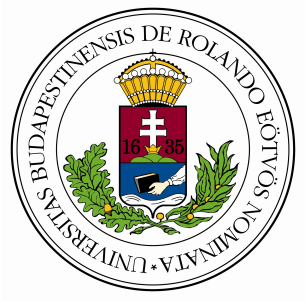
ANALYSIS OF ECONOMIC DYNAMICS USING DIFFERENTIAL EQUATIONS

Thesis

Áron Balázs Papp
Matematics BSc

Supervisor:

Dr. Petra Csomós
associate professor
Department of Applied Analysis and Computational Mathematics



Eötvös Loránd University, Faculty of Science

2025

Contents

Acknowledgement

First and foremost, I would like to express my gratitude to my supervisor, Dr. Petra Csomós, for her invaluable guidance, support, and encouragement throughout the entire process. Her dedication and readiness to discuss ideas, offer suggestions and help me improve my work have all greatly contributed to the development of this thesis. I am also sincerely grateful to Ágnes Backhausz for her helpful feedback on Itô integration.

I would also like to thank my family for their love, support and encouragement throughout my academic journey. Their constant belief in my abilities has inspired me to persevere. Finally, I am grateful to my friends for their companionship and support along the way.

Introduction

Over the past few decades, tools from probability theory and differential equations have become essential in modern economic modeling. In particular, Brownian motion, a concept originally developed in physics to describe random particle movement, has found important applications in economics and finance. It helps us model unpredictable changes over time, such as how stock prices fluctuate or how individuals face uncertain income in the future. When we build models based on this kind of randomness, we often use stochastic differential equations, which combine both randomness and deterministic change in a precise mathematical form.

A key idea in economic modeling is that individuals or firms are not the same. In reality, people differ in many ways: how productive they are, how much they save, or how they allocate their money. To better capture these differences, economists use heterogeneous agent models, where each agent (a person, a firm, or some other decision maker) follows their own path, which leads to more complex but more realistic models. For example, to understand how knowledge spreads through an economy or how inequality evolves over time, we use partial differential equations.

Another type of heterogeneous agent models focuses on problems that are about making the best possible decisions over time whether it is a household deciding how much to consume and save, or a government choosing the right economic policy. These types of problems are studied using the tools of optimal control theory. At the heart of this field lies the Hamilton—Jacobi—Bellman equation, which helps find the best decision-making strategy in uncertain and dynamic environments.

The thesis explores these themes in three main parts: First, we introduce stochastic calculus and its connections to finance through the famous Black—Scholes model. Next, we study a model of knowledge diffusion in the economy, based on a PDE inspired by biological population models. Finally, we turn to optimal control problems in economics, including an example model of individual decision-making under uncertainty.

The models presented in the thesis are selected from Achdou et al.: *Partial Differential Equations in Macroeconomics* [1]. Some aspects of the models are not discussed in detail in the original work, so we aim to elaborate on those points by refining its mathematical context and presenting further numerical results where relevant.

Chapter 1

Stochastic calculus in economics

This chapter deals with stochastic calculus in economics, starting with stochastic differential equations and their connection to PDEs. We then discuss their implications for the Black—Scholes model.

Notation 1.1. Let $\mathcal{M} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be an open set, where $n, m \in \mathbb{Z}^+$. If the function $f : \mathcal{M} \to \mathbb{R}$ is i times continuously differentiable in its first variable, and j times in the second, we denote it as $f \in C^{i,j}$.

Notation 1.2. Let $A \in \mathbb{R}^{n \times m}$ be a matrix. We denote by $\|A\|_F$ the Frobenius norm of A, i.e.:

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a^{i,j}|^2}.$$

1.1 Stochastic differential equations

Economic systems are inherently subject to uncertainty, driven by unpredictable shocks, policy changes, and external influences, therefore, capturing this randomness in differential equations is essential. In this chapter, we cover some basic concepts of stochastic calculus, which are essential for understanding the later developments in the thesis. The primary references for this topic are Michael Steele's *Stochastic Calculus and Financial Applications* [8] and Anne Carlstein's *Stochastic Calculus: Understanding Brownian Motion and Quadratic Variation* [9].

Definition 1.3. Let $(S, \mathcal{A})$ be a measurable space. A stochastic process is a set of random variables $\{X(t)\}_{t \in \mathbf{T}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbf{T} \neq \emptyset$, such that for each $t \in \mathbf{T}$, $X(t) : \Omega \to S$ is $(\mathcal{F}, \mathcal{A})$-measurable.
For each $\omega \in \Omega$, the trajectory corresponding to the realization ω of $X(t)$ is the function $t \mapsto X_\omega(t)$.

Assumption 1.4. In this thesis, we deal with continuous-time stochastic processes, so we assume $\mathbf{T} = [0, \infty)$ or $\mathbf{T} = [0, T]$ where $T > 0$. Furthermore, we assume that $S = \mathbb{R}$, and that it is equipped with the Borel sigma-algebra.

Assumption 1.5. From now on, we assume the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the increasing filtration $(\mathcal{F}_t)_{t\geq 0}$, such that $\mathcal{F}_t \subseteq \mathcal{F}$ for all $t \geq 0$, and satisfying the following conditions:

(i) $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous, i.e., for all $t \in T$, we have $\mathcal{F}_t = \bigcap_{s\geq t} \mathcal{F}_s$, and

(ii) $(\mathcal{F}_t)_{t\geq 0}$ is complete, i.e., for all $A \subseteq \Omega$ for which there exists a set $B \in \mathcal{F}$, such that $A \subseteq B$ and $\mathbb{P}(B) = 0$, we have $A \in \mathcal{F}_0$.

Assumption 1.5 means that under such a filtration, no new information can appear abruptly and events that do not change the outcome are accounted for in the initial knowledge.

Definition 1.6. A stochastic process $\{\mathbf{B}(t)\}_{t\geq 0} = \{(B^1(t), \ldots, B^d(t)\}_{t\geq 0}$ is called a d-dimensional Brownian motion (or Wiener process) if it satisfies the following properties:

(i) $\mathbf{B}(0) = 0$ almost surely,

(ii) $t \mapsto \mathbf{B}(t)$ is continuous almost surely,

(iii) The increments of $\{\mathbf{B}(t)\}_{t\geq 0}$ are normally distributed: for any $t \geq 0$ and $s > 0$, the increment $\mathbf{B}(t+s) - \mathbf{B}(t)$ follows a d-dimensional normal distribution with mean 0 and covariance matrix sI_d, i.e.,

$$\mathbf{B}(t+s) - \mathbf{B}(t) \sim \mathcal{N}(0, sI_d),$$

(iv) The increments of $\{\mathbf{B}(t)\}_{t\geq 0}$ are independent: for any $t \geq 0$ and $s > 0$, the increment $\mathbf{B}(t+s) - \mathbf{B}(t)$ is independent of the sigma-algebra $\sigma(\mathbf{B}(u) : 0 \leq u \leq t)$ which is the filtration generated by the process up to time t.

In one dimension, Brownian motion can be understood as the continuous-time parameterization of a symmetric random walk, describing the asymptotic behavior of its trajectories. It serves as a fundamental stochastic process and has widespread applications, particularly in finance and economics.

Proposition 1.7. *For any $t \in \mathbf{T}$, the trajectories of Brownian motion are not differentiable at t almost surely.*

Proof. Let $h > 0$. Since we know that $B(t+h) - B(t) \sim \sqrt{h}\mathcal{N}(0,1)$, we can write the difference quotient as:

$$\frac{B(t+h) - B(t)}{h} \sim \frac{\sqrt{h}\mathcal{N}(0,1)}{h} = \frac{\mathcal{N}(0,1)}{\sqrt{h}}.$$

Since $\mathcal{N}(0,1)$ is nonzero almost surely, as $h \to 0$, the difference quotient diverges with probability 1, therefore, the derivative does not exist almost surely. □

As we observed, Brownian motion is nowhere differentiable almost surely. However, calculus with Brownian motion can still be developed through Itô's framework. Since the full theoretical formulation of Itô integrals is beyond the scope of this thesis, we will instead use an equivalent, more practical definition based on a Riemann sum-like representation. For a detailed treatment of the formal construction, we recommend Chapters 6–7 of J. Michael Steele's book [8].

Definition 1.8. We say that a stochastic process $\{X(t)\}_{t\in\mathbf{T}}$ is

(i) adapted if for every $t \in \mathbf{T}$, the random variable $X(t)$ is $\mathcal{F}_t$-measurable,

(ii) càdlàg, if for each fixed ω, the trajectory $t \mapsto X_\omega(t)$ is everywhere right-continuous and has left limits everywhere,

(iii) progressively measurable, if for any $t \in \mathbf{T}$, the function defined by $(s, \omega) \to X_\omega(s)$ is measurable with respect to $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$ on $\mathbf{T} \times \Omega$.

Adapted means that the process can only depend on the information up to time t and not on future values, while the càdlàg property ensures that $X(t)$ is well-behaved.
Progressively measurable stochastic processes are of key importance to us. For those unfamiliar with them, we present a theorem that provides an easier way to think about such processes.

Proposition 1.9. *If a stochastic process is càdlàg and adapted, it is progressively measurable.*

When dealing with stochastic processes, the classical Riemann or Lebesgue integral is no longer adequate. As a result, a new kind of integral is required, which in our case is the Itô integral.

Definition 1.10. Let $t > 0$. We say that the random variable denoted by $\int_0^t H(s)\,\mathrm{d}B(s)$ is the Itô integral of a progressively measurable stochastic process $\{H(s)\}_{s\in[0,t]}$ with respect to a Brownian motion $\{B(s)\}_{s\in[0,t]}$, if

$$\mathbb{P}\left(\int_0^t H^2(s)\,\mathrm{d}s < \infty\right) = 1,$$

and for the partition $t_i = it/n$, where $0 \le i \le n$, we have

$$\mathbb{P}\left(\left|\sum_{i=1}^{n} H(t_{i-1})\big(B(t_i) - B(t_{i-1})\big) - \int_0^t H(s)\,\mathrm{d}B(s)\right| > \varepsilon\right) \to 0 \quad \text{as } n \to \infty,$$

for all $\varepsilon > 0$.

Remark 1.11. If the process $\{H(s)\}_{s\ge0}$ satisfies $\int_0^\infty H^2(s)\,\mathrm{d}s < \infty$ almost surely, then we can define the Itô integral on an infinite time horizon as:

$$\int_0^\infty H(s)\,\mathrm{d}B(s) := \lim_{t\to\infty} \int_0^t H(s)\,\mathrm{d}B(s).$$

It is easy to see that the Itô integral is not monotonous, however, it has other great properties, some of which we present below.

Proposition 1.12. *If* $\mathbb{E}\left[\int_0^\infty H^2(s)\,\mathrm{d}s\right] < \infty$, *then* $\mathbb{E}\left[\int_0^\infty H(s)\,\mathrm{d}B(s)\right] = 0$.

Proposition 1.13. *Let* $a, b \in \mathbb{R}$. *If the Itô integrals* $\int_0^\infty H_1(s)\,\mathrm{d}B(s)$ *and* $\int_0^\infty H_2(s)\,\mathrm{d}B(s)$ *exist, then so does* $\int_0^\infty \big(aH_1(s) + bH_2(s)\big)\,\mathrm{d}B(s)$, *and*

$$\int_0^\infty \big(aH_1(s) + bH_2(s)\big)\,\mathrm{d}B(s) = a\int_0^\infty H_1(s)\,\mathrm{d}B(s) + b\int_0^\infty H_2(s)\,\mathrm{d}B(s).$$

To model systems influenced by both deterministic trends and random noise, we work with special types of stochastic processes. In financial mathematics, for example, asset prices often evolve with some predictable drift while being constantly affected by unpredictable market shocks, modeled as noise via Brownian motion. For this reason, we define the so-called Itô process.

Definition 1.14. Let $\{\mathbf{B}(t)\}_{t\in\mathbf{T}} = \{(B^1(t),\dots,B^d(t))\}_{t\in\mathbf{T}}$ be a d-dimensional Brownian motion. We say that the $\{\mathbf{X}(t)\}_{t\in\mathbf{T}} = \{(X^1(t),\dots,X^n(t))\}_{t\in\mathbf{T}}$ process is an n-dimensional Itô process driven by a d-dimensional Brownian motion, if for almost every $\omega\in\Omega$, we have for all $1\le i\le n$, $t\in\mathbf{T}$:

$$X^i(t) = X^i(0) + \int_0^t \mu^i\left(s,\mathbf{X}(s)\right)\,\mathrm{d}s + \sum_{j=1}^{d}\int_0^t \sigma^{i,j}\left(s,\mathbf{X}(s)\right)\,\mathrm{d}B^j(s), \tag{1.1}$$

which we also write as:

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \boldsymbol{\mu}\left(s,\mathbf{X}(s)\right)\,\mathrm{d}s + \int_0^t \boldsymbol{\sigma}\left(s,\mathbf{X}(s)\right)\,\mathrm{d}\mathbf{B}(s),$$

where $\mathbf{X}(0)$ is $\mathcal{F}_0$-measurable, $\mu^i : \mathbf{T}\times\mathbb{R}^n\to\mathbb{R}$, $\sigma^{i,j} : \mathbf{T}\times\mathbb{R}^n\to\mathbb{R}$ are progressively measurable for all $1\le i\le n$, $1\le j\le d$, and $\boldsymbol{\mu} = (\mu^1,\dots,\mu^n)\in\mathbb{R}^n$, $\boldsymbol{\sigma} = (\sigma^{i,j})_{1\le i\le n,1\le j\le d}\in\mathbb{R}^{n\times d}$, satisfying

$$\int_0^t \left\|\boldsymbol{\mu}\big(s,\mathbf{X}(s)\big)\right\|\,\mathrm{d}s < \infty, \text{ and } \quad \int_0^t \left\|\boldsymbol{\sigma}\big(s,\mathbf{X}(s)\big)\right\|_F^2\,\mathrm{d}s < \infty$$

almost surely, for all $t\in\mathbf{T}$.

Remark 1.15. For a shorter notation, we formally write (1.1) in the differential form:

$$\mathrm{d}\mathbf{X}(t) = \boldsymbol{\mu}\big(t,\mathbf{X}(t)\big)\,\mathrm{d}t + \boldsymbol{\sigma}\big(t,\mathbf{X}(t)\big)\,\mathrm{d}\mathbf{B}(t). \tag{SDE}$$

The functions $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are usually referred to as drift and volatility, respectively. These names come from finance, where drift represents the expected return of an asset, and volatility represents the variance of the log-returns. We also call equation (SDE) a stochastic differential equation and $\mathbf{X}(t)$ a solution of (SDE).

Remark 1.16. We dissect (SDE) into two important categories:

(i) The functions $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are not deterministic, i.e., $\boldsymbol{\mu}(t,x) = \boldsymbol{\mu}(t,x,\omega)$ and $\boldsymbol{\sigma}(t,x) = \boldsymbol{\sigma}(t,x,\omega)$. We will not tackle this case in the thesis.

(ii) The functions $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are in the form $\widetilde{\mu}(t,x,\alpha_\omega(t))$ and $\widetilde{\sigma}(t,x,\alpha_\omega(t))$, where $\widetilde{\mu}$ and $\widetilde{\sigma}$ are deterministic functions, and α_ω is progressively measurable function called a control. This case will be elaborated on in Chapter 3.

Proposition 1.17. *If $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are Lipschitz continuous in their second variable, and satisfy the so-called "growth condition", i.e., there exists a constant $L>0$, such that*

$$\|\boldsymbol{\mu}(t,\underline{x})\|^2 + \|\boldsymbol{\sigma}(t,\underline{x})\|_F^2 \le L\big(1+\|\underline{x}\|^2\big), \quad \textit{for all } t\in\mathbf{T},\ \underline{x}\in\mathbb{R}^n,$$

then (SDE) *has a unique solution.*

Remark 1.18. If a solution of (SDE) exists, then it is continuous almost surely.

Itô's lemma is one of the fundamental tools in stochastic calculus. It plays a similar role to the Taylor expansion in classical calculus. The version below applies to one-dimensional Itô processes.

Lemma 1.19. *(Itô's lemma for Itô processes, one dimension)*
For a given function $f : \mathbf{T} \times \mathbb{R} \to \mathbb{R}$, $f \in C^{1,2}$, and an $\{X(t)\}_{t\in\mathbf{T}}$ 1-dimensional Itô process, the following holds:

$$\begin{aligned} f(t, X(t)) = f(0, X(0)) &+ \int_0^t \partial_x f(s, X(s)) \cdot \mu(s, X(s)) \,\mathrm{d}s \\ &+ \int_0^t \partial_t f(s, X(s)) \,\mathrm{d}s + \frac{1}{2}\int_0^t \partial_{xx} f(s, X(s)) \cdot \sigma^2(s, X(s)) \,\mathrm{d}s \\ &+ \int_0^t \partial_x f(s, X(s)) \cdot \sigma(s, X(s)) \,\mathrm{d}B(s). \end{aligned}$$

Using the short notation, we can write it as

$$\mathrm{d}f = \left(\partial_t f + \partial_x f\, \mu + \frac{1}{2}\partial_{xx} f\, \sigma^2\right) \mathrm{d}t + \partial_x f\, \sigma \,\mathrm{d}B(t).$$

Itô's lemma 1.19 can be generalized in higher dimensions as well.

Lemma 1.20. *(Itô's lemma for Itô processes, higher dimensions)*
For a given function $f : \mathbf{T} \times \mathbb{R}^n \to \mathbb{R}$, $f \in C^{1,2}$, and an $\{\mathbf{X}(t)\}_{t\in\mathbf{T}}$, n-dimensional Itô process driven by a d-dimensional Brownian motion, the following holds:

$$\begin{aligned} f(t, \mathbf{X}(t)) = f(0, \mathbf{X}(0)) &+ \int_0^t \partial_t f(s, \mathbf{X}(s)) \,\mathrm{d}s + \sum_{i=1}^n \int_0^t \partial_{x_i} f(s, \mathbf{X}(s)) \cdot \mu^i(s, \mathbf{X}(s)) \,\mathrm{d}s \\ &+ \frac{1}{2}\sum_{i,j=1}^n \int_0^t \partial_{x_i x_j} f(s, \mathbf{X}(s)) \cdot \sum_{k=1}^d \left(\sigma^{i,k}(s, \mathbf{X}(s))\right)\left(\sigma^{j,k}(s, \mathbf{X}(s))\right) \,\mathrm{d}s \\ &+ \sum_{i=1}^n \sum_{k=1}^d \int_0^t \partial_{x_i} f(s, \mathbf{X}(s)) \cdot \sigma^{i,k}(s, \mathbf{X}(s)) \,\mathrm{d}B^k(s). \end{aligned}$$

Using the short notation, we can write it as

$$\mathrm{d}f = \left(\partial_t f + \nabla f^\top \boldsymbol{\mu} + \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}\,\boldsymbol{\sigma}^\top \nabla^2 f\right]\right) \mathrm{d}t + \nabla f^\top \boldsymbol{\sigma} \,\mathrm{d}\mathbf{B}(t). \tag{1.2}$$

Remark 1.21. We may use Lemma 1.20 on a function f that does not depend on t explicitly, only through $\mathbf{X}(t)$, in which case the $\partial_t f$ term in (1.2) is not present and only $f \in C^2$ needs to be assumed. Similarly, if f does not depend on $\mathbf{X}(t)$, only on t, we get the fundamental theorem of calculus (Newton–Leibniz formula) back.

The final version of Itô's lemma we use in the thesis applies to functions of two Itô processes. This form is important to us mostly because it leads to a widely used result known as Itô's product rule.

Lemma 1.22. *(Itô's lemma for two Itô processes)*
For a given function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f \in C^{2,2}$, and the $\{X(t)\}_{t\in\mathbf{T}}, \{Y(t)\}_{t\in\mathbf{T}}$ Itô processes driven by the same Brownian motion $\{B(t)\}_{t\in\mathbf{T}}$, with drifts μ_X, μ_Y and volatilities σ_X, σ_Y respectively, we have

$$\begin{aligned} f(X(t),Y(t)) = f(X(0),Y(0)) &+ \int_0^t \partial_x f(X(s),Y(s))\,\mu_X(s,X(s))\,\mathrm{d}s \\ &+ \int_0^t \partial_y f(X(s),Y(s))\,\mu_Y(s,Y(s))\,\mathrm{d}s + \int_0^t \partial_x f(X(s),Y(s))\,\sigma_X(s,X(s))\,\mathrm{d}B(s) \\ &+ \int_0^t \partial_y f(X(s),Y(s))\,\sigma_Y(s,Y(s))\,\mathrm{d}B(s) + \frac{1}{2}\int_0^t \partial_{xx} f(X(s),Y(s))\,\sigma_X^2(s,X(s))\,\mathrm{d}s \\ &+ \frac{1}{2}\int_0^t \partial_{yy} f(X(s),Y(s))\,\sigma_Y^2(s,Y(s))\,\mathrm{d}s \\ &+ \int_0^t \partial_{xy} f(X(s),Y(s))\,\sigma_X(s,X(s))\,\sigma_Y(s,Y(s))\,\mathrm{d}s. \end{aligned}$$

Or using the short notation:

$$\begin{aligned} \mathrm{d}f = &\left(\partial_x f\mu_X + \partial_y f\mu_Y + \frac{1}{2}\partial_{xx} f\sigma_X^2 + \frac{1}{2}\partial_{yy} f\sigma_Y^2 + \partial_{xy} f\sigma_X\sigma_Y\right)\mathrm{d}t + \\ &+ \left(\partial_x f\sigma_X + \partial_y f\sigma_Y\right)\mathrm{d}B(t). \end{aligned}$$

Consequence 1.23. *By choosing $f(x,y) = xy$, we get a generalization of the formula for Leibniz's integration by parts, which we call Itô's product rule:*

$$\begin{aligned} X(t)Y(t) = X(0)Y(0) &+ \int_0^t Y(s)\,\mu_X(s,X(s))\,\mathrm{d}s \\ &+ \int_0^t X(s)\,\mu_Y(s,Y(s))\,\mathrm{d}s + \int_0^t Y(s)\,\sigma_X(s,X(s))\,\mathrm{d}B(s) \\ &+ \int_0^t X(s)\,\sigma_Y(s,Y(s))\,\mathrm{d}B(s) \\ &+ \int_0^t \sigma_X(s,X(s))\,\sigma_Y(s,Y(s))\,\mathrm{d}s. \end{aligned}$$

Or using the short notation:

$$\mathrm{d}(XY) = \left(Y\mu_X + X\mu_Y + \sigma_X\sigma_Y\right)\mathrm{d}t + \left(Y\sigma_X + X\sigma_Y\right)\mathrm{d}B(t).$$

1.2 Geometric Brownian motion

We now show a well-known example of (SDE) in one dimension, the so-called geometric Brownian motion (GBM for short), which we can explicitly solve using Itô's lemma. It describes the evolution of a variable whose logarithm follows a Brownian motion with drift. We will see an example of this in the following section, and also in Chapter 2.

Definition 1.24. We say that the stochastic process $\{X(t)\}_{t\in\mathbf{T}}$ follows a geometric Brownian motion if it satisfies the following stochastic differential equation:

$$\mathrm{d}X(t) = \mu X(t)\,\mathrm{d}t + \sigma X(t)\,\mathrm{d}B(t), \quad \text{where } \mu, \sigma \in \mathbb{R},\ \sigma > 0. \tag{GBM}$$

Theorem 1.25. *There exists a unique solution to* (GBM) *in the following form:*

$$X(t) = X(0)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right).$$

Proof. The existence and uniqueness of the solution follow directly from the assumptions and Proposition 1.17. Let $\{Y(t)\}_{t\in\mathbf{T}}$ such that:

$$Y(t) = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B(t),$$

and let $X(t) = X(0)\exp\left(Y(t)\right)$. We can easily see that $Y(t)$ satisfies the following stochastic differential equation:

$$\mathrm{d}Y(t) = \left(\mu - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma \mathrm{d}B(t)$$

By applying Itô's lemma to $f(y) = X(0)\exp\left(y\right) \in C^2$, we get:

$$\mathrm{d}\big(f(Y(t))\big) = \mathrm{d}\big(X(0)\exp\left(Y(t)\right)\big) = X(0)\exp\left(Y(t)\right)\big(\mu\,\mathrm{d}t + \sigma\,\mathrm{d}B(t)\big).$$

Since $f(Y(t)) = X(t)$, we have:

$$\mathrm{d}X(t) = \mu X(t)\,\mathrm{d}t + \sigma X(t)\,\mathrm{d}B(t),$$

which means that $X(t)$ satisfies (GBM) and this is the only solution. □

Since not all stochastic differential equations have closed-form solutions, numerical methods are widely used. Such a method is the Euler–Maruyama scheme, which is a stochastic extension of the classical Euler method for ODE-s. We show an example of this, by solving Equation (GBM), then comparing the numerical result to the analytical solution in Theorem 1.25.

The numerical solution can be obtained by partitioning the interval $[0, T]$ into J intervals of width $\Delta t = T/J$. Therefore, we have the partition $0 = t_0 < t_1 < \cdots < t_J = T$, where $t_n = n\Delta t$.

We simulate Brownian motion with independent normal increments, i.e, for every $n = 0, 1, \ldots, J-1$, we define $\Delta B_n = B(t_{n+1}) - B(t_n)$, which satisfies $\Delta B_n \sim \mathcal{N}(0, \Delta t)$ independently for each n. We denote the approximations by $X(t_n) \approx X_n$ for $n = 0, 1, \ldots, J$.

The update rule of the numerical scheme is:

$$X_{n+1} = X_n + \mu X_n \Delta t + \sigma X_n \Delta B_n.$$

The sequence of points (t_n, X_n) for $n = 0, 1, \ldots, J$ represents the numerical approximation of the solution at the discrete time points. These points can be plotted to visualize a simulated path of the stochastic process $X(t)$. Figure 1 illustrates a few simulated paths of the geometric Brownian motion using the Euler–Maruyama method. Figure 1 shows that the simulated paths are very similar to the analytical solution, which suggests that the numerical method gives a good approximation of the geometric Brownian motion.

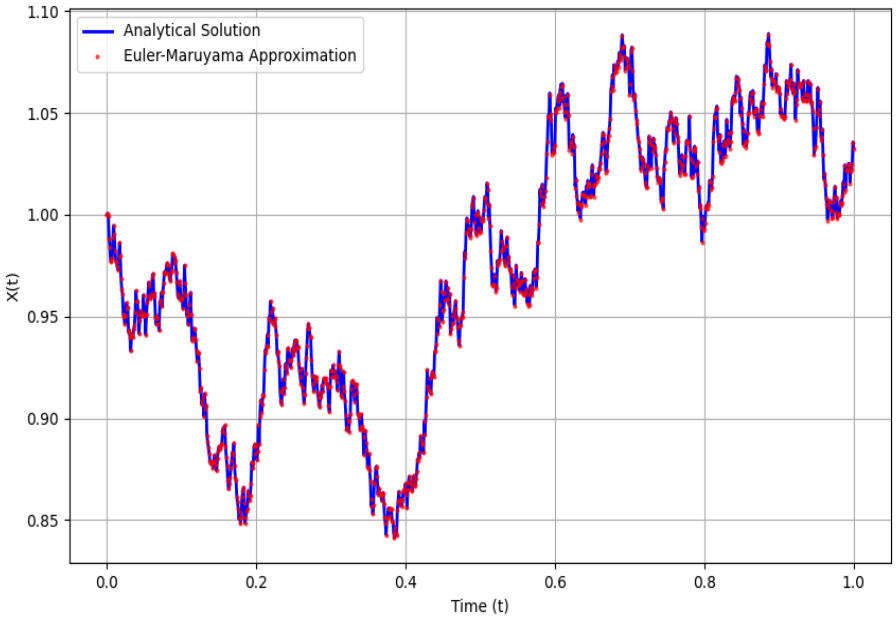


Figure 1: Analytical and numerical solution of (GBM) with initial condition $X(0) = 1$, drift $\mu = 0.1$, volatility $\sigma = 0.2$, $T = 1$ and $J = 1000$.

1.3 Connection to parabolic PDEs

Here we give some insight into how stochastic differential equations are related to parabolic differential equations by Itô's lemma. This section is based on Chapter 15 of J. Michael Steele's *Stochastic Calculus and Financial Applications* [8].

In order to simplify the proof while keeping it insightful, we omit a technical step that involves showing that Itô integrals can be considered as "local martingales". We also skip the argument that bounded local martingales are martingales. This is summarized in the lemma below.

Lemma 1.26. *Let $t > 0$ be fixed, and suppose that $\{H(s)\}_{0 \le s \le t}$ is a bounded, progressively measurable stochastic process. If the Itô integral $\int_0^\tau H(s)\,\mathrm{d}B(s)$ exists for all $\tau \in [0, t]$ in the sense of Definition 1.10, then the following process is a martingale:*

$$M(\tau) := \int_0^\tau H(s)\,\mathrm{d}B(s), \quad 0 \le \tau \le t.$$

Theorem 1.27. *(Feynman–Kac formula)*
Let $f : \mathbb{R} \to \mathbb{R}$, $q : \mathbb{R} \to \mathbb{R}$, $u : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be bounded functions, where $u \in C^{1,2}$. Furthermore, let $\mu : \mathbb{R} \to \mathbb{R}$, $\sigma : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous functions that satisfy

$$\mu^2(x) + \sigma^2(x) \leq K_0(1 + x^2), \quad x \in \mathbb{R} \tag{1.3}$$

for some $K_0 > 0$ constant. If $u(t, x)$ is the unique solution to the PDE

$$\partial_t u(t, x) = \frac{1}{2}\sigma^2(x)\partial_{xx}u(t, x) + \mu(x)\partial_x u(t, x) + q(x)u(t, x), \tag{1.4}$$

with the initial condition $u(0, x) = f(x)$, then $u(t, x)$ can be expressed as

$$u(t, x) = \mathbb{E}\left[f\big(x + X(t)\big) \exp\left(\int_0^t q\big(x + X(s)\big)\,\mathrm{d}s\right)\right],$$

where $\{X(t)\}_{t\geq 0}$ is the unique solution to

$$\mathrm{d}X(t) = \mu\big(X(t)\big)\,\mathrm{d}t + \sigma\big(X(t)\big)\,\mathrm{d}B(t), \quad X(0) = 0. \tag{1.5}$$

Proof. We introduce the stochastic process $\{M(s)\}_{0\leq s\leq t}$ such that:

$$M(s) = u\big(t - s, X_x(s)\big) \exp\left(\int_0^s q\big(X_x(v)\big)\,\mathrm{d}v\right) := U(s) \cdot V(s),$$

where $X_x(t) = x + X(t)$, given that $\{X(t)\}_{t\geq 0}$ is the unique solution to (1.5), which exists due to condition (1.3). Note that, due to Remark 1.18, X is continuous and so is X_x. Therefore, U and V are continuous as well.
Using Itô's lemma (Lemma 1.19) for U, we get:

$$\begin{aligned}\mathrm{d}U(s) = \mathrm{d}u\big(t - s, X_x(s)\big) = \Big(&\partial_x u(t - s, X_x(s)) \cdot \mu(X_x(s)) - \\ &- \partial_t u(t - s, X_x(s)) + \frac{1}{2}\partial_{xx}u(t - s, X_x(s)) \cdot \sigma^2(X_x(s))\Big)\,\mathrm{d}s \\ &+ \partial_x u(t - s, X_x(s)) \cdot \sigma(X_x(s))\,\mathrm{d}B(s).\end{aligned}$$

Since u is the unique solution to (1.4), we have:

$$\mathrm{d}u\big(t - s, X_x(s)\big) = -q(X_x(s))u(t - s, X_x(s))\,\mathrm{d}s + \partial_x u(t - s, X_x(s)) \cdot \sigma(X_x(s))\,\mathrm{d}B(s).$$

Furthermore, for V, we may also use Itô's lemma:

$$\begin{aligned}\mathrm{d}V(s) = \mathrm{d}\left(\exp\left(\int_0^s q\big(X_x(v)\big)\,\mathrm{d}v\right)\right) = \exp\left(\int_0^s q\big(X_x(v)\big)\,\mathrm{d}v\right) q\big(X_x(s)\big)\,\mathrm{d}s = \\ = V(s)q\left(X_x(s)\right)\,\mathrm{d}s.\end{aligned}$$

We now use Consequence 1.23, Itô's product rule:

$$\begin{aligned}\mathrm{d}M(s) = \mathrm{d}\big(U(s)V(s)\big) &= V(s)(-q(X_x(s))u(t - s, X_x(s)))\,\mathrm{d}s + \\ &+ U(s)V(s)q(X_x(s))\,\mathrm{d}s + V(s)\partial_x u(t - s, X_x(s))\sigma(X_x(s))\,\mathrm{d}B(s) \\ &= V(s)\partial_x u(t - s, X_x(s))\sigma(X_x(s))\,\mathrm{d}B(s).\end{aligned}$$

From our assumptions, we have:

$$\sup_{0\le s\le t} |M(s)| \le \|u\|_\infty \exp\left(t\|q\|_\infty\right) = K_1, \quad \text{where } K_1 \in \mathbb{R}.$$

Lemma 1.26 implies that $M(s)$ is a martingale. The martingale property means that for any $0 \le s \le t$, we have:

$$\mathbb{E}\left[M(t)|\mathcal{F}_s\right] = M(s).$$

Therefore, $\mathbb{E}[M(t)] = \mathbb{E}[M(0)]$ also holds, which means that:

$$u(t,x) = \mathbb{E}\big[M(0)\big] = \mathbb{E}\big[M(t)\big] = \mathbb{E}\left[f\big(x + X(t)\big) \exp\left(\int_0^t q(x + X(s))\,\mathrm{d}s\right)\right].$$

This completes the proof. □

To further illustrate how impactful Theorem 1.27 is, let us consider the case when $\mu(t,x) \equiv q(x) \equiv 0$ and $\sigma(t,x) \equiv 2D$, where $D > 0$. In this case, (1.4) becomes the heat equation

$$\partial_t u(t,x) = D \cdot \partial_{xx} u(t,x), \tag{1.6}$$

with $\mathrm{d}X(t) = D \cdot \mathrm{d}B(t)$, which in the integral form is:

$$X(t) = X(0) + D\int_0^t 1\,\mathrm{d}B(t) = DB(t).$$

Therefore, the solution to the heat equation (1.6) becomes:

$$u(t,x) = \mathbb{E}\big[f\big(x + DB(t)\big)\big],$$

which means that the solution of the heat equation can be thought of as the expected value of the initial condition evaluated along random paths generated by the Brownian motion, scaled by the diffusion parameter D. In other words, solving the heat equation corresponds to averaging the initial profile f over all the possible places a particle could diffuse to after time t, starting from position x.

The methods used in the proof of Theorem 1.27 open the door to deriving a variety of representation theorems of similar form. We show an example of this for arguably the most well-known problem in finance, i.e., option pricing.

1.4 Implications for the Black–Scholes model

The goal is to determine a fair price for a financial asset called an option. We will focus on a specific type of option, called a European call option, which gives the purchaser the right to buy an underlying asset from the seller at a fixed price $K > 0$ at a specific expiration time $T > 0$.
We assume that the stock price $\mathcal{S}(t)$ follows a geometric Brownian motion, while the

bond price $\mathcal{B}(t)$ evolves deterministically in the following manner, where $\sigma, \mu, r \geq 0$ are constants:

$$\mathrm{d}\mathcal{S}(t) = \mu\mathcal{S}(t)\,\mathrm{d}t + \sigma\mathcal{S}(t)\,\mathrm{d}B(t), \quad \mathrm{d}\mathcal{B}(t) = r\mathcal{B}(t)\,\mathrm{d}t, \quad t \in [0,T]. \tag{1.7}$$

To determine the fair price of the option, we define a function $u(t,x)$ as:

$$u(t,x) := \big\{\text{the fair price at time } t \text{ of a European call option, assuming } \mathcal{S}(t) = x\big\}.$$

Theorem 1.28. *Let $h(x)$ and $u(t,x)$ be bounded functions, where $u(t,x)$ is the unique solution to the below terminal value problem, called the Black–Scholes equation:*

$$\begin{cases} \partial_t u(t,x) = -\dfrac{1}{2}\sigma^2\partial_{xx}u(t,x) - rx\partial_x u(t,x) + ru(t,x), \\ u(T,x) = \max\{x-K,0\}, \end{cases} \tag{1.8}$$

and let $\{X_{t,x}(s)\}_{s\in[0,T]}$ be a stochastic process, defined as:

$$X_{t,x}(s) = \begin{cases} x & \text{if } s \in [0,t], \\ \text{the solution of (1.9)} & \text{if } s \in [t,T], \end{cases}$$

with

$$\mathrm{d}X_{t,x}(s) = rX_{t,x}(s)\,\mathrm{d}s + \sigma\,\mathrm{d}B_s, \quad X_{t,x}(t) = x. \tag{1.9}$$

Then, $u(t,x)$ can be expressed as:

$$u(t,x) = \exp\big(-r(T-t)\big)\mathbb{E}\big[\max\big\{X_{t,x}(T) - K, 0\big\}\big]. \tag{1.10}$$

In other words, the value of the option today is the average amount we expect it to pay at maturity, assuming the stock evolves under risk-neutral dynamics.

The proof is analogous to the proof of Theorem 1.27. Theorem 1.28 is of key significance, as the expected value in (1.10) can be calculated explicitly, giving us an analytical solution to (1.8). Furthermore, if we generalize the underlying processes in (1.7) with $\mu\big(t,\mathcal{S}(t)\big)$, $\sigma\big(t,(\mathcal{S}(t)\big)$ and $r\big(t,\mathcal{S}(t)\big)$, we can easily prove analogous representation theorems for this case, opening the door for simulations, even if the calculation of the expected value is not feasible.

Note that the drift coefficient μ does not appear in the Black–Scholes equation (1.8). This is a direct consequence of using a change of measure to move from the real-world probability measure to a theoretical risk-neutral measure. Throughout the thesis, we assume such a measure for convenience purposes. For more insight, see the relevant literature on Girsanov's theorem, in Chapter 13 of *Stochastic Calculus and Financial Applications* [8] for example.

Chapter 2

The Lucas–Alvarez model

In this chapter, our goal is to analyze one of the models presented in the paper by Achdou et al. [1], and gain a deeper understanding about the special type of parabolic differential equation introduced in this model.

2.1 The Fisher–KPP equation

This section explores the origins of reaction-diffusion equations and clarifies the terminology associated with them. The primary references are the papers by Fisher [4] and Kolmogorov et al. [5].

Definition 2.1. We say that a parabolic differential equation is a reaction-diffusion equation, if it is of the form

$$\partial_t u = D \cdot \Delta u + f(v) \tag{2.1}$$

where $\Omega \subseteq \mathbb{R}^n$, $u : [0, \infty) \times \Omega \to \mathbb{R}^m$ is the unknown function, $f : \mathbb{R}^m \to \mathbb{R}^m$, $f \in C^1$ and $D \in \mathbb{R}^{m \times m}$ is a diagonal matrix with positive entries only.

To fully define the problem, the equation is typically accompanied by the initial condition

$$u(0, x) = u_0(x), \quad x \in \mathbb{R}$$

where $u_0 : \Omega \to \mathbb{R}^m$ is a given initial state, as well as suitable boundary conditions, which depend on the specific problem and domain.

Remark 2.2. In this chapter, we investigate the one-space-dimensional case ($n = m = 1$).

Reaction-diffusion equations describe how diffusion spreads substances while reactions change their concentrations. These equations originally emerged in chemistry, where they were used to model the interplay between chemical reactions and diffusion processes.
One of the earliest and most influential examples is the Fisher–KPP equation, introduced in population genetics to describe the spread of advantageous traits in a population.

In 1937, slightly different forms of the equation were independently studied by Fisher [4] and by Kolmogorov, Petrovsky, and Piskunov (hence the name KPP) [5] in their respective works.

Definition 2.3. Let $r > 0$. We say that equation (2.1) is called a KPP equation if the function f satisfies the following properties:

(i) $f(0) = f(1) = 0$,

(ii) $f(u) > 0$ for $0 < u < 1$,

(iii) $f'(0) = r$,

(iv) $f'(u) < r$ for $0 < u \leq 1$.

Definition 2.4. Let $r > 0$, $K > 0$. We say that (2.1) is a Fisher equation if it is in the form

$$\partial_t u = D\,\partial_{xx} u + ru\left(1 - \frac{u}{K}\right), \tag{2.2}$$

that is, $f(u) = ru\left(1 - \frac{u}{K}\right)$.

Remark 2.5. By choosing

$$\widetilde{t} = rt, \quad \widetilde{x} = x\sqrt{\frac{r}{D}}, \quad \widetilde{u}(\widetilde{t}, \widetilde{x}) = \frac{u(t,x)}{K},$$

we get:

$$\partial_t u(t,x) = Kr\partial_{\widetilde{t}}\widetilde{u}(\widetilde{t}, \widetilde{x}), \quad \partial_{xx} u(t,x) = K\frac{r}{D}\partial_{\widetilde{x}\widetilde{x}}\widetilde{u}(\widetilde{t}, \widetilde{x}). \tag{2.3}$$

After substituting (2.3) into (2.2) and dividing both sides of the equation by $Kr \neq 0$, we end up with

$$\partial_{\widetilde{t}}\widetilde{u} = \partial_{\widetilde{x}\widetilde{x}}\widetilde{u} + \widetilde{u}(1 - \widetilde{u}). \tag{2.4}$$

Remark 2.6. Since the transformed Fisher equation (2.4) satisfies the KPP constraints, we usually refer to (2.2) as the Fisher–KPP equation.

In the original formulations the function u represents a population density ratio, so it is naturally assumed to take values in the interval $[0, 1]$. Within our model in this chapter, it will represent the proportion of individuals below a given productivity level.

2.2 The model's setup

An important category of heterogeneous agent models focuses on the dynamics of the distribution of knowledge and productivity. The Lucas–Alvarez model falls into this category, describing how knowledge spreads across a population of individuals with different initial levels of productivity. In this model, the evolution of the knowledge distribution follows a Fisher–KPP-type equation, capturing how learning and innovation shape long-term economic growth.

We assume that we have a population that consists of a continuum of people

$$\mathcal{P} = \{i : i \in \mathbb{R}\}.$$

We track the productivity for every individual i with the stochastic process $\{Z_i(t)\}_{t\geq 0}$, where

$$Z_i(t) : [0, \infty) \to [0, \infty).$$

We can describe the economy's strength with a cumulative distribution function $G : [0, \infty) \times \mathbb{R} \to [0, 1]$, defined by

$$G(t, z) = \mathbb{P}(Z_i(t) \leq z).$$

From now on, we are going to consider $X_i(t) = \ln(Z_i(t))$ and $\widetilde{F} : [0, \infty) \times \mathbb{R} \to [0, 1]$ with $\widetilde{F}(t, x) = \mathbb{P}(X_i(t) \leq x)$ and assume that $\widetilde{F}$ is once differentiable in its first variable and twice in its second.
The changes in knowledge of a given individual can occur in multiple ways, presented in the following.

Innovation

We introduce a $\theta > 0$ parameter, which we will refer to as the innovation parameter, and it represents the rate at which every individual's knowledge increases. We can write this as:

$$X_i(t) - X_i(0) = \int_0^t \theta \, \mathrm{d}s = \theta t \quad \big(\text{or formally: } \mathrm{d}X_i(t) = \theta \, \mathrm{d}t\big).$$

Since this source of productivity change is homogeneous among the population, our intuition suggests that it will not be significant when describing inequality and knowledge distributions. This observation will be verified in Remark 2.7.

Meetings between people

In this model, there are encounters between individuals that occur in the form of a Poisson process with parameter $\alpha > 0$. This means that for each individual i, the number of meetings N_i in the time interval $[0, \tau]$ follows a Poisson distribution:

$$\mathbb{P}(N_i(\tau) = k) = \frac{(\alpha\tau)^k e^{-\alpha\tau}}{k!}.$$

Accordingly, the expected number of meetings for individual i over this interval is $\mathbb{E}[N_i(\tau)] = \alpha\tau$. When an encounter occurs between two people, with productivity levels $X_i(t)$ and $X_j(t)$, both update their productivity to the higher of the two, that is, $\max\{X_i(t), X_j(t)\}$. To capture this idea, we can formally represent it with the below equation:

$$\mathrm{d}X_i(t) = \Delta_i(t) \, \mathrm{d}N_i(t).$$

where $\Delta_i(t)$ is the gain in productivity from a meeting at time t, which as established, occurs if the individual i meets someone with higher productivity.

Research

An individual's knowledge evolves not only through innovation or interactions, but also through research. This model treats research as an unpredictable source of knowledge, therefore, its effect can be described by a stochastic differential equation, defined in Chapter 1:

$$\mathrm{d}X_i(t) = \sigma\,\mathrm{d}B_i(t),$$

where $\sigma > 0$ is a constant. Note that this preserves the non-negativity of $Z_i(t)$.

2.3 Combined effect

We may consider a general version of the model, where all three of the above effects influence the productivity levels, i.e.:

$$\mathrm{d}X_i(t) = \theta\,\mathrm{d}t + \Delta_i(t)\,\mathrm{d}N_i(t) + \sigma\,\mathrm{d}B_i(t).$$

As a combination of these effects, we can model the idea flow with the following equation:

$$\partial_t \widetilde{F}(t,x) = -\theta\partial_x \widetilde{F}(t,x) + \frac{\sigma^2}{2}\partial_{xx}\widetilde{F}(t,x) - \alpha\widetilde{F}(t,x)\big(1-\widetilde{F}(t,x)\big). \tag{2.5}$$

Remark 2.7. By choosing $F(t,x) = \widetilde{F}(t, x+\theta t)$, we get:

$$\partial_t F(t,x) = \partial_t\widetilde{F}(t,x+\theta t) + \theta\partial_x\widetilde{F}(t,x+\theta t), \quad \partial_{xx}F(t,x) = \partial_{xx}\widetilde{F}(t,x+\theta t). \tag{2.6}$$

After substituting (2.6) into (2.5), we end up with a de-trended version of the equation, which will be in the center of our attention for this chapter:

$$\partial_t F(t,x) = \frac{\sigma^2}{2}\partial_{xx}F(t,x) - \alpha F(t,x)\big(1-F(t,x)\big). \tag{LA}$$

Equation (LA) is accompanied by initial and asymptotic boundary conditions:

$$F_0(x) = F(0,x), \tag{2.7}$$
$$\lim_{x\to-\infty} F(t,x) = 0, \tag{2.8}$$
$$\lim_{x\to\infty} F(t,x) = 1, \tag{2.9}$$

where $F_0(x)$ is the initial cumulative distribution function. We refer to equation (LA) as such, following Lucas and Alvarez, who first introduced the model. Note that (LA) is structurally similar to the Fisher–KPP equation. The reaction term $-\alpha F(1-F)$ captures the effect of the individuals' interactions with each other, and the diffusion term $(\sigma^2/2)\,\partial_{xx}F$ comes from research. But there is a crucial difference: the reaction term contains a negative coefficient, making it a "reverse" Fisher–KPP equation. This difference arises because ideas diffuse but do not self-replicate in the same manner as biological populations, therefore, not leading to exponential increase.

2.4 Traveling wave solutions

Although traveling wave solutions are typically associated with hyperbolic partial differential equations, it is interesting to note that some parabolic equations, such as (LA), can also admit such solutions.

Theorem 2.8. *Let $F(t,x) \in C^{1,2}$ be a solution to* (LA) *with asymptotic boundary conditions* (2.8)-(2.9)*. Then for any wave speed $c \geq \sigma\sqrt{2\alpha}$, there exists a unique, monotonically increasing function $\phi : \mathbb{R} \to [0,1]$, $\phi \in C^2$, such that $F(t,x) = \phi(x-ct)$, $\lim_{\xi\to-\infty} \phi(\xi) = 0$, and $\lim_{\xi\to\infty} \phi(\xi) = 1$.*

Proof. We seek a traveling wave solution of the form

$$F(t,x) = \phi(x-ct).$$

We compute the following partial derivatives:

$$\partial_t F(t,x) = -c\phi'(x-ct), \quad \partial_{xx} F(t,x) = \phi''(x-ct).$$

Substituting these into (LA), and writing $\xi = x - ct$ we get

$$-c\phi'(\xi) = \frac{\sigma^2}{2}\phi''(\xi) - \alpha\phi(\xi)\left(1-\phi(\xi)\right). \tag{2.10}$$

We can write this second-order ordinary differential equation as a nonlinear, autonomous system of two first-order ordinary differential equations:

$$\begin{cases} \phi'(\xi) = \psi(\xi), \\ \psi'(\xi) = \dfrac{2}{\sigma^2}\alpha\phi(\xi)\big(1-\phi(\xi)\big) - \dfrac{2}{\sigma^2}c\psi(\xi). \end{cases} \tag{2.11}$$

We now examine the phase plane of (2.11). The equilibrium points can be found by solving the following system of equations:

$$\begin{aligned} 0 &= \psi(\xi), \\ 0 &= \frac{2}{\sigma^2}\alpha\phi(\xi)\big(1-\phi(\xi)\big) - \frac{2}{\sigma^2}c\psi(\xi). \end{aligned}$$

From this, we get:

$$\frac{2}{\sigma^2}\alpha\phi(\xi)\big(1-\phi(\xi)\big) = 0 \quad \text{for all } \xi \in \mathbb{R}.$$

Therefore, the two equilibrium points of (2.11) are: $\mathbf{E}_1 = (0,0)$, and $\mathbf{E}_2 = (1,0)$.
The Jacobian matrix corresponding to system (2.11) is

$$J(\phi,\psi) = \begin{bmatrix} 0 & 1 \\ (2/\sigma^2)\alpha(1-2\phi) & -(2/\sigma^2)c \end{bmatrix}.$$

To compute the eigenvalues at each equilibrium, we need to find the roots of

$$\det\big(J(0,0) - \lambda I\big) = \begin{vmatrix} -\lambda & 1 \\ (2/\sigma^2)\alpha & -(2/\sigma^2)c - \lambda \end{vmatrix} = \lambda^2 + \frac{2c}{\sigma^2}\lambda - \frac{2\alpha}{\sigma^2}, \text{ and} \tag{2.12}$$

$$\det\big(J(1,0) - \lambda I\big) = \begin{vmatrix} -\lambda & 1 \\ -(2/\sigma^2)\alpha & -(2/\sigma^2)c - \lambda \end{vmatrix} = \lambda^2 + \frac{2c}{\sigma^2}\lambda + \frac{2\alpha}{\sigma^2}. \tag{2.13}$$

The roots of (2.12) and (2.13) are:

$$\lambda_{1,2} = -\frac{c \pm \sqrt{c^2 + 2\sigma^2\alpha}}{\sigma^2}, \quad \lambda_{3,4} = -\frac{c \pm \sqrt{c^2 - 2\sigma^2\alpha}}{\sigma^2} \quad \text{respectively.} \tag{2.14}$$

We can easily see that λ_1 and λ_2 are real numbers and of opposite sign, therefore, $\mathbf{E}_1$ is a saddle point. Since we assumed $c^2 \geq 2\sigma^2\alpha$, λ_3 and λ_4 are both negative real numbers, therefore, $\mathbf{E}_2$ is a stable node. If we had $c^2 - 2\sigma^2\alpha < 0$, then $\mathbf{E}_2$ would be a stable spiral. This could result in $\phi > 1$, and $\phi' < 0$ which would contradict our assumptions.
To prove the existence of a traveling wave solution, we need to find a trajectory from the saddle point $\mathbf{E}_1$ to the stable node $\mathbf{E}_2$ on the phase plane (ϕ, ϕ'). If exists, it corresponds to a $\phi(\xi), \psi(\xi)$ pair, with the following properties:

$$\lim_{\xi \to -\infty} \phi(\xi) \to 0, \quad \lim_{\xi \to \infty} \phi(\xi) = 1 \quad \text{and} \quad \lim_{\xi \to -\infty} \psi(\xi) = \lim_{\xi \to \infty} \psi(\xi) = 0.$$

Finding the nullclines are of interest to us, as they are the curves where the derivative of one of the variables is zero. The two nullclines of (2.11) are:

$$\psi = 0 \quad \text{and} \quad \psi = -\frac{\alpha}{c}\phi^2 + \frac{\alpha}{c}\phi.$$

To make our argument more visible, we sketch the phase plane of (2.11), which can be seen below in Figure 1.

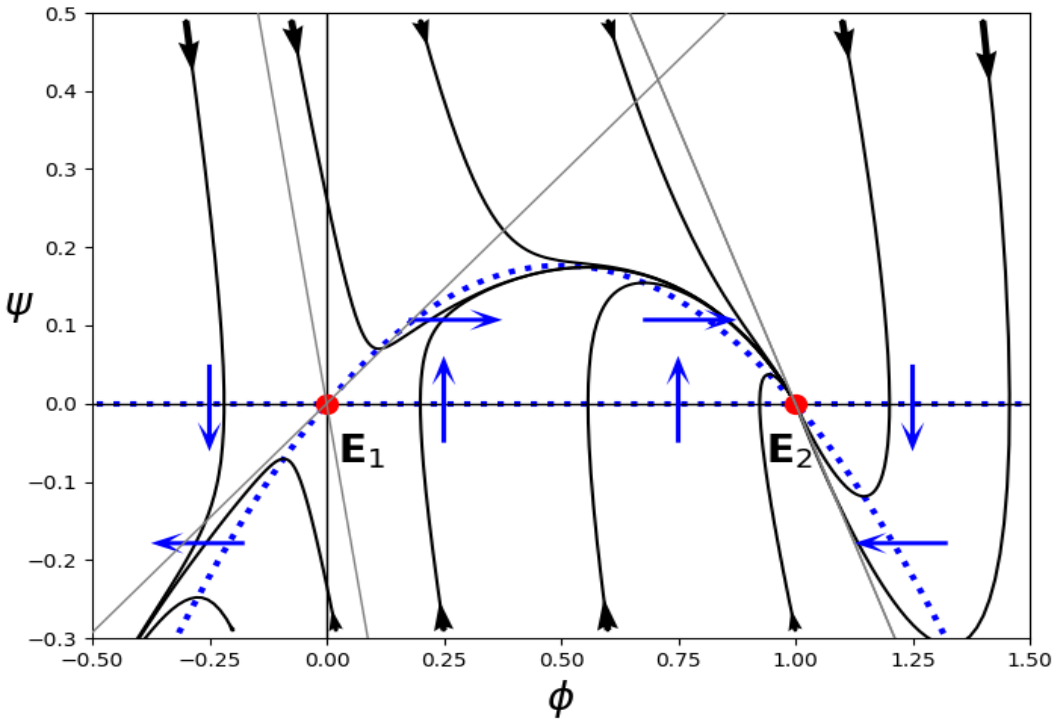


Figure 1: The (ϕ, ϕ') phase plane with nullclines, eigenvectors and some trajectories.

The directions of eigenvectors corresponding to eigenvalues (2.14) are also plotted on Figure 1. From our assumptions that $\phi \in [0, 1]$ and ϕ is monotonically increasing, we

know that the trajectory can only be in the first quadrant. We can see that if we start at some $\varepsilon > 0$ vicinity of $\mathbf{E}_1$ in this quadrant along the eigenvector associated with the positive eigenvector λ_1, the trajectory will eventually approach the equilibrium point $\mathbf{E}_2$. This trajectory is the unique traveling wave solution we are looking for. □

Consequence 2.9. *Let $F(t,x)$ be a solution of* (LA) *with the initial value condition* (2.7) *and the boundary conditions* (2.8)-(2.9)*, where $F_0(x) \in C^2$. Then, there exists some $c \geq \sigma\sqrt{2\alpha}$, for which the unique solution to* (LA) *with initial value and boundary conditions* (2.7)-(2.9) *is $F(t,x) = F_0(x - ct)$.*

The existence of a traveling wave solution means that under some appropriate initial distributions, productivity grows at a constant rate which is directly proportional to the rate at which people meet each other and do research.
While Theorem 2.8 guarantees that the traveling wave solution of (LA) exists, it does not provide a practical formula for computing it. For this, we use the so-called perturbation method.

The perturbation method

Perturbation methods are analytical techniques used to approximate solutions for problems that are difficult or impossible to solve exactly.
These methods are especially useful when the system contains a small parameter, which introduces a deviation from simpler, well-understood solution. This section is motivated by J. David Logan's book [6] on nonlinear partial differential equations.

A perturbation series of an unknown function g is an expansion in powers of a parameter ε_0. The series is typically written as

$$g(s) = g_0(s) + \varepsilon_0 g_1(s) + \varepsilon_0^2 g_2(s) + \dots. \tag{2.15}$$

In this series, g_0 is the unperturbated solution, while the higher-order terms $\varepsilon_0^n g_n$ represent corrections to the solution of $O(\varepsilon_0^n)$. From a numerical perspective, this means that the exact solution can be well-approximated, given a desired accuracy.
Since (2.10) is autonomous, we can assume that $\phi(0) = 1/2$. Let $\varepsilon > 0$, given by $\varepsilon = 1/c^2$. We persist with the assumption that $c^2 \geq 2\alpha\sigma^2$, therefore, $\varepsilon \leq 1/(2\alpha\sigma^2)$. We can write (2.10) as:

$$0 = \frac{\sigma^2}{2}\sqrt{\varepsilon}\phi''(\xi) + \phi'(\xi) - \alpha\sqrt{\varepsilon}\phi(\xi)\big(1 - \phi(\xi)\big). \tag{2.16}$$

By introducing $\eta = \xi\sqrt{\varepsilon}$ and $\Theta(\eta) = \phi(\xi/\sqrt{\varepsilon})$, we can rewrite (2.16) as:

$$0 = \frac{\sigma^2}{2}\varepsilon\Theta''(\eta) + \Theta'(\eta) - \alpha\Theta(\eta)\big(1 - \Theta(\eta)\big), \tag{2.17}$$

with the conditions

$$\lim_{\eta\to-\infty}\Theta(\eta) = 0, \quad \lim_{\eta\to\infty}\Theta(\eta) = 1, \quad \Theta(0) = \frac{1}{2}. \tag{2.18}$$

By expressing Θ as a perturbation series, i.e., in the form (2.15), we can write (2.17) as:

$$0 = \frac{\sigma^2}{2}\varepsilon\left(\Theta_0'' + \varepsilon\Theta_1'' + \varepsilon^2\Theta_2'' + \dots\right) + \left(\Theta_0' + \varepsilon\Theta_1' + \varepsilon^2\Theta_2' + \dots\right) \\ - \alpha\left(\Theta_0 + \varepsilon\Theta_1 + \varepsilon^2\Theta_2 + \dots\right)\Big(1 - \left(\Theta_0 + \varepsilon\Theta_1 + \varepsilon^2\Theta_2 + \dots\right)\Big).$$

After rearranging for powers of ε, and keeping only the zero- and first-order terms, we get:

$$0 = \left(\Theta_0' - \alpha\Theta_0 + \alpha\Theta_0^2\right) + \left(\frac{\sigma^2}{2}\Theta_0'' + \Theta_1' - \alpha\Theta_1 + 2\Theta_0\Theta_1\right)\varepsilon + O\left(\varepsilon^2\right). \tag{2.19}$$

We do the same procedure for conditions (2.18). After matching the coefficients of the polynomials in (2.19), we get the following ordinary differential equations:

$$\Theta_0' - \alpha\Theta_0 + \alpha\Theta_0^2 = 0, \quad \lim_{\eta\to-\infty}\Theta_0 = 0,\ \lim_{\eta\to\infty}\Theta_0 = 1,\ \Theta_0(0) = \frac{1}{2}; \tag{2.20}$$

$$\frac{\sigma^2}{2}\Theta_0'' + \Theta_1' - \alpha\Theta_1 + 2\alpha\Theta_0\Theta_1 = 0, \quad \lim_{\eta\to-\infty}\Theta_1 = \lim_{\eta\to\infty}\Theta_1 = \Theta_1(0) = 0. \tag{2.21}$$

The conditions in (2.21) make sense, as the boundary and midpoint values for Θ are the sum of those of Θ_0, Θ_1, and so on.
In Chapter 5.4 of *An Introduction to Nonlinear Partial Differential Equations* by J. David Logan [6], there seems to be a mistake in the rearrangement of terms, resulting in a different ODE for what in our case is Θ_1. Although the source's original equation is slightly different, we suggest a correction below, solving the correct differential equation.
We start by solving (2.20). Since (2.20) is a separable ordinary differential equation, we can solve it easily:

$$\begin{aligned}
\Theta_0'(\eta) - \alpha\Theta_0(\eta) + \alpha\Theta_0^2(\eta) &= 0 \\
\frac{\Theta_0'(\eta)}{\Theta_0(\eta) - \Theta_0^2(\eta)} &= \alpha \\
\left(\ln\left(\frac{\Theta_0(\eta)}{1-\Theta_0(\eta)}\right)\right)' &= \alpha \\
\ln\left(\frac{\Theta_0(\eta)}{1-\Theta_0(\eta)}\right) &= \alpha\eta + C_1 \\
\frac{\Theta_0(\eta)}{1-\Theta_0(\eta)} &= \exp\left(\alpha\eta + C_1\right) \\
\Theta_0(\eta) &= \frac{\exp\left(\alpha\eta + C_1\right)}{1 + \exp\left(\alpha\eta + C_1\right)}.
\end{aligned}$$

From (2.20), we have

$$\Theta_0(0) = \frac{\exp\left(C_1\right)}{1 + \exp\left(C_1\right)} = \frac{1}{2} \Rightarrow C_1 = 0,$$

thus, for Θ_0, we get:

$$\Theta_0(\eta) = \frac{\exp\left(\alpha\eta\right)}{1 + \exp\left(\alpha\eta\right)}. \tag{2.22}$$

The boundary conditions indeed hold:

$$\lim_{\eta\to-\infty}\Theta_0(\eta)=\lim_{\eta\to-\infty}\frac{\exp(\alpha\eta)}{1+\exp(\alpha\eta)}=0,\quad \lim_{\eta\to\infty}\Theta_0(\eta)=\lim_{\eta\to\infty}\frac{1}{1+\exp(-\alpha\eta)}=1.$$

We now turn to (2.21). The derivatives of Θ_0 are:

$$\big(\Theta_0(\eta)\big)'=\frac{\alpha\exp(\alpha\eta)\big(1+\exp(\alpha\eta)\big)-\alpha\exp(2\alpha\eta)}{\big(1+\exp(\alpha\eta)\big)^2}=\frac{\alpha\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2},$$

$$\begin{aligned}(\Theta_0(\eta))''&=\frac{\alpha^2\exp(\alpha\eta)\big(1+\exp(\alpha\eta)\big)^2-2\alpha^2\exp(2\alpha\eta)\big(1+\exp(\alpha\eta)\big)}{\big(1+\exp(\alpha\eta)\big)^4}=\\&=\frac{\alpha^2\exp(\alpha\eta)\big(1-\exp(\alpha\eta)\big)}{\big(1+\exp(\alpha\eta)\big)^3}.\end{aligned}\tag{2.23}$$

Substituting (2.22) and (2.23) into (2.21), we get the following ordinary differential equation:

$$\frac{\sigma^2\alpha^2}{2}\frac{\exp(\alpha\eta)\big(1-\exp(\alpha\eta)\big)}{\big(1+\exp(\alpha\eta)\big)^3}+\Theta_1'(\eta)-\alpha\Theta_1(\eta)+2\alpha\frac{\exp(\alpha\eta)}{1+\exp(\alpha\eta)}\Theta_1(\eta)=0,$$

which can be written as a first-order linear ODE:

$$\begin{aligned}&\Theta_1'(\eta)=a(\eta)\Theta_1(\eta)+b(\eta),\ \text{ where}\\&a(\eta)=\alpha-2\alpha\frac{\exp(\alpha\eta)}{1+\exp(\alpha\eta)},\quad b(\eta)=\frac{\sigma^2\alpha^2}{2}\frac{\exp(\alpha\eta)\big(1-\exp(\alpha\eta)\big)}{\big(1+\exp(\alpha\eta)\big)^3}.\end{aligned}\tag{2.24}$$

We compute $\int a(\eta)\,\mathrm{d}\eta$:

$$\int a(\eta)\,\mathrm{d}\eta=\int\left(\alpha-2\alpha\frac{\exp(\alpha\eta)}{1+\exp(\alpha\eta)}\right)\mathrm{d}\eta=\alpha\eta-2\ln(\exp(\alpha\eta)+1)+C_2.$$

By multiplying both sides of (2.24) by $\exp\big(-\alpha\eta+2\ln(\exp(\alpha\eta)+1)\big)$, we get:

$$\Theta_1(\eta)=\frac{\exp(\alpha\eta)}{\big(1+\exp(\alpha\eta)\big)^2}\int b(\eta)\frac{\big(\exp(\alpha\eta)+1\big)^2}{\exp(\alpha\eta)}\,\mathrm{d}\eta.\tag{2.25}$$

We can compute the integral in (2.25):

$$\begin{aligned}&\int b(\eta)\frac{\big(1+\exp(\alpha\eta)\big)^2}{\exp(\alpha\eta)}\,\mathrm{d}\eta=\frac{\sigma^2\alpha^2}{2}\int\frac{1-\exp(\alpha\eta)}{1+\exp(\alpha\eta)}\,\mathrm{d}\eta=\\&=\frac{\sigma^2\alpha^2}{2}\left(\eta-\frac{2\ln\big(1+\exp(\alpha\eta)\big)}{\alpha}+C_3\right).\end{aligned}$$

Therefore, we get:

$$\Theta_1(\eta) = \frac{\exp(\alpha\eta)}{\big(1+\exp(\alpha\eta)\big)^2}\frac{\sigma^2\alpha^2}{2}\bigg(\eta - \frac{2\ln\big(1+\exp(\alpha\eta)\big)}{\alpha} + C_3\bigg).$$

Since we know that

$$\Theta_1(0) = \frac{\sigma^2\alpha^2}{8}\bigg(-\frac{2\ln(2)}{\alpha} + C_3\bigg) = 0,$$

therefore, $C_3 = \dfrac{2\ln(2)}{\alpha}$.

We end up with:

$$\Theta_1(\eta) = \frac{\exp(\alpha\eta)}{\big(1+\exp(\alpha\eta)\big)^2}\frac{\sigma^2\alpha^2}{2}\bigg(\eta - \frac{2\ln\big(1+\exp(\alpha\eta)\big)}{\alpha} + \frac{2\ln(2)}{\alpha}\bigg).$$

We need to show that Θ_1 also satisfies the boundary conditions in (2.21):

$$\lim_{\eta\to\pm\infty}\frac{\sigma^2\alpha^2}{2}\bigg(\frac{\eta\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2} - \frac{2\ln(\exp(\alpha\eta)+1)\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2\alpha} + \frac{2\ln(2)\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2\alpha}\bigg) =$$

$$= \frac{\sigma^2\alpha^2}{2}\bigg(\lim_{\eta\to\pm\infty}\frac{\eta\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2} - \frac{2}{\alpha}\lim_{\eta\to\pm\infty}\frac{\ln(1+\exp(\alpha\eta))\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2} + \\ + \frac{2\ln(2)}{\alpha}\lim_{\eta\to\pm\infty}\frac{\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2}\bigg)$$

For the second limit, we use L'Hôpital's rule when $\eta\to+\infty$:

$$\lim_{\eta\to\infty}\frac{\ln(1+\exp(\alpha\eta))\exp(\alpha\eta)}{(1+\exp(\alpha\eta))^2} = \lim_{\eta\to\infty}\frac{\ln(1+\exp(\alpha\eta))}{\exp(\alpha\eta)+2+\exp(-\alpha\eta)} =$$
$$= \lim_{\eta\to\infty}\frac{(\alpha\exp(\alpha\eta))/(1+\exp(\alpha\eta))}{\alpha\exp(\alpha\eta)-\alpha\exp(-\alpha\eta)} = \lim_{\eta\to\infty}\frac{1}{(1-\exp(-2\alpha\eta))(1+\exp(\alpha\eta))} =$$
$$= \lim_{\eta\to\infty}\frac{1}{\exp(\alpha\eta)+1-\exp(-\alpha\eta)-\exp(-2\alpha\eta)} = 0.$$

It is easy to see that the rest of the limits are also equal to zero. Writing the solution using the original variables, we get the final form of our approximation:

$$\phi(\xi) = \frac{\exp\big(\alpha\xi c^{-1}\big)}{1+\exp\big(\alpha\xi c^{-1}\big)} + \frac{1}{c^2}\frac{\exp\big(\alpha\xi c^{-1}\big)}{\big(\exp\big(\alpha\xi c^{-1}\big)+1\big)^2}\frac{\sigma^2\alpha^2}{2}\cdot$$
$$\cdot\bigg(\frac{\xi}{c} - \frac{2\ln\big(\exp\big(\alpha\xi c^{-1}\big)+1\big)}{\alpha} + \frac{2\ln(2)}{\alpha}\bigg) + O\bigg(\frac{1}{c^4}\bigg).$$

We obtained an approximation to the solution of (LA). Conversely, we also described the initial distributions for which (LA) admits traveling wave solutions. We show some examples of such cumulative distribution functions in Figure 2, where $c_{\min} = \sigma\sqrt{2\alpha}$ is the minimal theoretical wave speed.

To further investigate the problem, we seek answers to the following questions in the next section.

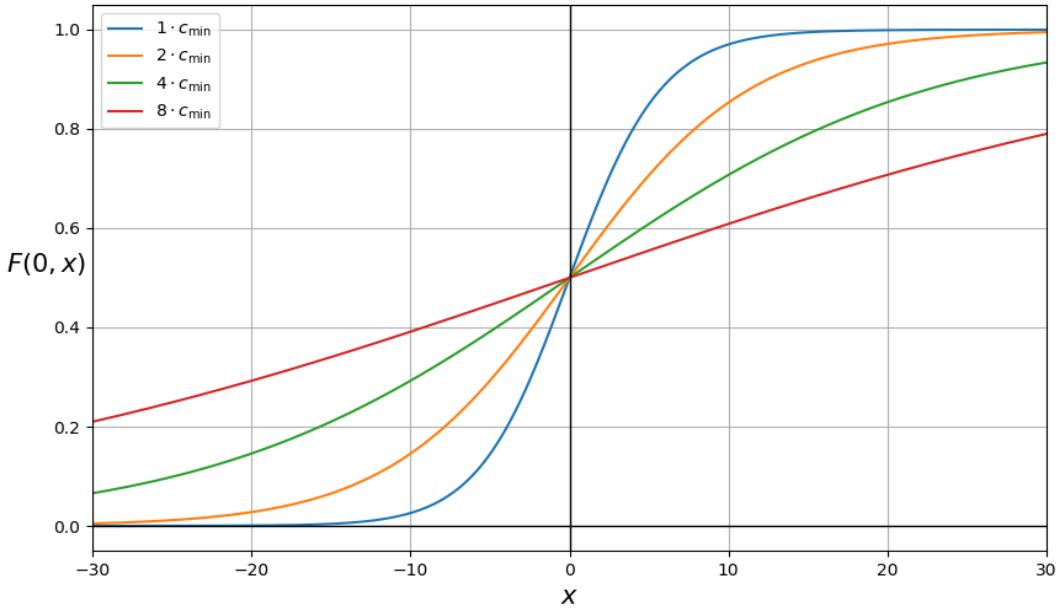


Figure 2: Possible initial conditions for different c values.

(i) Does (LA) admit traveling wave solutions with different initial conditions?

(ii) What do we know about the traveling wave's speed in those cases?

2.5 Solving the PDE numerically

In many complex systems, exact solutions to differential equations are often difficult or impossible to obtain due to their nonlinearity, high dimensionality, or other factors. In such cases, we can use numerical methods to approximate the solutions. These methods involve discretizing the equations and solving them iteratively, allowing for the analysis of complex problems where analytical solutions are not feasible. Below, we present and apply two methods to numerically solve equation (LA).

Finite difference method

The first approach to solving (LA) is to use a finite difference method. We solve (LA) on a discretized grid. The space interval $[-L, L]$ is discretized as follows:

$$x_j = x_0 + jh, \quad j = 0, 1, \ldots, N,$$

where $h > 0$ is the spatial step size given by $h = 2L/N$. Similarly, the time domain $[0, T]$ is discretized as

$$t_n = t_0 + n\tau, \quad n = 0, 1, \ldots, J,$$

where τ is the time step given by $\tau = T/J$. From now on, $y_j^n \approx F(t_n, x_j)$ denotes the approximate solution at t_n and x_j. We initialize the solution with the initial condition:

$$y_j^0 = F_0(x_j), \quad \text{for } j = 0, 1, \ldots, N.$$

We approximate the derivatives as follows:

$$\partial_t F(t_n, x_j) \approx \frac{y_j^{n+1} - y_j^n}{\tau}, \quad \text{and}$$
$$\partial_{xx} F(t_n, x_j) \approx \frac{y_{j+1}^n - 2y_j^n + y_{j-1}^n}{h^2}.$$

Substituting these approximations into (LA), we get the update rule:

$$y_j^{n+1} = \tau\Big(\frac{\sigma^2}{2}\frac{y_{j+1}^n - 2y_j^n + y_{j-1}^n}{h^2} - \alpha y_j^n(1 - y_j^n)\Big) + y_j^n,$$

for $n = 0, 1, \dots J-1$ and $j = 1, 2, \dots N-1$. In this implementation, we use Dirichlet boundary conditions, setting the values at the boundaries as follows:

$$y_0^n = 0, \quad y_N^n = 1 \quad \text{for all } n \in \{0, 1, \dots J\},$$

imitating the original boundary conditions (2.8)-(2.9). We note that this gives us a reliable approximation if L is large enough.

Operator splitting

As for the second numerical approach, we use operator splitting, a method in which we divide (LA) into two subproblems. The advantage of using this method in this case is that the exact solution is known for both subproblems.
Using the same discretization as in the finite difference method, we consider the two initial value problems at each time step $k = 1, 2, \dots, J$, we have:

$$\begin{cases} \partial_t y_1^{(k)}(t,x) = \dfrac{\sigma^2}{2}\partial_{xx} y_1^{(k)}(t,x), \quad t \in \big((k-1)\tau, k\tau\big] \\ y_1^{(k)}\big((k-1)\tau, x\big) = y_2^{(k-1)}\big((k-1)\tau, x\big). \end{cases} \tag{2.26}$$

$$\begin{cases} \partial_t y_2^{(k)}(t,x) = -\alpha y_2^{(k)}(t,x)\big(1 - y_2^{(k)}(t,x)\big), \quad t \in \big((k-1)\tau, k\tau\big] \\ y_2^{(k)}\big((k-1)\tau, x\big) = y_1^{(k)}(k\tau, x). \end{cases} \tag{2.27}$$

In order to initialize the method, we set $y_2^{(0)}(0,x) = F_0(x)$. To solve the initial value problem (2.26) we impose the following boundary conditions:

$$y_1^{(k)}(t, -L) = 0, \quad y_1^{(k)}(t, L) = 1, \tag{2.28}$$

which are consistent with the Dirichlet boundary conditions used in the finite difference method. However, this introduces some errors to our numerical solution, but this will not jeopardize our approach. We can solve (2.26) with the boundary conditions (2.28) by a Fourier series expansion. Due to the inhomogeneous boundary conditions, we first perform a homogenization step. To this end, let $\zeta : [-L, L] \to \mathbb{R}$, $\zeta(x) := (x + L)/(2L)$. We are looking for a solution in the following form:

$$y_1^{(k)}(t,x) = u(t,x) + \zeta(x), \tag{2.29}$$

where $u : [0,T] \times [-L,L] \to \mathbb{R}$. Substituting (2.29) into (2.26), we get the following initial-boundary value problem:

$$\begin{cases} \partial_t u(t,x) = \dfrac{\sigma^2}{2} \partial_{xx} u(t,x), \\ u(t,-L) = 0, \quad u(t,L) = 0, \\ u(0,x) = y_2^{(k-1)}\big((k-1)\tau, x\big) - \zeta(x). \end{cases} \tag{2.30}$$

We know that the Fourier-series expansion of the solution of (2.30) is:

$$u(t,x) = \sum_{n=1}^{\infty} b_n \exp\left(-\frac{\sigma^2}{2}\left(\frac{n\pi}{2L}\right)^2 (t-(k-1)\tau)\right) \sin\left(\frac{n\pi(x+L)}{2L}\right),$$

where the coefficients b_n are given by

$$b_n = \frac{1}{L}\int_{-L}^{L} \Big(y_2^{(k-1)}\big((k-1)\tau, x\big) - \zeta(x)\Big) \sin\left(\frac{n\pi(x+L)}{2L}\right) \mathrm{d}x.$$

Therefore, the solution of (2.26) with boundary conditions (2.28) takes the following form:

$$y_1^{(k)}(t,x) = \zeta(x) + \sum_{n=1}^{\infty} b_n \exp\left(-\frac{\sigma^2}{2}\left(\frac{n\pi}{2L}\right)^2 (t-(k-1)\tau)\right) \sin\left(\frac{n\pi(x+L)}{2L}\right).$$

As for the reaction step (2.27), we can solve it explicitly since it is a separable ordinary differential equation, which yields:

$$y_2^{(k)}(t,x) = \frac{y_1^{(k)}(k\tau,x)}{y_1^{(k)}(k\tau,x) + \big(1 - y_1^{(k)}(k\tau,x)\big)\exp\big(\alpha(t-(k-1)\tau)\big)}.$$

After completing the J-th step, we stop the iteration. The numerical solution obtained at this stage serves as an approximation of y at time T:

$$y(T,x) \approx y_2^{(J)}(T,x).$$

2.6 Implementation and comparison

Section 2.4 provided an analytical approximation of the traveling wave solutions of (LA) with initial conditions that resemble the sigmoid function. In this section, we analyze the numerical solutions obtained by using the finite difference method and operator splitting. The numerical methods in Section 2.5 were implemented in Python. The parameters of (LA) were set to $\alpha = 1$ and $\sigma = 1$ throughout this section.

In Section 2.4 we expressed the traveling wave solution of (LA) as a series in powers of c^{-2}. Therefore, a natural candidate for the initial cumulative distribution function would be:

$$F_0(x) = \frac{\exp\big(\alpha x c^{-1}\big)}{1 + \exp\big(\alpha x c^{-1}\big)},$$

as it is the leading-order term in the approximation obtained in Section 2.4. The results are plotted in Figure 3, shown over the interval $x \in [-40, 40]$ for clarity. The exact solution of (LA) was implemented simply as a shift of the initial condition. To respect the boundary conditions, in the implementation, (LA) was solved in a slightly larger domain, the central portion already illustrates the behavior of the solution well.

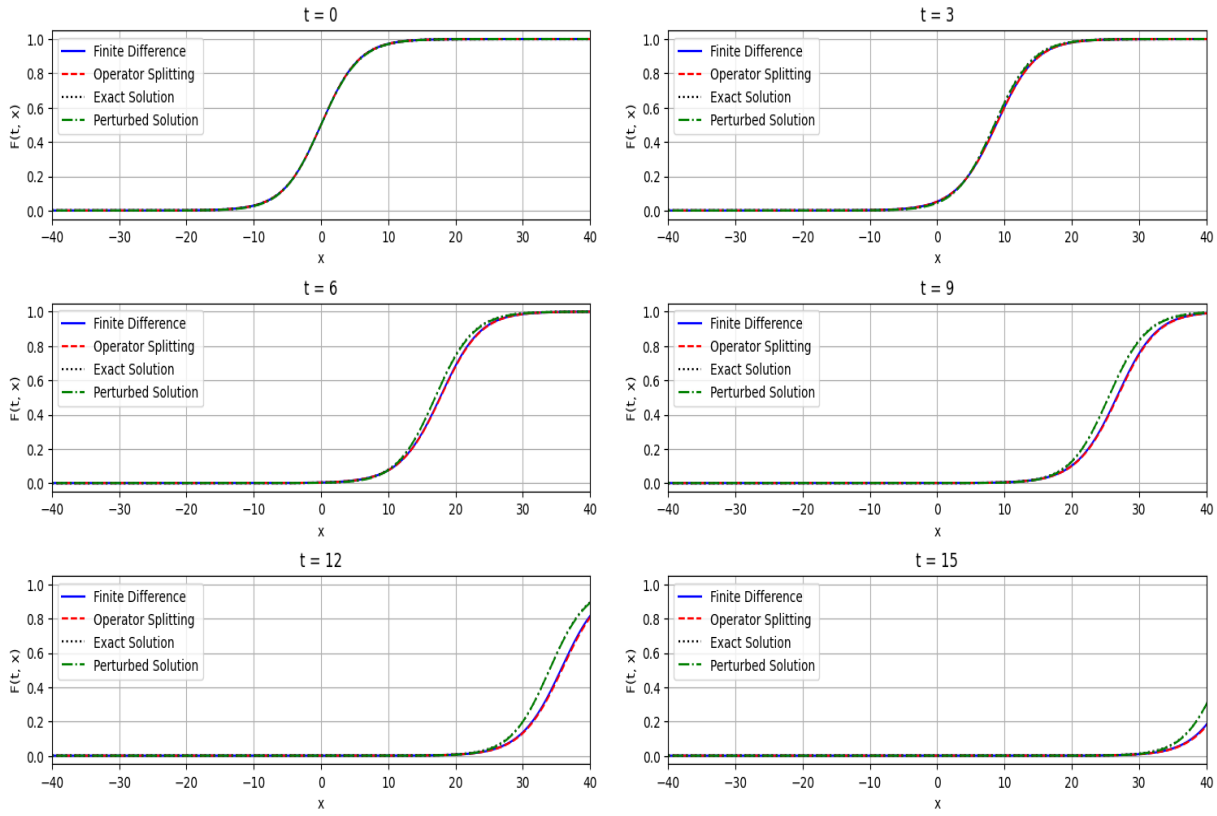


Figure 3: Traveling wave solution of (LA) with $F_0(x) = \exp\left((\sqrt{2}/4)x\right)/\left(1 + \exp\left((\sqrt{2}/4)x\right)\right)$ and wave speed $c^* = 2\sqrt{2}$, using $L = 50$, $T = 15$, $h = 0.1$, and $\tau = 0.01$.

We can see that our approximation obtained by the perturbation method in Section 2.4 is very accurate even for small c values. As we showed earlier, the solution of (LA) is a traveling wave.

We may also try to experiment with different initial distributions. First, we consider an initial condition, for which we did not prove the existence of a traveling wave solution. The results for the KPP equation in the paper of Kolmogorov et al. [5] suggest that if the initial knowledge distribution is the Dirac delta, there exists a traveling wave solution, and the wave speed is equal to the minimal wave speed, i.e., $c^* = c_{\min} = \sigma\sqrt{2\alpha}$ in our case. Although equation (LA) is slightly different from the KPP equation, we can still check if the numerical result yields a traveling wave. Since we are testing for the Dirac delta as the initial distribution, the corresponding cumulative distribution function is:

$$F_0(x) = \mathbb{1}_{x \geq 0}.$$

The true wave speed in question is $c^* \stackrel{?}{=} \sqrt{2}$. The results of the two numerical methods are presented in Figure 4, again using a slightly smaller interval than $[-L, L]$.

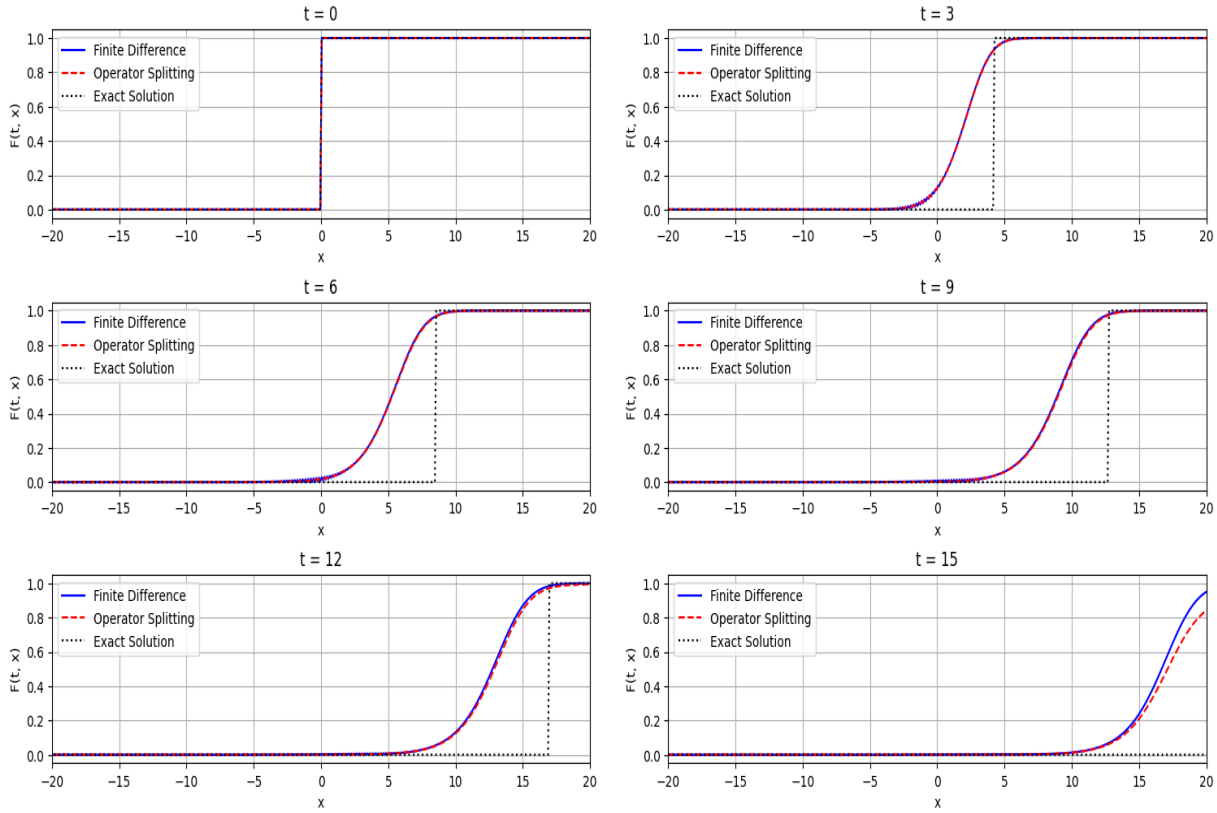


Figure 4: Traveling wave solution of (LA) with $F_0(x) = \mathbb{1}_{x \geq 0}$ and wave speed $c^* = \sqrt{2}$, using $L = 30$, $T = 15$, $h = 0.1$, and $\tau = 0.01$.

As we can see, the numerical solutions are traveling wave solutions and they indeed follow the shifted initial condition. Note that the wave profiles of the numerical solutions differ from those of the exact solution, because these numerical methods tend to smooth out sharp discontinuities over the iteration steps. This is because of the errors that arise from discretizing both the space and time variables.

Lastly, we show an example of another initial distribution that seems to be a viable initial condition to (LA) for traveling wave solutions, i.e., the exponential distribution. The results can be seen in Figure 5. Although we do not give a definitive proof to this observation, but the traveling wave speed seems to be $c_{\min}$ under this initial condition as well.

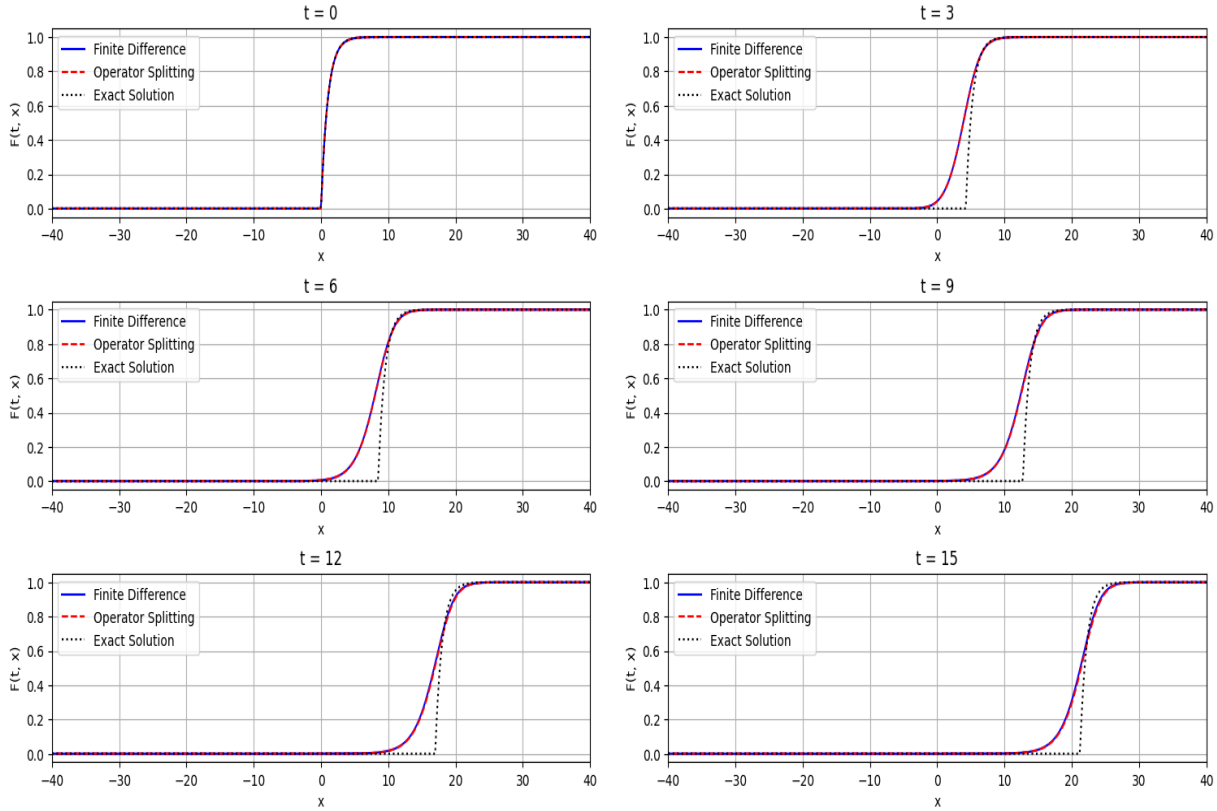


Figure 5: Traveling wave solution of (LA) with $F_0(x)$ being the cumulative distribution function of the exponential distribution, and wave speed $c^* = \sqrt{2}$, using $L = 50$, $T = 15$, $h = 0.1$, and $\tau = 0.01$.

Note that these results need to be taken with a grain of salt, regardless of how convincing they may appear. Our goal is not to prove the validity of our hypotheses on the value of c^*, but to give us a hint as to whether our assumptions are true or not. Additionally, it gives us an opportunity to compare the numerical methods themselves against each other.

Chapter 3

Optimal control problems

This chapter can be thought of as an enhancement of the foundation we built in Chapter 1. Optimal control theory provides a mathematical framework for solving decision-making problems over time. These methods are crucial in economics and finance, as people face decisions about their purchases every day. What makes this problem particularly intriguing is that, in most cases, we face evolving constraints that are often difficult to predict.

3.1 Preliminaries

The definitions presented in this section are primarily based on Lawrence C. Evans's work [10], and on Chapters 1-3 of Huyên Pham's book [11] on stochastic optimization.

In the context of optimal control theory, space (and time) variables are typically accompanied by control variables that represent the decision-making processes of the agents within the system. To motivate the mathematical setup for these equations, we first introduce the concept of controls.

Definition 3.1. Let $A \subset \mathbb{R}^m$, be a non-empty bounded set called the control set. The stochastic process $\{\alpha(s)\}_{s\geq 0}$ is called a control (or control process) if $\alpha : [0,\infty) \to A$ is progressively measurable. We denote the set of all controls by $\mathcal{A}_0$.

We consider the following controlled stochastic differential equation on $\mathbb{R}^n$:

$$\begin{cases} \mathrm{d}\mathbf{X}(s) = \boldsymbol{\mu}(\mathbf{X}(s), \alpha(s))\,\mathrm{d}s + \boldsymbol{\sigma}\big(\mathbf{X}(s), \alpha(s)\big)\,\mathrm{d}\mathbf{B}(s), \\ \mathbf{X}(t) = \underline{x}, \end{cases} \tag{CSDE}$$

where $\underline{x} \in \mathbb{R}^n$, $t \in [0, s]$ and the functions $\boldsymbol{\mu} : \mathbb{R}^n \times A \to \mathbb{R}^n$, $\boldsymbol{\sigma} : \mathbb{R}^n \times A \to \mathbb{R}^{n\times d}$ are measurable functions that are Lipschitz continuous in their first variable. Note that we could define $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ such that they depend on s as well, however, since our focus is on the case of an infinite time horizon, we do not extend our setup in this direction.

Proposition 3.2. *If there exists a constant $L > 0$, such that*

$$\|\boldsymbol{\mu}(\underline{x}, \underline{a})\|^2 + \|\boldsymbol{\sigma}(\underline{x}, \underline{a})\|_F^2 \leq L\left(1 + \|\underline{x}\|^2\right) \quad \textit{for all } \underline{x} \in \mathbb{R}^n,\ \underline{a} \in A, \tag{3.1}$$

then (CSDE) *has a unique solution, which we denote by $\{\mathbf{X}_{\underline{x}}(s)\}_{s\geq 0}$. From now on, condition* (3.1) *will be assumed.*

Assumption 3.3. From now on, we consider controls in the form $\alpha(s) = \mathbb{A}(\mathbf{X}_{\underline{x}}(s))$, where $\mathbb{A} : [0, \infty) \times \mathbb{R}^n \to A$ is a measurable function. This means that the actions we take only depend on the current state.

Note that for finite-horizon problems, the control would also explicitly depend on time as well, meaning that we would have controls in the form $\alpha(s) = \mathbb{A}(s, \mathbf{X}_{\underline{x}}(s))$.

The next step of our setup is to define a payoff function $\mathbf{P}_{\underline{x}} : \mathcal{A}_0 \to \overline{\mathbb{R}}$ that describes how "good" a control is if we start from the state $\underline{x}$. In order to align with our models in this chapter, we define it as:

$$\mathbf{P}_{\underline{x}}(\alpha) := \mathbb{E}\left[\int_0^\infty \exp(-qs) f\big(\mathbf{X}_{\underline{x}}(s), \alpha(s)\big)\, \mathrm{d}s\right] \quad \text{for all } \underline{x} \in \mathbb{R}^n,$$

where $q > 0$ and $f : \mathbb{R}^n \times A \to \mathbb{R}$ is a given measurable function.

Definition 3.4. We say that the control $\alpha \in \mathcal{A}_0$ is admissible if the following holds:

$$\mathbb{E}\left[\int_0^\infty \exp(-qs) \big| f\big(\mathbf{X}_{\underline{x}}(s), \alpha(s)\big)\big|\, \mathrm{d}s\right] < \infty.$$

We denote the set of admissible controls as $\mathcal{A}_{\underline{x}}$.

Assumption 3.5. We assume a quadratic growth condition for the function f, i.e., there exists a constant $K > 0$, such that:

$$|f(\underline{z}, \underline{a})| \leq K\left(1 + \|\underline{z}\|^2\right) \quad \text{for all } \underline{z} \in \mathbb{R}^n, \underline{a} \in A.$$

Proposition 3.6. *Under Assumption 3.5, every control in $\mathcal{A}_0$ is admissible.*

Definition 3.7. We say that the control $\alpha_{\text{opt}} \in \mathcal{A}_{\underline{x}}$ is optimal if for any admissible control $\alpha \in \mathcal{A}_{\underline{x}}$, we have

$$\mathbf{P}_{\underline{x}}(\alpha_{\text{opt}}) \geq \mathbf{P}_{\underline{x}}(\alpha).$$

This is of great importance for optimization, as α_{opt} determines the best course of action for the agent, given the dynamics and uncertainties in the model. This motivates us to define a special function, called the value function $v : \mathbb{R}^n \to \mathbb{R}$, which tells us what the best possible payoff starting from state x is, defined as:

$$v(\underline{x}) = \sup_{\alpha \in \mathcal{A}_{\underline{x}}} \{\mathbf{P}_{\underline{x}}(\alpha)\}.$$

From now, we assume that v is sufficiently smooth, i.e., $v \in C^2$.
We now introduce a concept within controlled differential equations that is essential for the later developments in this chapter: the dynamic programming principle. The basic idea is to separate a difficult problem into subproblems that are easier to solve. This is one of the reasons we keep track of the initial states in our notation. Intuitively, we expect an optimal strategy to have the property that, no matter the initial state and time, the remaining decisions must be optimal as well. The dynamic programming principle is a rigorous formulation of this idea.
The proof of the following theorem is based on Huyeñ's work in [11], with the modification that we now consider an infinite time horizon.

Theorem 3.8. *Let $\underline{x} \in \mathbb{R}^n$, and $\mathcal{T}$ be a set of stopping times on $[0,\infty)$. Then the value function v satisfies the following:*

$$v(\underline{x}) = \sup_{\alpha\in\mathcal{A}_{\underline{x}}}\left\{\sup_{\tau\in\mathcal{T}}\left\{\mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)v(\mathbf{X}_{\underline{x}}(\tau))\right]\right\}\right\} =$$
$$= \sup_{\alpha\in\mathcal{A}_{\underline{x}}}\left\{\inf_{\tau\in\mathcal{T}}\left\{\mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)v(\mathbf{X}_{\underline{x}}(\tau))\right]\right\}\right\},$$

with $\exp(-q\theta(\omega)) = 0$ *if* $\theta(\omega) = \infty$.

Proof. Let $\alpha \in \mathcal{A}_{\underline{x}}$, and $\tau \in \mathcal{T}$ be fixed.
Observe that for all $\tau \in \mathcal{T}$, we have the Markov property $\mathbf{X}_{\underline{x}}(s) = \mathbf{X}_{\mathbf{X}_{\underline{x}}(\tau)}(s-\tau)$ if $s \geq \tau$. This means that if we stop at some time t, and we continue from that state up to time s, we end up in the same state as if we had gone to $\mathbf{X}_{\underline{x}}(s)$ directly, under the same control. We split the integral into two parts, then we use the tower rule:

$$\mathbf{P}_{\underline{x}}(\alpha) = \mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \int_\tau^\infty \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s\right] =$$
$$= \mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s\right] +$$
$$+ \mathbb{E}\left[\mathbb{E}\left[\int_\tau^\infty \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s\middle|\mathcal{F}_\tau\right]\right].$$

By the definition of $\mathbf{P}_{\underline{x}}$ and the Markov behavior we observed earlier, we have

$$\mathbb{E}\left[\int_\tau^\infty \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s\middle|\mathcal{F}_\tau\right] = \exp(-q\tau)\mathbf{P}_{\mathbf{X}_{\underline{x}}(\tau)}(\alpha),$$

therefore, we can express $\mathbf{P}_{\underline{x}}$ as:

$$\mathbf{P}_{\underline{x}}(\alpha) = \mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)\mathbf{P}_{\mathbf{X}_{\underline{x}}(\tau)}(\alpha)\right]$$

Since α is an arbitrary control, we get the following inequality:

$$\mathbf{P}_{\underline{x}}(\alpha) \leq \mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)v\big(\mathbf{X}_{\underline{x}}(\tau)\big)\right],$$

for all $\tau \in \mathcal{T}$. Therefore, can write:

$$\mathbf{P}_{\underline{x}}(\alpha) \leq \inf_{\tau\in\mathcal{T}}\left\{\mathbb{E}\left[\int_0^\tau \exp(-qs)f\big(\mathbf{X}_{\underline{x}}(s),\alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)v\big(\mathbf{X}_{\underline{x}}(\tau)\big)\right]\right\}. \tag{3.2}$$

Furthermore, we know that for all $\varepsilon > 0$, there is an α_ε control, that is suboptimal only by ε, i.e.:

$$v\big(\mathbf{X}_{\underline{x}}(\tau)\big) - \varepsilon \leq \mathbf{P}_{\mathbf{X}_{\underline{x}}(\tau)}(\alpha_\varepsilon). \tag{3.3}$$

We now define the following process:

$$\alpha^*(s) = \begin{cases}\alpha(s) & \text{if } s \in [0,\tau],\\ \alpha_\varepsilon(s) & \text{if } s \in [\tau,\infty).\end{cases}$$

An important detail for the correctness of this proof is that $\alpha^*(s)$ is also an admissible control. This technicality will not be proved here, but for those interested in this lemma, we recommend Chapter 7 of *Continuous-time Stochastic Control and Optimization with Financial Applications* [11].

With a similar argument as before and using (3.3), for all τ we get:

$$\begin{aligned}\mathbf{P}_{\underline{x}}(\alpha^*) &= \mathbb{E}\left[\int_0^\tau \exp(-qs) f\big(\mathbf{X}_{\underline{x}}(s), \alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau)\mathbf{P}_{\mathbf{X}_{\underline{x}}(\tau)}(\alpha_\varepsilon)\right] \geq \\ &\geq \mathbb{E}\left[\int_0^\tau \exp(-qs) f\big(\mathbf{X}_{\underline{x}}(s), \alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau) v(\mathbf{X}_{\underline{x}}(\tau))\right] - \varepsilon.\end{aligned}$$

Since this holds for every $\tau \in \mathcal{T}$, we get the following:

$$\mathbf{P}_{\underline{x}}(\alpha) \geq \sup_{\tau\in\mathcal{T}} \left\{\mathbb{E}\left[\int_0^\tau \exp(-qs) f\big(\mathbf{X}_{\underline{x}}(s), \alpha(s)\big)\,\mathrm{d}s + \exp(-q\tau) v\big(\mathbf{X}_{\underline{x}}(\tau)\big)\right]\right\}. \tag{3.4}$$

If we take the supremum over all $\alpha \in \mathcal{A}_{\underline{x}}$ in equations (3.2) and (3.4), we end up with the desired result. □

3.2 The Hamilton–Jacobi–Bellman equation

Our aim is to leverage some of the results in Chapter 1 in the context of controlled stochastic differential equations (CSDE). Namely, we are going to see another application of Itô's lemma (Lemma 1.20), as it helps us to derive a nonlinear partial differential equation for the value function v.

The following steps are analogous to those in Chapter 3 of Huyên's book [11], however, similarly as we did at the dynamic programming principle, we provide a proof for the infinite time horizon case.

Assumption 3.9. For every $t \in \mathbf{T}$, we assume

$$\int_0^t \left\|\exp(-qs)\nabla v(\mathbf{X}_{\underline{x}}(s))\boldsymbol{\sigma}(\mathbf{X}_{\underline{x}}(s), \underline{a})\right\|^2 \mathrm{d}s < \infty.$$

Let $v \in C^2$ be the value function. Let us fix some $t \geq 0$, $h > 0$ and a constant control $\underline{a} \in A$ on $[t, t+h]$. From the dynamic programming principle, Theorem 3.8 we have:

$$v(\underline{x}) \geq \mathbb{E}\left[\int_t^{t+h} \exp(-qs) f(\mathbf{X}_{\underline{x}}(s), \underline{a})\,\mathrm{d}s + \exp(-q(t+h)) v(\mathbf{X}_{\underline{x}}(t+h))\right]. \tag{3.5}$$

Using Itô's lemma (Lemma 1.20) to $\exp(-qs)v(\mathbf{X}_{\underline{x}}(s))$ on the interval $[t, t+h]$ according

to (CSDE), and multiplying both sides by $\exp(-qt)$, we get:

$$\begin{aligned}\exp(-qh)v(\mathbf{X}_{\underline{x}}(t+h)) = v(x) &+ \int_t^{t+h} \exp(-qs)(-q)v(\mathbf{X}_{\underline{x}}(s))\,\mathrm{d}s + \\ &+ \int_t^{t+h} \exp(-qs)\nabla v(\mathbf{X}_{\underline{x}}(s))\boldsymbol{\sigma}(\mathbf{X}_{\underline{x}}(s),\underline{a})\,\mathrm{d}\mathbf{B}(s) + \\ &+ \int_t^{t+h} \exp(-qs)\nabla v(\mathbf{X}_{\underline{x}}(s))\boldsymbol{\mu}(\mathbf{X}_{\underline{x}}(s),\underline{a})\,\mathrm{d}s + \\ &+ \int_t^{t+h} \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\mathbf{X}_{\underline{x}}(s),\underline{a})\boldsymbol{\sigma}^\top(\mathbf{X}_{\underline{x}}(s),\underline{a})\exp(-qs)\nabla^2 v(\mathbf{X}_{\underline{x}}(s))\right]\,\mathrm{d}s.\end{aligned}$$

We take the expected value on both sides. Because of Proposition 1.12 and Assumption 3.9, the Itô integral term vanishes, leaving us with:

$$\begin{aligned}\mathbb{E}\left[\exp(-q(t+h))v(\mathbf{X}_{\underline{x}}(t+h))\right] = v(\underline{x}) &+ \mathbb{E}\left[\int_t^{t+h} \exp(-qs)(-q)v(\mathbf{X}_{\underline{x}}(s))\,\mathrm{d}s\right] + \\ &+ \mathbb{E}\left[\int_t^{t+h} \exp(-qs)\nabla v\big(\mathbf{X}_{\underline{x}}(s)\big)\boldsymbol{\mu}(\mathbf{X}_{\underline{x}}(s),\underline{a})\,\mathrm{d}s\right] + \\ &+ \mathbb{E}\left[\int_t^{t+h} \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\mathbf{X}_{\underline{x}}(s),\underline{a})\boldsymbol{\sigma}^\top(\mathbf{X}_{\underline{x}}(s),\underline{a})\exp(-qs)\nabla^2 v\big(\mathbf{X}_{\underline{x}}(s)\big)\right]\,\mathrm{d}s\right].\end{aligned}$$

By substituting $\exp(-q(t+h))\mathbb{E}\left[v(\mathbf{X}_{\underline{x}}(t+h))\right]$ into (3.5) and subtracting $v(\underline{x})$ form both sides, we get:

$$\begin{aligned}0 \geq \mathbb{E}\left[\int_t^{t+h} \exp(-qs)f(\mathbf{X}_{\underline{x}}(s),\underline{a})\,\mathrm{d}s\right] &+ \mathbb{E}\left[\int_t^{t+h} \exp(-qs)(-q)v(\mathbf{X}_{\underline{x}}(s))\,\mathrm{d}s\right] + \\ &+ \mathbb{E}\left[\int_t^{t+h} \exp(-qs)\nabla v\big(\mathbf{X}_{\underline{x}}(s)\big)\boldsymbol{\mu}(\mathbf{X}_{\underline{x}}(s),\underline{a})\,\mathrm{d}s\right] + \\ &+ \mathbb{E}\left[\int_t^{t+h} \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\mathbf{X}_{\underline{x}}(s),\underline{a})\boldsymbol{\sigma}^\top(\mathbf{X}_{\underline{x}}(s),\underline{a})\exp(-qs)\nabla^2 v(\mathbf{X}_{\underline{x}}(s))\right]\,\mathrm{d}s\right].\end{aligned}$$

By dividing both sides of the inequality by h, and taking the limit as $h \to 0$, we end up with:

$$0 \geq f(\underline{x},\underline{a}) - qv(\underline{x}) + \nabla v\big(\underline{x}\big)\boldsymbol{\mu}(\underline{x},\underline{a}) + \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x},\underline{a})\boldsymbol{\sigma}^\top(\underline{x},\underline{a})\nabla^2 v(\underline{x})\right].$$

Note that $\underline{a}$ was arbitrary in A, therefore, we can take the supremum over all $\underline{a} \in A$ and we get:

$$qv(\underline{x}) - \sup_{\underline{a}\in A}\left\{f(\underline{x},\underline{a}) + \nabla v\big(\underline{x}\big)\boldsymbol{\mu}(\underline{x},\underline{a}) + \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x},\underline{a})\boldsymbol{\sigma}^\top(\underline{x},\underline{a})\nabla^2 v(\underline{x})\right]\right\} \geq 0. \quad (3.6)$$

Since the optimal control α_t^* satisfies (3.5) with equality, we can use the same steps as above to get:

$$qv(\underline{x}) - \left(f(\underline{x},\alpha_t^*) + \nabla v\big(\underline{x}\big)\boldsymbol{\mu}(\underline{x},\alpha_t^*) + \frac{1}{2}\operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x},\alpha_t^*)\boldsymbol{\sigma}^\top(\underline{x},\alpha_t^*)\nabla^2 v(\underline{x})\right]\right) = 0. \quad (3.7)$$

Equations (3.6) and (3.7) together imply that the inequality in (3.6) must hold with equality when the optimal control α_t^* is applied. In other words, the supremum in (3.6) is attained by choosing the optimal control, in which case the value function v satisfies (3.7). This motivates the definition of the Hamilton–Jacobi–Bellman equation.

Definition 3.10. Let $\underline{x} \in \mathbb{R}^n$, and $v \in C^2$ be the value function. The nonlinear elliptic partial differential equation below is called the Hamilton–Jacobi–Bellman equation:

$$qv(\underline{x}) - \sup_{\underline{a} \in A} \left\{ f(\underline{x}, \underline{a}) + \nabla v(\underline{x}) \boldsymbol{\mu}(\underline{x}, \underline{a}) + \frac{1}{2} \operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x}, \underline{a}) \boldsymbol{\sigma}^\top(\underline{x}, \underline{a}) \nabla^2 v(\underline{x}) \right] \right\} = 0. \quad \text{(HJB)}$$

Remark 3.11. Throughout this section, we considered the infinite time horizon case, which gave us the time-independent equation (HJB). However, we could also work through the same steps in the finite time horizon setup, which would yield a similar version of (HJB).

3.3 Verification step

Our next goal is to provide sufficient conditions under which a solution to (HJB) coincides with the value function v. These conditions are given by the so-called verification theorem.

Theorem 3.12. *Let $w : \mathbb{R}^n \to \mathbb{R}$, $w \in C^2$, such that there exists some $K > 0$, for which $|w(\underline{x})| \leq K(1 + \|\underline{x}\|^2)$ for all $\underline{x} \in \mathbb{R}^n$. If for all $\underline{x} \in \mathbb{R}^n$, w satisfies the conditions below:*

(i) $qw(\underline{x}) - \sup_{\underline{a} \in A} \left\{ f(\underline{x}, \underline{a}) + \nabla w(\underline{x}) \boldsymbol{\mu}(\underline{x}, \underline{a}) + \frac{1}{2} \operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x}, \underline{a}) \boldsymbol{\sigma}^\top(\underline{x}, \underline{a}) \nabla^2 w(\underline{x}) \right] \right\} \geq 0$, *and*

(ii) $\limsup_{T \to \infty} \left\{ \exp(-qT) \mathbb{E}\left[w\big(\mathbf{X}_{\underline{x}}(T)\big) \right] \right\} \geq 0$,

then, we have $w \geq v$.
Furthermore, if there exists a measurable function $\widehat{\alpha}(\underline{x}) : \mathbb{R}^n \to A$ that satisfies the below conditions for all $\underline{x} \in \mathbb{R}^n$:

(iii) $$0 = qw(\underline{x}) - \sup_{\underline{a} \in A} \left\{ f(\underline{x}, \underline{a}) + \nabla w(\underline{x}) \boldsymbol{\mu}(\underline{x}, \underline{a}) + \frac{1}{2} \operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x}, \underline{a}) \boldsymbol{\sigma}^\top(\underline{x}, \underline{a}) \nabla^2 w(\underline{x}) \right] \right\} =$$
$$= qw(\underline{x}) - f(\underline{x}, \widehat{\alpha}(\underline{x})) + \nabla w(\underline{x}) \boldsymbol{\mu}(\underline{x}, \widehat{\alpha}(\underline{x})) + \frac{1}{2} \operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x}, \widehat{\alpha}(\underline{x})) \boldsymbol{\sigma}^\top(\underline{x}, \widehat{\alpha}(x)) \nabla^2 w(\underline{x}) \right],$$

(iv) the stochastic differential equation

$$\begin{cases} \mathrm{d}\mathbf{X}(s) = \boldsymbol{\mu}(\mathbf{X}(s), \widehat{\alpha}(\mathbf{X}(s))\, \mathrm{d}s + \boldsymbol{\sigma}(\mathbf{X}(s), \widehat{\alpha}(\mathbf{X}(s))\, \mathrm{d}\mathbf{B}(s), \\ \mathbf{X}(0) = \underline{x}, \end{cases}$$

has a unique solution denoted by $\widehat{\mathbf{X}}_{\underline{x}}(s)$,

(v) $\liminf_{T \to \infty} \left\{ \exp(-qT) \mathbb{E}\left[w(\widehat{\mathbf{X}}_{\underline{x}}(T)) \right] \right\} \leq 0$,

(vi) $\left\{ \widehat{\alpha}(\widehat{\mathbf{X}}_{\underline{x}}(s)) \right\}_{s \geq 0} \in \mathcal{A}_{\underline{x}}$,

then, we have $w = v$.

We do not prove Theorem 3.12 here, but the steps are similar to those of the earlier proofs. We need to use Itô's lemma for $\exp(-qt)w(\mathbf{X}_{\underline{x}}(t))$ on a growing set of finite intervals, then use the dominated convergence theorem. Theorem 3.12 is a powerful tool to check whether a candidate control is optimal or not.

Theorem 3.12 can also be thought of as an infinite horizon version of the Feynman–Kac formula introduced in Section 1.3. This analogy would be more apparent if we had a finite time horizon, as in that case, the HJB equation would be a parabolic PDE. Nonetheless, this scenario is also worth exploring.

Remark 3.13. Let $A = \{\underline{a}_0\}$, $w \in C^2$, such that there exists some $K > 0$, for which $|w(\underline{z})| \leq K(1 + \|\underline{z}\|^2)$ for all $\underline{z} \in \mathbb{R}^n$. If w is the solution to the linear elliptic partial differential equation

$$qw(\underline{x}) - \left(f(\underline{x}, \underline{a}_0) + \nabla w(\underline{x}) \boldsymbol{\mu}(\underline{x}, \underline{a}_0) + \frac{1}{2} \operatorname{Tr}\left[\boldsymbol{\sigma}(\underline{x}, \underline{a}_0) \boldsymbol{\sigma}^\top(\underline{x}, \underline{a}_0) \nabla^2 w(\underline{x}) \right] \right) = 0, \quad \underline{x} \in \mathbb{R}^n,$$

with the condition that $\exp(-qT)\mathbb{E}[w(X_{\underline{x}}(T)] \to 0$ as $T \to \infty$, then w has the following representation:

$$w(\underline{x}) = \mathbb{E}\left[\int_0^\infty \exp(-qt) f\left(\mathbf{X}_{\underline{x}}(t), \underline{a}_0\right) \mathrm{d}t \right],$$

where $\mathbf{X}_{\underline{x}}(t)$ is the unique solution to (CSDE) with $\alpha(s) = \underline{a}_0$.

3.4 Aiyagari–Bewley–Huggett model

This section offers a more detailed and in-depth extension of the material presented in Chapter 2 of *Partial Differential Equations in Macroeconomics* [1].

In this model, agents optimize consumption and savings dynamically in response to stochastic income fluctuations. Our goal is to examine the model's long-run, steady-state behavior, i.e., when quantities and distributions do not change over time. This will lead to a system where a time-independent HJB equation governs individual choices and a so-called Fokker–Planck equation describes wealth distribution in equilibrium.

We can define a consumption function $c_i(t) : [0, \infty) \to [0, \infty)$ for each individual i in our economy, which gives us the agent's consumption at time t.
Each agent has different preferences, risk tolerance, and consumption needs, which is why we use a utility function $U : [0, \infty) \to \mathbb{R}$ to represent their individual satisfaction gained from a purchase. A commonly used utility function is the CRRA (Constant Relative Risk Aversion), given by

$$U(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \text{if } \gamma \in (0, \infty) \setminus \{1\}, \\ \ln(x), & \text{if } \gamma = 1. \end{cases}$$

We call γ the risk-aversion parameter, which quantifies the agent's preference toward risk and uncertainty in decision-making. Since we assume that all investors are risk-averse, we

choose $\gamma > 0$ in order to ensure that the utility function is strictly increasing and concave. A consumer makes decisions about consumption $c_i(t)$ over time, which only depends on the current time and state. Here, the process $\{c_i(t)\}_{t\geq 0}$ is a control introduced in Section 3.1. The goal of each agent is to maximize their lifetime utility, which can be expressed as:

$$\max_{c_i} \mathbb{E}\left[\int_0^\infty \exp\left(-\rho t\right) U\big(c_i(t)\big)\,\mathrm{d}t\right],$$

where $\exp\left(-\rho t\right)$ is the discount factor with the discount rate ρ, reflecting the agent's preference for current consumption over future consumption.
Similarly to the Lucas–Alvarez model introduced in Chapter 2, we measure one's productivity with $Z_i(t) \geq 0$ at time $t \geq 0$. We denote the value of the individual's assets by $A_i(t)$. For $Z_i(t)$ amount of work, the individual gets a wage of $W_i(t)$.
The value of the agent's assets increases with its labor income $W_i(t)Z_i(t)$. They also earn a return on these assets with a rate of $r(t)$. Meanwhile, the agent consumes for $c_i(t)$ amount, reducing their wealth. As a result, the agent's wealth dynamics follow:

$$A_i(t) - A_i(0) = \int_0^t \big(Z_i(s) + r(t)A_i(s) - c_i(s)\big)\,\mathrm{d}s,$$

formally written as:

$$\mathrm{d}A_i(t) = \big(W_i(t)Z_i(t) + r(t)A_i(t) - c_i(t)\big)\,\mathrm{d}t + 0 \cdot \mathrm{d}B_i(t).$$

Furthermore, the agent's productivity evolves stochastically over time, improving due to skill accumulation, but is also subject to "random shocks" such as technological changes or layoffs:

$$Z_i(t) - Z_i(t) = \int_0^t \mu\big(Z_i(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(Z_i(s)\big)\,\mathrm{d}B_i(s),$$

formally written as:

$$\mathrm{d}Z_i(t) = \mu\big(Z_i(t)\big)\,\mathrm{d}t + \sigma\big(Z_i(t)\big)\,\mathrm{d}B(t).$$

We also impose a lower bound on the agents' wealth:

$$a_{\min} \leq A_i(t), \quad \text{for all } t \geq 0, \text{where } a_{\min} \in (-\infty, 0].$$

If the constraint is set at $a_{\min} = 0$, the agent is not allowed to borrow and must always maintain non-negative wealth. When $a_{\min} < 0$, limited borrowing is allowed, but only up to a specified debt ceiling. This restriction prevents agents from accumulating unbounded liabilities. While our theoretical setup does not impose explicit upper bounds on $A_i(t)$ or $Z_i(t)$, in a numerical implementation these variables would necessarily be bounded.

To describe the distribution of households by their wealth a and productivity z, we introduce the probability density function $g(a, z)$, meaning that it satisfies:

$$\int_{a_{\min}}^\infty \int_{-\infty}^\infty g(a, z)\,\mathrm{d}z\,\mathrm{d}a = 1, \quad \text{and } g(a, z) \geq 0 \text{ for all } a \in [a_{\min}, \infty), z \in \mathbb{R}. \tag{3.8}$$

Furthermore, we set the following constraint on g:

$$\int_{a_{\min}}^{\infty} \int_{-\infty}^{\infty} a \cdot g(a,z)\,\mathrm{d}z\,\mathrm{d}a = 0. \tag{3.9}$$

Equation (3.9) means that the total amount of wealth held by agents with positive assets exactly offsets the total debt of those with negative assets.

As workers decide how much to consume and save, the overall distribution of their wealth and work productivity changes over time. This process can be described using two time-dependent partial differential equations: an (HJB) equation, which gives us the best consumption path, and a Fokker–Planck equation, which explains how the distribution of wealth and productivity changes. Since our focus is on the long-run behavior of the system, we analyze the time-independent versions of these equations, which represent the stationary equilibrium of the Aiyagari–Bewley–Huggett economy.

Definition 3.14. Let $\{\mathbf{X}(t)\}_{t\geq 0}$ be an n-dimensional Itô process driven by a d-dimensional Brownian motion, and let $p(\underline{x})$ be the probability density function of the random variable $\mathbf{X}(t)$. Then, the equation

$$-\sum_{i=1}^{n} \partial_{x_i}\left[\mu^i(\underline{x})p(\underline{x})\right] + \sum_{i=1}^{n}\sum_{j=1}^{n} \partial_{x_i x_j}\left[p(\underline{x})\frac{1}{2}\sum_{k=1}^{d}\sigma^{i,k}(\underline{x})\sigma^{j,k}(\underline{x})\right] = 0$$

is called a (stationary) Fokker–Planck equation.

Remark 3.15. The Fokker–Planck equation is also called the Kolmogorov forward equation, as it determines the future wealth distribution, given the current state and decisions. On the other hand, equation (HJB) operates backward in time, as it considers how a worker values assets and productivity tomorrow and works backward to determine optimal consumption decisions in the present.

Remark 3.16. Similarly to Remark 3.11, we note that we could consider a time-dependent probability density function $p^*(t,\underline{x})$, for which we could define the time-dependent Fokker–Planck equation as:

$$-\sum_{i=1}^{n} \partial_{x_i}\left[\mu^i(t,\underline{x})p^*(t,\underline{x})\right] + \sum_{i=1}^{n}\sum_{j=1}^{n} \partial_{x_i x_j}\left[p^*(t,\underline{x})\frac{1}{2}\sum_{k=1}^{d}\sigma^{i,k}(t,\underline{x})\sigma^{j,k}(t,\underline{x})\right] = \partial_t p^*(t,\underline{x}).$$

Analogously to the value function defined in Section 3.1, we now define the value function as the best attainable expected lifetime utility, if the agent's initial asset value is a, and the initial productivity is z:

$$v(a,z) = \sup_{c_i\geq 0} \mathbb{E}\left[\int_0^{\infty} \exp\left(-\rho t\right) U\big(c_i(t)\big)\,\mathrm{d}t\right].$$

Note that due to Assumption 3.3, the value function $v(a,z)$ and the optimal consumption policy $c(a,z)$ only depend on the agent's state variables, so it no longer is necessary to keep neither the agent's index nor the time variable.

Definition 3.17. We say that the constants $r, w > 0$ and functions $g(a,z), v(a,z)$ describe a stationary equilibrium if they satisfy conditions (3.8)-(3.9) and the following HJB and Fokker–Planck equations:

$$\rho v(a,z) - \Big(\frac{1}{2}\sigma^2(z)\partial_{zz}v(a,z) + \mu(z)\partial_z v(a,z) + (wz+ra)\partial_a v(a,z) + \\ + \max_{c\geq 0}\{(-\partial_a v(a,z))\cdot c(a,z) + U(c(a,z))\}\Big) = 0, \tag{3.10}$$

$$-\partial_a\big[s(a,z)g(a,z)\big] - \partial_z\left[\mu(z)g(a,z)\right] + \frac{1}{2}\partial_{zz}\left[\sigma^2(z)g(a,z)\right] = 0, \tag{3.11}$$

where $s(a,z) = wz + ra - c^*(a,z)$ represents how much the agent saves given their current state (a,z), and c^* is the optimal control given by $c^* = (U')^{-1}\partial_a v$.

Remark 3.18. The optimal consumption policy $c^*(a,z)$ is obtained by maximizing the term $-\partial_a v \cdot c + U(c)$ with respect to $c \geq 0$. The first-order condition yields $U'(c^*) = \partial_a v(a,z)$, because U is strictly concave and differentiable. Hence, the optimal control is given by

$$c^*(a,z) = (U')^{-1}\big(\partial_a v(a,z)\big).$$

Remark 3.19. Equation (3.10) and more generally equation (HJB) do not always have smooth solutions. Furthermore, it is not guaranteed that there is only one solution to (HJB) either. In such cases, the weak solution that corresponds to a given optimization problem is called the "viscosity solution". For a detailed treatment of the viscosity solutions of the HJB equation, we recommend Chapter 4 of Huyeñ Pham's book [11].

In *Heterogeneous Agent Models in Continuous Time* [2], Achdou et al. give a sufficient condition for the existence of a stationary equilibrium, which we present below.

Proposition 3.20. *If for all $z \in \mathbb{R}$ we have $0 < \sigma(z) < \infty$, there exists a stationary equilibrium, i.e., a solution to equations* (3.8)-(3.11).

Furthermore, Achdou et al. [2] suggest important theorems regarding the characterization of stationary equilibria in their economic model.

Proposition 3.21. *For a stationary equilibrium, if $r < \rho$, the following statements hold:*

(i) There exists a so-called "cutoff productivity level" denoted by $\widehat{z}$, for which

$$s(a_{min}, z) = 0, \quad \text{if } z \leq \widehat{z},$$
$$s(a_{min}, z) > 0, \quad \text{if } z > \widehat{z}.$$

(ii) If an agent i has a constant productivity $z_{const} \leq \widehat{z}$, then $A_i(t) \to a_{min}$ as $t \to \infty$.

(iii) If $z_{const} \leq \widehat{z}$, then $\partial_a\big(wz + ra - c^*(a,z)\big) = r - \partial_a c^*(a,z) \to -\infty$ as $a \to a_{min}$.

Proposition 3.21 has deep implications. It assumes that individuals discount future consumption more heavily than the rate at which their savings can grow through interest, making them consume in the present rather than save for the future. Under this

assumption, individuals with productivity at or below the threshold $\widehat{z}$ have zero savings at the borrowing limit $a_{\min}$. Furthermore, the wealth of agents below the threshold eventually ends up near $a_{\min}$. Even a tiny increase in their wealth near the limit would be met with a very large increase in consumption, quickly driving them back down.

Conversely, those with higher productivity than $\widehat{z}$ will have positive savings even at the borrowing limit, allowing them to accumulate wealth. These dynamics collectively lead to a mass of agents accumulating exactly at the borrowing limit, thus segregating the individuals in the economy based on their productivity.

Summary

In this thesis, we explored a broad range of phenomena within economics and finance. We began by developing a foundational understanding of Brownian motion and its role in modeling uncertainty in economic systems. Building on this, we studied Itô calculus, and highlighted the usefulness of Itô's lemma through concrete examples. We solved the geometric Brownian motion model and illustrated how the Feynman—Kac formula provides a bridge between stochastic differential equations and partial differential equations, most notably in the heat equation and in context of the Black—Scholes option pricing model.

We then turned our attention to knowledge diffusion in macroeconomics. Using the Fisher–KPP equation, we introduced an abstract model of how knowledge spreads across an economy, based on the Lucas–Alvarez framework. We showed the existence of traveling wave solutions and applied a perturbation method to approximate the initial conditions. We implemented both a finite difference scheme and an operator splitting technique to estimate wave speeds and observe the model with different initial distributions.

In the final chapter, we investigated optimal control problems through the theory of controlled stochastic differential equations. We introduced the concepts of a payoff function and value function, and proved the dynamic programming principle. By using Itô's lemma we derived the Hamilton–Jacobi–Bellman equation, then presented the verification theorem. As an application, we examined the Aiyagari–Bewley–Huggett model, which describes individual saving behavior. We also introduced Fokker–Planck equations, which together with the HJB equation, describe the optimal individual behavior and the population distribution in a stationary equilibrium.

Overall, we established deep connections between stochastic processes, partial differential equations and control theory, and showed how these tools can be applied to understand complex economic dynamics. Through both analytical results and numerical experiments, we provided insight into the mathematical structure of decision-making, financial markets, and knowledge diffusion.

All Python codes for this thesis, including the numerical implementations in Sections 1.2 and 2.6 and the phase plane illustration in Section 2.4, can be found on the following GitHub site: *github.com/PappAron/thesis_2025*.

Bibliography

[1] Yves Achdou and Francisco J. Buera and Jean-Michel Lasry and Pierre-Louis Lions and Benjamin Moll: *PDE Models in Macroeconomics*, The Royal Society Publishing, 2014.

[2] Yves Achdou, Jean-Michel Lasry, Pierre-Louis Lions, Benjamin Moll: *Heterogeneous Agent Models in Continuous Time*, Princeton, 2013.

[3] Simon L. Péter: *Differenciálegyenletek és dinamikai rendszerek*, Typotex, 2012.

[4] Ronald Aylmer Fisher: *The wave of advance of advantageous genes*, Annals of Eugenics, 1937.

[5] Andrej Nyikolajevics Kolmogorov, Ivan Petrovsky, Nikolai Piskunov: *Studies of the Diffusion with the Increasing Quantity of the Substance*, Moscow University Bulletin of Mathematics, 1937.

[6] J. David Logan: *An Introduction to Nonlinear Partial Differential Equations (Second Edition)*, Wiley, 2008.

[7] Faragó István, Horváth Róbert: Numerikus Módszerek, Typotex, 2013.

[8] J. Michael Steele: *Stochastic Calculus and Financial Applications*, Springer, 2001.

[9] Anne Carlstein: *Stochastic Calculus: Understanding Brownian Motion and Quadratic Variation*, Cambridge University Press, 2020.

[10] Lawrence C. Evans: *An Introduction to Mathematical Optimal Control Theory*, Springer, 2024 version, University of California, Berkeley.

[11] Huyên Pham: *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer, 1999.

[12] Morton I. Kamien and Nancy L. Schwartz: *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2nd ed., Elsevier Science, 1991.

Statement of AI usage

I, *Áron Balázs Papp*, hereby declare that during the preparation of my thesis, I used the AI-based tools listed below to perform the following tasks:

Task	**Tool Used**	**Used For**	**Note**
LaTeX suggestions	GPT-4o	-	Syntax help for LaTeX
Python assistance	GPT-4o	Chapter 2.3	Python code quality standards, library suggestions
Grammar check	GPT-4o	-	-

Apart from the ones listed above, I did not use any other AI-based tools.