Partial Differential Equations and Diffusion Processes

James Nolen[1]
Department of Mathematics, Stanford University

[1]Email: nolen@math.stanford.edu.

Contents

Chapter 1

Introduction to PDE

An **ordinary differential equation (ODE)** is an equation involving an unknown function $y(t)$, its ordinary derivatives, and the (single) independent variable. For example: $y'(t) = -3y(t) + t$. A **partial differential equation (PDE)** is an equation involving an unknown function, its partial derivatives, and the (multiple) independent variables. PDE's are ubiquitous in science and engineering; the unknown function might represent such quantities as temperature, electrostatic potential, value of a financial security, concentration of a material, velocity of a fluid, displacement of an elastic material, population density of a biological species, acoustic pressure, etc. These quantities may depend on many variables, and one would like to understand *how* the unknown quantity depends on these variables. Typically a partial differential equation can be derived from physical laws (like Newton's laws of motion) and/or modeling assumptions (like the no-arbitrage principle) that specify the relationship between the unknown quantity and the variables on which it depends. So, often we are given a model in the form of a PDE which embodies physical laws and modeling assumptions, and we want to find a solution and study its properties.

1.1 Notation

For a given **domain** $\Omega \subset \mathbb{R}^d$, and a function $u : \Omega \to \mathbb{R}$ we will use $u_{x_1},\ u_{x_2},\ u_{x_2x_2}\ u_{x_1x_1}, u_{x_1x_2}\ \ldots$ to denote the partial derivatives of u with respect to independent variables $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$. We may also write

$$\frac{\partial u}{\partial x_1},\ \frac{\partial u}{\partial x_2},\ \frac{\partial^2 u}{\partial x_2^2},\ \frac{\partial^2 u}{\partial x_1^2},\ \frac{\partial^2 u}{\partial x_1 \partial x_2},\ \ldots, \tag{1.1}$$

respectively. When $d = 2$, we may use (x, y) instead of (x_1, x_2), and we may use

$$u_x,\ u_y,\ u_{yy}\ u_{xx}, u_{xy}, \ldots \quad \text{or} \quad \frac{\partial u}{\partial x},\ \frac{\partial u}{\partial y},\ \frac{\partial^2 u}{\partial y^2},\ \frac{\partial^2 u}{\partial x^2},\ \frac{\partial^2 u}{\partial x \partial y},\ \ldots \tag{1.2}$$

to denote the partial derivatives.

For much of the course, we will consider equations involving a variable representing time, which we denote by t. In this case, we will often distinguish the temporal domain from the domain for the other variables, and we generally will use $\Omega \subset \mathbb{R}^d$ to refer to the domain for the other variables only. For example, if $u = u(x, t)$, the domain for the function u is a subset of $\mathbb{R}^{d+1}$, perhaps $\Omega \times \mathbb{R}$ or $\Omega \times [0, \infty)$. In many physical applications, the other variables represent spatial coordinates. For

example, $u(x,t)$ might represent the temperature at position x at time t. In financial applications, however, the variable x might represent non-spatial quantities, like the price of a call option, and in this case we might use different notation. For example, $C(s,t)$ may denote the value of a European call option if the underlying stock price is s at time t. Even though the stock price s does not correspond to physical space, I will typically use the term "spatial variables" to refer to all of the independent variables except the variable representing time. In fact, one of the most fascinating points of this course is that the Black-Scholes equation for the price of a European call option can be transformed to the "heat equation" which models the dissipation of heat in a physical body, even though heat transfer and the fluctuation of stock prices are very different phenomena.

Normally we will use the notation Du or ∇u to refer to the gradient of u with respect to the "spatial variables" only. So, Du is the vector $Du = (u_{x_1}, u_{x_2}, \ldots, u_{x_d})$. We will use D^2u to refer to the collection of second partial derivatives of u with respect to x. The term Δu will always refer to the Laplacian with respect to the spatial coordinates. So, if $u = u(x,t)$ and $x \in \mathbb{R}^d$,

$$\Delta u = \sum_{j=1}^{d} u_{x_j x_j} = \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2} \tag{1.3}$$

Some people use $\nabla^2 u$ for the Laplacian instead of Δu. In some PDE literature it is common to use **multi-indices** to denote partial derivatives. A multi-index is a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ with integers $\alpha_i > 0$, and we define

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \tag{1.4}$$

where $|\alpha| = \alpha_1 + \cdots + \alpha_n$ is the order of the index. The notation $D^k u$ is used to refer to the collection of k^{th}-order partial derivatives of u.

1.2 Examples

The **order** of a PDE is the degree of the highest order derivatives appearing in the equation. For example, if there are two independent variables (x,y), a second-order PDE has the general form

$$F(u_{xx}, u_{yy}, u_{xy}, u_x, u_y, u, x, y) = 0.$$

while a first-order PDE has the general form

$$F(u_x, u_y, u, x, y) = 0.$$

In multi-index notation these could be written as $F(D^2u, Du, u, x, y) = 0$ or $F(Du, u, x, y) = 0$. In this course we will discuss only first and second order equations. Here are some examples of second-order equations:

- The heat equation
$$u_t = u_{xx} + u_{yy} \qquad (\text{or, } \; u_t = \Delta u) \tag{2.5}$$
- Black-Scholes' equation
$$u_t = -\frac{1}{2}\sigma^2(x,t)x^2 u_{xx} - rxu_x + ru \tag{2.6}$$

- Dupire's equation
$$u_t = \frac{1}{2}\sigma^2(x,t)x^2u_{xx} - rxu_x \tag{2.7}$$

- Laplace's equation
$$u_{xx} + u_{yy} = 0 \qquad (\text{or, } \Delta u = 0) \tag{2.8}$$

- Poisson's equation
$$u_{xx} + u_{yy} = f(x,y) \qquad (\text{or, } \Delta u = f) \tag{2.9}$$

- Reaction-diffusion equation
$$u_t = \Delta u + f(u) \tag{2.10}$$

- The wave equation
$$u_{tt} - c^2u_{xx} = 0 \qquad (\text{or, } u_{tt} = \Delta u) \tag{2.11}$$

Here are a few first order equations:

- Transport equation
$$u_t + v(x)\cdot Du = 0 \tag{2.12}$$

- Burgers' equation
$$u_t + uu_x = 0 \tag{2.13}$$

- A Hamilton-Jacobi equation
$$u_t + |Du|^2 = 0 \tag{2.14}$$

All of these examples involve a single function u. There are also many interesting **systems** of equations involving multiple unknown quantities simultaneously. For example, the famous Navier-Stokes equations are a second-order system of equations that model the velocity $u = \{u^{(i)}(x,t)\}_{i=1}^d$ of an incompressible fluid:

$$\begin{aligned} \partial_t u^{(i)} + u\cdot\nabla u^{(i)} &= \nu\Delta u^{(i)} - \nabla p, \quad i = 1,\dots,d \\ \sum_i u^{(i)}_{x_i} &= 0 \end{aligned} \tag{2.15}$$

1.3 Solutions and Boundary Conditions

We say that a function u **solves** a PDE if the relevant partial derivatives exist and if the equation holds at every point in the domain when you plug in u and its partial derivatives. For example, the function $u(x,y) = e^x\cos(y)$ solves the PDE $u_{xx} + u_{yy} = 0$ in the domain $\Omega = \mathbb{R}^2$. This definition of a solution is often called a **classical solution**, and we will use this definition unless stated otherwise. However, not every PDE has a solution in this sense, and it is sometimes useful to define a notion of **weak solution**.

The independent variables vary in a domain Ω, which is an open set that may or may not be bounded. A PDE will often be accompanied by **boundary conditions**, **initial conditions**, or **terminal conditions** that prescribe the behavior of the unknown function u at the boundary of

the domain under consideration. There are many boundary conditions, and the type of condition used in an application will depend on modeling assumptions.

Not all solutions to the PDE will satisfy the boundary conditions. For example, there are many smooth functions that satisfy the Laplace equation (2.8) for all (x, y) in a given smooth domain Ω. However, if we require that a solution u also satisfies the boundary condition $u(x, y) = g(x, y)$ for all points (x, y) in the boundary of the domain, with a prescribed function $g(x, y)$, then the solution will be unique. This is similar to the situation with ODEs: the equation $f'(x) = f(x)$ for $x \geq 0$ has many solutions ($f(x) = Ce^x$), but if we require that $f(0) = 27$, there is a unique solution $f(x) = 27e^x$.

The boundary of a domain will be denoted by $\partial\Omega$. One example of a **boundary value problem** (BVP) for the Laplace equation might be:

$$\begin{aligned} u_{xx} + u_{yy} &= 0, \quad \text{for all } (x, y) \in \Omega \quad \text{(PDE)} \\ u(x, y) &= g(x, y), \quad \text{for all } (x, y) \in \partial\Omega \quad \text{(boundary condition).} \end{aligned} \tag{3.16}$$

Prescribing the value of the solution $u = g$ on the boundary is called the **Dirichlet boundary condition**. If $g = 0$, we say that the boundary condition is **homogeneous**. There are many other types of boundary conditions, depending on the equation and on the application. Some boundary conditions involve derivatives of the solution. For example, instead of $u = g(x, y)$ on the boundary, we might impose $\nu \cdot \nabla u = g(x, y)$ for all $(x, y) \in \partial\Omega$. Here, the vector $\nu = \nu(x, y)$ is the exterior unit normal vector. This is called the **Neumann boundary condition**.

An example of an **initial value problem** (IVP) for the heat equation might be:

$$\begin{aligned} u_t &= u_{xx} + u_{yy}, \quad \text{for all } (x, y, t) \in \Omega \times (0, \infty) \quad \text{(PDE)} \\ u(x, y, t) &= g(x, y), \quad \text{for all } (x, y, t) \in \partial\Omega \times (0, \infty) \quad \text{(boundary condition)} \\ u(x, y, 0) &= h(x, y), \quad \text{for all } (x, y) \in \Omega \quad \text{(initial condition).} \end{aligned} \tag{3.17}$$

An initial condition is really a boundary condition on the $d + 1$ dimensional domain, the space-time domain. Since t will be interpreted as time, we use the term "initial condition" to refer to a boundary condition imposed on u at an initial time, often $t = 0$.

There are also applications for which it is interesting to consider terminal conditions imposed at a future time. For example, **terminal value problems** (TVP) arise in finance because a financial contract made at the present time may specify a certain "payoff" at a future time. The Black-Scholes model for the price of a European call option is the terminal value problem

$$\begin{aligned} u_t = -\frac{1}{2}\sigma^2 x^2 u - rxu_x + ru, \quad t < T \\ u(x, T) = \max(0, x - K) \end{aligned} \tag{3.18}$$

Here K is the strike price of the option which expires at time T (in the future). Another example of a terminal value problem is the following Hamilton-Jacobi equation:

$$u_t + |\nabla u|^2 = 0, \quad \text{for all } x \in \mathbb{R}^d,\ t < T \tag{3.19}$$

$$u(x, T) = g(x), \quad \text{for all } x \in \mathbb{R}^d. \tag{3.20}$$

Solving a BVP means finding a function that satisfies both the PDE and the boundary conditions. In many cases we cannot find an explicit representation for the solution, so "solving" the problem sometimes means showing that a solution exists or approximating it numerically.

1.4 Linear vs. Nonlinear

An equation is said to be **linear** if it has the form

$$\sum_{|\alpha|\leq k} a_\alpha(x) D^\alpha u = f(x) \tag{4.21}$$

(linear combination of u and its derivatives) = (function of the independent variables)

A **semilinear** equation has the form

$$\sum_{|\alpha|=k} a_\alpha(x) D^\alpha u + a_0(D^{k-1}u, \ldots, Du, u, x) = f(x) \tag{4.22}$$

A **quasilinear** equation has the form

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u, \ldots, Du, u, x) D^\alpha u + a_0(D^{k-1}u, \ldots, Du, u, x) = f(x) \tag{4.23}$$

An equation that depends in a nonlinear way on the highest order derivatives is called **fully nonlinear**. For example, the Laplace equation (2.8), the Poisson equation (2.9), the heat equation (2.5), Dupire's equation (2.7), Black-Scholes equation (0.1), the wave equation (2.11), and the transport equation (2.12) are all linear equations. The reaction-diffusion equation (2.10) is semilinear. Burgers' equation (2.13) is a quasilinear equation. The Hamilton-Jacobi equation (3.19) is fully nonlinear.

If an equation is linear and $f \equiv 0$ in the expression (4.21), the equation is called **homogeneous**. Otherwise it is called **inhomogeneous**. For example, the Laplace equation (2.8) is homogeneous, while the Poisson equation (2.9) is the inhomogeneous variety.

Generally linear equations are "easier" than nonlinear equations. One reason for this is that if u and v solve a linear, homogeneous equation, then so does any linear combination of u and v. For example, if u and v both solve the heat equation $u_t = u_{xx}$ and $v_t = v_{xx}$ then the function $w(x,t) = \alpha u(x,t) + \beta v(x,t)$ also solves the heat equation. This fact will be used many times in our analysis of linear equations. If the equation were nonlinear, it might not be so easy to find a nice PDE solved by w.

1.5 Some Important Questions

Here are some important questions to consider when studying a PDE.

- Is there a solution? This is not always easy to answer. Sometimes there may not be a classical solution, so one tries to find a **weak solution** by relaxing some conditions in the problem or weakening the notion of solution. For some problems, we may be able to construct a solution only in a region near part of the domain boundary. For time-dependent problems, sometimes there may be a classical solution only for a short period of time. The ODE analog is what happens to the solution of the problem $\dot{u} = u^2$, $u(0) = 1$: it has a unique solution $u(t) = 1/(1-t)$ that remains finite only until the time $t = 1$.

- If so, is the solution unique? Often a PDE will have infinitely many solutions, but only one of them satisfies the boundary condition under consideration. Nevertheless, verifying uniqueness of the solution may be difficult, especially when the equation is nonlinear.

- Does the solution depend continuously on the data (e.g. boundary conditions, initial conditions, terminal conditions)? That is, if we vary the data a little, does the solution behave in a stable manner or will it change completely? This is a very important property to have if you want to approximate a solution numerically, or if you want to quantify uncertainty in a simulation when there may be small fluctuations in the data.

- How can we represent the solution? Often we cannot find an explicit formula for the solution, even if the solution is unique.

- How regular is the solution? **Regularity** refers to the smoothness of the solution with respect to the independent variables. If the data in the problem have a certain regularity (for instance, is twice differentiable), what can we say about the solution? Will the solution be smoother than the data? Will the solution lose regularity in time? Both can happen: for instance, the heat equation has a regularizing effect: solution is "better" than the data, while Hamilton-Jacobi equations have solutions that may form singularities even if the data is smooth. Regularity also has practical implications. Roughly speaking, the ability to efficiently and accurately numerically approximate a solution to a PDE is directly related to the solution's regularity: the smoother the solution the easier it is to obtain it numerically.

- What are the qualitative/quantitative properties of the solution? How does the solution change as parameters in the equation change?

- How can we approximate the solution numerically? There is no universal numerical method that can be used for every PDE. Our theoretical understanding of the solutions helps in the development of efficient and convergent numerical solution methods.

- Given a solution (or measurements of a real system modeled by a particular PDE) can one reconstruct parameters in the equation? For instance, can we recover the heat conductivity if we measure the solution of the heat equation?This is an inverse problem.

Chapter 2

The Heat Equation, Part I

References:

- Evans, Section 2.3
- Strauss, 2.3-2.5, 3.3, 3.5

2.1 Physical Derivation and Interpretation

For $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$, the **heat equation** is

$$u_t = \Delta u \tag{1.1}$$

In the case of one dimension spatial dimension, $d = 1$, this is just $u_t = u_{xx}$. The heat equation models diffusion or heat transfer in a system out of equilibrium. The function $u(x, t)$ might represent temperature or the concentration of some substance, a quantity which may vary with time t.

Here is a derivation of the equation based on physical reasoning. Let $F(u)$ denote the **flux** of the quantity represented by u; the flux is a vector quantity representing the flow per unit surface area per unit time. From time t_1 to time t_2, the net change in "the amount of u" in a region $D \subset \mathbb{R}^d$ is determined by the net flux through the boundary ∂D. This is a conservation assumption – no material or heat energy is created or destroyed. This means that for any $t_1 < t_2$,

$$\int_D u(x, t_2)\, dx - \int_D u(x, t_1)\, dx = -\int_{t_1}^{t_2} \int_{\partial D} F(u) \cdot \nu \, dx \, dt.$$

The flux may be modeled as a linear function of ∇u: $F(u) = -a\nabla u$, where $a > 0$ is a constant. If u represents temperature, this assumption is known as Fourier's law of heat conduction; if u represents the concentration of a diffusing material, this is known as Fick's law. Therefore, the

function $u(x,t)$ should satisfy

$$\begin{aligned}\int_D u(x,t_2)\,dx - \int_D u(x,t_1)\,dx &= -\int_{t_1}^{t_2}\int_{\partial D} F(u)\cdot\nu\,dx\,dt \quad \text{(the conservation assumption)}\\ &= -\int_{t_1}^{t_2}\int_D \nabla\cdot F(u)\,dx\,dt \quad \text{(using the Divergence theorem)}\\ &= \int_{t_1}^{t_2}\int_D \nabla\cdot(a\nabla u)\,dx\,dt \quad \text{(assuming } F(u)=-a\nabla u)\\ &= \int_{t_1}^{t_2}\int_D a\Delta u\,dt\,dx \end{aligned} \tag{1.2}$$

Taking $D = B_r(x)$, a ball of radius r centered at a point x, dividing by $|D|$, and letting $r \to 0$, we conclude that u must satisfy

$$u(x,t_2) - u(x,t_1) = \int_{t_1}^{t_2} a\Delta u(x,t)\,dt \tag{1.3}$$

for all x. This is just the integral (in time) form of the heat equation (1.1). Here we have used the fact that if $w(x,t)$ is a continuous function, then

$$\lim_{r\to 0}\frac{1}{|B_r(x)|}\int_{B_r(x)} w(y,t)\,dy = w(x,t) \tag{1.4}$$

where $B_r(x)$ denotes the ball of radius r centered at the point x, and $|B_r(x)|$ denotes the volume of the ball.

If there is an external volumetric **source** (heat source, injection of material, etc.) or **sink** (cold bath, depletion of material, etc.) represented by a function $f(x,t)$, then we have

$$\int_D u(x,t_2)\,dx - \int_D u(x,t_1)\,dx = -\int_{t_1}^{t_2}\int_{\partial D} F(u)\cdot\nu\,dx\,dt + \int_{t_1}^{t_2}\int_D f(x,t)\,dx\,dt.$$

and the equation for u becomes inhomogeneous:

$$u_t = a\Delta u + f(x).$$

The case $f > 0$ represents an inflow of material or a heat source. The case $f < 0$ models outflow of material or a heat sink. In general f might not have a constant sign, representing the presence of both sources and sinks. In some models, the source or sink might depend on u itself: $f = f(u)$. For example, a simple model of an exothermic reaction might be $f = cu$, where the parameter $c > 0$ models a reaction rate.

In physical applications, the parameter $a > 0$ is sometimes called the **thermal conductivity** or the **diffusivity**. Notice that large values of a model rapid heat transfer or rapid diffusion; small values of a model slow heat transfer or slow diffusion. In some applications, the constant a is replaced by a matrix $a_{ij}(x)$ modeling a situation where the conductivity is variable, as in a composite material, for example. In this case, the derivation above produces the equation

$$u_t = \nabla\cdot(a(x)\nabla u) + f \tag{1.5}$$

The simple heat equation corresponds to $a_{ij}(x) \equiv Id$ (the identity matrix) and $f \equiv 0$. Assuming that $a(x)$ is differentiable, we could also write this as

$$u_t = \sum_{i,j} a_{ij}(x) u_{x_i x_j} + b(x) \cdot \nabla u + f \tag{1.6}$$

where $b(x) = (b_j(x))$, $b_j(x) = \sum_i \partial_{x_i} a_{ij}(x)$. This equation is said to be in **non-divergence form**, while the equation (1.5) is said to be in **divergence form**. As we will see later, the Black-Scholes equation has a similar form:

$$u_t = -\frac{1}{2}\sigma^2(x,t)x^2 u_{xx} - rxu_x + ru \tag{1.7}$$

Here the coefficient in front of the u_{xx} term might depend on both x and t, in general.

We may consider the heat equation for $x \in \mathbb{R}^d$, or in a bounded domain $\Omega \subset \mathbb{R}^d$ with appropriate boundary conditions. For example, we will consider the initial value problem with Dirichlet boundary condition

$$\begin{aligned} u_t &= \Delta u(x) + f(x,t), \quad x \in \Omega, \ \ t > 0 \\ u(x,t) &= h(x,t), \quad x \in \partial\Omega, \ \ t > 0 \\ u(x,0) &= g(x), \quad x \in \Omega, \ \ t = 0 \end{aligned} \tag{1.8}$$

where $\Omega \subset \mathbb{R}^d$ is some smooth, bounded domain. The **Dirichlet boundary condition** $u(x,t) = h(x,t)$ on $\partial\Omega$ may be interpreted as fixing the temperature at the boundary. Alternatively, the **Neumman boundary condition** $\nu \cdot \nabla u(x,t) = g(x,t)$ corresponds to prescribing the heat flux at the boundary (perhaps via an insulating layer, which means that $g = 0$).

2.2 The Heat Equation and a Random Walker

The preceding derivation of the heat equation was based on physical reasoning. Here is a rather different derivation and interpretation of the heat equation which illuminates its connection with probability and stochastic processes. Consider a simple random walk $X_x(n)$ on the integers $\mathbb{Z}$. We suppose that at each time step, the process moves independently either to the left, or to the right with probability 1/2, and that $X_x(0) = x$ initially. That is

$$X_x(n) = x + \sum_{j=1}^{n} s_j \tag{2.9}$$

where the steps s_j are independent, identically distributed random variables with $P(s_j = +1) = 1/2$ and $P(s_j = -1) = 1/2$ for each j.

The exit time

Let us assume that the starting point $x \in \mathbb{Z}$ lies between two integers $a, b \in \mathbb{Z}$: $a \leq x \leq b$. Consider the random time $s(x)$ that the walker spends before it hits either a or b if it started at x, and let

$\tau(x)$ be its expected value: $\tau(x) = E(s(x))$. As the walker moves initially either to the right or to the left,with equal probabilities, and spends a unit time to do so, we have the simple relation

$$\tau(x) = \frac{1}{2}\tau(x-1) + \frac{1}{2}\tau(x+1) + 1, \tag{2.10}$$

that may be re-written in the form of a discrete Poisson equation:

$$-\frac{\tau(x+1) + \tau(x-1) - 2\tau(x)}{2} = 1, \tag{2.11}$$

which is supplemented by the boundary conditions $\tau(a) = \tau(b) = 0$.

The expected prize value

Now suppose that $f(x)$ is a prescribed smooth function and that after n steps, we evaluate the expectation

$$u(x,n) = E\left[f(X_x(n))\right].$$

You may think of $u(x,n)$ as the expected "payoff" for a random walker starting at point x and walking for time n. The prize for landing at point $X_x(n) = y$ at time n is $f(y)$, and $u(x,n)$ is the expected prize value. How does u depend on the starting point x, time n, and prize distribution f?

At the initial time $n = 0$, the walker has not moved from its starting point, so we must have $u(x,0) = E\left[f(X_x(0))\right] = f(x)$. Now consider $n > 0$. Since $X_x(n)$ is a Markov process, we see that

$$E\left[f(X_x(n))\right] = E\left[E(f(X_x(n))|X_x(n-1))\right]$$

for each $n > 0$. The term $E(f(X_x(n))|X_x(n-1))$ is the conditional expectation of the payoff at time n, given the position at time $n-1$. We also know that starting from the point $X_x(n-1)$, the walker then moves to either $X_x(n-1) - 1$ (left) or to $X_x(n-1) + 1$ (right), each with probability $1/2$. Therefore, we can evaluate this conditional expectation explicitly:

$$E[f(X_x(n))|X_x(n-1)] = \frac{1}{2}f(X_x(n-1)-1) + \frac{1}{2}f(X_x(n-1)+1).$$

Taking the expected value on both sides, we conclude that

$$\begin{aligned} E\left[f(X_x(n))\right] &= E\left[E(f(X_x(n))|X_x(n-1))\right] \\ &= \frac{1}{2}E[f(X_x(n-1)-1)] + \frac{1}{2}E[f(X_x(n-1)+1)] \end{aligned}$$

Now subtract $E\left[f(X_x(n-1))\right]$ from both sides and we find that

$$\begin{aligned} u(x,n) - u(x,n-1) &= E\left[f(X_x(n))\right] - E\left[f(X_x(n-1))\right] \\ &= \frac{1}{2}E[f(X_x(n-1)-1)] + \frac{1}{2}E[f(X_x(n-1)+1)] - E\left[f(X_x(n-1))\right] \end{aligned}$$

Observe that $X_x(n-1) \pm 1 = X_{x\pm1}(n-1)$ almost surely. This observation allows us to express the change of u with respect to n in terms of changes in u with respect to x:

$$\begin{aligned} u(x,n) - u(x,n-1) &= \frac{1}{2}E[f(X_x(n-1)-1)] + \frac{1}{2}E[f(X_x(n-1)+1)] - E\left[f(X_x(n-1))\right] \\ &= \frac{1}{2}E[f(X_{x-1}(n-1))] + \frac{1}{2}E[f(X_{x+1}(n-1))] - E\left[f(X_x(n-1))\right] \\ &= \frac{1}{2}\left(u(x-1,n-1) - 2u(x,n-1) + u(x+1,n-1)\right) \end{aligned} \tag{2.12}$$

Observe! The relationship (2.12) looks very much like a discrete version of the heat equation! Let us explore this further. If we had let the spatial jumps be of size $h > 0$ instead of size 1 and let the time steps be of size $\delta > 0$ instead of size 1, the same derivation would lead us to the equation

$$u(x,t) - u(x,t-\delta) = \frac{1}{2}\left(u(x-h,t-\delta) - 2u(x,t-\delta) + u(x+h,t-\delta)\right)$$

Here we use $t \in \delta\mathbb{Z}$ to denote a point of the form $t = n\delta$ for some integer n. Now suppose we make the clever choice $\delta = h^2/2$. Then after dividing both sides by δ we have

$$\frac{u(x,t) - u(x,t-\delta)}{\delta} = \frac{\left(u(x-h,t-\delta) - 2u(x,t-\delta) + u(x+h,t-\delta)\right)}{h^2} \tag{2.13}$$

Taylor's theorem tells us that for differentiable v,

$$\frac{v(x,t) - v(x,t-\delta)}{\delta} = v_t(x,t) + O(\delta).$$

Also,

$$\begin{aligned} v(x+h) &= v(x) + hv_x(x) + \frac{h^2}{2}v_{xx}(x) + O(h^3) \\ v(x-h) &= v(x) - hv_x(x) + \frac{h^2}{2}v_{xx}(x) + O(h^3) \end{aligned}$$

so that

$$\frac{v(x+h) - 2v(x) + v(x-h)}{h^2} = v_{xx}(x) + O(h^2).$$

From this we see that (2.13) is a discrete version of the equation $u_t = u_{xx}$. This suggests that if we let the step size $h \to 0$ (with $\delta = h^2/2$), the function u (which depends on h) converges to some function $v(x,t)$ which satisfies $v_t = v_{xx}$.

How does the random walk behave under this scaling limit? If $\delta = 1/N$, $h = \sqrt{2}/\sqrt{N}$, and $t = k\delta$, then the scaled random walk may be written as

$$X_x(t) = x + \frac{\sqrt{2}}{\sqrt{N}}\sum_{j=1}^{Nt} s_j. \tag{2.14}$$

For continuous time $t > 0$, define the piecewise linear function $\tilde{X}_x(t)$ by interpolating the points $\{(n, X_x(n))\}_n$ like this:

$$\tilde{X}_x(t) = (n+1-t)X_x(n) + (t-n)X_x(n+1), \quad \text{if } t \in [n, n+1). \tag{2.15}$$

Since $E[s_j] = 0$ and $E[s_j^2] = 1$, the functional central limit theorem implies that the process $\tilde{X}_x(t)$ converges weakly, as $h \to 0$, to $\sqrt{2}W_x(t)$ where $W_x(t)$ is a Brownian motion with $W_x(0) = x$ (for example, see Karatsas and Shreve, section 2.4). Therefore, for any fixed t, $\tilde{X}_x(t)$ converges weakly to a Gaussian random variable with mean $\mu = x$ and variance $\sigma^2 = 2t$. Hence,

$$\lim_{h\to 0} E[f(X_x(t))] = \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi t}} e^{\frac{-|x-y|^2}{4t}} f(y)\, dy \tag{2.16}$$

In summary, we have observed that in the limit of small step sizes, the "value function" $u(x,t) = E[f(X_x(t))]$ representing the payoff at time t for a walk started at x satisfies the heat equation with initial data $u(x,0) = f(x)$. Alternatively, the formula $u(x,t) = E[f(X_x(t))]$ identifies the solution to the initial value problem for the heat equation $u_t = u_{xx}$ with a functional of Brownian motion. Later we will use Itô's formula to study this connection in much more generality. For now, notice how the following ideas appeared in this formal derivation:

- The Markov property of the random walk
- The fact that the walker steps to the left or right with equal probability (related to the Martingale property)
- The use of the space-time scaling ratio $\delta/h^2 = \text{constant}$

Exercises: 1. Perform similar analysis for dimension $d = 2$. That is, show that the expected payoff $E[f(X_{x,y}(n))]$ for the a random walk on the lattice $\mathbb{Z}^2$ satisfies a discrete version of the heat equation $u_t = u_{xx} + u_{yy}$.
2. Consider a random walker on integers with a bias: it jumps to the left with probability p and to the right with a probability $(1-p)$. What are the discrete equations for the exit time and the "prize distribution"? Find their continuous limits as well.
3. Consider a random walker on $\mathbb{Z}^2$ with asymmetry: assume that probability to go up or down is $1/4 - p$ while the probability to go left or right is $1/4 + p$. What is the corresponding discrete equation for the exit time from a square on the integer lattice, and what is the continuous limit?

2.3 The Fundamental Solution

Observe that the heat equation is a **linear equation**. Therefore, if u and v are both solutions to the heat equation, then so is any linear combination of u and v. This fact will be used frequently in our analysis.

Here we define a very special solution which allows us to construct solutions to initial value problems. The **fundamental solution** for the heat equation is the function

$$\Phi(x,t) = \frac{1}{(4\pi t)^{d/2}} e^{-\frac{|x|^2}{4t}}, \tag{3.17}$$

defined by $t > 0$. We have already seen this function in our derivation of the relation (2.16). Now we take this function as the starting point and show how it can be used to solve the heat equation. This function is also called the **heat kernel**, and it has the following properties:

(P0) For $t > 0$, $\Phi(x,t) > 0$ is an infinitely differentiable function of x and t.

(P1) $\Phi_t = \Delta\Phi$ for all $x \in \mathbb{R}^d$ and $t > 0$.

(P2) $\int_{\mathbb{R}^d} \Phi(x,t)\, dx = 1$ for all $t > 0$. Also, for each $t > 0$, $\Phi(x,t)$ is the probability density for a multivariate Gaussian random variable $x \in \mathbb{R}^d$ with mean $\mu = 0$ and covariance matrix $\Sigma_{ij} = 2t\delta_{ij}$ (in one dimension, $\sigma^2 = 2t$).

(P3) For any function $g(x)$ that is continuous and satsifies $|g(x)| \leq C_1 e^{C_2|x|}$, for some $C_1, C_2 > 0$,

$$\lim_{t\to 0} \int_{\mathbb{R}^d} \Phi(x,t) g(x)\, dx = g(0).$$

In particular, this holds for any continuous and bounded function.

Property P1 is easy if slightly tediously to verify directly by taking deriviatives of $\Phi(x,t)$. Property P2 says that the integral of Φ is invariant in t (remember, no heat created or destroyed). This is easy to verify by using a change of variables and the following basic fact:

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-x^2}\, dx &= \left(\left(\int_{-\infty}^{\infty} e^{-x^2}\, dx\right)^2\right)^{1/2} = \left(\int_{-\infty}^{\infty} e^{-x^2-y^2}\, dx\, dy\right)^{1/2} \\ &= \left(\int_0^{\infty}\int_0^{2\pi} r e^{-r^2}\, d\theta\, dr\right)^{1/2} = \left(2\pi \int_0^{\infty} r e^{-r^2}\, dr\right)^{1/2} = \sqrt{\pi}, \qquad (3.18)\end{aligned}$$

so that

$$\begin{aligned}\int_{\mathbb{R}^d} \Phi(x,t) dx &= \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{|x|^2}{4t}} dx = \frac{1}{\pi^{d/2}} \int_{\mathbb{R}^d} e^{-y^2} dy = \frac{1}{\pi^{d/2}} \int_{\mathbb{R}^d} e^{-y_1^2 - y_2^2 - \cdots - y_d^2} dy \\ &= \frac{1}{\pi^{d/2}} \left(\int_{\mathbb{R}} e^{-z^2} dz\right)^d = 1.\end{aligned}$$

Because the integral of $\Phi(x,t) > 0$ is 1 for all $t > 0$, the function $\Phi(x,t)$ defines a probability density for each $t > 0$ fixed. In fact, this is just the density for a multivariate Gaussian random variable with mean zero and covariance matrix $\Sigma_{ij} = 2t\delta_{ij}$ (in one dimension, $\sigma^2 = 2t$). So as $t \to \infty$, the variance grows linearly, and the standard deviation is proportional to $\sqrt{t}$.

Property P3 is a very interesting property which says that as $t \to 0$ the function $\Phi(x,t)$ concentrates at the origin. If Φ represents the density of a diffusing material at point x at time t, then P3 says that all of the mass concentrates at $x = 0$ as $t \to 0$. Mathematically this means that that Φ converges to a Dirac delta function (δ_0) in the sense of distributions as $t \to 0$. Since $\Phi(x,t) > 0$, you may think of the integral

$$\int_{\mathbb{R}^d} \Phi(x,t) g(x)\, dx \qquad (3.19)$$

as a weighted average of the function $g(x)$. In fact, this integral is an expectation with respect to the probability measure defined by Φ. As $t \to 0$, all of the weight concentrates near the origin where $g = g(0)$. In order to verify P3 we write:

$$\int_{\mathbb{R}^d} \Phi(x,t) g(x) dx = \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{|x|^2}{4t}} g(x) dx = \frac{1}{\pi^{d/2}} \int_{\mathbb{R}^d} e^{-y^2} g(y\sqrt{4t}) dy \to \frac{1}{\pi^{d/2}} \int_{\mathbb{R}^d} e^{-y^2} g(0) dy = g(0),$$

as $t \to 0$. In the last step we used the Lebesgue Dominated convergence theorem, since for all $t \in (0,1)$ we have a bound for the integrand by an integrable function independent of $t \in (0,1)$:

$$e^{-y^2} |g(y\sqrt{4t})| \leq C_1 e^{-y^2} e^{C_2|y|\sqrt{4t}} \leq C_1 e^{-y^2 + 2C_2|y|},$$

which is integrable.

Using these properties one may show the following:

Theorem 2.3.1 *For any function $g(x)$ that is continuous and satisfies $|g(x)| \leq C_1 e^{C_2|x|}$ for some $C_1, C_2 > 0$, the function*

$$u(x,t) = \int_{\mathbb{R}^d} \Phi(x-y,t)g(y)\,dy \tag{3.20}$$

satisfies

(i) $u \in C^\infty(\mathbb{R}^d \times (0,\infty))$ *(u is smooth in x and t for all positive times)*

(ii) $u_t = \Delta u$, *for all* $x \in \mathbb{R}^d$ *and* $t > 0$

(iii)

$$\lim_{(x,t)\to(x_0,0^+)} u(x,t) = g(x_0) \tag{3.21}$$

So, the function $u(x,t)$ defined by (3.20) solves the initial value problem in $\mathbb{R}^d$ with initial data $g(x)$. The values at $t = 0$ are defined by continuity, since the formula (3.20) is ill-defined for $t = 0$. Nevertheless, property (iii) says that the limit as $t \to 0^+$ is well defined and equal to g. Here is a very interesting point: even if $g(x)$ is merely continuous (not necessarily differentiable), we have a solution to the heat equation which is actually infinitely differentiable for all positive times! This is sometimes referred to as the smoothing property of the heat equation. Obviously, not all PDE's have this property. For instance, the simple transport equation

$$u_t + u_x = 0, \quad u(x,0) = g(x)$$

has the solution $u(x,t) = g(x-t)$ which is not at all smoother than the the initial data. The qualitative difference between the smoothing properties of the heat equation and the transport equation lies in the fact that the heat equation has a genuinely stochastic representation that produces the regularizing effect.

Consider the convolution formula:

$$u(x,t) = \int_{\mathbb{R}^d} \Phi(x-y,t)g(y)\,dy.$$

Since $\Phi(y,t)$ is the density for a probability measure, then so is $\Phi(x-y,t)$, although the mean is shifted to the point x. Therefore, this convolution formula is really an expectation

$$u(x,t) = E\left[g(X_x(t))\right],$$

where $\{X_x(t)\}_{t\geq 0}$ denotes a family of normal random variables with mean x and variance $2t$. This is precisely the conclusion (2.16) of our earlier derivation of the heat equation using the discrete random walk.

Note that if the growth condition $|g(x)| \leq C_1 e^{C_2|x|}$ were not satisfied, the integral $\int \Phi(x,t)g(x)\,dx$ might not even be finite. For example, if $g(x) = e^{x^2}$ the integral is not finite if t is large enough.

Proof of Theorem 2.3.1: First, property (iii) is a simple consequence of P3. Indeed, let $g_x(y) = g(x-y)$, then P3 implies that

$$\lim_{t\downarrow 0} \int \Phi(x-y,t)g(y)dy = \lim_{t\downarrow 0} \int \Phi(y,t)g(x-y)dy = \lim_{t\downarrow 0} \int \Phi(y,t)g_x(y)dy = g_x(0) = g(x-0) = g(x).$$

Properties (i) and (ii) follow from the fact that we may take derivatives of u by interchanging integration and differentiation. In general, one cannot do this. However, for $t > t_0 > 0$. the function $\Phi(x,t)$ is smooth with uniformly bounded and integrable derivatives of all orders (their seize is bounded by constants depending on t_0). Therefore, one can compute derivatives, as follows, invoking the dominated convergence theorem. The partial derivative u_t is defined by the limit

$$\frac{\partial u}{\partial t} = \lim_{h\to 0} \int_{\mathbb{R}^d} \frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} g(y)\, dy.$$

We know that as $h \to 0$,

$$\lim_{h\to 0} \frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} g(y) = \Phi_t(x-y,t)g(y), \tag{3.22}$$

so we'd like to say that

$$\lim_{h\to 0} \int_{\mathbb{R}^d} \frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} g(y)\, dy = \int_{\mathbb{R}^d} \Phi_t(x-y,t)g(y)\, dy \tag{3.23}$$

also holds. If h is sufficiently small (say $|h| < \epsilon$), then $t \pm h > 0$, and Taylor's theorem implies

$$\frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} = \Phi_t(x-y,t) + R(x,y,t,h), \tag{3.24}$$

where the remainder R satisfies the bound

$$R(x,y,t,h) \leq h \max_{|s|\leq \epsilon} |\Phi_{tt}(x-y,t+s)|. \tag{3.25}$$

Therefore, we see that for each x

$$\left| \frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} g(y) \right| \leq (\max_z |g(z)|) \left(|\Phi_t(x-y,t)| + \epsilon \max_{|s|\leq\epsilon} |\Phi_{tt}(x-y,t+s)| \right) \tag{3.26}$$

By computing Φ_t and Φ_{tt} directly, we see that the right hand side of (3.26) is integrable in y. Therefore, the dominated convergence theorem implies that

$$\frac{\partial u}{\partial t} = \lim_{h\to 0} \int_{\mathbb{R}^d} \frac{\Phi(x-y,t+h) - \Phi(x-y,t)}{h} g(y)\, dy = \int_{\mathbb{R}^d} \Phi_t(x-y,t)g(y)\, dy.$$

That is, using the dominated convergence theorem, we may justify bringing the limit inside the integral in (3.23). Using a similar argument with the dominated convergence theorem, one can show that

$$\Delta u = \int_{\mathbb{R}^d} \Delta\Phi(x-y,t)g(y)\, dy, \tag{3.27}$$

also holds, so that

$$u_t - \Delta u = \int_{\mathbb{R}^d} \left(\Phi_t(x-y,t) - \Delta\Phi(x-y,t)\right) g(y)\, dy = 0 \tag{3.28}$$

The last equality holds since Φ is itself a solution to $\Phi_t - \Delta\Phi = 0$.

In the same way, using the dominated convergence theorem, one may also take higher derivatives of $u(x,t)$, since Φ is infinitely differentiable, and each derivative is integrable (for $t > 0$). This shows that $u(x,t) \in C^\infty(\mathbb{R}^d \times (0,\infty))$, even if the initial data $g(x)$ is not smooth! In the case $d = 1$, Strauss works this out in section 3.5 (see Theorem 1, p. 79).

2.4 Duhamel's principle

So far we have derived a representation formula for a solution to the homogeneous heat equation in the whole space $x \in \mathbb{R}^d$ with given initial data. With the fundamental solution we may also solve the **inhomogeneous** heat equation using a principle called **Duhamel's principle**. Roughly speaking, the principle says that we may solve the inhomogeneous equation by regarding the source at time s as an initial condition at time s, an instantaneous injection of heat. The solution u is obtained by adding up (integrating) all of the infinitesimal contributions of this heating.

A steady problem converted to time-dependence

A simply example of how a steady problem can be converted to a time-dependent one is an elliptic problem of the form

$$-\nabla \cdot (a(x)\nabla u) = f(x), \quad \text{for } x \in \Omega, \tag{4.29}$$

$$u(x) = 0 \text{ for } x \in \partial\Omega. \tag{4.30}$$

This problem is posed in a smooth domain $\Omega \subset \mathbb{R}^d$, and $a(x)$ is the (possibly varying in space) diffusion coefficient. If $a(x) = 1$ then (4.29 is the standard Poisson equation

$$-\Delta u = f.$$

The Dirichlet boundary condition (4.30) mean that the boundary is cold (if we think of u as temperature). Here is how solution of (4.29)-(4.30) may be written in terms of a time-dependent problem without a source. Let $\phi(t, x)$ be the solution of the initial-boundary-value problem

$$\phi_t = \nabla \cdot (a(x)\nabla\phi), \quad \text{for } x \in \Omega, \tag{4.31}$$

$$\phi(x, t) = 0 \text{ for } x \in \partial\Omega \text{ and all } t > 0, \tag{4.32}$$

$$\phi(x, 0) = f(x) \quad \text{for } x \in \Omega. \tag{4.33}$$

The function $\phi(t, x)$ goes to zero as $t \to +\infty$, uniformly in $x \in \Omega$ – there are various ways to see that but we will take this for granted at the moment. Consider

$$\bar{u}(x) = \int_0^\infty \phi(x, t)dt, \tag{4.34}$$

this function satisfies the boudnary condition $\bar{u}(x) = 0$ for $x \in \Omega$, and, in addition, if we integrate (4.31) in time from $t = 0$ to $t = +\infty$ we get

$$-f(x) = \nabla \cdot (a(x)\nabla\bar{u}), \tag{4.35}$$

which is nothing but (4.31). Hence, $\bar{u}$ solves (4.31), so that solution of (4.31) can be represented as in (4.34) in terms of solutions of the initial-boundary-value problem which may be sometimes more convenient to solve than the elliptic problem directly. Furthermore, as in reality $\phi(x, t)$ goes to zero exponentially fast in time, a good approximation to $\bar{u}(x)$ may be obtained by integration not from 0 to $+\infty$ but rather over a short initial time period $[0, T]$.

The time-dependent Duhamel's principle

Suppose we wish to solve

$$\begin{aligned} u_t &= \Delta u + f(x,t), \quad x \in \mathbb{R}^d, \;\; t > 0 \\ u(x,0) &= g(x), \quad x \in \mathbb{R}^d. \end{aligned} \tag{4.36}$$

First, for $s \geq 0$, we define the family of functions $\tilde{w}(x,t;s)$ solving

$$\begin{aligned} \tilde{w}_t &= \Delta \tilde{w}, \quad x \in \mathbb{R}^d, \;\; t > s \\ \tilde{w}(x,s;s) &= f(x,s), \quad x \in \mathbb{R}^d, t = s. \end{aligned}$$

Notice that for each s, $\tilde{w}(\cdot,\cdot;s)$ solves an initial value problem with initial data prescribed at time $t = s$, instead of $t = 0$. Then set

$$w(x,t) = \int_0^t \tilde{w}(x,t;s)\, ds. \tag{4.37}$$

So, $\tilde{w}(x,t;s)$ represents the future influence (at time t) of heating at time $s \in (0,t)$, and $w(x,t)$ may be interpreted as the accumulation of all the effects from heating in the past. Duhamel's principle says that the solution $u(x,t)$ of the initial value problem (4.36) is given by

$$u(x,t) = u^h(x,t) + w(x,t) = u^h(x,t) + \int_0^t \tilde{w}(x,t;s)\, ds, \tag{4.38}$$

where $u^h(x,t)$ solves the homogeneous problem:

$$\begin{aligned} u^h_t &= \Delta u^h, \quad x \in \mathbb{R}^d, \;\; t > 0 \\ u^h(x,0) &= g(x), \quad x \in \mathbb{R}^d. \end{aligned} \tag{4.39}$$

In fact (we will prove below), the function $w(x,t)$ is the solution to the inhomogeneous problem with zero initial data:

$$\begin{aligned} w_t &= \Delta w + f(x,t), \quad x \in \mathbb{R}^d, \;\; t > 0 \\ w(x,0) &= 0, \quad x \in \mathbb{R}^d \end{aligned} \tag{4.40}$$

Since the PDE is linear, the combination of u^h and w solves (4.36).

Now, by Theorem 2.3.1 we may represent both of the functions $\tilde{w}$ and u^h in terms of the fundamental solution. Specifically,

$$\begin{aligned} \tilde{w}(x,t;s) &= \int_{\mathbb{R}^d} \Phi(x-y, t-s) f(y,s)\, dy \\ u^h(x,t) &= \int_{\mathbb{R}^d} \Phi(x-y,t) g(y)\, dy \end{aligned} \tag{4.41}$$

Combining this with the Duhamel formula (4.38), we see that

$$u(x,t) = \int_{\mathbb{R}^d} \Phi(x-y,t) g(y)\, dy + \int_0^t \int_{\mathbb{R}^d} \Phi(x-y,t-s) f(y,s)\, dy\, ds \tag{4.42}$$

Theorem 2.4.1 (see Evans Theorem 2, p. 50) *Suppose $f \in C_1^2(\mathbb{R}^d \times [0,\infty))$, then the function defined by (4.42) satisfies*

(i) $u_t = \Delta u + f(x,t)$ *for all* $t > 0$, $x \in \mathbb{R}^d$.

(ii) $u \in C_1^2(\mathbb{R}^d \times [0,\infty))$

(iii)

$$\lim_{(x,t)\to(x_0,0^+)} u(x,t) = g(x_0). \tag{4.43}$$

Proof: By the above analysis and Theorem 2.3.1, the only thing left to prove is that the function

$$w(x,t) = \int_0^t \int_{\mathbb{R}^d} \Phi(x-y,t-s) f(y,s)\, dy\, ds \tag{4.44}$$

solves the inhomogeneous problem (1.30). We compute derivatives:

$$\begin{aligned} \frac{w(x,t+h)-w(x,t)}{h} &= \frac{1}{h}\int_0^{t+h} \tilde{w}(x,t+h;s)\, ds - \frac{1}{h}\int_0^t \tilde{w}(x,t+h;s)\, ds \\ &= \frac{1}{h}\int_t^{t+h} \tilde{w}(x,t+h;s)\, ds \\ &\quad + \int_0^t \frac{\tilde{w}(x,t+h;s)-\tilde{w}(x,t;s)}{h}\, ds \end{aligned}$$

Using the properties of f and Φ and integrating by parts, one can show that

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} \tilde{w}(x,t+h;s)\, ds = \tilde{w}(x,s;s) = f(x,s) \tag{4.45}$$

and that

$$\begin{aligned} \lim_{h\to 0} \int_0^t \frac{\tilde{w}(x,t+h;s)-\tilde{w}(x,t;s)}{h}\, ds &= \int_0^t \tilde{w}_t(x,t;s)\, ds \\ &= \int_0^t \Delta \tilde{w}(x,t;s)\, ds = \Delta w(x,t) \end{aligned} \tag{4.46}$$

Therefore, $w_t = \Delta w + f(x,t)$. The initial condition is satisfied since

$$\lim_{t\to 0} \left| \int_0^t \int_{\mathbb{R}^d} \Phi(x-y,t-s) f(y,s)\, dy\, ds \right| \le \lim_{t\to 0} \int_0^t \max_y |f(y)|\, ds = 0. \tag{4.47}$$

So, $w(x,0) = 0$. See Evans Theorem 2, p. 50 for more details. $\square$

Relation to ODE's

You may have encountered Duhamel's principle already in the context of inhomogeneous ODEs. Let us point out the formal connection between the results above and what you may have seen already. Consider the following homogeneous ODE:

$$\eta'(t) = -A\eta(t) \tag{4.48}$$

$$\eta(0) = \eta_0. \tag{4.49}$$

Here $\eta : [0, \infty) \to \mathbb{R}$, and $A \in \mathbb{R}$ is some positive constant. The solution is the exponential

$$\eta(t) = e^{-tA}\eta_0$$

Now consider the inhomogeneous ODE:

$$\zeta'(t) = -A\zeta(t) + F(t) \tag{4.50}$$

$$\zeta(0) = \zeta_0$$

The solution is:

$$\zeta(t) = e^{-tA}\zeta_0 + \int_0^t e^{-(t-s)A}F(s)\,ds. \tag{4.51}$$

Exercise: Verify this.

You may think of $S(t) = e^{-tA}$ as a **solution operator** for the homogeneous equation (4.48). It maps the initial point η_0 to the value of the solution of (4.48) at time t: $S(t)\eta_0 \mapsto e^{-tA}\eta_0$. With this definition, the solution to the inhomogeneous equation (4.51) may be written as

$$\zeta(t) = S(t)\zeta_0 + \int_0^t S(t-s)F(s)\,ds. \tag{4.52}$$

Formally, the PDE (4.36) has the same structure as the ODE system (4.50). Letting $\zeta(t)$ denote the *function* $u(\cdot, t)$, we may write a formal equation

$$\zeta'(t) = -A\zeta + F(t) \tag{4.53}$$

where A is now an operator acting on the function $\zeta(t) = u(\cdot, t)$ according to $A : u(\cdot, t) \mapsto -\Delta u(\cdot, t)$, and $F(t)$ is the *function* $f(\cdot, t)$. This idea of defining an ODE for a function that takes values in a space of functions can be made mathematically rigorous using semigroup theory (for example, see Evans section 7.4, or the book *Functional Analysis* by K. Yosida).

Suppose that we know the solution operator $S(t)$ for the homogeneous equation $\eta' = -A\eta$, corresponding to the homogeneous heat equation $u_t = \Delta u$. Then the representation (4.52) suggests that the solution to the inhomogeneous equation should be

$$u(x,t) = S(t)g(x) + \int_0^t S(t-s)f(x,s)\,ds \tag{4.54}$$

We have already computed the solution operator $S(t)$ – it is given by convolution with the heat kernel:

$$S(t)g(x) = \int_{\mathbb{R}^d} \Phi(x-y,t)g(y)\,dy \tag{4.55}$$

and

$$S(t-s)f(x,s) = \int_{\mathbb{R}^d} \Phi(x-y,t-s)f(y,s)\,dy. \tag{4.56}$$

Combining this with (4.54) gives us the solution formula (4.42).

2.5 Boundary Value Problems

On the Half-Line

We now demonstrate a technique for solving boundary value problems for the heat equation on the half-line:

$$\begin{aligned} u_t = \Delta u, \quad x > 0, \;\; t > 0 \\ u(x,0) = g(x), \quad x > 0 \\ u(0,t) = 0, t > 0. \end{aligned} \tag{5.57}$$

The boundary condition imposed at $x = 0$ is the homogeneous Dirichlet condition. The convolution formula (3.20) gives a solution on the entire line $x \in \mathbb{R}$, but this function will not necessarily satisfy the boundary condition at $x = 0$. In fact, g is not yet defined for $x < 0$, so for this problem the convolution formula does not make sense immediately. So, we need to modify our approach to solving the problem.

The idea we demonstrate here is to construct a solution on the *whole line* in such a way that the condition $u(0,t) = 0$ is satisfied for all t. Then the restriction of this function to the right half-line will be a solution to our problem (5.57). To construct a solution on the whole line, we need to define the initial data for $x < 0$. The key observation is that if the initial data on the whole line has odd symmetry, then the heat equation preserves this symmetry. Moreover, any continuous function $f(x)$ that has odd symmetry (i.e. $f(-x) = -f(x)$) must satisfy $f(0) = 0$. Therefore, if $u(x,t)$ has odd symmetry for all t, then $u(0,t) = 0$ will be satisfied for all t.

We begin by extending the function $g(x)$ on $(-\infty, 0)$ by odd-reflection:

$$g^{ex}(x) = g(x), \;\; x \geq 0, \qquad g^{ex}(x) = -g(-x), \;\; x < 0 \tag{5.58}$$

This function has odd symmetry: $g^{ex}(-x) = -g^{ex}(x)$. Then we solve the extended problem

$$\begin{aligned} \bar{u}_t = \Delta \bar{u}, \quad x \in \mathbb{R}, \;\; t > 0 \\ \bar{u}(x,0) = g^{ex}(x), \quad x \in \mathbb{R} \end{aligned}$$

Using the convolution formula, our solution is

$$\bar{u}(x,t) = \int_{\mathbb{R}} \Phi(x-y,t) g^{ex}(y)\,dy.$$

Using a change of variables and the fact that Φ has even symmetry, it is easy to see that $\bar{u}$ has odd-symmetry: $\bar{u}(-x,t) = -\bar{u}(x,t)$ for all $x \in \mathbb{R}$. Therefore, $\bar{u}(0,t) = 0$ for all $t > 0$, and the

restriction of $\bar{u}(x,t)$ to the half-line satisfies (5.57). So, our solution is (for $x \geq 0$):

$$\begin{aligned} u(x,t) = \bar{u}(x,t) &= \int_{\mathbb{R}} \Phi(x-y,t) g^{ex}(y)\, dy \\ &= \frac{1}{(4\pi t)^{1/2}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4t}} g^{ex}(y)\, dy \\ &= \frac{1}{(4\pi t)^{1/2}} \int_0^\infty \left(e^{-\frac{|x-y|^2}{4t}} - e^{-\frac{|x+y|^2}{4t}} \right) g(y)\, dy \end{aligned} \tag{5.59}$$

Inhomogeneous boundary conditions, shifting the data

Suppose we modify the above problem to become

$$\begin{aligned} u_t = \Delta u, \quad x > 0, \;\; t > 0 \\ u(x,0) = g(x), \quad x > 0 \\ u(0,t) = h(t), t > 0 \end{aligned} \tag{5.60}$$

Now the boundary condition at the origin is $u(0,t) = h(t)$ which may be non-zero in general. Therefore, the reflection technique won't work without modification, since odd reflection guaranteed that $u = 0$ at the boundary. One way to solve boundary value problems with inhomogeneous boundary conditions is to "shift the data". That is, we subtract something from u that satisfies the boundary condition (but maybe not the PDE). In the present case, suppose we have a function $\hat{h}(x,t) : [0,\infty) \times [0,\infty) \to \mathbb{R}$ such that $\hat{h}(0,t) = h(t)$. This function $\hat{h}$ extends h off the axis $x = 0$. Then let $v(x,t) = u(x,t) - \hat{h}(x,t)$. This function v satisfies the homogeneous boundary condition: $v(0,t) = u(0,t) - \hat{h}(0) = h(t) - h(t) = 0$. However, v solves a different PDE. Since $u = v + \hat{h}$, we compute

$$\partial_t (v + \hat{h}) = \Delta (v + \hat{h}) \tag{5.61}$$

so that v satisfies

$$v_t = \Delta v + \Delta \hat{h} - \hat{h}_t \tag{5.62}$$

Putting this all together, we see that $u = v + \hat{h}$ where v solves

$$\begin{aligned} v_t = \Delta v + f(x,t), \quad x > 0, \;\; t > 0 \\ v(x,0) = g(x) - \hat{h}(x,0), \quad x > 0 \\ v(0,t) = 0, t > 0 \end{aligned} \tag{5.63}$$

and $f(x,t) = \Delta \hat{h} - \hat{h}_t$. The price to pay for shifting the data is that now we may have an inhomogeneous equation and different initial conditions. The key fact that makes this solution technique possible is the fact that the equation is linear; thus we can easily derive and solve an equation for the shifted function v.

Another example

Here we illustrate this shifting technique and the reflection technique together. Let us solve the inhomogeneous equation with inhomogeneous Dirichlet boundary condition:

$$\begin{aligned} u_t = u_{xx} + k(x,t), \quad t > 0, \;\; x > 0 \\ u(x,0) = g(x), \quad x > 0 \\ u(0,t) = 1, \quad t > 0 \end{aligned}$$

We also have added a source term $k(x,t)$ just for illustration. Here we suppose that $g(x) \geq 0$ is smooth, bounded, and $g(0) = 1$. Let's find an integral representation formula for $u(x,t)$ involving g, k, and the fundamental solution $\Phi(x,t)$.

The boundary condition is inhomogeneous, so we first shift the function u to transform the boundary condition into the homogeneous condition. One way to do this would be setting $v(x,t) = u(x,t) - 1$. Then $v(0,t) = u(0,t) - 1 = 0$, so v satisfies the homogenous boundary condition. There are other ways to do this, as well. The function v satisfies the modified problem:

$$\begin{aligned} v_t = v_{xx} + k(x,t), \quad t > 0, \;\; x > 0 \\ v(x,0) = g(x) - 1 := \tilde{g}(x), \quad x > 0 \\ v(0,t) = 0, \quad t > 0 \end{aligned}$$

We now solve for v and set $u = v + 1$.

To solve for v, we extend the problem onto the entire line and solve using the fundamental solution and Duhamel's principle. To obtain the boundary condition, we extend $\tilde{g}$ and k by odd reflection:

$$\begin{aligned} \tilde{g}^{ex}(x) = \tilde{g}(x) = g(x) - 1, \;\; x \geq 0, \qquad & \tilde{g}^{ex}(x) = -\tilde{g}(-x) = 1 - g(-x), \;\; x < 0 \\ k^{ex}(x) = k(x), \;\; x \geq 0, \qquad & k^{ex}(x) = -k(-x), \;\; x < 0 \end{aligned}$$

Therefore, using the Duhamel formula (4.42), we construct a solution

$$v(x,t) = \int_{\mathbb{R}} \Phi(x-y,t)\tilde{g}^{ex}(y)\,dy + \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s)k^{ex}(y,s)\,dy\,ds \tag{5.64}$$

so that

$$\begin{aligned} u(x,t) &= 1 + v(x,t) \\ &= 1 + \int_{\mathbb{R}} \Phi(x-y,t)\tilde{g}^{ex}(y)\,dy + \int_0^t \int_{\mathbb{R}} \Phi(x-y,t-s)k^{ex}(y,s)\,dy\,ds \qquad (5.65) \\ &= 1 + \frac{1}{(4\pi t)^{1/2}} \int_0^\infty \left(e^{-\frac{|x-y|^2}{4t}} - e^{-\frac{|x+y|^2}{4t}} \right) (g(y) - 1)\,dy \\ &\quad + \int_0^t \int_0^\infty \frac{1}{(4\pi s)^{1/2}} \left(e^{-\frac{|x-y|^2}{4(t-s)}} - e^{-\frac{|x+y|^2}{4(t-s)}} \right) k(y,s)\,dy\,ds \qquad (5.66) \end{aligned}$$

2.6 Uniqueness of solutions: the energy method

Using the fundamental solution we have constructed one solution to the problem

$$\begin{aligned} u_t = \Delta u + f(x,t), \quad x \in \mathbb{R}^d, \;\; t > 0 \\ u(x,0) = g(x), \quad x \in \mathbb{R}^d \end{aligned} \tag{6.67}$$

where $f \in C_1^2(\mathbb{R}^d \times [0,\infty))$ and $|g(x)| \leq C_1 e^{C_2|x|}$. Is this the only solution? If there were another solution v, then their difference $w = u - v$ would satisfy

$$\begin{aligned} w_t = \Delta w, \quad x \in \mathbb{R}^d, \;\; t > 0 \\ w(x,0) = 0, \quad x \in \mathbb{R}^d \end{aligned} \tag{6.68}$$

since the equation is linear. We'd like to say that $w \equiv 0$ for all $t > 0$ since the initial data is zero. This would imply that $u = v$ so that the solution is unique. However, it turns out (surprise!) that there are non-trivial solutions to this initial value problem (6.68). So the solution to (6.67) is not unique. Nevertheless, the non-trivial solutions to (6.68) must grow very rapidly as $|x| \to \infty$, and if we restrict our attention to solutions satisfying a certain growth condition, then the only solution of (6.68) is the trivial solution $w \equiv 0$. Therefore, under a certain growth restriction, the solution to (6.67) must be unique:

Theorem 2.6.1 (See Evans Theorem 7, p. 58) *There exists at most one classical solution to the initial value problem (6.67) satisfying the growth estimate*

$$|u(x,t)| \leq A e^{a|x|^2}, \quad \forall\, x \in \mathbb{R}^d, t \in [0,T] \tag{6.69}$$

for constants $A, a > 0$.

From now on, we will always assume that our solutions to the heat equation in the whole space satisfy this growth condition. Notice that the condition $|g(x)| \leq C_1 e^{C_2|x|}$ is within the limits of the this growth condition.

For boundary value problems in a bounded domain, this technical issue does not arise, and solutions may be unique. For example, consider the initial value problem with Dirichlet boundary conditions:

$$\begin{aligned} u_t &= \Delta u + f(x,t), \quad x \in \Omega, \;\; t > 0 \\ u(x,t) &= h(x,t), \quad x \in \partial\Omega, \;\; t > 0 \\ u(x,0) &= g(x), \quad x \in \Omega, \;\; t = 0 \end{aligned} \tag{6.70}$$

Theorem 2.6.2 *There is at most one solution to the initial value problem (6.70).*

Proof: If there were two classical solutions to this problem, then their difference $w = u - v$ would satisfy (since the equation is linear!):

$$\begin{aligned} w_t &= \Delta w, \quad x \in \Omega, \;\; t > 0 \\ w(x,t) &= 0, \quad x \in \partial\Omega, \;\; t > 0 \\ w(x,0) &= 0, \quad x \in \Omega, \;\; t = 0 \end{aligned}$$

We wish to show that $w(x,t) = 0$ for all $t \geq 0$ and $x \in \Omega$, implying that $u = v$. To see this, multiply the equation by w and integrate in x and t:

$$\int_0^T \int_\Omega w_t(x,t)w(x,t)\,dx\,dt = \int_0^T \int_\Omega (\Delta w(x,t))w(x,t)\,dx\,dt \tag{6.71}$$

We will use this equality to show that the quantity $E(T) := \int_\Omega w^2(x,T)\,dx$ must be zero for all T. The left hand side is:

$$\begin{aligned}
\int_0^T \int_\Omega w_t(x,t)w(x,t)\,dx\,dt &= \int_0^T \int_\Omega \frac{1}{2}\frac{\partial}{\partial t} w^2\,dx\,dt \\
&= \frac{1}{2}\int_\Omega \left(\int_0^T \frac{\partial}{\partial t} w^2\,dt \right) dx \\
&= \frac{1}{2}\int_\Omega (w^2(x,T) - w^2(x,0))\,dx \quad \text{(FTC)} \\
&= \frac{1}{2}\int_\Omega w^2(x,T)\,dx = \frac{1}{2}E(T)
\end{aligned}$$

We may evaluate the right hand side of (6.71) using the fact that $w(\Delta w) = w\nabla\cdot\nabla w = \nabla\cdot(w\nabla w) - \nabla w\cdot\nabla w$, so that

$$\int_0^T \int_\Omega (\Delta w(x,t))w(x,t)\,dx\,dt = \int_0^T \int_\Omega \nabla\cdot(w\nabla w) - |\nabla w|^2\,dx\,dt.$$

The first integral on the right side vanishes, by the divergence theorem and the fact that $w = 0$ on the boundary:

$$\int_0^T \int_\Omega \nabla\cdot(w\nabla w)\,dx\,dt = \int_0^T \int_{\partial\Omega} \nu\cdot(w\nabla w)\,dS(y)\,dt = 0.$$

Therefore,

$$\int_0^T \int_\Omega (\Delta w(x,t))w(x,t)\,dx\,dt = -\int_0^T \int_\Omega |\nabla w|^2\,dx\,dt \leq 0.$$

Now returning to (6.71) we see that

$$\frac{1}{2}E(T) \leq 0.$$

Obviously, $E(T) \geq 0$. Therefore, $E(T) = 0$ for all T. This implies that $w(x,T) \equiv 0$ for all $T > 0$ and $x \in \Omega$. $\square$

Exercise: Suppose that in the boundary value problem (6.70) we replace the Dirichlet condition ($u = h$ on $\partial\Omega$) with the Neumann condition $\nu\cdot\nabla u = g$ for $x \in \partial\Omega$. Here ν is the exterior normal vector at the boundary. Would Theorem 2.6.2 hold in this case?

2.7 Properties of Solutions

Infinite speed of propagation

Consider the function

$$u(x,t) = \int_{\mathbb{R}^d} \Phi(x-y,t)g(y)\,dy \tag{7.72}$$

Suppose that $g(y) \geq 0$. Then, since $\Phi(x,t) > 0$ for all $x \in \mathbb{R}$ and $t > 0$, the function $u(x,t)$ will be non-negative for all $x \in \mathbb{R}$ and $t > 0$. In fact, if $g(y) \geq 0$ is actually positive in some small region D, then since $\Phi(x,t) > 0$ for all $x \in \mathbb{R}$ and $t > 0$, the function $u(x,t)$ given by (7.72) will be positive for **all** $x \in \mathbb{R}$ and $t > 0$! This phenomenon is called **infinite speed of propagation**. Even if the initial data $g \geq 0$ is supported only in a small ball, the solution will be nonzero *everywhere* for any small time $t > 0$. This surprising property is quite different from the behavior of other time-dependent PDE like the wave equation.

Exercise. The heat equation was obtained as a limit of discrete equations appearing from the random walker model. The random walker, obviously, has a finite speed of propagation. Explain why there is no contradiction to the infinite speed of propagation of the heat equation and show that the random walker speed of propagation tends to infinity in the continuum limit.

Comparison principle, Maximum principle

This analysis also shows that the ordering of initial data is preserved by the corresponding solutions, in the following sense. Suppose that u and v both solve the heat equation (and satisfy the growth conditions) with initial data

$$u(x,0) = g_1(x) \leq g_2(x) = v(x,0) \tag{7.73}$$

Then the function $w = v - u$ solves the heat equation with initial data $g_2(x) - g_1(x) \geq 0$. So, w is non-negative for all $t > 0$, implying that $v \geq u$ for all x and $t > 0$.

Even in a bounded domain, solutions to the heat equation obey a comparison principle or **maximum principle**. To state this, we define the sets

$$\Omega_T = \Omega \times (0,T] \tag{7.74}$$

for some $T > 0$. This is an open set in $\mathbb{R}^d \times \mathbb{R}$. If Ω is a ball in $\mathbb{R}^2$, this set is a cylinder. In general, however, Ω_T is called the **parabolic cylinder**. Then the set

$$\Gamma_T = \overline{\Omega_T} \setminus \Omega_T \tag{7.75}$$

is the boundary portion on the bottom and sides of the cylinder (but not the top!). So, Γ_T resembles a cup, and it is called the **parabolic boundary**.

Theorem 2.7.1 (Weak Maximum Principle) *Suppose that $u \in C^{2,1}(\Omega_T) \cap C(\overline{\Omega}_T)$.*

(i) Suppose that $u_t \leq \Delta u$ for all $(x,t) \in \Omega_T$. Then

$$\max_{(x,t)\in\overline{\Omega}_T} u(x,t) = \max_{(x,t)\in\Gamma_T} u(x,t) \tag{7.76}$$

That is, the maximum of u must be attained on the boundary.

(ii) Suppose that $u_t \geq \Delta u$ for all $(x,t) \in \Omega_T$. Then

$$\min_{(x,t)\in\overline{\Omega}_T} u(x,t) = \min_{(x,t)\in\Gamma_T} u(x,t) \tag{7.77}$$

That is, the minimum of u must be attained on the boundary.

The assumption $u \in C^{2,1}(\Omega_T) \cap C(\overline{\Omega}_R)$ means that u has two space derivatives and one time derivative that are continuous for $(x,t) \in \Omega_T$. Also, u is continuous up to the boundary.

Proof: We prove only (i); the proof of (ii) is similar. Suppose that u attains its maximum at an interior point $(x_0, t_0) \in \Omega_T$:

$$u(x_0,t_0) = \max_{(x,t)\in\overline{\Omega}_T} u(x,t) \tag{7.78}$$

At a local maximum, we must have $u_t = 0$ and $\Delta u \leq 0$. Therefore, $u_t \geq \Delta u$ must be satisfied at this point, which means that $u_t < \Delta u$ could **not** hold at the maximum point (x_0, t_0). Therefore, if $u_t < \Delta u$ at all points in Ω_T, u cannot have a local maximum in the interior Ω_T.

For the general case, $u_t \leq \Delta u$, consider the function $w^\epsilon = u - \epsilon t$. This function satisfies

$$w^\epsilon_t = u_t - \epsilon \leq \Delta u - \epsilon = \Delta w^\epsilon - \epsilon < \Delta w^\epsilon \tag{7.79}$$

Therefore, we may apply the preceding argument to w^ϵ to conclude that

$$\max_{(x,t)\in\overline{\Omega}_T} w^\epsilon(x,t) \leq \max_{(x,t)\in\Gamma_T} w^\epsilon(x,t) \tag{7.80}$$

Letting $\epsilon \to 0$, $w^\epsilon \to u$ uniformly, and therefore

$$\max_{(x,t)\in\overline{\Omega}_T} u(x,t) = \lim_{\epsilon\to 0} \max_{(x,t)\in\overline{\Omega}_T} w^\epsilon(x,t) \leq \lim_{\epsilon\to 0} \max_{(x,t)\in\Gamma_T} w^\epsilon(x,t) = \max_{(x,t)\in\Gamma_T} u(x,t) \tag{7.81}$$

This proves (i). □

Corollary 2.7.1 (Comparison) *Suppose that u and v both solve the heat equation in the bounded domain Ω_T with $u(x,t) \geq v(x,t)$ for all $(x,t) \in \Gamma_T$. Then $u(x,t) \geq v(x,t)$ for all $x \in \overline{\Omega}_T$. That is, if u is greater than v on the parabolic boundary, then u is greater than v everywhere in the domain.*

Proof: The function $w = u - v$ satisfies the heat equation and is non-negative on the parabolic boundary Γ_T. The weak maximum principle implies that

$$\min_{(x,t)\in\overline{\Omega}_T} w(x,t) = \min_{(x,t)\in\Gamma_T} w(x,t) \geq 0. \tag{7.82}$$

So, $u \geq v$ for all $(x,t) \in \overline{\Omega}_T$. □

The maximum principle, as we will see later, is a very powerful tool that allows to obtain many qualitative results for parabolic equations in a very elegant and often effortless manner.

2.8 Non-constant coefficients

Let Ω be a bounded connected domain in $\mathbb{R}^n$. Consider the operator

$$Lu = a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + b_i(x)\frac{\partial u}{\partial x_i} + c(x)u$$

with $u \in C^2(\Omega) \cap C(\overline{\Omega})$. The functions a_{ij}, b_i and c are always assumed to be continuous in $\overline{\Omega}$ while L is assumed to be uniformly elliptic:

$$a_{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2 \text{ for all } x \in \Omega \text{ and all } \xi \in \mathbb{R}^n$$

with a positive constant $\lambda > 0$.

Such operators also obey the weak maximum principle. We will give an elliptic version but the parabolic version is verbatim the same as for the heat equation.

Theorem 2.8.1 *(The Weak Maximum Principle) Suppose that $u \in C^2(\Omega) \cap C(\overline{\Omega})$ satisfies $u \geq 0$ and $Lu \geq 0$ in Ω with $c(x) \leq 0$ in Ω. Then u attains its maximum on $\partial\Omega$.*

Proof. The proof is very close to that for the heat equation. The main observation is that if x_0 is an interior maximum then $\nabla u(x_0) = 0$ while the matrix $D^2u = \left(\frac{\partial^2 u}{\partial x_i \partial x_j}\right)$ is non-positive semi-definite. Hence the left side of $Lu \geq 0$ may not be positive at an interior maximum, as $c \leq 0$ and $u \geq 0$. This would be a contradiction were it not for the possibility $Lu = 0$ that arises if $D^2u(x_0)$ is degenerate – the rest of the proof fights this.

Given $\epsilon > 0$ define $w(x) = u(x) + \epsilon e^{\alpha x_1}$ with α to be determined – this will ensure that $D^2w(x_0)$ is non-degenerate. Then we have

$$Lw = Lu + \epsilon e^{\alpha x_1}\left(a_{11}\alpha^2 + b_1\alpha + c\right).$$

Recall that $a_{11} \geq \lambda > 0$ and $|b_1|, |c| \leq$ const. Thus we may choose $\alpha > 0$ so that

$$a_{11}\alpha^2 + b_1\alpha + c > 0$$

and thus $Lw > 0$ in Ω. Therefore the function w attains its maximum in $\bar{\Omega}$ at the boundary $\partial\Omega$. Indeed, as before, if w attains its (non-negative) maximum at $x_0 \notin \partial\Omega$ then $\nabla w(x_0) = 0$ and the matrix $D_{ij} = \dfrac{\partial^2 w}{\partial x_i \partial x_j}$ is non-positive definite. Hence we would have

$$Lw(x_0) = a_{ij}(x_0)D^2_{ij}w(x_0) + c(x_0)w(x_0) \leq 0$$

which is a contradiction. Thus $x_0 \in \partial\Omega$ and we obtain

$$\sup_\Omega u \leq \sup_\Omega w \leq \sup_{\partial\Omega} w \leq \sup_{\partial\Omega} u + \epsilon \sup_{\partial\Omega} e^{\alpha x_1} \leq C\epsilon + \sup_{\partial\Omega} u$$

with the constant C independent of ϵ. We let $\epsilon \to 0$ and finish the proof. $\square$

Corollary 2.8.1 *The Dirichlet problem*

$$Lu = f \text{ in } \Omega$$
$$u = \phi \text{ on } \partial\Omega$$

with $f \in C(\Omega)$, $\phi \in C(\partial\Omega)$ has a unique solution if $c(x) \leq 0$.

Remark 2.8.1 We first note that the assumption of non-negativity of u is not needed if $c = 0$. However, non-positivity of $c(x)$ is essential: otherwise the Dirichlet problem may have a non-unique solution, as $u(x, y) = \sin x \sin y$ solves

$$\Delta u + 2u = 0 \text{ in } \Omega$$
$$u = 0 \text{ on } \partial\Omega$$

with $\Omega = [0, \pi] \times [0, \pi] \subset \mathbb{R}^2$. Second, the assumption that Ω is bounded is also essential both for uniqueness and for the maximum principle to hold: the function $u(x) = \log |x|$ solves

$$\Delta u = 0 \text{ in } \Omega$$
$$u = 0 \text{ on } \partial\Omega$$

with $\Omega = \{|x| > 1\}$.

The most immediate application of the maximum principle is to obtain uniform a priori estimates on the solutions. We still assume that the matrix a_{ij} is uniformly elliptic in $\bar{\Omega}$: $a_{ij}(x)\xi_i\xi_j \geq \lambda|\xi|^2$, $\lambda > 0$, and a_{ij}, b_i and c are continuous in $\bar{\Omega}$. We assume in addition that

$$\sup_{\Omega} |a_{ij}| + \sup_{\Omega} |b_i| \leq \Lambda.$$

The first result deals with the Dirichlet boundary conditions.

Theorem 2.8.2 *Assume that $u \in C^2(\Omega) \cap C(\bar{\Omega})$ satisfies*

$$Lu = f \text{ in } \Omega$$
$$u = \phi \text{ on } \partial\Omega$$

for some $f \in C(\bar{\Omega})$ and $\phi \in C(\partial\Omega)$. There exists a constant $C(\lambda, \Lambda, \text{diam}(\Omega))$ so that

$$|u(x)| \leq \max_{\partial\Omega} |\phi| + C \max_{\Omega} |f| \text{ for all } x \in \Omega \tag{8.83}$$

provided that $c(x) \leq 0$.

Proof. Let us denote $F = \max_\Omega |f|$ and $\Phi = \max_{\partial\Omega} |\phi|$ and assume that Ω lies inside a strip $\{0 < x_1 < d\}$. Define $w(x) = \Phi + \left(e^{\alpha d} - e^{\alpha x_1}\right) F$ with $\alpha > 0$ to be chosen so as to ensure

$$Lw \leq -F \text{ in } \Omega \tag{8.84}$$
$$w \geq \Phi \text{ on } \partial\Omega.$$

We calculate $w \geq \Phi$ on $\partial\Omega$ and

$$-Lw = (a_{11}\alpha^2 + b_1\alpha)Fe^{\alpha x_1} - c\Phi - c\left(e^{\alpha d} - e^{\alpha x_1}\right) F \geq (a_{11}\alpha^2 + b_1\alpha)F \geq (\lambda\alpha^2 + b_1\alpha)F \geq F$$

when α is large enough. Hence w satisfies (8.84). The comparison principle implies that $-w \leq u \leq w$ in Ω and in particular

$$\sup_{\Omega} |u| \leq \Phi + \left(e^{\alpha d} - 1\right) F$$

so that (8.83) holds. □

Chapter 3

Stochastic Processes and Stochastic Calculus

We already learned that the function

$$u(x,t) = \int_{\mathbb{R}} \Phi(x-y,t)g(y)\,dy = \int_{\mathbb{R}} \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy = E[g(X_x(t))] \tag{0.1}$$

satisfies the heat equation for $x \in \mathbb{R}$, $t > 0$ with initial data $u(x,0) = g(x)$ (defined by continuity as $t \to 0$). The expectation in (0.1) is with respect to a Gaussian random variable $X_x(t)$ with mean $\mu = x$ and variance $\sigma^2 = 2t$. Now we begin to explore more deeply the connection between second order linear partial differential equations and stochastic processes.

These notes review basic concepts related to continuous time stochastic processes, Brownian motion, and stochastic calculus, which are relavent to our study of PDEs. I presume that you have had some previous exposure to the topics in Section 3.1 of these notes (in Stat219, for example). The more advanced topics are covered in much greater detail in the Math 236 course; here we only review and sketch the main ideas that we will need later. There are several excellent texts covering these topics. Here are a few:

- Course notes for Math136/Stat219, available online.
- B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Berlin: Springer-Verlag 2003.
- I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, New York: Springer 1991.

3.1 Continuous Time Stochastic Processes and Brownian Motion

3.1.1 Measure theory

Given a probability space $(\Omega, \mathcal{F}, P)$, a **continous time stochastic process** is a collection of random variables $\{X_t(\omega)\}_{t\in[0,T]}$ indexed by the "time variable" t. Here Ω is the sample space, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and P is a probability measure. Recall that a **σ-algebra** is a collection of subsets of Ω that is closed under countable unions or intersections. That is,

(i) $\Omega \in \mathcal{F}, \quad \emptyset \in \mathcal{F}$

(ii) $A^c \in \mathcal{F}$ whenever $A \in \mathcal{F}$, where $A^c = \Omega \setminus A$.

(iii) $\bigcup_{i=1}^{\infty} A_i \subset \mathcal{F}$ whenever $A_i \in \mathcal{F}$ for each $i = 1, 2, 3, \ldots$.

(iii) $\bigcap_{i=1}^{\infty} A_i \subset \mathcal{F}$ whenever $A_i \in \mathcal{F}$ for each $i = 1, 2, 3, \ldots$.

A **probability measure** $P : \mathcal{F} \to [0, 1]$ assigns to each set $A \in \mathcal{F}$ a number $P(A) \geq 0$ in such a way that

(i) $P(\Omega) = 1$

(ii) $P(\emptyset) = 0$.

(iii) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ whenever $\{A_i\}_{i=1}^{\infty}$ is a countable collection of disjoint sets in $A_i \in \mathcal{F}$.

So, the collection of sets $\mathcal{F}$ are called the **measureable sets**; these are the sets to which we can assign a probability. If A is not measureable, we run into problems trying to define $P(A)$.

For each t, $X_t(\omega)$ is a measurable random variable taking values in $\mathbb{R}$, or $\mathbb{R}^d$. Measureable means that for each $\alpha \in \mathbb{R}$, the set $\{\omega \mid \ X_t(\omega) > \alpha\}$ is a measurable set (i.e. contained in $\mathcal{F}$). Thus we can assign a probability to such sets. If an event $A \in \mathcal{F}$ satisfies $P(A) = 1$, we say that A occurs "with probability 1" or "almost surely".

For each ω, we say that the function $t \mapsto X_t(\omega)$ is a **sample path** or **trajectory** or **realization** of the stochastic process. A nondecreasing family of σ-algebras $\{\mathcal{F}_t\}_{t\in[0,T]}$, $\mathcal{F}_t \subset \mathcal{F}$ is called a **filtration**. This means that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ whenever $s \leq t$. We say that the process $X_t(\omega)$ is **adapted** to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ if for each $t \geq 0$, $X_t(\omega)$ is an $\mathcal{F}_t$ measurable random variable (such a process is also said to be **non-anticipating**). Typically, $\mathcal{F}_t$ is defined as

$$\mathcal{F}_t = \sigma\left(X_s(\omega), \quad s \in [0, t]\right) \tag{1.2}$$

The notation on the right means the σ-algebra **generated** by the process up to time t. It is the smallest σ-algebra contained in $\mathcal{F}$ with respect to which $X_s(\omega)$ is measurable for each $s \in [0, t]$. One may think of $\mathcal{F}_t$ as containing all "information" about the process up to time t. We will always be working with processes that are also measureable with respect to t, for almost every ω. That is, the function $t \mapsto X_t(\omega)$ defines a measurable function of $[0, T]$ with respect to the Borel σ-algebra $\mathcal{B}([0, T])$.

3.1.2 Conditional Expectation

Sometimes we want to estimate the probability of an event given some partial information about the process. For example, given the history of the process up to time s, we might want to compute $E[f(X_t(\omega))]$, given that $X_s(\omega) = x$ is known. This is the idea behind conditional expectation. The problem is that the quantity $f(X_t(\omega))$ is measurable with respect to $\mathcal{F}_t$, but if $t > s$ then this quantity may **not** be measurable with respect to $\mathcal{F}_s$. The conditional expectation of $f(X_t(\omega))$ with respect to $\mathcal{F}_s$ (the "history up to time s") is written as

$$E[f(X_t(\omega)) \mid \ \mathcal{F}_s] \tag{1.3}$$

This is a random variable, in general. If $\eta(\omega)$ is a $\mathcal{F}_t$ measureable random variable (for example $\eta(\omega) = f(X_t(\omega))$), then the **conditional expectation** $E[\eta|\mathcal{F}_s]$ is the $\mathcal{F}_s$ measureable random variable $\xi(\omega)$ satisfying

$$E[\xi(\omega)\mathbb{I}_A(\omega)] = E[\eta(\omega)\mathbb{I}_A(\omega)] \tag{1.4}$$

for all $A \in \mathcal{F}_s$. Here $\mathbb{I}_A(\omega)$ is the **indicator function** of the set A: $\mathbb{I}_A(\omega) = 1$ if $\omega \in A$, $\mathbb{I}_A(\omega) = 0$ otherwise. We can also write this relationship in integral notation:

$$\int_A \xi(\omega)dP(\omega) = \int_A \eta(\omega)dP(\omega) \tag{1.5}$$

The point is that ξ is actually $\mathcal{F}_s$ measurable, and from the point of view of integration over $\mathcal{F}_s$ measurable sets, ξ is indistinguishable from η. That is, integration of ξ over any set $A \in \mathcal{F}_s$ is the same as integration of η over that set. Although η may not be $\mathcal{F}_s$ measureable, the integral on the right hand side of (1.5) makes sense since $A \in \mathcal{F}_t$, as well. One can show that the conditional expectation is uniquely defined, up to modification on a set of measure zero. That is, if $\xi_1(\omega)$ and $\xi_2(\omega)$ both satisfy the definition, then $\xi_1(\omega) = \xi_2(\omega)$ almost surely.

One can also show that among all $\mathcal{F}_s$ measurable random variables, the conditional expectation $E[\eta|\mathcal{F}_s]$ is the random variable ξ that minimizes

$$E[|\xi(\omega) - \eta(\omega)|^2], \tag{1.6}$$

provided that these random variables have bounded second moments. So, in this sense the conditional expectation $\xi = E[\eta|\mathcal{F}_s]$ is the best approximation to η among $\mathcal{F}_s$ measurable random variables.

Here are a few very useful properties of conditional expectation. We suppose that η is $\mathcal{F}_t$ measureable with $\mathcal{F}_s \subset \mathcal{F}_t$ (i.e. $s < t$):

- $E[\eta \mid \mathcal{F}_s] \geq 0$ if $\eta(\omega) \geq 0$.
- For constants α_1 and α_2, $E[\alpha_1\eta_1(\omega) + \alpha_2\eta_2(\omega) \mid \mathcal{F}_s] = \alpha_1 E[\eta_1 \mid \mathcal{F}_s] + \alpha_2 E[\eta_2 \mid \mathcal{F}_s]$, .
- If $\phi(\omega)$ is a $\mathcal{F}_s$ measurable random variable, then $E[\phi \mid \mathcal{F}_s] = \phi$, almost surely.
- If $\phi(\omega)$ is a $\mathcal{F}_s$ measurable random variable, then $E[\phi\eta \mid \mathcal{F}_s] = \phi E[\eta \mid \mathcal{F}_s]$, assuming these quantities are finite.
- $E[\eta] = E[E[\eta \,|\mathcal{F}_s]]$
- If $\phi(\omega)$ is $\mathcal{F}_s$ measurable, then

$$E[\phi\eta] = E[E[\phi\eta \,|\mathcal{F}_s]] = E[\phi E[\eta \,|\mathcal{F}_s]] \tag{1.7}$$

Suppose η and ξ are two random variables, $\mathcal{F}$ measureable. We may define the conditional expectation of η with respect to ξ as follows. Let $\mathcal{F}_\xi$ be the σ-algebra generated by $\xi(\omega)$. This is the smallest σ-algebra contained in $\mathcal{F}$ with respect to which ξ is measureable. Then we define $E[\eta|\xi] = E[\eta|\mathcal{F}_\xi]$. Typically we will encounter quantities like $E[X_{t_2}|X_{t_1}]$ for $t_2 > t_1$, where X_t is a stochastic process.

We may also define the **conditional probability** of an event with respect to a σ-algebra. Given an event $A \in \mathcal{F}$, we define the conditional probability $P(A|\mathcal{F}_s)$ to be the conditional expectation

$E[\mathbb{I}_A(\omega)|\mathcal{F}_s]$, where $\mathbb{I}_A(\omega)$ is the indicator function of the set A. Similarly, we define the conditional probability of one event with respect to another as $P(A|C) = E[\mathbb{I}_A(\omega)|\mathcal{F}_C]$ where $\mathcal{F}_C$ is the σ-algebra generated by the random variable $\mathbb{I}_C(\omega)$. Notice that this σ-algebra is particularly simple, containing only the sets $\mathcal{F}_C = \{\Omega, \emptyset, C, C^c\}$.

3.1.3 Brownian Motion

A real-valued continuous time stochastic process $B_t(\omega)$, $t \in [0, T]$, is called a **Brownian motion** or **Wiener process** if

(i) $B_t(\omega)$ is a Gaussian process

(ii) $E[B_t] = 0$ for all t

(iii) $E[B_t B_s] = \min(t, s)$

(iv) For almost every $\omega \in \Omega$, the sample path $t \to B_t(\omega)$ is a continuous function of $t \in [0, T]$.

In many texts, Brownian motion is denoted by B_t, instead of B_t. This definition says that for each t, $B_t(\omega)$ is a Gaussian random variable with mean zero and variance t:

$$P(B_t(\omega) \in [a, b]) = \frac{1}{\sqrt{2\pi t}} \int_a^b e^{-\frac{x^2}{2t}}\, dx \tag{1.8}$$

Notice also that these properties imply that increments of the process (i.e. $B_t - B_s$) are also mean zero Gaussian random variables with variance $t - s$ since

$$E[(B_t - B_s)^2] = E[(B_t)^2] - 2E[B_t B_s] + E[B_s^2] = t - 2s + s = t - s \tag{1.9}$$

Moreover, if $0 \le t_1 \le t_2 \le t_3 \le t_4 \le T$, the increments $B_{t_4} - B_{t_3}$ and $B_{t_2} - B_{t_1}$ are mutually independent since these random variables are jointly Guassian and uncorrelated:

$$\begin{aligned} E[(B_{t_4} - B_{t_3})(B_{t_2} - B_{t_1})] &= E[B_{t_4} B_{t_2}] + E[B_{t_3} B_{t_1}] - E[B_{t_4} B_{t_1}] - E[B_{t_3} B_{t_2}] \\ &= t_2 + t_1 - t_1 - t_2 = 0 \end{aligned} \tag{1.10}$$

Theorem 3.1.1 *There exists a complete probability space $(\Omega, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$, and a stochastic process $B_t(\omega)$ such that $B_t(\omega)$ is a Brownian motion.*

There are various ways to construct Brownian motion. For examples and proofs of this theorem, see the references listed above. The first proof of this result was given by Norbert Wiener (J. Math. Phys., 1923)

Brownian motion in $\mathbb{R}^d$. Similarly, we may define Brownian motion in $\mathbb{R}^d$. A d-dimensional Brownian motion is an $\mathbb{R}^d$-valued continuous time stochastic process $B_t(\omega) = (B_t^1(\omega), \ldots, B_t^d(\omega))$, $t \in [0, T]$ such that the components $B_t^i(\omega)$ are independent one-dimensional Brownian motions. Therefore,

(i) $B_t^i(\omega)$ is a Gaussian process for each i

(ii) $E[B_t^i] = 0$ for all t, $i = 1, \ldots, d$

(iii) $E[B_t^i B_s^j] = \min(t,s)\delta_{ij}$, for all i and j.

(iv) For almost every $\omega \in \Omega$, the sample path $t \to B_t(\omega)$ is a continuous function from $[0,T]$ to $\mathbb{R}^d$.

3.1.4 Properties of Brownian Motion

Next, we describe some important properties of one-dimensional Brownian motion. Proofs and further discussion of these statements may be found in the references listed above.

Markov Property and Independent Increments. Brownian Motion is a **Markov process**. This implies that if $s < t$, the increment $B_t(\omega) - B_s(\omega)$ is independent of the σ-algebra $\mathcal{F}_s$. In terms of conditional probability, this means that for any Borel set A,

$$P(B_t(\omega) - B_s(\omega) \in A \mid \mathcal{F}_s) = P(B_t(\omega) - B_s(\omega) \in A \mid B_s) \tag{1.11}$$

Also, for any continuous bounded function f

$$E[f(B_t - B_s)|\mathcal{F}_s] = E[f(B_t - B_s)|B_s] \tag{1.12}$$

That is, if $B_s(\omega)$ is known, then $B_t(\omega) - B_s(\omega)$ depends on no other information in the past before time s. In fact,

$$P(B_t(\omega) - B_s(\omega) \in A \mid B_s) = \frac{1}{\sqrt{2\pi(t-s)}} \int_A e^{-\frac{|B_s - y|^2}{2(t-s)}}\, dy \tag{1.13}$$

so that

$$E\left[B_t(\omega) - B_s(\omega) \mid B_s\right] = 0 \tag{1.14}$$

almost surely. For $t_1 < t_2 < t_3 < t_4 \leq T$, the increments $B_{t_4} - B_{t_3}$ and $B_{t_2} - B_{t_1}$ are mutually independent so that

$$P\left(B_{t_4} - B_{t_3} \in A,\ \ B_{t_2} - B_{t_1} \in C\right) = P\left(B_{t_4} - B_{t_3} \in A\right) P\left(B_{t_2} - B_{t_1} \in C\right) \tag{1.15}$$

Finite Dimensional Distributions. For any $t_1 < t_2 < t_3 < \cdots < t_n$, the random variables $\{B_{t_k}\}_{k=1}^n$ are jointly Guassian with covariance determined by properties (ii) and (iii) above. Given any collection of Borel sets $\{A_k\}_{k=1}^n$, let us compute

$$P(B_{t_k} \in A_k, \quad k = 1, \ldots n) \tag{1.16}$$

A set of the form $\{B_{t_k} \in A_k, \quad k = 1, \ldots n\} \subset C([0,T])$ is called a **cyllinder set**. The increments of Brownian motion are independent. Therefore, $B_2 = B_1 + (B_1 - B_2)$ expresses B_2 as a sum of independent Gaussian random variables. Similarly, writing $X_j = B_j - B_{j-1}$ for $j = 1, \ldots, n$ ($X_0 = 0$), we may write $B_k = \sum_{j=1}^k X_j$ as a sum of independent Gaussian random variables, each having mean zero and variance $t_j - t_{j-1}$. Therefore,

$$\begin{aligned} P\left(B_{t_1} \in A_1,\ B_{t_2} \in A_2\right) &= P\left(X_1 \in A_1,\ X_1 + X_2 \in A_2\right) \\ &= \int_{A_1} \int_{A_2 - x_1} p(x_1; 0, t_1) p(x_2; 0, t_2 - t_1)\, dx_2\, dx_1 \\ &= \int_{A_1} \int_{A_2} p(x_1; 0, t_1) p(x_2; x_1, t_2 - t_1)\, dx_2\, dx_1 \end{aligned} \tag{1.17}$$

where $p(x;y,t)$ is the Gaussian kernel

$$p(x;y,t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{|x-y|^2}{2t}} \tag{1.18}$$

Proceeding inductively in this way we find that

$$P(B_{t_k} \in A_k, \quad k = 1, \dots n) = \int_{A_1} \cdots \int_{A_n} \prod_{k=1}^{n} p(x_k; x_{k-1}, t_k - t_{k-1})\, dx_n \cdots dx_1$$

and for any continuous and bounded $f : \mathbb{R}^n \to \mathbb{R}$

$$\begin{aligned} &E\left[f(B_1, \dots, B_n)\right] \\ &\quad = \int_{A_1} \cdots \int_{A_n} f(x_1, x_2, \dots, x_{n-1}, x_n) \prod_{k=1}^{n} p(x_k; x_{k-1}, t_k - t_{k-1})\, dx_n \cdots dx_1 \end{aligned}$$

When $k = 1$, we define $x_{k-1} = x_0 := 0$ in this formula.

The Martingale Property. An $\mathcal{F}_t$-adapted stochastic process $X_t(\omega)$ is called a martingale, sub-martingale, or super-martingale in the following cases:

- **Martingale:**
$$E[X_t|\mathcal{F}_s] = X_s \quad \text{for all } s \in [0,t], \text{ almost surely} \tag{1.19}$$
- **Sub-martingale:**
$$E[X_t|\mathcal{F}_s] \geq X_s \quad \text{for all } s \in [0,t], \text{ almost surely} \tag{1.20}$$
- **Super-martingale:**
$$E[X_t|\mathcal{F}_s] \leq X_s \quad \text{for all } s \in [0,t], \text{ almost surely} \tag{1.21}$$

A one-dimensional Brownian motion $B_t(\omega)$ is a martingale, since

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= E[X_t - X_s + X_s|\mathcal{F}_s] &(1.22)\\ &= E[(X_t - X_s)|X_s] + E[X_s|X_s] \\ &= E[(X_t - X_s)] + X_s = 0 + X_s. &(1.23)\end{aligned}$$

One may think of a martingale as a "fair game." If X_t represents a gambler's account balance at time t, then the condition $E[X_t|\mathcal{F}_s] = X_s$ says that the expected future balance, given the current balance, is unchanged. So the game favors neither the gambler nor the house. Of course, the change $X_t - X_s$ may be positive or negative, but its expected value conditioned on X_s is zero. On the other hand, if X_t is a sub-martingale, then given the gambler's current account balance, he may expect his earnings to increase. If X_t is a super-martingale, then given the gambler's current account balance, he may expect his earnings to decrease (this seems to be the most realistic model, given the success of many casinos).

A useful observation is that if X_s is an $\mathcal{F}_s$-martingale and a function $\phi(x)$ is convex, then $\phi(X_s)$ is a sub-martingale. This follows from Jensen's inequality. For instance, if if X_s is a martingale, so are $|X_s|$, X_s^2, and all X_s^{2m} with $m \in \mathbb{N}$. Continuous martingales satisfy a remarkable property that estimates the maximum of a process by the terminal time statistics.

Theorem 3.1.2 *(Continuous Doob inequality) If M_t is a continuous in time martingale, then for all $p \geq 1$, $T \geq 0$ and $\lambda > 0$ we have*

$$P\left[\sup_{0\leq t\leq T} |M_t| \geq \lambda\right] \leq \frac{1}{\lambda^p}\mathbb{E}(|M_T|^p).$$

We will not prove this result here but rather prove it only for discrete martingales. A sequence X_j is a martingale with respect to a sequence of σ-algebras $\mathcal{F}_j$ if (i) $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, (ii) X_n is $\mathcal{F}_n$-measurable, (iii) $E[|X_n|] < +\infty$, and (iv) $E(X_{n+1}|\mathcal{F}_n) = X_n$ almost surely. It follows that $E(X_m|\mathcal{F}_n) = X_n$ almost surely for all $m \geq n$. The discrete Doob's inequality is the following estimate that bounds the supremum of X_j in terms of the expectation of the last element:

Theorem 3.1.3 *(Discrete Doob's inequality) Suppose $(X_j, \mathcal{F}_j)$ is a martingale sequenc, $1 \leq j \leq n$, then for any $l > 0$ and any $p \geq 1$we have*

$$P\Big\{\omega : \sup_{1\leq j\leq n} |X_j| \geq l\Big\} \leq \frac{1}{l^p}E(|X_n|^p).$$

Proof. Let us define $S(\omega) = \sup_{1\leq j\leq n} |X_j(\omega)$. Then the event $E = \{\omega : \ S(\omega) \geq l\}$ can be decomposed as a disjoint union of the sets

$$E_j = \{\omega : \ |X_1(\omega)| < l, \ldots, |X_{j-1}(\omega) < l, |X_j(\omega) \geq l\},$$

that is, $E = \bigcup_{j=1}^n E_j$ and $E_j \cap E_m = \emptyset$ for $j \neq m$. Note that, as $|X_j| \geq l$ on the set E_j we have an inequality

$$P(E_j) \leq \frac{1}{l^p}\int_{E_j} |X_j|^p dP.$$

The function $\phi(x) = |x|^p$ is convex for $p \geq 1$, hence, as we mentioned above, the sequence $|X_j|^p$ is a sub-martingale, thus $|X_j|^p \leq \mathbb{E}(|X_n|^p|\mathcal{F}_j)$, hence

$$P(E_j) \leq \frac{1}{l^p}\int_{E_j} |X_j|^p dP \leq \frac{1}{l^p}\int_{E_j} \mathbb{E}(|X_n|^p|\mathcal{F}_j)dP$$

Moreover, the set E_j is $\mathcal{F}_j$-measurable as follows immediately from the way E_j is defined, hence

$$P(E_j) \leq \frac{1}{l^p}\int_{E_j} \mathbb{E}(|X_n|^p|\mathcal{F}_j)dP = \frac{1}{l^p}\int_{E_j} |X_n|^p dP,$$

simply from the definition of $\mathbb{E}(|X_n|^p|\mathcal{F}_j)$. now, summing over all j and using the fact that E_j are disjoint we obtain

$$P(E) = \sum_{j=1}^n P(E_j) \leq \frac{1}{l^p}\sum_{j=1}^n \int_{E_j} |X_n|^p dP \leq \frac{1}{l^p}\int_E |X_n|^p dP \leq \frac{1}{l^p}\int_\Omega |X_n|^p dP = \frac{1}{l^p}\mathbb{E}(|X_n|^p),$$

and we are done. □

Path space $C([0,T])$ and Wiener measure. A Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ induces a probability measure P^* on $C([0,T])$, the space of real-valued continuous functions on the interval $[0,T]$. The σ-algebra on which this measure is defined is the Borel σ-algebra

(denoted by $\mathcal{B}(C([0,T]))$) that is generated by the uniform norm. This is the σ-algebra generated by open and closed sets determined by the metric $d(f,g) = \|f-g\|_0 = \sup_{s\in[0,T]}|f(s)-g(s)|$. Consequently, the probability space $(C([0,T]), \mathcal{B}(C([0,T])), P^*)$ is refered to as the **canonical probability space** for Brownian motion, and the measure P^* is called the **Wiener measure**. The condition $B_0(\omega) = 0$ almost surely means that this measure concentrates on functions that are 0 for $t = 0$.

Regularity. Although sample paths of Brownian motion are continuous, almost surely, they are not smooth. In fact, with probability one, the sample paths are nowhere differentiable in t. Thus $\frac{d}{dt}B_t(\omega)$ is ill-defined for all t. It can be shown, however, that sample paths are almost surely Hölder continuous with exponent α for any $\alpha \in [0, 1/2)$ (this follows from the Kolmogorov Čentsov Theorem). This means that if $\alpha \in [0, 1/2)$ is a fixed constant, then the quantity

$$\sup_{\substack{s,t\in[0,T]\\ s\neq t}} \frac{|B_t(\omega) - B_s(\omega)|}{|t-s|^\alpha} \tag{1.24}$$

if finite, almost surely, where the bound depends on ω. In other words, for almost every ω, there is a constant $C = C(\omega)$ such that

$$|B_t(\omega) - B_s(\omega)| \leq C|t-s|^\alpha \tag{1.25}$$

holds for all $t, s \in [0,T]$. This gives us some measure of the continuity of $B_t(\omega)$. If $B_t(\omega)$ were differentiable in t, then this condition would be satisfied with $\alpha = 1$.

3.1.5 Stopping times

A real-valued, $\mathcal{F}$ measurable random variable $\tau(\omega)$ is called a **stopping time** with respect to a filtration $\{\mathcal{F}_t\}_t$ if the event $\{\tau \leq t\}$ is $\mathcal{F}_t$ measureable for each t. That is $\{\tau \leq t\} \in \mathcal{F}_t$. This means that at each time t, we may determine whether or not the even has occured. Typically we will consider stopping times of the form,

$$\tau(\omega) = \inf\{s \geq 0 |\ B_s(\omega) \in A\} \tag{1.26}$$

where A is some Borel set in $\mathbb{R}$ (or $\mathbb{R}^d$ if B_t is multi-dimensional). This stopping time is refered to as the **first hitting time** to the set A. On the other hand, the random variable

$$\tau(\omega) = \sup\{s \geq 0 |\ B_s(\omega) \in A\}$$

is **not** a stopping time, since we cannot determine whether the process will re-enter A at some point in the future (i.e. $\{\tau \leq t\}$ is not $\mathcal{F}_t$ measureable).

3.2 The Itô Integral and Stochastic Differential Equations

In this section we describe a method of giving rigorous meaning to a differential equation of the form

$$\begin{aligned} \frac{d}{dt}y(t) &= g(y(t),t) + f(y(t),t)\cdot(\text{random noise}) \\ &= g(y(t),t) + f(y(t),t)B'(t). \end{aligned} \tag{2.27}$$

In integral form, such and equation would look like

$$y(T) = y(0) + \int_0^T g(y(s), s)\, ds + \int_0^T f(y(t), s)B'(s)\, ds. \tag{2.28}$$

Of course, $B(s)$ is not differentiable for any s, so neither (2.27) nor (2.28) makes any sense.

Total Variation, Riemann-Stieltjes Integrals

Suppose $f(t)$ is a continuous function, then the **Riemann-Stieltjes integral** of f with respect to $\phi(s)$ over $[0, T]$ is defined by

$$I(T) = \int_0^T f(s)d\phi(s) := \lim_{|\Gamma|\to 0} \sum_{i=1}^m f(p_i)\left(\phi(s_i) - \phi(s_{i-1})\right) \tag{2.29}$$

where $\Gamma = \{0 = s_0 < s_1 < \cdots < s_m = T\}$ is a partition of the interval $[0, T]$, and each p_i is a point satisfying $s_{i-1} \le p_i \le s_i$. By $|\Gamma| \to 0$ we mean that the maximum distance between two consecutive points in Γ goes to zero. The limit (2.29) may not exist, in general. However, the following condition on ϕ will be sufficient to make sense of this limit. We say that a function $\phi : [0, T] \to \mathbb{R}$ has **bounded total variation** if

$$TV(\phi) = \sup \sum_{\Gamma} |\phi(s_i) - \phi(s_{i-1})| < \infty \tag{2.30}$$

where the supremum is taken over all possible partitions Γ. The expression TV stands for **total variation**.

Theorem 3.2.1 *If $f(s)$ is continuous over $[0, T]$, and ϕ has bounded total variation, then the integral $I(T)$ is well-defined and finite.*

Notice that if $\phi(s) = s$, then the integral defined by (2.29) is just the usual Riemann-integral. More generally, it is easy to see that if ϕ is continuously differentiable, then

$$\int_0^T f(s)d\phi(s) = \int_0^T f(s)\phi'(s)\, ds.$$

Nevertheless, the Riemann-Stieltjes integral may be defined for functions that have bounded variation but are not necessarily continuously differentiable. Unfortunately, Brownian motion does not even fit into this class.

Theorem 3.2.2 *The sample paths of Brownian motion have unbounded total variation, almost surely.*

Indeed, if Γ is a partition of length m with evenly spaced points $s_i = iT/m$, then we may compute explicitly

$$E\left[\sum_{\Gamma} |B_{s_i} - B_{s_{i-1}}|\right] = \sum_{i=1}^m E[|B_{s_i} - B_{s_{i-1}}|] = C\sum_{i=1}^m \sqrt{|s_i - s_{i-1}|} = Cm\sqrt{\frac{T}{m}} \tag{2.31}$$

This quantity diverges as $m \to \infty$. This shows that the mean variation is infinite. One can show that the same is true pathwise, almost surely. Consequently, $\int_0^T f(s)dB_s$ is not necessarily well-defined as a Riemann-Stieltjes integral.

The Itô Integral

In order to define the integral $\int_0^T f(s)dB_s$, we first define it for a simple class of processes. Then for a more general class of processes, we define the integral by approximation. The resulting definition is called the **Itô stochastic integral**.

Definition 3.2.1 *A stochastic process $f(s,\omega)$ is called a* **simple process** *or* **step process** *if there is a partition $\Gamma = \{0 = s_0 < s_1 < \cdots < s_m = T\}$ such that*

$$f(s) = \xi_i(\omega) \quad \text{if } s \in [s_i, s_{i+1}), \;\; i = 0, \ldots, m-1 \tag{2.32}$$

and the random variable $\xi_i(\omega)$ is $\mathcal{F}_{s_i}$ measurable for each $i = 0, \ldots, m-1$, and $E[|\xi_i|^2] < \infty$.

So a step process is piecewise constant with finitely many pieces, and it is easy to see that a step process is adapted to $\mathcal{F}_t$ (non-anticipating). For such a process we define the integral

$$\int_0^T f(s,\omega)dB_s(\omega) := \sum_{i=0}^{m-1} f(s_i)(B_{s_{i+1}} - B_{s_i}) = \sum_{i=0}^{m-1} \xi_i(\omega)(B_{s_{i+1}} - B_{s_i}) \tag{2.33}$$

There are only finitely many terms in the sum, thus there is no problem with convergence of the sum. Now we extend the definition to a larger class of processes. Let $\mathcal{L}^2([0,T])$ denote the class of $\mathcal{F}_t$-adapted processes $g(s,\omega)$ satisfying

$$\|g\|_{\mathcal{L}^2} := \left(E\left[\int_0^T g(t,\omega)^2 \, dt \right] \right)^{1/2} < \infty \tag{2.34}$$

Clearly any step process is contained in $\mathcal{L}^2([0,T])$. In fact, the set of all step processes is a dense subset of $\mathcal{L}^2([0,T])$ in the following sense:

Theorem 3.2.3 *For any process $g(s,\omega) \in \mathcal{L}^2([0,T])$, there is a sequence of bounded step processes $f_k(s,\omega) \in \mathcal{L}^2([0,T])$ such that*

$$\lim_{k\to\infty} E\left[\int_0^T (g(s,\omega) - f_k(s,\omega))^2 \, ds \right] = 0 \tag{2.35}$$

Proof: This is proved in the book by B. Oksendal (see steps 1,2,3 of section 3.1). The proof is not difficult, and you should first try to sketch a proof on your own. □

The theorem says that we may approximate g by a sequence of simple processes $\{f_k\}_k$. Since each f_k is a simple process, we may integrate each f_k as defined by (2.33). Then we may define the Itô integral for g by approximation – the integral of g is the limit of the integrals of the f_k.

Definition 3.2.2 *For $g(t,\omega) \in \mathcal{L}^2([0,T])$, the Itô integral of g is defined by*

$$\int_0^T g(s,\omega)dB_s(\omega) := \lim_{k\to\infty} \int_0^T f_k(s,\omega)dB_s(\omega) \tag{2.36}$$

where $f_k(s,\omega)$ is a sequence of bounded step functions approximating g, in the sense of (2.35). The integrals on the right are defined by (2.33).

The convergence in (2.36) holds in $L^2((\Omega, \mathcal{F}, P))$. This means that there exists a square-integrable random variable $\eta(\omega)$ such that

$$E[\eta^2] < \infty \qquad \text{and} \qquad \lim_{k\to\infty} E\left[\left|\eta(\omega) - \int_0^T f_k(s,\omega)dB_s(\omega)\right|^2\right] = 0 \tag{2.37}$$

This random variable is what we call $\int_0^T g(s,\omega)dB_s(\omega)$. Moreover, the definition is independent of the choice of $\{f_k\}$; if we choose a different approximating sequence, we obtain the same limit (after modification on a set of zero measure)! Proofs of these statements may be found in the Oksendal book.

Properties of the Itô Integral

Here are some important properties of the Itô integral:

(i) $\int_0^T dB_s(w) = B_T(\omega)$.

(ii) The integral is a linear functional of the process g:

$$\int_0^T (\alpha g_1(s,\omega) + \beta g_2(s,\omega))\, dB_s(\omega) = \alpha \int_0^T g_1(s,\omega)dB_s(\omega) + \beta \int_0^T g_2(s,\omega)dB_s(\omega) \tag{2.38}$$

(iii) **Itô Isometry**: for any $f, g \in \mathcal{L}^2([0,T])$

$$E[|\int_0^T g(s,\omega)\, dB|^2] = \int_0^T E[|g(s,\omega)|^2]\, ds \tag{2.39}$$

and

$$E[\int_0^T g(s,\omega)\, dB \cdot \int_0^T f(s,\omega)\, dB] = \int_0^T E[g(s,\omega)f(s,\omega)\, ds. \tag{2.40}$$

(iv) $E[\int_0^T g(s,\omega)\, dB] = 0$.

(v) The indefinite integral $I(t,\omega) := \int_0^t g(s,\omega)\, dB_s(\omega)$ defines a $\mathcal{F}_t$ measurable random variable.

(vi) The indefinite integral $I(t,\omega) := \int_0^t g(s,\omega)\, dB_s(\omega)$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.

(vii) There is a version of the stochastic process $I(t,\omega) := \int_0^t g(s,\omega)\, dB_s(\omega)$ that has continuous sample paths.

Proof: Property (i) is obvious, since this is the integral of the step process $f(s,\omega) \equiv 1$. Property (ii) is also easy to check.

Now we prove (iii). Suppose that g is a simple process (the general case $g \in \mathcal{L}^2([0,T])$ follows by approximation). Then

$$E[|\int_0^T g(s,\omega)\, dB|^2] \;=\; \sum_{j=0}^{m-1}\sum_{i=0}^{m-1} E\left[g(s_j,\omega)g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})(B_{s_{j+1}} - B_{s_j})\right] \tag{2.41}$$

To evaluate each term in the sum, we condition on $\mathcal{F}_{s_k}$ where $s_k = \max(s_i, s_j)$. Then if $i \neq j$,

$$\begin{aligned} E\left[g(s_j,\omega)g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})(B_{s_{j+1}} - B_{s_j})\right] \\ &= E\left[E[g(s_j,\omega)g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})(B_{s_{j+1}} - B_{s_j})|\mathcal{F}_{s_k}]\right] \\ &= E\left[g(s_j,\omega)g(s_i,\omega)E[(B_{s_{i+1}} - B_{s_i})(B_{s_{j+1}} - B_{s_j})|\mathcal{F}_{s_k}]\right] \\ &= 0 \end{aligned} \tag{2.42}$$

since the increments $(B_{s_{i+1}} - B_{s_i})$ and $(B_{s_{j+1}} - B_{s_j})$ are independent. However, if $i = j$, then

$$\begin{aligned} E\left[(g(s_i,\omega))^2(B_{s_{i+1}} - B_{s_i})^2\right] \\ &= E\left[E[(g(s_i,\omega))^2(B_{s_{i+1}} - B_{s_i})^2|\mathcal{F}_{s_i}]\right] \\ &= E\left[(g(s_i,\omega))^2E[(B_{s_{i+1}} - B_{s_i})^2|\mathcal{F}_{s_i}]\right] \\ &= E\left[(g(s_i,\omega))^2(s_{i+1} - s_i)\right] \\ &= E\left[\int_0^T (g(s_i,\omega))^2\, ds\right] \end{aligned} \tag{2.43}$$

This fact is the key that allows us to pass to the limit (2.36). Suppose that g_k is a sequence of step functions converging to f in the sense of (2.35). Then without loss of generality, we may assume that g_k is a Cauchy sequence. Therefore, from property (iii),

$$\begin{aligned} \lim_{k,j\to\infty} E\left[|\int_0^T g_k(s,\omega)dB - \int_0^T g_j(s,\omega)dB|^2\right] &= \lim_{k,j\to\infty} E\left[|\int_0^T (g_k(s,\omega) - g_j(s,\omega))dB|^2\right] \\ &= \lim_{k,j\to\infty} E\left[\int_0^T |g_k(s,\omega) - g_j(s,\omega)|^2 ds\right] \\ &= 0 \end{aligned} \tag{2.44}$$

Thus the random variables $\int_0^T g_k(s,\omega)dB$ are also a Cauchy sequence, so there exists a limit in $L^2(\Omega,\mathcal{F},P)$ (this is a complete metric space).

For property (iv), notice that if g is a step process, then

$$\begin{aligned} E[\int_0^T g(s,\omega)\, dB] &= E\left[\sum_{i=0}^{m-1} g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})\right] \\ &= \sum_{i=0}^{m-1} E\left[g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})\right] \\ &= \sum_{i=0}^{m-1} E\left[E[g(s_i,\omega)(B_{s_{i+1}} - B_{s_i})|\mathcal{F}_{s_i}]\right] \\ &= \sum_{i=0}^{m-1} E\left[g(s_i,\omega)E[(B_{s_{i+1}} - B_{s_i})|\mathcal{F}_{s_i}]\right] \\ &= \sum_{i=0}^{m-1} 0 \end{aligned} \tag{2.45}$$

The equality at (2.45) follows from the fact that $g(s_i)$ is $\mathcal{F}_{s_i}$ measureable. By approximation the same result holds for general $g \in \mathcal{L}^2([0,T])$. That is, this property is preserved under the limit (2.36).

The proof of property (vi) is similar to the proof of properties (iii) and (iv) using the conditioning idea. Property (vii) can be proved using the maximal martingale inequality and the Borel-Cantelli lemma. See the references for details. □

Having defined the Itô integral, can now give a rigorous meaning to the "differential equation with noise" (2.27) and (2.28). For adapted processes $G(t,\omega), F(t,\omega) \in \mathcal{L}^2([0,T])$, we will now use the notation

$$dX_t = G(t,\omega)dt + F(t,\omega)dB_t, \quad t \in [0,T] \tag{2.46}$$

to mean that the process $X_t(\omega)$ satisfies the integral equation

$$X_t(\omega) = X_0(\omega) + \int_0^t G(s,\omega)ds + \int_0^t F(s,\omega)dB_s(\omega) \tag{2.47}$$

for all $t \in [0,T]$. Here $X_0(\omega)$ is some square-integrable random variable. The first integral is a regular Riemann integral if G is continuous. In general, however, it is a Lebesgue integral. The second integral is the Itô integral.

The terms in the equation might depend on X_t itself. That is, we say that a process X_t satisfies the differential equation

$$dX_t = H(X_t,t)dt + K(X_t,t)dB_t \tag{2.48}$$

if $X_t(\omega)$ satisfies the integral equation

$$X_t(\omega) = X_0(\omega) + \int_0^t H(X_t,t)ds + \int_0^t K(X_t,t)dB_s(\omega) \tag{2.49}$$

Theorem 3.2.4 *Suppose that $H(x,t)$ and $K(x,t)$ are continuous in both variables and that*

$$|H(x,t) - H(y,t)| \leq C|x-y|, \;\; x,y \in \mathbb{R}, \;\; t \in [0,T] \tag{2.50}$$

$$|K(x,t) - K(y,t)| \leq C|x-y|, \;\; x,y \in \mathbb{R}, \;\; t \in [0,T] \tag{2.51}$$

for some constant $C > 0$. Then there exists a unique solution $X_t(\omega) \in \mathcal{L}^2([0,T])$ to the stochastic equation

$$\begin{aligned} dX_t &= H(X_t,t)dt + K(X_t,t)dB_t \\ X_0(\omega) &= x_0 \end{aligned} \tag{2.52}$$

The conditions (2.50) and (2.51) say that the coefficients are uniformly Lipschitz continuous in the x variable. This is a natural requirement, as Lipschitz continuity is generally a necessary condition for uniqueness of solutions to ODE's of the form $y'(t) = f(y(t))$. This theorem may be proved by Picard Iteration, as in the case of ODE's. For details, see the references listed above.

3.3 Itô's Formula

In stochastic calculus, Itô's formula plays the role of the chain rule, which is one of the most fundamental ideas in ordinary calculus. Suppose that a differentiable function $x(t)$ satisfies the ode

$$x'(t) = g(x(t), t).$$

The integral form this equation is

$$x(t) = x(0) + \int_0^t g(x(s), s)\, ds.$$

Now the ordinary chain rule tells us that

$$\frac{d}{dt}\phi(x(t)) = \phi'(x(t))x'(t) = \phi'(x(t))g(x(t), t) \tag{3.53}$$

or in integral form,

$$\phi(x(t)) = \phi(x(0)) + \int_0^t \phi'(x(s))g(x(s), s)\, ds \tag{3.54}$$

If ϕ is smooth, then Taylor's theorem tells us that

$$\phi(x(t)) = \phi(x(0)) + \phi'(x(0))(x(t) - x(0)) + \frac{1}{2}\phi''(x(0))(x(t) - x(0))^2 + O(|x(t) - x(0)|^3), \tag{3.55}$$

even if $x(t)$ is not differentiable. If $x(t)$ is differentiable, the term $\phi'(x(0))(x(t) - x(0))$ is $O(t)$, while the next term $\frac{1}{2}\phi''(x(0))(x(t) - x(0))^2$ is $O(t^2)$, smaller than $O(t)$ when t is small. Now, for the purpose of illustration, let us suppose that $x(t) = B_t$ is a Brownian motion (not differentiable). Plugging in to (3.55) we find that

$$\phi(B_t) = \phi(B_0) + \phi'(B_0)(B_t - B_0) + \frac{1}{2}\phi''(B_0)(B_t - B_0)^2 + h.o.t.,$$

The increments of Brownian motion are $N(0, t)$. Therefore, the first term $\phi'(B_0)(B_t - B_0)$ is zero on average, while the second term corresponding to $(B_t - B_0)^2$ contributes $O(t)$ on average, unlike the case when $x(t)$ was differentiable. This suggests that the "chain rule" for $\phi(B_t)$ should be (in integral form):

$$\phi(B_t) = \phi(B_0) + \int_0^t \phi'(s)\, dB_s + \frac{1}{2}\int_0^t \phi''(s)\, ds \tag{3.56}$$

If B_t were actually differentiable, the normal chain rule would imply that

$$\phi(B_t) = \phi(B_0) + \int_0^t \phi'(s)B'(s)\, ds \tag{3.57}$$

So, in (3.56) there is an extra term due to the fact that the increments of B_t are $O(t)$ in means square sense. That is $E[(B_t - B_0)^2] = t$. Roughly speaking, we may say that $(dB_t)^2 \approx dt$. The following theorem makes these ideas more precise.

Theorem 3.3.1 (Itô's formula) *Suppose $\phi(x,t)$ is twice differentiable in x, differentiable in t. Suppose that the process $X(t,\omega)$ satisfies*

$$dX_t = F(t,\omega)\,dt + G(t,\omega)dB_t \tag{3.58}$$

Then the process $Y_t(\omega) = \phi(X_t(\omega),t)$ satisfies

$$dY_t = \left(\phi_x(X_t(\omega),t)F(t,\omega) + \phi_t(X_t(\omega),t) + \frac{1}{2}\phi_{xx}(X_t(\omega),t)G^2(t,\omega)\right)\,dt + \phi_x(X_t(\omega),t)G(t,\omega)dB_t$$

or in integral form,

$$\begin{aligned}\phi(X_t(\omega),t) &= \phi(X_0(\omega),0) + \int_0^t \left(\phi_x(X_s(\omega),s)F(s,\omega) + \phi_t(X_s(\omega),s) + \frac{1}{2}\phi_{xx}(X_s(\omega),s)G^2(s,\omega)\right)\,ds \\ &\quad + \int_0^t \phi_x(X_s(\omega),s)G(s,\omega)dB_s \end{aligned} \tag{3.59}$$

Proof: See references listed above.□

The differential form of Itô's formula may also be written as

$$dY_t = \phi_x(X_t(\omega),t)dX_t + \phi_t(X_t(\omega),t)dt + \frac{1}{2}\phi_{xx}(X_t(\omega),t)G^2(t,\omega)\,dt \tag{3.60}$$

You can remember this formula by considering Taylor's formula:

$$\phi(x+dx,t+dt) \approx \phi(x,t) + \phi_x dx + \phi_t dt + \frac{1}{2}\phi_{xx}(dx)^2 \tag{3.61}$$

then using $dx = Fdt + GdB$ you get Itô's formula by keeping the low order terms, and remembering that $dB \sim (dt)^{1/2}$ and $(GdB)^2 \sim G^2dt$, roughly speaking.

Example 1: Let $X_t = B_t$ and $\phi(x) = x^2$. Then by applying Itô's formula to the process $Y_t = (B_t)^2$ we find that

$$(B_t)^2 = \int_0^t 2B_s\,dB_s + t \tag{3.62}$$

Here $F \equiv 0$ and $G \equiv 1$. Notice that this is different from what ordinary calculus would predict, because if $\phi(t)$ is differentiable with $\phi(0) = 0$, then

$$\int_0^t 2\phi(s)d\phi(s) = \int_0^t 2\phi(s)\phi'(s)\,ds = \int_0^2 \frac{d}{ds}(\phi(s))^2\,ds = (\phi(t))^2 \tag{3.63}$$

Example 2: Let $X_t = B_t$ and $\phi(x) = e^{\alpha x}$. Then by applying Itô's formula to the process $Y_t = e^{\alpha B_t}$ we find that

$$Y_t = 1 + \int_0^t \alpha Y_s\,dB_s + \frac{\alpha^2}{2}\int_0^t Y_s\,ds \tag{3.64}$$

which may be expressed as

$$dY_t = \alpha Y_t\,dB_t + \frac{\alpha^2}{2}Y_t dt \tag{3.65}$$

Similarly, the function $Z_t = e^{\alpha B_t - t\alpha^2/2}$ satisfies

$$dZ_t = \alpha Z_t \, dB_t \tag{3.66}$$

This shows that Z_t is a martingale since

$$Z_t = 1 + \int_0^t \alpha Z_s \, dB_s \tag{3.67}$$

and the Itô integral is a martingale. (Actually one can easily compute directly that Z is a martingale without the help of stochastic calculus. Try this on your own.)

Theorem 3.3.2 (Itô Product Rule) *Suppose that $X_t(\omega)$ and $Y_t(\omega)$ two stochastic processes satisfying*

$$\begin{aligned} dX_t &= F(X_t,t)dt + G(X_t,t)dB_t \\ dY_t &= H(Y_t,t)dt + K(Y_t,t)dB_t \end{aligned} \tag{3.68}$$

Then the process $Z_t(\omega) = X_t(\omega)Y_t(\omega)$ satisfies

$$\begin{aligned} dZ_t &= (F(X_t,t)Y_t + H(Y_t,t)X_t + G(X_t,t)K(Y_t,t))\,dt \\ &\quad +(G(X_t,t)Y_t + K(Y_t,t)X_t)dB_t \\ &= Y_t dX_t + X_t dY_t + G(X_t,t)K(Y_t,t)dt \end{aligned} \tag{3.69}$$

Itô's Formula in Multiple Dimensions

We can also define vector-valued stochastic integrals using a m-dimensional Brownian motion. Suppose that $G(s,\omega)$ is a matrix valued process such that

$$G^{ij}(s,\omega) \in \mathcal{L}^2([0,T]), \quad i = 1,\dots d, \;\; j = 1,\dots,m \tag{3.70}$$

If B_t is a m-dimensional Brownian motion, then

$$X_t = \int_0^t G(s,\omega)\, dB_t \tag{3.71}$$

defines a d-dimensional stochastic process whose components are

$$X_t^{(i)} = \sum_{j=1}^m \int_0^t G^{ij}(s,\omega)\, dB_t^{(j)}, \quad i = 1,\dots,d \tag{3.72}$$

Itô's formula extends to multiple dimensions in the following way.

Theorem 3.3.3 *Suppose that B_t is a m-dimensional Brownian motion and that $X_t(\omega) = (X_t^{(i)}(\omega))_i$ is a d-dimensional stochastic process satisfying*

$$X_t^{(i)}\omega) = X_0^{(i)}(\omega) + \int_0^t F^{(i)}(s,\omega)\, ds + \sum_{j=1}^m \int_0^t G^{ij}(s,\omega)\, dB_t^{(j)}, \tag{3.73}$$

If $\phi(x_1, \ldots, x_d, t)$ is twice-differentiable in the spatial variables, differentiable in t, then the one-dimensional process $Y_t := \phi(X_t(\omega), t)$ satisfies

$$\begin{aligned} dY_t &= [F(t,\omega)\cdot\nabla\phi(X_t(\omega),t) + \phi_t(X_t(\omega),t)]\,dt \\ &\quad + \sum_{j=1}^{m}\sum_{i=1}^{d}\frac{\partial\phi}{\partial x_i}(X_t(\omega),t)\; G^{ij}(t,\omega)dB_t^{(j)} \\ &\quad + \frac{1}{2}\left(\sum_{k=1}^{m}\sum_{i,j=1}^{d}\phi_{x_ix_j}(X_t(\omega),t)G^{(ik)}(t,\omega)G^{(jk)}(t,\omega)\right)\,dt \end{aligned}$$

You can remember the last term by Taylor's formula and the heuristic formula

$$dB_t^{(i)}dB_t^{(j)} \sim dt, \quad \text{if } i=j, \qquad (\sim 0, \text{ otherwise}) \tag{3.74}$$

so that

$$(G^{(kh)}\,dB_t^{(h)})(G^{(qp)}\,dB_t^{(p)}) \sim \delta_{hp}G^{(kh)}G^{(qp)}dt \tag{3.75}$$

Thus, off-diagonal terms ($p \neq h$) vanish in the formula.

3.4 Girsanov Transformation

Suppose that a stochastic process $X_t(\omega)$ satisfies the stochastic differential equation

$$dX_t = a(t,\omega)dt + dB_t \tag{4.76}$$

with $X_0 = 0$, where F is an element of $\mathcal{L}^2([0,T])$. In integral form this is

$$X_t = \int_0^t a(s,\omega)\,ds + B_t \tag{4.77}$$

We may think of this as "Brownian motion shifted by $\int a\,ds$." Indeed, this stochastic process induces a measure on $C([0,T])$ which is absolutely continuous with respect to the Wiener measure.

Example. Let us first consider the simple example when a is a constant so that $X_t = at + B_t$, that is, obviously, not a Brownian motion. The idea of the Girsanov theorem is to re-weigh the paths in $C[0,T]$ in such a way that the process X_t becomes a Brownian motion with respect to the new measure. This is done as follows. Consider the weight

$$Z_t = e^{-aB_t - a^2t/2}$$

and the new measure $dQ_t = Z_t dP_t$, where P_t is the law of the Brownian motion. Let $f(x)$ be a scalar function and let us compute

$$\begin{aligned} &\mathbb{E}_Q(f(X_t)) = \int f(y+at)e^{-ay-a^2t/2}e^{-y^2/(2t)}\frac{dy}{\sqrt{2\pi t}} \\ &= \int f(z)e^{-a(z-at)-a^2t/2-z^2/(2t)+2azt/(2t)-a^2t^2/(2t)}\frac{dz}{\sqrt{2\pi t}} = \int f(z)e^{-z^2/2}\frac{dz}{\sqrt{2\pi t}} = \mathbb{E}(f(B_t)), \end{aligned}$$

where we used the change of variables $z = y + at$. Hence, the law of X_t with respect to the measure Q is the same as the law of the Brownian motion.

In the general case when a is not necessarily a constant define the stochastic process

$$Z_t = e^{-\int_0^t a(t,\omega)\, dW_s - \frac{1}{2}\int_0^t |a(s,\omega)|^2\, ds}$$

Theorem 3.4.1 (Novikov's Theorem) *Suppose that*

$$E\left[e^{\frac{1}{2}\int_0^t |a(s,\omega)|^2\, ds}\right] < \infty \tag{4.78}$$

for each $t \geq 0$. Then $E[Z_t] = 1$ and Z_t is a martingale (with respect to $(\mathcal{F}_t)_{t\geq 0}$).

Let us verify that Z_t is, indeed, a martingale. We write it as $Z_t = e^{Y_t}$, where

$$dY_t = -\frac{a^2(t,\omega)}{2}dt - a(t,\omega)dB_t.$$

Ito's formula implies that

$$dZ_t = e^{Y_t}dY_t + \frac{1}{2}e^{Y_t}(dY_t)^2 = Z_t\left[-\frac{a^2(t,\omega)}{2}dt - a(t,\omega)dB_t + \frac{a^2(t.\omega)}{2}dt\right] = -a(t,\omega)Z_t dB_t,$$

hence

$$Z_t = -\int_0^t a(s,\omega)Z_s dB_s$$

is a martingale.

Using the process Z_t we define a new probability measure Q on $C([0,T])$ given by

$$Q(A) = E[Z_T \mathbb{I}_A(\omega)] \tag{4.79}$$

where $\mathbb{I}_A(\omega)$ denotes the indicator function of the set A. Now we define a new process $\tilde{B}_t(\omega)$ by

$$\tilde{B}_t(\omega) = B_t(\omega) - \int_0^t F(s,\omega)\, ds \tag{4.80}$$

Theorem 3.4.2 *[Girsanov's theorem] The process $X_t(\omega)$ given by (4.76) is a Brownian motion under the measure Q.*

Proof. We start with the following lemma.

Lemma 3.4.1 *Let μ and ν be two probability measures on a probability space $(\Omega, \mathcal{G})$ such that $d\nu(\omega) = f(\omega)d\mu(\omega)$ for some $f \in L^1(d\mu)$. Let X be a random variable such that*

$$\mathbb{E}_\nu(|X|) = \int |X(\omega)|f(\omega)d\mu(\omega) < +\infty,$$

and let $\mathcal{H}$ be a subalgebra of $\mathcal{G}$. Then

$$\mathbb{E}_\mu(fX|\mathcal{H}) = \mathbb{E}_\nu(X|\mathcal{H})\mathbb{E}_\mu(f|\mathcal{H}). \tag{4.81}$$

Proof. Let $S \in \mathcal{H}$ be an $\mathcal{H}$-measurable set, then

$$\int_S \mathbb{E}_\nu(X|\mathcal{H}) f d\mu = \int_S \mathbb{E}_\nu(X|\mathcal{H}) d\nu = \int_S X d\nu = \int_S X f d\mu = \int_S \mathbb{E}_\mu(fX|\mathcal{H}) d\mu.$$

On the other hand, we have, once again, using $\mathcal{H}$-measurability of S:

$$\int_S \mathbb{E}_\nu(X|\mathcal{H}) f d\mu = \mathbb{E}_\mu(\mathbb{E}_\nu(X|\mathcal{H}) f \chi_S) = \mathbb{E}_\mu(\mathbb{E}_\mu(\mathbb{E}_\nu(X|\mathcal{H}) f \chi_S|\mathcal{H})) = \mathbb{E}_\mu(\chi_S \mathbb{E}_\mu(\mathbb{E}_\nu(X|\mathcal{H}) f|\mathcal{H}))$$
$$= \mathbb{E}_\mu(\chi_S \mathbb{E}_\nu(X|\mathcal{H}) \mathbb{E}_\mu(f|\mathcal{H})) = \int_S \mathbb{E}_\nu(X|\mathcal{H}) \mathbb{E}_\mu(f|\mathcal{H}) d\mu,$$

and (4.81) follows.

Now, we prove Theorem 3.4.2. We will assume for simplicity that the function $a(s, \omega)$ is bounded by a deterministic constant and will use Levy's characterization of the Brownian motion. It says that if $Y(t)$ is a continuous stochastic process such that both $Y(t)$ is a martingale and the process $Y^2(t) - t$ is a martingale then $Y(t)$ is a Brownian motion. This is what we verify for the process $X(t)$ given by (4.76). Actually, we will check only that $X(t)$ is a martingale with respect to the measure Q, and leave verifying this for $X^2(t) - t$ to the reader. Let us put $K(t) = Z_t X(t)$, then by Ito's formula:

$$dK_t = Z_t dX_t + X_t dZ_t + (dY_t dZ_t) = Z_t(adt + dB_t) + X_t Z_t(-adB_t) + (dB_t)(-Z_t a dB_t)$$
$$= Z_t(dB_t - X_t a dB_t),$$

hence K_t is a martingale with respect to the measure P. Hence, Lemma 3.4.1 implies that

$$\mathbb{E}_Q(X_t|\mathcal{F}_s) = \frac{\mathbb{E}_P(Z_t X_t|\mathcal{F}_s)}{\mathbb{E}_P(Z_t|\mathcal{F}_s)} = \frac{\mathbb{E}_P(K_t|\mathcal{F}_s)}{Z_s} = \frac{K_s}{Z_s} = X_s,$$

hence X_t is a Q-martingale.

Chapter 4

Second order linear PDEs and the Feynman-Kac Formula

References:

- Karatzas and Shreve, Sections 4.4, 5.7
- Oksendal, Chapter 8
- R. Feynman, Rev. Mod. Phys. Vol. 20 (1948), pp. 367-387.
- M. Kac, Trans. Amer. Math. Soc. Vol. 65 (1949), pp. 1-13.
- M. Kac, Proc. of the 2nd Berkeley Symp. on Math. Stats. and Prob. (1951), pp. 189-215.
- M. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton University Press, 1985. Chapters 1-2.
- R. Bass, *Diffusions and Elliptic Operators*, Springer, 1998.

4.1 The Feynman-Kac Formula

In these notes we study a stochastic representation formula for solutions to certain second-order linear partial differential equations. This representation is called the Feynman-Kac formula, named after the physicist Richard Feynman (1918-1988) and mathematician Mark Kac (1914-1984) (see references above). You may think of the formula as a generalization of the convolution formula for the solution to the heat equation, which represents a solution to the heat equation as an expectation with respect to a Gaussian random variable.

The basic result

Here is the basic idea. In what follows, we will work with solutions to initial value problems and with solutions to terminal value problems. One can switch between these two perspectives through

a simple change of variables: $t \to T - t$. Suppose that $w(x,t) \in C^{2,1}(\mathbb{R} \times [0,\infty))$ solves the initial value problem

$$w_t = \frac{\sigma^2(x)}{2} w_{xx} + b(x) w_x \quad x \in \mathbb{R}, \;\; t > 0 \tag{1.1}$$

with initial data $w(x,0) = f(x)$, which is smooth and compactly supported. We also assume that $b(x)$ and $\sigma(x)$ are Lipschitz continuous and bounded. We also assume that w is bounded. Then for $t > 0$ fixed, the function $u(x,s) = w(x, t-s)$ satisfies the terminal value problem

$$u_s + \frac{\sigma^2(x)}{2} u_{xx} + b(x) u_x = 0, \quad x \in \mathbb{R}, \;\; s < t \tag{1.2}$$

with terminal condition $u(x,t) = f(x)$. Moreover, $u \in C^{2,1}(\mathbb{R} \times (-\infty, t])$. Now let $B_s(\omega)$ be a standard Brownian motion with filtration $(\mathcal{F}_s)_{s \geq 0}$. Suppose that, $X_s(\omega)$ is an $\mathcal{F}_s$-adapted solution to the stochastic ODE

$$dX_s^x = b(X_s^x)\, ds + \sigma(X_s^x) dB_s \tag{1.3}$$

with $X_0^x = x$. The existence and uniqueness of such a solution is guaranteed by our assumptions about b and σ.

Now, a direct application of Ito's formula shows us that,

$$\begin{aligned} u(X_t^x, t) - u(X_0, 0) &= \int_0^t \left(u_s(X_s^x, s) + b(X_s^x) u_x(X_s^x, s) + \frac{\sigma^2(X_s^x)}{2} u_{xx}(X_s^x, s) \right) ds \\ &\quad + \int_0^t \sigma(X_s^x) u_x(X_s^x, s)\, dB_s \\ &= \int_0^t \sigma(X_s^x) u_x(X_s^x, s)\, dB_s \qquad \text{(due to (1.2))} \end{aligned}$$

Therefore, taking the expectation, we find that

$$E\left[u(X_t^x, t)\right] = E\left[u(X_0^x, 0)\right] = u(x, 0), \tag{1.4}$$

since the Itô integral has zero mean. In terms of w, this shows that

$$w(x,t) = u(x,0) = E\left[u(X_t^x, t)\right] = E\left[f(X_t^x)\right] \tag{1.5}$$

In summary, these arguments demonstrate the following:

Theorem 4.1.1 *(i)* **Initial value problem:** *Suppose that $w(x,t) \in C^{2,1}(\mathbb{R} \times [0,\infty))$ is bounded and satisfies*

$$w_t = \frac{\sigma^2(x)}{2} w_{xx} + b(x) w_x \quad x \in \mathbb{R}, \;\; t > 0$$

with initial condition $w(x,0) = f(x) \in C_0^2(\mathbb{R})$. Then $w(x,t)$ is represented by

$$w(x,t) = E\left[f(X_t^x)\right]$$

where $dX_s^x = b(X_s^x)\, ds + \sigma(X_s^x) dB_s$ for $s \geq 0$ and $X_0^x(\omega) = x$.

(ii) **Terminal value problem:** *Suppose that* $u(x,t) \in C^{2,1}(\mathbb{R}\times(-\infty,T])$ *is bounded and satisfies*

$$u_t + \frac{\sigma^2(x)}{2}u_{xx} + b(x)u_x = 0 \quad x \in \mathbb{R}, \;\; t < T$$

with terminal condition $u(x,T) = f(x) \in C_0^2(\mathbb{R})$. *Then* $u(x,t)$ *is represented by*

$$u(x,t) = E\left[f(X^x_{T-t})\right]$$

where $dX^x_s = b(X^x_s)\,ds + \sigma(X^x_s)dB_s$ *for* $s \geq 0$ *and* $X^x_0(\omega) = x$.

Generalizations

To avoid technical difficulties, we have been rather conservative in our assumptions about the initial conditions and the coefficients. In fact, these representations hold under milder conditions on the initial data and the coefficients. Now let us suppose that $u(x,t)$ satisfies the second-order linear PDE

$$u_t + \sum_{i,j=1}^{d} \frac{1}{2}a_{ij}(x,t)u_{x_ix_j} + \sum_{j=1}^{d} b_j(x,t)u_{x_j} + c(x,t)u = 0, \quad x \in \mathbb{R}^d, \;\; t < T \tag{1.6}$$

with terminal condition $u(x,T) = f(x)$ which is continuous (but not necessarily differentiable or bounded). We also assume

- The matrix a_{ij} is given by $a_{ij} = \sum_k \sigma_{ik}\sigma_{kj} = \sigma\sigma^T$ for some matrix $\sigma_{jk}(x,t)$.
- The matrix $a_{ij} = a_{ij}(x,t)$ is uniformly positive definite: $\sum_{ij} a_{ij}\xi_j\xi_i \geq \mu|\xi|^2$ for some constant $\mu > 0$, independent of (x,t).
- Both $\sigma_{ij}(x,t)$ and $b(x,t) = (b_j(x,t))$ are Lipschitz continuous in x, continuous in t, and grow at most linearly in x.
- The funtion $c(x,t)$ is continuous in (x,t) and bounded in x.
- The terminal condition $f(x)$ satisfies the growth condition $|f(x)| \leq Ce^{p|x|^2}$ for some constant $p > 0$ sufficiently small.
- $u(x,t)$ satisfies the growth condition $|u(x,t)| \leq Ce^{p|x|^2}$ for $x \in \mathbb{R}$, $t \in [t_0,T]$ and some constant $p > 0$ sufficiently small.

Suppose that for a given (x,t), the process $X^{x,t}_s(\omega) : [t,T]\times\Omega \to \mathbb{R}^d$ satisfies

$$dX^{x,t}_s = b(X^{x,t}_s, s)\,ds + \sum_j \sigma_{ij}(X^{x,t}_s, s)\,dB^{(j)}_s, \;\; s \in [t,T] \tag{1.7}$$

with $X^{x,t}_t(\omega) = x$. The superscripts indicate that the process $X^{x,t}_s$ starts at the point x at time t. Notice that $X^{x,t}_s$ is a vector, and $b(x,s)$ is also a vector. Then one can prove:

Theorem 4.1.2 *Under the assumptions given above,* $u(x,t)$ *satisfies*

$$u(x,t) = E\left[f(X^{x,t}_T)e^{\int_t^T c(X^{x,t}_s,s)\,ds}\right]. \tag{1.8}$$

Sketch of proof: To prove this statement, one may apply Itô's formula and the product rule to the process defined by

$$H_r(\omega) = u(X_r^{x,t}, r)e^{\int_t^r c(X_s^{x,t},s)\,ds}, \quad r \in [t,T]. \tag{1.9}$$

The fact that the terminal data $f(x)$ may not be smooth or bounded causes some difficulty that may be overcome by using Itô's formula with stopping times. For $n > 0$, let $S_n(\omega)$ be the stopping time $S_n = \inf\{s \geq t | \, |X_s^{x,t}| \geq n\}$. Then we conclude that for $r \in (t,T)$,

$$\begin{aligned} H_{r\wedge S_n} - H_t &= u(X_{r\wedge S_n}^{x,t}, r\wedge S_n)e^{\int_t^{r\wedge S_n} c(X_s^{x,t},s)\,ds} - u(X_t^{x,t},t) \\ &= \int_t^{r\wedge S_n} e^{\int_t^s c(X_\tau^{x,t},\tau)\,d\tau}\left(u_s + \sum_j b_j u_{x_j} + \frac{1}{2}a_{ij}u_{x_ix_j} + c(X_s^{x,t},s)u\right)ds \\ &\quad + \int_t^{r\wedge S_n} e^{\int_t^s c(X_\tau^{x,t},\tau)\,d\tau}\sum_{i,j} u_{x_i}\sigma_{ij}dB_s^{(j)} \\ &= \int_t^{r\wedge S_n} e^{\int_t^s c(X_\tau^{x,t},\tau)\,d\tau}\sum_{i,j} u_{x_i}\sigma_{ij}dB_s^{(j)} \quad \text{(using (1.6))} \end{aligned}$$

Notice that arguments inside the integrals are evaluated at $(X_s^{x,t}, s)$. Taking the expectation as before, we conclude that

$$u(x,t) = E\left[u(X_t^{x,t},t)\right] = E\left[u(X_{r\wedge S_n}^{x,t}, r)e^{\int_t^{r\wedge S_n} c(X_s^{x,t},s)\,ds}\right]. \tag{1.10}$$

Notice that if u itself is not bounded, then the expectation on the right is not obviously finite. This explains our use of the stopping time – the stopping time restricts $X_{r\wedge S_n}^{x,t}$ to a bounded region, over which u must be bounded since u is continuous. The next step is to take $n \to \infty$. Using the growth assumptions on u and the coefficients one can show that as $n \to \infty$, the above expression remains finite since $P(S_n < r) = O(e^{-\alpha n^2})$ as $n \to \infty$. This shows that

$$u(x,t) = E\left[u(X_r^{x,t}, r)e^{\int_t^r c(X_s^{x,t},s)\,ds}\right] \tag{1.11}$$

Then we let $r \to T$. If we knew that u were sufficiently smooth and bounded at $r = T$, then we could apply Itô's formula with $r = T$ in the above formula. This was our approach in the first section, since we assumed the initial (or terminal) data was C^2. In general, however, this is not the case. Nevertheless, one may use the dominated convergence theorem to show that as $r \to T$,

$$\lim_{r\to T} E\left[u(X_r^x, r)e^{\int_t^r c(X_s^x,s)\,ds}\right] = E\left[f(X_T^x)e^{\int_t^T c(X_s^x,s)\,ds}\right] \tag{1.12}$$

even when f is merely continuous and satisfies a growth condition (see Karatzas and Shreve for more details). □

Next we formulate a similar result for the initial value problem. Suppose that $w(x,t)$ satisfies

$$w_t = \sum_{i,j=1}^d \frac{1}{2}a_{ij}(x,t)u_{x_ix_j} + \sum_{j=1}^d b_j(x,t)u_{x_j} + c(x,t)u, \quad x \in \mathbb{R}^d, \;\; t > 0 \tag{1.13}$$

with initial condition $w(x,0) = f(x)$. Then the function $\tilde{u}(x,-t) := w(x,t)$ satisfies (1.7) with $T = 0$, and coefficients given by $\tilde{a}_{ij}(x,t) = a_{ij}(x,-t)$, $\tilde{b}(x,t) = b(x,-t)$, $\tilde{c}(x,t) = c(x,-t)$. For given (x,t) let $X_s^{x,t}(\omega)$ satisfy

$$dX_s^{x,t} = b(X_s^{x,t}, t-s)\,ds + \sum_j \sigma_{ij}(X_s^{x,t}, t-s)\,dB_s^{(j)}, \quad s \in [0,t] \tag{1.14}$$

Then the analysis above shows that

$$w(x,t) = E\left[f(X_t^{x,t}) e^{\int_0^t c(X_s^{x,t}, t-s)\,ds}\right]. \tag{1.15}$$

In particular, if $c \equiv 0$, then

$$w(x,t) = E\left[f(X_t^{x,t})\right]. \tag{1.16}$$

These are very elegant formulas which have a natural physical interpretation. Here is how I think about it. The equation (1.13) models the diffusion, transport, and reaction of a scalar quantity $w(x,t)$. The vector field b is the "drift" or wind. The matrix a_{ij} determines the rates of diffusion in a given direction. The process $X_t^{x,t}$ may be thought of as the paths of particles diffusing in this velocity field. The function $c(x,t)$ represents a reaction rate. So, imagine hot, reactive particles being carried in the wind. Now, consider the formula (1.16) for the case $c \equiv 0$ (no reaction). What determines the temperature at a point (x,t)? The temperature at this point is determined by which particles arrive at point x at time t and how hot those particles were initially. The quantity $f(X_t^{x,t})$ represents the initial "temperature" evaluated at the "end" of the path $f(X_t^{x,t})$. Notice that $X_s^{x,t}$ actually runs backwards in the time-frame associated with the PDE. Roughly speaking, $f(X_t^{x,t})$ tells us what information propagates to the point x at time t. The paths are random; formula (1.16) says that the solution is determined by the expectation over all such particles. In the case that $c \neq 0$, formula (1.15) tells us that the reaction heats up each particle along its trajectory, increasing (or decreasing) its temperature by a factor of $e^{\int_0^t c(X_s^{x,t}, t-s)\,ds}$. Notice that when a_{ij}, b, and c are independent of t, we can replace $t-s$ in the above expressions with s.

4.1.1 Poisson's equation

Here we use Itô's formula to derive a representation for solutions to **Poisson's equation**. Suppose that $w(x)$ is C^2 and bounded, and satisfies

$$\sum_{i,j} \frac{1}{2} a_{ij}(x) w_{x_i x_j} + \sum_j b_j(x) w_{x_j} - c(x) w = f(x), \quad x \in \mathbb{R}^d \tag{1.17}$$

with $c(x) \geq c_0 > 0$ for some constant $c_0 > 0$. As before, we assume $a_{ij} = \sigma\sigma^T$ is uniformly positive definite, and that a, b, and c satsify the continuity criteria given earlier.

Theorem 4.1.3 *Suppose that $X_t^x(\omega)$ solves the stochastic differential equation*

$$dX_t^x = b(X_t^x)\,dt + \sum_j \sigma_{ij}(X_t^x)\,dB_t^{(j)}, \quad t \geq 0$$

with $X_0^x(\omega) = x \in \mathbb{R}^d$, almost surely. The solution $w(x)$ is represented by

$$w(x) = E\left[\int_0^\infty e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds\right]. \tag{1.18}$$

Proof: Now apply Itô's formula and the product rule to the process

$$H_t(\omega) = e^{-\int_0^t c(X_s)ds} w(X_t^x). \tag{1.19}$$

We compute:

$$\begin{aligned} H_t - H_0 &= w(X_t^x)e^{-\int_0^t c(X_s^x)\,ds} - w(X_0^x) \\ &= \int_0^t e^{-\int_0^s c(X_\tau^x)\,d\tau}\left(\sum_j b_j w_{x_j} + \frac{1}{2}\sum_{i,j} a_{ij} w_{x_i x_j} - c(X_s^x)w\right) ds \\ &\quad + \int_0^t e^{-\int_0^s c(X_\tau^x)\,d\tau}\sum_{i,j} w_{x_i}\sigma_{ij} dB_s^{(j)} \\ &= \int_0^t e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds + \int_0^t e^{-\int_0^s c(X_\tau^x)\,d\tau}\sum_{i,j} w_{x_i}\sigma_{ij} dB_s^{(j)} \end{aligned} \tag{1.20}$$

Now we take the expectation of both sides and let $t \to \infty$. Due to the lower bound on $c(x)$,

$$\lim_{t\to\infty}\left|E\left[w(X_t^x)e^{-\int_0^t c(X_s^x)\,ds}\right]\right| \le \lim_{t\to\infty} e^{-c_0 t}\|w\|_\infty = 0 \tag{1.21}$$

Therefore,

$$w(x) = -E\int_0^\infty e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds \tag{1.22}$$

□

4.2 Boundary Value Problems

So far we have considered solutions to partial differential equations posed in the whole space $x \in \mathbb{R}^d$. Itô's formula also leads to representation formulas for solutions to PDE's posed in a bounded domain with appropriate boundary conditions. We consider two types of problems: boundary value problems for elliptic equations and initial value/terminal value problems for parabolic equations.

Boundary value problems

Suppose that $D \subset \mathbb{R}^d$ is a smooth, bounded domain. Suppose that $w(x) \in C^2(\bar{D})$ (thus it is bounded) and satisfies

$$\sum_{i,j}\frac{1}{2}a_{ij}(x)w_{x_i x_j} + \sum_j b_j(x)w_{x_j} - c(x)w = f(x), \quad x \in D \tag{2.23}$$

and boundary condition $w(x) = g(x)$ for $x \in \partial D$. Here we suppose that $c(x) \ge 0$. How can we represent the solution? If $X_t^x(\omega)$ solves the stochastic differential equation

$$dX_t^x = b(X_t^x)\,dt + \sum_j \sigma_{ij}(X_t^x)\,dB_t^{(j)}, \quad t \ge 0$$

with $X_0^x(\omega) = x \in D$, then the trajectories will travel outside the set D, where w is not defined. To overcome this difficulty, we define the stopping time $\gamma_D(\omega) = \inf\{t \mid X_t^x \in \mathbb{R}^d \setminus D\}$. This is the first hitting time to the boundary ∂D. Then define the process

$$H_t(\omega) = e^{-\int_0^{t\wedge\gamma_D} c(X_s)ds} w(X^x_{t\wedge\gamma_D}). \tag{2.24}$$

(Recall the notation $t \wedge \gamma_D := \min(t, \gamma_D)$.) Itô's formula and the product rule then imply that

$$\begin{aligned} H_t - H_0 &= w(X^x_{t\wedge\gamma_D})e^{-\int_0^{t\wedge\gamma_D} c(X_s^x)\,ds} - w(X_0^x) \\ &= \int_0^{t\wedge\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau}\left(\sum_j b_j w_{x_j} + \frac{1}{2}\sum_{i,j} a_{ij} w_{x_i x_j} - c(X_s^x)w\right) ds \\ &\quad + \int_0^{t\wedge\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau}\sum_{i,j} w_{x_i}\sigma_{ij} dB_s^{(j)} \\ &= \int_0^{t\wedge\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds + \int_0^{t\wedge\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau}\sum_{i,j} w_{x_i}\sigma_{ij} dB_s^{(j)} \end{aligned} \tag{2.25}$$

As before, we now take the expectation of both sides and let $t \to \infty$. Let us suppose that $E[\gamma_D(\omega)] < \infty$. Therefore, $\lim_{t\to\infty} \gamma_D(\omega) \wedge t = \gamma_D(\omega)$, almost surely. Consequently, the fact that w is bounded and that $c \geq 0$, we may use the dominated convergence theorem to show that

$$\lim_{t\to\infty} E\left[w(X^x_{t\wedge\gamma_D})e^{-\int_0^{t\wedge\gamma_D} c(X_s^x)\,ds}\right] = E\left[w(X^x_{\gamma_D})e^{-\int_0^{\gamma_D} c(X_s^x)\,ds}\right] = E\left[g(X^x_{\gamma_D})e^{-\int_0^{\gamma_D} c(X_s^x)\,ds}\right], \tag{2.26}$$

Similarly, using the fact that f is bounded and $E[\gamma_D] < \infty$, we may use the dominated convergence theorem to show that

$$\lim_{t\to\infty} E\left[\int_0^{t\wedge\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds\right] = E\left[\int_0^{\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds\right] \tag{2.27}$$

Therefore, taking $t \to \infty$, we obtain a representation for $w(x)$:

$$w(x) = E\left[g(X^x_{\gamma_D})e^{-\int_0^{\gamma_D} c(X_s^x)\,ds}\right] - E\left[\int_0^{\gamma_D} e^{-\int_0^s c(X_\tau^x)\,d\tau} f(X_s^x)\,ds\right]. \tag{2.28}$$

Notice that with the stronger assumtion $c(x) \geq c_0 > 0$, we could lift the condition that $E[\gamma_D] < \infty$, which was used in the application of the dominated convergence theorem to obtain (2.27). We could also lift the restriction that $w \in C^2(\bar{D})$, and require only that $w \in C^2(D) \cap C(\bar{D})$ (thus, the second derivatives might blow up at that boundary). To handle this case, stop the process when it is distance ϵ from the boundary. Then let $\epsilon \to 0$.

Example 1: In particular, this representation shows that if $w(x)$ solves $\Delta w = 0$ in D with $w(x) = g(x)$ for $x \in \partial D$, then

$$w(x) = E\left[g(x + \sqrt{2}B_{\gamma_D})\right] \tag{2.29}$$

The quantity $g(x+\sqrt{2}B_{\gamma_D})$ is the boundary function evaluated at the point where the process first hits the boundary. The solution to the PDE is the expectation of these values.

Initial value problems

Suppose that $D \subset \mathbb{R}^d$ is a smooth bounded domain. Let $D_T = D \times (0, T]$ denote the parabolic cylinder. Suppose that $w(x,t) \in C^{2,1}(D_T) \cap C(\bar{D}_T)$ satisfies the initial value problem

$$w_t = \sum_{i,j} \frac{1}{2} a_{ij}(x,t) w_{x_i x_j} + \sum_j b_j(x,t) w_{x_j} + c(x,t)w, \quad x \in D, \;\; t > 0$$
$$w(x,0) = f(x) \quad x \in D$$
$$w(x,t) = g(x,t) \quad x \in \partial D, t \geq 0$$

Here we assume $c(x,t)$ is bounded and continuous. For given $(x,t) \in D_T$, let $X_s^{x,t}(\omega)$ satisfy

$$dX_s^{x,t} = b(X_s^{x,t}, t-s)\, ds + \sum_j \sigma_{ij}(X_s^{x,t}, t-s)\, dB_s^{(j)}, \quad s \in [0,t] \tag{2.30}$$

Define the stopping time $\gamma_D^{x,t} = \inf\{s \geq 0 \,|\, X_s^{x,t} \in \mathbb{R} \setminus D\}$. This is the first time the process hits the boundary of the set D. Then define $\gamma^{x,t} = \gamma_D^{x,t} \wedge t$ This is also a stopping time, and it represents the time at which the process $(X_s^{x,t}, t-s)$ hits the parabolic boundary $(D \times \{0\}) \cup (\partial D \times [0,T])$, which is the boundary of the set D_T. For convient notation, let us define the function

$$k(x,t) = \begin{cases} f(x), & \text{if } t = 0, \; x \in \bar{D} \\ g(x,t), & \text{if } t > 0, \; x \in \partial D \end{cases} \tag{2.31}$$

This function is equal to $f(x)$ at the base of the parabolic boundary, and it is equal to $g(x,t)$ on the sides of the parabolic boundary.

Theorem 4.2.1 *Under the above assumptions, $w(x,t)$ satisfies*

$$w(x,t) = E\left[k(X_{\gamma^{x,t}}^{x,t}, \gamma^{x,t}) e^{\int_0^{\gamma^{x,t}} c(X_s^{x,t}, t-s)\, ds} \right] \tag{2.32}$$

Proof: I leave this as an exercise. It may be proved as in the other cases. □

4.3 Transition Densities

Consider the vector-valued stochastic process defined by

$$dX_t = b(X_t)\, dt + \sigma^{ij}(X_t)\, dW_t^j \quad \text{for } \; t > 0, \qquad X_0(\omega) = x. \tag{3.33}$$

Suppose that $a_{ij} = \sigma\sigma^T$ is uniformly positive. Suppose also that a and b satisfy the continuity conditions described previously. Because of the Markov property of Brownian motion, one can show that X_t is a Markov process satisfying

$$P(X_t \in A \,|\, \mathcal{F}_s) = P(X_t \in A \,|\, X_s), \quad \forall\, s \in [0,t). \tag{3.34}$$

Suppose that X_t has a smooth transition density $p(x,s;y,t)$. This means that

$$P(X_t \in A \,|\, X_s = x) = \int_A p(x,s;y,t)\, dy \tag{3.35}$$

and

$$E\left[f(X_t) \mid \ X_s = x\right] = \int_{\mathbb{R}^d} f(y)p(x,s;y,t)\,dy \tag{3.36}$$

for suitable functions f. What equation does $p(x,s;y,t)$ satisfy?

Here is a formal computation that can be made rigorous under suitable smoothness and growth assumptions on the coefficients b and σ^{ij}. If $f(x)$ is smooth and compactly supported, then Itô's formula tells us that

$$f(X_t) - f(X_s) = \int_s^t \mathcal{A}f(X_r)\,dr + \int_s^t \sum_{ij} \frac{\partial f}{\partial x_i}(X_r)\sigma^{ij}(X_r)dW_r^j \tag{3.37}$$

where $\mathcal{A}$ denotes the differential operator

$$\mathcal{A}f(y) := \frac{1}{2}\sum_{ij} a_{ij}(y)f_{y_iy_j} + b(y)\cdot\nabla_y f$$

Conditioning on the event $X_s = x$ and taking the expectation, we obtain

$$E[f(X_t) \mid X_s = x] - E[f(X_s) \mid X_s = x] = \int_s^t E[\mathcal{A}f(X_r) \mid X_s = x]\,dr.$$

Now using the definition of the transition density, we may write this expression as

$$\int_{\mathbb{R}^d} f(y)p(x,s;y,t)\,dy - f(x) = \int_s^t \int_{\mathbb{R}^d} (\mathcal{A}f(y))p(x,s;y;r)\,dy\,dr \tag{3.38}$$

Formally differentiating both sides with respect to t, we obtain the equation

$$\int_{\mathbb{R}^d} f(y)p_t(x,s;y,t)\,dy = \int_{\mathbb{R}^d} (\mathcal{A}f(y))p(x,s;y;t)\,dy \tag{3.39}$$

Now on the right hand side, integrate by parts:

$$\begin{aligned}
\int_{\mathbb{R}^d} (\mathcal{A}f(y))p(x,s;y;t)\,dy &= \int_{\mathbb{R}^d} (\frac{1}{2}\sum_{ij} a_{ij}(x)f_{y_iy_j} + b(x)\cdot\nabla_y f)p(x,s;y;t)\,dy \\
&= \int_{\mathbb{R}^d} f(y)\left(\frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial_{y_i}\partial_{y_j}}\left(a_{ij}(x)p(x,s;y;t)\right) - \nabla_y\cdot\left(b(y)p(x,s;y;t)\right)\right)dy \\
&= \int_{\mathbb{R}^d} f(y)\left(\mathcal{A}_y^* p(x,s;y,t)\right)\,dy
\end{aligned} \tag{3.40}$$

where $\mathcal{A}_y^*$ is the **adjoint operator** defined by

$$\mathcal{A}_y^* g(y) := \frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial_{y_i}\partial_{y_j}}\left(a_{ij}(y)g(y)\right) - \nabla_y\cdot\left(b(y)g(y)\right) \tag{3.41}$$

In the integration by parts step, the boundary terms vanish since f has compact support. Therefore, $p(x,s;y,t)$ should satisfy

$$\int_{\mathbb{R}^d} f(y)\left(p_t(x,s;y,t) - \mathcal{A}_y^* p(x,s;y,t)\right)\,dy = 0 \tag{3.42}$$

Since $f(y)$ is chosen arbitrarily, and since we assume p to be sufficiently smooth, this implies that for each fixed x and s, the function $u(y,t) := p(x,s;y,t)$ satisfies $u_t = \mathcal{A}_y^* u$. That is,

$$\frac{\partial}{\partial t} p(x,s;y,t) = \mathcal{A}_y^* p(x,s;y,t). \tag{3.43}$$

As $t \searrow s$, $p(x,s;y,t)$ as a function of y converges to a delta distribution centered at $y = x$. This equation (3.43) is often called the **Kolmogorov forward equation** for the transition density $p(x,s;y,t)$. The term "forward" is applied since it describes the forward evolution of the probability density for X_t.

For fixed y and t, the function $u(x,s) = p(x,s;y,t)$ satisfies a different equation. To derive this equation, suppose that f is again smooth and compactly supported. We have already shown that the solution to the terminal value problem

$$w_s + \mathcal{A}_x w = 0, \quad s < t, \;\; x \in \mathbb{R}^d \tag{3.44}$$

with terminal data $w(x,t) = f(x)$ has the representation

$$w(x,s) = E\left[f(X_t) \mid X_s = x\right] = \int_{\mathbb{R}^d} f(y) p(x,s;y,t)\, dy \tag{3.45}$$

Formally differentiating the integral expression with respect to s and x and using (3.44) we find that

$$\int_{\mathbb{R}^d} f(y)(p_s(x,s;y,t)\, dy + \mathcal{A}_x p(x,s;y,t))\, dy = 0 \tag{3.46}$$

Since f was arbitrarily chosen this implies that for each y, $p_s(x,s;y,t) + \mathcal{A}_x p(x,s;y,t) = 0$. Since x and s were also arbitrarily chosen, this suggests that for each y and t fixed,

$$\frac{\partial}{\partial s} p(x,s;y,t) + \mathcal{A}_x p(x,s;y,t) = 0. \tag{3.47}$$

Since the coefficients defining the process X_t are independent of t, the transition density is a function of $t - s$:

$$p(x,s;y,t) = \rho(x,y,t-s) \tag{3.48}$$

for some function $\rho(x,y,r)$. Then (3.47) shows that for fixed y, $\rho(x,y,t)$ satisfies

$$\frac{\partial}{\partial t} \rho(x,y,t) = \mathcal{A}_x \rho(x,y,t) \tag{3.49}$$

This equation is often called the **Kolmogorov backward equation**.

For rigorous proofs of these results, see the following references:

- A. Friedman, *Stochastic Differential Equations and Applications*, Vol. 1, Academic Press, 1975, Section 6.4, 6.5, especially Thm. 4.7, Thm 5.4, and related discussion.
- A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, 1964., Chapter 9, on construction of fundamental solutions by the parametrix method. See also chapter 1, section 8, Thm. 15.

Chapter 5

The Heat Equation, Part II

References:

- W. Strauss, 4.1, 4.2, Chapter 5.
- L.C. Evans, Section 7.1.

When we studied the connection between stochastic calculus and second-order linear PDE, we learned that solutions to the heat equation in a bounded domain could be represented in terms of Brownian motion and stopping times of Brownian motion. For example, suppose that $D \subset \mathbb{R}^d$ is a smooth bounded domain and that $u(x,t)$ solves the initial value problem

$$\begin{aligned} u_t = \Delta u \quad x \in D,\ t > 0 \qquad (0.1)\\ u(x,0) = \phi(x), \quad x \in D \\ u(x,t) = 0, \quad x \in \partial D,\ \ t > 0 \qquad (0.2)\end{aligned}$$

The solution is represented by

$$u(x,t) = E\left[\phi(X_t^x)\mathbb{I}_{\gamma^x > t}(\omega)\right] \qquad (0.3)$$

where $X_t^x = x + \sqrt{2}B_t$ and γ^x is the first time X_t^x hits the boundary. In these notes we describe an analytical construction of solutions based on the idea of superposition: the sum of two solutions to the heat equation is also a solution. As you will see, this idea gives us a method for constructing and approximating solutions to a wide variety of equations, not just the simple heat equation.

5.1 Separation of Variables and Eigenfunction Expansion

Consider the homogeneous boundary value problem with homogeneous Dirichlet boundary condition

$$\begin{aligned} u_t = \Delta u, \quad x \in D,\ t > 0 \\ u(x,0) = \phi(x), \quad x \in D \qquad (1.4)\\ u(x,t) = 0, \quad x \in \partial D,\ t > 0 \end{aligned}$$

The idea behind the method of separation of variables is to write the solution as a **superposition** of functions of the form $v(x)\theta(t)$:

$$u(x,t) = \sum_{n=1} v_n(x)\theta_n(t) \tag{1.5}$$

where each term in the series is a solution to the heat equation and boundary condition.

Let us first try to find a function of the form $w(x,t) = v(x)\theta(t)$ which satisfies both the PDE and the boundary condition. The functions v and θ are unknown. Plugging $v(x)\theta(t)$ into the PDE we find:

$$\theta'(t)v(x) = \theta(t)\Delta v \tag{1.6}$$

Therefore,

$$-\frac{\theta'(t)}{\theta(t)} = -\frac{\Delta v(x)}{v(x)} \tag{1.7}$$

Since the left side depends only on x, and the right side depends only on t, the the two sides must be equal to some constant which we will call λ. Therefore, θ and v must satisfy

$$\theta'(t) = -\lambda\theta(t) \qquad \text{and} \qquad -\Delta v = \lambda v \tag{1.8}$$

The solution must also satisfy the boundary conidition: $v(x)\theta(t) = 0$ for all $x \in \partial D$. Therefore, either $v(x) = 0$ for all $x \in \partial D$, or $\theta(t) \equiv 0$. The latter possibility cannot hold if w is to be nontrivial. So, we find that v must satisfy the boundary value problem

$$\begin{aligned} -\Delta v(x) &= \lambda v(x), \quad x \in D \\ v(x) &= 0, \quad x \in \partial D \end{aligned} \tag{1.9}$$

and θ must satisfy the linear ODE $\theta'(t) = -\lambda\theta(t)$. We now have learned that if we were able to solve (1.9) for some constant λ, then we could solve the ODE for $\theta(t)$, and the function $w(x,t) = v(x)\theta(t)$ would satisfy

$$\begin{aligned} w_t &= \Delta w, \quad x \in D \\ w(x,t) &= 0, \quad x \in \partial D. \end{aligned}$$

At $t = 0$, $w(x,0) = v(x)\theta(0)$, and $w_t(x,0) = v(x)\theta'(0)$.

Now suppose that for each integer $n \geq 0$, $v_n(x)$ satisfies

$$\begin{aligned} -\Delta v_n(x) &= \lambda_n v_n(x), \quad x \in D \\ v_n(x) &= 0, \quad x \in \partial D \end{aligned} \tag{1.10}$$

for some constants $\{\lambda_n\}_n$, and $\theta_n(t)$ satisfies

$$\theta_n'(t) = -\lambda_n\theta_n(t) \tag{1.11}$$

Then, since the equation is linear, we should hope that the sum

$$u(x,t) = \sum_{n} v_n(x)\theta_n(t) \tag{1.12}$$

satisfies

$$u_t = \Delta u, \quad x \in D$$
$$u(x,t) = 0, \quad x \in \partial D$$

with initial conditions

$$u(x,0) = \sum_n v(x)\theta(0). \tag{1.13}$$

If it happens that $\phi(x) = \sum_n v(x)\theta(0)$, then we have solved the initial value problem (1.4)!

General Strategy

This analysis suggests the following strategy for solving the initial value problem:

1. Find a set of functions $v_n(x)$ and constants λ_n satisfying (1.10).
2. Represent the initial data as a linear combination of the functions $v_n(x)$ by finding numbers b_n such that
$$\phi(x) = \sum_n b_n v_n(x) \tag{1.14}$$
3. For each n, solve the ODE $\theta_n'(t) = -\lambda_n \theta_n(t)$ with initial data $\theta_n(0) = b_n$.
4. Construct the solution: $u(x,t) = \sum_n v_n(x)\theta_n(t)$.

This raises several mathematical issues. First, when can we solve (1.10)? Second, supposing that we have a large set of solutions $\{v_n(x), \lambda_n\}$ to (1.10), is it possible to write ϕ as a linear combination of the functions $v_n(x)$? That is, under what conditions might we find constants $\{b_n\}$ such that

$$\phi(x) = \sum_n b_n v_n(x) \tag{1.15}$$

also hold? These sums may be infinite series, so we must take care to make sense of the limit and the mode of convergence. Even if ϕ has this representation, it is not obvious that the series $\sum_n v_n(x)\theta_n(t)$ will solve the equation in the classical sense. We might not be able to differentiate the series term by term.

It turns out that this program *is* possible in many practical scenarios. The functions $v_n(x)$ and the constants λ_n which satisfy (1.10) are called **eigenfunctions** and **eigenvalues** of the operator $(-\Delta)$. The eigenvalues are discrete, and there are infinitely many of them. Moreover, there is a large class of functions $\phi(x)$ that can be represented as a linear combination of the eigenfunctions, as expressed in (1.15). In general, there is not an explicit formula for the eigenfunctions $v_n(x)$ or the eigenvalues λ_n, but for certain ideal domains like rectangles and spheres, $v_n(x)$ and λ_n can be computed explicitly.

One may think of the functions $v_n(x)$ as resonant modes which are determined by the domain and the boundary condition. The technique of separation of variables identifies the solution with a superposition of these modes. This idea originates from the work of Joseph Fourier on heat conduction and has very far-reaching applications in mathematics, not just in the solution of the heat equation.

Heat equation in a bounded interval

Here we demonstrate the strategy in the case that the domain $D = [0, L]$ is an interval in $\mathbb{R}$. This is one case where we can compute the functions $v_n(x)$ explicitly. The PDE (1.10) is now an ODE

$$\begin{aligned} -v_n''(x) &= \lambda_n v_n(x), \quad x \in [0, L] \\ v_n(0) &= v_n(L) = 0. \end{aligned} \tag{1.16}$$

Using the boundary conditions, we see for any integer $n > 0$ there is a solution of the form

$$v_n(x) = \sin(\frac{\pi n x}{L}) \tag{1.17}$$

with $\lambda_n = \left(\frac{\pi n}{L}\right)^2 > 0$. Now the condition (1.15) may be written as:

$$\phi(x) = \sum_{n=1}^{\infty} b_n \sin(\frac{\pi n x}{L}). \tag{1.18}$$

For each n the solution to the ODE $\theta'(t) = -\lambda_n \theta_n(t)$ with initial data $\theta_n(0) = b_n$ is:

$$\theta_n(t) \quad = \quad b_n e^{-\lambda_n t}$$

Therefore, we expect that the infinite series

$$u(x,t) = \sum_{n=1}^{\infty} \sin(\frac{\pi n x}{L}) \theta_n(t) = \sum_{n=1}^{\infty} b_n e^{-(\frac{n\pi}{L})^2 t} \sin(\frac{\pi n x}{L}) \tag{1.19}$$

solves the heat equation. So for each t, the solution is a linear combination of sine functions. Notice that the coefficient for each mode, decreases exponentially fast in t, with a rate that increases with the index n.

Now, we need to understand the meaning of the infinite series (1.19). Does the series converge? Does it define a differentiable function? In what sense does $u(x,t)$ satisfy the initial condition? Before we can address these questions, let's review some important properties of trigonometric series.

5.2 Fourier Series

The preceding analysis in the case that D is an interval leads us to the general question: when can a function be represented as a linear combination of trigonometric functions? A **Fourier Series** is an infinite series of the form

$$\phi(x) = \frac{1}{2} A_0 + \sum_{n=1}^{\infty} \left(A_n \cos(\frac{n\pi x}{L}) + B_n \sin(\frac{n\pi x}{L}) \right). \tag{2.20}$$

Assuming the series converges, the function defined by the series is periodic on the interval $[-L, L]$, but it may not be continuous. The coefficients $\{A_n\}_n$, $\{B_n\}_n$ are called the **Fourier coefficients** of the function ϕ. We will consider the following questions:

- Which functions ϕ can be represented in this way?
- How can one compute the coefficients of the series for such a function $\phi(x)$?
- In what sense does the series converge to a function $\phi(x)$?

Important Note: To avoid confusion later, we turn our attention now to the interval $D = [-L, L]$. Notice that (1.18) has the same form as (2.20) with $A_0 = 0$, $A_n = 0$, and $B_n = b_n$.

Orthogonality

For $D = [-L, L]$, we consider the function space

$$L^2(D) = \left\{ f : D \to \mathbb{R} \mid \int_D |f(x)|^2 \, dx < \infty \right\}. \tag{2.21}$$

This is an infinite dimensional vector space (called a Hilbert space) with **inner product**

$$(g, f)_2 = \int_D g(x) f(x) \, dx. \tag{2.22}$$

We say that two functions f and g are **orthogonal** in the space $L^2(D)$ if $(f, g)_2 = 0$. The orthogonality of f and g is analogous to orthogonality of two vectors in R^d. The inner product extends the notion of a dot product in $\mathbb{R}^d$ to infinite dimensional spaces. The **norm** of a function $f \in L^2$ is given by $\|f\| = (f, f)_2^{1/2}$. Recall that in $\mathbb{R}^d$, the Euclidean norm of a vector is given by $\|r\| = (r \cdot r)^{1/2}$. So, the norm $\|f\|$ defined on $L^2(D)$ generalizes the idea of a norm in $\mathbb{R}^d$.

It is not hard to see that for $D = [-L, L]$,

$$\left(\cos(\frac{n\pi x}{L}), \sin(\frac{m\pi x}{L})\right)_2 = \int_{-L}^{L} \cos(\frac{n\pi x}{L}) \sin(\frac{m\pi x}{L}) \, dx = 0 \tag{2.23}$$

for all integers n and m. Also, if $n \neq m$ (n and m are integers),

$$\begin{aligned} \left(\sin(\frac{n\pi x}{L}), \sin(\frac{m\pi x}{L})\right)_2 &= 0 \\ \left(\cos(\frac{n\pi x}{L}), \cos(\frac{m\pi x}{L})\right)_2 &= 0. \end{aligned} \tag{2.24}$$

In the case $n = m$, however,

$$\begin{aligned} \left(\sin(\frac{n\pi x}{L}), \sin(\frac{n\pi x}{L})\right)_2 &= L \\ \left(\cos(\frac{n\pi x}{L}), \cos(\frac{n\pi x}{L})\right)_2 &= L \end{aligned} \tag{2.25}$$

This shows that the terms appearing in the Fourier series are mutually orthogonal.

Completeness

Here is a truly amazing fact:

Theorem 5.2.1 *The set of functions* $\{1\}\cup\{\sin(n\pi x/L)\}_n\cup\{\cos(n\pi_x/L)\}_n$ *is complete in* $L^2([-L,L])$.

This means that the set of functions $\{1\} \cup \{\sin(n\pi x/L)\}_n \cup \{\cos(n\pi_x/L)\}_n$ forms a **basis** for the infinite dimensional space $L^2([-L,L])$ just as a set of d mutually orthogonal vectors in $\mathbb{R}^d$ forms a basis for $\mathbb{R}^d$. These functions are mutually orthogonal, and the span of these functions is **dense** in the space L^2. In other words, given a function $\phi \in L^2([-L,L])$, we can find finite linear combinations of the basis functions that approximate ϕ with arbitrary accuracy (measured in norm $\|\cdot\|_2$). Again, given a $\phi \in L^2$ and any $\epsilon > 0$, one can find coefficients A_n and B_n and N sufficiently large such that

$$\|\phi(x) - \left(\frac{1}{2}A_0 + \sum_{n=1}^{N}\left(A_n\cos(\frac{n\pi x}{L}) + B_n\sin(\frac{n\pi x}{L})\right)\right)\|_2 \leq \epsilon. \tag{2.26}$$

Therefore, for a function $\phi(x) \in L^2(D)$, the coefficients of the Fourier series (which are uniquely determined) give us a representation of ϕ in this infinite-dimensional basis. One might think of the Fourier coefficients as the "coordinates" of the function ϕ in this infinite dimensional space. It certainly is not obvious that any function L^2 can be expressed as a sum of sines and cosines.

How can we compute the coefficients A_n and B_n? Suppose there only finitely many terms in the sum (2.20):

$$\phi(x) = \frac{1}{2}A_0 + \sum_{n=1}^{N}\left(A_n\cos(\frac{n\pi x}{L}) + B_n\sin(\frac{n\pi x}{L})\right) \tag{2.27}$$

The orthogonality property allows us to compute:

$$\begin{aligned}\left(\phi(x), \sin(\frac{k\pi x}{L})\right)_2 &= \int_{-L}^{L}\phi(x)\sin(\frac{k\pi x}{L})\,dx \\ &= \sum_{n=1}^{N}A_n\left(\cos(\frac{n\pi x}{L}), \sin(\frac{k\pi x}{L})\right)_2 + \sum_{n=1}B_n\left(\sin(\frac{n\pi x}{L}), \sin(\frac{k\pi x}{L})\right)_2\end{aligned}$$

Only one term is nonzero:

$$\left(\phi(x), \sin(\frac{k\pi x}{L})\right)_2 = B_k\left(\sin(\frac{k\pi x}{L}), \sin(\frac{k\pi x}{L})\right)_2 = B_kL \tag{2.28}$$

Therefore,

$$B_k = \frac{1}{L}\left(\phi(x), \sin(\frac{k\pi x}{L})\right)_2 = \frac{1}{L}\int_{-L}^{L}\phi(x)\sin(\frac{k\pi x}{L})\,dx, \quad k > 0. \tag{2.29}$$

Similarly, we may multiply the sum by $\cos(\frac{k\pi x}{L})$, integrate, and conclude that

$$A_k = \frac{1}{L}\left(\phi(x), \cos(\frac{k\pi x}{L})\right)_2 = \frac{1}{L}\int_{-L}^{L}\phi(x)\cos(\frac{k\pi x}{L})\,dx, \quad k \geq 0. \tag{2.30}$$

Now suppose we are given *any* function $\phi \in L^2$. We can compute the integrals (2.29) and (2.30), and then form the infinite series (2.20). Does the series converge? If so, in what sense is $\phi(x)$ equal to its Fourier series?

Theorem 5.2.2 (L^2 Convergence of Fourier Series) *Let $\phi(x) \in L^2([-L, L])$ and define A_k and B_k by (2.30) and (2.29), respectively. If*

$$\phi_N(x) = \frac{1}{2}A_0 + \sum_{n=1}^{N} \Big(A_n \cos(\frac{n\pi x}{L}) + B_n \sin(\frac{n\pi x}{L}) \Big) \tag{2.31}$$

then

$$\lim_{N\to\infty} \|\phi - \phi_N\|_2 = 0 \tag{2.32}$$

Proof: See Strauss, Theorem 3, p. 124. □

Theorem 5.2.3 (Pointwise Convergence of Fourier Series) *Suppose that $\phi(x)$ and $\phi'(x)$ are both piecewise continuous on $[-L, L]$ then for every $x \in [-L, L]$,*

$$\lim_{N\to\infty} \phi_N(x) = \frac{1}{2}\left(\phi(x^+) + \phi(x^-)\right) \tag{2.33}$$

In particular, if ϕ is actually continuous while ϕ' is piecewise continuous, then ϕ_N converges pointwise to ϕ for all $x \in (-L, L)$.

Proof: See Strauss, Theorems 2 and 4, p. 124-125, and see section 5.5. □

Theorem 5.2.4 (Parseval's Equality) *For $f \in L^2([-L, L])$,*

$$\|f\|_2^2 = \int_{-L}^{L} |f(x)|^2\, dx = \frac{L}{2}A_0^2 + L\sum_{n=1}^{\infty} \left(|A_n|^2 + |B_n|^2\right) \tag{2.34}$$

This shows that for $\phi(x) \in L^2([-L, L])$ and $\phi_N(x)$ defined by the partial sum (2.31), the error in approximating ϕ by ϕ_N has L^2 norm

$$\|\phi - \phi_N\|_2^2 = L \sum_{n=N+1}^{\infty} \left(|A_n|^2 + |B_n|^2\right) \tag{2.35}$$

So if the coefficients decay to zero very rapidly as $n \to \infty$, this error will converge to zero very rapidly as $N \to \infty$. In this case, ϕ may be well-approximated in L^2 by only a few terms in its Fourier series. This has great significance for numerical approximation methods (for example, in solving PDE's numerically, in image processing, and many other applications). Under what conditions do the coefficients decay to zero rapidly?

Food for thought: Show that if $f(x) \in C^1([-L, L])$ and $f(L) = f(-L)$, then $|A_n| \le Cn^{-1}$ and $|B_n| \le Cn^{-1}$ for some constant $C > 0$, for all n. What if $f \in C^k$ for $k > 1$?

5.3 Solving the heat equation

Now we return to the boundary value problem (1.4) with $D = [0, L]$:

$$\begin{aligned} u_t = u_{xx}, \quad x \in D = [0, L] \qquad \text{(PDE)} \\ u(0,t) = 0 = u(L,t), \qquad \text{(Boundary Condition)} \\ u(x,0) = \phi(x) \qquad \text{(Initial Condition)} \end{aligned} \tag{3.36}$$

(Note: we now use $D = [0, L]$.) From the method of separation of variables, we expect that the solution is given by the infinite series

$$u(x,t) = \sum_{n=1}^{\infty} \sin(\frac{\pi n x}{L})\theta_n(t) = \sum_{n=1}^{\infty} b_n e^{-(\frac{n\pi}{L})^2 t} \sin(\frac{\pi n x}{L}) \tag{3.37}$$

where b_n are the Fourier coefficients for ϕ. We'll show that the function defined by this series does indeed solve the equation.

If $\phi(x) \in L^2(D)$, then its Fourier sine series is well-defined, and the coefficients satisfy

$$\|\phi\|_{L^2}^2 = \sum_{n=1}^{n} |b_n|^2 < \infty \tag{3.38}$$

This also implies that for each $t > 0$, $u(\cdot, t) \in L^2(D)$ since

$$\|u(\cdot,t)\|_{L^2}^2 = \sum_{n=1}^{\infty} |b_n|^2 |e^{-2\lambda_n t}| \leq e^{-2\lambda_1 t} \sum_{n=1}^{n} |b_n|^2 \leq e^{-2\lambda_1 t} \|\phi\|_{L^2}^2 \tag{3.39}$$

In fact, we can show that for each $t > 0$, u is actually smooth! To see this, we consider the partial sums defining u:

$$S_N(x,t) = \sum_{n=1}^{N} b_n e^{-(\frac{n\pi}{L})^2 t} \sin(\frac{\pi n x}{L}) \tag{3.40}$$

for each N, this is a smooth function of (x, t) for $t > 0$, $x \in D$. Moreover, the partial sums must converge uniformly. To see this, notice that for any $M < N$

$$\begin{aligned} |S_N - S_M| &\leq \sum_{n=M+1}^{N} |b_n| e^{-(\frac{n\pi}{L})^2 t} \\ &\leq \max_n |b_n| \sum_{n=M+1}^{N} e^{-(\frac{n\pi}{L})^2 t} \end{aligned} \tag{3.41}$$

Because the terms in the last sum are exponentially small in n, this difference converges to zero as $N, M \to \infty$, uniformly in $x \in D$ and $t \in [\epsilon, \infty)$ for any $\epsilon > 0$. Therefore, S_N defines a Cauchy sequence. To see that it is bounded, notice that

$$|S_N(x,t)| \leq \sum_{n=1}^{N} |b_n| e^{-(\frac{n\pi}{L})^2 t} \leq \max_n |b_n| \sum_{n=1}^{N} e^{-(\frac{n\pi}{L})^2 t}$$

This is uniformly bounded in N, uniformly in $x \in D$ and $t \in [\epsilon, \infty)$ for any $\epsilon > 0$. Therefore, the partial sums $S_N(x,t)$ (which are continuous) converge uniformly to a function $u(x,t)$:

$$\lim_{N\to\infty} \max_{\substack{x\in D \\ t\in[\epsilon,\infty)}} |u(x,t) - S_N(x,t)| = 0 \tag{3.42}$$

Therefore, the function $u(x,t)$ defined by this series is a continuous function.

We can perform the same analysis on the derivatives of S_N. Since the partial sums are smooth functions, we can take derivatives term by term. By applying the same analysis as above to the functions $\frac{\partial}{\partial x}S_N$ and $\frac{\partial}{\partial t}S_N$, and to all higher derivatives, we find the function u is actually smooth for $t > 0$! The derivatives of S_N converge uniformly to the derivatives of u for $x \in D$ and $t \in [\epsilon, \infty)$ for any $\epsilon > 0$. This fact depends crucially on the exponential decay of the terms in the Fourier series for u and does not hold for general Fourier series. Since each partial sum satisfies the heat equation and the Dirichlet boundary condition, this convergence analysis shows that $u_t = u_{xx}$ as well, for all $x \in D$, $t > 0$. Moreover, $u(0,t) = u(L,t) = 0$.

What happens as $t \to 0$? It is not hard to show that as $t \to 0$,

$$\lim_{t\to 0} \|u(\cdot,t) - \phi(\cdot)\|_{L^2(D)} = 0 \tag{3.43}$$

Notice that the function ϕ need not be continuous, since ϕ can be any function in $L^2(D)$. In summary we have shown the following

Theorem 5.3.1 *Let $\phi(x) \in L^2([0,L])$. There exists a unique function $u(x,t) \in C^\infty((0,L)\times(0,\infty))$ such that*

$$\begin{aligned} u_t = u_{xx} \quad & x \in [0,L],\ t > 0 \\ u(0,t) = u(L,t) = 0, \quad & t > 0 \end{aligned} \tag{3.44}$$

and

$$\lim_{t\to 0} \|u(\cdot,t) - \phi(\cdot)\|_{L^2(D)} = 0 \tag{3.45}$$

The function u is defined by the Fourier sine series (3.37).

The fact that u is unique was proved earlier in the course, assuming existence of a solution. If ϕ is continuous and $\phi(0) = \phi(L) = 0$, then one can show (by other methods) that $u(x,t) \to \phi(x)$ as $t \to 0$ uniformly in x, not just in the L^2 sense.

A remark on boundary conditions

In problem (0.1) we have imposed homogeneous Dirichlet boundary conditions. This sort of problem would arise in the Black-Scholes model for the price of a double-barrier option after the equation is transformed to the heat equation. There also are other boundary conditions that are of interest for various applications. Many common linear boundary conditions have the form

$$\alpha(x)u + \beta(x)\frac{\partial u}{\partial \nu} = g(x), \quad x \in \partial D \tag{3.46}$$

where α, β, and g are prescribed functions. The boundary condition has a different interpretation, depending on the choice of α, β, g, and on the physical model.

The **Dirichlet boundary condition** specifies the value of u on the boundary and corresponds to $\beta \equiv 0$. We will assume $\alpha \equiv 1$, so that the condition may be expressed as:

$$u(x,t) = g(x), \quad x \in \partial D \tag{3.47}$$

where g is some function that we will assume to be smooth. If u models temperature, then the Dirichlet condition corresponds to fixing the temperature at the boundary. A common Dirichlet condition is $g \equiv 0$. In fact, if $g \neq 0$, we may transform the problem so that the boundary condition becomes $g \equiv 0$, as follows. Suppose that u satisfies

$$\begin{aligned} u_t &= \Delta u, \quad x \in D \\ u(x,t) &= g(x), \quad x \in \partial D \\ u(x,0) &= \phi(x) \quad x \in D \end{aligned} \tag{3.48}$$

and that $h(x) \in C^2(\overline{D})$ is any function such that $h(x) = g(x)$ for all $x \in \partial D$. Then, the function $w(x,t) = u(x,t) - h(x)$ solves

$$\begin{aligned} w_t &= \Delta w + f, \quad x \in D \\ w(x,t) &= 0, \quad x \in \partial D \\ w(x,0) &= \phi(x) - h(x), \quad x \in D \end{aligned} \tag{3.49}$$

where $f(x) = \Delta h(x)$. Thus, if we can solve a problem of the form (3.49) where $g \equiv 0$, we can solve problem (3.48) by setting $u(x,t) = w(x,t) + h(x,t)$.

The **Neumann boundary condition** specifies the normal derivative of u on the boundary and corresponds to $\alpha \equiv 0$, $\beta \neq 0$. We will assume $\beta \equiv 1$, so that the condition may be expressed as:

$$\frac{\partial u}{\partial \nu} = \nu \cdot \nabla u = g(x), \quad x \in \partial D \tag{3.50}$$

where ν is an exterior unit normal. As in the Dirichlet case, g is often chosen to be identically zero. If u models heat transfer, this models a situation where the heat flux across the boundary is zero (because of an insulating material), which is different from setting the temperature to zero.

The **Robin boundary condition** is the third case where $\alpha \neq 0$ and $\beta \neq 0$. The Robin condition models the case where the heat flux through the boundary depends on the temperature itself.

The boundary conditions are called **homogeneous** if $g \equiv 0$.

Source terms

Suppose we wish to solve the inhomogeneous equation:

$$\begin{aligned} w_t &= w_{xx} + h(x,t), \quad x \in D = [0, L] \\ w(0,t) &= 0, \quad w(L,t) = 0, \\ w(x,0) &= \phi(x). \end{aligned} \tag{3.51}$$

As before, we expand the solution in terms of the eigenfunctions v_n:

$$u(x,t) = \sum_n \theta_n(t) v_n(x) \tag{3.52}$$

Assuming $f(x,t) \in L^2(D)$ for all t, we may write

$$f(x,t) = \sum_n f_n(t) v_n(x) \tag{3.53}$$

By formally plugging these series into the equation and equating coefficients, we see that the ODE's for $\theta_n(t)$ are:

$$\theta_n'(t) = -\lambda_n \theta_n(t) + f_n(t) \tag{3.54}$$

with initial conditions $\theta_n(0) = a_n$, where a_n are the Fourier coefficients of ϕ. The differences now is that the ODE for θ_n is inhomogenous.

Exercise: Using the method of separation of variables, find a representation for the solution to (3.36) if the Dirichlet boundary condition is replaced by the homogeneous Neumann boundary condition: $u_x(0) = u_x(L) = 0$.

Other generalizations

This separation of variables technique extends to multiple dimensions and to equations with variable coefficients. For example, consider the problem

$$\begin{aligned} u_t = \nabla \cdot (a(x) \nabla u) + f(x,t), \quad & x \in D \\ u(x,t) = 0, \quad & x \in \partial D \\ u(x,0) = \phi(x). & \end{aligned} \tag{3.55}$$

where $D \subset \mathbb{R}^d$ is a smooth bounded domain. We suppose that $0 \leq a_* \leq a(x) \leq a^*$ for some constants a^* and a_*, and that $a(x)$ is sufficiently smooth (for example, $a(x) \in C^2(\bar{D})$). We look for a solution of the form

$$u(x,t) = \sum_n v_n(x) \theta_n(t) \tag{3.56}$$

where the functions v_n are eigenfunctions satisfying $-\nabla \cdot (a(x) \nabla v_n) = \lambda_n v_n$ with eigenvalues λ_n, and $v(x) = 0$ for $x \in \partial D$. As before, we express the initial condition in terms of the functions v_n:

$$\phi(x) = \sum_n a_n v_n(x) \tag{3.57}$$

and we expand the forcing term f in terms of v_n:

$$f(x,t) = \sum_n f_n(t) v_n(x) \tag{3.58}$$

The functions $\theta_n(t)$ satisfy a second-order ODE, as before. If $f \neq 0$, then these ODE's will be inhomogeneous.

When can we do this?

Theorem 5.3.2 *There exists a set of eigenfunctions $\{v_n(x)\}_{n=1}^{\infty}$ normalized by $\int_D |v_n(x)|^2 \, dx = 1$ which forms an orthonormal basis for $L^2(D)$.*

Proof: See Evans Theorem 1, section 6.5. □

Therefore, if $f, \phi \in L^2(D)$, the expansions (3.57) and (3.58) are valid, the series converging in the L^2 sense. In general, there is no explicit formula for the eigenfunctions v_n or the eigenvalues λ_n, as in the case of the Laplacian (Δ) when the domain is an interval, rectangle, or sphere.

Whether this series solution satisfies the PDE in the classical sense depends on the smoothness of the coefficients and on the data. One may show that

Theorem 5.3.3 *If the coefficients $a(x)$ are sufficiently smooth and the functions f and ϕ are sufficiently smooth with derivatives vanishing at the boundary, then this construction produces a classical solution to the problem (3.55). In particular, if $a(x) \in C^\infty(D)$, and if $f(x), \phi(x) \in C_0^\infty(D')$ for some $D' \subset\subset D$, then $u \in C^\infty$.*

Proof: See Evans, section 7.1. □

.

An excellent reference for Fourier series is: A. Zygmund, *Trigonometric Series*, Vols. I and II, Third Edition, Cambridge UP, 2002.

Chapter 6

The Black-Scholes Equation

References:

- F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Political Economy **81** (1973), p. 637-654.
- R. Merton, *Theory of rational option pricing* , Bell Journal of Economics and Management Science **4** (1973), pp. 141-183.
- M. Harrison and S. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stoch. Proc. Appls. **11** (1981), pp. 215-260.
- R. Korn and E. Korn, *Option Pricing and Portfolio Optimization*, AMS, 2001.
- P. Wilmott, S. Howison, J. Dewynne, *The mathematics of financial derivatives*, Cambridge UP, 1995.

In these notes we study the **Black-Scholes equation**:

$$u_t + \frac{\sigma^2}{2}x^2 u_{xx} + rxu_x - ru = 0, \quad t < T, \quad x > 0 \tag{0.1}$$

with terminal condition $u(x,T) = g(x)$. The function $u(x,t)$ models the fair price of a contingent claim or "option" based on the market price of an underlying security. The variable x represents the market price of the security, and the variable t represents time. The fixed time T is called the **expiry** or time of maturity. The solution $u(x,t)$ depends on the current time, the current stock price, the expiry time, and the boundary condition $g(x)$ which is determined by the type of option. The solution also depends on the parameters σ and r which represent the **volatility** of the underlying asset and the **risk-free interest rate**, respectively.

There are many types of options. A **European call option** is a contract that entitles the holder of the option to buy a given stock at price $K > 0$, called the **strike price**, at a future time T. If the option is exercised, the seller of the option is obligated to sell the stock at price K regardless of the market price of the stock. At the expiry time T (in the future), the market price of the stock might be greater than or less than K. Therefore, if the price of the stock at time T is greater than K, then the option holder can buy the stock at K (from the option seller) and immediately sell the

stock at the higher price (on the market), making a profit. On the other hand, if the market price of the stock at time T is less than K, then at time T the holder of the option gains nothing from exercising the option. So, no action is taken; the option is worthless at this point. The question that Black, Scholes, and Merton were trying to answer in their seminal papers (1973) is: "How should one set the price of this option contract, given that the stock price now at time $t < T$ is x?"

6.1 Deriving the Black-Scholes equation via risk-neutral pricing

Here we derive the equation (0.1). We suppose that there are two asset classes: a risk-free bond (treasury note) and a risky stock. Let Y_t denote the price of the bond and X_t denote the price of the stock. We assume that Y_t and X_t are governed by the following equations:

$$\begin{aligned} dY_t &= rY_t dt \\ dX_t &= \mu X_t dt + \sigma X_t dB_t \end{aligned} \tag{1.2}$$

The equation for Y_t is just an ODE with solution $Y_t = Y_0 e^{rt}$. The constant $r > 0$ is the risk-free interest rate. The equation for X_t is a stochastic equation modeling random fluctuations in the stock price. This model of the stock price is called **geometric Brownian motion**. Using Itô's formula we find that the function $H_t = \log(X_t)$ satisfies the stochastic equation

$$\begin{aligned} dH_t &= \frac{1}{X_t} dX_t - \frac{1}{2}\frac{1}{X_t^2}\sigma^2 X_t^2 dt \\ &= \frac{1}{X_t}\left(\mu X_t dt + \sigma X_t dB_t\right) - \frac{1}{2}\frac{1}{X_t^2}\sigma^2 X_t^2 dt \\ &= (\mu - \frac{\sigma^2}{2})dt + \sigma dB_t \end{aligned}$$

Therefore, $H_t = qt + \sigma B_t$ with $q = \mu - \sigma^2/2$, and

$$X_t = e^{H_t} = e^{qt+\sigma B_t} = e^{(\mu-\sigma^2/2)t+\sigma B_t}$$

Notice that $X_t > 0$ with probability one; the stock price cannot become negative! The expected stock price is:

$$E[X_t] = e^{qt}E[e^{\sigma B_t}] = e^{qt+t\sigma^2/2} = e^{\mu t}$$

So, for this model, $\mu = \frac{1}{t}\log E[X_t]$ is the growth rate of the mean stock price. This paramater μ is called the **mean rate of return** for the stock.

For a European call option, the holder of the option may buy at price K at time T and then sell at the current market price, realizing a profit of $X_T - K$, assuming $X_T \geq K$. If $X_T < K$, then the holder of the option does not exercise (since he is under no obligation). So he profits 0. Putting this together, we say that the **payoff** of a European call option is

$$g(X_T) = \max\left(X_T - K, 0\right) \tag{1.3}$$

Sometimes this payoff is denoted by $g(X_T) = (X_T - K)^+$, as well. Notice that the payoff is non-negative.

What is the right price for the option contract? For a fixed strike K and expiry T, we use $u(x,t)$ to denote the price of the option at time t if the current market price is $X_t = x$. Taking into account the possibility of investing in the risk-free bond with interest rate r, one possible pricing strategy would be to set

$$u(x,t) \stackrel{?}{=} E\left[e^{-r(T-t)}g(X_T) \mid X_t = x\right] = E\left[e^{-r(T-t)}\left(X_T - K\right)^+ \mid X_t = x\right]. \tag{1.4}$$

The factor $e^{-r(T-t)}$ is a discount factor, converting future value at time T to present value at time $t < T$. Therefore, under this definition the option price would be the expected net present value of the payoff $g(X_T)$. However, this seemingly natural definition turns out to be the **wrong price!**

The correct price is determined by taking the expectation (1.4) with respect to a different measure called the **risk-neutral measure**. Under this measure, the discounted stock price $\hat{X}_t := e^{-rt}X_t$ is a martingale. The process $\hat{X}_t$ is the net present value (at $t = 0$) of the stock if the price at time t is X_t, discounted according to the risk-free interest rate r. It satisfies the SDE

$$d\hat{X}_t = (\mu - r)\hat{X}_t\,dt + \sigma\hat{X}_t\,dB_t \tag{1.5}$$

We know that

$$\hat{X}_t = e^{-rt}X_t = e^{(\mu-r)t}e^{\sigma B_t - t\sigma^2/2}.$$

The second term, $e^{\sigma B_t - t\sigma^2/2}$, is a martingale with respect to the filtration $\{\mathcal{F}\}_t$ (can you show why?). Therefore, for any $s \le t$,

$$\begin{aligned} E\left[\hat{X}_t \middle| \mathcal{F}_s\right] = e^{(\mu-r)t}E\left[e^{\sigma B_t - t\sigma^2/2} \middle| \mathcal{F}_s\right] &= e^{(\mu-r)t}e^{\sigma B_s - s\sigma^2/2} \\ &= e^{(\mu-r)(t-s)}e^{(\mu-r)s}e^{\sigma B_s - s\sigma^2/2} \\ &= e^{(\mu-r)(t-s)}\hat{X}_s \end{aligned}$$

This shows that if $\mu > r$, the discounted price $\hat{X}_t$ is a sub-martingale; if $\mu < r$ the discounted price is a super-martingale. Only in the special case that $\mu = r$ will the discounted price be a martingale with respect to the original measure.

So, computing the fair price of the contingent claim boils down to determining the risk-neutral measure, with respect to which $\hat{X}_t$ is a martingale. We can determine this measure using Girsanov's theorem, as follows. Girsanov's Theorem tells us that the shifted process

$$\tilde{B}_t = B_t - at$$

is a Brownian motion under the new measure on $(C([0,T]), \mathcal{F})$ defined by

$$\tilde{P}(A) = E\left[e^{aB_T - a^2T/2}\mathbb{I}_A(\omega)\right]$$

where $\mathbb{I}_A(\omega)$ denotes the indicator function for the set $A \in \mathcal{F}$. Returning to (1.5) we see that

$$d\hat{X}_t = (\mu - r + \sigma a)\hat{X}_t\,dt + \sigma\hat{X}_t\,d\tilde{B}_t.$$

So, if we choose $a = (r - \mu)/\sigma$, we have

$$d\hat{X}_t = \sigma\hat{X}_t\,d\tilde{B}_t$$

Thus, under the measure $\tilde{P}$, the discounted price $\hat{X}_t$ is a martingale, since it satisfies

$$\hat{X}_t = \hat{X}_0 + \int_0^t \sigma \hat{X}_s \, d\tilde{B}_s$$

where $\tilde{B}_s$ is a Brownian motion under $\tilde{P}$. Therefore, under this new measure, the stock price (not discounted) and the bond price satisfy

$$\begin{aligned} dY_t &= rY_t dt \\ dX_t &= rX_t dt + \sigma X_t d\tilde{B}_t \end{aligned} \tag{1.6}$$

so that both assets have the same mean rate of return, even though the stock price may have a mean rate of return $\mu \neq r$ under the original measure.

We have now determined the risk-neutral measure to be used in pricing the contingent claim. Therefore, the fair price of the claim is

$$u(x,t) = \tilde{E}\left[e^{-r(T-t)} g(X_T) \mid X_t = x\right]. \tag{1.7}$$

Pause for a moment and consider this formula. The term $e^{-r(T-t)}g(X_T)$ is the discounted payoff. However, we are taking the expectation with respect to the risk-neutral measure. Under the measure $\tilde{P}$, the process $\tilde{B}_t$ is a standard Brownian motion. Thus, considering (0.2) and (1.6) we see that fair price is completely independent of the original mean return rate μ!

Now consider $u(x,t)$ as a function of x and t, where x is the known value of the underlying stock at time $t < T$. We have already seen that if $\tilde{B}_t$ is a Brownian motion, the stochastic process defined by

$$dX_t = rX_t dt + \sigma X_t d\tilde{B}_t, \quad X_0 = x > 0$$

is associated with the differential operator

$$\mathcal{A}u := \frac{\sigma^2 x^2}{2} u_{xx} + rxu_x$$

In particular, the function $u(x,t)$ is the stochastic representation for a solution to the terminal value PDE

$$u_t + \mathcal{A}u + c(x,t)u = 0, \quad x > 0, \;\; t < T \tag{1.8}$$

with terminal condition $u(x,T) = g(x)$, where $c(x,t)$ is just the constant $c(x,t) = -r$ in the present case. Thus, (0.2) represents a solution to the terminal value problem

$$\begin{gathered} u_t + \frac{\sigma^2 x^2}{2} u_{xx} + rxu_x - ru = 0, \quad x > 0, \;\; t < T \\ u(x,T) = g(x), \quad x \in \mathbb{R} \end{gathered} \tag{1.9}$$

This PDE is called the **Black-Scholes equation**. It is a terminal value problem, not an initial value problem. Notice that the PDE itself does not depend on the specific form of the payoff at time T. While deriving the fair price (and the resulting equation) we only assumed that the option may be exercised only at expiry T. The function $g(x)$, however, need not be $(x-K)^+$ as in the case of a European call.

6.2 Transformation to heat equation and dimensionless variables

Although the equation (1.9) has variable coefficients and low-order terms, we may transform the equation into the heat equation through a sequence of simple transformations, as we now demonstrate. At the same time, we will reduce the number of parameters in the model by transforming to "dimensionless variables." This procedure demonstrates an important technique in many other applications of PDEs, as well. Although there may be many parameters in a PDE model, the solution to the underlying model may depend on only a few fundamental quantities which are independent of the system of units chosen to define the parameters. So, it is often useful to change variables to normalize the units and eliminate unnecessary degrees of freedom.

The basic steps in our transformation of the Black-Scholes equation are those learned in Homework 1. Suppose that a function $w(x,t)$ satisfies the heat equation $w_t = w_{xx}$ for $x \in \mathbb{R}$ and $t > 0$:

- The function $v(x,t) = w(x, T-t)$ satisfies
$$v_t + v_{xx} = 0, \quad t < T$$
- For $x > 0$, the function $v(x,t) = w(\log(\lambda^{-1}x), t)$ satisfies
$$v_t = x^2 v_{xx} + x v_x, \quad x > 0$$
- The function $v(x,t) = w(\gamma x, \delta t)$ satisfies
$$v_t = \frac{\delta}{\gamma^2} v_{xx}$$
- The function $v(x,t) = e^{\alpha x} w(x,t)$ satisfies
$$v_t = v_{xx} - 2\alpha v_x + \alpha^2 v$$
- The function $v(x,t) = e^{\beta t} w(x,t)$ satisfies
$$v_t = v_{xx} + \beta v$$

Using these observations, we now start from the Black-Scholes PDE and work toward the heat equation. Suppose that $u(x,t)$ satisfies the Black-Scholes PDE

$$u_t + \frac{\sigma^2}{2} x^2 u_{xx} + r x u_x - r u = 0, \quad t < T, \;\; x > 0$$

with the terminal condition $u(x,T) = g(x)$. There are at least three parameters in the equation: r, T, and σ and perhaps parameters defining the terminal data g (the strike price K, for example). Moreover, the variables x and t have specific units (e.g. Dollars, Euros, Yen / seconds, hours, etc.). First we make a time-change: define the function $v^1(x,\tau) = u(x, T - \delta\tau)$ for $\tau > 0$. Then $v^1(x,\tau)$ satisfies the initial value problem

$$v^1_\tau = \delta \frac{\sigma^2}{2} x^2 v^1_{xx} + \delta r x v^1_x - \delta r v^1, \quad \tau > 0, \;\; x > 0$$

So, if we choose $\delta = 2/\sigma^2$, we have

$$v^1_\tau = x^2 v^1_{xx} + \frac{2r}{\sigma^2} x v^1_x - \frac{2r}{\sigma^2} v^1, \quad \tau > 0, \;\; x > 0$$

The initial condition for v^1 is $v^1(x,0) = u(x,T) = g(x)$. Not only have we transformed the problem to an initial value problem, we have also normalized the temporal units, by scaling with the factor δ. For convenience, let us define $k := \frac{2r}{\sigma^2}$. Then

$$v^1_\tau = x^2 v^1_{xx} + kx v^1_x - kv^1, \quad \tau > 0, \;\; x > 0$$

Now we transform the terms on the right hand side, working with the highest order terms first. The variable x lies in the domain $[0,\infty)$, and has a specific monetary unit. For $\lambda > 0$ a parameter to be chosen, we change variables according to $x = \lambda e^y$ for $y \in \mathbb{R}$ a dimensionless variable. That is, we define a function $v^2(y,\tau)$ for $y \in \mathbb{R}$ by $v^2(\log(\lambda^{-1}x),\tau) = v^1(x,\tau)$. Thus, $v^2(y,\tau) = v^1(\lambda e^y,\tau)$, and a simple computation shows that v^2 satsifes

$$\begin{aligned} v^2_\tau &= (v^2_{yy} - v^2_y) + k(v^2_y) - kv^2 \\ &= v^2_{yy} + (k-1)v^2_y - kv^2, \quad \tau > 0, \;\; y \in \mathbb{R} \end{aligned}$$

with initial condition $v^2(y,0) = g(\lambda e^y)$.

Now define $v^3(y,\tau)$ by $v^3(y,\tau) = e^{\alpha y} v^2(y,\tau)$. Thus, $v^2 = e^{-\alpha y} v^3$ so that

$$v^2_y = v^3_y e^{-\alpha y} - \alpha v^3 e^{-\alpha y}, \qquad v^2_{yy} = v^3_{yy} e^{-\alpha y} + \alpha^2 v^3 e^{-\alpha y} - 2\alpha v^3_y e^{-\alpha y}$$

and

$$\begin{aligned} v^3_\tau &= v^3_{yy} + \alpha^2 v^3 - 2\alpha v^3_y + (k-1)(v^3_y - \alpha v^3) - kv^3 \\ &= v^3_{yy} + ((k-1) - 2\alpha)\, v^3_y + \left(\alpha^2 - \alpha(k-1) - k\right) v^3 \end{aligned}$$

By choosing $\alpha = \frac{1}{2}(k-1)$, this reduces to

$$v^3_\tau = v^3_{yy} - (\alpha^2 + k) v^3, \quad t > 0, \;\; y \in \mathbb{R}$$

The initial condition for v^3 is $v^3(x,0) = e^{\alpha y} v^2(x,0) = e^{\alpha y} g(\lambda e^y)$.

Next, let $v^4(y,\tau) = e^{\beta\tau} v^3(y,\tau)$, with $\beta = (\alpha^2 + k) = \frac{1}{4}(k+1)^2$. We see that $v^4(y,\tau)$ satisfies the heat equation $v^4_\tau = v^4_{yy}$, with initial condition $v^4(y,0) = e^{\alpha y} g(\lambda e^y)$. The function v^4 still has a dimensional value (this is the price of the option in dollars/euros/etc.) So, the last step is to normalize these units by defining $v^5(y,\tau) = p v^4(y,\tau)$, where p is some scaling factor. Thus, v^5 satisfies

$$v^4_\tau = v^4_{yy}, \quad \tau > 0, \;\; x \in \mathbb{R}$$

with initial condition $v^5(y,0) = p e^{\alpha y} g(\lambda e^y)$.

In summary we made the following transformations:

- $v^1(x,\tau) = u(x, T - \delta\tau), \qquad \delta = 2/\sigma^2, \qquad \tau = \delta^{-1}(T-t)$
- $v^2(y,\tau) = v^1(\lambda e^y,\tau), \qquad y = \log(\lambda^{-1}x)$

- $v^3(y,\tau) = e^{\alpha y} v^2(y,\tau), \qquad \alpha = \frac{1}{2}(k-1), \;\; k = 2r/(\sigma^2)$
- $v^4(y,\tau) = e^{\beta \tau} v^3(y,\tau), \qquad \beta = \frac{1}{4}(k+1)^2, \;\; k = 2r/(\sigma^2)$
- $v^5(y,\tau) = p v^4(y,\tau).$

From now on we drop the super-script and use $v(y,\tau) = v^5(y,\tau)$ to denote the transformed solution. Working through these transformations in reverse, we see that

$$u(x,t) = p^{-1} e^{-\alpha \log(\lambda^{-1}x)} e^{-\beta \delta^{-1}(T-t)} v(\log(\lambda^{-1}x), \delta^{-1}(T-t)) \tag{2.10}$$

An explicit solution formula for the European call option

The terminal data for a European call option is

$$g(x) = \max(x - K, 0)$$

where K is the strike price (in monetary units). Considering the transformations above, we see that this terminal condition is transformed to the initial condition

$$v(y,0) = p e^{\alpha y} g(\lambda e^y) = p e^{\alpha y} \max(\lambda e^y - K, 0) = \max(p e^{\alpha y} \lambda e^y - p e^{\alpha y} K, 0)$$

So, if we choose $\lambda = p^{-1} = K$, this becomes

$$v(y,0) = \max\left(e^{(\alpha+1)y} - e^{\alpha y}, 0\right) = \max\left(e^{\frac{1}{2}(k+1)y} - e^{\frac{1}{2}(k-1)y}, 0\right)$$

Thus, we have transformed the terminal value problem for the price of a European call option to an initial value problem

$$\begin{aligned} v_\tau &= v_{yy}, \quad \tau > 0, \;\; y \in \mathbb{R} \\ v(y,0) = v_0(y) &:= \max\left(e^{\frac{1}{2}(k+1)y} - e^{\frac{1}{2}(k-1)y}, 0\right), \quad y \in \mathbb{R} \end{aligned} \tag{2.11}$$

Notice that although there were originally four parameters in the model (σ, r, T, K) there is now only one fundamental parameter k. So if we could solve this initial value problem for every k, we could solve for all possible choices of σ, r, T, and K by reversing the transformations above.

Consider the initial data associated with (2.11). For $y = 0$, $v_0(0) = 0$. For $y < 0$, $e^{\frac{1}{2}(k+1)y} < e^{\frac{1}{2}(k-1)y}$, so $v_0(y) = 0$ for $y \le 0$. For $y > 0$, $v_0(y) > 0$ and v_0 grows exponentially in y. Thus, a solution to (2.11) is given by the convolution formula:

$$\begin{aligned} v(y,\tau) = \Phi * v_0 &= \int_{\mathbb{R}} \Phi(y-z,\tau) v_0(z)\, dz \\ &= \frac{1}{\sqrt{4\pi\tau}} \int_0^\infty e^{-\frac{|y-z|^2}{4\tau}} \left(e^{\frac{1}{2}(k+1)z} - e^{\frac{1}{2}(k-1)z}\right) dz \end{aligned} \tag{2.12}$$

where Φ is the fundamental solution for the heat equation. This solution is unique among the class of functions satisfying the usual growth condition.

Because of the special form of the initial data in (2.11), this convolution integral can be computed explicitly! Then we may reverse the transformations to compute $u(x,t)$, the price of the option. To evaluate the integral, we first "complete the square":

$$\begin{aligned} -\frac{(y-z)^2}{4\tau} + cz &= -\frac{1}{4\tau}\left((y-z)^2 - cz4\tau\right) \\ &= -\frac{1}{4\tau}\left((z-(y+2c\tau))^2 - 4c^2\tau^2 - 4cy\tau\right) \\ &= -\frac{(z-(y+2c\tau))^2}{4\tau} + c^2\tau + cy \end{aligned} \tag{2.13}$$

Thus,

$$\begin{aligned} \int_0^\infty \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{(y-z)^2}{4\tau}} e^{cz}\, dz &= e^{c^2\tau+cy}\int_0^\infty \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{(z-(y+2c\tau))^2}{4\tau}}\, dz \\ &= e^{c^2\tau+cy}\left(1-\Psi(-\frac{y+2c\tau}{\sqrt{2\tau}})\right) \\ &= e^{c^2\tau+cy}\Psi(\frac{y+2c\tau}{\sqrt{2\tau}}) \end{aligned} \tag{2.14}$$

where $\Psi(r) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{r} e^{-z^2/2}\, dz$ (this function Ψ is called the **cumulative normal distribution function**). Therefore, the solution to the transformed equation is

$$v(y,\tau) = e^{c_1^2\tau+c_1 y}\Psi(\frac{y+2c_1\tau}{\sqrt{2\tau}}) - e^{c_2^2\tau+c_2 y}\Psi(\frac{y+2c_2\tau}{\sqrt{2\tau}}) \tag{2.15}$$

where c_1 and c_2 are the constants

$$\begin{aligned} c_1 &= \frac{1}{2}(k+1) = \frac{1}{\sigma^2}\left(r+\frac{\sigma^2}{2}\right) = \sqrt{\beta} = \alpha+1, \\ c_2 &= \frac{1}{2}(k-1) = \frac{1}{\sigma^2}\left(r-\frac{\sigma^2}{2}\right) = \alpha. \end{aligned} \tag{2.16}$$

Now, we combine this computation with our transformations to compute $u(x,t)$:

$$u(x,t) = Ke^{-\alpha y}e^{-\beta\tau}v(y,\tau) \tag{2.17}$$

with $\tau = \delta^{-1}(T-t) = \sigma^2(T-t)/2$ and $y = \log(x/K)$. Notice that in the original variables,

$$d_1 := \frac{y+2c_1\tau}{\sqrt{2\tau}} = \frac{\log(x/K)+(r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{2.18}$$

and

$$d_2 := \frac{y+2c_2\tau}{\sqrt{2\tau}} = \frac{\log(x/K)+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{2.19}$$

Therefore, in view of (2.15), (2.16), and (2.17), the option price is given by the explicit formula

$$u(x,t) = x\Psi(d_1) - Ke^{-r(T-t)}\Psi(d_2)$$

This formula is called the **Black-Scholes formula** for the price of a European call option (see Black and Scholes, p. 644). Notice that d_1 and d_2 depend on x, t, σ, T, K, but the function Ψ does not.

Barrier options

A **barrier option** is one that either expires or becomes active when the underlying stock price crosses a threshold or **barrier**. One type of barrier option is a **down-and-out** call option. This is a call option that expires if the value of the stock price drops below the barrier before the expiration time T. An **up-and-out** call option expires if the stock rises above the barrier before time T. These are also called **knock-out options**.

The price of a such a barrier option may be computed explicitly for the simple Black-Scholes model. Suppose that $u(x,t)$ is the price of a down-and-out European call option with knock-out price of $x_0 < K$. According to the Black-Scholes model, the risk-neutral price is:

$$u(x,t) = \tilde{E}\left[\mathbb{I}_{\{\gamma^{x,t} \geq T\}}\, e^{-r(T-t)} g(X_T) \mid X_t = x\right]. \tag{2.20}$$

where the stopping time $\gamma^{x,t}$ is the first time the stock price hits the barrier:

$$\gamma^{x,t}(\omega) = \inf\{s > t \mid X_s(\omega) \leq x_0\} \tag{2.21}$$

and $\mathbb{I}_A$ is the indicator function of the set A. So, the expectation is taken only over those paths that do not cross the barrier (over the set $\{\gamma^{x,t} \geq T\}$), since the option expires and is worthless whenever X_t hits x_0 before time T. This function $u(x,t)$ is the solution to the terminal value problem

$$\begin{gathered} u_t + \frac{\sigma^2 x^2}{2} u_{xx} + rxu_x - ru = 0, \quad x > x_0, \;\; t < T \\ u(x,T) = g(x), \quad x > x_0 \\ u(x_0,t) = 0, t < T \end{gathered} \tag{2.22}$$

This is the Black-Scholes terminal value problem with an additional boundary condition, corresponding to the knock-out price. The domain for x is now the half-line: $x > x_0$.

As in the case of the "plain vanilla" European call, we may transform the equation to the heat equation via a logarithmic change of variables. Changing variables as before, we may write

$$u(x,t) = Ke^{-\alpha y - \beta\tau} v(y,\tau) \tag{2.23}$$

where y, τ, α, and β are defined as before and v solves the initial value problem

$$\begin{gathered} v_\tau = v_{yy}, \quad y > y_0, \; \tau > 0 \\ v(y,0) = \max\left(e^{\frac{1}{2}(k+1)y} - e^{\frac{1}{2}(k-1)y}, 0\right), \quad y > y_0 \\ v(y_0,t) = 0, \quad t < T \end{gathered}$$

The point $y_0 \in \mathbb{R}$ is the transformed barrier $y_0 = \log(\lambda^{-1}x_0) = \log(x_0/K)$, and $y_0 < 0$ under the assumption that $x_0 < K$.

So, we must solve the initial value problem for the heat equation on the half line $\{y > y_0\}$. This we can do using the reflection technique. Extend the initial data by odd-reflection about the point $y = y_0$, and solve the problem on the entire line. The corresponding solution will have odd symmetry about the barrier: $v(y_0 + h, t) = -v(y_0 - h, t)$. Consequently, the boundary condition $v(y_0, t) = 0$ is achieved.

Notice that the extended initial data may be written as:

$$v^{ex}(y) = v_0(y) - v_0(2y_0 - y), \quad y \in \mathbb{R} \tag{2.24}$$

where $v_0(y) = \max\left(e^{\frac{1}{2}(k+1)y} - e^{\frac{1}{2}(k-1)y}, 0\right)$. Suppose that v^1 and v^2 satisfy the heat equation on the whole line with initial data $v^1(y,0) = v_0(y)$ and $v^2(y,0) = v_0(2y_0 - y)$, then the function $v(y,t)$ is given by

$$v(y,t) = v^1(y,t) - v^2(y,t) = v^1(y,t) - v^1(2y_0 - y, t) \tag{2.25}$$

Thus,

$$u(x,t) \quad = \quad Ke^{-\alpha y - \beta\tau} v^1(y,t) - Ke^{-\alpha y - \beta\tau} v^1(2y_0 - y, t) \tag{2.26}$$

Notice that the first term is $C(x,t)$, the price of the plain vanilla European call with no barrier. Consider the second term. Under the relation $x = Ke^y$, the point $2y_0 - y$ corresponds to x_0^2/x. Therefore,

$$Ke^{-\alpha y - \beta\tau} v^1(2y_0 - y, t) = e^{-\alpha y} e^{\alpha(2y_0 - y)} Ke^{-\alpha(2y_0 - y) - \beta\tau} v^1(2y_0 - y, t) = e^{-2\alpha(y - y_0)} C(\frac{x_0^2}{x}, t) \tag{2.27}$$

and the price of the barrier option is

$$\begin{aligned} u(x,t) &= C(x,t) - e^{-2\alpha(y-y_0)} C(\frac{x_0^2}{x}, t) \\ &= C(x,t) - \left(\frac{x}{x_0}\right)^{-(k-1)} C(\frac{x_0^2}{x}, t) \end{aligned} \tag{2.28}$$

since $2\alpha = (k-1)$. Recall that $k = 2r/\sigma^2$. The up-and-out option may be treated in a similar manner.

Chapter 7

Volatility Estimation and Dupire's Equation

References:

- B. Dupire, *Pricing and hedging with smiles*, in Mathematics of Derivative Securities, Cambridge University Press, 1997, pp. 103-111.
- I. Bouchouev and V. Isakov, *The inverse problem of option pricing*, Inverse Problems **13** (1997), pp. L11-L17
- O. Pironneau, *Dupire-like identities for complex option*, C.R. Acad. Sci. Paris, Ser. I, **344** (2007), pp. 127-133.
- Y. Achdou and O. Pironneau, *Volatility smile by multilevel least square*, Int. J. Theor. Appl. Finance, **5** (2002), pp. 619-643.

In our derivation of the Black-Scholes equation we assumed that the volatility was a constant, independent of time and of the stock price itself. This is a significant simplification of the problem which enabled us to compute the price of some options explicitly. In general, we might model the volatility as a deterministic function of both the stock price and time $\sigma = \sigma(x,t)$, so that under the risk-neutral measure, the stock price satisfies:

$$dX_t = r(t)X_t dt + \sigma(X_t, t)\, d\tilde{B}_t \tag{0.1}$$

where $\tilde{B}_t$ is a Brownian motion under the risk-neutral measure $\tilde{P}$. In such a model $\sigma(x,t)$ is sometimes called the **local volatility**. Here we also allow the interest rate to change with time. Then the price of the option is

$$u(x,t) = \tilde{E}\left[e^{-\int_t^T r(s)\,ds} g(X_T) \mid X_t = x\right]. \tag{0.2}$$

where $g(x)$ is the payoff. So, under suitable conditions on σ, we see from the Feynman-Kac representation formula that this represents the solution to the terminal value problem

$$u_t + \frac{\sigma^2(x,t)x^2}{2} u_{xx} + r(t)xu_x - r(t)u = 0, \quad t < T, x > 0 \tag{0.3}$$

with $u(x,T) = g(x)$. Now the coefficients in the equation depend on x and t. Suppose that $g(x) = (x-K)^+$ is the payoff for a European call option. Let us refer to the corresponding solution by $u(x,t;T,K,\sigma)$ to emphasise that the solution depends on the strike K, the expiration time T, and the volatility function σ.

In this generality, one cannot obtain explicit formulas for option prices $u(x,t;T,K,\sigma)$ as in the case of constant volatility. So, one must settle for a numerical solution of the terminal value problem. What should be used for $\sigma(x,t)$? Our goal is to obtain a good estimate of σ given market data. While we don't know the future history of the stock, we may have knowledge of the market prices of various options for various strikes and maturities. Presumably these prices are given by the solution of the terminal value problem above. So, this suggests the following mathematical question:

> Given a set of times $\{t_i\}_i$, stock prices $\{X_i\}_i$, strikes $\{K_i\}_i$, and expiration times $\{T_i\}_i$, and given the market prices $\{C_i\}_i$ of call options corresponding to these parameters, can we find a function $\sigma(x,t)$ such that
>
> $$u(X_i,t_i;T_i,K_i,\sigma) = C_i \tag{0.4}$$
>
> holds for each i? That is, can we find σ so that our model fits the observed market prices of call options?

This sort of problem is called an inverse problem. If we knew the local volatility $\sigma(x,t)$ for all x and t then we could solve the terminal value problem (perhaps, numercally), and knowing market prices of options is somehow related to such solutions (0.3), assuming our model is a good one. So, given such "observations of a solution", can we reconstruct the parameters in the PDE? This sort of problem arises in many applications of PDEs, not just in financial applications.

7.1 Dupire's equation

For simplicity we assume $r(t) = r$ is constant in this section.

Theorem 7.1.1 *Suppose that the function $u(x,t;K,T)$ solves (0.3) with terminal data $u(x,T) = g(x) = (x-K)^+$. Then for fixed $x > 0$ and t, the function $v(K,T) = u(x,t;K,T)$ satisfies the initial value problem*

$$v_T = \frac{\sigma^2(K,T)K^2}{2} v_{KK} - rKv_K, \quad T > t \tag{1.5}$$

with initial condition

$$v(K,t) = (x-K)^+ \tag{1.6}$$

In particular,

$$\sigma(K,T) = \sqrt{2\left(\frac{v_T + rKv_K}{K^2 v_{KK}}\right)} \tag{1.7}$$

The equation (1.5) is called **Dupire's equation** and the formula (1.7) is called **Dupire's formula**. The marvelous result is due to Bruno Dupire (see references above). This tells us that *if*, for

fixed x and t, we know the option prices $u(x,t;K,T)$ for all strikes K and all expiration times T, then (1.7) gives us a unique local volatility function $\sigma(K,T)$. Notice this tells us the volatility in the future! Although we can't observe the stock in the future, observing option prices for various K and T tells us something about the volatility, assuming that those prices are based on model (0.3).

Proof: Before proving the theorem, we first recall some results about transition densities for diffusion processes. For a given $x \in \mathbb{R}^d$ and $s \in \mathbb{R}$, suppose that $Y_t^{x,s}$ is a stochastic process satisfying

$$\begin{aligned} dY_t &= b(Y_t,t)\,dt + \sigma(Y_t,t)\,dB_t, \quad t \geq s \\ Y_s^{x,s} &= x \end{aligned} \tag{1.8}$$

Under suitable assumptions on the coefficients b and σ, this process has a smooth transition density $p(x,s;y,t)$ that satisfies

$$P(Y_t^{x,s} \in A) = \int_A p(x,s;y,t)\,dy \tag{1.9}$$

for any Borel set $A \subset \mathbb{R}^d$.

For fixed y and t, $p(\cdot,\cdot;y,t)$ satisfies the **backward equation**

$$\partial_s p + \mathcal{A}_{x,s} p = 0, \quad t > s \tag{1.10}$$

where $\mathcal{A}_x$ is the differential operator defined by

$$\mathcal{A}_{x,s} f := \frac{1}{2}\sum_{ij} a_{ij}(x,s) f_{x_i x_j} + b(x,s)\cdot \nabla_x f$$

Moreover,

$$\lim_{s \nearrow t} p(x,s;y,t) = \delta_y(x), \tag{1.11}$$

the Dirac delta distribution centered at y. If Y_t represents a diffusing particle, then roughly speaking, this backward equation describes "how the particles got to where they are today."

For fixed x and s, $p(x,s;\cdot,\cdot)$ satisfies the **forward equation**

$$\partial_t p = \mathcal{A}_{y,t}^* p, \quad t > s \tag{1.12}$$

where $\mathcal{A}_y^*$ is the **adjoint operator** defined by

$$\mathcal{A}_{y,t}^* g := \frac{1}{2}\sum_{ij} \frac{\partial^2}{\partial_{y_i}\partial_{y_j}}\left(a_{ij}(y,t) g\right) - \nabla_y \cdot \left(b(y,t) g\right) \tag{1.13}$$

and $a_{ij}(y,t) = \sigma\sigma^T$. This describes the evolution of the density forward in time. Moreover,

$$\lim_{t \searrow s} p(x,s;y,t) = \delta_x(y) \tag{1.14}$$

Now we prove the result of Dupire. Consider the function $w(x,t;K,T) = e^{r(T-t)}u(x,t;K,T) = \tilde{E}[g(X_T)|X_t = x]$, which satisfies

$$w_t + \frac{\sigma^2(x,t)x^2}{2}w_{xx} + rxw_x = 0, \quad t < T, x > 0 \tag{1.15}$$

with terminal data $w(x,T) = g(x) = (x-K)^+$. The solution may be expressed in terms of the transition density for X_t (under the risk-neutral measure):

$$\begin{aligned} w(x,t;K,T) = \tilde{E}[g(X_T)|X_t = x] &= \int_0^\infty g(y)p(x,t;y,T)\,dy \\ &= \int_K^\infty (y-K)p(x,t;y,T)\,dy \end{aligned} \tag{1.16}$$

Now differentiate this expression twice with respect to K:

$$\partial_K w = -\int_K^\infty p(x,t;y,T)\,dy, \tag{1.17}$$

and

$$\partial_K^2 w = p(x,t;K,T). \tag{1.18}$$

So, w_{KK} is the transition density $p(x,t;K,T)$ corresponding to the stock price process (under the risk-neutral measure). Therefore, as a function of K and T, $p = w_{KK}$ must satisfy the forward equation

$$p_T = \frac{\partial^2}{\partial K^2}\left(\frac{\sigma^2(K,T)K^2}{2}p\right) - \frac{\partial}{\partial K}(rKp), \quad T > t \tag{1.19}$$

Here x and t are fixed. Thus,

$$(w_{KK})_T = \frac{\partial^2}{\partial K^2}\left(\frac{\sigma^2(K,T)K^2}{2}w_{KK}\right) - \frac{\partial}{\partial K}(rKw_{KK}), \quad T > t \tag{1.20}$$

Now we want to obtain an equation for w not for w_{KK}, so we integrate twice with respect to K to obtain

$$w_T = \frac{\sigma^2(K,T)K^2}{2}w_{KK} - rKw_K + rw + aK + c \tag{1.21}$$

for some constants a and c. In this step, notice that

$$\begin{aligned} \int_0^K \int_0^{K_2} \frac{\partial}{\partial K_1}(rK_1 w_{K_1K_1})\,dK_1\,dK_2 &= \int_0^K rK_2 w_{K_2K_2}\,dK_2 \\ &= \int_0^K \partial_{K_2}(rK_2w_{K_2}) - rw_{K_2}\,dK_2 \\ &= rKw_K - rw + const \end{aligned} \tag{1.22}$$

Supposing that as $K \to \infty$, w and its derivatives go to zero, this implies that $a = c = 0$. Thus,

$$w_T = \frac{\sigma^2(K,T)K^2}{2}w_{KK} - rKw_K + rw \tag{1.23}$$

Recalling that $w = e^{r(T-t)}u$, we now see that u satisfies Dupire's equation

$$u_T = \frac{\sigma^2(K,T)K^2}{2}u_{KK} - rKu_K, \quad T > t, K > 0. \tag{1.24}$$

This is an initial value problem. The initial condition is $u(x,t;K,t) = (x-K)^+$ (in the initial condition, K is the variable, x is fixed).

As $K \to \infty$, $u(x,t;K,T) \to 0$, and as $K \to 0$, $u(x,t;K,T) \to x$, since

$$\lim_{K\to 0} u(x,t;K,T) = \tilde{E}[e^{-r(T-t)}X_T|X_t = x] = x \tag{1.25}$$

(Recall that the discounted price is a martingale under $\tilde{P}$). □

Digital options

A **digital option** or **binary option** is one with payoff

$$g_b(x) = \begin{cases} 0, & x < K, \\ b & x \geq K \end{cases} \tag{1.26}$$

at the expiration time T. Here $b > 0$ is some fixed constant. One can also calibrate the model (0.3)) using these options.

Theorem 7.1.2 *Suppose that the function $u(x,t;K,T)$ solves (0.3) with terminal data*

$$g_b(x) = \begin{cases} 0, & x < K, \\ b & x \geq K \end{cases} \tag{1.27}$$

Then for fixed $x > 0$ and t, the function $v(K,T) = u(x,t;K,T)$ satisfies the initial value problem

$$v_T = \frac{\partial}{\partial K}\left(\frac{\sigma^2(K,T)K^2}{2}\frac{\partial}{\partial K}v\right) - rKv_K - rv, \quad T > t, K > 0 \tag{1.28}$$

with initial condition

$$v(x,t;K,t) = \begin{cases} 0, & x < K, \\ b & x \geq K \end{cases} \tag{1.29}$$

This was proved in the paper by O. Pironneau, cited above. Here we present a slightly different argument following the one presented above for the plain vanilla call.

Proof: Consider the function $w(x,t;K,T) = e^{r(T-t)}u(x,t;K,T) = \tilde{E}[g_d(X_T)|X_t = x]$, which satisfies

$$w_t + \frac{\sigma^2(x,t)x^2}{2}w_{xx} + rxw_x = 0, \quad t < T, x > 0 \tag{1.30}$$

with terminal data $w(x,T) = g_b(x)$, defined above. The solution may be expressed in terms of the transition density for X_t (under the risk-neutral measure):

$$\begin{aligned} w(x,t;K,T) = \tilde{E}[g_b(X_T)|X_t = x] &= \int_0^\infty g_b(y)p(x,t;y,T)\,dy \\ &= \int_K^\infty bp(x,t;y,T)\,dy \end{aligned} \tag{1.31}$$

Now differentiate this expression *once* with respect to K:

$$\partial_K w = -bp(x,t;K,T). \tag{1.32}$$

So, $-b^{-1}w_K$ is the transition density $p(x,t;K,T)$ corresponding to the stock price process (under the risk-neutral measure). Therefore, as a function of K and T, $p = -b^{-1}w_K$ must satisfy the forward equation

$$p_T = \frac{\partial^2}{\partial K^2}\left(\frac{\sigma^2(K,T)K^2}{2}p\right) - \frac{\partial}{\partial K}(rKp), \quad T > t \tag{1.33}$$

Here x and t are fixed. Thus,

$$(w_K)_T = \frac{\partial^2}{\partial K^2}\left(\frac{\sigma^2(K,T)K^2}{2}w_K\right) - \frac{\partial}{\partial K}(rKw_K), \quad T > t \tag{1.34}$$

Now we want to obtain an equation for w not for w_K, so we integrate **once** with respect to K to obtain

$$w_T = \frac{\partial}{\partial K}\left(\frac{\sigma^2(K,T)K^2}{2}w_K\right) - rKw_K + c \tag{1.35}$$

for some constants a and c.

Supposing that as $K \to \infty$, w and its derivatives go to zero, this implies that $c = 0$. Thus,

$$w_T = \frac{\partial}{\partial K}\left(\frac{\sigma^2(K,T)K^2}{2}\frac{\partial}{\partial K}w\right) - rKw_K \tag{1.36}$$

Recalling that $w = e^{r(T-t)}u$, we now see that u satisfies the equation

$$u_T = \frac{\partial}{\partial K}\left(\frac{\sigma^2(K,T)K^2}{2}\frac{\partial}{\partial K}u\right) - rKu_K - ru, \quad T > t, K > 0. \tag{1.37}$$

This is an initial value problem. The initial condition is

$$u(x,t;K,t) = \begin{cases} 0, & x < K, \\ b & x \geq K \end{cases} \tag{1.38}$$

(in the initial condition, K is the variable, x is fixed).

As $K \to \infty$, $u(x,t;K,T) \to 0$, and as $K \to 0$, $u(x,t;K,T) \to x$, since

$$\lim_{K\to 0} u(x,t;K,T) = \tilde{E}[e^{-r(T-t)}b|X_t = x] = e^{-r(T-t)}b \tag{1.39}$$

□

7.2 Model Calibration with Dupire's Equation

Here we explain one approach to model calibration using Dupire's equation. This is the approach presented in the paper by Yves Achdou and Olivier Pironneau, cited above. These notes are intended to give a high-level explanation and motivation of the main steps of the calibration scheme.

By **model calibration**, I mean solving the following problem: given current stock price x_0 at time t_0, and given strikes $\{K_i\}_{i=1}^N$, expiration times $\{T_i\}_{i=1}^N$, and market prices $\{c_i\}_{i=1}^N$ of European call options corresponding to these parameters, can we find a local volatility function $\sigma(x,t)$ such that

$$u(x_0, t_0; T_i, K_i, \sigma) = c_i \tag{2.40}$$

holds for each i? In other words, can we make the model fit the observed data? The function $u(x_0, t_0; K_i, T_i, \sigma)$ denotes a solution to the model (0.3). Since x_0 and t_0 are fixed, we will use $u(K_i, T_i)$ or just u_i for shorter notation. If the number of data points N is large, then working directly with (0.3) (or a discretized version) might be impractical. Just to compare u_i with c_i for each i requires solving the partial differential equation N times! Dupire's equation offers an efficient alternative. Recall that $u(K,T)$ as a function of K and T solves Dupire's equation

$$u_T = \frac{\sigma^2(K,T)K^2}{2}u_{KK} - rKu_K, \quad T > t_0 \tag{2.41}$$

with initial condition

$$u(K,t) = (x-K)^+ \tag{2.42}$$

Right away we see how Dupire's equation makes model calibration more efficient: comparing u_i with c_i for each $i = 1, \ldots, N$ requires solving Dupire's equation only once, and evaluating the solution at the N points (K_i, T_i).

For simplicity let us suppose that $t_0 = 0$; otherwise we could just replace T with $T' = T - t_0$. The domain for $u(K,T)$ is infinite. In practice, we truncate the domain to be $(K,T) \in Q := [0, \bar{K}] \times [0, \bar{T}]$ and solve the boundary value problem

$$\begin{aligned} u_T &= \frac{\sigma^2(K,T)K^2}{2}u_{KK} - rKu_K, \qquad (K,T) \in Q \\ u(K,0) &= (x-K)^+, \quad K \in [0,\bar{K}] \quad \text{(initial condition)} \\ u(0,T) &= x, \quad T \in [0,\bar{T}] \quad \text{(boundary condition)} \\ u(\bar{K},T) &= 0, \quad T \in [0,\bar{T}] \quad \text{(boundary condition)} \end{aligned} \tag{2.43}$$

The boundary conditions are motivated by (1.25) and preceding comments. Suppose we can solve Dupire's equation by some numerical scheme and compare our model to market data . . . what next? How do we change σ to get a better fit to the data?

The method described in the Achdou-Pironneau paper is called **regularized least squares approximation**. Defining $\eta = \sigma^2$, we compute the optimal volatility function $\sigma^*(K,T) = \sqrt{\eta^*}$ by solving a minimization problem:

$$J(\eta^*) = \min_{\eta} J(\eta) \tag{2.44}$$

where $J(\eta)$ is the objective function

$$J(\eta) = \left[\sum_{i=1}^{N}(u(K_i,T_i) - c_i)^2\right] + J_r(\eta). \tag{2.45}$$

The terms $u(K_i, T_i)$ in the objective function depend on $\eta = \sigma^2$ in a nonlinear way (through solution of the PDE). The term inside the sum penalizes discrepancies between the model and the observed data. The functional $J_r(\eta)$ is called a regularization term. A simple choice for this term might be

$$J_r(\eta) = \int_Q \epsilon |\nabla \eta|^2 \, dK \, dT, \tag{2.46}$$

but there are many other possibilities. The regularization is needed to make the problem tractable. Since we have only finitely many data points c_i there could be multiple solutions to the inverse problem. Not all of the possible choices of σ^* will be "nice functions"– smooth and bounded. So, the regularization term penalizes the bad possibilities. With the regularization term, minimizing the objective functional is a trade-off between having a good fit to the data and having a solution that has nice properties. This technique is used in many inverse problems and is sometimes called Tychonov regularization.

Here is the optimization algorithm used to solve (2.44). We construct a sequence of functions η^m, $m = 1, 2, \ldots$ by the following procedure:

1. Start with $\eta^m(K, T)$. This is a function of K and T, perhaps discretized in some way.
2. Compute $J(\eta^m)$ by computing $u(K_i, T_i, \eta^m)$ for each $i = 1, \ldots, N$. This requires only *one* solution of (2.43).
3. Find a function $d\eta$ so that $J(\eta^m + d\eta) < J(\eta^m)$.
4. Repeat with $\eta^{m+1} = \eta^m + d\eta$ until you are satisfied with the result.

In practice, the function $\eta(K, T)$ is discretized. For example, we might represent η as a continuous, piecewise linear function, determined by its values at finitely many pre-determined points. In this way, we may represent η as a vector in $\mathbb{R}^n$ for some n, perhaps very large. Furthermore, we restrict η to be uniformly bounded and uniformly positive.

One method of choosing the step direction $d\eta$ is the **method of steepest descent**: choose $d\eta$ to be a multiple of the gradient $\nabla_\eta J(\eta)$. This means using the iteration scheme

$$\eta^{m+1} = \eta^m - \rho \nabla_\eta J(\eta^m) \tag{2.47}$$

where $\rho > 0$ is a constant which could be chosen at each step to guarantee that the iterations lead to a decrease in the objective function. What do we really mean by $\nabla_\eta J$? If η is represented by a finite-dimensional vector (for example, the values of η at the nodal points in a finite-element representation), then we may interpret $\nabla_\eta J(\eta^m)$ to mean the gradient of the nonlinear functional J with respect to this vector. On the other hand, one can make sense of $\nabla_\eta J$ even when η is infinite dimensional (via the Frechét derivative). In this case, we say $\nabla_\eta J(\eta)$ is the function of (K, T) such that

$$J(\eta + d\eta) = J(\eta) + \int_Q (\nabla_\eta J(\eta) d\eta) \, dK \, dT + o(\|d\eta\|) \tag{2.48}$$

holds for all $d\eta$, where $\|d\eta\|$ denotes a norm on the function space (Hilbert space) in which η and $d\eta$ reside. In other words, to leading order $J(\eta + d\eta)$ is approximately the linear function of $d\eta$ given by

$$J(\eta + d\eta) \approx J(\eta) + \langle \nabla_\eta J(\eta), d\eta \rangle \tag{2.49}$$

where we use $\langle f, g \rangle$ to denote the inner product $\int_Q fg \, dK dT$.

Computing $\nabla_\eta J(\eta)$

We need to compute

$$\nabla_\eta J(\eta) = \sum_i 2(u_i - c_i)\nabla_\eta u_i + \nabla_\eta J_r(\eta). \tag{2.50}$$

or for a given "step direction" $d\eta$

$$\langle \nabla_\eta J(\eta), d\eta\rangle = \langle \sum_i 2(u_i - c_i)\nabla_\eta u_i, d\eta\rangle + \langle \nabla_\eta J_r(\eta), d\eta\rangle. \tag{2.51}$$

For each i, $u_i = u(K_i, T_i, \eta)$ is a function of $\eta = \eta(K,T)$, and $\nabla_\eta u_i$ is the gradient (or Frechét derivative) of this function with respect to η. For each i, $u_i(\eta)$ is the value of a solution to a PDE that depends on η. So $\nabla_\eta u_i$ tells us how the solution at the point (K_i, T_i) changes with respect to the parameter η. By definition,

$$u(K_i, T_i, \eta + d\eta) = u(K_i, T_i, \eta) + \langle \nabla_\eta u_i, d\eta\rangle + (\text{higher order terms}) \tag{2.52}$$

Since the equation solved by u is linear we see that

$$\begin{aligned} \mathcal{L}_D^\eta \left(u(K,T,\eta + d\eta) - u(K,T,\eta)\right) &= \frac{d\eta}{2}K^2\frac{\partial^2}{\partial K^2}u(K,T,\eta + d\eta) \\ &= \frac{d\eta}{2}K^2\frac{\partial^2}{\partial K^2}u(K,T,\eta) + h.o.t. \end{aligned} \tag{2.53}$$

where $\mathcal{L}_D^\eta$ is the linear differential operator

$$\mathcal{L}_D^\eta u := u_T - \frac{\eta(K,T)K^2}{2}u_{KK} + rKu_K \tag{2.54}$$

appearing in Dupire's equation. Also, $u(K,T,\eta + d\eta) - u(K,T,\eta) = 0$ on the boundaries and initially. Therefore, $\langle \nabla_\eta u_i, d\eta\rangle = w(K_i, T_i)$, where w solves

$$\begin{aligned} w_T - \frac{\eta(K,T)K^2}{2}w_{KK} + rKw_K &= \frac{d\eta}{2}K^2\frac{\partial^2}{\partial K^2}u(K,T,\eta) \\ w(K,0) = 0, \quad K \in [0,\bar{K}] &\quad \text{(initial condition)} \\ w(0,T) = 0, \quad T \in [0,\bar{T}] &\quad \text{(boundary condition)} \\ w(\bar{K},T) = 0, \quad T \in [0,\bar{T}] &\quad \text{(boundary condition)} \end{aligned} \tag{2.55}$$

This could be problematic: Given a small change $d\eta$, we can compute the leading order change in u_i, $\langle \nabla_\eta u_i, d\eta\rangle$, by solving (2.55). However, we have to solve a PDE for each choice of $d\eta$. So if η is high-dimensional, computing the gradient $\nabla_\eta u_i$ can be very expensive, since we must know how u changes as we step in many possible directions $d\eta$. The solution to this problem is to use what is called an **adjoint** method.

To motivate this method, consider the linear algebra problem of solving $Ax = b$, for many possible right-hand-sides, $b_1, b_2, \ldots$. Here x is a vector and A is an invertible square matrix. Suppose that what we really care about is not x itself but some functional of x, say $g \cdot x = \langle g, x\rangle$ for some given vector g. Letting $G = A^{-1}$, the solution is $x = Gb$. So, the functional of interest is

$$\langle g, x\rangle = \langle g, Gb\rangle = \langle G^T g, b\rangle \tag{2.56}$$

So if we are able to compute $G^T g$ one time, then evaluating the quantity $\langle g, x\rangle$ for $M > 1$ choices of b means computing M matrix-vector products $\langle (G^T g), b\rangle$ (rather than solving the equation $Ax = b$, M times).

The method described in the Achdou-Pironneau paper is exactly this. Instead of a matrix equation we have a PDE involving a linear differential operator. The "functional" of interest is:

$$\langle \sum_i 2(u_i - c_i)\nabla_\eta u_i, d\eta\rangle = \langle \sum_i 2(u_i - c_i)\delta_{K_i,T_i}(K,T), w\rangle \tag{2.57}$$

where $\delta_{K_i,T_i}(K,T)$ denotes the Dirac-delta centered at (K_i, T_i). So, $\sum_i 2(u_i - c_i)\delta_{K_i,T_i}$ corresponds to the g in the simple linear algebra example. $G = A^{-1}$ corresponds to solving the PDE, while G^T corresponds to solving the adjoint problem and w corresponds to Gb. Specifically, we solve the adjoint problem

$$\begin{aligned} \partial_T P + \frac{\partial^2}{\partial K^2}\left(\frac{\eta K^2}{2}P\right) + \frac{\partial}{\partial K}(rKP) &= \sum_i 2(u_i - c_i)\delta_{K_i,T_i} \\ P(\bar{T}, K) &= 0 \\ P(T, \bar{K}) &= 0 \end{aligned} \tag{2.58}$$

Then

$$\begin{aligned} \langle \nabla_\eta J, d\eta\rangle &= \langle \sum_i 2(u_i - c_i)\delta_{K_i,T_i}, w\rangle + \langle \nabla J_r, d\eta\rangle \\ &= \int_Q \left(\partial_T P + \frac{\partial^2}{\partial K^2}\left(\frac{\eta K^2}{2}P\right) + \frac{\partial}{\partial K}(rKP)\right) w(K,T)\, dK dT + \langle \nabla J_r, d\eta\rangle \\ &= -\int_Q P\left(w_T - \frac{\eta(K,T)K^2}{2}w_{KK} + rKw_K\right) dK dT + \langle \nabla J_r, d\eta\rangle \\ &= -\int_Q \frac{d\eta}{2} P K^2 \frac{\partial^2}{\partial K^2} u(K,T,\eta)\, dK dT + \langle \nabla J_r, d\eta\rangle \end{aligned} \tag{2.59}$$

The last two steps are integration by parts and using the equation solved by w and the boundary conditions. Notice that $u = u(K,T,\eta)$ and $P = P(K,T,\eta)$ do not depend on $d\eta$!

The algorithm

Consequently, the steepest decent algorithm becomes:

1. Compute $u(K,T,\eta^m)$ by solving (2.43) with $\sigma^2 = \eta$.
2. Compute $P(K,T,\eta^m)$ by solving (2.58).
3. Compute $\nabla_\eta J$ by solving

$$\langle \nabla_\eta J, d\eta\rangle = -\int_Q \frac{d\eta}{2} P K^2 \frac{\partial^2}{\partial K^2} u(K,T,\eta)\, dK dT + \langle \nabla J_r, d\eta\rangle$$

(computation of the last terms depends on the type of regularization chosen). When η is discretized, this becomes a linear system of equations.

4. Take a step:
$$\eta^{m+1} = \pi(\eta^m - \rho \nabla_\eta J(\eta^m))$$

5. Measure $e = \|\nabla_\eta J\|$. If e is smaller than a predetermined tolerance, stop – we are close to a minimizer. Otherwise, repeat using η^{m+1}.

The operator π is a projection onto the space in which η should lie. For example, π might ensure that the new iteration is bounded and uniformly positive.

See the paper by Achdou and Pironneau for more details about numerical discretization of the equations, the regularization term, and some numerical results.

Chapter 8

Optimal Control and the HJB equation

References:

- L.C. Evans, *Partial Differential Equations*, AMS 1998, Section 10.3
- W. Fleming and R. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, 1975.
- W. Fleming and M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer, 1993.
- M. Bardi, I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkäuser, 1997.
- M. Soner, *Controlled Markov processes, viscosity solutions and applications to mathematical finance*, Lecture Notes in Mathematics **1660**, Springer, Berlin, 1997.
- T. Leung, R. Sircar, and T. Zariphopoulou, *Credit Derivatives and Risk Aversion*, Preprint, 2007.

8.1 Deterministic Optimal Control

Consider the following abstract optimization problem. Let $y(s) : [t, T] \to \mathbb{R}^d$ denote the **state of a system** at time $s \in [t, T]$. This vector function could represent many things like the position and orientation of an aircraft, the amount of capital available to a government, or the wealth of an individual investor. We'll suppose that $y(s)$ satisfies the ordinary differential equation

$$\begin{aligned} y'(s) &= f(y(s), \alpha(s)), \quad s \in [t, T] \\ y(t) &= x \in \mathbb{R}^d \end{aligned} \tag{1.1}$$

(If $d > 1$, this is a system of ODEs.) The function $f(y, \alpha) : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$ models the system dynamics. We'll suppose that f is bounded and Lipschitz continuous. The function $\alpha(s) : [t, T] \to A$

is called a **control**. The control takes values in the set A, a compact subset of $\mathbb{R}^m$. The set of all possible controls or **admissible controls** will be denoted by $\mathcal{A}_{t,T}$:

$$\mathcal{A}_{t,T} = \{\alpha(s) : [t,T] \to A \mid \ \alpha(s) \text{ is measureable}\}. \tag{1.2}$$

When the dependence on t and T is clear from the context, we will simply use $\mathcal{A}$ instead. By choosing α we have some control over the course of the system $y(t)$. For example, in a mechanics application $\alpha(s)$ might represent a throttle control, which determines how much thrust comes from an engine. Or, in an economics application α might represent the rate at which economic resources are consumed.

So, we choose a control $\alpha(\cdot)$ and the system evolves according to (1.1); we'd like to control the system in an optimal way, in the following sense. Suppose that the function $g(x) : \mathbb{R}^d \to \mathbb{R}$ represents a **final payoff**; this is a reward which depends on the final state of the system at time T. Also, the function $r(x,\alpha) : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ represents a **running payoff** or **running cost**. If $r > 0$, this would represent additional path-dependent payoff; if $r < 0$, this would represent operational costs incurred before the final payoff at time T. Given the initial state of the system $y(t) = x$, the optimization problem is to find an optimal control $\alpha^*(\cdot)$ that maximizes net profit:

$$J_{x,t}(\alpha^*) = \max_{\alpha(\cdot)\in\mathcal{A}} J_{x,t}(\alpha) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \tag{1.3}$$

If $r < 0$, this may be thought of as the optimal balance between payoff and running costs. Although, we may be able to control the system (by choosing α) so that the final payoff $g(y(T))$ is large, it may be very expensive to arrive at this state. So, we want to find the optimal control that balances these competing factors.

Even if an optimal control does not exist, we may study the function

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} J_{x,t}(\alpha) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \tag{1.4}$$

This function is called the **value function** associated with the control problem. It depends on x and t through the initial conditions defining $y(s)$. There are many interesting mathematical questions related to this optimization problem. For example:

1. For given (x,t), is there an optimal control α^*?
2. If so, how can we compute it?
3. How does the value function u depend on x and t? Does it satisfy a PDE?

Some Examples

Here are two examples which fit into this abstract framework.

Example 1: Suppose you want to drive a boat from position x_0 at time t to a position x_f at time T. Let $x(s)$ denote the position of the boat, $v(s)$ denote the velocity, then a simple model for

the boat dynamics might be

$$\begin{aligned} x'(s) &= v(s) \\ v'(s) &= \frac{\alpha(s)}{(m_1 + m(s))} - \beta(v(s)) \\ m'(s) &= -k|\alpha(s)| \end{aligned}$$

Here $m(s)$ is the mass of the boat's fuel, m_1 is the boat's dry weight, the function $\beta(v) \geq 0$ models drag as the boat moves through the water. The vector $\alpha(s)$ represents a throttle and direction control, and its magnitude is proportional to the rate of fuel consumption (k is a proportionality constant). The acceleration is also proportional to the throttle control parameter.

How should we steer the boat in order to minimize fuel consumption? In this setting, the system state $y(s)$ is the vector $y(s) = (x(s), v(s), m(s))$. One way to model this problem would be to find

$$\max_{\alpha} J_{x_0,t}(\alpha) = \max_{\alpha} \left[m(T) + p(x(T)) \right] \tag{1.5}$$

Here $p \leq 0$ might be a function satisfying $p(x) = 0$ if x is close to x_f and $p(x) \ll -1$ if x is far from x_f. So, although we don't need to land precisely at x_f, there is a big penalty for leaving the boat far from x_f. There is no "running cost" in this model. Notice that it is possible for $m(s)$ to become negative, which is non-physical. We could fix this modeling issue by modifying the equations appropriately or by applying an additional **state constraint** of the form $0 \leq m(s)$.

Example 2: Here is a variant of a classic example studied by F. P. Ramsey (see *The Economic Journal*, Vol. 38, No. 152, 1928). The problem is to determine how much of a nation's resources should saved and how much should be consumed. Let $c(s)$ denote the rate of capital consumption, let $p(s)$ denote the rate of capital production, and let $k(s)$ denote the amount of capital at time s. Then the rate change in capital is the difference between the rates of production and consumption:

$$k'(s) = p(s) - c(s). \tag{1.6}$$

Suppose that the production is related to capital and consumption as $p(s) = P(c(s), k(s))$. Suppose also that consumption is related to capital according to $c(s) = \alpha(s)C(k(s))$, where $\alpha(s)$ is a control. Therefore,

$$k'(s) = P(\alpha(s)C(k(s)), k(s)) - \alpha(s)C(k(s)). \tag{1.7}$$

Given current level of capital $k(t) = k_0$, we'd like to choose a level of consumption (by choosing α) which maximizes the total utility; this goal might be modeled by the optimal control problem

$$\begin{aligned} \max_{\alpha} J_{k_0,t}(\alpha) &= \max_{\alpha} \left[\int_t^T U(c(s))\, ds + U_T(k(T)) \right] \\ &= \max_{\alpha} \left[\int_t^T U(\alpha(s)C(k(s)))\, ds + U_T(k(T)) \right] \end{aligned} \tag{1.8}$$

Here U is a utility function, and U_T models some payoff representing the utility of having left-over capital $k(T)$ at time T.

8.2 The Dynamic Programming Principle

Theorem 8.2.1 *Let $u(x,t)$ be the value function defined by (1.4). If $t < \tau \leq T$, then*

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^\tau r(y(s),\alpha(s))\, ds + u(y(\tau),\tau) \right] \tag{2.9}$$

The relation (4.48) is called the **Dynamic Programming Principle**, and it is a fundamental tool in the analysis of optimal control problems. It says that if we know the value function at time $\tau > t$, we may determine the value function at time t by optimizing from time t to time τ and using $u(\cdot,\tau)$ as the payoff. Roughly speaking, this is reminiscent of the Markov property of a stochastic process, in the sense that if we know $u(x,\tau)$ we can determing $u(\cdot,t)$ for $t < \tau$ without any other information about the control problem beyond time τ (ie. times $s \in [\tau, T]$). (More precisely, it means that $u(x,t)$ satisfies what is called a semi-group property.)

Proof of Theorem 8.2.1: At the heart of this proof of the Dynamic Programming Principle is the observation that any admissible control $\alpha \in \mathcal{A}_{t,T}$ is the combination of a control in $\mathcal{A}_{t,\tau}$ with a control in $\mathcal{A}_{\tau,T}$. We will express this relationship as

$$\mathcal{A}_{t,T} = \mathcal{A}_{t,\tau} \oplus \mathcal{A}_{\tau,T} \tag{2.10}$$

This notation $\oplus$ means that if $\alpha_t(s) \in \mathcal{A}_{t,\tau}$ and $\alpha_\tau(s) \in \mathcal{A}_{\tau,T}$, then the control defined by splicing α_t and α_τ according to

$$\alpha(s) = (\alpha_t \oplus \alpha_\tau)(s) := \begin{cases} \alpha_t(s), & s \in [t,\tau] \\ \alpha_\tau(s), & s \in [\tau, T] \end{cases} \tag{2.11}$$

is an admissible control in $\mathcal{A}_{t,T}$. On the other hand, if we have $\alpha \in \mathcal{A}_{t,T}$, then by restricting the domain of α to $[t,\tau]$ we obtain an admissible control in $\mathcal{A}_{t,\tau}$. Similarly, by restricting the domain of α to $[\tau, T]$ we obtain an admissible control in $\mathcal{A}_{\tau,T}$.

The function u is defined as

$$\begin{aligned} u(x,t) &= \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \\ &= \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^\tau r(y(s),\alpha(s))\, ds + \int_\tau^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \end{aligned} \tag{2.12}$$

Notice that the first integral on the right depends only on y and α up to time τ, while the last two terms depend on the values of y and α after time τ. Since a control $\alpha \in \mathcal{A}_{t,T}$ may be decomposed as $\alpha = \alpha_1 \oplus \alpha_2$ with $\alpha_1 \in \mathcal{A}_{t,\tau}$ and $\alpha_2 \in \mathcal{A}_{\tau,T}$, we may maximize over each component in the decomposition:

$$\begin{aligned} u(x,t) &= \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^\tau r(y(s),\alpha(s))\, ds + \int_\tau^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \\ &= \max_{\substack{\alpha_1\in\mathcal{A}_{t,\tau}, \\ \alpha_2\in\mathcal{A}_{\tau,T}, \\ \alpha=\alpha_1\oplus\alpha_2}} \left[\int_t^\tau r(y(s),\alpha(s))\, ds + \int_\tau^T r(y(s),\alpha(s))\, ds + g(y(T)) \right] \end{aligned}$$

On the right hand side, the system state $y(t)$ is determined by (1.1) with $\alpha = \alpha_1 \oplus \alpha_2 \in \mathcal{A}_{t,T}$. Therefore, we may decompose the system state as $y(s) = y_1 \oplus y_2$ where $y_1(s) : [t, \tau] \to \mathbb{R}^d$ and $y_2(s) : [\tau, T] \to \mathbb{R}^d$ are defined by

$$\begin{aligned} y_1'(s) &= f(y_1(s), \alpha_1(s)), \quad s \in [t, \tau] \\ y_1(t) &= x \end{aligned}$$

and

$$\begin{aligned} y_2'(s) &= f(y_2(s), \alpha_2(s)), \quad s \in [\tau, T] \\ y_2(\tau) &= y_1(\tau) = y(\tau) \end{aligned}$$

Here we use $\oplus$ to denote the splicing or gluing of y_1 and y_2 to create $y(t) : [t, T] \to \mathbb{R}^d$. Therefore,

$$u(x,t) = \max_{\alpha_1 \in \mathcal{A}_{t,\tau}} \max_{\substack{\alpha_2 \in \mathcal{A}_{\tau,T} \\ y_2(\tau) = y_1(\tau)}} \left[\int_t^\tau r(y_1(s), \alpha_1(s))\, ds + \int_\tau^T r(y_2(s), \alpha_2(s))\, ds + g(y_2(T)) \right]$$

where the initial point for $y_2(\tau)$ is $y_2(\tau) = y_1(\tau)$. Observe that y_1 depends only on x and α_1, not on y_2 or α_2. Since the first integral depends only on α_1 and y_1, this may be rearranged as

$$\begin{aligned} u(x,t) &= \max_{\alpha_1 \in \mathcal{A}_{t,\tau}} \max_{\substack{\alpha_2 \in \mathcal{A}_{\tau,T} \\ y_2(\tau) = y_1(\tau)}} \left[\int_t^\tau r(y_1(s), \alpha_1(s))\, ds + \int_\tau^T r(y_2(s), \alpha_2(s))\, ds + g(y_2(T)) \right] \\ &= \max_{\alpha_1 \in \mathcal{A}_{t,\tau}} \left[\int_t^\tau r(y_1(s), \alpha_1(s))\, ds + \max_{\substack{\alpha_2 \in \mathcal{A}_{\tau,T} \\ y_2(\tau) = y_1(\tau)}} \left(\int_\tau^T r(y_2(s), \alpha_2(s))\, ds + g(y_2(T)) \right) \right] \\ &= \max_{\alpha_1 \in \mathcal{A}_{t,\tau}} \left[\int_t^\tau r(y_1(s), \alpha_1(s))\, ds + u(y_1(\tau), \tau) \right] \qquad \text{(using the definition of } u\text{)} \\ &= \max_{\alpha(\cdot) \in \mathcal{A}} \left[\int_t^\tau r(y(s), \alpha(s))\, ds + u(y(\tau), \tau) \right] \end{aligned} \tag{2.13}$$

This completes the proof. □

Notice that in this proof we have not assumed that an optimal control exists.

8.3 The Hamilton-Jacobi-Bellman Equation

How does the value function depend on x and t? Is it continuous in (x, t)? Is it differentiable? Does it satisfy a PDE? Unfortunately, the value function may not be differentiable, even in simple examples! Here is one interesting example of this fact. Suppose that $f(x, \alpha) = \alpha$, $g \equiv 0$, and $r(x, \alpha)$ is defined by

$$r(x, \alpha) = -\mathbb{I}_D(x) = \begin{cases} -1, & x \in D \\ 0, & x \in \mathbb{R}^d \setminus D \end{cases} \tag{3.14}$$

where $D \subset \mathbb{R}^d$ is some bounded set. Suppose that the set of admissible controls is defined by (5.53) with $A = \{|\alpha| \leq 1\}$. In this case, $y'(s) = \alpha(s)$ and $|y'(s)| \leq 1$. Therefore, the value function may be written as

$$u(x,t) = \max_{\substack{y:[t,T]\to\mathbb{R}^d \\ |y'|\leq 1,\, y(t)=x}} \left[\int_t^T -\mathbb{I}_D(y(s))\, ds \right] \tag{3.15}$$

Clearly $u(x,t) \leq 0$, and the optimum is obtained by paths that spend the least amount of time in the set D. If $x \in \mathbb{R}^d \setminus D$, then $u(x,t) = 0$, because we could take $y(s) = x$ for all $s \in [t,T]$. In this case, the system state doesn't change, so the integral is zero, which is clearly optimal. On the other hand, if $x \in D$ then the optimal control moves $y(s)$ to $\mathbb{R}^d \setminus D$ as quickly as possible and then stays outside D. Since $|y'(s)| \leq 1$, this implies that the value function is given explicitly by

$$u(x,t) = -\min\left((T-t), dist(x, \mathbb{R} \setminus D)\right) \tag{3.16}$$

where

$$dist(x, \mathbb{R} \setminus D) = \inf_{y \in \mathbb{R}\setminus D} |x - y|, \tag{3.17}$$

is the Euclidean distance from x to the outside of D. Albeit continuous, this function may not be differentiable! For example, suppose that $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ is the unit disk. In this case,

$$u(x,t) = \begin{cases} |x| - 1, & |x| \leq 1 \\ 0, & |x| \geq 1 \end{cases} \tag{3.18}$$

for $t \leq T - 1$. Thus $u(x,t)$ is not differentiable at the origin $x = (x_1, x_2) = (0,0)$ for $t < T - 1$.

So, in general, the value function may not be differentiable. However, one can still derive a PDE satisfied by the value function. If the value function is differentiable, this equation is satisfied in the classical sense. At points where the value function is not differentiable, one can show that the value function (assuming it is at least continuous) satisfies the PDE in a weaker sense. This weaker notion of "solution" is called a "viscosity solution" of the PDE. The theory of viscosity solutions is beyond the scope of this course, so we will formally compute as if the value function were actually differentiable. For more details about viscosity solutions see the references cited above.

For now, let us use the Dynamic Programming Principle to formally derive an equation solved by the value function $u(x,t)$. The Dynamic Programming Principle does not require differentiability of the value function; however, in our computations we assume that the value function is continuous and differentiable with respect to both x and t. The Dynamic Programming Principle tells us that

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^\tau r(y(s), \alpha(s))\, ds + u(y(\tau), \tau) \right] \tag{3.19}$$

To formally derive a PDE for u, we let $h \in (0, T-t)$ and set $\tau = t + h < T$. Then

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^{t+h} r(y(s), \alpha(s))\, ds + u(y(t+h), t+h) \right] \tag{3.20}$$

We'll assume that nearly optimal controls are approximately constant for $s \in [t, t+h]$.

First, consider the term $u(y(t+h), t+h)$. From the chain rule and our assumption that u is continuously differentiable in x and t, we conclude that

$$u(y(t+h), t+h) = u(y(t), t) + hy'(t) \cdot \nabla u(y(t), t) + hu_t(y(t), t) + o(h) \tag{3.21}$$

Recall that a function $\phi(h)$ is said to be $o(h)$ ("little oh of h") if $\lim_{h\to 0}(\phi(h)/h) = 0$. So, (3.21) says that $u(y(t+h), t+h)$ is equal to a linear funciton of h plus something that is $o(h)$ (i.e. higher order than h, but not necessarily $O(h^2)$). Therefore,

$$\begin{aligned} u(y(t+h), t+h) &= u(y(t), t) + hf(y(t), \alpha(t)) \cdot \nabla u(y(t), t) + hu_t(y(t), t) + o(h) \\ &= u(x, t) + hf(x, \alpha(t)) \cdot \nabla u(x, t) + hu_t(x, t) + o(h) \end{aligned} \tag{3.22}$$

Now, plug this into (3.20):

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^{t+h} r(y(s), \alpha(s))\, ds + u(x,t) + hf(x, \alpha(t)) \cdot \nabla u(x,t) + hu_t(x,t) + o(h) \right] \tag{3.23}$$

The term $u(x,t)$ may be pulled out of the maximum, so that it cancels with the left hand side:

$$0 = hu_t(x,t) + o(h) + \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^{t+h} r(y(s), \alpha(s))\, ds + hf(x, \alpha(t)) \cdot \nabla u(x,t) \right] \tag{3.24}$$

Now divide by h and let $h \to 0$.

$$0 = u_t(x,t) + \frac{o(h)}{h} + \max_{\alpha(\cdot)\in\mathcal{A}} \left[\frac{1}{h} \int_t^{t+h} r(y(s), \alpha(s))\, ds + f(x, \alpha(t)) \cdot \nabla u(x,t) \right] \tag{3.25}$$

If $\alpha(s)$ is continuous at t, then as $h \to 0$,

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} r(y(s), \alpha(s))\, ds = r(y(t), \alpha(t)) = r(x, \alpha(t)) \tag{3.26}$$

So, if the nearly optimal controls are approximately constant for $s \in [t, t+h]$, then by letting $h \to 0$ in (3.25) we conclude,

$$u_t(x,t) + \max_{a\in A} [r(x,a) + f(x,a) \cdot \nabla u(x,t)] = 0, \quad x \in \mathbb{R}^d,\, t < T \tag{3.27}$$

This equation is called the **Hamilton-Jacobi-Bellman equation**. The function $u(x,t)$ also satisfies the terminal condition

$$u(x,T) = g(x). \tag{3.28}$$

Notice that the HJB equation if a first-order, fully nonlinear equation, having the form

$$u_t + H(\nabla u, x) = 0$$

where the function H is defined by

$$H(p,x) = \max_{a\in A} [r(x,a) + f(x,a) \cdot p], \quad p \in \mathbb{R}^d \tag{3.29}$$

H is sometimes called the **Hamiltonian**. Observe that the essence of this derivation is the approximation of the system state by

$$y(t+h) \approx y(t) + hf(y(t), \alpha(t)) \tag{3.30}$$

and therefore, the "payoff" appearing in the dynamic programming principle is approximately

$$\begin{aligned} u(y(t+h), t+h) &\approx u(x,t) + (\Delta y) \cdot \nabla u(x,t) + h u_t(x,t) \\ &\approx u(x,t) + hf(x, \alpha(t)) \cdot \nabla u(x,t) + h u_t(x,t) \end{aligned} \tag{3.31}$$

In addition to telling us how the value function depends on x and t, this PDE suggests what the optimal control should be. Suppose $u(x,t)$ is differentiable and solves the PDE in the classical sense. Then

$$u_t + H(\nabla u, x) = 0 \tag{3.32}$$

where $H(p,x)$ is defined by (3.29). Then the optimal control is computed by finding $(y^*(s), \alpha^*(s))$ which satisfies

$$H(\nabla u(y^*(s), s), y^*(s)) = r(y^*(s), \alpha^*(s)) + f(y^*(s), a^*(s)) \cdot \nabla u(y^*(s), s) \tag{3.33}$$

and

$$\begin{aligned} \frac{d}{dt} y^*(s) = f(y^*(s), \alpha^*(s)) \quad s > t \\ y^*(t) = x \end{aligned} \tag{3.34}$$

A simple discrete approximation scheme would be to compute $\alpha^*(t)$ by maximizing:

$$r(x, a^*) + f(x, a^*) \cdot \nabla u(x,t) = \max_a \left[r(x,a) + f(x,a) \cdot \nabla u(x,t) \right] \tag{3.35}$$

Then set $\alpha^*(t) = a^*$, and $y^*(t+h) = x + hf(x, a^*)$. Then, replacing x with $y^*(t+h)$, t with $t+h$, we maximize again to compute $\alpha^*(t+h)$, and so forth. Of course, this scheme breaks down if u is not differentiable.

Examples

Example 1: In this example we want to maximize utility from consumption of our resources over time interval $[t, T]$. If $c(s)$ is the rate of consumption, then we model the utility gained from this consumption as

$$U = \int_t^T e^{-\mu s} \psi(c(s))\, ds \tag{3.36}$$

For example, we might choose ψ to be an increasing, concave function of c like $\psi(c) = c^\nu$ for some power $\nu \in (0,1)$. The factor $e^{-\mu s}$ is a discount factor. Let us suppose that the rate of consumption is $c(s) = \alpha(s) y(s)$, where y is our total wealth and α is a control. So, α is approximately the proportion of total wealth consumed in a unit of time. Instead of consuming resources, we might invest them in such a way that our total wealth satisfies the ode

$$y'(s) = qy(s) - c(s) = qy(s) - \alpha(s) y(s) \tag{3.37}$$

Without consumption, our wealth would grow exponentially at rate $q > 0$. So, investing some of our wealth might allow us to consume more in the long run. The control $\alpha(s)$ should satisfy some realistic constraints, for example $\alpha \in A = [0, \bar{a}]$. Thus, if x denotes the current wealth $y(t) = x$, the control problem is to find

$$u(x,t) = \max_{\alpha \in A} \left[\int_t^T e^{-\mu s} (\alpha(s) y(s))^\nu \, ds + g(y(T)) \right] \tag{3.38}$$

The function g models some utility of leaving an amount of unused wealth $y(T)$ at the final time. In this example, we can actually compute $y(s)$ explicitly:

$$y(s) = y(t) e^{q(t-s) - \int_t^s \alpha(\tau)\, d\tau} \tag{3.39}$$

To determine the associated HJB equation, notice that $f(x,a) = (q-a)x$ and $r(x,a,s) = e^{-\mu s}(ax)^\nu$. Therefore, we expect the value function to satisfy

$$u(x,t) + \max_{a \in A} \left[e^{-\mu s}(ax)^\nu + (q-a)x \cdot \nabla u \right] = 0. \tag{3.40}$$

perhaps in a weak sense (viscosity solutions) rather than a classical sense.

Example 2: Here's an example from engineering. Suppose that $r \equiv 0$, $f(x,a) = -v(x) + a$, and $A = \{|a| \leq \mu_0\}$. In this case, the HJB equation is

$$u_t(x,t) + \max_{|a| \leq \mu_0} \left[-v(x) \cdot \nabla u(x,t) + a \cdot \nabla u(x,t) \right] = 0, \quad x \in \mathbb{R}^d,\, t < T \tag{3.41}$$

It is easy to see that the optimal a is $a = \mu_0 (\nabla u)/|\nabla u|$, so that the equation becomes

$$u_t - v(x) \cdot \nabla u + \mu_0 |\nabla u| = 0 \tag{3.42}$$

The function $G(x,t) = u(x, T-t)$ satisfies

$$G_t + v(x) \cdot \nabla G = \mu_0 |\nabla G|. \tag{3.43}$$

In the combustion community, this equation is called the "G-equation" and is used in computational models of premixed turbulent combustion. The level set $\{G = 0\}$ is considered to be the flame surface. The parameter μ_0 corresponds to the laminar flame speed (i.e. the flame speed without the turbulent velocity field v).

8.4 Infinite Time Horizon

So far we have considered a deterministic control problem with **finite time horizon**. This means that the optimization involves a finite time interval and may involve a terminal payoff. One might also consider an optimization posed on an infinite time interval. Suppose that $y : [t, \infty) \to \mathbb{R}^d$ satisfies

$$\begin{aligned} y'(s) &= f(y(s), \alpha(s)), \quad s \in [t, \infty) \\ y(t) &= x \in \mathbb{R}^d \end{aligned} \tag{4.44}$$

Now the domain for the control is also $[t, \infty)$. We'll use $\mathcal{A} = \mathcal{A}_{t,\infty}$ for the set of admissible controls. For $x \in \mathbb{R}^d$, define the value function

$$u(x,t) = \max_{\alpha(\cdot)\in\mathcal{A}} J_{x,t}(\alpha) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_t^\infty e^{-\lambda s} r(y(s), \alpha(s))\, ds \right] \tag{4.45}$$

The exponential term in the integral is a discount factor; without it, the integral might be infinite. Notice that there is no terminal payoff, only running payoff. This optimal control problem is said to involve an **infinite time horizon**. Notice also that the value functions depends on t in a trivial way. Since r and f do not depend on t, we may change variables to see that

$$u(x,t) \quad = \quad e^{-\lambda t} u(x,0) \tag{4.46}$$

So, to find $u(x,t)$ it suffices to compute

$$u(x) = \max_{\alpha(\cdot)\in\mathcal{A}} J_x(\alpha) = \max_{\alpha(\cdot)\in\mathcal{A}} \left[\int_0^\infty e^{-\lambda s} r(y(s), \alpha(s))\, ds \right] \tag{4.47}$$

where $\mathcal{A} = \mathcal{A}_{0,\infty}$.

Theorem 8.4.1 (Dynamic Programming Principle) *Let $u(x)$ be the value function defined by (4.47). For any $x \in \mathbb{R}^d$ and $h > 0$,*

$$u(x) = \max_{\alpha(\cdot)\in\mathcal{A}_{0,h}} \left[\int_0^h e^{-\lambda s} r(y(s), \alpha(s))\, ds + e^{-\lambda h} u(y(h)) \right] \tag{4.48}$$

Proof: Exercise. □

Using the Dynamic Programming Principle, one can formally derive the HJB equation for the infinite horizon control problem. The equation is:

$$-\lambda u + \max_{a\in A} [r(x,a) + f(x,\alpha)\cdot \nabla u] = 0 \tag{4.49}$$

which has the form

$$-\lambda u + H(\nabla u, x) = 0 \tag{4.50}$$

with the Hamiltonian $H(p,x)$ defined by

$$H(p,x) = \max_{a\in A} [r(x,a) + f(x,a)\cdot p] \tag{4.51}$$

Exercise: Check these computations.

8.5 Brief Introduction to Stochastic Optimal Control

Thus far, we have considered deterministic optimal control in which the dynamic behaviour of the system state is deterministic. In a stochastic optimal control problem, the system state $y(s)$ is a stochastic process. Consequently, the controls also will be stochastic, since we may want to steer

the system in a manner that depends on the system's stochastic trajectory. To this end, we now suppose that the **system state** $Y_s(\omega) : [t,T] \times \Omega \to \mathbb{R}^d$ now satisfies the stochastic differential equation (or system of equations)

$$\begin{aligned} dY_s &= f(Y_s, \alpha_s, s)ds + \sigma(Y_t, \alpha_s, s)dB_s, \quad s \geq t \\ Y_t &= x, \quad a.s. \end{aligned} \tag{5.52}$$

where B_s is a n-dimensional Brownian motion defined on probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_s\}_{s\geq t}, P)$, and σ is a $d \times n$ matrix.

We control the system state through the **control process** $\alpha_s(\omega) : [t,T] \times \Omega \to \mathbb{R}^m$ which is adapted to the filtration $\{\mathcal{F}_s\}_{s\geq t}$. The set of admissible controls is now

$$\mathcal{A}_{t,T} = \{\alpha_s(\omega) : [t,T] \times \Omega \to A \mid \ \alpha_s \ \text{ is adapted to the filtration } \ \{\mathcal{F}_s\}_{s\geq t}\}. \tag{5.53}$$

The assumption that the controls are adapted means that we cannot look into the future; the control can only be chosen on the basis of information known up to the present time. Supposing that σ and f satisfy the usual bounds and continuity conditions, the stochastic process $Y_s(\omega)$ is uniquely determined by the initial condition $Y_t = x$ and the control process $\alpha_s(\omega)$.

Given a time $T > t$, the abstract stochastic optimal control problem is to maximize

$$\max_{\alpha \in \mathcal{A}_{t,T}} J_{x,t}(\alpha(\cdot)) = \max_{\alpha \in \mathcal{A}_{t,T}} E\left[\int_t^T r(Y_s, \alpha_s, s)\, ds + g(Y_T) \mid Y_t = x\right]$$

As before, the function $r(y, \alpha, s)$ represents a running payoff (or running cost, if $r < 0$), and g represents a terminal payoff (or terminal cost, if $g < 0$). Since the system state is a stochastic process, the net payoff is a random variable, and our goal is to maximize the expected payoff. Even if an optimal control process does not exist, we may define the value function to be

$$u(x,t) = \max_{\alpha \in \mathcal{A}_{t,T}} E\left[\int_t^T r(Y_s, \alpha_s, s)\, ds + g(Y_T)| \ Y_t = x\right] \tag{5.54}$$

Notice that the value function is *not* random.

Variations

There are variations on this theme. For example, we might add the possibility of a payoff based on a stopping criteria. In this case, we want to maximize:

$$\max_{\alpha \in \mathcal{A}_{t,T}} E\left[\int_t^{\gamma \wedge T} r(Y_s, \alpha_s, s)\, ds + g(Y_T)\mathbb{I}_{\{\gamma \geq T\}} + h(Y_\gamma)\mathbb{I}_{\{\gamma < T\}} \mid Y_t = x\right]$$

Here, the random variable $\gamma(\omega)$ is a stopping time. The function h represents some payoff that is incurred if $\gamma < T$ (or, this may represent a penalty if $h < 0$).

In (5.54) the time horizon is finite. One could also pose an optimal control problem on an infinite time horizon. For example, one might consider maximizing

$$\max_{\alpha \in \mathcal{A}} J_{x,t}(\alpha(\cdot)) = \max_{\alpha \in \mathcal{A}} E\left[\int_t^\gamma e^{-\lambda s} r(Y_s, \alpha_s)\, ds + e^{-\lambda \gamma} h(Y_\gamma)\right] \tag{5.55}$$

where γ is a stopping time.

8.5.1 Dynamic Programming Principle for Stochastic Control

For the stochastic control problem there is a Dynamic Programming Principle that is analogous to the DPP for deterministic control. Using the Markov Property of the stochastic process Y_t, one can easily prove the following:

Theorem 8.5.1 *Let $u(x,t)$ be the value function defined by (5.54). If $t < \tau \leq T$, then*

$$u(x,t) = \max_{\alpha \in \mathcal{A}_{t,\tau}} E\left[\int_t^\tau r(Y_s, \alpha_s, s)\, ds + u(Y_\tau, \tau) \mid Y_t = x\right] \tag{5.56}$$

Proof: Exercise. The idea is the same as in the case of deterministic control. Split the integral into two pieces, one over $[t,\tau]$ and the other over $[\tau, T]$. Then condition on $\mathcal{F}_\tau$ and use the Markov property, so that the second integral and the payoff may be expressed in terms of $u(Y_\tau, \tau)$.□

8.5.2 HJB equation

Using the Dynamic Programming Principle, one can formally derive a PDE for the value function $u(x,t)$. As in the case of deterministic optimal control, one must assume that the value function is sufficiently smooth. Because the dynamics are stochastic, we want to apply Itô's formula in the way that we used the chain rule to derive the HJB equation for deterministic control. Thus, this formal computation requires that the value function by twice differentiable.

From Itô's formula we see that

$$\begin{aligned} u(Y_\tau,\tau) - u(x,t) &= \int_t^\tau \left[u_t(Y_s,s) + f(Y_s,\alpha_s,s)\cdot \nabla u(Y_s,s)\right] ds \\ &\quad + \int_t^\tau \frac{1}{2}\sum_k \sum_{i,j} u_{x_i x_j}(Y_s,s)\sigma^{jk}(Y_s,\alpha_s,s)\sigma^{ik}(Y_s,\alpha_s,s)\, ds \\ &\quad + \int_t^\tau (\nabla u(Y_s,s))^T \sigma(Y_s,\alpha_s,s)\, dB_s \\ &= \int_t^\tau u_t(Y_s,s) + \mathcal{L}^\alpha u(Y_s,s)\, ds + \int_t^\tau (\nabla u(Y_s,s))^T \sigma(Y_s,\alpha_s,s)\, dB_s \end{aligned} \tag{5.57}$$

where $\mathcal{L}$ is the second order differential operator

$$\begin{aligned} \mathcal{L}^\alpha u &= f(y,\alpha,s)\cdot\nabla u(y,s) + \frac{1}{2}\sum_k\sum_{i,j} u_{y_i y_j}(y,s)\sigma^{jk}(y,\alpha_s,s)\sigma^{ik}(y,\alpha_s,s) \\ &= f(y,\alpha,s)\cdot\nabla u(y,s) + \frac{1}{2}\mathrm{tr}(D^2 u(y,s)\sigma(y,\alpha,s)\sigma^T(y,\alpha,s)) \end{aligned} \tag{5.58}$$

and $D^2 u$ is the matrix of second partial derivatives. Now we plug this into the DPP relation (5.56) and use the fact the martingale term in (5.57) has zero mean. We obtain:

$$\begin{aligned} 0 &= \max_{\alpha\in\mathcal{A}_{t,\tau}} E\left[\int_t^\tau r(Y_s,\alpha_s,s)\, ds + u(Y_\tau,\tau) - u(x,t) \mid \ Y_t = x\right] \\ &= \max_{\alpha\in\mathcal{A}_{t,\tau}} E\left[\int_t^\tau r(Y_s,\alpha_s,s)\, ds + \int_t^\tau u_t(Y_s,s) + \mathcal{L}^\alpha u(Y_s,s)\, ds \mid \ Y_t = x\right] \end{aligned} \tag{5.59}$$

Finally, let $\tau = t + h$, divide by h and let $h \to 0$, as in the deterministic case. We formally obtain the HJB equation

$$u_t(x,t) + \max_{a \in A} [r(x,a,t) + \mathcal{L}^a u(x,t)] = 0. \tag{5.60}$$

This may be written as

$$u_t(x,t) + \max_{a \in A} \left[r(x,a,t) + \frac{1}{2}\text{tr}(D^2 u(x,t)\sigma(x,a,t)\sigma^T(x,a,t)) + f(x,a,t) \cdot \nabla u(x,t) \right] = 0 \tag{5.61}$$

which is, in general, a fully-nonlinear, *second order* equation of the form

$$u_t + H(D^2 u, Du, x, t) = 0 \tag{5.62}$$

Notice that the equation is deterministic. The set of possible control values $A \subset \mathbb{R}^m$ is a subset of Euclidean space, and the maximum in the HJB equation (5.70) is over this deterministic set, not over the set $\mathcal{A}$.

HJB for the infinite horizon problem

Deriving the HJB for the infinite horizon problem is similar. Suppose $r = r(y,a)$ and $f = f(y,a)$. Suppose that u is the value function defined by

$$u(x) = \max_{\alpha \in \mathcal{A}} E\left[\int_0^\infty e^{-\lambda s} r(Y_s, \alpha_s)\, ds \mid \ Y_0 = x \right] \tag{5.63}$$

Then the Dynamic Programming Principle shows that for any $\tau > 0$

$$u(x) = \max_{\alpha \in \mathcal{A}} E\left[\int_0^\tau e^{-\lambda s} r(Y_s, \alpha_s)\, ds + e^{-\lambda \tau} u(Y_\tau) \mid \ Y_0 = x \right] \tag{5.64}$$

Using Itô's formula as before, we formally derive the second order equation equation

$$-\lambda u(x) + \max_{a \in A} [r(x,a) + \mathcal{L}^a u(x)] = 0 \tag{5.65}$$

8.5.3 Examples

Example 1: In this example, we consider the problem of portfolio optimization. We already considered the problem of maximizing utility from consumption of resources. In that model, we assumed that in the absense of consumption the individual's total wealth grows exponentially according to a deterministic rate (interest rate). Now we suppose that the individual may invest money in various asset classes and or consume resources. For simplicity we suppose that the individual may invest in either a stock (risky) or a bond (risk free growth). Suppose that the stock and bond satsfiy the equations

$$\begin{aligned} db_s &= r b_s\, ds \\ dP_s &= \mu P_s\, ds + \sigma P_s\, dB_s \end{aligned} \tag{5.66}$$

Here B_s is a standard Brownian motion. The individual may decide how much money to consume, how much money to invest in stock, and how much money to invest in bonds. Let c_s denote the

consumption rate. Let π_s denote the proportion of investments to put in stocks. Now, if Y_s is the individual's wealth at time s, Y_s satisfies the equation

$$\begin{aligned} dY_s &= (1-\pi_s)rY_s\,ds + \pi_s\mu Y_s\,ds + \pi_s\sigma Y_s\,dB_s - c_s\,ds \\ &= [((1-\pi_s)r + \pi_s\mu)Y_s - c_s]\,ds + \pi_s\sigma Y_s\,dB_s \end{aligned} \tag{5.67}$$

Lets suppose that $c_s = \eta_s Y_s$ with $\eta \geq 0$, so that the consumption rate cannot be negative. Then the control has two components: $\alpha_s = (\eta_s, \pi_s)$. Now our goal is to maximize the discounted utility

$$J_x(\alpha(\cdot)) := \int_0^\infty e^{-\lambda s}U(c_s)\,ds = \int_0^\infty e^{-\lambda s}U(\eta_s Y_s)\,ds \tag{5.68}$$

where $U(c)$ is a utility function, typically increasing and concave. So the value function is

$$u(x) = \max_{\alpha\in\mathcal{A}} \int_0^\infty e^{-\lambda s}U(\eta_s Y_s)\,ds \tag{5.69}$$

Notice that u depends on π_s through the definition of Y_s.

The HJB equation has the form

$$-\lambda u(x) + \max_{\eta,\,\pi}\left[U(\eta y) + ((1-\pi)r + \pi\mu - \eta)xu_x(x) + \frac{\pi^2\sigma^2}{2}x^2u_{xx}(x)\right] = 0. \tag{5.70}$$

This example is based on the paper by Robert Merton, *Optimal Consumption and Portfolio Rules in a Continuous-Time Model*, J. Econ. Theory, **3**, 1971 pp. 373-413. The model can be solved explicitly. See lecture notes by M. Soner for details.

Example 2: This example comes from the interesting paper by Leung, Sircar, and Zariphopoulou cited above. An investor want to optimize her portfolio, investing in either stock or bond, with no consumption. The problem is that the company in which she invests may default. If this occurs, she must sell her stock at the market price and put her money in the bond (she can't re-invest in another stock for the time interval under consideration). The stock price is modeled as:

$$dP_s = \mu P_s ds + \sigma P_s dB_s^1 \tag{5.71}$$

where B_s^1 is a Brownian motion. The value of the firms assets is modeled as

$$dV_s = \nu V_s ds + \eta V_s(\rho dB_s^1 + \rho' dB_s^2) \tag{5.72}$$

where B_s^2 is an independent Brownian motion, and $\rho \in (-1,1)$, $\rho' = \sqrt{1-\rho^2}$. Default occurs if the firm's price P_s drops below a boundary $\tilde{D}_s = De^{-\beta(T-s)}$, $s \in [0,T]$, $D > 0$. Choosing to invest a ratio π_s in the stock (as in the preceding example), the investor's wealth is Y_s where

$$dY_s = (1-\pi_s)rY_s\,ds + \pi_s\mu Y_s\,ds + \pi_s\sigma Y_s\,dB_s \tag{5.73}$$

and r is the interest rate for the bond. The initial condition is $Y_t = y$. We define the default time, a stopping time, by

$$\gamma_t = \inf\{\tau \geq t \mid \ V_\tau \leq \tilde{D}_\tau\} \tag{5.74}$$

If $\gamma_t > T$, no default occurs over the time interval in question.

Therefore, the investor wants to optimize

$$u(x,y,t) = \max_{\pi_s(\cdot)} E\left[U(Y_T)\mathbb{I}_{\{\gamma_t > T\}} + U(Y_{\gamma_t}e^{r(T-\gamma_t)})\mathbb{I}_{\{\gamma_t \le T\}} \mid V_t = x, Y_t = y\right] \tag{5.75}$$

Here $U(y)$ is a utility function ($U(y) = -e^{-hy}$ is used in the paper, $h > 0$). In this case, the HJB equation is

$$u_t + \mathcal{L}u + rxu_x + \max_{\pi}\left[\frac{1}{2}\sigma^2\pi^2 u_{xx} + \pi(\rho\sigma\eta u_{xy} + (\mu - r)u_x)\right] = 0 \tag{5.76}$$

where $\mathcal{L}$ is the operator

$$\mathcal{L}u = \frac{\eta^2 y^2}{2}u_{yy} + \nu y u_y.$$

The domain for u is $\{(t,y,x)\,|t \in [0,T], x \in \mathbb{R},\ y \in [\tilde{D}(t), \infty)\}$, and the boundary condition is

$$u(x,y,T) = U(x), \qquad u(x, De^{-\beta(T-t)}, t) = U(xe^{r(T-t)}).$$

There are other examples in this paper illustrating techniques for valuation of credit derivatives.