})\Bigr)\,dW_x(\omega), \tag{4.6}$$

for all $S \in \mathfrak{B}_s$. It suffices to verify (4.6) for all S of the form

$$S = \{\omega \in \mathfrak{P}_0 : \omega(t_1) \in E_1, \dots, \omega(t_K) \in E_K\},$$

given $t_j \in [0,s]$, E_j Borel sets in $\mathbb{R}^n$. For such S, (4.6) follows directly from the characterization of the Wiener integral given in §1, that is, from (1.6)–(1.9) in the case $x = 0$, together with the identity

$$\int f(\widetilde{\omega}(t))\,dW_y(\widetilde{\omega}) = E\bigl(f(y+\omega(t))\bigr) \tag{4.7}$$

used to define (1.36).

We can easily extend (4.5) to

$$E_x\bigl(F(\omega(s+t_1),\dots,\omega(s+t_k))\big|\mathfrak{B}_s\bigr) = E_{\omega(s)}\bigl(F(\omega(t_1),\dots,\omega(t_k))\bigr), \tag{4.8}$$

for W_x-almost all ω, given $t_1,\dots,t_k > 0$, and F continuous on $\prod_1^k \dot{\mathbb{R}}^n$, as in (1.8). Also, standard limiting arguments allow us to enlarge the class of functions F for which this works. We then get the following more definitive statement of the Markov property.

Proposition 4.2. *For $s > 0$, define the map*

$$\sigma_s : \mathfrak{P}_0 \longrightarrow \mathfrak{P}_0, \quad (\sigma_s\omega)(t) = \omega(t+s). \tag{4.9}$$

Then, given φ bounded and $\mathfrak{B}_\infty$-measurable, we have

$$E_x\bigl(\varphi \circ \sigma_s\big|\mathfrak{B}_s\bigr) = E_{\omega(s)}(\varphi), \quad \text{for } W_x\text{-almost all } \omega. \tag{4.10}$$

The following is a useful restatement of Proposition 4.2.

Corollary 4.3. *For $s > 0$, define the map*

$$\vartheta_s : \mathfrak{P}_0 \to \mathfrak{P}_0, \quad (\vartheta_s\omega)(t) = \omega(t+s) - \omega(s). \tag{4.11}$$

Then, given $\varphi \in L^1(\mathfrak{P}_0, dW_0)$, we have

$$E_x\bigl(\varphi \circ \vartheta_s\big|\mathfrak{B}_s\bigr) = E_0(\varphi). \tag{4.12}$$

In particular,

$$E_x\bigl(f\bigl(\vartheta_s\omega(t)\bigr)\big|\mathfrak{B}_s\bigr) = E_0\bigl(f(\omega(t))\bigr). \tag{4.13}$$

Note that (4.12) implies ϑ_s is measure preserving, in the sense that

$$W_x\bigl(\vartheta_s^{-1}(S)\bigr) = W_0(S), \tag{4.14}$$

for W_0-measurable sets S. The map ϑ_s is not one-to-one, of course, but it is *onto* the set of paths in $\mathfrak{P}_0$ satisfying $\omega(0) = 0$.

The Markov property also implies certain independence properties. A function $\varphi \in L^1(\mathfrak{P}_0, dW_x)$ is said to be independent of the σ-algebra $\mathfrak{B}_t$ provided that, for all continuous F,

$$\int_S F\big(\varphi(\omega)\big)\, dW_x(\omega) = W_x(S) E_x(F \circ \varphi), \quad \forall\, S \in \mathfrak{B}_t. \tag{4.15}$$

An equivalent condition is

$$E_x\big(F(\varphi)\psi\big) = E_x\big(F(\varphi)\big) E_x(\psi), \quad \forall\, \psi \in L^1(\mathfrak{P}_0, \mathfrak{B}_t, dW_x), \tag{4.16}$$

given $F(\varphi)\psi \in L^1(\mathfrak{P}_0, dW_x)$, and another equivalent condition is

$$E_x\big(F(\varphi)|\mathfrak{B}_t\big) = E_x\big(F(\varphi)\big). \tag{4.17}$$

In turn, this identity holds whenever the left side is constant. From Corollary 4.3 we deduce:

Corollary 4.4. *For* $s \ge 0$, $\vartheta_s\omega(t) = \omega(t+s) - \omega(s)$ *is independent of* $\mathfrak{B}_s$.

Proof. By (4.13),

$$E_x\big(F(\omega(t+s) - \omega(s))\big|\mathfrak{B}_s\big) = E_0\big(F(\omega(t))\big), \tag{4.18}$$

which is constant.

The Markov property gives rise to martingales. By definition (valid in general for an increasing family $\mathfrak{B}_t$ of σ-fields), a *martingale* is a family $F_t \in L^1(\mathfrak{P}_0, \mathfrak{B}_t, dW_{x,t})$ such that

$$E_x\big(F_t\big|\mathfrak{B}_s\big) = F_s \quad \text{when } s < t. \tag{4.19}$$

If $E_x(F_t|\mathfrak{B}_s) \ge F_s$ for $s < t$, $\{F_t\}$ is called a *submartingale* over $\mathfrak{B}_t$. The following is a very useful class of martingales.

Proposition 4.5. *Let* $h(t,x)$ *be smooth in* $t \ge 0, x \in \mathbb{R}^n$, *and satisfy* $|h(t,x)| \le C_\varepsilon e^{\varepsilon|x|^2}$ *for all* $\varepsilon > 0$, *and the backward heat equation*

$$\frac{\partial h}{\partial t} = -\Delta h. \tag{4.20}$$

Then $\mathfrak{h}_t(\omega) = h(t, \omega(t))$ *is a martingale over* $\mathfrak{B}_t$.

Proof. The hypothesis on $h(t,x)$ implies that, for $t, s > 0$,

$$h(s,x) = \int p(t,y) h(t+s, x-y)\, dy, \tag{4.21}$$

where $p(t,x) = e^{t\Delta}\delta(x)$ is given by (1.5). Now

$$\begin{aligned} E_x\big(\mathfrak{h}_{t+s}\big|\mathfrak{B}_s\big) &= E_x\big(h(t+s,\omega(t+s))\big|\mathfrak{B}_s\big) \\ &= E_{\omega(s)}\big(h(t+s,\omega(t))\big), \end{aligned} \tag{4.22}$$

for W_x-almost all ω, by (4.5). This is equal to

$$\int p(t,y-\omega(s))\ h(t+s,y)\ dy, \tag{4.23}$$

by the characterization (1.9) of expectation, adjusted as in (1.36), and by (4.21) this is equal to $h(s,\omega(s)) = \mathfrak{h}_s(\omega)$.

Corollary 4.6. *For one-dimensional Brownian motion, the following are martingales over* $\mathfrak{B}_t$*:*

$$\mathfrak{x}_t(\omega) = \omega(t), \quad \mathfrak{q}_t(\omega) = \omega(t)^2 - 2t, \quad \mathfrak{z}_t(\omega) = e^{a\omega(t)-a^2t}, \tag{4.24}$$

given $a > 0$.

One important property of martingales is the following martingale maximal inequality.

Proposition 4.7. *If* F_t *is a martingale over* $\mathfrak{B}_t$*, then, given any countable set* $\{t_j\} \subset \mathbb{R}^+$*, the "maximal function"*

$$F^*(\omega) = \sup_j\ F_{t_j}(\omega) \tag{4.25}$$

satisfies, for all $\lambda > 0$,

$$W_x\big(\{\omega : F^*(\omega) > \lambda\}\big) \le \frac{1}{\lambda}\|F_t\|_{L^1(\mathfrak{P}_0,dW_x)}. \tag{4.26}$$

Of course, the assumption that F_t is a martingale implies that $\|F_t\|_{L^1}$ is independent of t.

Proof. It suffices to demonstrate this for an arbitrary finite subset $\{t_j\}$ of $\mathbb{R}^+$. Thus we can work with $f_j(\omega) = F_{t_j}(\omega), \mathfrak{B}_j = \mathfrak{B}_{t_j}, 1 \le j \le N$, and take $t_1 < t_2 < \cdots < t_N$, and the martingale hypothesis is that $E_x(f_k\big|\mathfrak{B}_j) = f_j$ when $j < k$. There is no loss in assuming $f_N(\omega) \ge 0$, so all $f_j(\omega) \ge 0$. Now consider

$$S_\lambda = \{\omega : f^*(\omega) > \lambda\} = \{\omega : \text{some } f_j(\omega) > \lambda\}. \tag{4.27}$$

There is a pairwise-disjoint decomposition

$$S_\lambda = \bigcup_{j=1}^N S_{\lambda j}, \quad S_{\lambda j} = \{\omega : f_j(\omega) > \lambda \text{ but } f_\ell(\omega) \le \lambda \text{ for } \ell < j\}. \tag{4.28}$$

Note that $S_{\lambda j}$ is $\mathfrak{B}_j$-measurable. Consequently, we have

$$
\begin{aligned}
&\int_{S_\lambda} f_N(\omega)\, dW_x(\omega) \\
&= \sum_{j=1}^{N} \int_{S_{\lambda j}} f_N(\omega)\, dW_x(\omega) = \sum_{j=1}^{N} \int_{S_{\lambda j}} f_j(\omega)\, dW_x(\omega) \\
&\geq \sum_{j=1}^{N} \lambda\, W_x\big(S_{\lambda j}\big) = \lambda\, W_x(S_\lambda).
\end{aligned} \tag{4.29}
$$

This yields (4.26), in this special case, and the proposition is hence proved.

Applying the martingale maximal inequality to $\mathfrak{z}_t(\omega) = e^{a\omega(t)-a^2 t}$, we obtain the following.

Corollary 4.8. *For one-dimensional Brownian motion, given* $t > 0$,

$$
W_0\big(\{\omega \in \mathfrak{P}_0 : \sup_{0\le s\le t} \omega(s) - as > \lambda\}\big) \le e^{-a\lambda}. \tag{4.30}
$$

Proof. The set whose measure is estimated in (4.30) is

$$
\{\omega \in \mathfrak{P}_0 : \sup_{0\le s\le t} e^{a\omega(s)-a^2 s} > e^{a\lambda}\}.
$$

Since paths in $\mathfrak{P}_0$ are continuous, one can take the sup over $[0,t]\cap\mathbb{Q}$, which is countable, so (4.26) applies. Note that $E_0(\mathfrak{z}_t) = 1$.

We turn to a discussion of the *strong* Markov property of Brownian motion. For this, we need the notion of a stopping time. A function τ on $\mathfrak{P}_0$ with values in $[0,+\infty]$ is called a *stopping time* provided that, for each $t \ge 0$, $\{\omega \in \mathfrak{P}_0 : \tau(\omega) < t\}$ belongs to the σ-field $\mathfrak{B}_t$. It follows from (3.39) that $\vartheta_{\overline{\mathcal{O}}}$ is a stopping time. So is $\vartheta_{\mathcal{O}}$.

Given a stopping time τ, define $\mathfrak{B}_{\tau+}$ to be the σ-algebra of sets $S \in \mathfrak{B}_\infty$ such that $S \cap \{\omega : \tau(\omega) < t\}$ belongs to $\mathfrak{B}_t$ for each $t \ge 0$. Note that τ is measurable with respect to $\mathfrak{B}_{\tau+}$. The hypothesis that τ is a stopping time means precisely that the whole set $\mathfrak{P}_0$ satisfies the criteria for membership in $\mathfrak{B}_{\tau+}$. We note that any $t \in [0,\infty)$, regarded as a constant function on $\mathfrak{P}_0$, is a stopping time and that, in this case, $\mathfrak{B}_{t+} = \bigcap_{s>t} \mathfrak{B}_s$.

The following analogue of Propositions 4.1 and 4.2 is one statement of the strong Markov property.

Proposition 4.9. *If* τ *is a stopping time such that* $\tau(\omega) < \infty$ *for* W_x-*almost all* ω, *and if* $t > 0$, *then*

$$
E_x\Big(f\big(\omega(\tau+t)\big)\big|\mathfrak{B}_{\tau+}\Big) = E_{\omega(\tau)}\big(f(\omega(t))\big), \tag{4.31}
$$

for W_x-almost all ω. More generally, with

$$(\sigma_\tau \omega)(t) = \omega(t+\tau),$$

and φ bounded and $\mathfrak{B}_\infty$-measurable, we have

$$E_x\big(\varphi \circ \sigma_\tau \big| \mathfrak{B}_{\tau+}\big) = E_{\omega(\tau)}(\varphi), \tag{4.32}$$

for W_x-almost all ω.

As in (4.6), the content of (4.31) is that

$$\int_S f\big(\omega(\tau+t)\big)\, dW_x(\omega) = \int_S \Big(\int f\big(\omega^\#(t)\big)\, dW_{\omega(\tau)}(\omega^\#)\Big)\, dW_x(\omega), \tag{4.33}$$

given $S \in \mathfrak{B}_{\tau+}$. In other words, given that $S \cap \{\omega : \tau(\omega) < t'\} \in \mathfrak{B}_{t'}$, for each $t' \geq 0$. There is no loss in taking $x = 0$, and we can rewrite (4.33) as

$$\int_S f\big(\omega(\tau+t)\big)\, dW(\omega) = \int_S \int f\big(\omega^\#(t) + \omega(\tau)\big)\, dW(\omega^\#)\, dW(\omega). \tag{4.34}$$

It is useful to approximate τ by discretization:

$$\tau_\nu(\omega) = 2^{-\nu} k, \ \text{ if } 2^{-\nu}(k-1) \leq \tau(\omega) < 2^{-\nu} k. \tag{4.35}$$

Thus

$$\{\omega : \tau_\nu(\omega) < t\} = \{\omega : \tau(\omega) < 2^{-\nu} k\} \in \mathfrak{B}_t, \tag{4.36}$$

so each τ_ν is a stopping time. Note that

$$\begin{aligned} A_{\nu k} &= \{\omega : \tau_\nu(\omega) = 2^{-\nu} k\} \\ &= \{\omega : \tau(\omega) < 2^{-\nu} k\} \setminus \{\omega : \tau(\omega) < 2^{-\nu}(k-1)\} \end{aligned} \tag{4.37}$$

belongs to $\mathfrak{B}_{2^{-\nu}k}$.

If τ is replaced by τ_ν, the left side of (4.34) becomes

$$\sum_{\nu,k} \int_{S \cap A_{\nu k}} f\big(\omega(t + 2^{-\nu} k)\big)\, dW(\omega), \tag{4.38}$$

and the right side of (4.34) becomes

$$\sum_{\nu,k} \int_{S \cap A_{\nu k}} \int f\big(\omega^\#(t) + \omega(2^{-\nu} k)\big)\, dW(\omega^\#)\, dW(\omega). \tag{4.39}$$

Note that if $S \in \mathfrak{B}_{\tau+}$, then $S \cap A_{\nu k} \in \mathfrak{B}_{2^{-\nu}k}$. Thus, the fact that each term in the sum (4.38) is equal to the corresponding term in (4.39) follows from (4.6). Consequently, we have

$$\int_S f\big(\omega(\tau_\nu + t)\big)\, dW(\omega) = \int_S \int f\big(\omega^\#(t) + \omega(\tau_\nu)\big)\, dW(\omega^\#)\, dW(\omega), \tag{4.40}$$

for all ν, if $S \in \mathfrak{B}_{\tau+}$. The desired identity (4.34) follows by taking $\nu \to \infty$, if $f \in C(\dot{\mathbb{R}}^n)$. Passing from this to (4.32) is then done as in the proof of Proposition 4.2.

In particular, the extension of (4.31) analogous to (4.8), in the special case $F(x_1, x_2) = f(x_2 - x_1)$, yields the identity

$$
\begin{aligned}
\int_S f\big(\omega(\tau+t) - \omega(\tau)\big)\, dW(\omega) &= \int_S \int f\big(\omega^\#(t)\big)\, dW(\omega^\#)\, dW(\omega) \\
&= E\big(f(\omega(t))\big) \cdot W(S),
\end{aligned}
\tag{4.41}
$$

given $S \in \mathfrak{B}_{\tau+}$. This, together with the extension to $F(x_1, \dots x_K)$, says that $\omega(\tau+t) - \omega(\tau) = \beta(t)$ has the probability distribution of a Brownian motion, independent of $\mathfrak{B}_{\tau+}$. This is a common form in which the strong Markov property is stated.

It is sometimes useful to consider stopping times for which $\{\omega : \tau(\omega) = \infty\}$ has positive measure. In such a case, the extension of Proposition 4.9 is that (4.32) holds for W_x-almost ω in the set $\{\omega : \tau(\omega) < \infty\}$. Thus, for example, (4.33) and (4.34) hold, given $S \in \mathfrak{B}_{\tau+}$ and $S \subset \{\omega : \tau(\omega) < \infty\}$.

We next look at some operator-theoretic properties of

$$
\begin{aligned}
Q_t &: L^2(\mathfrak{P}_0, dW_0) \to L^2(\mathfrak{P}_0, dW_0), \quad Q_t\varphi = E_0(\varphi|\mathfrak{B}_t), \\
\Theta_t &: L^2(\mathfrak{P}_0, dW_0) \to L^2(\mathfrak{P}_0, dW_0), \quad \Theta_t\varphi(\omega) = \varphi(\vartheta_t\omega),
\end{aligned}
\tag{4.42}
$$

where ϑ_t is given by (4.11). For each $t \ge 0$, Q_t is an orthogonal projection, and $Q_sQ_t = Q_tQ_s = Q_s$, for $s \le t$. Note that (4.13) implies

$$
Q_t\Theta_t = Q_0, \tag{4.43}
$$

since Q_0 is the orthogonal projection of $L^2(\mathfrak{P}_0, dW_0)$ onto

$$
\mathcal{R}(Q_0) = \text{ set of constant functions.} \tag{4.44}
$$

Proposition 4.10. *The family* $\Theta_t,\ t \in [0, \infty)$, *is a strongly continuous semigroup of isometries of* $L^2(\mathfrak{P}_0, dW_0)$, *with*

$$
\mathcal{R}(\Theta_t) \subset \ \mathit{Ker}(Q_t - Q_0) = \{\varphi : E_0(\varphi|\mathfrak{B}_t) = \ \mathit{const.}\}. \tag{4.45}
$$

Proof. That Θ_t is an isometry follows from the measure-preserving property (4.14). If we apply Q_0 to (4.43), we get $Q_0\Theta_t = Q_0$; hence $(Q_t - Q_0)\Theta_t = 0$, which yields (4.45).

The semigroup property follows from a straightforward calculation:

$$
\vartheta_\sigma\vartheta_s\omega = \vartheta_{\sigma+s}\omega \Longrightarrow \Theta_{s+\sigma} = \Theta_s\Theta_\sigma. \tag{4.46}
$$

The convergence

$$
\Theta_s\varphi \to \Theta_t\varphi \quad \text{in } L^2(\mathfrak{P}_0, dW_0), \text{ as } s \to t, \tag{4.47}
$$

is easy to demonstrate for $\varphi(\omega)$ of the form (1.8), that is,

$$\varphi(\omega) = f\big(\omega(t_1), \dots, \omega(t_k)\big), \tag{4.48}$$

with f continuous on $\dot{\mathbb{R}}^n \times \cdots \times \dot{\mathbb{R}}^n$ (k factors). In fact, $\varphi(\vartheta_s(\omega)) = \varphi_s(\omega) \to \varphi_t(\omega)$ boundedly and pointwise on $\mathfrak{P}_0$ for such φ. Since the set of φ of the form (4.48) is dense in $L^2(\mathfrak{P}_0, dW_0)$, (4.47) follows.

Proposition 4.11. *The family of orthogonal projections Q_t is strongly continuous in $t \in [0, \infty)$.*

Proof. It is easy to verify that, for any $\varphi \in L^2(\mathfrak{P}_0, dW_0)$,

$$Q_s\varphi \to Q_{t-}\varphi = E_0(\varphi|\mathfrak{B}_{t-}), \quad \text{as } s \nearrow t, \tag{4.49}$$

provided $t > 0$, and

$$Q_s\varphi \to Q_{t+}\varphi = E_0(\varphi|\mathfrak{B}_{t+}), \quad \text{as } s \searrow t, \tag{4.50}$$

where

$$\mathfrak{B}_{t-} = \sigma\Big(\bigcup_{s<t} \mathfrak{B}_s\Big), \quad \mathfrak{B}_{t+} = \bigcap_{s>t} \mathfrak{B}_s. \tag{4.51}$$

It is also easy to verify that $\mathfrak{B}_{t-} = \mathfrak{B}_t$, for $t > 0$, so $Q_s\varphi \to Q_t\varphi$ as $s \nearrow t$. On the other hand, it is not true that $\mathfrak{B}_{t+} = \mathfrak{B}_t$, so the continuity of $Q_t\varphi$ from above requires more work.

Suppose $t_j \in \mathbb{Q}^+$ and

$$0 \le t_1 < t_2 < \cdots < t_\ell \le t < t_{\ell+1} < \cdots < t_{\ell+k}. \tag{4.52}$$

Let $f_j \in C\big(\dot{\mathbb{R}}^n\big)$. Consider any function on $\mathfrak{P}$ of the form

$$\begin{aligned} \varphi(\omega) &= A_\ell(\omega) B_{k\ell}(\omega) \\ &= f_1\big(\omega(t_1)\big) \cdots f_\ell\big(\omega(t_\ell)\big) \cdot f_{\ell+1}\big(\omega(t_{\ell+1})\big) \cdots f_{\ell+k}\big(\omega(t_{\ell+k})\big). \end{aligned} \tag{4.53}$$

Denote by $\mathcal{C}^\wedge$ the linear span of the set of such functions. For φ of the form (4.53), we have

$$E_0(\varphi|\mathfrak{B}_t) = A_\ell(\omega) E_0(B_{k\ell}|\mathfrak{B}_t). \tag{4.54}$$

If $t_{\ell+\nu} = t + s_\nu,\ 1 \le \nu \le k$, we have, by (4.8),

$$E_0(B_{k\ell}|\mathfrak{B}_t) = E_{\omega(t)}\big(f_{\ell+1}(\omega(s_1)) \cdots f_{\ell+k}(\omega(s_k))\big), \quad \text{a.e. on } \mathfrak{P}_0. \tag{4.55}$$

Now, if $t \le t + h < t_{\ell+1}$, we also have

$$\begin{aligned} E_0(\varphi|\mathfrak{B}_{t+h}) &= A_\ell(\omega) E_0(B_{k\ell}|\mathfrak{B}_{t+h}) \\ &= A_\ell(\omega) E_{\omega(t+h)}(\psi_\ell), \end{aligned} \tag{4.56}$$

where

$$\psi_\ell(\omega) = f_{\ell+1}\big(\omega(s_1 - h)\big) \cdots f_{\ell+k}\big(\omega(s_k - h)\big). \tag{4.57}$$

Now, as in (1.9),

$$E_x(\psi_\ell) = \int \cdots \int p(s_1 - h, x_1)p(s_2 - s_1, x_2 - x_1) \cdots p(s_k - s_{k-1}, x_k - x_{k-1}) \cdot f_{\ell+1}(x_1 + x) \cdots f_{\ell+k}(x_k + x)\, dx_k \cdots dx_1. \tag{4.58}$$

The continuity in (x, h) is clear. Since paths in $\mathfrak{P}_0$ are continuous, we have, by linearity, that

$$\varphi \in \mathcal{C}\hat{} \Longrightarrow E_0(\varphi|\mathfrak{B}_t) = \lim_{h\searrow 0} E_0(\varphi|\mathfrak{B}_{t+h}), \quad W_0\text{-a.e.} \tag{4.59}$$

Now the Stone-Weierstrass theorem implies that $\mathcal{C}\hat{}$ is dense in $C(\mathfrak{P})$, which is dense in $L^2(\mathfrak{P}, dW_0) = L^2(\mathfrak{P}_0, dW_0)$. Thus we have

$$E_0(\varphi|\mathfrak{B}_{t+}) = E_0(\varphi|\mathfrak{B}_t), \quad W_0\text{-a.e.}, \tag{4.60}$$

for every $\varphi \in L^2(\mathfrak{P}_0, dW_0)$, and the proposition is proved.

Exercises

1. Show that the martingale maximal inequality applied to $\mathfrak{x}_t(\omega) = \omega(t)$ yields
$$W_0\Big(\Big\{\omega \in \mathfrak{P}_0 : \sup_{0\le s\le t} \omega(s) > b\sqrt{4t/\pi}\Big\}\Big) \le \frac{1}{b}.$$
Compare with the precise result in (3.29).
2. With $\mathfrak{B}_{t-}$ characterized by (4.51), show that $\mathfrak{B}_{t-} = \mathfrak{B}_t$, as stated in the proof of Proposition 4.11. (*Hint*: In the characterization (4.1) of $\mathfrak{B}_t$, one can restrict attention to E *open* in $\mathbb{R}^n$.)
3. Using (4.60), show that
$$S \in \mathfrak{B}_{0+} \Longrightarrow W_0(S) = 0 \text{ or } 1.$$
This is called Blumenthal's 01 law. If $E \in \mathbb{R}^n$ is a closed set, show that
$$\{\omega \in \mathfrak{P}_0 : \omega(t_\nu) \in E \text{ for some } t_\nu \searrow 0\}$$
is a set in $\mathfrak{B}_{0+}$. (*Hint*: Consider $\{\omega \in \mathfrak{P}_0 : \text{dist}(\omega(t), E) \ge \delta > 0 \text{ for } t \in [2^{-\nu}\varepsilon, \varepsilon] \cap \mathbb{Q}\} = S(E, \delta, \varepsilon, \nu)$.)
4. Let $\mathcal{N}$ be the collection of (W_0-outer measurable) subsets of $\mathfrak{P}_0$ with W_0-measure zero. Form the family of σ-algebras $\mathfrak{B}_t^{\#} = \mathfrak{B}_t \cup \mathcal{N}$, called the *augmentation* of $\mathfrak{B}_t$. Show that $\mathfrak{B}_t^{\#} \supset \mathfrak{B}_{t+}$ and, with notation parallel to (4.51),
$$\mathfrak{B}_{t-}^{\#} = \mathfrak{B}_t^{\#} = \mathfrak{B}_{t+}^{\#}.$$
Note: The augmentation of $\mathfrak{B}_t$ is bigger than the completion of $\mathfrak{B}_t$.
5. Let $\widetilde{\mathfrak{F}}_t$ be the σ-algebra of subsets of $\mathfrak{P}_0$ generated by sets of the form (4.1) for $s \ge t$, and set $\mathcal{A}_\infty = \bigcap_{t>0} \widetilde{\mathfrak{F}}_t$. Using Blumenthal's 01 law and Exercise 2 of §1, show that
$$S \in \mathcal{A}_\infty \Longrightarrow W_0(S) = 0 \text{ or } 1.$$

If $E \subset \mathbb{R}^n$ is a closed set, show that

$$\{\omega \in \mathfrak{P}_0 : \omega(t_\nu) \in E \text{ for some } t_\nu \nearrow \infty\}$$

is a set in $\mathcal{A}_\infty$.

5. First exit time and the Poisson integral

At the end of §3 we produced a formula for $\mathrm{PI}\, f$, giving the solution u to

$$\Delta u = 0 \text{ in } \Omega, \quad u = f \text{ on } \partial\Omega, \tag{5.1}$$

at least in case Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary. Here we produce a formula that is somewhat neater than (3.36) and that is also amenable to extension to general bounded, open $\Omega \subset \mathbb{R}^n$, with no smoothness assumed on $\partial\Omega$. In the smooth case, the formula is

$$\mathrm{PI}\, f(x) = E_x\big(f(\omega(\vartheta_{\overline{\Omega}}))\big), \quad x \in \overline{\Omega}, \tag{5.2}$$

where $\vartheta_{\overline{\Omega}}(\omega)$ is the first exit time defined by (3.24).

From an intuitive point of view, the formula (5.2) has a very easy and natural justification. To show that the right side of (5.2), which we denote by $u(x)$, is harmonic on Ω, it suffices to verify the mean-value property. Let $x \in \Omega$ be the center of a closed ball $B \subset \Omega$. We claim that $u(x)$ is equal to the mean value of $u|_{\partial B}$. Indeed, a continuous path ω starting from x and reaching $\partial\Omega$ must cross ∂B, say at a point $y = \omega(\vartheta_B)$. The future behavior of such paths is independent of their past, so the probability distribution of the first contact point $\omega(\vartheta_{\overline{\Omega}})$, when averaged over starting points in ∂B, should certainly coincide with the probability distribution of such a first contact point in $\partial\Omega$, for paths starting at x (the distribution of whose first contact point with ∂B must be constant, by symmetry).

The key to converting this into a mathematical argument is to note that the time $\vartheta_B(\omega)$ is not constant, so one needs to make use of the strong Markov property as a tool to establish the mean-value property of the function $u(x)$ defined by the right side of (5.2).

Let us first make some comments on the right side $u(x)$ of (5.2). By (3.40) we have

$$\Big| u(x) - \int\limits_{\mathfrak{P}_{\overline{\Omega},N}} f\big(\omega(\vartheta_{\overline{\Omega}})\big)\, dW_x(\omega) \Big| \le C \|f\|_{L^\infty(\partial\Omega)}\, N^{-1}. \tag{5.3}$$

Let us extend $f \in C(\partial\Omega)$ to an element $f \in C_0(\mathbb{R}^n)$, without increasing the sup norm. By (3.41), we have

$$f\big(\omega(\vartheta_{\overline{\Omega}})\big) = \lim_{\nu\to\infty} f\big(\omega(\vartheta_{\nu,N}(\omega))\big), \quad \text{for } \omega \in \mathfrak{P}_{\overline{\Omega},N}, \tag{5.4}$$

where $\vartheta_{\nu,N}(\omega) = \min\big(N, \inf\{s \in 2^{-\nu}\mathbb{Z}^+ : \omega(s) \notin \overline{\Omega}\}\big)$. Thus, if the integral in (5.3) is denoted by $u_N(x)$, then

$$u_N(x) = \lim_{\nu\to\infty} u_{N\nu}(x) = \lim_{\nu\to\infty} \int\limits_{\mathfrak{P}_{\overline{\Omega},N}} f\big(\omega(\vartheta_{\nu,N}(\omega))\big)\, dW_x(\omega). \tag{5.5}$$

Here the limit exists pointwise in $x \in \Omega$. Now each $u_{N\nu}$ is continuous on Ω, indeed on $\mathbb{R}^n$. Consequently, $u(x)$ given by the right side of (5.2) is at least a bounded, measurable function of x.

To continue the analysis, given $x \in \Omega$, we define a probability measure $\nu_{x,\Omega}$ on $\partial\Omega$ by

$$E_x\big(f(\omega(\vartheta_{\overline{\Omega}}))\big) = \int\limits_{\Omega} f(y)\, d\nu_{x,\Omega}(y), \tag{5.6}$$

for $f \in C(\partial\Omega)$.

Lemma 5.1. *If $x \in \mathcal{O} \subset\subset \Omega$ and $\mathcal{O}$ and Ω are open, then*

$$\nu_{x,\Omega} = \int\limits_{\partial\mathcal{O}} \nu_{y,\Omega}\, d\nu_{x,\mathcal{O}}(y). \tag{5.7}$$

Proof. The identity (5.4) is equivalent to the statement that, for $f \in C(\partial\Omega)$,

$$E_x\big(f(\omega(\vartheta_{\overline{\Omega}}))\big) = \int\limits_{\partial\mathcal{O}} E_y\big(f(\omega(\vartheta_{\overline{\Omega}}))\big)\, d\nu_{x,\mathcal{O}}(y). \tag{5.8}$$

The right side is equal to

$$E_x\big(g(\omega(\vartheta_{\overline{\mathcal{O}}}))\big), \quad g(y) = E_y\big(f(\omega(\vartheta_{\overline{\Omega}}))\big). \tag{5.9}$$

In other words,

$$g(\omega(\vartheta_{\overline{\mathcal{O}}})) = E_{\omega(\vartheta_{\overline{\mathcal{O}}})}(\varphi), \quad \varphi(\omega) = f\big(\omega(\vartheta_{\overline{\Omega}}(\omega))\big). \tag{5.10}$$

Now we use the strong Markov property, in the form (4.22), namely,

$$E_{\omega(\tau)}(\varphi) = E_x\big(\varphi \circ \sigma_\tau \big| \mathfrak{B}_{\tau+}\big),$$

for W_x-almost all ω, where $(\sigma_\tau\omega)(t) = \omega(t+\tau)$ and τ is a stopping time. This implies

$$\int\limits_{\mathfrak{P}_0} E_{\omega(\tau)}(\varphi)\, dW_x(\omega) = \int\limits_{\mathfrak{P}_0} E_x\big(\varphi\circ\sigma_\tau\big|\mathfrak{B}_{\tau+}\big)\, dW_x(\omega) = E_x(\varphi\circ\sigma_\tau). \tag{5.11}$$

Applied to $\tau = \vartheta_{\overline{\mathcal{O}}}$, this shows that (5.9) is equal to $E_x(\varphi\circ\sigma_{\vartheta_{\overline{\mathcal{O}}}})$. Now, with $\widetilde{\omega}(t) = \sigma_{\vartheta_{\overline{\mathcal{O}}}}\omega(t) = \omega(t + \vartheta_{\overline{\mathcal{O}}}(\omega))$, we have, for $\mathcal{O} \subset\subset \Omega$, $\vartheta_{\overline{\Omega}}(\widetilde{\omega}) = \vartheta_{\overline{\Omega}}(\omega) - \vartheta_{\overline{\mathcal{O}}}(\omega)$, as long as ω is a continuous path starting in $\mathcal{O}$. Hence

$$\varphi(\widetilde{\omega}) = f\big(\widetilde{\omega}(\vartheta_{\overline{\Omega}}(\omega) - \vartheta_{\overline{\mathcal{O}}}(\omega))\big) = f\big(\omega(\vartheta_{\overline{\Omega}}(\omega))\big) = \varphi(\omega). \tag{5.12}$$

Thus (5.9) is equal to $E_x(\varphi)$, which is the left side of (5.6), and the lemma is proved.

Consequently, the right side $u(x)$ of (5.2) is a bounded, measurable function of x satisfying the mean-value property. An integration yields that such $u(x)$ is equal to the mean value of u over any *ball* $\mathcal{D} \subset \Omega$, centered at x, from which it follows that $u(x)$ is continuous in Ω. Then the mean-value property guarantees that u is harmonic on Ω. To verify (5.2), it remains to show that $u(x)$ has the correct boundary values.

Lemma 5.2. *Assume $\partial\Omega$ is smooth. Given $y \in \partial\Omega$, we have $u(y) = f(y)$, and u is continuous at $y \in \overline{\Omega}$.*

Proof. That $u(y) = f(y)$ follows from the fact that $\vartheta_{\overline{\Omega}}(\omega) = 0$ for W_y-almost all ω, according to Proposition 3.7. To show that $u(x) \to u(y)$ as $x \to y$ from within Ω, we argue as follows.

By (3.23), for $x \in \Omega$, $E_x(\vartheta_{\overline{\Omega}}) = -\Delta^{-1}1(x)$. Hence this quantity approaches 0 as $x \to y$. Thus, given $\varepsilon_j > 0$, there exists $\delta > 0$ such that

$$|x - y| \le \delta \Longrightarrow W_x\big(\{\omega : \vartheta_{\overline{\Omega}}(\omega) > \varepsilon_1\}\big) < \varepsilon_2. \tag{5.13}$$

Meanwhile, in a short time, $0 \le s \le \varepsilon_1$, a path $\omega(s)$ is not likely to wander far. In fact, by (3.28) plus a scaling argument,

$$\begin{aligned} \mathcal{W}_{\varepsilon_1} = \{\omega \in \mathfrak{P}_0 : &\sup_{0 \le s \le \varepsilon_1} |\omega(s) - \omega(0)| \ge \varepsilon_1^{1/3}\} \\ &\Longrightarrow W_x(\mathcal{W}_{\varepsilon_1}) \le \psi(\varepsilon_1), \end{aligned} \tag{5.14}$$

where $\psi(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Thus, if $|x - y| \le \delta$, with probability $> 1 - \varepsilon_2 - \psi(\varepsilon_1)$, a path starting at x will, within time ε_1, hit $\partial\Omega$, without leaving the ball $B_{\varepsilon_1^{1/3}}(x)$ of radius $\varepsilon_1^{1/3}$ centered at x. Now, a given $f \in C(\partial\Omega)$ varies only a little over $\{z \in \partial\Omega : |z - y| \le \varepsilon_1^{1/3} + \delta\}$ if ε_1 and δ are small enough. Therefore, indeed $u(x) \to u(y)$, as $x \to y$.

We have completed the demonstration of the following.

Proposition 5.3. *If Ω is a bounded region in $\mathbb{R}^n$ with smooth boundary and $f \in C(\partial\Omega)$, then PI f is given by (5.2).*

Recall from §5 of Chapter 5 the construction of

$$\mathrm{PI} : C(\partial\Omega) \longrightarrow L^\infty(\Omega) \cap C^\infty(\Omega) \tag{5.15}$$

when Ω is an arbitrary bounded, open subset of $\mathbb{R}^n$, with perhaps a very nasty boundary. As shown there, we can take

$$\Omega_1 \subset\subset \Omega_2 \subset\subset \cdots \subset\subset \Omega_j \nearrow \Omega \tag{5.16}$$

such that each boundary $\partial\Omega_j$ is smooth, and, if f is extended from $\partial\Omega$ to an element of $C_o(\mathbb{R}^n)$, then

$$x \in \Omega \Longrightarrow \mathrm{PI}\, f(x) = \lim_{j\to\infty} u_j(x), \tag{5.17}$$

where $u_j \in C(\overline{\Omega}_j)$ is the Poisson integral of $f|_{\partial\Omega_j}$. In (5.17) one has uniform convergence on compact sets $K \subset \Omega$, the right side being defined for $j \geq j_0$, where $K \subset \Omega_{j_0}$. The details were carried out in Chapter 5 for $f \in C^\infty(\mathbb{R}^n)$, but approximation by smooth functions plus use of the maximum principle readily extends this to $f \in C_o(\mathbb{R}^n)$.

If we apply Proposition 5.3 to Ω_j, we conclude that, for $f \in C_o(\mathbb{R}^n), x \in \Omega$,

$$\mathrm{PI}\, f(x) = \lim_{j\to\infty} E_x\Big(f\big(\omega(\vartheta_{\overline{\Omega}_j})\big)\Big). \tag{5.18}$$

On the other hand, it is straightforward from the definitions that

$$\vartheta_{\overline{\Omega}_j}(\omega) \nearrow \vartheta_\Omega(\omega), \quad \text{for all } \omega \in \mathfrak{P}_0. \tag{5.19}$$

Therefore, via the dominated convergence theorem, we can pass to the limit in (5.18), proving the following.

Proposition 5.4. *If Ω is any bounded, open region in $\mathbb{R}^n$ and $f \in C(\partial\Omega)$, then*

$$PI\, f(x) = E_x\Big(f\big(\omega(\vartheta_\Omega)\big)\Big), \quad x \in \Omega. \tag{5.20}$$

We recall from Chapter 5 the notion of a *regular* boundary point. A point $y \in \partial\Omega$ is regular provided PI f is continuous at y, for all $f \in C(\partial\Omega)$. We discussed several criteria for a boundary point to be regular, particularly in Propositions 5.11–5.16 of Chapter 5. Here is another criterion.

Proposition 5.5. *If $\Omega \subset \mathbb{R}^n$ is a bounded open set, $y \in \partial\Omega$, then y is a regular boundary point if and only if*

$$E_x(\vartheta_\Omega) \to 0, \quad \text{as } x \to y, \ x \in \Omega. \tag{5.21}$$

Proof. Recall from (3.20) that $E_x(\vartheta_\Omega) = -\Delta^{-1}1(x)$. Thus (5.21) holds if and only if this function is a weak barrier at $y \in \partial\Omega$, as defined in Chapter 5, right after (5.26). Therefore, (5.21) here implies y is a regular point. On the other hand, $\Delta^{-1}1(x)$ can be written as the sum $x_1^2/2 + u_0(x)$, where $u_0 = -(1/2)\,\mathrm{PI}\,(x_1^2\big|_{\partial\Omega})$, so if (5.21) fails, y is not a regular point.

One might both compare and contrast this proof with that of Lemma 5.2. In that case, where $\partial\Omega$ was assumed smooth, the known regularity of each boundary point was exploited to guarantee that $E_x(\vartheta_{\overline{\Omega}}) \to 0$ as $x \to y \in \partial\Omega$, which then was exploited to show that $u(x) \to u(y)$ as $x \to y$.

In the next section, we will derive another criterion for y to be regular, in terms of "capacity."

Exercises

1. Explore connections between the formulas for PI $f(x)$, for $f \in C(\partial\Omega)$, when Ω is bounded and $\partial\Omega$ smooth, given by (3.36) and by (5.2), respectively.

6. Newtonian capacity

The (Newtonian) capacity of a set is a measure of size that is very important in potential theory and closely related to the probability of a Brownian path hitting that set. In our development here, we restrict attention to the case $n \geq 3$ and define the capacity of a compact set $K \subset \mathbb{R}^n$. We first assume that K is the closure of an open set with smooth boundary.

Proposition 6.1. *Assume $n \geq 3$. If $K \subset \mathbb{R}^n$ is compact with smooth boundary ∂K, then there exists a unique function U_K, harmonic on $\mathbb{R}^n \setminus K$, such that $U_K(x) \to 1$ as $x \to K$ and $U_K(x) \to 0$ as $|x| \to \infty$.*

Proof. We can assume that the origin $0 \in \mathbb{R}^n$ is in the interior of K. Then the inversion $\psi(x) = x/|x|^2$ interchanges 0 and the point at infinity, and the transformation

$$v(x) = |x|^{-(n-2)} w(|x|^{-2}x) \tag{6.1}$$

preserves harmonicity. We let w be the unique harmonic function on the bounded domain $\psi(\mathbb{R}^n \setminus K)$, with boundary value $w(x) = |x|^{-(n-2)}$ on $\psi(\partial K)$. Then v, defined by (6.1), is the desired solution. The uniqueness is immediate, via the maximum principle.

Note that the construction yields

$$|U_K(x)| \leq C|x|^{-(n-2)}, \quad |\partial_r U_K(x)| \leq C|x|^{-(n-1)}, \quad |x| \to \infty. \tag{6.2}$$

The $n = 3$ case of this result was done in §1 of Chapter 9.

Another approach to the proof of Proposition 6.1 would be to represent $U_K(x)$ as a single-layer potential, as in (11.44) of Chapter 7. This was noted in a remark after the proof of Proposition 11.5 in that chapter.

Now that we have established the existence of such U_K, Exercises 7–9 of §3 apply, to yield

$$U_K^t(x) \nearrow U_K(x), \quad \text{as } t \nearrow \infty, \tag{6.3}$$

where, for $x \in \mathcal{O} = \mathbb{R}^n \setminus K$,

$$U_K^t(x) = 1 - e^{t\Delta_O} 1(x) \\ = W_x\big(\{\omega : \vartheta_{\mathcal{O}}(\omega) \le t\}\big). \tag{6.4}$$

Here, Δ_O is the Laplace operator on $\mathcal{O}$, with Dirichlet boundary condition. The last identity follows from (3.27). We can replace the first exit time $\vartheta_{\mathcal{O}}$ by the first hitting time:

$$\mathfrak{h}_K(\omega) = \vartheta_{\mathbb{R}^n \setminus K}(\omega). \tag{6.5}$$

Consequently,

$$U_K(x) = W_x\big(\{\omega : \mathfrak{h}_K(\omega) < \infty\}\big); \tag{6.6}$$

that is, for $x \in \mathcal{O}$, $U_K(x)$ is the probability that a Brownian path ω, starting at x, eventually hits K.

We set $U_K(x) = 1$ for $x \in K$. Then (6.6) holds for $x \in K$ also. It follows that $U_K \in C_o(\mathbb{R}^n)$, and ΔU_K is a distribution supported on ∂K. In fact, Green's formula yields, for $\varphi \in C_0^\infty(\mathbb{R}^n)$,

$$(U_K, \Delta\varphi) = -\int_{\partial K} \varphi(y) \frac{\partial}{\partial \nu} U_K(y)\, dS(y), \tag{6.7}$$

where ν is the unit normal to ∂K, pointing into K. By Zaremba's principle, $\partial_\nu U_K(y) > 0$, for all $y \in \partial K$, so we see that $\Delta U_K = -\mu_K$, where μ_K is a positive measure supported on ∂K. The total mass of μ_K is called the *capacity* of K:

$$\text{cap } K = \int_K d\mu_K(x). \tag{6.8}$$

Since, with $C_n = (n-2) \cdot \text{Area}(S^{n-1})$,

$$U_K(x) = -\Delta^{-1}\mu_K = C_n \int |x-y|^{-(n-2)}\, d\mu_K(y), \tag{6.9}$$

we have

$$C_n \iint \frac{d\mu_K(x)\, d\mu_K(y)}{|x-y|^{n-2}} = \int U_K(x)\, d\mu_K(x) = \text{ cap } K, \tag{6.10}$$

the left side being proportional to the potential energy of a collection of charged particles, with density $d\mu_K$, interacting by a repulsive force with potential $C_n|x-y|^{-(n-2)}$. The function $U_K(x)$ is called the *capacitary potential* of K. Note that we can also use Green's theorem to get

$$\|\nabla U_K\|^2_{L^2(\mathbb{R}^n)} = \int_K U_K(x)\, d\mu_K(x) = \text{ cap } K. \tag{6.11}$$

Note that if $K_1 \subset K_2$ have capacitary potentials U_j, $\Delta U_j = -\mu_j$, then $U_2 = 1$ on K_1, so

$$\begin{aligned} \text{cap } K_1 &= \int U_2(x)\, d\mu_1(x) = -(U_2, \Delta U_1) \\ &= \int U_1(x)\, d\mu_2(x) \le \text{ cap } K_2, \end{aligned} \tag{6.12}$$

since $U_1(x) \le 1$. Thus capacity is a monotone set function.

Before establishing more formulas involving capacity, we extend it to general compact $K \subset \mathbb{R}^n$. We can write $K = \bigcap K_j$, where $K_1 \supset\supset K_2 \supset\supset \cdots \supset\supset K_j \searrow K$, each K_j being compact with smooth boundary. Clearly, $U_j = U_{K_j}$ is a decreasing sequence of functions ≤ 1, and by (6.11), ∇U_j is bounded in $L^2(\mathbb{R}^n)$. Furthermore, $\Delta U_j = -\mu_j$, where μ_j is a positive measure supported on ∂K_j, of total mass cap K_j, which is nonincreasing, by (6.12). Consequently, we have a limit:

$$\lim_{j\to\infty} U_j = U_K, \tag{6.13}$$

defined a priori pointwise, but also holding in various topologies, such as the weak* topology of $L^\infty(\mathbb{R}^n)$. We have $U_K \in L^\infty(\mathbb{R}^n)$, $0 \le U_K(x) \le 1$; $\nabla U_K \in L^2(\mathbb{R}^n)$, and $\Delta U_K = -\mu$, where μ is a positive measure, supported on K. Furthermore, $\mu_j \to \mu$ in the weak* topology, and $U_K = -\Delta^{-1}\mu$. Any neighborhood of K contains some K_j. Thus, if $K_1' \supset\supset K_2' \supset\supset \cdots \supset\supset K_j' \searrow K$ is another choice, one is seen to obtain the same limit U_K, hence the same measure μ, which we denote as μ_K. We set

$$\text{cap } K = \int d\mu_K(x). \tag{6.14}$$

Note that, as in (6.12), cap $K = \int U_j(x)\, d\mu_K(x)$, for each j. Thus, as before, cap $K = \int U_K(x)\, d\mu_K(x)$, this time by the monotone convergence theorem. Consequently,

$$U_K(x) = 1 \quad \mu_K\text{-almost everywhere.} \tag{6.15}$$

Clearly, cap $K \le \inf$ cap K_j. In fact, we claim

$$\text{cap } K = \inf \text{ cap } K_j. \tag{6.16}$$

This is easy to see; μ_j converges to μ_K pointwise on $C_o(\mathbb{R}^n)$; choose $g \in C_o(\mathbb{R}^n)$, equal to 1 on K_1; then

$$\text{cap } K = (g, \mu_K) = \lim\, (g, \mu_j) = \lim \text{ cap} K_j, \tag{6.17}$$

proving (6.16). We consequently extend the monotonicity property:

Proposition 6.2. *For general compact $K \subset L$, we have cap $K \le$ cap L.*

Proof. We can take compact approximants with smooth boundary, $K_j \searrow K$, $L_j \searrow L$, such that $K_j \subset L_j$. By (6.12) we have cap $K_j \le$ cap L_j, and this persists in the limit by (6.16). We also have $U_K(x) \le U_L(x)$ for all x. Using (6.15), we obtain

$$\text{cap } K = \int U_L(x)\, d\mu_K(x). \tag{6.18}$$

One possibility is that cap $K = 0$. This happens if and only if $\mu_K = 0$, thus if and only if $U_K = 0$ almost everywhere. If cap $K > 0$, we continue to call U_K the capacitary potential of K.

We record some more ways in which $U_j \to U_K$. First, it certainly holds in the weak* topology on $L^\infty(\mathbb{R}^n)$. Hence $\nabla U_j \to \nabla U_K$ in $\mathcal{D}'(\mathbb{R}^n)$. By (6.11), ∇U_j is bounded in $L^2(\mathbb{R}^n)$; hence $\nabla U_j \to \nabla U_K$ weakly in $L^2(\mathbb{R}^n)$. Since also $U_j \in C_o(\mathbb{R}^n)$, we have

$$\begin{aligned} \|\nabla U_K\|_{L^2}^2 &= \lim_{j\to\infty} (\nabla U_j, \nabla U_K) = \lim_{j\to\infty}\; -(U_j, \Delta U_K) \\ &= \lim_{j\to\infty} \int U_j(x)\, d\mu(x) = \text{ cap } K, \end{aligned} \tag{6.19}$$

the last identity holding as in the derivation of (6.15). Thus (6.11) is extended to general compact K. Furthermore, this implies

$$\nabla U_j \longrightarrow \nabla U_K \quad \text{in } L^2(\mathbb{R}^n)\text{-norm}. \tag{6.20}$$

Hence

$$\mu_j \longrightarrow \mu_K \quad \text{in } H^{-1}(\mathbb{R}^n)\text{-norm}. \tag{6.21}$$

We now extend the identities (6.3) and (6.6) to general compact K, in reverse order.

Proposition 6.3. *The identity (6.6) holds for general compact $K \subset \mathbb{R}^n$.*

Proof. Since (6.6) has been established for the compact K_j with smooth boundary, we have

$$1 - U_j(x) = W_x(\mathfrak{A}_{K_j}), \quad \mathfrak{A}_{K_j} = \{\omega \in \mathfrak{P}_0 : \omega(\mathbb{R}^+) \subset \mathbb{R}^n \setminus K_j\}. \tag{6.22}$$

Clearly, if $K_j \searrow K$, $\mathfrak{A}_{K_1} \subset \mathfrak{A}_{K_2} \subset \cdots \subset \mathfrak{A}_{K_j} \nearrow \widetilde{\mathfrak{A}}_K$, where $\widetilde{\mathfrak{A}}_K$ is a proper subset of $\mathfrak{A}_K = \{\omega \in \mathfrak{P}_0 : \omega(\mathbb{R}^+) \subset \mathbb{R}^n \setminus K\}$. However, for $n \ge 3$, Brownian motion is nonrecurrent, as was established in Exercise 10 of §3. Thus $|\omega(t)| \to \infty$ as $t \to \infty$, for W_x-almost all ω, so in fact $W_x(\mathfrak{A}_K \setminus \widetilde{\mathfrak{A}}_K) = 0$, and hence $1 - U_K(x) = W_x(\mathfrak{A}_K)$, which is equivalent to (6.6).

Proposition 6.4. *The identity (6.3) holds for general compact $K \subset \mathbb{R}^n$.*

Proof. We define $U_K^t(x)$ to be $1 - e^{t\Delta_O}1(x)$, as in (6.4); the second identity in (6.4) continues to hold, by (3.27). Now, clearly, the family of sets $S_t =$

$\{\omega \in \mathfrak{P}_0 : \mathfrak{h}_K(\omega) \le t\}$ is increasing as $t \nearrow \infty$, with union

$$\bigcup S_t = \{\omega \in \mathfrak{P}_0 : \mathfrak{h}_K(\omega) < \infty\},$$

and this gives (6.3).

We next establish the subadditivity of capacity.

Proposition 6.5. *If K and L are compact, then*

$$U_{K\cup L}(x) \le U_K(x) + U_L(x) \tag{6.23}$$

and

$$\text{cap}(K \cup L) \le \big(\text{cap } K\big) + \big(\text{cap } L\big). \tag{6.24}$$

Proof. The inequality (6.23) follows directly from (6.6) and the subadditivity of Wiener measure. Now, as in (6.12), we have

$$\begin{aligned} \int U_K(x)\, d\mu_{K\cup L}(x) &= -(U_K, \Delta U_{K\cup L}) \\ &= \int U_{K\cup L}(x)\, d\mu_K(x) \\ &= \text{cap } K, \end{aligned} \tag{6.25}$$

the last identity by (6.18), with L replaced by $K \cup L$. Hence

$$\text{cap } K + \text{cap } L = \int \big[U_K(x) + U_L(x)\big]\, d\mu_{K\cup L}(x),$$

so the estimate (6.23) implies (6.24).

Note that even if K and L are disjoint, typically there is inequality in (6.23), hence in (6.24). In fact, if K and L are disjoint compact sets,

$$\begin{aligned} (\text{cap } K) + (\text{cap } L) &= \text{cap}(K \cup L) + R, \\ R &= \int_L U_K(x)\, d\mu_{K\cup L}(x) + \int_K U_L(x)\, d\mu_{K\cup L}(x), \end{aligned} \tag{6.26}$$

the quantity R being > 0 unless either cap $K = 0$ or cap $L = 0$. Unlike measures, the capacity is *not* an additive set function on disjoint compact sets.

We began this section with the statement that the capacity of K is closely related to the probability of a Brownian path hitting K. We have directly tied $U_K(x)$ to this probability, via (6.6). We now provide a two-sided estimate on $U_K(x)$ in terms of cap K.

Proposition 6.6. *Let* $\delta(x) = \sup\{|x-y| : y \in K\}$, *and let* $d(x)$ *denote the distance of* $x \in \mathbb{R}^n$ *from* K. *Then*

$$\frac{C_n}{\delta(x)^{n-2}}\big(\text{cap } K\big) \le U_K(x) \le \frac{C_n}{d(x)^{n-2}}\big(\text{cap } K\big). \tag{6.27}$$

Proof. The formula $U_K(x) = C_n \int |x-y|^{-(n-2)} d\mu_K(y)$ represents $U_K(x)$ as $C_n(\text{cap } K)$ times a weighted average of $|x-y|^{-(n-2)}$ over K. Now, for $y \in K$, $d(x) \le |x-y| \le \delta(x)$, so (6.27) follows.

We want to compare this with the probability that a Brownian path hits ∂K in the interval $[0,t]$. It t is large, we know that $|\omega(t)|$ is probably large, given that $n \ge 3$, and hence $\omega(s)$ probably will not hit K for any $s > t$. Thus we expect this probability (which is equal to $U_K^t(x)$) to be close to $U_K(x)$. We derive a quantitative estimate as follows. Since $1 - U_K^t(x) = e^{t\Delta_{\mathcal{O}}}1(x)$, we have, for $s \ge 0$,

$$U_K^{t+s}(x) - U_K^t(x) = e^{t\Delta_{\mathcal{O}}}1(x) - e^{(t+s)\Delta_{\mathcal{O}}}1(x) = e^{t\Delta_{\mathcal{O}}}U_K^s(x), \tag{6.28}$$

and taking $s \nearrow \infty$, we get

$$U_K(x) - U_K^t(x) = e^{t\Delta_{\mathcal{O}}}U_K(x). \tag{6.29}$$

Hence, if we denote the heat kernel on $\mathcal{O} = \mathbb{R}^n \setminus K$ by $p_{\mathcal{O}}(t,x,y)$, and that on $\mathbb{R}^n$ by $p(t,x-y)$, as in (1.5),

$$\begin{aligned} &U_K(x) - U_K^t(x) \\ &\quad = \int p_{\mathcal{O}}(t,x,y)U_K(y)\,dy \le \int p(t,x-y)U_K(y)\,dy \\ &\quad = C_n \iint \frac{p(t,x-y)}{|y-z|^{n-2}}\,dy\,d\mu_K(z) \le (\text{cap } K)\sigma_K(t,x), \end{aligned} \tag{6.30}$$

where

$$\sigma_K(t,x) = C_n \sup_{z\in K} \int \frac{p(t,x-y)}{|y-z|^{n-2}}\,dy = \sup_{z\in K} \int_t^\infty p(s,x-z)\,ds, \tag{6.31}$$

the last integral being another way of writing $e^{t\Delta}(-\Delta)^{-1}\delta(x-z)$ when $n \ge 3$. An upper bound on $\sigma_K(t,x)$ is $\int_t^\infty (4\pi s)^{-n/2}\,ds$, so we have

$$0 \le U_K(x) - U_K^t(x) \le \frac{2}{n-2}\big(4\pi\big)^{-n/2} t^{-n/2+1}\big(\text{cap } K\big). \tag{6.32}$$

There is an interesting estimate on the smallest eigenvalue of $-\Delta$ on the complement of a compact set K, in terms of cap K, which we now describe. Let $Q = \{x \in \mathbb{R}^n : 0 \le x_j \le 1\}$ be the closed unit cube in $\mathbb{R}^n$, and let $K \subset Q$ be compact. We consider the boundary condition on functions on $Q \setminus K$:

$$u = 0 \text{ on } \partial K, \quad \frac{\partial u}{\partial \nu} = 0 \text{ on } \partial Q \setminus \partial K. \tag{6.33}$$

To define this precisely, let $H^1(Q,K)$ denote the closure in $H^1(Q)$ of the set of functions in $C^\infty(Q)$ vanishing on a neighborhood of K. Then the quadratic form $(du,dv)_{L^2}$ restricted to $H^1(Q,K)\times H^1(Q,K)$ defines an unbounded, self-adjoint operator L, which we denote $-\Delta_{Q,K}$, with $\mathcal{D}(L^{1/2}) = H^1(Q,K)\subset H^1(Q)$. Hence $-\Delta_{Q,K}$ has compact resolvent and thus a discrete spectrum. Let $\lambda_0(K)$ be its smallest eigenvalue.

Proposition 6.7. *The smallest eigenvalue $\lambda_0(K)$ of $-\Delta$ on $Q\setminus K$, with boundary condition (6.33), satisfies the estimate*

$$\lambda_0(K)\ge \gamma_n \text{ cap } K, \tag{6.34}$$

for some $\gamma_n > 0$.

Proof. Let $p_{Q,K}(t,x,y)$ denote the heat kernel of $\Delta_{Q,K}$. With $\mathcal{O}=\mathbb{R}^n\setminus K$, let $p_{\mathcal{O}}(t,x,y)$ denote the heat kernel of Δ on $\mathcal{O}$, with Dirichlet boundary condition, as in (6.30). We claim that

$$\int_Q p_{Q,K}(t,x,y)\,dy\le\int_{\mathbb{R}^n} p_{\mathcal{O}}(t,x,y)\,dy,\quad x\in Q. \tag{6.35}$$

To see this, define $\widetilde{K}$ by the method of images, so in each unit cube with integer vertices we have a reflected image of K, and, with $\widetilde{\mathcal{O}}=\mathbb{R}^n\setminus\widetilde{K}$,

$$p_{Q,K}(t,x,y)=\sum_j p_{\widetilde{\mathcal{O}}}(t,x,R_jy),\quad x,y\in Q, \tag{6.36}$$

where the transformations R_j are appropriate reflections. Then (6.35) follows from the obvious pointwise estimate $p_{\widetilde{\mathcal{O}}}(t,x,y)\le p_{\mathcal{O}}(t,x,y)$. Now, if we set

$$M(t)=\sup_{x\in Q}\int_{\mathbb{R}^n} p_{\mathcal{O}}(t,x,y)\,dy, \tag{6.37}$$

it follows that

$$\sup_x\int_Q p_{Q,K}(t,x,y)\,dy\le M(t),\quad \sup_y\int_Q p_{Q,K}(t,x,y)\,dx\le M(t), \tag{6.38}$$

the latter by symmetry. It is well known that the operator norm of $e^{t\Delta_{Q,K}}$ is bounded by the quantities (6.38). (See Proposition 5.1 in Appendix A.) Thus

$$\|e^{t\Delta_{Q,K}}\|\le M(t). \tag{6.39}$$

To relate this to capacity, note that

$$M(t)=\sup_{x\in Q}\bigl(1-U_K^t(x)\bigr). \tag{6.40}$$

Now, applying the first estimate of (6.27), in concert with the estimate (6.32), we have

$$M(t) \le 1 - C_n n^{-n/2+1}\bigl(\text{cap } K\bigr) + \frac{2}{n-2}(4\pi)^{-n/2} t^{-n/2+1}\bigl(\text{cap } K\bigr). \tag{6.41}$$

In particular, there exists a finite $T = T_n$ and $\kappa > 0$ such that

$$M(T) \le 1 - \kappa(\text{cap } K) \le e^{-\kappa \text{ cap } K}. \tag{6.42}$$

Since this is an upper bound on $\|e^{T\Delta_{Q,K}}\|$, we have $\lambda_0(K) \ge (\kappa/T) \text{ cap } K$, proving (6.34).

As an application of this, we establish the following result of Molchanov on a class of Dirichlet problems with compact resolvent.

Proposition 6.8. *Let Ω be an unbounded, open subset of $\mathbb{R}^n$, with complement S. Suppose that there exists $\psi(a) \nearrow \infty$ as $a \searrow 0$, such that, for each $a \in (0,1]$, if $\mathbb{R}^n$ is tiled by cubes Q_{aj} of edge a, we have*

$$\text{cap}(Q_{aj} \cap S) \ge \psi(a) a^{2(n-2)}, \tag{6.43}$$

for all but finitely many j. Then the Laplace operator Δ on Ω, with Dirichlet boundary condition, has compact resolvent.

Proof. By scaling Q_{aj} to a unit cube, we see that if (6.43) holds, then $-\Delta$ on $Q_{aj} \setminus S$, with Dirichlet boundary condition on ∂S, Neumann on $\partial Q_{aj} \setminus S$, has smallest eigenvalue $\ge \gamma_n(\text{cap } Q_{aj} \cap S) a^{-2(n-2)}$, which, by hypothesis (6.43) is $\ge \gamma_n \psi(a)$ for all but finitely many j. The variational characterization of the spectrum implies that the spectral subspace of $L^2(\Omega)$ on which $-\Delta$ has spectrum in $[0, \gamma_n \psi(a)]$ is finite-dimensional, for each $a > 0$, and this implies that Δ has compact resolvent.

In our continued study of which boundary points of a region Ω are regular, it will be useful to have the following variant of Proposition 6.6. Here, B_r is the ball of radius r centered at the origin in $\mathbb{R}^n$; see Fig. 6.1.

Proposition 6.9. *Let K be a compact subset of the ball B_1. Let $V_K(x)$ denote the probability that a Brownian path, starting at $x \in \mathbb{R}^n$, hits K before hitting the shell $\partial B_4 = \{x : |x| = 4\}$. Then there is a constant $\widetilde{\gamma}_n > 0$ such that*

$$x \in B_1 \Longrightarrow V_K(x) \ge \widetilde{\gamma}_n\bigl(\text{cap } K\bigr). \tag{6.44}$$

Proof. Note that, by (5.20), V_K is also defined by

$$\Delta V_K = 0 \text{ on } B_4 \setminus K, \quad V_K = 1 \text{ on } K, \quad V_K = 0 \text{ on } \partial B_4. \tag{6.45}$$

FIGURE 6.1

We will compare $V_K(x)$ with $U_K(x)$. By (6.27), we have

$$x \in B_1 \Longrightarrow U_K(x) \geq 2^{-(n-2)} C_n \bigl(\text{cap } K\bigr) \tag{6.46}$$

and

$$x \in \partial B_4 \Longrightarrow U_K(x) \leq 3^{-(n-2)} C_n \bigl(\text{cap } K\bigr). \tag{6.47}$$

By (6.47) and the maximum principle, we have, for $x \in B_4 \setminus K$,

$$V_K(x) \geq \frac{U_K(x) - q(K)}{1 - q(K)}, \quad q(K) = 3^{-(n-2)} C_n \bigl(\text{cap } K\bigr). \tag{6.48}$$

Now $C_n(\text{cap } K) \leq C_n(\text{cap } B_1) = 1$ (compare with Exercise 1 at the end of this section), so using (6.46) we readily obtain (6.44), with

$$\widetilde{\gamma}_n = \Bigl(1 - 3^{-(n-2)}\Bigr)^{-1} \Bigl(2^{-(n-2)} - 3^{-(n-2)}\Bigr) C_n. \tag{6.49}$$

In particular, $\widetilde{\gamma}_3 = C_3/4 = \pi$.

Of course, since $V_K(x) \leq U_K(x)$, we also have

$$x \in B_4, \ \text{dist}(x, K) \geq \rho \Longrightarrow V_K(x) \leq C_n \rho^{-(n-2)} \bigl(\text{cap } K\bigr). \tag{6.50}$$

This upper bound is valid for $K \subset B_4$; we don't need $K \subset B_1$.

Now suppose $y \in K$ is the center of concentric balls B_j, of radius $2^{-j} r$, where $r > 0$ is fixed, $0 \leq j \leq \nu$. See Fig. 6.2. Pick $x \in B_\nu$. We want to estimate the probability that a Brownian path starting at x will exit B_0 before hitting K. Let's call the probability $p_{\text{miss}}(x, K)$. Using Proposition 6.9 and scaling, we see that, given $x \in B_j$, the probability that it hits ∂B_{j-2} before hitting $K \cap B_j$ is $\leq 1 - \widetilde{\gamma}_n r_j^{-(n-2)} \cdot \text{cap}(K \cap B_j)$, where $r_j = 2^{-j} r$. Using the independence of this event and of the event that,

FIGURE 6.2

given $x \in \partial B_{j-2}$, the path will hit ∂B_{j-4} before hitting $K \cap B_{j-2}$, which follows from the strong Markov property, we have an upper bound

$$p_{\text{miss}}(x, K) \leq \prod_{j \in \mathcal{S}_\nu} \Big(1 - \widetilde{\gamma}_n r^{-(n-2)} 2^{(n-2)j} \cdot \ \operatorname{cap}\big(K \cap B_j\big)\Big), \tag{6.51}$$

where $\mathcal{S}_\nu = \{j : 0 \leq j \leq \nu,\ j = \nu \bmod 2\}$. A similar argument dominates $p_{\text{miss}}(x, K)$ by a product over $\{1, \dots \nu\} \setminus \mathcal{S}_\nu$, so

$$p_{\text{miss}}(x, K)^2 \leq \prod_{j=0}^{\nu} \Big(1 - \widetilde{\gamma}_n r^{-(n-2)} 2^{(n-2)j} \cdot \ \operatorname{cap}\big(K \cap B_j\big)\Big). \tag{6.52}$$

Note that, as $\nu \to \infty$, the right side of (6.52) tends to zero, precisely when the sum

$$\sum_{j=0}^{\infty} 2^{(n-2)j} \cdot \operatorname{cap}\big(K \cap B_j\big) \tag{6.53}$$

is *infinite*. We are now ready to state the Wiener criterion for regular points.

Proposition 6.10. *Let Ω be a bounded, open set in $\mathbb{R}^n$, and let $y \in \partial\Omega$. If Ω is inside a ball $\widetilde{B}$, set $K = \widetilde{B} \setminus \Omega$. Then y is a regular point for Ω if and only if the infinite series (6.53) is divergent, where $B_j = \{x \in \mathbb{R}^n : |x - y| \leq 2^{-j}\}$.*

Proof. First suppose (6.53) is divergent. Fix $f \in C(\partial\Omega)$, and look at

$$u(x) = \operatorname{PI} f(x) = E_x\Big(f\big(\omega(\mathfrak{h}_K)\big)\Big). \tag{6.54}$$

Given $\varepsilon > 0$, fix $r > 0$ so that f varies by less than ε on $\{z \in \partial\Omega : |z-y| \le r\}$. By (6.52), if $\delta > 0$ is small enough and $|x-y| \le \delta$, then the probability that a Brownian path $\omega(t)$, starting at x, crosses $\partial B_0 = \{z : |z-y| = r\}$ before hitting K is $< \varepsilon$. Consequently,

$$|x-y| \le \delta \Longrightarrow \Big| E_x\Big(f\big(\omega(\mathfrak{h}_K)\big)\Big) - f(y)\Big| \le \varepsilon + \varepsilon \cdot \sup\, |f|. \tag{6.55}$$

This shows that $PI\ f(x) \to f(y)$ as $x \to y$, for any $f \in C(\partial\Omega)$, so y is regular.

For the converse, if (6.53) converges, we claim there is a $J < \infty$ such that there exist points in $\Omega \cap B_J$, arbitrarily close to y, which are starting points of Brownian paths whose probability of hitting K before exiting B_J is $\le 1/2$.

Consider the shells $A_j = \{x : 2^{-j-1} \le |x-y| \le 2^{-j}\}$; $B_j = \bigcup_{\ell \ge j} A_\ell$. We will estimate the probability that a point picked at random in A_ℓ is the starting point of a Brownian path that hits K before exiting B_J, where ℓ is chosen $> J$. Since we are assuming $n \ge 3$, by the analysis behind nonrecurrence in Exercises 7–10 of §3, the probability that a path starting in A_ℓ *ever* hits $B_{\ell+3}$ is $\le 1/4$. Thus if we alter K to $K_\ell = K \setminus B_{\ell+3}$, the probability that a Brownian path starting in A_ℓ hits K_ℓ before ∂B_J is not decreased by more than 1/4. We aim to show that this new probability is $\le 1/4$ if J is chosen large enough.

Now there is no further decrease in probability that the path hits K_ℓ before ∂B_J if we instead have it start at a random point in $B_{\ell+5}$, since almost all such paths will pass into A_ℓ, in a uniformly distributed fashion through its inner boundary. So we deal with the modified problem of estimating the probability $\widetilde{p}$ that a Brownian path, starting at a random point in $B_{\ell+5}$, hits $K_\ell = K \setminus B_{\ell+3}$ before exiting B_J.

We partition the set $\{j : J \le j \le \ell+3\}$ into two sets, where j is even or odd; call these subsets $\mathcal{J}_0$ and $\mathcal{J}_1$, respectively. Then form

$$\mathcal{A}_0 = \bigcup_{j \in \mathcal{J}_0} A_j, \quad \mathcal{A}_1 = \bigcup_{j \in \mathcal{J}_1} A_j. \tag{6.56}$$

We estimate the probability p_μ that a path starting in $B_{\ell+5}$ hits $K_\ell \cap \mathcal{A}_\mu$ before hitting ∂B_J. We have

$$p_\mu(x) \le \sum_{j \in \mathcal{J}_\mu} p_{\mu j}, \tag{6.57}$$

where $p_{\mu j}$ is the probability that, given $|x-y| = (3/4) \cdot 2^{-j-1}$ (i.e., x is on a shell S_{j+1} halfway between the two boundary components of A_{j+1}), then a path starting at x hits $K \cap A_j$ before hitting S_{j-1}. By (6.50) and a dilation argument, we have an estimate of the form

$$p_{\mu j} \le \gamma_n' 2^{(n-2)j}\ \mathrm{cap}(K \cap A_j). \tag{6.58}$$

Thus the probability $\widetilde{p}$ that we want to estimate satisfies

$$\widetilde{p} \leq \gamma_n' \sum_{j=J}^{\ell+3} 2^{(n-2)j} \text{ cap}(K \cap A_j). \tag{6.59}$$

Of course, $\text{cap}(K \cap A_j) \leq \text{ cap}(K \cap B_j)$, so if (6.53) is assumed to converge, we can pick J sufficiently large that the right side of (6.59) is guaranteed to be $\leq 1/4$.

From here it is easy to pick $f \in C(\partial\Omega)$ such that $f(y) = 1$ but (6.54) does not converge to 1 as $x \to y$. This completes the proof of Proposition 6.10 and also shows that the hypothesis of convergence or divergence of (6.53) can be replaced by such a hypothesis on

$$\sum_{j=0}^{\infty} 2^{(n-2)j} \cdot \text{cap}\big(K \cap A_j\big). \tag{6.60}$$

We can extend capacity to arbitrary sets $S \subset \mathbb{R}^n$. The *inner capacity* $\text{cap}^-(S)$ is defined by

$$\text{cap}^-(S) = \sup \{\text{cap } K : K \text{ compact}, K \subset S\}. \tag{6.61}$$

Clearly, $\text{cap}^-(K) = \text{cap } K$ for compact K. If $U \subset \mathbb{R}^n$ is open, we also set $\text{cap } U = \text{ cap}^-(U)$. Now the *outer capacity* $\text{cap}^+(S)$ is defined by

$$\text{cap}^+(S) = \inf \{\text{cap } U : U \text{ open}, S \subset U\}. \tag{6.62}$$

It is easy to see that $\text{cap}^+(S) \geq \text{ cap}^-(S)$ for all S. If $\text{cap}^+(S) = \text{ cap}^-(S)$, then S is said to be capacitable, and the common quantity is denoted cap S. The analysis leading to (6.16) shows that every compact set is capacitable; also, by definition, every open set is capacitable. G. Choquet proved that every Borel set is capacitable; in fact, his capacitability theorem extends to a more general class of sets, known as Souslin sets. We refer to [Mey] for a detailed presentation of this result.

The outer capacity can be shown to satisfy the property that, for any increasing sequence of sets $S_j \subset \mathbb{R}^n$,

$$S_j \nearrow S \Longrightarrow \text{ cap}^+(S_j) \nearrow \text{ cap}^+(S).$$

We establish a useful special case of this.

Proposition 6.11. *If U_j and U are open and $U_j \nearrow U$, then*

$$\text{cap } U_j \nearrow \text{ cap } U.$$

Proof. Given $\varepsilon > 0$, pick a compact $K \subset U$ such that $\text{cap } K \geq \text{ cap } U - \varepsilon$. Then $K \subset U_j$ for large j, so $\text{cap } U_j \geq \text{ cap } U - \varepsilon$ for large j.

We next present a result, due to M. Brelot, to the effect that the set of irregular boundary points of a given bounded, open set is rather small.

Proposition 6.12. *If $\Omega \subset \mathbb{R}^n$ is open and bounded, the set I of irregular boundary points in $\partial\Omega$ has inner capacity zero.*

Proof. The claim is that if $K \subset I$ is compact, then cap $K = 0$. By subadditivity, it suffices to show the following: Given $y \in \partial\Omega$, there is a neighborhood B of y in $\mathbb{R}^n$ such that any compact $K \subset I \cap B$ has capacity zero.

We prove the result in the case that Ω is connected. Let $L = \overline{B} \setminus \Omega$, and consider the capacitary potential $U_L(x)$. In this case, $\mathbb{R}^n \setminus L$ is connected. The function $1 - U_L(x)$ is a weak barrier at any $z \in L \cap \partial\Omega$ with the property that $U_L(x) \to 1$ as $x \to z,\ x \in \mathbb{R}^n \setminus L$. Thus it suffices to show that the set $J = \{z \in L : U_L(z) < 1\}$ has inner capacity zero.

Let $K \subset J$ be compact. We know that $U_K(x) \le U_L(x)$ for all $x \in \mathbb{R}^n$. Thus $U_K(x) < 1$ on K. Now, by (6.15), $U_K(x) = 1$ for μ_K-almost all x, so we conclude that $\mu_K = 0$, hence cap $K = 0$. This completes the proof when Ω is connected.

The general case can be done as follows. If Ω is not connected, it has at most countably many connected components. One can connect the various components via little tubes whose total (inner) capacity can be arranged, via Proposition 6.11, to be arbitrarily small, say $< \varepsilon$. Then the set of irregular points is decreased by a set of inner capacity $< \varepsilon$. The reader is invited to supply the details.

As noted in Proposition 5.5, the set of irregular points of $\partial\Omega$ can be characterized as the set of points of discontinuity of a function E, defined on $\overline{\Omega}$ to be $-\Delta^{-1}1(x)$ for $x \in \Omega$ and to be 0 on $\partial\Omega$. Such a set of points of discontinuity is a Borel subset of Ω, in fact an $\mathcal{F}_{\sigma\delta}$-set. Thus the capacitability theorem applies: If $\Omega \subset \mathbb{R}^n$ is a bounded open set, the set of irregular points of $\partial\Omega$ has capacity zero. This sharpening of Proposition 6.12 was first established by H. Cartan.

As we stated at the beginning of this section, we have been working under the assumption that $n \ge 3$. Two phenomena that we have exploited fail when $n = 2$. One is that Δ has a fundamental solution ≤ 0 on all of $\mathbb{R}^n$. The other is that Brownian motion is nonrecurrent. (Of course, these two phenomena are related.) There is a theory of logarithmic capacity of planar sets. One way to approach things is to consider capacities only of subsets of some fixed disk, of large radius R, and use the Laplace operator on this disk, with the Dirichlet boundary condition. Then one looks at Brownian paths only up to the first exit time from this disk. The results of this section extend. In particular, the Wiener criterion for $n = 2$ is the convergence or divergence of

$$\sum_{j=1}^{\infty} j \cdot \operatorname{cap}\bigl(K \cap A_j\bigr). \tag{6.63}$$

Exercises

1. If $K \subset \mathbb{R}^n$ is compact, show that
$$\lim_{|x|\to\infty} |x|^{n-2} U_K(x) = C_n \text{ cap } K.$$
If $K = B_a$ is a ball of radius a, show that cap $B_a = a^{n-2}/C_n$.
Show generally that if $a > 0$ and $K_a = \{ax : x \in K\}$, then cap $K_a = a^{n-2}$ cap K.
2. Show that cap $K =$ cap ∂K. Show that the identity cap $\partial B_a = a^{n-2}/C_n$ follows from (6.27), with x the center of B_a.
3. Let C_{ar} be the union of two balls of radius a, with centers separated by a distance r. Show that
$$\text{cap } C_{ar} \nearrow 2 \text{ cap } B_a, \quad \text{as } r \to \infty.$$
Estimate the rate of convergence.
4. The task here is to estimate the capacity of a cylinder in $\mathbb{R}^n$, of height b and radius a. Suppose $\mathcal{C}(a,b) = \{x \in \mathbb{R}^n : 0 \le x_n \le b,\ x_1^2 + \cdots + x_{n-1}^2 \le a^2\}$. Show that there are positive constants α_n and β_n such that
$$\begin{aligned} \text{cap } \mathcal{C}(a,1) &\sim \alpha_n a^{n-3}, \quad a \to 0,\ n \ge 4, \\ \text{cap } \mathcal{C}(a,1) &\sim \beta_n a^{n-2}, \quad a \to \infty,\ n \ge 3. \end{aligned}$$
Derive an appropriate result for $n = 3,\ a \to 0$.
5. Let ν be a positive measure supported on a compact set $K \subset \mathbb{R}^n$, such that
$$U_\nu(x) = -\Delta^{-1}\nu(x) = C_n \int \frac{d\nu(x)}{|x-y|^{n-2}} \le 1.$$
Show that $U_\nu(x) \le U_K(x)$ for all $x \in \mathbb{R}^n$. Taking the limit as $|x| \to \infty$, deduce from the asymptotic behavior of $U_\nu(x)$ and $U_K(x)$ (as in Exercise 1) that $\int d\nu(x) \le$ cap K.

6. Show that, for compact $K \subset \mathbb{R}^n$,
$$\text{cap } K = \inf\Big\{ \int |\nabla f(x)|^2 \, dx : f \in C_0^\infty(\mathbb{R}^n),\ f = 1 \text{ on nbd of } K \Big\}. \tag{6.64}$$
(*Hint*: Show that a minimizing sequence f_j approaches U_K.)
Show that the condition $f = 1$ on a neighborhood of K can be replaced by $f \ge 1$ on K. Show that if $f \in C_0^1(\mathbb{R}^n),\ \lambda > 0$,
$$\text{cap}\Big(\{x \in \mathbb{R}^n : |f(x)| \ge \lambda\}\Big) \le \lambda^{-2} \|\nabla f\|_{L^2}^2. \tag{6.65}$$
7. Show that, for compact $K \subset \mathbb{R}^n$,
$$\frac{1}{\text{cap } K} = \inf \Big\{ C_n \iint \frac{d\lambda(x)\ d\lambda(y)}{|x-y|^{n-2}} : \lambda \in \mathcal{P}_K^+ \Big\}, \tag{6.66}$$
where $\mathcal{P}_K^+$ denotes the space of probability measures supported on K.
(*Hint*: Consider the sesquilinear form
$$\gamma(\mu,\lambda) = C_n \iint |x-y|^{-n+2} \ d\mu(x)\ d\overline{\lambda}(y) = -(\Delta^{-1}\mu, \lambda)$$

as a (positive-definite) inner product on the Hilbert space $H_K^{-1}(\mathbb{R}^n) = \{u \in H^{-1}(\mathbb{R}^n) : \text{supp } u \subset K\}$. Thus

$$|\gamma(\mu, \lambda)| \le \gamma(\mu, \mu)^{1/2}\, \gamma(\lambda, \lambda)^{1/2}.$$

Take $\mu = (\text{cap } K)^{-1}\mu_K \in \mathcal{P}_K^+$, where μ_K is the measure in (6.8)–(6.10). Show that (at least, when ∂K is smooth),

$$\lambda \in \mathcal{P}_K^+ \cap H_K^{-1}(\mathbb{R}^n) \Longrightarrow \gamma(\mu, \lambda) = \frac{1}{\text{cap } K} \int U_K(y)\, d\lambda(y) = \frac{1}{\text{cap } K},$$

and conclude that $\gamma(\lambda, \lambda) \ge 1/(\text{cap } K)$. Then use some limiting arguments.)

8. If $K \subset \mathbb{R}^3$ is compact, relate cap K to the zero frequency limit of the scattering amplitude, defined in Chapter 9, §1.
9. Try to establish directly the equivalence between the regularity criteria given by Propositions 5.5 and 6.10.
10. In Chapter 5, §5, a compact set $K \subset \mathbb{R}^n$ was called "negligible" provided there is no nonzero $u \in H^{-1}(\mathbb{R}^n)$ supported on K. Show that if K is negligible, then cap $K = 0$. Try to prove the converse.
11. Sharpen the subadditivity result (6.24) to

$$\text{cap}(K \cup L) + \text{cap}(K \cap L) \le (\text{cap } K) + (\text{cap } L),$$

for compact sets K and L. This property is called "strong subadditivity." (*Hint*: By (6.6), $U_K(x) = W_x(S_K)$, where $S_K = \{\omega : \mathfrak{h}_K(\omega) < \infty\}$. Show that $S_{K\cup L} = S_K \cup S_L$ and $S_{K\cap L} = S_K \cap S_L$, and deduce that

$$U_{K\cup L}(x) + U_{K\cap L}(x) \le U_K(x) + U_L(x).$$

Extending the reasoning used in the proof of Proposition 6.5, deduce that

$$\begin{aligned} \text{cap } K + \text{cap } L &= \int \Big[U_K(x) + U_L(x)\Big]\, d\mu_{K\cup L}(x) \\ &\ge \int \Big[U_{K\cup L}(x) + U_{K\cap L}(x)\Big]\, d\mu_{K\cup L}(x) \\ &= \text{cap}(K \cup L) + \text{cap}(K \cap L).) \end{aligned}$$

7. Stochastic integrals

We will motivate the introduction of the stochastic integral by modifying the Feynman-Kac formula, to produce a formula for the solution operator $e^{t(\Delta+X)}$ to

$$\frac{\partial u}{\partial t} = \Delta u + Xu, \quad u(0) = f; \quad Xu = \sum X_j(x) \frac{\partial u}{\partial x_j}. \tag{7.1}$$

As in (2.2), we use the Trotter product formula to write

$$e^{t(\Delta+X)} f = \lim_{k\to\infty} \Big(e^{(t/k)X} e^{(t/k)\Delta}\Big)^k f. \tag{7.2}$$

If we assume that each coefficient X_j of the vector field X is bounded and uniformly Lipschitz, then Proposition A.2 applies to (7.2), given $f \in$

$L^p(\mathbb{R}^n)$, $1 \le p < \infty$, or $f \in C_o(\mathbb{R}^n)$, in view of Proposition 9.13 in Appendix A. Now, for any k, $\left(e^{(t/k)X}e^{(t/k)\Delta}\right)^k f$ can be expressed as a k-fold integral:

$$\begin{aligned}&\left(e^{(t/k)X}e^{(t/k)\Delta}\right)^k f(x)\\ &\quad = \int \cdots \int f(x_k)p\big(\frac{t}{k}, x_k - x_{k-1} - \frac{t}{k}\xi_{k-1}\big)\\ &\qquad \cdots p\big(\frac{t}{k}, x_2 - x_1 - \frac{t}{k}\xi_1\big)p\big(\frac{t}{k}, x_1 - x - \frac{t}{k}\xi_0\big)\, dx_1 \cdots dx_k,\end{aligned} \tag{7.3}$$

where (with $x_0 = x$)

$$\xi_j = X(x_j) + r_j, \quad r_j = O(k^{-1}). \tag{7.4}$$

Now we can write

$$p\big(\frac{t}{k}, x_{j+1} - x_j - \frac{t}{k}\xi_j\big) = p\big(\frac{t}{k}, x_{j+1} - x_j\big)\, e^{\xi_j \cdot (x_{j+1}-x_j)/2 - (t/k)|\xi_j|^2/4}. \tag{7.5}$$

Consequently, parallel to (2.4),

$$\left(e^{(t/k)X}e^{(t/k)\Delta}\right)^k f(x) = E_x(\varphi_k), \tag{7.6}$$

where

$$\varphi_k(\omega) = f\big(\omega(t)\big)e^{A_k(\omega) - B_k(\omega)}, \tag{7.7}$$

with

$$\begin{aligned} A_k(\omega) &= \frac{1}{2}\sum_{j=0}^{k-1}\Big[X\big(\omega(\frac{j}{k}t)\big) + r_j\Big] \cdot \Big[\omega\big(\frac{j+1}{k}t\big) - \omega\big(\frac{j}{k}t\big)\Big]\\ B_k(\omega) &= \frac{1}{4}\frac{t}{k}\sum_{j=0}^{k-1}\Big[X\big(\omega(\frac{j}{k}t)\big) + r_j\Big]^2. \end{aligned} \tag{7.8}$$

Thus we expect to establish a formula of the form

$$e^{t(\Delta + X)}f(x) = E_x\left(f\big(\omega(t)\big)e^{A(t,\omega) - B(t,\omega)}\right), \tag{7.9}$$

where

$$B(t,\omega) = \frac{1}{4}\int_0^t X\big(\omega(s)\big)^2\, ds, \tag{7.10}$$

and

$$A(t,\omega) = \frac{1}{2}\lim_{k\to\infty}\sum_{j=0}^{k-1} X\big(\omega(\frac{j}{k}t)\big) \cdot \Big[\omega\big(\frac{j+1}{k}t\big) - \omega\big(\frac{j}{k}t\big)\Big]. \tag{7.11}$$

In (7.10), $X(\omega)^2$ denotes $\sum X_j(\omega)^2$. If the coefficients X_j are real-valued, this is equal to $|X(\omega)|^2$.

Certainly $B_k(\omega) \to B(t,\omega)$ nicely for all $\omega \in \mathfrak{P}_0$. The limit we now need to investigate is (7.11), which we would like to write as

$$A(t,\omega) = \frac{1}{2}\int_0^t X\big(\omega(s)\big)\cdot d\omega(s). \tag{7.12}$$

However, $\omega(s)$ has unbounded variation for W_x-almost all ω, so there remains some analysis to be done on this object, which is a prime example of a stochastic integral.

We aim to make sense out of stochastic integrals of the form

$$\int_0^t g\big(s,\omega(s)\big)\cdot d\omega(s), \tag{7.13}$$

beginning with

$$\int_0^t g(s)\cdot d\omega(s) = \lim_{k\to\infty}\sum_{j=0}^{k-1} g\big(\frac{j}{k}t\big)\cdot\Big[\omega\big(\frac{j+1}{k}t\big)-\omega\big(\frac{j}{k}t\big)\Big]. \tag{7.14}$$

This is readily seen to be well defined in $L^2(\mathfrak{P}_0, dW_x)$, in view of the fact that the terms $\theta_j(\omega) = \omega\big((j+1)t/k\big) - \omega\big(jt/k\big)$ satisfy

$$\|\theta_j\|^2_{L^2(\mathfrak{P}_0,dW_x)} = 2\frac{t}{k}, \quad (\theta_j,\theta_\ell)_{L^2(\mathfrak{P}_0,dW_x)} = 0, \quad \text{for } j\neq \ell, \tag{7.15}$$

the first by (1.38). Thus

$$\Big\|\sum_{j=0}^{k-1} g\big(\frac{j}{k}t\big)\Big[\omega\big(\frac{j+1}{k}t\big)-\omega\big(\frac{j}{k}t\big)\Big]\Big\|^2_{L^2(\mathfrak{P}_0,dW_x)} = 2\sum_{j=0}^{k-1}\frac{t}{k}\big|g\big(\frac{j}{k}t\big)\big|^2. \tag{7.16}$$

For continuous g, this is a Riemann sum approximating $\int_0^t |g(s)|^2\, ds$, as $k\to\infty$. Thus we obtain the following:

Proposition 7.1. *Given $g \in C\big([0,t]\big)$, the right side of (7.14) converges in $L^2(\mathfrak{P}_0, dW_x)$. The resulting correspondence*

$$g \mapsto \int_0^t g(s)\, d\omega(s)$$

extends uniquely to $\sqrt{2}$ times an isometry of $L^2\big([0,t],dt\big)$ into $L^2(\mathfrak{P}_0, dW_x)$.

We next consider

$$S_k(\omega) = \sum_{j=0}^{k-1} g\big(t_j,\omega(t_j)\big)\cdot\big[\omega(t_{j+1})-\omega(t_j)\big] = \sum_{j=0}^{k-1} g_j(\omega)\cdot\theta_j(\omega), \tag{7.17}$$

where $\theta_j(\omega) = \omega(t_{j+1}) - \omega(t_j),\ t_j = (j/k)t$. Following [Si], Chapter V, we compute

$$\|S_k\|^2_{L^2(\mathfrak{P}_0,dW_x)} = \sum_{j,\ell} E_x\Big(g_j(\omega)\theta_j(\omega)g_\ell(\omega)\theta_\ell(\omega)\Big). \tag{7.18}$$

If $\ell > j$, $\theta_\ell(\omega) = \omega(t_{\ell+1}) - \omega(t_\ell)$ is independent of the other factors in parentheses on the right side of (7.18), so the expectation of the product is equal to $E_x(g_j\theta_j g_\ell)E_x(\theta_\ell) = 0$ since $E_x(\theta_\ell) = 0$. Similarly the terms in the sum in (7.18) vanish when $\ell < j$, so

$$\begin{aligned} \|S_k\|^2_{L^2(\mathfrak{P}_0,dW_x)} &= \sum_j E_x\big(|g_j(\omega)|^2\big)E_x\big(|\theta_j|^2\big) \\ &= 2\sum_j E_x\big(|g(t_j,\omega(t_j))|^2\big)(t_{j+1}-t_j). \end{aligned} \tag{7.19}$$

If g and ω are continuous, this is a Riemann sum approximating the integral $2\int_0^t E_x\big(|g(s,\omega(s))|^2\big)\,ds$, and we readily obtain the following result.

Proposition 7.2. *Given $g \in BC\big([0,t]\times\mathbb{R}^n\big)$, the expression (7.17) converges as $k\to\infty$, in $L^2(\mathfrak{P}_0,dW_x)$, to a limit we denote by (7.13). Furthermore, the map*

$$g \mapsto \int_0^t g\big(s,\omega(s)\big)\cdot d\omega(s)$$

is $\sqrt{2}$ times an isometry into $L^2(\mathfrak{P}_0,dW_x)$, when g has the square norm

$$Q_x(g) = \int_0^t E_x\Big(\big|g\big(s,\omega(s)\big)\big|^2\Big)\,ds. \tag{7.20}$$

Note that $Q_x(g) = \int_0^t\int_{\mathbb{R}^n}|g(s,y)|^2\,p(s,x-y)\,dy\,ds$. In case $g = g(\omega(s))$, we have $Q_x(g)$ given as the square of a weighted L^2-norm:

$$Q_x(g) = \int\limits_{\mathbb{R}^n}|g(y)|^2\,r_t(x-y)\,dy = R_t(D)|g|^2(x), \tag{7.21}$$

where

$$R_t(D) = \Delta^{-1}(e^{t\Delta}-I),\quad r_t(x) = R_t(D)\delta(x). \tag{7.22}$$

We see that $R_t(D)\in OPS^{-2}(\mathbb{R}^n)$. The convolution kernel $r_t(x)$ is smooth on $\mathbb{R}^n\setminus 0$ and rapidly decreasing as $|x|\to\infty$. More precisely, one easily verifies that

$$r_t(x) \le C(n,t)|x|^{-2}e^{-|x|^2/4t},\ \text{ for } |x|\ge\frac{1}{2}, \tag{7.23}$$

and

$$r_t(x)\le C(n,t)|x|^{2-n},\ \text{ for } |x|\le\frac{1}{2},\quad n\ge 3, \tag{7.24}$$

with $|x|^{2-n}$ replaced by $\log 1/|x|$ for $n=2$ and by 1 for $n=1$. Of course, $r_t(x)>0$ for all $t>0, x\in\mathbb{R}^n\setminus 0$.

In particular, the integral in (7.21) is absolutely convergent and $Q_x(g)$ is a continuous function of x provided

(7.25)
$$g \in L^p_{\text{loc}}(\mathbb{R}^n), \text{ for some } p > n, \quad \text{and} \quad g \in L^2\big(\mathbb{R}^n, \langle x\rangle^{-2} e^{-|x|^2/4t}\, dx\big).$$

Proposition 7.2 is adequate to treat the case where the coefficients X_j are in $BC(\mathbb{R}^n)$ and purely imaginary. Since $A_k(\omega) \to A(t,\omega)$ in $L^2(\mathfrak{P}_0, dW_x)$,

$$e^{A_k(\omega)} \longrightarrow e^{A(t,\omega)} \text{ in measure,} \tag{7.26}$$

and boundedly, since the terms in (7.26) all have absolute value 1. Then convergence of (7.6) follows from the dominated convergence theorem. In such a case, $X(\omega)^2$ in (7.10) is equal to $-|X(\omega)|^2$. We have the following.

Proposition 7.3. *If $X = iY$ is a vector field on $\mathbb{R}^n$ with coefficients that are bounded, continuous, and purely imaginary, then*

(7.27)
$$e^{t(\Delta+iY)} f(x) = E_x\Big(f\big(\omega(t)\big) e^{(i/2)\int_0^t Y(\omega(s))\cdot d\omega(s) + (1/4)\int_0^t |Y(\omega(s))|^2\, ds}\Big).$$

One final ingredient is required to prove Proposition 7.3, since in this case e^{tX} is not a semigroup of bounded operators, so we cannot apply Proposition A.2. However, we can apply Proposition A.3, with

$$S(t)f(x) = \int f(y) p\big(t, y - x - tX(x)\big)\, dy.$$

If $X = iY$ is purely imaginary, then, parallel to (7.5), we have

$$p\big(t, y - x - itY(x)\big) = p(t, y-x) e^{iY(x)\cdot(y-x)/2 + t|Y(x)|^2/4}.$$

If V is bounded and continuous, a simple modification of the analysis above, combining techniques of §2, yields

$$e^{t(\Delta+X-V)} f(x) = E_x\Big(f\big(\omega(t)\big) e^{A(t,\omega)/2 - B(t,\omega)/4 - \int_0^t V(\omega(s))\, ds}\Big) \tag{7.28}$$

when X is purely imaginary. For another interpretation of this, consider

$$\begin{aligned} H &= \sum_j \Big(-i\frac{\partial}{\partial x_j} - A_j(x)\Big)^2 + V \\ &= -\sum_j \Big(\frac{\partial^2}{\partial x_j^2} - 2iA_j \frac{\partial}{\partial x_j} - i\frac{\partial A_j}{\partial x_j} - A_j^2\Big) + V. \end{aligned} \tag{7.29}$$

Assume each A_j is real-valued, and $A_j, \partial A_j/\partial x_j \in BC(\mathbb{R}^n)$. Then

$$
\begin{aligned}
e^{-tH} f(x) &= E_x\Big(f\big(\omega(t)\big)e^{S(t,\omega)}\Big), \\
S(t,\omega) &= i\int_0^t A\big(\omega(s)\big)\cdot d\omega(s) \\
&\qquad - i\int_0^t (\operatorname{div} A)\big(\omega(s)\big)\,ds - \int_0^t V\big(\omega(s)\big)\,ds.
\end{aligned}
\tag{7.30}
$$

Compare with the derivation in [Si], Chapter V.

If the coefficients of X are not assumed to be purely imaginary, we need some more estimates. More generally, we will derive further estimates on the approximants $S_k(\omega)$ to $\int_0^t g(s,\omega(s))\cdot d\omega(s)$, defined by (7.17).

Lemma 7.4. *If g is bounded and continuous, then*

$$
E_x\big(e^{S_k}\big) \le e^{t\gamma^2}, \quad \gamma = \|g\|_{L^\infty}, \tag{7.31}
$$

and

$$
E_x\big(e^{\lambda|S_k|}\big) \le 2e^{t\lambda^2\gamma^2}. \tag{7.32}
$$

Proof. The left side of (7.31) is

$$
\begin{aligned}
E_x\big(e^{g_0\theta_0}\cdots e^{g_{k-1}\theta_{k-1}}\big) &= \sum_{\nu=0}^{\infty}\frac{1}{\nu!}E_x\Big(e^{g_0\theta_0}\cdots e^{g_{k-2}\theta_{k-2}}g_{k-1}^{\nu}\theta_{k-1}^{\nu}\Big) \\
&\le \sum_{\nu=0}^{\infty}\frac{\gamma^\nu}{\nu!}E_x\big(e^{g_0\theta_0}\cdots e^{g_{k-2}\theta_{k-2}}\big)E_x\big(\theta_{k-1}^{\nu}\big) \\
&= E_x\big(e^{g_0\theta_0}\cdots e^{g_{k-2}\theta_{k-2}}\big)E_x\big(e^{\gamma\theta_{k-1}}\big),
\end{aligned}
\tag{7.33}
$$

by independence arguments such as used in the analysis of (7.18). Note that the sums over ν above have terms that vanish for odd ν. Now $E_x(e^{\gamma\theta_j}) = e^{(t_{j+1}-t_j)\gamma^2}$. An inductive argument leads to (7.31), and (7.32) follows from this plus $e^{|u|} \le e^u + e^{-u}$.

We next estimate the $L^2(\mathfrak{P}_0, dW_x)$-norm of $S_{2k} - S_k$. Another calculation, parallel to (7.18)–(7.19), yields

$$
\begin{aligned}
&\|S_{2k} - S_k\|^2_{L^2(\mathfrak{P}_0,dW_x)} \\
&= \sum_j E_x\Big(\big|g\big(t_{j+1/2},\omega(t_{j+1/2})\big) - g\big(t_j,\omega(t_j)\big)\big|^2\Big)(t_{j+1} - t_{j+1/2}),
\end{aligned}
\tag{7.34}
$$

where $t_j = jt/k$ as in (7.17), and $t_{j+1/2} = (j+1/2)t/k$. If we assume a Lipschitz condition on g, we obtain the following estimate.

Lemma 7.5. *Assume that*

$$|g(t,x) - g(s,y)|^2 \le C_0|t-s|^2 + C_1|x-y|^2. \tag{7.35}$$

Then

$$\|S_{2k} - S_k\|^2_{L^2(\mathfrak{P}_0,dW_x)} \le C_0\frac{t^2}{k^2} + 2C_1\frac{t}{k}. \tag{7.36}$$

Proof. This follows from (7.34) plus $E_x\big(|\omega(t)-\omega(s)|^2\big) = 2|t-s|$.

We can now make an estimate directly relevant to the limiting behavior of (7.7).

Lemma 7.6. *Given the bound* $\|g\|_{L^\infty} \le \gamma$, *we have*

$$\|e^{S_{2k}} - e^{S_k}\|_{L^1(\mathfrak{P}_0,dW_x)} \le \sqrt{2}\|S_{2k} - S_k\|_{L^2(\mathfrak{P}_0,dW_x)}e^{32t\gamma^2}. \tag{7.37}$$

Proof. Using $e^u - e^v = (u-v)\Phi(u,v)$, with $|\Phi(u,v)| \le e^{2|u|+2|v|}$, we have

$$\|e^{S_{2k}} - e^{S_k}\|_{L^1(\mathfrak{P}_0)} \le \|S_{2k} - S_k\|_{L^2(\mathfrak{P}_0)} \cdot \|e^{4|S_{2k}|+4|S_k|}\|^{1/2}_{L^1(\mathfrak{P}_0)}; \tag{7.38}$$

and the estimate (7.32), plus $2e^{u+v} \le e^{2u} + e^{2v}$, then yields (7.37).

With these estimates, we can pass to the limit in (7.6)–(7.7), obtaining the following result.

Proposition 7.7. *If X is a real vector field on $\mathbb{R}^n$ whose coefficients are bounded and uniformly Lipschitz, and if $f \in C_0^\infty(\mathbb{R}^n)$, then*

(7.39)
$$e^{t(\Delta+X)}f(x) = E_x\Big(f\big(\omega(t)\big)e^{(1/2)\int_0^t X(\omega(s))\cdot d\omega(s) - (1/4)\int_0^t |X(\omega(s))|^2\,ds}\Big).$$

Now that the identity (7.39) is established for X and f such as described above, one can use limiting arguments to extend the identity to more general cases. Such extensions are left to the reader.

We now evaluate the stochastic integral $\int_0^t \omega(s)\,d\omega(s)$ in the case of one-dimensional Brownian motion. One might anticipate that it should be $\omega(t)^2/2 - \omega(0)^2/2$. However, guesses based on what should happen if ω had bounded variation can be misleading, and the truth is a little stranger. Let us begin with

$$\begin{aligned} \omega(t)^2 - \omega(0)^2 &= \sum_{j=0}^{k-1}\big[\omega(t_{j+1})^2 - \omega(t_j)^2\big] \\ &= \sum_j \big[\omega(t_{j+1}) + \omega(t_j)\big]\cdot\big[\omega(t_{j+1}) - \omega(t_j)\big], \end{aligned} \tag{7.40}$$

where $t_j = (j/k)t$, as in (7.17). We also use $\theta_j(\omega) = \omega(t_{j+1}) - \omega(t_j)$ below. Recalling that $\int_0^t \omega(s)\, d\omega(s)$ is the limit of $\sum \omega(t_j)[\omega(t_{j+1}) - \omega(t_j)]$, we write (7.40) as

$$\omega(t)^2 - \omega(0)^2 = 2\sum_{j=0}^{k-1} \omega(t_j)\theta_j(\omega) + \sum_{j=0}^{k-1} \theta_j(\omega)^2. \tag{7.41}$$

The next result is the key to the computation.

Lemma 7.8. *Given* $t > 0$,

$$\Theta_k(\omega) = \sum_{j=0}^{k-1} \Big[\omega\Big(\frac{j+1}{k}t\Big) - \omega\Big(\frac{j}{k}t\Big)\Big]^2 \longrightarrow 2t \ \text{ in } L^2(\mathfrak{P}_0, dW_x), \tag{7.42}$$

as $k \to \infty$.

Proof. We have

$$\begin{aligned} E_x\big(|\Theta_k - 2t|^2\big) &= E_x\Big(\Big|\sum_j \Big[\theta_j(\omega)^2 - 2\frac{t}{k}\Big]\Big|^2\Big) \\ &= \sum_j E_x\Big(\Big[\theta_j(\omega)^2 - 2\frac{t}{k}\Big]^2\Big), \end{aligned} \tag{7.43}$$

the last identity by independence of the different θ_j. Now we know that $E_x(\theta_j^2) = 2t/k$; furthermore, generally $E_x\big([F - E_x(F)]^2\big) \le E_x(F^2)$, so it follows that

$$E_x\big(|\Theta_k - t|^2\big) \le \sum_j E_x(\theta_j^4) = 12\frac{t^2}{k}. \tag{7.44}$$

This proves the lemma.

Thus, as $k \to \infty$, the right side of (7.41) converges in $L^2(\mathfrak{P}_0, dW_x)$ to $\int_0^t \omega(s)\ d\omega(s) + t$. This gives the identity

$$\int_0^t \omega(s)\ d\omega(s) = \frac{1}{2}\big[\omega(t)^2 - \omega(0)^2 - 2t\big], \tag{7.45}$$

for W_x-almost all ω.

More generally, for sufficiently smooth f, we can write

$$f\big(\omega(t)\big) - f\big(\omega(0)\big) = \sum_{j=0}^{k-1} \big[f\big(\omega(t_{j+1})\big) - f\big(\omega(t_j)\big)\big] \tag{7.46}$$

and use the expansion

$$\begin{aligned} &f\big(\omega(t_{j+1})\big) - f\big(\omega(t_j)\big) \\ &\quad = \theta_j(\omega) f'\big(\omega(t_j)\big) + \frac{1}{2}\theta_j(\omega)^2 f''\big(\omega(t_j)\big) + O\big(|\theta_j(\omega)|^3\big) \end{aligned} \tag{7.47}$$

to generalize (7.45) to Ito's fundamental identity:

$$f\big(\omega(t)\big) - f\big(\omega(0)\big) = \int_0^t f'\big(\omega(s)\big)\ d\omega(s) + \int_0^t f''\big(\omega(s)\big)\ ds, \tag{7.48}$$

for one-dimensional Brownian motion. For n-dimensional Brownian motion and functions of the form $f = f(t, x)$, this generalizes to

$$\begin{aligned} f\big(t,\omega(t)\big) &- f\big(0,\omega(0)\big) \\ &= \int_0^t \big(\nabla_x f\big)\big(s,\omega(s)\big)\cdot d\omega(s) \\ &\quad + \int_0^t \big(\Delta f\big)\big(s,\omega(s)\big)\ ds + \int_0^t f_t\big(s,\omega(s)\big)\ ds. \end{aligned} \tag{7.49}$$

Another way of writing this is

$$df\big(t,\omega(t)\big) = (\nabla_x f)\cdot d\omega + (\Delta f)\,dt + f_t\,dt. \tag{7.50}$$

We remind the reader that our choice of $e^{t\Delta}$ rather than $e^{t\Delta/2}$ to define the transition probabilities for Brownian paths leads to formulas that sometimes look different from those arising from the latter convention, which for example would replace $(\Delta f)\,dt$ by $(1/2)(\Delta f)\,dt$ in (7.50).

Note in particular that

$$d\big(e^{\lambda\omega(t)-\lambda^2 t}\big) = \lambda\ e^{\lambda\omega(t)-\lambda^2 t}\ d\omega(t);$$

in other words, we have a solution to the "stochastic differential equation":

$$d\mathfrak{X} = \lambda\mathfrak{X}\ d\omega(t), \quad \mathfrak{X}(t) = e^{\lambda\omega(t)-\lambda^2 t}, \tag{7.51}$$

for W_0-almost all ω. Recall from (4.16) that this is the martingale $\mathfrak{z}_t(\omega)$.

We now discuss a dynamical theory of Brownian motion due to Langevin, whose purpose was to elucidate Einstein's work on the motion of a Brownian particle. Langevin produced the following equation for the *velocity* of a small particle suspended in a liquid, undergoing the sort of random motion investigated by R. Brown:

$$\frac{dv}{dt} = -\beta v + \omega'(t), \quad v(0) = v_0. \tag{7.52}$$

Here, the term $-\beta v$ represents the frictional force, tending to slow down the particle as it moves through the fluid. The term $\omega'(t)$, which contributes to the force, is due to "white noise," a random force whose statistical properties identify it with the time derivative of ω, which is defined, not classically, but through Propositions 7.1 and 7.2. Thus we rewrite (7.52) as the stochastic differential equation

$$dv = -\beta v\ dt + d\omega, \quad v(0) = v_0. \tag{7.53}$$

As in the case of ODE, we have $d(e^{\beta t}v) = e^{\beta t}(dv + \beta v\, dt)$, so (7.50) yields $d(e^{\beta t}v) = e^{\beta t} d\omega$, which integrates to

$$\begin{aligned} v(t) &= v_0 e^{-\beta t} + \int_0^t e^{-\beta(t-s)}\, d\omega(s) \\ &= v_0 e^{-\beta t} + \omega(t) - \beta \int_0^t e^{-\beta(t-s)} \omega(s)\, ds. \end{aligned} \tag{7.54}$$

The actual path of such a particle is given by

$$x(t) = x_0 + \int_0^t v(s)\, ds. \tag{7.55}$$

In the case $x_0 = 0, v_0 = 0$, we have

$$\begin{aligned} x(t) &= \int_0^t \int_0^s e^{-\beta(s-r)}\, d\omega(r)\, ds \\ &= \frac{1}{\beta} \int_0^t \bigl[1 - e^{-\beta(t-s)}\bigr]\, d\omega(s). \end{aligned} \tag{7.56}$$

Via the identity in (7.54), we have

$$x(t) = \int_0^t e^{-\beta(t-s)} \omega(s)\, ds. \tag{7.57}$$

Of course, the path $x(t)$ taken by such a particle is not the same as the "Brownian path" $\omega(t)$ we have been studying, but it is approximated by $\omega(t)$ in the following sense. It is observed experimentally that the frictional force component in (7.52) acts to slow down a particle in a very short time ($\sim 10^{-8}$ sec.). In other words, the dimensional quantity β in (7.52) is, in terms of units humans use to measure standard macroscopic quantities, "large." Now (7.57) implies

$$\lim_{\beta \to \infty} \beta x_\beta(t) = \omega(t), \tag{7.58}$$

where $x_\beta(t)$ denotes the path (7.57).

There has been further work on the dynamics of Brownian motion, particularly by L. Ornstein and G. Uhlenbeck [UO]. See [Nel3] for more on this, and references to other work.

Exercises

1. If $g \in C^1\bigl([0,t]\bigr)$, show that the integral of Proposition 7.1 is given by

$$\int_0^t g(s)\, d\omega(s) = g(t)\omega(t) - g(0)\omega(0) - \int_0^t g'(s)\omega(s)\, ds.$$

Show that this yields the second identity in (7.54) and the implication (7.56) $\Rightarrow$ (7.57).

2. With θ_j as in (7.15), show that

$$E_x\left(\sum_{j=0}^{k-1}|\theta_j(\omega)|^3\right) \rightarrow 0, \quad \text{as } k \rightarrow \infty.$$

(*Hint*: Use $2|\theta_j|^3 \le \varepsilon|\theta_j|^2 + \varepsilon^{-1}|\theta_j|^4$ and (7.44).)

3. Making use of Exercise 2, give a detailed proof of Ito's formula (7.48). Assume $f \in C^2(\mathbb{R})$ and

$$|D^\alpha f(x)| \le C_\varepsilon e^{\varepsilon|x|^2}, \quad \forall\, \varepsilon > 0, \quad |\alpha| \le 2.$$

More generally, establish (7.49).
Warning: The estimate of the remainder term in (7.47) is valid only when $|\omega(t_{j+1} - \omega(t_j)|$ is bounded (say $\le K$). But the probability that $|\omega(t_{j+1}) - \omega(t_j)|$ is $\ge K$ is very small.

4. Show that (7.42) implies that W_x-almost all paths ω have locally unbounded variation, on any interval $[s,t] \subset [0,\infty)$.
5. If $\psi(t,\omega) = \int_0^t g\big(s,\omega(s)\big)\cdot d\omega(s)$ is a stochastic integral given by Proposition 7.2, show that

$$E_x\Big(\psi(t,\cdot)\Big) = 0.$$

Show that $\psi(t,\cdot)$ is a martingale, that is, $E_x\Big(\psi(t,\cdot)|\mathfrak{B}_s\Big) = \psi(s,\cdot)$, for $s \le t$. Compare Exercise 2 of §8.

8. Stochastic integrals, II

In §7 we considered stochastic integrals of the form

$$h(t,\omega) = \int_0^t g\big(s,\omega(s)\big)\cdot d\omega(s), \tag{8.1}$$

where g is defined on $[0,\infty)\times\mathbb{R}^n$. This is a special case of integrals of the form

$$\psi(t,\omega) = \int_0^t \varphi(s,\omega)\cdot d\omega(s), \tag{8.2}$$

where φ is defined on $[0,\infty)\times\mathfrak{P}_0$. There are important examples of such φ which are not of the form $\varphi(s,\omega) = g\big(s,\omega(s)\big)$, such as the function h in (8.1), typically. It is important to be able to handle more general integrals of the form (8.2), for a certain class of functions φ on $[0,\infty)\times\mathfrak{P}_0$ called "adapted," which will be defined below.

To define (8.2), we extend the analysis in (7.17)–(7.19). Thus we consider

$$S_k(t,\omega) = \sum_{j=0}^{k-1}\varphi(t_j,\omega)\cdot\big[\omega(t_{j+1}) - \omega(t_j)\big] = \sum_{j=0}^{k-1}\varphi_j(\omega)\cdot\theta_j(\omega), \tag{8.3}$$

where, as before, $\theta_j(\omega) = \omega(t_{j+1}) - \omega(t_j),\ t_j = (j/k)t$. As in (7.18), we want to compute

$$\|S_k(t,\cdot)\|^2_{L^2(\mathfrak{P}_0,dW_x)} = \sum_{j,\ell} E_x(\varphi_j\theta_j\varphi_\ell\theta_\ell). \tag{8.4}$$

Following the analysis of (7.18), we want θ_ℓ to be independent of the other factors in the parentheses on the right side of (8.4) when $\ell > j$. Thus we demand of φ that

$$\varphi(s,\cdot) \text{ is independent of } \omega(t+h)-\omega(t), \quad \forall\, t \ge s, \quad h > 0. \tag{8.5}$$

Granted this, we see that the terms in the sum in (8.4) vanish when $j \neq \ell$, and

$$\begin{aligned}\|S_k(t,\cdot)\|^2_{L^2(\mathfrak{P}_0,dW_x)} &= \sum_j E_x\big(|\varphi_j|^2\big)E_x\big(|\theta_j|^2\big)\\ &= 2\sum_j E_x\big(|\varphi(t_j,\cdot)|^2\big)(t_{j+1}-t_j).\end{aligned} \tag{8.6}$$

If $\varphi \in C\big(\mathbb{R}^+, L^2(\mathfrak{P}_0, dW_x)\big)$, this is a Riemann sum approximating

$$2\int_0^t E_x\big(|\varphi(s,\cdot)|^2\big)\, ds = 2\|\varphi\|^2_{L^2([0,t]\times\mathfrak{P}_0)}.$$

We use the following spaces:

$$\begin{aligned} C\big(I,\mathcal{R}(Q)\big) &= \big\{\varphi \in C\big(I, L^2(\mathfrak{P}_0, dW_x)\big) : \varphi(t) = Q_t\varphi(t), \forall t \in I\big\},\\ L^2\big(I,\mathcal{R}(Q)\big) &= \big\{\varphi \in L^2\big(I, L^2(\mathfrak{P}_0, dW_x)\big) : \varphi(t) = Q_t\varphi(t), \forall t \in I\big\},\end{aligned} \tag{8.7}$$

where $I = [0,T]$, and, as in §4, $Q_t\varphi = E_x(\varphi|\mathfrak{B}_t)$. Elements of these spaces satisfy (8.5), by Corollary 4.4.

Proposition 8.1. *Given $\varphi \in C\big(I,\mathcal{R}(Q)\big)$, the expression (8.3) converges as $k = 2^\nu \to \infty$, in the space $C\big(I,\mathcal{R}(Q)\big)$, to a limit we denote (8.2). Furthermore, $\psi = \mathfrak{I}(\varphi)$ extends uniquely to a linear map*

$$\mathfrak{I} : L^2\big(I,\mathcal{R}(Q)\big) \to C\big(I,\mathcal{R}(Q)\big), \tag{8.8}$$

satisfying

$$\|\mathfrak{I}(\varphi)(t,\cdot)\|_{L^2(\mathfrak{P}_0,dW_x)} = \sqrt{2}\,\|\varphi\|_{L^2([0,t)\times\mathfrak{P}_0,dt\ dW_x)}. \tag{8.9}$$

Regarding continuity, note that

$$\|\mathfrak{I}(\varphi)(t+h,\cdot)-\mathfrak{I}(\varphi)(t,\cdot)\|_{L^2(\mathfrak{P}_0,dW_x)} = \sqrt{2}\,\|\varphi\|_{L^2([t,t+h]\times\mathfrak{P}_0,dt\ dW_x)}. \tag{8.10}$$

We need to verify that $\mathfrak{I}(\varphi)(t,\cdot) \in \mathcal{R}(Q_t)$. But clearly, each term $\varphi(t_j,\omega)\cdot[\omega(t_{j+1})-\omega(t_j)]$ in (8.3) belongs to $\mathcal{R}(Q_t)$ in this case, so we have the desired result.

We mention an approach to (8.8) just slightly different from that described above. Define a simple function to be a function $\varphi(t,\omega)$ that is constant in t for t in intervals of the form $[\ell 2^{-\nu},(\ell+1)2^{-\nu})$, with values in $\mathcal{R}(Q_s)$, $s=\ell 2^{-\nu}$, for some $\nu\in\mathbb{Z}^+$. For a simple function φ, the stochastic integral has a form similar to (8.3), namely,

$$\begin{aligned}\int_0^t \varphi(s,\omega)\cdot d\omega(s) = \sum_{j=0}^{\ell-1}\varphi(t_j,\omega)\cdot\bigl[\omega(t_{j+1})-\omega(t_j)\bigr] \\ +\varphi(t_\ell,\omega)\cdot\bigl[\omega(t)-\omega(t_\ell)\bigr],\end{aligned} \tag{8.11}$$

where $t_j=j2^{-\nu}$ and $t\in\bigl[\ell 2^{-\nu},(\ell+1)2^{-\nu}\bigr)$. An identity similar to (8.6), together with the denseness of the set of simple functions in $L^2(I,\mathcal{R}(Q))$, yields (8.8).

There is the following generalization of Ito's formula (7.49)–(7.50). Suppose

$$\mathfrak{X}(t)=\mathfrak{X}_0+\int_{t_0}^t u(s,\omega)\,ds+\int_{t_0}^t v(s,\omega)\,d\omega(s), \tag{8.12}$$

where $u,v\in L^2\bigl(I,\mathcal{R}(Q)\bigr)$. Then $\mathfrak{X}\in C\bigl(I,\mathcal{R}(Q)\bigr)$. We write

$$d\mathfrak{X}=u\,dt+v\,d\omega. \tag{8.13}$$

We might assume $\mathfrak{X},u$, and ω take values in $\mathbb{R}^n$ and v is $n\times n$ matrix-valued. More generally, let ω take values in $\mathbb{R}^n$, $\mathfrak{X}$ and u in $\mathbb{R}^m$, and v in $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$.

If $\mathfrak{Y}(t)=g\bigl(t,\mathfrak{X}(t)\bigr)$, with $g(t,x)$ real-valued and smooth in its arguments, then

$$\begin{aligned}d\mathfrak{Y}(t)=(\nabla_x g)\bigl(t,\mathfrak{X}(t)\bigr)\cdot d\mathfrak{X}(t) \\ +(D^2g)\bigl(t,\mathfrak{X}(t)\bigr)\bigl(d\mathfrak{X}(t),d\mathfrak{X}(t)\bigr)+g_t\bigl(t,\mathfrak{X}(t)\bigr)\,dt,\end{aligned} \tag{8.14}$$

where $(D^2g)(d\mathfrak{X},d\mathfrak{X})=\sum(\partial^2 g/\partial x_j\partial x_k)\,d\mathfrak{X}_j\cdot d\mathfrak{X}_k$ is computed, via (8.13), by the rules

$$dt\cdot dt=dt\cdot d\omega_j=d\omega_j\cdot dt=0,\quad d\omega_j\cdot d\omega_k=\delta_{jk}\,dt. \tag{8.15}$$

There is also an integral formula for $g\bigl(t,\mathfrak{X}(t)\bigr)-g(t_0,\mathfrak{X}_0)$, parallel to (7.49):

$$\begin{aligned}g\bigl(t,\mathfrak{X}(t)\bigr)=g(t_0,\mathfrak{X}_0)+\int_{t_0}^t\Bigl(\frac{\partial^2 g}{\partial x_j\partial x_k}\Bigr)v_{j\ell}v_{k\ell}\,ds \\ +\int_{t_0}^t g_t\bigl(s,\mathfrak{X}(s)\bigr)\,ds+\int_{t_0}^t\frac{\partial g}{\partial x_j}\bigl(u_j\,ds+v_{j\ell}\,d\omega_\ell\bigr).\end{aligned} \tag{8.16}$$

Here, we sum over repeated indices. The formulas (7.49) and (7.50) cover the special case $u=0,\ v=I$. The proof of (8.16) is parallel to that of (7.49).

If we apply (8.14) to $g(x) = e^{\lambda x},\ m = 1$, we obtain for

$$\mathfrak{Y}(t) = \exp\Big(\lambda \mathfrak{X}(t) - \lambda^2 \int_{t_0}^t |v(s,\omega)|^2\ ds\Big), \tag{8.17}$$
$$\mathfrak{X}(t) = \int_{t_0}^t v(s,\omega)\cdot d\omega(s),$$

the stochastic differential equation

$$d\mathfrak{Y} = \lambda \mathfrak{Y}\ v\cdot d\omega, \tag{8.18}$$

generalizing the identity (7.51).

There is another important property that $\mathfrak{Y}(t)$, defined by (8.17), has in common with $\mathfrak{z}_t(\omega) = e^{\lambda\omega(t)-\lambda^2 t}$.

Proposition 8.2. *Given $v \in L^2\big(I, \mathcal{R}(Q)\big)$, with values in $\mathbb{R}^n$, the function $\mathfrak{Y}(t)$ defined by (8.17) is a supermartingale; that is, for $s \le t$,*

$$E_x\big(\mathfrak{Y}(t)\big|\mathfrak{B}_s\big) \le \mathfrak{Y}(s), \quad W_x\text{-a.e. on } \mathfrak{P}_0. \tag{8.19}$$

Proof. We treat the case $t_0 = 0$. First suppose v_ν is a simple function, constant as a function of t on intervals of the form $[\ell 2^{-\nu}, (\ell+1)2^{-\nu})$, with values in $\mathcal{R}(Q_{\ell 2^{-\nu}})$, and $\mathfrak{Y}_\nu$ is given by (8.17), with $v = v_\nu$. We claim that $\mathfrak{Y}_\nu$ is a martingale, that is,

$$E_x\big(\mathfrak{Y}_\nu(t)\big|\mathfrak{B}_s\big) = \mathfrak{Y}_\nu(s), \quad \text{for } s \le t. \tag{8.20}$$

Suppose, for example, that $0 \le t < 2^{-\nu}$, so $v_\nu(s) = v_\nu(0)$, for $s \le t$. Now $v_\nu(0)$ is independent of $\omega(t) - \omega(s)$, so in this case

$$\begin{aligned} E_x\big(\mathfrak{Y}_\nu(t)\big|\mathfrak{B}_s\big) &= E_x\big(e^{\lambda v_\nu(0)[\omega(t)-x]-\lambda^2 t|v_\nu(0)|^2}\big|\mathfrak{B}_s\big) \\ &= e^{\lambda v_\nu(0)[\omega(s)-x]-\lambda^2 s|v_\nu(0)|^2}\cdot E_x\Big(e^{\lambda v_\nu(0)[\omega(t)-\omega(s)]-\lambda^2(t-s)|v_\nu(0)|^2}\big|\mathfrak{B}_s\Big), \end{aligned}$$

and the last conditional expectation is 1. A similar argument in the case $\ell 2^{-\nu} \le s \le t \le (\ell+1)2^{-\nu}$, using (8.11), gives

$$E_x\big(\mathfrak{Y}_\nu(t)\big|\mathfrak{B}_s\big) = \mathfrak{Y}_\nu(t_{\nu\ell})E_x\Big(e^{\lambda v_{\nu\ell}[\omega(t)-\omega(t_{\nu\ell})]-\lambda^2(t-t_{\nu\ell})|v_{\nu\ell}|^2}\big|\mathfrak{B}_s\Big) = \mathfrak{Y}_\nu(s),$$

where $t_{\nu\ell} = \ell 2^{-\nu},\ v_{\nu\ell} = v_\nu(t_{\nu\ell})$. The identity (8.20), for general $s \le t$, follows easily from this.

For general $v \in L^2\big(I, \mathcal{R}(Q)\big)$, we can take simple v_ν converging to v in the norm of this space, and then $\mathfrak{X}_\nu \to \mathfrak{X}$ in $C\big(I, \mathcal{R}(Q)\big)$, where $\mathfrak{X}_\nu(t) = \int_0^t v_\nu(s,\omega)\cdot d\omega(s)$. Passing to a subsequence, we can assume (for fixed s, t) that $\mathfrak{X}_\nu(s) \to \mathfrak{X}(s)$ and $\mathfrak{X}_\nu(t) \to \mathfrak{X}(t),\ W_x$-a.e.; hence $\mathfrak{Y}_\nu(s) \to \mathfrak{Y}(s)$ and $\mathfrak{Y}_\nu(t) \to \mathfrak{Y}(t),\ W_x$-a.e. Then (8.19) follows, by Fatou's lemma.

The case of general $t_0 \ge 0$ is easily obtained from this; one can extend $v(s,\omega)$ to be 0 for $0 \le s < t_0$.

Note in particular that $s = 0$ in (8.19) implies

$$E_x\Big(e^{\lambda \mathfrak{X}(t) - \lambda^2 \int_{t_0}^t |v(s,\cdot)|^2\, ds}\Big) \le 1. \tag{8.21}$$

Using Cauchy's inequality, we deduce that

$$E_x\Big(e^{\lambda \mathfrak{X}(t)/2}\Big) \le E_x\Big(e^{\lambda^2 \int_{t_0}^t |v(s,\cdot)|^2\, ds}\Big)^{1/2}. \tag{8.22}$$

We get a similar estimate upon replacing $v(s,\omega)$ by $-v(s,\omega)$, which converts $\mathfrak{X}(t)$ to $-\mathfrak{X}(t)$. Since $e^{|x|} \le e^x + e^{-x}$, we have (replacing λ by 2λ)

$$E_x\Big(e^{\lambda |\mathfrak{X}(t)|}\Big) \le 2E_x\Big(e^{4\lambda^2 \int_{t_0}^t |v(s,\cdot)|^2\, ds}\Big)^{1/2}. \tag{8.23}$$

Compare with Lemma 7.4. Note that the convexity of the exponential function implies

$$E_x\Big(e^{t^{-1} \int_0^t F(s,\cdot)\, ds}\Big) \le \frac{1}{t}\int_0^t E_x\Big(e^{F(s,\cdot)}\Big)\, ds. \tag{8.24}$$

Therefore, (8.23) implies

$$\begin{aligned} E_x\big(e^{\lambda|\mathfrak{X}(t)|}\big) &\le 2\left[\frac{1}{t-t_0}\int_{t_0}^t E_x\Big(e^{4\lambda^2 t|v(s\cdot)|^2}\Big) ds\right]^{1/2} \\ &\le 2 \max_{t_0 \le s \le t} E_x\Big(e^{4\lambda^2 t|v(s,\cdot)|^2}\Big)^{1/2}. \end{aligned} \tag{8.25}$$

If we expand $\mathfrak{Y}_\nu(t) = e^{\lambda \mathfrak{X}_\nu(t) - \lambda^2 \int_{t_0}^t |v_\nu(s,\cdot)|^2\, ds}$ in powers of λ, the coefficient of each λ^j is a martingale. The coefficient of λ^4, for example, is

$$\frac{1}{24}|\mathfrak{X}_\nu(t)|^4 - \frac{1}{2}\mathfrak{X}_\nu(t)^2\Big(\int_{t_0}^t |v_\nu(s,\omega)|^2\, ds\Big) + \frac{1}{2}\Big(\int_{t_0}^t |v_\nu(s,\omega)|^2\, ds\Big)^2. \tag{8.26}$$

This has expectation zero; hence

$$\begin{aligned} \frac{1}{24}E_x\big(|\mathfrak{X}_\nu(t)|^4\big) &\le \frac{1}{2}E_x\Big(\mathfrak{X}_\nu(t)^2\Big(\int_{t_0}^t |v_\nu(s,\cdot)|^2\, ds\Big)\Big) \\ &\le \frac{1}{48}E_x\big(|\mathfrak{X}_\nu(t)|^4\big) + 48E_x\Big(\Big(\int_{t_0}^t |v_\nu(s,\cdot)|^2\, ds\Big)^2\Big), \end{aligned} \tag{8.27}$$

so

$$\begin{aligned} E_x\big(|\mathfrak{X}_\nu(t)|^4\big) &\le 48^2 E_x\Big(\Big(\int_{t_0}^t |v_\nu(s,\cdot)|^2\, ds\Big)^2\Big) \\ &\le \big(48|t-t_0|\big)^2 \frac{1}{t-t_0}\int_{t_0}^t E_x\big(|v_\nu(s,\cdot)|^4\big)\, ds \\ &\le \big(48|t-t_0|\big)^2 \max_{t_0 \le s \le t} E_x\big(|v_\nu(s,\cdot)|^4\big), \end{aligned} \tag{8.28}$$

where the second inequality here uses convexity, as in (8.24). Again a use of Fatou's lemma yields for

$$\mathfrak{X}(t) = \int_{t_0}^{t} v(s,\omega)\cdot d\omega(s) \tag{8.29}$$

the estimate

$$\|\mathfrak{X}(t)\|_{L^4(\mathfrak{P}_0)} \le \bigl(48|t-t_0|\bigr)^{1/2} \max_{t_0\le s\le t} \|v(s,\cdot)\|_{L^4(\mathfrak{P}_0)}. \tag{8.30}$$

Similarly we obtain, for $t_1 < t_2$,

$$\|\mathfrak{X}(t_1)-\mathfrak{X}(t_2)\|_{L^4(\mathfrak{P}_0)} \le C_1|t_1-t_2|^{1/2} \max_{t_1\le s\le t_2} \|v(s,\cdot)\|_{L^4(\mathfrak{P}_0)}, \tag{8.31}$$

with $C_1 = \sqrt{48}$, when $\mathfrak{X}(t)$ is given by (8.29). If $\mathfrak{X}(t)$ is given more generally by (8.12), we have

$$\begin{aligned}\|\mathfrak{X}(t_1)-\mathfrak{X}(t_2)\|_{L^4(\mathfrak{P}_0)} \le C_0|t_1-t_2| &\max_{t_1\le s\le t_2} \|u(s,\cdot)\|_{L^4(\mathfrak{P}_0)} \\ &+ C_1|t_1-t_2|^{1/2} \max_{t_1\le s\le t_2} \|v(s,\cdot)\|_{L^4(\mathfrak{P}_0)}.\end{aligned} \tag{8.32}$$

The martingale maximal inequality of Proposition 4.7 extends to submartingales, but it is not obvious that it applies to the supermartingale $\mathfrak{Y}(t)$. However, it does apply to $\mathfrak{Y}_\nu(t)$, so, for each $\nu \in \mathbb{Z}^+$, we have

$$\begin{aligned}W_x\Bigl(\Bigl\{\omega\in\mathfrak{P}_0 : \sup_{t\in I(t_0,t_1)} \mathfrak{X}_\nu(t)-\mathfrak{X}_\nu(t_0) - \lambda\int_{t_0}^{t} |v_\nu(s,\omega)|^2\,ds > \beta\Bigr\}\Bigr) \\ \le e^{-\lambda\beta},\end{aligned} \tag{8.33}$$

where $I(t_0,t_1) = [t_0,t_1]\cap\mathbb{Q}$. It follows that

$$\begin{aligned}W_x\Bigl(\Bigl\{\omega\in\mathfrak{P}_0 : \sup_{t\in I(t_0,t_1)} |\mathfrak{X}_\nu(t)-\mathfrak{X}_\nu(t_0)| > \lambda\int_{t_0}^{t_1} |v_\nu(s,\omega)|^2\,ds + \beta\Bigr\}\Bigr) \\ \le 2e^{-\lambda\beta}.\end{aligned} \tag{8.34}$$

Thus, if we have

$$\int_{t_0}^{t_1} |v_\nu(s,\omega)|^2\,ds < \frac{\beta}{\lambda}, \quad \text{for } \omega\in S, \tag{8.35}$$

then

$$W_x\Bigl(S\cap\bigl\{\omega\in\mathfrak{P}_0 : \sup_{t\in I(t_0,t_1)} |\mathfrak{X}_\nu(t)-\mathfrak{X}_\nu(t_0)| > 2\beta\bigr\}\Bigr) \le 2e^{-\lambda\beta}. \tag{8.36}$$

Now

$$W_x\Big(\Big\{\omega \in \mathfrak{P}_0 : \int_{t_0}^{t_1} |v_\nu(s,\omega)|^2\, ds \geq \frac{\beta}{\lambda}\Big\}\Big) \leq \frac{\lambda}{\beta}\int_{t_0}^{t_1} E_x\big(|v_\nu(s,\cdot)|^2\big)\, ds. \tag{8.37}$$

Taking $\beta = \delta,\ \lambda = 1/\delta^2$, we deduce that if

$$\int_{t_0}^{t_1} \|v_\nu(s,\cdot)\|^2_{L^2(\mathfrak{P}_0)}\, ds < \delta^3\varepsilon, \tag{8.38}$$

then

$$W_x\Big(\Big\{\omega \in \mathfrak{P}_0 : \sup_{t\in I(t_0,t_1)} |\mathfrak{X}_\nu(t) - \mathfrak{X}_\nu(t_0)| > 2\delta\Big\}\Big) \leq \varepsilon + e^{-1/\delta}. \tag{8.39}$$

Since $\mathfrak{X}_\nu(t)$ converges to $\mathfrak{X}(t)$ in measure, locally uniformly in t, we have

$$W_x\Big(\Big\{\omega \in \mathfrak{P}_0 : \sup_{t\in I(t_0,t_1)} |\mathfrak{X}(t) - \mathfrak{X}(t_0)| > 2\delta\Big\}\Big) \leq \varepsilon + e^{-1/\delta} \tag{8.40}$$

whenever

$$\int_{t_0}^{t} \|v(s,\cdot)\|^2_{L^2(\mathfrak{P}_0)}\, ds < \delta^3\varepsilon. \tag{8.41}$$

The estimate (8.40) enables us to establish the following important result.

Proposition 8.3. *Let $I = [0,T]$. Given $v \in L^2\big(I, \mathcal{R}(Q)\big)$, so $\int_0^t v(s,\omega)\cdot d\omega(s) = \mathfrak{X}(t)$ belongs to $C\big(I,\mathcal{R}(Q)\big)$, you can define $\mathfrak{X}(t,\omega)$ so that $t \mapsto \mathfrak{X}(t,\omega)$ is continuous in t, for W_x-a.e. ω.*

Proof. Start with any measurable function on $I \times \mathfrak{P}_0$ representing $\mathfrak{X}(t)$; call it $\mathfrak{X}^b(t,\omega)$, so for each $t \in I$, $\mathfrak{X}^b(t,\cdot) = \mathfrak{X}(t)$, W_x-a.e. on $\mathfrak{P}_0$. Set $\mathfrak{X}(t,\omega) = \mathfrak{X}^b(t,\omega)$, for $t \in I \cap \mathbb{Q}$. From (8.40)–(8.41) it follows that there is a set $N \subset \mathfrak{P}_0$ such that $W_x(N) = 0$ and $\sigma_\omega(t) = \mathfrak{X}(t,\omega)$ is uniformly continuous in $t \in I \cap \mathbb{Q}$ for each $\omega \in \mathfrak{P}_0 \setminus N$. Then, for $\omega \in \mathfrak{P}_0,\ t \in I \setminus \mathbb{Q}$, define $\mathfrak{X}(t,\omega)$ by continuity:

$$\mathfrak{X}(t,\omega) = \lim_{I\cap\mathbb{Q}\ni t_\nu \to t} \mathfrak{X}^b(t_\nu,\omega), \quad \omega \in \mathfrak{P}_0 \setminus N. \tag{8.42}$$

If $\omega \in N$, define $\mathfrak{X}(t,\omega)$ arbitrarily.

To show that this works, it remains to check that, for each $t \in I$,

$$\mathfrak{X}(t,\cdot) = \mathfrak{X}(t), \quad W_x\text{-a.e. on } \mathfrak{P}_0. \tag{8.43}$$

Indeed, since $\mathfrak{X}^b(t_\nu,\cdot) \to \mathfrak{X}(t)$ in L^2-norm, passing to a subsequence we have $\mathfrak{X}^b(t_{\nu_j},\cdot) \to \mathfrak{X}(t)$ W_x-a.e. Comparing with (8.42), we have (8.43).

Exercises

1. Generalize (8.30) to show that $\mathfrak{X}(t) = \int_{t_0}^t v(s,\omega)\cdot d\omega(s)$ satisfies
$$\|\mathfrak{X}(t)\|^{2k}_{L^{2k}(\mathfrak{P}_0)} \le C_k|t-t_0|^{k-1}\int_{t_0}^t \|v(s,\cdot)\|^{2k}_{L^{2k}(\mathfrak{P}_0)}\, ds,$$
for $k\in\mathbb{Z}^+$.
2. Given $\varphi\in L^2\Big([0,\infty),\mathcal{R}(Q)\Big)$, show that, for $t\ge s$,
$$E_x\left(\int_s^t \varphi(\tau,\omega)\cdot d\omega(\tau)\Big|\mathfrak{B}_s\right)=0.$$
Deduce that the stochastic integral $\psi(t,\omega)=\int_0^t\varphi(s,\omega)\cdot d\omega(s)$ is a martingale, so that, for $t\ge s$,
$$E_x\Big(\psi(t,\cdot)\Big|\mathfrak{B}_s\Big)=\psi(s,\cdot).$$
3. Show that if $v(s,\omega)$ satisfies the hypotheses of Proposition 8.2, then the supermartingale $\mathfrak{Y}(t)$ in (8.17) is a martingale if and only if
$$E_x\ \mathfrak{Y}(t)=1,\quad \forall\ t\ge 0.$$

9. Stochastic differential equations

In this section we treat stochastic differential equations of the form

$$d\mathfrak{X} = b(t,\mathfrak{X})\,dt + \sigma(t,\mathfrak{X})\,d\omega,\quad \mathfrak{X}(t_0)=\mathfrak{X}_0. \tag{9.1}$$

The function $\mathfrak{X}$ is an unknown function on $I\times\mathfrak{P}_0$, where $I=[t_0,T]$. We assume $t_0\ge 0$. As in the case of ordinary differential equations, we will use the Picard iteration method, to obtain the solution $\mathfrak{X}$ as the limit of a sequence of approximate solutions to (8.1), which we write as a stochastic integral equation:

$$\mathfrak{X}(t)=\mathfrak{X}_0+\int_{t_0}^t b\big(s,\mathfrak{X}(s)\big)\ ds+\int_{t_0}^t \sigma\big(s,\mathfrak{X}(s)\big)\ d\omega(s)=\Phi\mathfrak{X}(t). \tag{9.2}$$

The last identity defines the transformation Φ, and we look for a fixed point of Φ. As usual, $\mathfrak{X}(t)$ is shorthand for $\mathfrak{X}(t,\omega)$. If ω is a Brownian path in $\mathbb{R}^n$, we can let $\mathfrak{X}$ and $b(t,x)$ take values in $\mathbb{R}^m$ and let $\sigma(t,x)$ be an $m\times n$ matrix-valued function.

Let us assume that $\sigma(t,x)$ and $b(t,x)$ are continuous in their arguments and satisfy

$$\begin{aligned} &|b(t,x)|\le K_0(1+|x|), && |b(t,x)-b(t,y)|\le L_0|x-y|,\\ &|\sigma(t,x)|\le K_1(1+|x|^2)^{1/2}, && |\sigma(t,x)-\sigma(t,y)|\le L_1|x-y|. \end{aligned} \tag{9.3}$$

We will use results of §8 to show that

$$\Phi: L^2\big(I,\mathcal{R}(Q)\big)\longrightarrow C\big(I,\mathcal{R}(Q)\big), \tag{9.4}$$

where, as in (8.7),

$$C\big(I,\mathcal{R}(Q)\big)=\big\{\varphi\in C\big(I,L^2(\mathfrak{P}_0,dW_0)\big):\varphi(t)\in\mathcal{R}(Q_t),\ \forall t\in I\big\},$$

and $L^2\big(I,\mathcal{R}(Q)\big)$ is similarly defined. Note that $\mathfrak{X}(s)$ belongs to $\mathcal{R}(Q_s)$ if and only if $\mathfrak{X}(s)$ is (equal W_0-a.e. to) a $\mathfrak{B}_s$-measurable function on $\mathfrak{P}_0$, so if $\mathfrak{X}(s)\in\mathcal{R}(Q_s)$, then also $\sigma\big(s,\mathfrak{X}(s)\big)$ and $b\big(s,\mathfrak{X}(s)\big)$ belong to $\mathcal{R}(Q_s)$. Thus Proposition 8.1 applies to the second integral in (9.2), and if $\mathfrak{X}_0\in\mathcal{R}(Q_{t_0})$, we have (9.4).

Applying (8.9) to estimate the second integral in (9.2), we have

$$\begin{aligned}\|\Phi\mathfrak{X}(t)-\mathfrak{X}_0\|^2_{L^2(\mathfrak{P}_0)}\le 2K_0^2&\left(\int_{t_0}^t\big(1+\|\mathfrak{X}(s)\|_{L^2(\mathfrak{P}_0)}\big)\,ds\right)^2\\&+4K_1^2\int_{t_0}^t\big(1+\|\mathfrak{X}(s)\|^2_{L^2(\mathfrak{P}_0)}\big)\,ds.\end{aligned}\tag{9.5}$$

Also (8.9) applies to an estimate of the second integral in

$$\begin{aligned}\Phi\mathfrak{X}(t)-\Phi\mathfrak{Y}(t)=&\int_{t_0}^t\big[b\big(s,\mathfrak{X}(s)\big)-b\big(s,\mathfrak{Y}(s)\big)\big]\,ds\\&+\int_{t_0}^t\big[\sigma\big(s,\mathfrak{X}(s)\big)-\sigma\big(s,\mathfrak{Y}(s)\big)\big]\,ds.\end{aligned}\tag{9.6}$$

We get

$$\begin{aligned}\|\Phi\mathfrak{X}(t)-\Phi\mathfrak{Y}(t)\|^2_{L^2(\mathfrak{P}_0)}\le 2L_0^2&\left(\int_{t_0}^t\|\mathfrak{X}(s)-\mathfrak{Y}(s)\|_{L^2(\mathfrak{P}_0)}\,ds\right)^2\\&+4L_1^2\int_{t_0}^t\|\mathfrak{X}(s)-\mathfrak{Y}(s)\|^2_{L^2(\mathfrak{P}_0)}\,ds.\end{aligned}\tag{9.7}$$

To solve (9.2), we take $\mathfrak{X}_0(t,\omega)=\mathfrak{X}_0(\omega)$, the given initial value, and inductively define $\mathfrak{X}_{j+1}=\Phi\mathfrak{X}_j$. Note that

$$\mathfrak{X}_1(t,\omega)=\mathfrak{X}_0(\omega)+\int_{t_0}^t b\big(s,\mathfrak{X}_0(\omega)\big)\,ds+\int_{t_0}^t\sigma\big(s,\mathfrak{X}_0(\omega)\big)\,d\omega(s)\tag{9.8}$$

contains a stochastic integral of the form (7.14), provided $\mathfrak{X}_0(\omega)$ is constant. On the other hand, the stochastic integral yielding $\mathfrak{X}_2(t,\omega)$ is usually not even of the form (7.13), but rather of the more general form (8.2). The following estimate will readily yield convergence of the sequence $\mathfrak{X}_j$.

Lemma 9.1. *For some $M=M(T)<\infty$, we have*

$$\|\mathfrak{X}_{j+1}(t)-\mathfrak{X}_j(t)\|^2_{L^2(\mathfrak{P}_0)}\le\frac{(M|t-t_0|)^{j+1}}{(j+1)!},\quad t_0\le t\le T.\tag{9.9}$$

Proof. We establish this estimate inductively. For $j=0$, we can use (9.5), with $\mathfrak{X}=\mathfrak{X}_1$, and the $j=0$ case of (9.9) follows. Assume that (9.9) holds for $j = 0, \dots, k-1$; we need to get it for $j = k$. To do this, apply (9.7) with $\mathfrak{X} = \mathfrak{X}_k$, $\mathfrak{Y} = \mathfrak{X}_{k-1}$, to get

$$\begin{aligned} \|\mathfrak{X}_{k+1}(t) - \mathfrak{X}_k(t)\|^2_{L^2(\mathfrak{P}_0)} \le \frac{2L_0^2 M^k}{k!} \Bigl(\int_{t_0}^t |s-t_0|^{k/2}\, ds \Bigr)^2 \\ + \frac{4L_1^2 M^k}{k!} \int_{t_0}^t |s-t_0|^k\, ds. \end{aligned} \tag{9.10}$$

This is $\le \bigl(M|t-t_0|\bigr)^{k+1}/(k+1)!$ as long as M is sufficiently large for (9.9) to hold for $j = 0$ and also $M \ge 2L_0^2 \max(1,T) + 4L_1^2$.

These estimates immediately yield an existence theorem:

Theorem 9.2. *Given $0 \le t_0 < T < \infty$, $I = [t_0, T]$, if b and σ are continuous on $I \times \mathbb{R}^n$ and satisfy the estimates (9.3), and if $\mathfrak{X}_0 \in \mathcal{R}(Q_{t_0})$, then the equation (9.2) has a unique solution $\mathfrak{X} \in C\bigl(I, \mathcal{R}(Q)\bigr)$.*

Only the uniqueness remains to be demonstrated. But if $\mathfrak{X}$ and $\mathfrak{Y}$ are two such solutions, we have $\Phi\mathfrak{X} = \mathfrak{X}$ and $\Phi\mathfrak{Y} = \mathfrak{Y}$, so (9.7) implies

$$\|\mathfrak{X}(t) - \mathfrak{Y}(t)\|^2_{L^2(\mathfrak{P}_0)} \le \text{ right side of (9.7)},$$

and a Gronwall argument implies $\|\mathfrak{X}(t) - \mathfrak{Y}(t)\|_{L^2} = 0$, for all $t \in I$.

Of course, the hypothesis that b and σ are continuous in t can be weakened in ways that are obvious from an examination of (9.4)–(9.7). Allowing b and σ to be piecewise continuous in t, still satisfying (9.3), we can reduce (9.1) to the case $t_0 = 0$, by setting $b(t,x) = 0$ and $\sigma(t,x) = 0$ for $0 \le t < t_0$.

If $\mathfrak{X}_0$ has higher integrability, so does the solution $\mathfrak{X}(t)$. To see this, in case $\mathfrak{X}_0 \in L^4(\mathfrak{P}_0)$, we can exploit (8.26)–(8.30) to produce the following estimate, parallel to (9.7):

$$\begin{aligned} &\|\Phi\mathfrak{X}(t) - \Phi\mathfrak{Y}(t)\|^4_{L^4(\mathfrak{P}_0)} \le \\ &8L_0^4 \Bigl(\int_{t_0}^t \|\mathfrak{X}(s) - \mathfrak{Y}(s)\|_{L^4(\mathfrak{P}_0)}\, ds \Bigr)^4 \\ &+ 8(48^2) L_1^4 |t-t_0| \int_{t_0}^t \|\mathfrak{X}(s) - \mathfrak{Y}(s)\|^4_{L^4(\mathfrak{P}_0)}\, ds. \end{aligned} \tag{9.11}$$

Using this, assuming $\mathfrak{X}_0 \in L^4(\mathfrak{P}_0, dW_0)$, we can obtain the following analogue of (9.9):

$$\|\mathfrak{X}_{j+1}(t) - \mathfrak{X}_j(t)\|^4_{L^4(\mathfrak{P}_0)} \le \frac{\bigl(M|t-t_0|^2\bigr)^{j+1}}{(j+1)!}, \tag{9.12}$$

for $M = M(T)$, on any interval $t \in [t_0, T]$. We have the following:

Proposition 9.3. *Under the hypotheses of Theorem 9.2, if also $\mathfrak{X}_0 \in L^4(\mathfrak{P}_0, dW_0)$, then $\mathfrak{X} \in C\bigl(I, L^4(\mathfrak{P}_0, dW_0)\bigr)$.*

More generally, one can establish that $\mathfrak{X} \in C\big(I, L^{2k}(\mathfrak{P}_0)\big)$, provided $\mathfrak{X}_0 \in L^{2k}(\mathfrak{P}_0)$, $k \geq 1$. The case $2k = 4$ enables us to prove part of the following important result.

Proposition 9.4. *The solution $\mathfrak{X}(t)$ to (9.2) given by Theorem 9.2 can be represented as $\mathfrak{X}(t,\omega)$ such that, for W_0-a.e. $\omega \in \mathfrak{P}_0$, the map $t \mapsto \mathfrak{X}(t,\omega)$ is continuous in t.*

Proof. First we assume $\mathfrak{X}_0 \in L^4(\mathfrak{P}_0, dW_0)$ and give a demonstration that is somewhat parallel to that of Theorem 1.1. Given $\varepsilon > 0, \delta > 0$, and $s, t \in \mathbb{R}^+$ such that $|t-s| < \delta$, we estimate the probability that $|\mathfrak{X}(t) - \mathfrak{X}(s)| > \varepsilon$. We use the estimate

$$\|\mathfrak{X}(t) - \mathfrak{X}(s)\|^4_{L^4(\mathfrak{P}_0)} \leq C|t-s|^2, \tag{9.13}$$

$C = C(T)$, for $s, t \in [0,T]$, which follows (when $t > s$) from

$$\begin{aligned} \|\mathfrak{X}(t) - \mathfrak{X}(s)\|^4_{L^4(\mathfrak{P}_0)} \leq C\Big(\int_s^t \|b(\tau, \mathfrak{X}(\tau))\|_{L^4}\, d\tau\Big)^4 \\ + C\int_s^t \|\sigma(\tau, \mathfrak{X}(\tau))\|^4_{L^4}\, d\tau, \end{aligned} \tag{9.14}$$

together with the estimate $\|\mathfrak{X}(s)\|_{L^4} \leq C(\tau)$. Consequently, given $s, t \in \mathbb{R}^+$,

$$W_0\Big(\big\{\omega \in \mathfrak{P}_0 : |\mathfrak{X}(t,\omega) - \mathfrak{X}(s,\omega)| > \varepsilon\big\}\Big) \leq \frac{C}{\varepsilon^4}|t-s|^2. \tag{9.15}$$

Now an argument parallel to that of Lemma 1.2 gives

$$\begin{aligned} W_0\Big(\big\{\omega \in \mathfrak{P}_0 : |\mathfrak{X}(t_1,\omega) - \mathfrak{X}(t_j,\omega)| > \varepsilon, \text{ for some } j = 2, \dots, \nu\big\}\Big) \\ \leq Cr\big(\frac{\varepsilon}{2}, \delta\big), \end{aligned} \tag{9.16}$$

when $\{t_1, \dots, t_\nu\}$ is any finite set of numbers in $\mathbb{Q}^+$ such that $0 \leq t_1 < \cdots < t_\nu$ and $t_\nu - t_1 \leq \delta$, where

$$r(\varepsilon, \delta) = \min\big(1, C\delta^2\varepsilon^{-4}\big). \tag{9.17}$$

The function $r(\varepsilon,\delta)$ takes the place of $\rho(\varepsilon,\delta)$ in (1.23); as in (1.21), we have

$$\frac{r(\varepsilon,\delta)}{\delta} \to 0, \quad \text{as } \delta \to 0, \tag{9.18}$$

for each $\varepsilon > 0$. From here, one shows just as in the proof of Theorem 1.1 that, for some $Z \subset \mathfrak{P}_0$ such that $W_0(Z) = 0$, the map $t \mapsto \mathfrak{X}(t,\omega)$ is uniformly continuous on $t \in \mathbb{Q}^+$, for each $\omega \in \mathfrak{P}_0 \setminus Z$. the rest of the proof of Proposition 9.4 can be carried out just like the proof of Proposition 8.3.

We now give another demonstration of Proposition 9.4, not requiring $\mathfrak{X}_0$ to be in $L^4(\mathfrak{P}_0)$, but only in $L^2(\mathfrak{P}_0)$. In such a case, under the hypotheses, and conclusions, of Theorem 9.2, we have $\sigma\big(t, \mathfrak{X}(t)\big) \in C\big(I, \mathcal{R}(Q)\big)$. Hence Proposition 8.3 applies to the second integral in (9.2), so $\mathfrak{A}(t, \omega) = \int_{t_0}^t \sigma\big(s, \mathfrak{X}(s)\big)\, d\omega(s)$ can be represented as a continuous function of t, for W_0-a.e. $\omega \in \mathfrak{P}_0$. Furthermore, we have $b\big(t, \mathfrak{X}(t)\big) \in C\big(I, L^2(\mathfrak{P}_0)\big) \subset C\big(I, L^1(\mathfrak{P}_0)\big)$. Thus, by Fubini's theorem, the first integral in (9.2) is absolutely integrable, hence continuous in t, for W_0-a.e. ω. This establishes the desired property for the left side of (9.2).

We next investigate the dependence of the solution to (9.2) on the initial data $\mathfrak{X}_0$, in a fashion roughly parallel to the method used in §6 of Chapter 1. Thus, let $\mathfrak{Y}$ solve

$$\mathfrak{Y}(t) = \mathfrak{Y}_0 + \int_{t_0}^t b\big(s, \mathfrak{Y}(s)\big)\, ds + \int_{t_0}^t \sigma\big(s, \mathfrak{Y}(s)\big)\, d\omega(s). \tag{9.19}$$

Proposition 9.5. *Assume that $b(t, x)$ and $\sigma(t, x)$ satisfy the hypotheses of Theorem 9.2 and are also C^1 in x. If $\mathfrak{X}(t)$ and $\mathfrak{Y}(t)$ solve (9.2) and (9.19), respectively, then*

$$\|\mathfrak{X}(t) - \mathfrak{Y}(t)\|_{L^2(\mathfrak{P}_0)} \le C(t, L_0, L_1)\|\mathfrak{X}_0 - \mathfrak{Y}_0\|_{L^2(\mathfrak{P}_0)}. \tag{9.20}$$

Proof. Consider $\mathcal{Z}(t) = \mathfrak{X}(t) - \mathfrak{Y}(t)$, which satisfies the identity

$$\begin{aligned} \mathcal{Z}(t) = \mathcal{Z}_0 &+ \int_{t_0}^t b'\big(s, \mathfrak{X}(s), \mathfrak{Y}(s)\big)\mathcal{Z}(s)\, ds \\ &+ \int_{t_0}^t \sigma'\big(s, \mathfrak{X}, \mathfrak{Y}(s)\big)\mathcal{Z}(s)\, d\omega(s), \end{aligned} \tag{9.21}$$

with $\mathcal{Z}_0 = \mathfrak{X}_0 - \mathfrak{Y}_0$. Here

$$b'(s, x, y) = \int_0^1 D_x b\big(s, ux + (1-u)y\big)\, du, \tag{9.22}$$

so $b'(s, x, y)(x - y) = b(s, x) - b(s, y)$, and similarly

$$\sigma'(s, x, y) = \int_0^1 D_x \sigma\big(s, ux + (1-u)y\big)\, du. \tag{9.23}$$

We estimate the right side of (9.21) in $L^2(\mathfrak{P}_0)$. By (9.3), $|b'(s, x, y)| \le L_0$, so

$$\left\| \int_{t_0}^t b'\big(s, \mathfrak{X}(s), \mathfrak{Y}(s)\big)\mathcal{Z}(s)\, ds \right\|_{L^2} \le L_0 \int_{t_0}^t \|\mathcal{Z}(s)\|_{L^2}\, ds. \tag{9.24}$$

Since $|\sigma'(s, x, y)| \le L_1$ and $\sigma'(s, \mathfrak{X}(s), \mathfrak{Y}(s))\mathcal{Z}(s) \in \mathcal{R}(Q_s)$, we have

$$\left\| \int_{t_0}^t \sigma'\big(s, \mathfrak{X}(s), \mathfrak{Y}(s)\big)\mathcal{Z}(s)\, d\omega(s) \right\|_{L^2}^2 \le L_1^2 \int_{t_0}^t \|\mathcal{Z}(s)\|_{L^2}^2\, ds. \tag{9.25}$$

Thus the identity (9.21) implies

$$\|\mathcal{Z}(t)\|_{L^2}^2 \le 3\|\mathfrak{X}_0 - \mathfrak{Y}_0\|_{L^2}^2 + 3\big[L_0^2(t-t_0)^2 + L_1^2\big] \int_{t_0}^t \|\mathcal{Z}(s)\|_{L^2}^2\, ds. \tag{9.26}$$

Now Gronwall's inequality applied to this estimate yields (9.20).

Note that (9.21) is a linear stochastic equation for $\mathcal{Z}(t)$, of a form a little different from (9.2), if $\mathfrak{X}(s)$ and $\mathfrak{Y}(s)$ are regarded as given. On the other hand, we can regard $\mathfrak{X}, \mathfrak{Y}$, and $\mathcal{Z}$ as solving together a system of stochastic equations, of the same form as (9.2).

An important special case of (9.2) is the case $\mathfrak{X}_0 = x$, a given point of $\mathbb{R}^m$, so let us look at $\mathfrak{X}^{x,s}(t)$, defined for $t \ge s$ as the solution to

$$\mathfrak{X}^{x,s}(t) = x + \int_s^t b\big(r, \mathfrak{X}(r)\big)\, dr + \int_s^t \sigma\big(r, \mathfrak{X}(r)\big)\, d\omega(r). \tag{9.27}$$

In this case we have the following useful property, which is basically the Markov property. Let $\mathfrak{B}_s^t$ denote the σ-algebra of subsets of $\mathfrak{P}_0$ generated by all sets of the form

$$\{\omega \in \mathfrak{P}_0 : \omega(t_1) - \omega(s_1) \in A\}, \quad s \le s_1 \le t_1 \le t,\ A \subset \mathbb{R}^m \text{ Borel}, \tag{9.28}$$

plus all sets of W_0-measure zero.

Proposition 9.6. *For any fixed $t \ge s$, the solution $\mathfrak{X}^{x,s}(t)$ to (9.27) is $\mathfrak{B}_s^t$-measurable.*

Proof. By the proof of Theorem 9.2, we have $\mathfrak{X}^{x,s}(t) = \lim_{k\to\infty} \mathfrak{X}_k(t)$, where $\mathfrak{X}_0(t) = x$ and, for $k \ge 0$,

$$\mathfrak{X}_{k+1}(t) = x + \int_s^t b\big(r, \mathfrak{X}_k(r)\big)\, dr + \int_s^t \sigma\big(r, \mathfrak{X}_k(r)\big)\, d\omega(r).$$

It follows inductively that each $\mathfrak{X}_k(t)$ is $\mathfrak{B}_s^t$-measurable, so the limit also has this property.

The behavior of $\mathfrak{X}^{x,s}(t)$ will be important for the next section. We derive another useful property here.

Proposition 9.7. *For $s \le \tau \le t$, we have*

$$\mathfrak{X}^{x,s}(t,\omega) = \mathfrak{X}^{q,\tau}(t,\omega), \quad q = \mathfrak{X}^{x,s}(\tau,\omega), \tag{9.29}$$

for W_0-a.e. $\omega \in \mathfrak{P}_0$.

Proof. Let $\mathfrak{Y}(t)$ denote the right side of (9.29). Thus $\mathfrak{Y}(\tau) = \mathfrak{X}^{x,s}(\tau)$. The stochastic equation satisfied by $\mathfrak{X}^{x,s}(t)$ then implies

$$\mathfrak{Y}(t) = \mathfrak{X}^{x,s}(\tau) + \int_\tau^t b\big(r, \mathfrak{Y}(r)\big)\, dr + \int_\tau^t \sigma\big(r, \mathfrak{Y}(r)\big)\, d\omega(r).$$

Now (9.27) implies that $\mathfrak{X}^{x,s}(t)$ satisfies this same stochastic equation, for $t \geq \tau$. The identity $\mathfrak{Y}(t) = \mathfrak{X}^{x,s}(t)$ a.e. on $\mathfrak{P}_0$ follows from the uniqueness part of Theorem 9.2.

Exercises

1. Show that the solution to

$$d\mathfrak{X} = a(t)\mathfrak{X}(t)\ dt + b(t)\mathfrak{X}(t)\ d\omega(t),$$

in case $m = n = 1$, is given by

$$\mathfrak{X}(t) = \mathfrak{X}(0)\ \exp\left\{\int_0^t [a(s) - b(s)^2]\, ds + \int_0^t b(s)\, d\omega(s)\right\} = \mathfrak{X}(0)e^{\mathfrak{Z}(t)}. \tag{9.30}$$

In this problem and the following one, $\mathfrak{X}(t)$ depends on ω, but $a(t)$ and $b(t)$ do not depend on ω, nor do $f(t)$ and $g(t)$ below.

2. Show that the solution to

$$d\mathfrak{X}(t) = \Big[f(t) + a(t)\mathfrak{X}(t)\Big]\ dt + \Big[g(t) + b(t)\mathfrak{X}(t)\Big]\ d\omega(t),$$

in case $m = n = 1$, is given by $\mathfrak{X}(t) = e^{\mathfrak{Z}(t)}\mathfrak{Y}(t)$, where $e^{\mathfrak{Z}(t)}$ is as in (9.30) and

$$\mathfrak{Y}(t) = \mathfrak{X}(0) + \int_0^t \Big[e^{-\mathfrak{Z}(s)}f(s) - g(s)b(s)\Big]\ ds + \int_0^t g(s)e^{-\mathfrak{Z}(s)}\ d\omega(s).$$

3. Consider the system

$$d\mathfrak{X}(t) = \Big[A(t)\mathfrak{X}(t) + f(t)\Big]\, dt + g(t)\, d\omega(t), \tag{9.31}$$

where $A(t) \in \mathrm{End}(\mathbb{R}^m)$, $f(t) \in \mathbb{R}^m$, and $g(t) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. Suppose $S(t, s)$ is the solution operator to the linear $m \times m$ system of differential equations

$$\frac{dy}{dt} = A(t)y, \quad S(t, t) = I,$$

as considered in Chapter 1, §5. Show that the solution to (9.31) is

$$\mathfrak{X}(t) = S(t, 0)\mathfrak{X}(0) + \int_0^t S(t, s)f(s)\ ds + \int_0^t S(t, s)g(s)\ d\omega(s).$$

4. The following Langevin equation is more general than (7.52):

$$x''(t) = -\nabla V\Big(x(t)\Big) - \beta x'(t) + \omega'(t). \tag{9.32}$$

Rewrite this as a first-order system of the form (9.1). Using Exercise 3, solve this equation when $V(x)$ is the harmonic oscillator potential, $V(x) = ax^2$.

10. Application to equations of diffusion

Let $\mathfrak{X}^{x,s}(t)$ solve the stochastic equation

$$\mathfrak{X}^{x,s}(t) = x + \int_s^t b\big(\mathfrak{X}^{x,s}(r)\big)\ dr + \int_s^t \sigma\big(\mathfrak{X}^{x,s}(r)\big)\ d\omega. \tag{10.1}$$

As in (9.2), x and b can take values in $\mathbb{R}^m$ and σ values in $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. We want to study the transformations on functions on $\mathbb{R}^m$ defined by

$$\Phi_s^t f(x) = E_0 \; f\big(\mathfrak{X}^{x,s}(t)\big), \quad 0 \le s \le t. \tag{10.2}$$

Clearly, $\mathfrak{X}^{x,s}(s) = x$, so

$$\Phi_t^t f(x) = f(x). \tag{10.3}$$

We assume $b(x)$ and $\sigma(x)$ are bounded and satisfy the Lipschitz conditions of (9.3). For simplicity we have taken b and σ to be independent of t in (10.1). We claim this implies the following:

$$\Phi_0^t f(x) = \Phi_s^{t+s} f(x), \tag{10.4}$$

for $s, t \ge 0$. In fact, it is clear that

$$\mathfrak{X}^{x,s}(t+s,\omega) = \mathfrak{X}^{x,0}(t, \vartheta_s\omega), \tag{10.5}$$

where $\vartheta_s\omega(\tau) = \omega(\tau+s) - \omega(s)$, as in (4.11). The measure-preserving property of the map $\vartheta_s : \mathfrak{P}_0 \to \mathfrak{P}_0$ then implies

$$E_0 \; f\big(\mathfrak{X}^{x,0}(t,\vartheta_s\omega)\big) = E_0 \; f\big(\mathfrak{X}^{x,0}(t)\big) = \Phi_0^t f(x),$$

so we have established (10.4). Let us set

$$P^t f = \Phi_0^t f = E_0 \; f\big(\mathfrak{X}^x(t)\big), \tag{10.6}$$

where for notational convenience we have set $\mathfrak{X}^x(t) = \mathfrak{X}^{x,0}(t)$.

We will study the action of P^t on the Banach space $C_o(\mathbb{R}^m)$ of continuous functions on $\mathbb{R}^m$ that vanish at infinity.

Proposition 10.1. *For each $t \ge 0$,*

$$P^t : C_o(\mathbb{R}^m) \longrightarrow C_o(\mathbb{R}^m), \tag{10.7}$$

and P^t forms a strongly continuous semigroup of operators on $C_o(\mathbb{R}^m)$.

Proof. If $f \in C_o(\mathbb{R}^m)$, then f is uniformly continuous, that is, it has a modulus of continuity:

$$|f(x) - f(y)| \le \omega_f\big(|x-y|\big), \tag{10.8}$$

where $\omega_f(\delta)$ is a bounded, continuous function of δ such that $\omega_f(\delta) \to 0$ as $\delta \to 0$. Then

$$\begin{aligned} \big|P^t f(x) - P^t f(y)\big| &\le E_0 \; \big|f\big(\mathfrak{X}^x(t)\big) - f\big(\mathfrak{X}^y(t)\big)\big| \\ &\le E_0 \; \omega_f\big(\big|\mathfrak{X}^x(t) - \mathfrak{X}^y(t)\big|\big). \end{aligned} \tag{10.9}$$

Now if x is fixed and $y = x_\nu \to x$, then, for each $t \ge 0$, $\mathfrak{X}^x(t) - \mathfrak{X}^{x_\nu}(t) \to 0$ in $L^2(\mathfrak{P}_0)$, by Proposition 9.5. Hence $\mathfrak{X}^x(t) - \mathfrak{X}^{x_\nu}(t) \to 0$ in measure on $\mathfrak{P}_0$, so the Lebesgue dominated convergence theorem implies that (10.9) tends to 0 as $y \to x$. This shows that $P^t f \in C(\mathbb{R}^m)$ if $f \in C_o(\mathbb{R}^m)$.

To show that $P^t f(x)$ vanishes at infinity, for each $t \ge 0$, we note that, for most $\omega \in \mathfrak{P}_0$ (in a sense that will be quantified below), $|\mathfrak{X}^x(t) - x| \le C\langle t\rangle$ if C is large, so if $f \in C_o(\mathbb{R}^m)$ and $|x|$ is large, then $f\big(\mathfrak{X}^x(t,\omega)\big)$ is small for most $\omega \in \mathfrak{P}_0$.

In fact, subtracting x from both sides of (10.1) and estimating L^2-norms, we have

$$\|\mathfrak{X}^x(t) - x\|^2_{L^2(\mathfrak{P}_0)} \le 2B^2t^2 + 2S^2t, \quad B = \sup\ |b|,\ S = \sup\ |\sigma|. \tag{10.10}$$

Hence

$$W_0\Big(\big\{\omega \in \mathfrak{P}_0 : |\mathfrak{X}^x(t,\omega) - x| > \lambda\big\}\Big) \le \frac{2B^2t^2 + 2S^2t}{\lambda^2}. \tag{10.11}$$

The mapping property (10.7) follows.

We next examine continuity in t. In fact, parallel to (10.9), we have

$$\big|P^t f(x) - P^s f(x)\big| \le E_0\ \omega_f\big(\big|\mathfrak{X}^x(t) - \mathfrak{X}^x(s)\big|\big). \tag{10.12}$$

We know from §9 that $\mathfrak{X}^x(t) \in C\big(\mathbb{R}^+, L^2(\mathfrak{P}_0)\big)$, and estimates from there readily yield that the modulus of continuity can be taken to be independent of x. Then the vanishing of (10.12), uniformly in x, as $s \to t$, follows as in the analysis of (10.9).

There remains the semigroup property, $P^s P^{t-s} = P^t$, for $0 \le s \le t$. By (10.4), this is equivalent to $\Phi_0^s \Phi_s^t = \Phi_0^t$. To establish this, we will use the identity

$$E_0\Big(f\big(\mathfrak{X}^{x,s}(t)\big)\big|\mathfrak{B}_s\Big) = E_0\ f\big(\mathfrak{X}^{x,s}(t)\big) = \Phi_s^t f(x), \tag{10.13}$$

which is an immediate consequence of Proposition 9.6. If we replace s by τ in (10.13), and then replace x by $\mathfrak{X}^{x,s}(\tau)$, with $s \le \tau \le t$, and use the identity

$$\mathfrak{X}^{q,\tau}(t) = \mathfrak{X}^{x,s}(t), \quad q = \mathfrak{X}^{x,s}(\tau), \tag{10.14}$$

established in Proposition 9.7, we obtain

$$E_0\Big(f\big(\mathfrak{X}^{x,s}(t)\big)\big|\mathfrak{B}_\tau\Big) = \Phi_\tau^t f\big(\mathfrak{X}^{x,s}(\tau)\big). \tag{10.15}$$

We thus have, for $s \le \tau \le t$,

$$\begin{aligned} \Phi_s^\tau \Phi_\tau^t f(x) &= E_0\Big(\Phi_\tau^t f\big(\mathfrak{X}^{x,s}(\tau)\big)\big|\mathfrak{B}_s\Big) \\ &= E_0\Big(E_0\Big(f\big(\mathfrak{X}^{q,\tau}(t)\big)\big|\mathfrak{B}_\tau\Big)\Big|\mathfrak{B}_s\Big) \\ &= E_0\Big(f\big(\mathfrak{X}^{q,\tau}(t)\big)\big|\mathfrak{B}_s\Big), \end{aligned} \tag{10.16}$$

and again using (10.14) we see that this is equal to the left side of (10.13), hence to $\Phi_s^t f(x)$, as desired. This completes the proof of Proposition 10.1.

We want to identify the infinitesimal generator of P^t. Assume now that $D^\alpha f$, for $|\alpha| \le 2$, are bounded and continuous on $\mathbb{R}^m$. Then Ito's formula implies

$$f\big(\mathfrak{X}^x(t)\big) = f(x) + \int_0^t \Big(\frac{\partial^2 f}{\partial x_j \partial x_k}\Big)\sigma_{j\ell}\sigma_{k\ell}\; dr + \int_0^t \frac{\partial f}{\partial x_j}\big(b_j\; dr + \sigma_{j\ell}\; d\omega_\ell\big), \tag{10.17}$$

using the summation convention. Let us apply E_0 to both sides. Now

$$E_0\Big(\int_0^t \frac{\partial f}{\partial x_j}\sigma_{j\ell}\; d\omega_\ell\Big) = 0, \tag{10.18}$$

so we have

$$E_0\big(f(\mathfrak{X}^x(t))\big) = f(x) + \int_0^t E_0\Big(\frac{\partial^2 f}{\partial x_j \partial x_k}A_{jk}\Big)\; dr + \int_0^t E_0\Big(\frac{\partial f}{\partial x_j}b_j\Big)\; dr, \tag{10.19}$$

where A_{jk} in the first integral is given by

$$A_{jk}(y) = \sum_\ell \sigma_{j\ell}(y)\sigma_{k\ell}(y), \quad y = \mathfrak{X}^x(r). \tag{10.20}$$

In matrix notation,

$$A = \sigma\sigma^t. \tag{10.21}$$

We can take the t-derivative of the right side of (10.16), obtaining

$$\frac{\partial}{\partial t}P^t f(x) = E_0\Big(A_{jk}\big(\mathfrak{X}^x(t)\big)\,\partial_j\partial_k f\big(\mathfrak{X}^x(t)\big) + b_j\big(\mathfrak{X}^x(t)\big)\,\partial_j f\big(\mathfrak{X}^x(t)\big)\Big). \tag{10.22}$$

In particular,

$$\frac{\partial}{\partial t}P^t f(x)\big|_{t=0} = \sum_{j,k} A_{jk}(x)\,\partial_j\partial_k f(x) + \sum_j b_j(x)\,\partial_j f(x) = Lf(x), \tag{10.23}$$

where the last identity defines the second-order differential operator L, acting on functions of x. This is known as Kolmogorov's diffusion equation. We have shown that the infinitesimal generator of the semigroup P^t, acting on $C_o(\mathbb{R}^m)$, is a closed extension of the operator

$$L = \sum A_{jk}(x)\,\partial_j\partial_k + \sum b_j(x)\,\partial_j, \tag{10.24}$$

defined initially, let us say, on $C_0^2(\mathbb{R}^m)$.

It is clear from (10.6) that $\|P^t f\|_{L^\infty} \le \|f\|_{L^\infty}$ for each $f \in C_o(\mathbb{R}^m)$, so P^t is a contraction semigroup on $C_o(\mathbb{R}^m)$. It is also clear that

$$f \ge 0 \Longrightarrow P^t f \ge 0 \;\text{ on } \mathbb{R}^m, \tag{10.25}$$

that is, P^t is "positivity preserving." For given $x \in \mathbb{R}^n,\ t \geq 0,\ f \mapsto P^t f(x)$ is a positive linear functional on $C_o(\mathbb{R}^m)$. Hence there is a uniquely defined positive Borel measure $\mu_{x,t}$ on $\mathbb{R}^m$, of mass ≤ 1, such that

$$P^t f(x) = \int f(y)\ d\mu_{x,t}(y). \tag{10.26}$$

In fact, by the construction (10.6),

$$\mu_{x,t} = F_{(x,t)*} W_0, \tag{10.27}$$

where $F_{(x,t)}(\omega) = \mathfrak{X}^x(t,\omega)$, and (10.27) means $\mu_{x,t}(U) = W_0\big(F^{-1}_{(x,t)}(U)\big)$ for a Borel set $U \subset \mathbb{R}^m$. This implies that, for each $x, t,\ \mu_{x,t}$ is a probability measure on $\mathbb{R}^m$, since $|\mathfrak{X}^x(t)|$ is finite for W_0-a.e. $\omega \in \mathfrak{P}_0$.

We will use the notation

$$P(s,x,t,U) = \mu_{x,t-s}(U), \quad 0 \leq s \leq t,\ U \subset \mathbb{R}^m,\ \text{Borel}. \tag{10.28}$$

We can identify $P(s,x,t,U)$ with the probability that $\mathfrak{X}^{x,s}(t)$ is in U. We can rewrite (10.26) as

$$P^t f(x) = \int f(y)\ P(0,x,t,dy) \tag{10.29}$$

or

$$\Phi^t_s f(x) = \int f(y)\ P(s,x,t,dy). \tag{10.30}$$

The semigroup property on P^t implies

$$P(s,x,t,U) = \int P(s,x,\tau,dy)\ P(\tau,y,t,U), \quad 0 \leq s \leq \tau \leq t, \tag{10.31}$$

which is known as the Chapman-Kolmogorov equation.

Let us denote by $\mathcal{L}$ the extension of (10.24) that is the infinitesimal generator of P^t. If V is a bounded, continuous function on $\mathbb{R}^m$, then $\mathcal{L} - V$ generates a semigroup on $C_o(\mathbb{R}^m)$, and an application of the Trotter product formula similar to that done in §2 yields

$$e^{t(\mathcal{L}-V)} f(x) = E_0\left(f\big(\mathfrak{X}^x(t)\big)\ e^{-\int_0^t V(\mathfrak{X}^x(s))\,ds} \right). \tag{10.32}$$

This furnishes an existence result for weak solutions to the initial-value problem

$$\begin{aligned} \frac{\partial u}{\partial t} &= \sum A_{jk}(x)\,\partial_j\partial_k u + \sum b_j(x)\,\partial_j u - Vu, \\ u(0) &= f \in C_o(\mathbb{R}^m), \end{aligned} \tag{10.33}$$

under the hypotheses that V is bounded and continuous, the coefficients b_j are bounded and uniformly Lipschitz, and A_{jk} has the form (10.20), with $\sigma_{j\ell}$ bounded and uniformly Lipschitz. As for the last property, we record the following fact:

Proposition 10.2. *If $A(x)$ is a C^2 positive-semidefinite, matrix-valued function on $\mathbb{R}^m$ with $D^\alpha A(x)$ bounded on $\mathbb{R}^m$ for $|\alpha| \le 2$, then there exists a bounded, uniformly Lipschitz, matrix-valued function $\sigma(x)$ on $\mathbb{R}^m$ such that $A(x) = \sigma(x)\sigma(x)^t$.*

This result is quite easy to prove in the elliptic case, that is, when for certain $\lambda_j \in (0, \infty)$,

$$\lambda_0 |\xi|^2 \le \sum A_{jk}(x)\xi_j\xi_k \le \lambda_1 |\xi|^2, \tag{10.34}$$

but a careful argument is required if $A(x)$ is allowed to degenerate. See the exercises for more on this.

If $A_{jk}(x)$ has bounded, continuous derivatives of order ≤ 2, we can form the formal adjoint of (10.24):

$$L^t f = \sum \partial_j \partial_k \big(A_{jk}(x) f\big) - \sum \partial_j \big(b_j(x) f\big) = \widetilde{L} f - V f, \tag{10.35}$$

where $\widetilde{L}$ has the same second-order derivatives as L, though perhaps a different first-order part, and $V(x) = -\sum \partial_j \partial_k A_{jk}(x) + \sum \partial_k b_j(x)$. Thus $\widetilde{L}$ has an extension, which we denote as $\widetilde{\mathcal{L}}$, generating a contraction semigroup on $C_o(\mathbb{R}^m)$, with the positivity-preserving property. Furthermore, $\widetilde{\mathcal{L}} - V$ generates a semigroup on $C_o(\mathbb{R}^m)$, and there is a formula for $e^{t(\widetilde{\mathcal{L}}-V)} f$ parallel to (10.32). Thus we obtain a weak solution to the initial-value problem

$$\frac{\partial u}{\partial t} = \sum \partial_j \partial_k \big(A_{jk}(x) u\big) - \sum \partial_j \big(b_j(x) u\big), \quad u(0) = f \in C_o(\mathbb{R}^m), \tag{10.36}$$

provided that $A_{jk}(x)$ satisfies the conditions of Proposition 10.2, and that each b_j is bounded, with bounded, continuous first derivatives. Equation (10.36) is called the Fokker-Planck equation.

To continue, we shall make a further simplifying hypothesis, namely that the ellipticity condition (10.34) hold. We will also assume $A_{jk}(x)$ and $b_j(x)$ are C^∞, and that $D^\alpha A_{jk}(x)$ and $D^\alpha b_j(x)$ are bounded for all α. In such a case, $(g_{jk}) = (A_{jk})^{-1}$ defines a Riemannian metric on $\mathbb{R}^m$, and if Δ_g denotes its Laplace operator, we have

$$Lf = \Delta_g f + Xf, \tag{10.37}$$

for some smooth vector field $X = \sum \xi_j(x)\, \partial_j$, such that $D^\alpha \xi_j(x)$ is bounded for $|\alpha| \le 1$. Note that if we use the inner product

$$(f, g) = \int f(x)\overline{g(x)}\; dV(x), \tag{10.38}$$

where dV is the Riemannian volume element determined by the Riemannian metric g_{jk}, then this puts the same topology on $L^2(\mathbb{R}^m)$ as the inner product $\int f(x)\overline{g(x)}\; dx$. We prefer the inner product (10.38), since Δ_g is then self-adjoint.

Now consider the closed operator $\mathcal{L}_2$ on $L^2(\mathbb{R}^m)$ defined by

$$\mathcal{L}_2 f = Lf \quad \text{on } \mathcal{D}(\mathcal{L}_2) = H^2(\mathbb{R}^m). \tag{10.39}$$

It follows from results on Chapter 6, §2, that $\mathcal{L}_2$ generates a strongly continuous semigroup $e^{t\mathcal{L}_2}$ on $L^2(\mathbb{R}^n)$. To relate this semigroup to the semigroup $P^t = e^{t\mathcal{L}}$ on $C_o(\mathbb{R}^m)$ described above, we claim that

$$e^{t\mathcal{L}_2} f = e^{t\mathcal{L}} f, \quad \text{for } f \in C_0^\infty(\mathbb{R}^m). \tag{10.40}$$

To see this, let $u_0(t,x)$ and $u_1(t,x)$ denote the left and right sides, respectively. These are both weak solutions to $\partial_t u_j = Lu_j$, for which one has regularity results. Also, estimates discussed in §2 of Chapter 6 imply that $u_0(t,x)$ vanishes as $|x| \to \infty$, locally uniformly in $t \in [0,\infty)$. Thus the maximum principle applies to $u_0(t,x) - u_1(t,x)$, and we have (10.40). From here a simple limiting argument yields

$$e^{t\mathcal{L}_2} f = e^{t\mathcal{L}} f, \quad \text{for } f \in C_o(\mathbb{R}^m) \cap L^2(\mathbb{R}^m). \tag{10.41}$$

Now the dual semigroup $(e^{t\mathcal{L}_2})^*$ is a strongly continuous semigroup on $L^2(\mathbb{R}^n)$, with infinitesimal generator $\mathcal{L}_2^t$ defined by

$$\mathcal{L}_2^t f = L^t f \quad \text{on } \mathcal{D}(\mathcal{L}_2^t) = H^2(\mathbb{R}^m), \tag{10.42}$$

where L^t is given by (10.35). An argument parallel to that used to establish (10.41) shows that

$$\big(e^{t\mathcal{L}_2}\big)^* f = e^{t\mathcal{L}_2^t} f = e^{t(\widetilde{\mathcal{L}}-V)} f, \quad \text{for } f \in C_o(\mathbb{R}^m) \cap L^2(\mathbb{R}^m). \tag{10.43}$$

On the other hand, $(P^t)^* = (e^{t\mathcal{L}})^*$ is a weak*-continuous semigroup of operators on $\mathfrak{M}(\mathbb{R}^m)$, the space of finite Borel measures on $\mathbb{R}^m$; it is not strongly continuous. Using (10.43), we see that

$$\big(f, e^{t\mathcal{L}_2} g\big) = \big(e^{t(\widetilde{\mathcal{L}}-V)} f, g\big), \quad \text{for } f, g \in C_0^\infty(\mathbb{R}^m), \tag{10.44}$$

and bringing in (10.40) we have

$$\big(e^{t\mathcal{L}}\big)^* f = e^{t(\widetilde{\mathcal{L}}-V)} f, \tag{10.45}$$

for $f \in C_0^\infty(\mathbb{R}^m)$, hence for $f \in C_0(\mathbb{R}^m) \cap L^1(\mathbb{R}^m)$. From here one can deduce that $\big(e^{t\mathcal{L}}\big)^*$ preserves $L^1(\mathbb{R}^m)$ and acts as a strongly continuous semigroup on this space.

Let us return to the family of measures $P(s,x,t,\cdot)$. Under our current hypotheses, regularity results for parabolic PDE imply that, for $s < t$, there is a smooth function $p(s,x,t,y)$ such that

$$P(s,x,t,U) = \int_U p(s,x,t,y)\, dy. \tag{10.46}$$

We have

$$\Phi_s^t f(x) = \int f(y)\, p(s,x,t,y)\, dy, \quad s < t, \tag{10.47}$$

and

$$(\Phi_s^t)^* f(y) = \int f(x)\, p(s,x,t,y)\, dx, \quad s < t. \tag{10.48}$$

Furthermore, we have for $p(s,x,t,y)$ the "backward" Kolmogorov equation

$$\frac{\partial p}{\partial s} = -\sum_{j,k} A_{jk}(x) \frac{\partial^2 p}{\partial x_j \partial x_k} - \sum_j b_j(x) \frac{\partial p}{\partial x_j} \tag{10.49}$$

and the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \sum_{j,k} \frac{\partial^2}{\partial y_j \partial y_k} \Bigl(A_{jk}(y)p\Bigr) - \sum_j \frac{\partial}{\partial y_j} \Bigl(b_j(y)p\Bigr). \tag{10.50}$$

While we have restricted attention to the smooth elliptic case for the last set of results, it is also interesting to relax the regularity required on the coefficients as much as possible, and to let the coefficients depend on t, and also to allow degeneracy. See [Fdln] and [SV] for more on this. Exercise 5 below illustrates the natural occurence of degenerate L.

We mention that, working with (10.32), we can obtain the solution to

$$\begin{gathered} \frac{\partial u}{\partial t} = Lu, \quad \text{for } t \geq 0,\ x \in \Omega, \\ u(t,x) = 0, \quad \text{for } x \in \partial\Omega, \quad u(0,x) = f(x), \end{gathered} \tag{10.51}$$

by considering a sequence $V_\nu \to \infty$ on $\mathbb{R}^m \setminus \Omega$, as in the analysis in §3, when Ω is an open domain in $\mathbb{R}^m$, with smooth boundary, or at least with the regularity property used in Proposition 3.3. In analogy with (3.8), we get

$$u(t) = E_0\Bigl(f\bigl(\mathfrak{X}^x(t)\bigr)\psi_{\overline{\Omega}}(\mathfrak{X}^x, t)\Bigr), \tag{10.52}$$

where

$$\begin{aligned} \psi_{\overline{\Omega}}(\mathfrak{X}^x, t) = 1 \quad & \text{if } \mathfrak{X}^x\bigl([0,t]\bigr) \subset \overline{\Omega}, \\ 0 \quad & \text{otherwise.} \end{aligned} \tag{10.53}$$

The proof can be carried out along the same lines as in the proof of Proposition 3.3, provided $\mathcal{L}_2$ (defined in (10.39)) is self-adjoint. Otherwise a different approach is required. Also, when $\mathcal{L}_2$ is self-adjoint, the analysis leading to Proposition 3.5 extends to (10.51), for any open $\Omega \subset \mathbb{R}^m$, with no boundary regularity required. For other approaches to these matters, and also to the Dirichlet problem for $Lu = f$ on Ω, in both the elliptic and degenerate cases, see [Fdln] and [Fr].

We end this section with a look at a special case of (10.1), namely when $\sigma = I$, so we solve

$$\mathfrak{X}^x(t) = x + \omega(t) + \int_0^t b\bigl(\mathfrak{X}^x(r)\bigr)\, dr. \tag{10.54}$$

Assume as before that b is bounded and uniformly Lipschitz. Then the analysis of (10.6) done above implies

$$e^{t(\Delta+X)}f(x) = E_0\ f\big(\mathfrak{X}^x(t)\big), \quad X = \sum b_j(x)\,\partial_j. \tag{10.55}$$

On the other hand, in §7 we derived the formula

$$e^{t(\Delta+X)}f(x) = E_x\Big(f\big(\omega(t)\big)e^{\mathcal{Z}(t)}\Big), \tag{10.56}$$

where

$$\mathcal{Z}(t) = \frac{1}{2}\int_0^t b\big(\omega(s)\big)\cdot d\omega(s) - \frac{1}{4}\int_0^t \big|b\big(\omega(s)\big)\big|^2\,ds.$$

We conclude that the right-hand sides of (10.55) and (10.56) coincide. We can restate this identity as follows. Given $x \in \mathbb{R}^n$, we have a map

$$\Xi^x : \mathfrak{P}_0 \to \mathfrak{P}_0, \quad \Xi^x(\omega)(t) = \mathfrak{X}^x(t). \tag{10.57}$$

Then Wiener measure W_0 on $\mathfrak{P}_0$ gives rise to a measure $\Xi^x_* W_0$ on $\mathfrak{P}_0$, by

$$\Xi^x_* W_0(S) = W_0\big((\Xi^x)^{-1}(S)\big). \tag{10.58}$$

For example, if $0 \le t_1 < \cdots < t_k$,

$$\int\limits_{\mathfrak{P}_0} F\big(\omega(t_1),\dots,\omega(t_k)\big)\ d\Xi^x_* W_0 = E_0\ F\big(\mathfrak{X}^x(t_1),\dots,\mathfrak{X}^x(t_k)\big). \tag{10.59}$$

Thus the identity of (10.55) and (10.56) can be written as

$$\int\limits_{\mathfrak{P}_0} f\big(\omega(t)\big)\ d\Xi^x_* W_0 = \int\limits_{\mathfrak{P}_0} f\big(\omega(t)\big)e^{\mathcal{Z}(t)}\ dW_x. \tag{10.60}$$

This is a special case of the following result of Cameron-Martin and Girsanov:

Proposition 10.3. *Given* $t \in (0,\infty)$, $\Xi^x_* W_0\big|_{\mathfrak{B}_t}$ *is absolutely continuous with respect to* $W_x\big|_{\mathfrak{B}_t}$, *with Radon-Nikodym derivative*

$$\frac{d\Xi^x_* W_0}{dW_x} = e^{\mathcal{Z}(t)}. \tag{10.61}$$

Note that by taking $f_\nu \nearrow 1$ in (10.56), we have $E_x\big(e^{\mathcal{Z}(t)}\big) = 1$, so the supermartingale $e^{\mathcal{Z}(t)}$ is actually a martingale in this case.

To prove the proposition, it suffices to show that, for $0 \le t_1 < \cdots < t_k \le t$, and a sufficiently large class of continuous functions f_j,

$$\begin{aligned} &E_0\Big(f_1\big(\mathfrak{X}^x(t_1)\big)\cdots f_k\big(\mathfrak{X}^x(t_k)\big)\Big) \\ &= E_x\Big(f_1\big(\omega(t_1)\big)\cdots f_k\big(\omega(t_k)\big)e^{\mathcal{Z}(t)}\Big). \end{aligned} \tag{10.62}$$

We will get this by extending (10.55) and (10.56) to formulas for the solution operators to time-dependent equations of the form

$$\frac{\partial u}{\partial t} = (\Delta + X)u - V(t,x)u, \quad u(0) = f. \tag{10.63}$$

Only the coefficient $V(t,x)$ depends on t; X does not. Parallel to (2.16), we can extend (10.55) to

$$u(t) = E_0\left(f\big(\mathfrak{X}^x(t)\big)e^{-\int_0^t V(s,\mathfrak{X}^x(s))\,ds}\right), \tag{10.64}$$

and we can extend (10.56) to

$$u(t) = E_x\left(e^{\mathcal{Z}(t)} f\big(\omega(t)\big)e^{-\int_0^t V(s,\omega(s))\,ds}\right). \tag{10.65}$$

Now we can pick $V(s,x)$ to be highly peaked, as a function of s, near $s = t_1, \dots, t_k$, in such a way as to get

$$e^{-\int_0^t V(s,\omega(s))\,ds} \approx e^{-V_1\left(\omega(t_1)\right)} \cdots e^{-V_k\left(\omega(t_k)\right)}. \tag{10.66}$$

Thus having the identity of (10.64) and (10.65) for a sufficiently large class of functions $V(s,x)$ can be seen to yield (10.62). We leave the final details to the reader.

For further material on the Cameron-Martin-Girsanov formula (10.61), see [Fr], [Kal], [McK], and [Øk].

Exercises

1. As an alternative derivation of (10.13), namely,

$$E_0\left(f\left(\mathfrak{X}^{x,s}(t)\right)\Big|\mathfrak{B}_s\right) = P^{t-s}f(x),$$

   via the Markov property, show that in light of the identity (10.5), it follows by applying (4.12) to $E_0\left(f\left(\mathfrak{X}^x(t-s,\vartheta_s\omega)\right)\Big|\mathfrak{B}_s\right)$.

2. Under the hypotheses of Proposition 10.1, show that, for $\lambda > 0$,

$$E_0\left(e^{\lambda|\mathfrak{X}^x(t)-x|}\right) \le 2e^{2\lambda^2 S^2 t + \lambda B t}.$$

   (*Hint*: If $\mathcal{Z}(t)$ denotes the last integral in (10.1), use (8.23) to estimate the quantity $E_0\left(e^{\lambda|\mathcal{Z}(t)|}\right)$.) Using this estimate in place of (10.10), get as strong a bound as you can on the behavior of $P^t f(x)$, for fixed $t \in \mathbb{R}^+$, as $|x| \to \infty$, given $f \in C_0(\mathbb{R}^n)$, that is, f continuous with compact support.

3. Granted the hypotheses under which the identity $\left(e^{t\mathcal{L}}\right)^* = e^{t(\widetilde{\mathcal{L}}-V)}$ on the space $C_o(\mathbb{R}^m) \cap L^1(\mathbb{R}^m)$ was established in (10.45), show that if $\widetilde{P}(t)$ denotes $\left(e^{t\mathcal{L}}\right)^*$ restricted to $L^1(\mathbb{R}^m)$, then $\mathcal{P}(t) = \widetilde{P}(t)^* : L^\infty(\mathbb{R}^m) \to L^\infty(\mathbb{R}^m)$ is

given by the same formula as (10.6):

$$\mathcal{P}(t)f(x) = E_0\ f\Big(\mathfrak{X}^x(t)\Big), \quad f \in L^\infty(\mathbb{R}^m).$$

Show that

$$P(s,x,t,U) = \mathcal{P}(t-s)\chi_U(x).$$

4. Assume $A(x)$ is real-valued, $A \in C^2(\mathbb{R}^m)$, and $A(x) \geq 0$ for all x. Show that

$$|\nabla A(x)|^2 \leq 4A(x)\ \sup\Big\{|D^2 A(y)| : |x-y| < \frac{2A(x)}{|\nabla A(x)|}\Big\}.$$

Use this to show that $\sqrt{A(x)}$ is uniformly Lipschitz on $\mathbb{R}^m$, establishing the scalar case of Proposition 10.2. (*Hint*: Reduce to the case $m = 1$; show that if $A'(c) > 0$, then A' must change by at least $A'(c)/2$ on an interval of length $\leq 2A(c)/A'(c)$, to prevent A from changing sign. Use the mean-value theorem to deduce $|A''(\zeta)| \geq |A'(c)|^2/4A(c)$ for some ζ in this interval.) For the general case of Proposition 10.2, see [Fdln], p. 189.

5. Suppose (10.1) is the system arising in Exercise 4 of §9, for $\mathfrak{X} = (x,v)$. Show that the generator L for P^t is given by

$$L = \frac{\partial^2}{\partial v^2} - \Big[\beta v + V'(x)\Big]\frac{\partial}{\partial v} + v\frac{\partial}{\partial x}. \tag{10.67}$$

6. Using methods produced in Chapter 8, §6, to derive Mehler's formula, compute the integral kernel for e^{tL} when L is given by (10.67), with $V(x) = ax^2$. *Remark*: This integral kernel is smooth for $t > 0$, reflecting the hypoellipticity of $\partial_t - L$. This is a special case of a general phenomenon analyzed in [Ho]. A discussion of this work can also be found in Chapter 15 of [T3].

A. The Trotter product formula

It is often of use to analyze the solution operator to an evolution equation of the form

$$\frac{\partial u}{\partial t} = Au + Bu$$

in terms of the solution operators e^{tA} and e^{tB}, which individually might have fairly simple behavior. The case where A is the Laplace operator and B is multiplication by a function is used in §2 to establish the Feynman-Kac formula, as a consequence of Proposition A.4 below.

The following result, known as the Trotter product formula, was established in [Tr].

Theorem A.1. *Let A and B generate contraction semigroups e^{tA} and e^{tB}, on a Banach space X. If $\overline{A+B}$ is the generator of a contraction semigroup $R(t)$, then*

$$R(t)f = \lim_{n\to\infty} \big(e^{(t/n)A}\, e^{(t/n)B}\big)^n f, \tag{A.1}$$

for all $f \in X$.

Here, $\overline{A+B}$ denotes the closure of $A+B$. A simplified proof in the case where $A+B$ itself is the generator of $R(t)$ is given in an appendix to [Nel2]. We will give that proof.

Proposition A.2. *Assume that* A, B, *and* $A+B$ *generate contraction semigroups* $P(t)$, $Q(t)$, *and* $R(t)$ *on* X, *respectively, where* $\mathcal{D}(A+B) = \mathcal{D}(A) \cap \mathcal{D}(B)$. *Then (A.1) holds for all* $f \in X$.

Proof. It suffices to prove (A.1) for $f \in \mathcal{D} = \mathcal{D}(A+B)$. In such a case, we have

$$P(h)Q(h)f - f = h(A+B)f + o(h), \tag{A.2}$$

since $P(h)Q(h)f - f = (P(h)f - f) + P(h)(Q(h)f - f)$. Also, $R(h)f - f = h(A+B) + o(h)$, so

$$P(h)Q(h)f - R(h)f = o(h) \text{ in } X, \quad \text{for } f \in \mathcal{D}.$$

Since $A+B$ is a closed operator, $\mathcal{D}$ is a Banach space in the norm $\|f\|_{\mathcal{D}} = \|(A+B)f\| + \|f\|$. For each $f \in \mathcal{D}$, $h^{-1}\big(P(h)Q(h) - R(h)\big)f$ is a bounded set in X. By the uniform boundedness principle, there is a constant C such that

$$\frac{1}{h}\big\|P(h)Q(h)f - R(h)f\big\| \le C\|f\|_{\mathcal{D}},$$

for all $h > 0$ and $f \in \mathcal{D}$. In other words, $\big\{h^{-1}\big(P(h)Q(h) - R(h)\big) : h > 0\big\}$ is bounded in $\mathcal{L}(\mathcal{D}, X)$, and the family tends strongly to 0 as $h \to 0$. Consequently,

$$\frac{1}{h}\big\|P(h)Q(h)f - R(h)f\big\| \longrightarrow 0$$

uniformly for f is a compact subset of $\mathcal{D}$.

Now, with $t \ge 0$ fixed, for any $f \in \mathcal{D}$, $\{R(s)f : 0 \le s \le t\}$ is a compact subset of $\mathcal{D}$, so

$$\big\|\big(P(h)Q(h) - R(h)\big)R(s)f\big\| = o(h), \tag{A.3}$$

uniformly for $0 \le s \le t$. Set $h = t/n$. We need to show that $\big(P(h)Q(h)\big)^n f - R(hn)f \to 0$, as $n \to \infty$. Indeed, adding and subtracting terms of the form $(P(h)Q(h))^j R(hn - hj)$, and using $\|P(h)Q(h)\| \le 1$, we have

$$\begin{aligned} &\big\|\big(P(h)Q(h)\big)^n f - R(hn)f\big\| \\ &\quad \le \big\|\big(P(h)Q(h) - R(h)\big)R(h(n-1))f\big\| \\ &\qquad + \big\|\big(P(h)Q(h) - R(h)\big)R(h(n-2))f\big\| \\ &\qquad\quad + \cdots + \big\|\big(P(h)Q(h) - R(h)\big)f\big\|. \end{aligned} \tag{A.4}$$

This is a sum of n terms that are uniformly $o(t/n)$, by (A.3), so the proof is done.

Note that the proof of Proposition A.2 used the contractivity of $P(t)$ and of $Q(t)$, but not that of $R(t)$. On the other hand, the contractivity of $R(t)$ follows from (A.1). Furthermore, the hypothesis that $P(t)$ and $Q(t)$ are contraction semigroups can be generalized to $\|P(t)\| \le e^{at}$, $\|Q(t)\| \le e^{bt}$. If $C = A + B$ generates a semigroup $R(t)$, we conclude that $\|R(t)\| \le e^{(a+b)t}$.

We also note that only certain properties of $S(h) = P(h)Q(h)$ play a role in the proof of Proposition A.2. We use

$$S(h)f - f = hCf + o(h), \quad f \in \mathcal{D} = \mathcal{D}(C), \tag{A.5}$$

where C is the generator of the semigroup $R(h)$, to get

$$S(h)f - R(h)f = o(h), \quad f \in \mathcal{D}. \tag{A.6}$$

As above, we have $h^{-1}\|S(h)f - R(h)f\| \le C\|f\|_{\mathcal{D}}$ in this case, and consequently $h^{-1}\|S(h)f - R(h)f\| \to 0$ uniformly for f in a compact subset of $\mathcal{D}$, such as $\{R(s)f : 0 \le s \le t\}$. Thus we have analogues of (A.3) and (A.4), with $P(h)Q(h)$ everywhere replaced by $S(h)$, proving the following.

Proposition A.3. *Let $S(t)$ be a strongly continuous, operator-valued function of $t \in [0, \infty)$, such that the strong derivative $S'(0)f = Cf$ exists, for $f \in \mathcal{D} = \mathcal{D}(C)$, where C generates a semigroup on a Banach space X. Assume $\|S(t)\| \le 1$ or, more generally, $\|S(t)\| \le e^{ct}$. Then, for all $f \in X$,*

$$e^{tC} f = \lim_{n\to\infty} S(n^{-1}t)^n \, f. \tag{A.7}$$

This result was established in [Chf], in the more general case where $S'(0)$ has *closure* C, generating a semigroup.

Proposition A.2 applies to the following important family of examples. Let $X = L^p(\mathbb{R}^n)$, $1 \le p < \infty$, or let $X = C_o(\mathbb{R}^n)$, the space of continuous functions vanishing at infinity. Let $A = \Delta$, the Laplace operator, and $B = -M_V$, that is, $Bf(x) = -V(x)f(x)$. If V is bounded and continuous on $\mathbb{R}^n$, then B is bounded on X, so $\Delta - V$, with domain $\mathcal{D}(\Delta)$, generates a semigroup, as shown in Proposition 9.12 of Appendix A. Thus Proposition A.2 applies, and we have the following:

Proposition A.4. *If $X = L^p(\mathbb{R}^n)$, $1 \le p < \infty$, or $X = C_o(\mathbb{R}^n)$, and if V is bounded and continuous on $\mathbb{R}^n$, then, for all $f \in X$,*

$$e^{t(\Delta - V)} f = \lim_{n\to\infty} \Big(e^{(t/n)\Delta} e^{-(t/n)V} \Big)^n f. \tag{A.8}$$

This is the result used in §2. If $X = L^p(\mathbb{R}^n)$, $p < \infty$, we can in fact take $V \in L^\infty(\mathbb{R}^n)$. See the exercises for other extensions of this proposition.

It will be useful to extend Proposition A.2 to solution operators for time-dependent evolution equations:

$$\frac{\partial u}{\partial t} = Au + B(t)u, \quad u(0) = f. \tag{A.9}$$

We will restrict attention to the special case that A generates a contraction semigroup and $B(t)$ is a continuous family of *bounded* operators on a Banach space X. The solution operator $S(t,s)$ to (A.9), satisfying $S(t,s)u(s) = u(t)$, can be constructed via the integral equation

$$u(t) = e^{tA}f + \int_0^t e^{(t-s)A}B(s)u(s)\,ds, \tag{A.10}$$

parallel to the proof of Proposition 9.12 in Appendix A on functional analysis. We have the following result.

Proposition A.5. *If A generates a contraction semigroup and $B(t)$ is a continuous family of bounded operators on X, then the solution operator to (A.9) satisfies*

(A.11)

$$S(t,0)f = \lim_{t\to\infty} \Bigl(e^{(t/n)A}\, e^{(t/n)B((n-1)t/n)}\Bigr) \cdots \Bigl(e^{(t/n)A}\, e^{(t/n)B(0)}\Bigr) f,$$

for each $f \in X$.

There are n factors in parentheses on the right side of (A.11), the jth from the right being $e^{(t/n)A}e^{(t/n)B((j-1)t/n)}$.

The proof has two parts. First, in close parallel to the derivation of (A.4), we have, for any $f \in \mathcal{D}(A)$, that the difference between the right side of (A.11) and

$$e^{(t/n)(A+B((n-1)t/n))} \cdots e^{(t/n)(A+B(0))} f \tag{A.12}$$

has norm $\le n \cdot o(1/n)$, tending to zero as $n \to \infty$, for t in any bounded interval $[0,T]$. Second, we must compare (A.12) with $S(t,0)f$. Now, for any fixed $t > 0$, define $v(s)$ on $0 \le s \le t$ by

$$\frac{\partial v}{\partial s} = Av + B\Bigl(\frac{j-1}{n}t\Bigr)v, \quad \frac{j-1}{n}t \le s < \frac{j}{n}t; \quad v(0) = f. \tag{A.13}$$

Thus (A.12) is equal to $v(t)$. Now we can write

$$\frac{\partial v}{\partial s} = Av + B(s)v + R(s)v, \quad v(0) = f, \tag{A.14}$$

where, for n large enough, $\|R(s)\| \le \varepsilon$, for $0 \le s \le t$. Thus

$$v(t) = S(t,0)f + \int_0^t S(t,s)R(s)v(s)\,ds, \tag{A.15}$$

and the last term in (A.15) is small. This establishes (A.11).

Thus we have the following extension of Proposition A.4. Denote by $BC(\mathbb{R}^n)$ the space of bounded, continuous functions on $\mathbb{R}^n$, with the sup norm.

Proposition A.6. *If $X = L^p(\mathbb{R}^n)$, $1 \le p < \infty$, or $X = C_o(\mathbb{R}^n)$, and if $V(t)$ belongs to $C\big([0,\infty), BC(\mathbb{R}^n)\big)$, then the solution operator $S(t,0)$ to*

$$\frac{\partial u}{\partial t} = \Delta u - V(t)u$$

satisfies

(A.16)

$$S(t,0)f = \lim_{n\to\infty} \Big(e^{(t/n)\Delta} e^{-(t/n)V((n-1)t/n)}\Big) \cdots \Big(e^{(t/n)\Delta} e^{-(t/n)V(0)}\Big) f,$$

for all $f \in X$.

To end this appendix, we give an alternative proof of the Trotter product formula when $Au = \Delta u$ and $Bu(x) = V(x)u(x)$, which, while valid for a more restricted class of functions $V(x)$ than the proof of Proposition A.4, has some desirable features. Here, we define $v_k = \Big(e^{(1/n)\Delta} e^{-(1/n)V}\Big)^k f$ and set

$$v(t) = e^{s\Delta} e^{-sV} v_k, \quad \text{for } t = \frac{k}{n} + s,\ 0 \le s \le \frac{1}{n}. \tag{A.17}$$

We use Duhamel's principle to compare $v(t)$ with $u(t) = e^{t(\Delta - V)} f$. Note that $v(t) \to v_{k+1}$ as $t \nearrow (k+1)/n$, and for $k/n < t < (k+1)/n$,

$$\begin{aligned} \frac{\partial v}{\partial t} &= \Delta v - e^{s\Delta} V e^{-sV} v_k \\ &= (\Delta - V)v + [V, e^{s\Delta}] e^{-sV} v_k. \end{aligned} \tag{A.18}$$

Thus, by Duhamel's principle,

$$v(t) = e^{t(\Delta - V)} f + \int_0^t e^{(t-s)(\Delta - V)} R(s)\, ds, \tag{A.19}$$

where

$$R(s) = [V, e^{\sigma\Delta}] e^{-\sigma V} v_k, \quad \text{for } s = \frac{k}{n} + \sigma,\ 0 \le \sigma < \frac{1}{n}. \tag{A.20}$$

We can write $[V, e^{\sigma\Delta}] = [V, e^{\sigma\Delta} - 1]$, and hence

$$R(s) = V(e^{\sigma\Delta} - 1) e^{-\sigma V} v_k - (e^{\sigma\Delta} - 1) V e^{-\sigma V} v_k. \tag{A.21}$$

Now, as long as

$$\mathcal{D}(\Delta - V) = \mathcal{D}(\Delta) = H^2(\mathbb{R}^n), \tag{A.22}$$

we have, for $0 \le \gamma \le 1$,

$$\big\| e^{t(\Delta - V)} \big\|_{\mathcal{L}(H^{-2\gamma}, L^2)} = \big\| e^{t(\Delta - V)} \big\|_{\mathcal{L}(L^2, H^{2\gamma})} \le C(T) t^{-\gamma}, \tag{A.23}$$

for $0 < t \le T$. Thus, if we take $\gamma \in (0,1)$ and $t \in (0,T]$, we have for

$$F(t) = \int_0^t e^{(t-s)(\Delta - V)} R(s)\, ds, \tag{A.24}$$

the estimate

$$\|F(t)\|_{L^2} \le C \int_0^t (t-s)^{-\gamma}\, \|R(s)\|_{H^{-2\gamma}}\, ds. \tag{A.25}$$

We can estimate $\|R(s)\|_{H^{-2\gamma}}$ using (A.21), together with the estimate

$$\big\|e^{\sigma\Delta} - 1\big\|_{\mathcal{L}(L^2, H^{-2\gamma})} \le C\, \sigma^{\gamma}, \quad 0 \le \gamma \le 1. \tag{A.26}$$

Since $\sigma \in [0, 1/n]$ in (A.21), we have

$$\begin{aligned} &\|R(s)\|_{H^{-2\gamma}} \le C n^{-\gamma} \varphi(V) \|f\|_{L^2}, \\ &\quad\varphi(V) = \Big(\|V\|_{\mathcal{L}(H^{2\gamma})} + \|V\|_{L^\infty}\Big) e^{s\|V\|_{L^\infty}}. \end{aligned} \tag{A.27}$$

Thus, estimating $v(t) = u(t)$ at $t = 1$, we have

$$\Big\| \Big(e^{(1/n)\Delta} e^{-(1/n)V}\Big)^n f - e^{(\Delta - V)} f \Big\|_{L^2} \le C_\gamma \varphi(V) \|f\|_{L^2} \cdot n^{-\gamma}, \tag{A.28}$$

for $0 < \gamma < 1$, provided multiplication by V is a bounded operator on $H^{2\gamma}(\mathbb{R}^n)$. Note that this holds if $D^\alpha V \in L^\infty(\mathbb{R}^n)$ for $|\alpha| \le 2$, and

$$\|V\|_{\mathcal{L}(H^{2\gamma})} \le C \sup_{|\alpha| \le 2} \|D^\alpha V\|_{L^\infty}. \tag{A.29}$$

One can similarly establish the estimate

$$\Big\| \Big(e^{(t/n)\Delta} e^{-(t/n)V}\Big)^n f - e^{t(\Delta - V)} f \Big\|_{L^2} \le C(t) \varphi(V) \|f\|_{L^2} \cdot n^{-\gamma}. \tag{A.30}$$

Exercises

1. Looking at Exercises 2–4 of §2, Chapter 8, extend Proposition A.4 to any V, continuous on $\mathbb{R}^n$, such that Re $V(x)$ is bounded from below and $|\text{Im } V(x)|$ is bounded.
   (*Hint*: First apply those exercises directly to the case where V is smooth, real-valued, and bounded from below.)
2. Let $H = L^2(\mathbb{R}),\ Af = df/dx,\ Bf = ixf(x)$, so $e^{tA} f(x) = f(x+t),\ e^{tB} f(x) = e^{itx} f(x)$. Show that Theorem A.1 applies to this case, but not Proposition A.2. Compute both sides of
   $$e^{pA+qB} f = \lim_{n\to\infty} \Big(e^{(p/n)A} e^{(q/n)B}\Big)^n f,$$
   and verify this identity directly.
   Compare with the discussion of the Heisenberg group, in §14 of Chapter 7.
3. Suppose A and B are *bounded* operators. Show that
   $$\Big\| e^{t(A+B)} - \Big(e^{(t/n)A} e^{(t/n)B}\Big)^n \Big\| \le \frac{Ct}{n}$$

and that

$$\left\| e^{t(A+B)} - \left(e^{(t/2n)A} e^{(t/n)B} e^{(t/2n)A} \right)^n \right\| \leq \frac{ct}{n^2}.$$

(*Hint*: Use the power series expansions for $e^{(t/n)A}$, and so forth.)

References

[Az] R. Azencott, Behavior of diffusion semi-groups at infinity, *Bull. Soc. Math. France* 102(1974), 193–240.

[B] D. Bell, *The Malliavin Calculus*, Longman, Essex, 1987.

[BG] R. Blumenthal and R. Getoor, *Markov Processes and Potential Theory*, Academic Press, New York, 1968.

[CM] R. Cameron and W. Martin, Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations, *Bull. AMS* 51(1945), 73–90.

[Chf] P. Chernoff, Note on product formulas for operator semigroups, *J. Func. Anal.* 2(1968), 238–242.

[CW] K. Chung and R. Williams, *Introduction to Stochastic Integration*, Birkhauser, Boston, 1990.

[Db1] J. Doob, The Brownian movements and stochastic equations, *Ann. Math.* 43(1942), 351–369.

[Db2] J. Doob, *Stochastic Processes*, J.Wiley, New York, 1953.

[Db3] J. Doob, *Classical Potential Theory and its Probabilistic Counterpart*, Springer-Verlag, New York, 1984.

[DS] N. Dunford and J. Schwartz, *Linear Operators*, Wiley, New York, 1958.

[Dur] R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth, Belmont, Calif., 1984.

[Ein] A. Einstein, *Investigations on the Theory of the Brownian Movement*, Dover, New York, 1956.

[El] K.Elworthy, *Stochastic Differential Equations on Manifolds*, LMS Lecture Notes #70, Cambridge Univ. Press, Cambridge, 1982.

[Em] M. Emery, *Stochastic Calculus in Manifolds*, Springer-Verlag, New York, 1989.

[Fdln] M. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton Univ. Press, Princeton, N. J., 1985.

[Fr] A. Friedman, *Stochastic Differential Equations and Applications*, Vols. 1 & 2, Academic Press, New York, 1975.

[HP] E. Hille and R. Phillips, *Functional Analysis and Semi-groups*, Colloq. Publ. AMS, Providence, R. I., 1957.

[Ho] L. Hörmander, Hypoelliptic second order differential equations, *Acta Math.* 119(1967), 147–171.

[IkW] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, Amsterdam, 1981.

[Ito] K. Ito, *On Stochastic Differential Equations*, Memoirs AMS #4, 1951.

[IMc] K. Ito and H. McKean, *Diffusion Processes and Their Sample Paths*, Springer-Verlag, New York, 1974.

[Kac] M. Kac, *Probability and Related Topics in Physical Sciences*, Wiley, New York, 1959.

[Kal] G. Kallianpur, Stochastic Filtering Theory, Springer-Verlag, New York, 1980.

[KS] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

[K] T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.

[Kol] A. Kolmogorov, Uber die analytishen Methoden in Wahrscheinlichkeitsrechnung, *Math. Ann.* 104(1931), 415–458.

[Lam] J. Lamperti, *Stochastic Processes*, Springer-Verlag, New York, 1977.

[Lev] P. Lévy, Random functions, *Univ. of Calif. Publ. in Statistics* I(12)(1953), 331–388.

[Mal] P. Malliavin, Stochastic calculus of variations and hypoelliptic operators, *Proc. Intern. Symp. SDE, Kyoto* (1976), 195–263.

[McK] H. McKean, *Stochastic Integrals*, Academic Press, New York, 1969.

[Mey] P. Meyer, *Probability and Potentials*, Blaisdell, Waltham, Mass., 1966.

[Nel] E. Nelson, *Operator Differential Equations*, Graduate Lecture Notes, Princeton Univ., Princeton, N. J., 1965.

[Nel2] E. Nelson, Feynman integrals and the Schrödinger equation, *J. Math. Phys.* 5(1964), 332–343.

[Nel3] E. Nelson, *Dynamical Theories of Brownian Motion*, Princeton Univ. Press, Princeton, N. J., 1967.

[Øk] B. Øksendal, *Stochastic Differential Equations*, Springer-Verlag, New York, 1989.

[Par] E. Pardoux, Stochastic partial differential equations, a review, *Bull. des Sciences Math.* 117(1993), 29–47.

[Pet] K. Petersen, *Brownian Motion, Hardy Spaces, and Bounded Mean Oscillation*, LMS Lecture Notes #28, Cambridge Univ. Press, Cambridge, 1977.

[PS] S. Port and C. Stone, *Brownian Motion and Classical Potential Theory*, Academic Press, New York, 1979.

[RaT] J. Rauch and M. Taylor, Potential and scattering theory on wildly perturbed domains, *J. Func. Anal.* 18(1975), 27–59.

[RS] M. Reed and B. Simon, *Methods of Mathematical Physics*, Academic Press, New York, Vols. 1,2, 1975; Vols. 3,4, 1978.

[Sch] Z. Schuss, *Theory and Applications of Stochastic Differential Equations*, Wiley, New York, 1980.

[Si] B. Simon, *Functional Integration and Quantum Physics*, Academic Press, New York, 1979.

[Stk] D. Stroock, The Kac approach to potential theory I, *J. Math. Mech.* 16(1967), 829–852.

[Stk2] D. Stroock, The Malliavin calculus, a functional analytic approach, *J. Func. Anal.* 44(1981), 212–257.

[StV] D. Stroock and S. Varadhan, *Multidimensional Diffusion Processes*, Springer-Verlag, New York, 1979.

[T] M. Taylor, Scattering length and perturbations of $-\Delta$ by positive potentials, *J. Math. Anal. Appl.* 53(1976), 291–312.

[T2] M. Taylor, Estimate on the fundamental frequency of a drum, *Duke Math.* J. 46(1979), 447–453.

[T3] M. Taylor, *Pseudodifferential Operators*, Princeton Univ. Press, Princeton, N. J., 1981.

[Tro] H. Trotter, On the product of semigroups of operators, *Proc. AMS* 10(1959), 545–551.

[Ts] M. Tsuji, *Potential Theory and Modern Function Theory*, Chelsea, New York, 1975.

[UO] G. Uhlenbeck and L. Ornstein, On the theory of Brownian motion, *Phys. Rev.* 36(1930), 823–841.

[Wal] J. Walsch, An introduction to stochastic partial differential equations, pp.265–439 in *Ecole d'été de Probabilité de Saint-Fleur XIV*, LNM #1180, Springer-Verlag, New York, 1986.

[Wie] N. Wiener, Differential space, *J. Math. Phys.* 2(1923), 131–174.

[Yo] K. Yosida, *Functional Analysis*, Springer-Verlag, New York, 1965.