

# Lecture 11: Some applications of the backward equation

## Readings

Recommended:

- Pavliotis [2014] 7.1-7.3 (first-passage times)
- Oksendal [2005] 7.3-7.4 (Dynkin formula), 8.2 (Feynman-Kac), 9.1 (Boundary-value problems)
- Gardiner [2009] 5.5 (first-passage times), 4.3.5 (Feynman-Kac)

Optional:

- Oksendal [2005] 9.2-9.3 (more on boundary-value problems)
- Korolov and Sinai [2010] 21.3, 21.4
- Pavliotis [2014] p.68 (Feynman-Kac)

In this lecture we'll look at some applications of the backward equation. First, we'll learn how to solve for first-passage time distributions. In particular, we'll show how to solve for the mean first passage time (MFPT) by solving an elliptic equation with Dirichlet boundary conditions. Then, we'll show how to solve more general boundary-value elliptic PDEs using stochastic processes. Finally, we'll learn about the Feynman-Kac formula, which generalises the backward equation to include a source term.

## 11.1 First-passage times

Consider a time-homogeneous process

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad (1)$$

which lives in a bounded domain  $\Omega \subset \mathbb{R}^d$  with boundary  $\partial\Omega$ . How long does it take to reach the boundary? This question has a number of applications, such as:

- an organism looking for food, where  $\partial\Omega$  is the boundary of the food source
- in a chemical reaction; where  $\partial\Omega$  is the boundary of another molecular conformation
- a ligand looking for a binding site at location  $\partial\Omega$
- in the stock market, e.g. the time it takes for an investment to double its value
- the time when average sea level has risen by 3 feet

and many others. In these applications we are interested in solving for the *first-passage time*, the time it takes a trajectory starting at  $x$  to hit  $\partial\Omega$  for the first time, defined mathematically as

$$T(x) = \inf\{t : X_t \in \partial\Omega | X_0 = x\}. \quad (2)$$

We can answer questions about first-passage times with a variant of the backward equation. In this section we outline how to do this, assuming all functions defined are nice enough that we can differentiate them as many times as we need and interchange derivatives and integrals; the assumptions required to make this true are touched upon in the next section.

We start by putting an *absorbing* boundary condition at  $\partial\Omega$ , so that trajectories are removed from the system as soon as they reach the boundary:  $p(y, t|x, s) = 0$  for  $x \in \partial\Omega$ . Let  $G(x, t)$  be the probability that a process is still in  $D$  at time  $t$ , given that it started at  $x$  at time 0. We can express  $G$  in terms of the transition density as

$$G(x, t) = P(T(x) \geq t) = \int_{\Omega} p(y, t|x, 0) dy. \quad (3)$$

To solve for  $G(x, t)$  we need an equation that relates its evolution in  $t$  to its evolution in  $x$ . The forward equation gives the evolution of  $p(y, t|x, s)$  in  $(y, t)$ , and the backward equation gives the evolution of  $p(y, t|x, s)$  in  $(x, s)$ . However, for a time-homogeneous process,  $p(y, t|x, 0) = p(y, 0|x, -t)$ , so from the backward equation, we get

$$\partial_t p(y, t|x, 0) = \mathcal{L}_x p(y, t|x, 0).$$

Integrate over  $y$  to get

$$\partial_t G = \mathcal{L}G \quad (\text{in } \Omega), \quad G(x, 0) = 1 \quad (x \in \Omega), \quad G(x, t) = 0 \quad (x \in \partial\Omega). \quad (4)$$

If we solve this parabolic equation for  $G$ , we can recover the full distribution of  $T$ , from (3).

To find moments of  $T$ , we can solve an elliptic equation instead. Consider the *mean first-passage time* (MFPT)  $T_1(x) = \mathbb{E}T(x)$ . Since the probability density of  $T(x)$  is  $-\partial_t G(x, t)$ , we have,

$$T_1(x) = \mathbb{E}T(x) = - \int_0^{\infty} t \partial_t G(x, t) dt = \int_0^{\infty} G(x, t) dt,$$

provided  $tG(x, t) \rightarrow 0$  as  $t \rightarrow \infty$ , which is a condition for  $\mathbb{E}T_1(x)$  to exist. Now apply  $\mathcal{L}$  to both sides of the equation to get

$$\mathcal{L}T_1 = \int_0^{\infty} \mathcal{L}G(x, t) dt = \int_0^{\infty} \partial_t G dt = -1.$$

We have just shown that  $T_1$ , if it exists, solves an elliptic equation

$$\mathcal{L}T_1(x) = -1 \quad (x \in \Omega), \quad T_1(x) = 0 \quad (x \in \partial\Omega). \quad (5)$$

**Example 11.1** (Brownian motion on a line) How long does it take a Brownian motion starting at the origin, to exit the interval  $[-R, R]$ ?

The generator of Brownian motion is  $\mathcal{L} = \frac{1}{2}\partial_{xx}$ , so we must solve

$$\frac{1}{2}\partial_{xx}T_1 = -1, \quad T_1(-R) = T_1(R) = 0.$$

You can easily find the solution is  $T_1(x) = -x^2 + R^2$ , so the MFPT starting at  $x = 0$  is  $T_1 = R^2$ . Notice that to find the MFPT from the origin, we have to solve for for all starting points  $x$ .

**Exercise 11.1.** Suppose in the example above that the Brownian motion is reflected at  $-R$ . What is the MFPT for it to reach  $R$ ?

**Example 11.2** (Arrhenius formula for reaction rate) [Gardiner, 2009]; [Pavliotis, 2014, Section 7.3]

One is often interested in calculating the transition rate between metastable states; for example, a molecule transitions from one conformation to another with a given rate, a protein denatures with some rate, the Ising model at low temperature has rare transitions between the +1 state and the -1 state. One model to calculate the transition rate between metastable states is to consider the overdamped Langevin equation for a process  $X$  moving on an energy surface with two deep minima. We consider here a one-dimensional equation,

$$dX_t = -U'(X_t)dt + \sqrt{2D}dW_t.$$

Here  $D$  is the diffusion coefficient and  $U(x)$  is the potential energy.

Consider a potential  $U(x)$  with two deep wells at  $x = a, c$ , and a saddle point in between at  $x = b$ , as shown in the figure. The stationary distribution is

$$\rho_s(x) = Z^{-1} e^{-U(x)/D}.$$

If  $D$  is small enough, then the stationary distribution is concentrated near each of the local minima, and there is a low probability of being found everywhere else. We expect the system to spend long times near  $a$  or  $c$ , and occasionally jump between them. What is the MFPT to go from  $a$  to  $c$ ?

The MFPT from some point  $x$  to some point  $z$  solves

$$-U' \partial_x T_1 + D \partial_{xx} T_1 = -1, \quad T_1'(-\infty) = 0, \quad T_1(z) = 0.$$

Solving the ODE above gives

$$T_1(x) = \frac{1}{D} \int_x^z dy \left[ e^{U(y)/D} \int_{-\infty}^y e^{-U(s)/D} ds \right]. \tag{6}$$

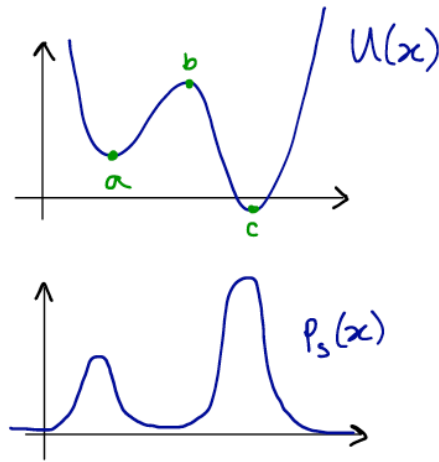
We would like to evaluate this expression with  $x = a$ , and we choose  $z$  to be some point near  $c$  but slightly before it, i.e. such that  $b < z < c$ . We don't put the absorbing point exactly at  $c$ , because when the process gets close to  $c$ , it spends a long time wandering around in the flat part of the potential energy landscape before it actually hits  $c$ , and we only care about the time to reach the basin of  $c$ , not the point  $c$  itself.

Expression (6) is exact, but it is complicated and doesn't tell us much about how  $T_1$  depends on the parameters in the problem. Let's try to find a simpler expression by approximating it when  $D$  is small.

The function  $e^{-U(s)/D}$  is only large near  $s = a$ ; for  $a < s < c$  it is close to zero so its contribution to the inner integral can be neglected. Therefore we can approximate the integral using Laplace asymptotics: Taylor-expand  $U(s)$  near  $s = a$ , and then send the limits of integration to  $\infty$ :

$$\text{For } y > a: \quad \int_{-\infty}^y e^{-U(s)/D} ds \approx \int_{-\infty}^{\infty} e^{-\frac{1}{D}(U(a) + \frac{1}{2}U''(a)(s-a)^2)} ds = \sqrt{\frac{2\pi D}{|U''(a)|}} e^{-U(a)/D}.$$

Now we can pull this factor out of the outer integrand in (6), since it is constant with respect to  $y$ . We can do a similar approximation for the outer integrand, since the first exponential factor is sharply peaked near



$y = b$ . Therefore we can approximate, for  $x < b$ ,

$$\int_x^z e^{U(y)/D} dy \approx \int_{-\infty}^{\infty} e^{\frac{1}{D}(U(b) - \frac{1}{2}U''(b)(y-b)^2)} dy = \sqrt{\frac{2\pi D}{|U''(b)|}} e^{U(b)/D}.$$

Putting this together gives

$$\mathbb{E}T_1(a \rightarrow c) \sim \frac{2\pi}{\sqrt{|U''(a)||U''(b)|}} e^{(U(b)-U(a))/D} \quad \text{as } D \rightarrow 0. \quad (7)$$

This is the *Arrhenius formula*; the rate of the reaction is approximated as  $1/T_1$ . The most important feature is the exponential dependence on the energy barrier,  $U(b) - U(a)$ , which shows that even a modest change in energy barrier, such as doubling it, will cause the MFPT to change by orders of magnitude.

## 11.2 Boundary-value problems

Consider the PDE for a function  $u : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$ :

$$\mathcal{L}u = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \quad (8)$$

with  $\mathcal{L} = b(x) \cdot \nabla + a(x) : \nabla^2$ , and  $\Omega \subset \mathbb{R}^d$  is open and bounded. This is similar to equation (5) for the MFPT, but with a more general source term and boundary condition. Can we similarly express its solution with a stochastic process of the form (1)?

We aim to express the solution to (8) using a stochastic process  $X$  that is stopped at the boundary  $\partial\Omega$ . To this end, let  $\tau_\Omega$  be the first passage time to  $\partial\Omega$ . Our main result is the following.

**Theorem** (Boundary-value problem). *Suppose  $f, a_{ij}, b_i$  are bounded and Lipschitz continuous in  $\Omega$ ,  $g$  is continuous on  $\partial\Omega$ , and  $\partial\Omega$  is  $C^2$ . Then the solution to (8) can be written as*

$$u(x) = \mathbb{E}^x \left( g(X_{\tau_\Omega}) - \int_0^{\tau_\Omega} f(X_t) dt \right), \quad (9)$$

where  $\tau_\Omega$  is the first passage time to  $\Omega$ ,  $X_{\tau_\Omega}$  is the first exit point, and where  $X$  solves (1).

See e.g. [Koralov and Sinai, 2010, Theorem 21.11], and [Friedman, 2004, Theorem 5.1.1].

*Remark.* The assumptions in this theorem ensure there is a unique strong solution  $u \in C^2(\Omega) \cap C(\bar{\Omega})$ . Given slightly stronger conditions we obtain a solution  $u \in C^2(\bar{\Omega})$ . We may weaken the conditions to obtain a solution that is less smooth. See Appendix, Section 11.4.1, for a summary of the relevant PDE theorems.

Before proving (9) we must first introduce the concept of integrals up to a random time  $\tau_\Omega$ . This is possible for random times which are *stopping times*. The definition of and some properties of stopping times are reviewed in the Appendix, Section 11.4.3; here, we only need to know that a first-passage time is a stopping time, and if  $\tau$  is a stopping time, then so is  $\tau \wedge T$  for any finite  $T > 0$ .

Given a stopping time  $\tau$ , the indicator function  $1_{\{t \leq \tau\}}$  is adapted, so we can define a stochastic integral with  $\tau$  in the limits of integration as

$$\int_0^\tau f(t, \omega) dW_t \equiv \int_0^\infty 1_{\{t \leq \tau\}} f(t, \omega) dW_t. \quad (10)$$

We would like the non-anticipating property to still hold; the following Lemma tells us when it does.

**Lemma** (Dynkin's formula). *If  $\tau$  is a stopping time with  $\mathbb{E}^x \tau < \infty$ , and  $f \in C^2$  with compact support, then*

$$\mathbb{E}^x f(X_\tau) = f(x) + \mathbb{E}^x \int_0^\tau \mathcal{L}f(X_s) ds. \quad (11)$$

*Proof.* Given a function  $f \in C^2$ , Itô's formula in integral form applied to  $f$  and then evaluated at  $\tau \wedge T$  (see (10)) where  $T > 0$  is a fixed time gives

$$f(X_{\tau \wedge T}) - f(X_0) = \int_0^{\tau \wedge T} \mathcal{L}f ds + \int_0^{\tau \wedge T} \nabla f a dW_s.$$

Taking the expectation and using the nonanticipating property of the Itô integral shows that

$$\mathbb{E}^x f(X_{\tau \wedge T}) - f(x) = \mathbb{E}^x \int_0^{\tau \wedge T} \mathcal{L}f ds.$$

Now, by assumption,  $f$  and its first two derivatives are bounded, so  $|\mathcal{L}f| \leq M$ . Therefore

$$\mathbb{E}^x \left| \int_0^T 1_{\{t \leq (\tau \wedge T)\}} \mathcal{L}f \right| \leq ME(\tau \wedge T) \leq M\mathbb{E}\tau < \infty,$$

so applying the Dominated Convergence Theorem to the identity above gives (11). □

*Remark.* The condition  $\mathbb{E}\tau < \infty$  is critical. See the homework for a case where this fails.

We would like to apply Dynkin's formula to the function  $u(X_t)$  in (8). To do so we must show  $\mathbb{E}^x \tau_\Omega$  is finite.

**Lemma.** *Given the assumptions in (8),*

$$\sup_{x \in \bar{\Omega}} \mathbb{E}^x \tau_\Omega < \infty. \quad (12)$$

*Proof.* [Friedman, 2004, part of Theorem 6.5.1] Let  $h(x) = -Ae^{\lambda x_1}$ . We can choose  $A, \lambda$  large enough so that

$$\mathcal{L}h = -a_{11}A\lambda^2 e^{\lambda x_1} - b_1 A \lambda e^{\lambda x_1} \leq -1 \quad \text{in } \Omega.$$

Therefore, by Itô's formula,

$$\mathbb{E}^x h(X_{\tau \wedge T}) - h(x) \leq -\mathbb{E}^x(\tau \wedge T).$$

Since  $|h(x)| \leq K$  in  $\Omega$ , using the triangle inequality we obtain  $\mathbb{E}^x(\tau \wedge T) \leq 2K$ . Take  $T \rightarrow \infty$  and use the monotone convergence theorem to obtain  $\mathbb{E}^x \tau \leq 2K$ . □

*Remark.* The proof above didn't require very many assumptions. In particular, it didn't require the uniform ellipticity of  $a$  nor any smoothness conditions on  $\partial D$  or the coefficients of  $\mathcal{L}$ . The proof only required that  $a_{11} > 0$  in  $\bar{\Omega}$ , and that the coefficients of  $\mathcal{L}$  be Lipschitz continuous in  $\Omega$  to ensure a unique process  $X$  exists.

*Remark.* The key part of the proof was to construct a function  $h$  satisfying  $\mathcal{L}h \leq -1$ . An alternative proof could choose  $h$  to be the solution to  $\mathcal{L}h = -1$ , which is known to exist under the assumptions given. However, a subtlety is that this equation is usually only guaranteed to hold in  $\Omega$ , and not (without further assumptions) at  $\partial\Omega$ . Therefore, one has to either extend the solution to a larger domain  $\Omega' \supset \overline{\Omega}$ , arguing that the PDE has sufficient regularity that one can do so, or else consider a sequence of domains  $\Omega_\varepsilon \subset \Omega$ , and consider the result in the limit  $\bigcup_\varepsilon \Omega_\varepsilon$ . See Korolov and Sinai [2010], Theorem 21.11 for the former approach, and Friedman [2004], Section 6.5 for an example of the latter.

Finally, we are ready to show our main result.

*Proof of Boundary-value theorem, (9).* [Korolov and Sinai, 2010, Theorem 21.11] We may use Dynkin's formula since  $\mathbb{E}^x \tau_\Omega < \infty$ . If  $u \in C^2(\overline{\Omega})$ , then it can be extended to a  $C^2$  function with compact support on  $\mathbb{R}^d$  (see Appendix 11.4.1, to which we may apply Dynkin's formula:

$$\mathbb{E}^x u(X_{\tau_\Omega}) - u(x) = \mathbb{E}^x \int_0^{\tau_\Omega} \mathcal{L}u(X_t) dt.$$

Solving for  $u(x)$  and noting that  $u(X_t) = f(X_t)$  for  $X_t \in \Omega$ , and  $u(X_{\tau_\Omega}) = g(X_{\tau_\Omega})$ , gives the desired result.

If  $u \notin C^2(\overline{\Omega})$ , then instead we construct a sequence of domains  $\Omega_1 \subset \Omega_2 \subset \dots \subset \Omega$  with smooth boundaries, such that  $\overline{\Omega}_n \subset \Omega$  and  $\bigcup_n \Omega_n = \Omega$ . Let  $\tau_n$  be the stopping times corresponding to domains  $\Omega_n$ . Then  $\lim_{n \rightarrow \infty} \tau_n = \tau_\Omega$  almost surely. We may apply Dynkin's formula to  $u$  in each  $\Omega_n$ , to obtain (9) with  $\tau_n$  instead of  $\tau_\Omega$ , and then use the Dominated Convergence Theorem to obtain (9). □

**Example 11.3** (MFPT) Let  $g(x) = 1$ ,  $f(x) = 0$ . Then

$$u(x) = \mathbb{E}^x \int_0^{\tau_\Omega} 1 dt = \mathbb{E}^x \tau_\Omega.$$

We recover (5), the equation we derived in the previous section for the MFPT.

**Example 11.4** (Mean-Value Theorem for harmonic functions) Suppose  $u : \mathbb{R}^d \rightarrow \mathbb{R}$  is harmonic:  $\Delta u = 0$ . Recall the Mean Value Theorem says that

$$u(x) = \frac{1}{\text{Area}(\partial B(x, r))} \int_{\partial B(x, r)} u(y) dS(y),$$

where  $B(x, r)$  is the ball of radius  $r$  centered at  $x$ ,  $\partial B(x, r)$  is its boundary, and  $dS(y)$  is the surface measure on  $\partial B(x, r)$ . Let's show this theorem using stochastic processes.

Let  $dX_t = dW_t$ , so the generator of  $X_t$  is  $\mathcal{L}f = \frac{1}{2} \Delta f$ . Suppose we are given a function  $u$  such that  $\Delta u = 0$  in some domain  $\Omega$ , and let  $B(x, r) \subset \Omega$ . By the boundary-value theorem,

$$u(x) = \mathbb{E}^x u(X_{\tau_B}),$$

where  $\tau_B$  is the first hitting time of  $\partial B(x, r)$  for  $X_t$ . Since Brownian motion is spatially isotropic, the distribution of first hitting points  $\{X_{\tau_B}\}$  is uniform on  $\partial B(x, r)$ , so

$$\mathbb{E}^x u(X_{\tau_B}) = \frac{1}{\text{Area}(\partial B(x, r))} \int_{\partial B(x, r)} u(y) dS(y).$$

*Remark.* It is possible to mix boundary conditions, for example to let  $u(x) = g(x)$  on  $\Gamma \subset \partial\Omega$ ,  $u_x(x) = 0$  on  $\partial\Omega \setminus \Gamma$ . This corresponds to stopping the process at  $\Gamma$ , and reflecting it at  $\partial\Omega \setminus \Gamma$ . The stopping time must be  $\tau_\Gamma$ , the first time the process hits  $\Gamma$ . See Evans [2013].

### 11.3 Feynman-Kac formula\*

Consider the following PDE for a function  $v(x, t) : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ :

$$\partial_t v = \mathcal{L}v - c(x)v, \quad v(x, 0) = f(x), \tag{13}$$

where  $\mathcal{L}v = b(x) \cdot \nabla v + a(x) : \nabla^2 v$ . We ask that  $c(x) \geq 0$ , which helps ensure the existence of a unique weak solution to (13) (Evans [2010]). This is similar to the backward equation, but with an extra term  $-c(x)v$ . We expressed the solution to the backward equation using an ensemble of trajectories of a diffusion process. Can we similarly express  $v$  using an ensemble of trajectories? Let's examine some special cases first.

- Without the term  $-c(x)v$ , (13) is simply the backward equation, whose solution can be represented as

$$v(x, t) = \mathbb{E}^x f(X_t), \tag{14}$$

where  $X_t$  solves the SDE (1).

- Without the diffusion term  $a : \nabla^2 v$ , (13) is a first-order PDE

$$\partial_t v = b(x) \cdot \nabla v - c(x)v, \quad v(x, 0) = f(x).$$

This equation can be solved by the method of characteristics. A characteristic  $x(s)$  starting at point  $x_0$  solves the ODE

$$\frac{dx}{ds} = b(x), \quad x(0) = x_0.$$

Let  $t > s$  and define  $z(s) = v(x(s), t - s)$ . This function evolves as

$$\frac{dz}{ds} = -\partial_t v + b(x) \cdot \nabla v = c(x)z, \quad z(0) = v(x, t).$$

We can solve to get  $z(t) = z(0)e^{\int_0^t c(x(s))ds}$ . Using that  $z(t) = f(x(t))$  gives

$$v(x, t) = z(0) = e^{-\int_0^t c(x(s))ds} f(x(t)). \tag{15}$$

From this expression for the solution along each characteristic  $x(s)$ , we can piece together the solution anywhere in space.

When all the terms are included, we get a combination of (14), (15), called the Feynman-Kac formula.

**Theorem** (Feynman-Kac, time-homogeneous). *Let  $f \in C_0^2(\mathbb{R}^d)$ ,  $c \in C(\mathbb{R}^d)$ ,  $c(x) \geq 0$ . The solution to the PDE*

$$\partial_t v = b(x) \cdot \nabla v + a(x) : \nabla^2 v - c(x)v, \quad v(x, 0) = f(x), \tag{16}$$

is given by

$$v(x, t) = \mathbb{E}^x \left[ e^{-\int_0^t c(X_s)ds} f(X_t) \right], \tag{17}$$

where  $X_t$  solves the SDE (1).

*Proof.* This proof is very similar to our proof of the backward equation from an earlier lecture. Let  $v(x, t)$  be the solution to the PDE (16), which we assume exists and is twice differentiable. Let

$$Z_1(s) = e^{-\int_0^s c(X_r) dr}, \quad Z_2(s) = v(X_s, t - s).$$

From the Itô product rule, we have  $d(Z_1 Z_2) = Z_1 dZ_2 + Z_2 dZ_1 + dZ_1 dZ_2$ . We calculate:

$$\begin{aligned} dZ_1(s) &= -Z_1 c ds \\ dZ_2(s) &= (-v_t + b \cdot \nabla v + a : \nabla^2 v) ds + \sigma \cdot \nabla v \cdot dW_s \\ &= (-v_t + \mathcal{L}v) + \sigma \cdot v \cdot dW_s. \\ dZ_1(s) dZ_2(s) &= 0. \end{aligned}$$

All functions above are evaluated at  $(X_s, t - s)$ , but we omit the function arguments for clarity. Therefore

$$d(Z_1(s)Z_2(s)) = Z_1 \underbrace{(-v_t + \mathcal{L}v - vc)}_{=0, \text{ since } v \text{ solves (16)}} ds + \sigma \cdot \nabla v dW_s = \sigma \cdot \nabla v dW_s.$$

Integrating in time and taking expectation gives

$$\mathbb{E}^x(Z_1(t)Z_2(t)) = \mathbb{E}^x Z_1(0)Z_2(0).$$

But  $Z_1(0)Z_2(0) = v(x, t)$ , and  $Z_1(t)Z_2(t) = e^{-\int_0^t c(X_r) dr} v(X_t, 0) = e^{-\int_0^t c(X_r) dr} f(X_t)$ . Therefore

$$\mathbb{E}^x[e^{-\int_0^t c(X_r) dr} f(X_t)] = \mathbb{E}^x(Z_1(t)Z_2(t)) = \mathbb{E}^x Z_1(0)Z_2(0) = v(x, t).$$

□

As we saw for the backward equation, trajectories of a diffusion process can be thought of as characteristics for the second-order, parabolic equation (17).

*Remark.* In many applications, particularly in finance, the Feynman-Kac formula is presented for a time-inhomogeneous process. In this case, one has to solve a PDE backwards in time. See Appendix, section 11.4.2 for the theorem statement and proof.

In the PDE (13), what does  $c(x)$  represent physically? We claim it represents the “killing rate,” i.e. the rate at which trajectories disappear, given they are at  $x$ . Physically, this could happen by chemicals being adsorbed to a surface with a rate depending on location, gamblers being kicked out of a casino with a rate that depends on how much money they have, etc. To why this interpretation is reasonable, suppose that we kill the process  $X_t$  at rate  $c(x)$ , so

$$P(X_t \text{ is killed in } [t, t+h]) = c(X_t)h + o(h).$$

Then, by dividing the time interval  $[0, t]$  into a fine partition  $\{t_0, t_1, \dots\}$  with  $t_i = ih$  for  $h$  small, we estimate

$$P(X_t \text{ survives until } t) \approx (1 - c(X_{t_1})h)(1 - c(X_{t_2})h) \cdots (1 - c(X_{t_n})h) \rightarrow e^{-\int_0^t c(X_s) ds} \text{ as } h \rightarrow 0.$$

Therefore,

$$v(x, t) = \mathbb{E}^x [f(X_t)P(X_t \text{ is alive at time } t)] = \mathbb{E}^x [f(X_t)1_{[0, \xi)}(t)],$$



where  $\xi(\omega)$  is the random time at which  $X_t(\omega)$  is killed. Therefore  $v(x, t)$  is the average of  $f(X_t)$ , weighted by the probability that the trajectory is still alive at time  $t$ . (See Oksendal [2005], Section 8.2 (p.145), or Evans [2013], Section 6.B.)

The Feynman-Kac formula was first developed in quantum mechanics to evaluate certain functionals of Brownian motion. It was constructed heuristically by Feynman and then rigorously by Kac, after he saw a lecture by Feynman and realized they were thinking about the same thing in different languages. Now, it is also heavily used in finance, where it has found use in pricing financial instruments, for example via the *Black-Scholes equation*.

**Example 11.5** The *Black-Scholes equation* comes from considering the proper price of an option to buy a stock. Let  $S(t)$  be the price of a stock at time  $t$ , which is assumed to behave as a geometric Brownian motion

$$dS = \mu S dt + \sigma S dW_t, \quad S(0) = s_0,$$

where  $\mu$  is the drift in price, and  $\sigma \neq 0$  is the volatility. A *European call option* is the right to buy one share of the stock at price  $p$  at time  $T$ . One is interested in determining the “proper price” of the option, i.e. the price where on average both the buyer and seller break even. The proper price at time  $t$  given  $S(t) = s$  is given by a function  $u(s, t)$ .

It is assumed that the interest rate is constant,  $r > 0$ , so that 1\$ in the bank at  $t = 0$  is worth  $e^{rT}$  at  $t = T$ , or conversely that 1\$ at  $t = T$  is worth  $e^{-rT}$  at  $t = 0$ . By asking that options be “self-financing” by a portfolio of bonds that grow as  $dB = rBdt$ , and by considering how to price options so as to create no arbitrage opportunities for others, one arrives at the Black-Scholes equation Evans [2013]

$$\begin{aligned} u_t + rsu_s + \frac{\sigma^2}{2} s^2 u_{ss} - ru &= 0, & 0 \leq t \leq T, \\ u(0, t) &= 0 & 0 \leq t \leq T, \\ u(s, T) &= (s - p)^+ & s > 0. \end{aligned}$$

This looks formally like the Feynman-Kac formula, although the derivation is quite different. Notably, the drift  $\mu$  does not appear in the equation.

## References

- Durrett, R. (1996). *Stochastic Calculus: A practical introduction*. CRC Press, Taylo & Francis Group.
- Durrett, R. (2005). *Probability: Theory and Examples*. Thomson, 3rd edition.
- Evans, L. C. (2010). *Partial differential equations*, volume 19. American Mathematical Soc.
- Evans, L. C. (2013). *An Introduction to Stochastic Differential Equations*. American Mathematical Society.
- Friedman, A. (2004). *Stochastic differential equations and applications*. Dover.
- Gardiner, C. (2009). *Stochastic methods: A handbook for the natural sciences*. Springer, 4th edition.
- Gilbarg, D., Trudinger, N. S., Gilbarg, D., and Trudinger, N. (1977). *Elliptic partial differential equations of second order*, volume 224. Springer.

Grimmett, G. and Stirzaker, D. (2001). *Probability and Random Processes*. Oxford University Press.

Koralov, L. B. and Sinai, Y. G. (2010). *Theory of Probability and Random Processes*. Springer.

Oksendal, B. (2005). *Stochastic Differential Equations*. Springer, 6 edition.

Pavliotis, G. A. (2014). *Stochastic Processes and Applications*. Springer.

## 11.4 Appendix

### 11.4.1 Collected results about PDEs

The following results concern an operator

$$Lu = a(x) : \nabla^2 u + b(x) \cdot \nabla u + c(x)f.$$

We are interested in the solution to the boundary-value problem

$$Lu = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega. \quad (18)$$

Let  $C^{2,1}(\Omega)$  be the space of functions on  $\Omega$  which are twice continuously differentiable whose derivatives up to 2nd order are Lipschitz continuous.

Recall that  $\Omega$  which satisfies the exterior sphere condition if for each  $\xi \in \partial\Omega$ , there exists a ball  $B = B_R(y)$  such that  $\bar{B} \cap \bar{\Omega} = \xi$ . If  $\partial\Omega$  is  $C^2$ , it satisfies this condition.

**Theorem** (Gilbarg et al. [1977], Theorem 6.13). *Given an open bounded domain  $\Omega$  which satisfies the exterior sphere condition. Suppose  $L$  is uniformly elliptic on  $\Omega$ , with  $c \leq 0$ . Further suppose  $f$  and the coefficients of  $L$  are bounded and Lipschitz continuous in  $\Omega$ , and  $g$  is continuous on  $\partial\Omega$ . Then there is a unique solution  $u \in C^{2,1}(\Omega) \cap C(\bar{\Omega})$  to (18).*

*Remark.* The conditions of this theorem can be weakened. Let  $C^{k,\alpha}(\Omega)$  denote the space of functions on  $\Omega$  that are  $k$  times continuously differentiable whose derivatives up to  $k$ th order are Hölder continuous with exponent  $\alpha \in (0, 1]$ . The theorem says that if  $f$  and the coefficients of  $L$  are bounded and in  $C^{0,\alpha}(\Omega)$  (with all other conditions the same), then  $u \in C^{2,\alpha}(\Omega) \cap C(\bar{\Omega})$ .

By strengthening the conditions on the coefficients, we obtain a solution that is twice continuously differentiable up to the boundary.

**Theorem** (Gilbarg et al. [1977], Theorem 6.14). *Given an open bounded domain  $\Omega$  which is a  $C^{2,1}$  domain. Suppose  $\mathcal{L}$  is uniformly elliptic on  $\Omega$  with  $c \leq 0$ . Further suppose  $f$  and the coefficients of  $L$  are Lipschitz continuous on  $\bar{\Omega}$ , and  $g \in C^{2,1}(\bar{\Omega})$ . Then there is a unique solution  $u \in C^{2,1}(\bar{\Omega})$  to (8).*

Given a solution up to the boundary, it can be extended a little bit beyond the boundary.

**Theorem** (Gilbarg et al. [1977], Lemma 6.37). *Let  $\Omega$  be a  $C^{2,1}$  domain and let  $\Omega'$  be an open set containing  $\bar{\Omega}$ . Suppose  $u \in C^{2,1}(\bar{\Omega})$ . Then there exists a function  $w \in C^{2,1}(\Omega')$  such that  $w = u$  in  $\Omega$ .*

### 11.4.2 Feynman-Kac, time-inhomogeneous

**Theorem** (Feynman-Kac, time-inhomogeneous). *The solution to the PDE*

$$u_t + \mathcal{L}u - c(x,t)u = 0, \quad u(x, T) = f(x), \quad (19)$$

where  $\mathcal{L}u = b(x,t) \cdot \nabla u + a(x,t) : \nabla^2 u$ , can be expressed as

$$u(x, t) = \mathbb{E}^{X_t=x} \left[ e^{-\int_t^T c(X_s, s) ds} f(X_T) \right].$$

Here  $X_t$  is a process that solves

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_t = x,$$

and  $\sigma(x, t)$  is a matrix such that  $\frac{1}{2}\sigma(x, t)\sigma^T(x, t) = a(x, t)$ .

*Proof.* This is essentially the same as for the time-homogeneous case. For simplicity, let's consider the one-dimensional case  $x \in \mathbb{R}$ ; the multi-dimensional case is very similar. Let  $u(x, t)$  be the solution to the above PDE. Let

$$Z_1(s) = e^{-\int_t^s c(X_r, r) dr}, \quad Z_2(s) = u(X_s, s).$$

From the Itô product rule, we have  $d(Z_1 Z_2) = Z_1 dZ_2 + Z_2 dZ_1 + dZ_1 dZ_2$ . We calculate:

$$\begin{aligned} dZ_1 &= -Z_1 c ds \\ dZ_2 &= (u_s + bu_x + au_{xx})ds + \sigma u_x dW_s \\ &= (u_s + \mathcal{L}u)ds + \sigma u_x dW_s. \\ dZ_1 dZ_2 &= 0. \end{aligned}$$

Inserting these into the product rule gives

$$\begin{aligned} d \left( e^{-\int_t^s c(X_r, r) dr} u(X_s, s) \right) &= Z_1 dZ_2 + Z_2 dZ_1 \\ &= Z_1 [(u_s + \mathcal{L}u)ds + \sigma u_x dW_s] - Z_1 u c ds \\ &= Z_1 \sigma u_x dW_s + Z_1 \underbrace{(u_s + \mathcal{L}u - cu)}_{=0, \text{ by construction}} ds \end{aligned}$$

The RHS has expected value 0, so the expectation of the product  $Z_1 Z_2$  is constant with time. Therefore

$$\mathbb{E}^{X_t=x} [Z_1(T)Z_2(T)] = \mathbb{E}^{X_t=x} [z_1(t)z_2(t)].$$

But  $\mathbb{E}^{X_t=x} [Z_1(T)Z_2(T)] = \mathbb{E}^{X_t=x} \left[ e^{-\int_t^T c(X_s, s) ds} f(X_T) \right]$ , and  $\mathbb{E}^{X_t=x} [z_1(t)z_2(t)] = \mathbb{E}^{X_t=x} u(X_t, t) = u(x, t)$ . Putting these together gives the result.  $\square$

### 11.4.3 Stopping times

Many of the applications we will consider involve random times. We need to know what kinds of random times we are allowed to work with, and how formulas such as Itô's formula, isometry, etc must be adapted when the time in the limits of integration is random. Some of this section is based on [Evans, 2013, Chapter 6], which contains some nice examples and simple proofs of the statements below.

**Definition.** A random variable  $\tau : \Omega \rightarrow [0, \infty]$  is called a *stopping time* with respect to a Brownian motion  $(W_t)_{t \geq 0}$  if the event  $\{\tau \leq t\}$  can be decided using only the values of  $(W_s)_{0 \leq s \leq t}$ .

*Remark.* The rigorous definition is that  $\tau$  is a stopping time with respect to the filtration  $\{\mathcal{F}_t\}$  if

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \geq 0,$$

that is, the set of all  $\omega \in \Omega$  such that  $\tau(\omega) \leq t$  is  $\mathcal{F}_t$ -measurable.

**Theorem** (Properties of stopping times). *Let  $\tau_1, \tau_2$  be two stopping times. Then*

- (i)  $\{\tau_i = t\} \in \mathcal{F}_t$ , for all times  $t \geq 0$ .
- (ii)  $\tau_1 \wedge \tau_2 = \min(\tau_1, \tau_2)$  and  $\tau_1 \vee \tau_2 = \max(\tau_1, \tau_2)$  are both stopping times.

*Proof.* (i)  $\{\tau_i = t\} = \{\tau_i \leq t\} - \{\tau_i < t\}$ , both of which are measurable in  $\mathcal{F}_t$  (**Exercise:** show that  $\{\tau_i < t\}$  is  $\mathcal{F}_t$ -measurable, by writing it as a countable union of measurable sets.)

- (ii)  $\{\tau_1 \wedge \tau_2 \leq t\} = \{\tau_1 \leq t\} \cup \{\tau_2 \leq t\} \in \mathcal{F}_t$ . Similar calculations hold for the max.

□

#### Examples 11.6

1.  $\tau = c$ , where  $c$  is a constant, is a stopping time.
2. Let  $X_t$  the the solution to an SDE, and let  $E$  be a nonempty open or closed subset of  $\mathbb{R}^d$ . Then

$$\tau = \inf\{t \geq 0 \mid X_t \in E\} \tag{20}$$

is a stopping time. (Set  $\tau = \infty$  for sample paths that never hit  $E$ .) For a proof, see [Evans, 2013, Section 6.A].

This is called the *first-passage time* to set  $E$ .

3. If  $\tau_E$  is a first-passage time, then  $\tau_E \wedge T$  is a stopping time. This is a technique used to construct stopping times with finite expected values.
4. The random variable

$$\sigma = \sup\{t \geq 0 \mid X_t \in E\},$$

representing the last time that  $X_t$  hits  $E$ , is *not* a stopping time. Heuristically, this is because  $\{\sigma \leq t\}$  depends on the whole future of the process, so cannot depend on only  $X_s$  for  $s \leq t$  / cannot be  $\mathcal{F}_t$ -measurable.