# Dynamic Programming

# and its

# Applications to Economic Theory

Juan Carlos Suárez Serrato

May 18, 2006

# Contents

# Dedication

In the past year, two of my guiding lights faded. From my grandfathers I learned the meaning of hard work, the value of being honest with one's self, and the determination to strive against adversity. They gave me the motivation to fulfill their high expectations and, through their example, the hope to follow in their footsteps.

# Acknowledgements

Throughout my college years I have enjoyed the support and encouragement of many people to whom I owe a great deal of respect and admiration. I would especially like to thank my adviser, Dr. Allen Holder, for the time he devoted to this project, and for being the most inspiring mathematician I know. His challenges have always been stimulating, and his insights invaluable.

I wish to acknowledge other mathematicians and economists who did not influence this project directly, but whose help and kindness shaped my experience the last four years. For his friendship and for nailing my feet to the ground when I needed it done, I thank Dr. Eugenio Suárez. I would like to thank Dr. John Huston, my economics adviser, for providing valuable counsel I repeatedly failed to take. For their perennial support and inspiration, I would like to thank Dr. Roberto Hasfura, Dr. Jorge González, Dr. Barry Hirsch, Dr. Julie DeCourcy, and Dr. Scott Chapman. I want to thank Dr. Sean Knabb for introducing me to this material, and to Dr. Charles Becker for allowing this to happen.

It is hard to overestimate the value of the Mathematics Department at Trinity University; one can not think of a better place to be inspired. The past four years would not have been possible were it not for the students of Mathematics at Trinity, whose friendship and encouragement made our struggles fun.

Finally, commencement nor completion of this project would have been possible without the understanding and support of my friends and family to whom I am enormously indebted.

# Chapter 1

# Introduction

The use of Dynamic Programming in economic modelling has revolutionized economic thought and has allowed the field to tackle interesting problems such as those approached by Edward Prescott and Finn Kydland, the 2004 Nobel Laureates in Economic Science. The award was presented "for their contributions to dynamic macroeconomics: the time consistency of economic policy and the driving forces behind business cycles" (Nobel, 2004). The models they pioneered are now ubiquitous and are an indispensable part of every economist's toolbox. Moreover, such topics are given high priority in most doctoral programs in Economics. It is an apparent contradiction that these models rarely manage to find their way into undergraduate curricula in Economics programs.

The mathematical tools available to undergraduates majoring in Economics is, in most cases, limited by the course offerings of the Mathematics department. In order to include dynamic models in undergraduate Economics programs, some treatment of dynamic programming must be introduced in the course offerings of Mathematics departments. Although the author's main interest is Economics, dynamic programming spans several disciplines in application including Astronomy, Physics, and Engineering. Students of these disciplines would benefit from understanding the applications of these methods to Economics as much as students of other disciplines benefit from examples drawn from Physics and Engineering that are presented in most Calculus courses.

This guide is an introduction to the field of dynamic programming that balances rigor and applications. Although it is not a comprehensive survey of the field, it encompasses a rigorous development of the theory that is sufficient to tackle interesting problems. The attempt is to make this material accessible to advanced undergraduate students who have knowledge of mathematical optimization or real analysis. The author shares the point of view of McShane (1989), who opines "ordinary undergraduate students of mathematics should be taught a form of control theory simple enough to be understood and general enough to be applicable to many problems."

The first objective is to contextualize some of the concepts used in the development of the theory behind dynamic programming in terms familiar to the target audience. These ideas include the use of correspondences as a generalized concept of a function and the characterization of continuity to such correspondences. The second objective is to present a rigorous development of the results that validate the use of these methods. The motivation to present this development is twofold. First, it is desirable to establish the results. Second, the approach followed is attractive in concept and provides some simple yet interesting ways to attack problems that might arise elsewhere. The third and final objective is to understand, by considering an application, the usefulness of the theory developed. Through this example we aim to understand why the development follows the path it does. We hope this guide serves as an introductory overview that will encourage the reader to further pursue the study of dynamic programming.

The remainder of the chapter is organized as follows. Section 1.1 gives a brief introduction to some of the main references associated with dynamic programming, outlines the texts that are central in our development, and suggests some preparatory reading. Section 1.2 presents a brief historical account of the growth of the field. Section 1.3 presents an outline of some applications of the methods of dynamic programming and gives some suggestions as to how the reader might delve into the field after understanding what is developed in this treatment.

## 1.1 Literature Review

Several approaches to the study of dynamic programming exist. These include the Calculus of Variations, which gave birth to Optimal Control Theory, as well as other recursive methods of Dynamic Programming. These methods are all related and coincide in the idea of "thinking in states," as Ljungqvist and Sargent (2004) note. By the concept of states, we mean that a problem of inter-temporal optimization will be affected, or have a different state, in a future period depending on the solution achieved in this period. The recursiveness of states links these problems through time making these methods a natural approach to problems in the fields of Physics and Economics. For our purposes we refer to dynamic programming, control theory, and recursive methods as the same concept giving preference to the term dynamic programming. One of the goals of this project is to contextualize these approaches both historically and topically and find a suitable approach to the material.

In exploring the field of dynamic programming from an applications point of view, the methods presented in most books like Hoy et al. (2001), Ljungqvist and Sargent (2004), and Shone (2002) are "naïve," as McShane (1989) notes, where applications are discussed with no existence theorems being established. Conversely, the books that develop the theory are overly dense and sometimes too involved to be approachable by an undergraduate student. Finally, recent developments in Stochastic Control Theory, such as those presented in Yong and Zhou (1999) and Young (2000), make an introduction to this field even more daunting because they rely on Measure Theory.

As an undergraduate seeking to understand the results that prove the methods of dynamic programming, the author would have liked to have found a treatment that combines both rigor and applications. Given that the field is young, the original developments of the material are still accessible and relevant. Richard Bellman first coined the title of dynamic programming to the study of these methods in his 1957 monograph (Bellman, 1957). It is a consensus between several authors writing in this field that his original piece still provides much insight and is an enjoyable read. Pontryagin et al. (1964) present the other classic work in the field, which was published shortly afterwards and uses an approach that relies on partial differential equations. The approaches used by the two classic treatments above view time as a continuous process and employ several methods such as those of partial differential equations to develop the material. These characteristics made these developments unappealing since most data and applications in Economics view time as a discrete process.

The treatment presented in Stokey et al. (1989) serves as main reference. We follow this development since it has several desirable attributes with respect to applications and development of the material. Stokey et al. (1989) first present the results of deterministic dynamic programming. Several examples of applications are then presented in abstract form. A basic treatment of Measure Theory is then introduced and is sufficient to develop the results of stochastic dynamic programming, followed by some applications of these methods to more realistic examples. This development serves the authors' purpose of having this book be a reference guide for researches in the area. As such, it is inadequate to provide an introduction to undergraduates with an interest in the field, but is a great source of results and insight that will drive this project.

The approach followed in Stokey et al. (1989) is useful since the reader may be familiarized with the idea of thinking in "states" as well as develop an appreciation for the field. In economics, the concept of states is analogue to the an agent's decision-set. At every time period, the agent being modelled chooses one decision from its decision-set. This decision affects the possibilities available to him in the next period. Following this process recursively, we have that the choices available to an agent at any point in time depend on the original state, and the choices made at every preceding time period. By developing the theory of deterministic dynamic programming, Stokey et al. (1989) habituate the reader to thinking in states through the use of correspondences.

In order to contextualize the material in Stokey et al. (1989), we rely on results presented in a typical undergraduate Mathematics program. Suggested prerequisites include the sections on mathematical optimization, metric spaces, and real analysis as presented in Holder (2005), Bryant (1985), Rudin (1976), and Royden (1988) respectively. Finally, other results more particular to economics are drawn from Mas-Colell et al. (1995) and Chiang (1984).

## 1.2 Historical Motivation

As mentioned above, the contributions of Bellman (1957) and Pontryagin et al. (1964) formally described the results that solidified the study of these methods. Although these contribution were made recently, the problems they address have been attacked by some of the first mathematicians in history. The contributions of these mathematicians provided some important results and influenced the work of their modern counterparts, although they did not establish the existence results formally of Pontryagin et al. (1964) or the insight of separating the infinite process into a one-stage control problem as developed by Bellman (1957).

A comprehensive treatment of the history of dynamic programming and its precursors is beyond the scope of this guide. However, a short review of the accomplishments is warranted. Ferguson (2004) traces the "development of the theory of the calculus of variations, from its roots in the work of Greek thinkers and continuing through to the Renaissance." This study commences by setting up some problems studies in antiquity that seem to be direct precedents of the study of the calculus of variations. Some of these include Hero's principle of least time and the isoparametric problems of Pappus. Ferguson (2004) then follows the contributions of great mathematicians such as Fermat, Newton, as well as Bernoulli's postulation of the brachistochrone problem, "the problem concerning the shortest distance between two [accelerating] points," (Fomin and Gelfand, 2000) in an infamous competition. The first real problem of this kind to be formulated and solved was posed by Newton in his "famous work on mechanics methods, *Philosophiae naturalis principia mathematica* (1685) [$\cdots$], thus marking the birth of the theory of the calculus of variations"(Ferguson, 2004). Several other advances would facilitate the development of the current theory of dynamic programming. One example of these is the conception of the first recursive function by Kurt Gödel (Goldstein, 2005).

Ferguson (2004) proceeds to tell the story of the correspondence between Euler and Lagrange and how their interaction fomented the derivation of the Euler equations, which is the goal of the final section in Chapter 3. The last topic in this study is the emanation of optimal control theory from the calculus of variations. McShane (1989) presents a different perspective regarding the historical development of dynamic programming. This study is centered on the history of optimal control theory, and has a less forgiving perspective.

McShane (1989) recognizes, as does Ferguson (2004), the lack of philosophical rigor in the calculus of variations and criticizes it more pungently. When commenting on the treatment of the calculus of variations in Fomin and Gelfand (2000) he notes that, "if the calculus of variations is mathematics, our conclusions must be deducible logically from the hypothesis, with no use of anything that is 'clear from the physical meaning' "(McShane, 1989). In the same way McShane (1989) chastises some of the early developments in the field of calculus of variations, he argues later developments were mislead. He states:

> The whole subject was introverted. We who were working in it were striving to advance the theory of the calculus of variations as an end in itself, without attention to its relation with other fields of activity (McShane, 1989).

In the perspective of McShane (1989), the contributions of Pontryagin et al. (1964) changed the focus of the calculus of variations by focusing on problems in Engineering and Economics. "In the process, they incidentally introduced new and important ideas into the calculus of variations" (McShane, 1989).

Earlier contributions by Bellman (1957) and others made these methods ripe for the application in Economics. As Edward Prescott noted in his address when receiving the Nobel Prize, it took several years before he and others would reinvent the existing fields of optimal control theory and dynamic optimization with an Economics flavor (Nobel, 2004).

## 1.3 Applications and Extensions

It is desirable to demonstrate the power of these methods by considering applications. Given the nature of this project, we examine in detail one problem, that of deterministic economic growth. The approach taken in the final chapter of this guide is fairly abstract, with the goal of this chapter being to elucidate some intuition into why the assumptions made to obtain the desired results are consistent with economic theory. However, there are numerous applications that are more concrete and that might help the reader establish a better understanding of the material.

Ljungqvist and Sargent (2004) present a plethora of applications in economics that might quench the reader's thirst for a less rigorous development. Fischer and Blanchard (1989) present a thorough treatment of modern macroeconomics, in which many of the models presented rely on recursive methods to provide insight. Similarly, Sala-i-Martin (2000) presents notes from graduate courses in Economic Growth that assumes that students have the ability to solve problems of recursive nature. Barro and Sala-i-Martin (2003) reach a high point in the study of topics related to Economic Growth and present several models that assume the student is familiar with the concepts of dynamic programming. Aghion and Griffith (2005) extend certain growth topics to include the idea of endogenous growth and technological development.

Other applications outside economics were considered for this guide. The most entertaining applies the Bellman equation to the study of optimal policies in sports. Romer (2002) applies the Bellman equation and finds an optimal policy for conversion plays in football. Several other applications in the fields of Engineering and Physics exist but were not considered.

The nature of this project makes a comprehensive treatment of the field of dynamic programming almost impossible. In this treatment, several interesting extensions of the material will not be included but are worth mentioning. First, the use of computational methods to find numerical solutions to stochastic optimal control problems in economics has become ubiquitous in recent years. Diáz-Giménez (2001) presents the method of linear quadratic approximations to evaluate dynamic programs with macroeconomic data. Another topic of interest is the stability properties of policy functions derived using the methods presented in this guide. Stokey et al. (1989) and Vohra (2005) present treatments of the study of dynamic stability and interesting applications of these results.

The following Chapters are organized as follows. Chapter 2 introduces the mathematical background necessary for the development of the theory of dynamic programming in Chapter 3, and it contextualizes it within other concepts advanced undergraduates should be familiar with. Chapter 3 develops the theory of deterministic dynamic programming following the development to the derivation of the Euler equations. Finally, Chapter 4 presents some insight into the necessity of the assumptions established in Chapter 3 through the application of dynamic programming to the analysis of deterministic economics growth in an abstract form.

# Chapter 2

# Continuities

In this Chapter we consider some material that is preliminary to the results that follow in Chapter 3, and that is requisite to the development of dynamic programming. The main focus of this Chapter is the study of several definitions of continuities for functions followed by the characterization of continuity for point-to-set mappings, or what we call correspondences. Much of the development that follows rests on a firm understanding of the concepts of upper and lower-hemi continuities.

We begin section 2.1 by considering a general definition of functional continuity. We then relate this classic definition of continuity to the stronger property of uniform continuity and upper and lower semi continuities for functions. Section 2.2 introduces the concept of correspondences and relates several versions of continuity of correspondences to the analog properties of functions. The section concludes with two theorems that are useful in the proceeding section.

## 2.1 Functional Continuity

The following is a classical definition of functional continuity and is typically introduced in a Real Analysis course.

**Definition 1.** *(Continuous Functions.) Let* $(\mathbb{X}, d_{\mathbb{X}})$ *and* $(\mathbb{Y}, d_{\mathbb{Y}})$ *be metric spaces; suppose* $E \subset \mathbb{X}$*, and* $f$ *maps* $E$ *into* $\mathbb{Y}$ *. Then* $f$ *is said to be continuous at* $p$ *if for every* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that*

$$d_{\mathbb{X}}(x, p) < \delta \text{ implies } d_{\mathbb{Y}}(f(x), f(p)) < \varepsilon.$$

As presented in Rudin (1976), Royden (1988), and other sources, continuity is more generally defined as a topological concept. For our purposes there is little gain from such a general development, and we restrict the generality of our development. A property that is stronger than continuity is that of uniform continuity.

**Definition 2.** *(Uniform Continuity.) Let $f$ be a mapping of a metric space $(\mathbb{X}, d_{\mathbb{X}})$ into a metric space $(\mathbb{Y}, d_{\mathbb{Y}})$. We say that $f$ is uniformly continuous on $\mathbb{X}$ if for every $\varepsilon > 0$there exists $\delta > 0$ such that*

$$d_{\mathbb{X}}(q,p) < \delta \text{ implies } d_{\mathbb{Y}}(f(q), f(p)) < \varepsilon.$$

Two differences exists between Definitions 1 and 2. First, the $\delta$ used in Definition 1 depend on the point $p$. Secondly, the $\delta$ used in Definition 2 can be used for every pair of points such that $d_{\mathbb{X}}(p,q) < \delta$. We say that uniform continuity is a stronger property since every function with this property also has the property of functional continuity. Theorem 1 below relates the concepts of functional continuity with uniform continuity through the requirement that the space $\mathbb{X}$ be compact.

**Theorem 1.** *Let $f$ be a continuous mapping of a compact metric space $(\mathbb{X}, d_{\mathbb{X}})$ into a metric space $(\mathbb{Y}, d_{\mathbb{Y}})$. Then $f$ is uniformly continuous in $\mathbb{X}$ .*

*Proof.* Let $\varepsilon > 0$. Then $\forall\ x \in E, \exists\ \delta_x > 0 \ni f(N_{\delta_x/2}(x)) \subseteq N_{\varepsilon/2}(f(x))$. Since $\{N_{\delta_x/2}(x) : x \in E\}$ is an open cover of $E$, and since $A$ is compact, there exists a finite subcover, say $\{N_{\delta_{x_i}/2}(x_i) : i = 1, 2, \cdots, n\}$. Define $\delta = \min\{\delta_{x_i}/2 : i = 1, 2, \cdots, n\}$. Let $p, q \in E$ be $\ni d(p,q) < \delta$. Then $p \in N_\delta(x_i)$ for some $1 \leq j \leq n$. Also,

$$\begin{aligned} d(q, x_j) &\leq d(p,q) + d(p, x_j) \\ &< \delta + \delta_{x_j}/2 \\ &< \delta_{x_j}. \end{aligned}$$

So,

$$d(f(p), f(q)) \leq d(f(p), f(x_j)) + d(f(x_j), f(q)) < \varepsilon.$$

$\boxdot$

We proceed by introducing the concept of Lipschitz Continuity. This continuity essentially states that the difference in function values is bounded proportionally to the distance between the arguments.

**Definition 3.** *(Lipschitz Continuity.) Let $f$ be a mapping of a metric space $(\mathbb{X}, d_{\mathbb{X}})$ into a metric space $(\mathbb{Y}, d_{\mathbb{Y}})$. We say that $f$ is Lipschitz continuous on $\mathbb{X}$ if $\exists\ \lambda > 0$ such that for every $x, y \in \mathbb{X}$,*

$$d_{\mathbb{Y}}(f(x), f(y)) < \lambda d_{\mathbb{X}}(x,y).$$

Theorem 2 relates the concepts introduced in Definition 3 to Definitions 1 and 2. Namely, it states that a function that is Lipschitz continuous is uniformly continuous and thus continuous.

**Theorem 2.** *Let f be a mapping of a metric space* $(\mathbb{X}, d_{\mathbb{X}})$ *into a metric space* $\mathbb{Y}$ *with the property of Lipschitz continuity. Then f is uniformly continuous.*

*Proof.* Let $f$ be as stated and define $\lambda > 0$ to be $\ni$

$$d_{\mathbb{Y}}(f(x), f(y)) < \lambda d_{\mathbb{X}}(x, y).$$

Let $\varepsilon > 0$ and define $\delta = \frac{\varepsilon}{\lambda}$. It follows that for $q \in N_{\delta}(p)$,

$$\begin{aligned} d_{\mathbb{Y}}(f(x), f(y)) &< \lambda d_{\mathbb{X}}(x, y) \\ &< \lambda\delta = \varepsilon. \end{aligned}$$

Since $\delta$ was chosen independently of $p$, we have the desired result. $\boxdot$

The following definitions of continuities are not directly related to Lipschitz continuity. Instead, we view these as relaxed instances of Definition 1. Moreover, by the definition of the lim sup and lim inf, these have the advantage of always existing in the extended real numbers.

**Definition 4.** *(Functional Upper Semi-Continuity.) The function f is said to be upper semi-continuous at* $x_0$ *if*

$$\limsup_{x \to x_0} f(x) \leq f(x_0).$$

**Definition 5.** *(Functional Lower Semi-Continuity.) The function f is said to be lower semi-continuous at* $x_0$ *if*

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

The claim above that the functional definitions of upper and lower semi-continuities are weaker than Definition 1 can be substantiated by verifying that both limitsinequalities in the definitions above hold when a function has the property of being continuous. The next theorem provides a way to appreciate this relationship.

**Theorem 3.** *Let f be a mapping of a metric space* $(\mathbb{X}, d_{\mathbb{X}})$ *into a metric space* $(\mathbb{Y}, d_{\mathbb{Y}})$ *be such that it is both upper and lower semi-continuous. Then, f is continuous.*

*Proof.* Let $x_0 \in \mathbb{X}$, and $f$ be as stated. From the definitions above, it follows that

$$\lim_{x \to x_0} \inf f(x) \geq f(x_0) \geq \lim_{x \to x_0} \sup f(x).$$

From the definitions of the lim inf and lim sup we always have the opposite relationship. We thus have that $\lim_{x \to x_0} \inf f(x) = \lim_{x \to x_0} \sup f(x)$, and the function $f$ is continuous at $x_0$. ⊡

## 2.2 Continuities of Correspondences

Our focus now shifts to the study of correspondences, or what is commonly known as point-to-set maps. In the same way we approached functional continuities, we study several definitions of continuities for correspondences. We further provide insight as to how these properties relate to those introduced in the previous section. We first provide a formal definition of correspondences.

**Definition 6.** *(Correspondence) A correspondence* $\Gamma : \mathbb{X} \to \mathbb{Y}$ *is a relation that assigns a set* $\Gamma(x) \subseteq \mathbb{Y}$ *to each* $x \in \mathbb{X}$.

Correspondences prove to be valuable in the study of dynamic programming since they provide a natural way to relate the state in the present period to the state in the future period. To acclimate the reader to the ideas of lower and upper hemi-continuity of a correspondence, we first introduce the concepts of lower and upper semi-continuity.

**Definition 7.** *(Upper Semi-Continuity for Correspondences) The correspondence* $\Gamma$ *is upper semi-continuous at* $x_0$ *if* $\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni$

$$x \in N_\delta(x) \text{ implies } \Gamma(x) \subseteq \bigcup_{y \in \Gamma(x_0)} N_\varepsilon(y).$$

The word upper in Definition 7 makes sense since the $\varepsilon$ neighborhood of the target set is an upper-approximation of the local images.

**Definition 8.** *(Lower Semi-Continuity for Correspondences) The correspondence* $\Gamma$ *is lower semi-continuous if* $\forall\, \varepsilon > 0, \exists\, \delta > 0 \ni$

$$x \in N_\delta(x) \text{ implies } \Gamma(x_0) \subseteq \bigcup_{y \in \Gamma(x)} N_\varepsilon(y).$$

Conversely, Definition 8 says that the target set is within an $\varepsilon$ neighborhood of the local images. The following theorem relates upper and lower semi-continuity of a function and correspondences is found in Holder (2005). The proof of this theorem is omitted for brevity.

**Theorem 4.** *The function $f : \mathbb{X} \to \mathbb{R}$ is upper or lower semi-continuous if and only if the correspondence $\Gamma : \mathbb{X} \to \mathbb{R}$ defined by $\Gamma(x) = \{y : y \leq f(x)\}$ is upper or lower semi-continuous, respectively.*

Next, we introduce the concept of a graph of a correspondence. The graph of a correspondence is often used to establish related properties of the correspondence.

**Definition 9.** *(Graph) The graph of a correspondence $\Gamma$ is the set*

$$A = \{(x, y) \in \mathbb{X} \times \mathbb{Y} : y \in \Gamma(x)\}.$$

Having introduced several continuities, we now introduce the concept of upper hemi-continuity. We establish the desired definition only for compact-valued correspondences. By compact valued we mean that for any $x \in \mathbb{X}$, the set $\Gamma(x)$ is compact. As Stokey et al. (1989) note, "a general definition of u.h.c. for all correspondences is available, but $[\cdots]$ its wider scope is never useful" for our purposes.

**Definition 10.** *(Upper Hemi-Continuity.) A compact-valued correspondence $\Gamma$ is* **upper hemi-continuous** *(u.h.c.) at $x$ if $\Gamma(x)$ is nonempty and if, for every sequence $x_n \to x$ and every sequence $\{y_n\} \ni y_n \in \Gamma(x_n)\ \forall\ n, \exists$ a convergent subsequence of $\{y_n\}$ whose limit point $y$ is in $\Gamma(x)$.*

The following theorem establishes a relationship between the concept of u.h.c. and functional continuity. We attribute the following theorem to Mas-Colell et al. (1995), who note that u.h.c. " can be thought of as a natural generalization of the notion of continuity for functions."

**Theorem 5.** *A correspondence $\Gamma$ is single-valued and u.h.c. if and only if it is also continuous in the functional sense.*

*Proof.* If $\Gamma$ is continuous as a function, we have that its graph is closed. Moreover, since the continuous mappings of compact sets is compact, and therefore bounded as well. Thus, $\Gamma$ is upper hemi-continuous as a correspondence.

Suppose now that $\Gamma$ is upper hemi-continuous as a correspondence and consider any sequence $x_n \to x \in A$ with $x_n \in A\ \forall\ n$. Define $S = \{x_n : n = 1, 2, \cdots\} \cup \{x\}$. Now, let $\varepsilon > 0$ and set $N \in \mathbb{N}$ to be such that $d_{\mathbb{X}}(x, x_n) < \varepsilon\ \forall\ n \geq N$. Define $r$ to be such that

$$r > \max\{||x_1||, ||x_2||, \cdots, ||x_N||, ||x|| + \varepsilon\}.$$

We have that $||x|| < r$, $\forall\ x \in S$, so that $S$ is bounded. Because $S$ is also closed, it follows that $S$ is compact, by the Heine-Borel theorem. By hypothesis, $\Gamma(S)$ is bounded, and so $\overline{\Gamma(S)}$ is compact. Assume for the sake of obtaining a contradiction that the sequence $\{\Gamma(x_n)\} \subset \overline{\Gamma(S)}$ does not converge to $\Gamma(x)$. Extract a subsequence $x_{n_k} \rightarrow x$ such that $\Gamma(x_{n_k}) \rightarrow y$ for some $y \in \overline{\Gamma(S)} \ni y \neq \Gamma(x)$. Then the graph of $\Gamma$ is not closed, which contradicts the property of u.h.c. as a correspondence. $\boxdot$

The use of single valued correspondences would destroy the motivation to introduce the concept of correspondences and in general we do no consider them as single valued. Nonetheless, the theorem above helps achieve the concept of u.h.c. The second type of continuity, lower hemi-continuity for correspondences, presents another generalization of the concept of functional continuity for correspondences.

**Definition 11.** *(Lower Hemi-Continuity.) A correspondence* $\Gamma$ *is* **lower hemi-continuous** *(l.h.c.) at* $x$ *if* $\Gamma(x)$ *is nonempty and if, for every* $y \in \Gamma(x)$ *and every* $x_n \rightarrow x$, $\exists\ N \geq 1$ *and a sequence* $\{y_n\}_{n=N}^{\infty}$ *such that* $y_n \rightarrow y$ *and* $y_n \in \Gamma(x)$, $\forall\ n \geq N$.

The theorem below establishes a result parallel to that relating u.h.c. to functional continuity presented in Theorem 5. For the purpose of brevity, a proof is omitted.

**Theorem 6.** *If a correspondence* $\Gamma$ *is single valued and l.h.c. is also continuous in the functional sense.*

The study of l.h.c. and u.h.c. on their own is required to address some of the topics covered in the following section. Figure 2.1 shows in panel (a) a lower hemi-correspondence that is not upper hemi-continuous, and in panel (b) an upper hemi-continuous correspondence that is not lower hemo-continuous. The images provide intuition about the corresponding definitions. A general concept of continuity of a correspondence is now introduced.

**Definition 12.** *(Continuity of Correspondences) A correspondence* $\Gamma :\ \mathbb{X} \rightarrow \mathbb{Y}$ *is continuous at* $x \in \mathbb{X}$ *if it is both l.h.c. and u.h.c.*

The proof of the following theorem uses the fact that if the graph of the correspondence is closed, then the graph is close-valued. That is, for every $x \in \mathbb{X}$, the set $\Gamma(x)$ is a closed set.

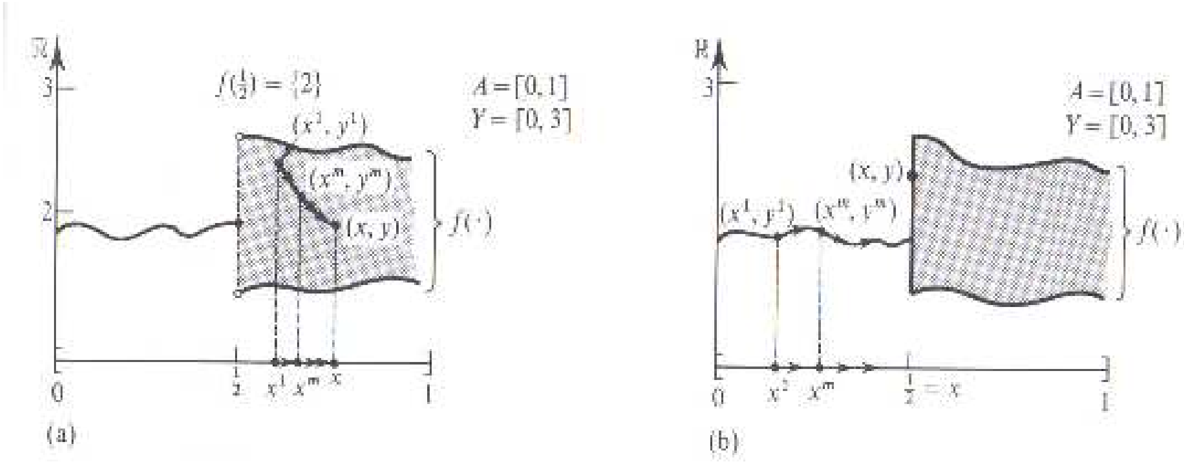


Figure 2.1: Upper and lower hemi-continuity Mas-Colell et al. (1995).

**Theorem 7.** *Let $\Gamma$ be a nonempty-valued correspondence, and let $A$ be the graph of $\Gamma$. Suppose that $A$ is closed, and that for a bounded set $\hat{\mathbb{X}} \subseteq \mathbb{X}$, the set $\Gamma(\hat{\mathbb{X}})$ is bounded. Then $\Gamma$ is compact-valued and u.h.c.*

*Proof.* For any $x \in \mathbb{X}$ we have that since $A$ is closed, it follows that $\Gamma(x)$ is closed as well. Moreover, since $\Gamma$ is compact-valued by hypothesis, the Heine-Borel theorem yields compactness when restricting $\mathbb{X}$ to subsets of $\mathbb{R}^k$.

Let $\hat{x} \in \mathbb{X}$, and let $\{x_n\} \subset \mathbb{X}$ be such that $x_n \to \hat{x}$. Since $\Gamma(x_n) \neq \emptyset$, choose $y_n \in \Gamma(x_n)\ \forall\, n$. Since $x_n \to x$, there is a bounded set $\hat{\mathbb{X}} \subseteq \mathbb{X}$ containing $\{x_n\}$ and $\hat{x}$. By hypothesis, $\Gamma(\hat{\mathbb{X}})$ is bounded. Hence, $\{y_n\} \subseteq \Gamma(\hat{\mathbb{X}})$ has a convergent subsequence, call it $\{y_n\}$; let $\hat{y}$ be the limits point of the subsequence. Then, $\{(x_n, y_n)\}$ is a convergent subsequence in $A$ converging to $(\hat{x}, \hat{y})$; since $A$ is closed, it follows that $(\hat{x}, \hat{y}) \in A$. Hence $\hat{y} \in \Gamma(\hat{x})$, so $\Gamma$ is u.h.c. at $\hat{x}$. Since $\hat{x}$ was chosen arbitrarily, we obtain the result. $\boxdot$

The last theorem in this section establishes another relationship between a correspondence and its graph. In this case, the convexity of the graph $A$ of the correspondence $\Gamma$ is used to show $\Gamma$ is u.h.c. under a set of circumstances.

**Theorem 8.** *Let $\Gamma$ be a nonempty-valued correspondence, and let $A$ be the graph of $\Gamma$. Suppose that $A$ is convex, and that for a bounded set $\hat{\mathbb{X}} \subseteq \mathbb{X}$, there is a bounded set $\hat{\mathbb{Y}} \subseteq \mathbb{Y} \ni \Gamma(x) \cap \hat{\mathbb{Y}} \neq \emptyset$, $\forall\, x \in \hat{\mathbb{X}}$. Then $\Gamma$ is l.h.c. at every interior point of $\mathbb{X}$ .*

*Proof.* Choose $\hat{x} \in \mathbb{X}\backslash\mathbb{X}'$; $\hat{y} \in \Gamma(\hat{x})$; and $\{x_n\} \subset \mathbb{X}$ with $x_n \to x$. Let $\varepsilon > 0$ be such that $\hat{X} = N_\varepsilon(\hat{x})$. Note that for some $N \geq 1, x_n \in \hat{\mathbb{X}}$, $\forall\, n$; without loss of generality we take $N = 1$.

Let $D$ denote the boundary of the set $\hat{\mathbb{X}}$. Every point $x_n$ has at least one representation as a convex combination of $\hat{x}$ and a point in $D$. For each $n$, choose $\alpha_n \in [0,1]$ and $d_n \in D$ such that

$$x_n = \alpha_n d_n + (1-\alpha_n)\hat{x}.$$

$D$ is a bounded set and $x_n \to x$, so $\alpha_n \to 0$. Choose $\hat{Y}$ such that $\Gamma(x) \cap \hat{Y} \neq \emptyset, \forall\, x \in \hat{\mathbb{X}}$. Then for each $n$, choose $\hat{y}_n \in \Gamma(d_n) \cap \hat{\mathbb{Y}}$, and define

$$y_n = \alpha_n \hat{y}_n + (1-\alpha_n)\hat{y}.$$

Since $(d_n, \hat{y}_n) \in A,\ \forall\, n, (\hat{x}, \hat{y}) \in A$, and $A$ is convex, it follows that $(x_n, y_n) \in A,\ \forall\, n$. Moreover, since $\alpha_n \to 0$ and all of the $\hat{y}_n$'s lie in the bounded set $\hat{\mathbb{Y}}$, it follows that $y_n \to \hat{y}$. Hence $\{(x_n, y_n)\}$ lies in $A$ and converges to $(\hat{x}, \hat{y})$. $\boxdot$

This chapter provided the reader with the concepts of upper and lower hemi-continuities for correspondences as well presenting some results that will prove valuable in the next section. We have studied correspondences and different types of continuities. The following section uses the concepts introduced so far to establish the foundations of dynamic programming.

# Chapter 3

# Dynamic Programming

In this Chapter, we establish the foundations of dynamic programming. In Section 3.1 we first consider topics of analytical nature like that of a contraction mapping and a fixed point. These topics, as well as those developed in prior sections, lead to the establishment of the Theorem of the Maximum in section 3.2. In Section 3.3 we define the sequence problem and the functional equation problem. We then establish the congruency of their solutions under the assumption that the return function $F$ is bounded. Finally, Section 3.4 explores the classical variational approach to dynamic problems and established the sufficiency of the Euler equations together with the transversality condition.

## 3.1 Contraction Mappings and Fixed Points

This section is devoted to the study of fixed points. Although fixed points can be guaranteed for many function types, we focus on those obtained from a contraction mapping. The formal definition of a fixed point and contraction follow.

**Definition 13.** *(Fixed Point) A function mapping $T : \mathbb{X} \to \mathbb{X}$ has a fixed point if for some $\bar{x} \in \mathbb{X}$, $T(\bar{x}) = \bar{x}$. Then, $\bar{x}$ is called a fixed point of $T$.*

**Definition 14.** *(Contraction) Let $(\mathbb{X}, d_{\mathbb{X}})$ be a metric space and $T : \mathbb{X} \to \mathbb{X}$ be a function mapping of $\mathbb{X}$ into itself. $T$ is a* **contraction mapping** *(with* modulus $\beta$*) if for some $\beta \in (0, 1)$,*

$$d_{\mathbb{X}}(T(x), T(y)) \leq \beta d_{\mathbb{X}}(x, y), \ \forall \ x, y \in \mathbb{X}.$$

The following theorem relates the two concepts introduced above by establishing the fact that contraction mappings always have fixed points. The theorem below examines complete spaces. That is, spaces in which every Cauchy sequence is a convergent sequence.

**Theorem 9.** *(Contraction Mapping Theorem) If $(\mathbb{X}, d_{\mathbb{X}})$ is a complete metric space and $T : \mathbb{X} \to \mathbb{X}$ is a contraction with modulus $\beta$, then*

1. *$T$ has exactly one fixed point $\hat{x} \in \mathbb{X}$, and*
2. *for any $x_0 \in \mathbb{X}, d_{\mathbb{X}}(T^n(x_0), x) \leq \beta^n d_{\mathbb{X}}(x_0, x), n = 1, 2, \cdots,$*

*where we define the iterates of $T$, the mappings $\{T^n\}$ by $T^0(x) = x$, and $T^n = T(T^{n-1}(x)), n = 1, 2, \cdots, n.$*

*Proof.* 1. We first prove the existence of a fixed point. Choose $v_0 \in \mathbb{X}$, and define $\{v_n\}_{n=0}^{\infty}$ by $v_{n+1} = T(v_n)$, so that $v_n = T^n(v)$. By the contraction mapping property of $T$,

$$d_{\mathbb{X}}(v_2, v_1) = d_{\mathbb{X}}(T(v_1), T(v_0)) \leq \beta d_{\mathbb{X}}(v_1, v_0).$$

By induction and the triangle inequality, we have that for any $m > n$

$$\begin{aligned} d_{\mathbb{X}}(v_m, v_n) &\leq d_{\mathbb{X}}(v_m, v_{m-1}) + \cdots + d_{\mathbb{X}}(v_{n+1}, v_n) \\ &\leq [\beta^{m-1} + \cdots + \beta^n] d_{\mathbb{X}}(v_1, v_0) \\ &= \beta^n [\beta^{m-n-1} + \cdots + \beta + 1] d_{\mathbb{X}}(v_1, v_0) \\ &\leq \frac{\beta^n}{1-\beta} d_{\mathbb{X}}(v_1, v_0), \end{aligned}$$

where the fourth inequality holds from

$$\beta^{m-n-1} + \cdots + \beta + 1 = \sum_{n=1}^{m-n-1} \beta^n \leq \sum_{n=1}^{\infty} \beta^n = \frac{1}{1-\beta}.$$

From the equation above, we gather $\{v_n\}_{n=0}^{\infty}$ is a Cauchy sequence. Since $\mathbb{X}$ is a complete metric space, it follows that $v_n \to v \in \mathbb{X}$.

We now show uniqueness of the fixed point by showing $T(v) = v$. Note first that $\forall\, n, v_0 \in \mathbb{X}$,

$$\begin{aligned} d_{\mathbb{X}}(T(v), v) &\leq d_{\mathbb{X}}(T(v), T^n(v_0)) + d_{\mathbb{X}}(T^n(v_0), v) \\ &\leq \beta d_{\mathbb{X}}(v, T^{n-1}(v_0)) + d_{\mathbb{X}}(T^n(v_0), v). \end{aligned}$$

Since both terms above converge to zero, we have that $d_{\mathbb{X}}(Tv, v) \to 0$ as $n \to \infty$.

Finally, assume for the sake of obtaining a contradiction that $\hat{v} \neq v$ is such that $T(\hat{v}) = \hat{v}$. Then,

$$0 < a = d_{\mathbb{X}}(v, \hat{v}) = d_{\mathbb{X}}(T(\hat{v}), T(v)) \leq \beta d_{\mathbb{X}}(\hat{v}, v) = \beta a,$$

which cannot hold since $a > 0, \beta < 1$. Thus, there is not another $v$ such that $T(v) = v$ and we have the desired uniqueness property.

2. Observe that for $n \geq 1$,

$$d_{\mathbb{X}}(T^n(v_0), v) = d_{\mathbb{X}}[T(T^{n-1}(v_0)), T_v] \leq \beta d_{\mathbb{X}}(T^{n-1}(v_0), v),$$

so that the result follows by induction.

$\boxdot$

For our purposes, contraction mappings are restricted to metric spaces with desirable properties such as completeness. The following corollaries establish results about contraction mappings defined over complete metric spaces.

**Corollary 1.** *Let* $(\mathbb{X}, d_{\mathbb{X}})$ *be a complete metric space and* $T : \mathbb{X} \to \mathbb{X}$ *be a contraction mapping with fixed point* $x \in \mathbb{X}$*. If* $\mathbb{X}'$ *is a closed subset of* $\mathbb{X}$ *and* $T(\mathbb{X}') \subseteq \mathbb{X}'$*, then* $x \in \mathbb{X}'$*. If in addition* $T(\mathbb{X}') \subseteq \mathbb{X}'' \subseteq \mathbb{X}'$*, then* $x \in \mathbb{X}''$*.*

*Proof.* Choose $v_0 \in \mathbb{X}'$ and note that $\{T^n(v_0)\}$ is a sequence in $\mathbb{X}'$ converging to $v$. Since $\mathbb{X}'$ is a closed subset of a compact space $\mathbb{X}$ it is itself closed. Thus, we have that $v \in \mathbb{X}'$. If in addition we have that $T(\mathbb{X}') \subseteq \mathbb{X}''$, then it follows that $v = Tv \in \mathbb{X}''$. $\boxdot$

**Corollary 2.** *Let* $(\mathbb{X}, d_{\mathbb{X}})$ *be a complete metric space and* $T : \mathbb{X} \to \mathbb{X}$*, and suppose that for some* $N \in \mathbb{N}$*,* $T^N : \mathbb{X} \to \mathbb{X}$ *is a contraction mapping with modulus* $\beta$*. Then*

1. *$T$ has exactly one fixed point in $\mathbb{X}$ , and*
2. *for any $x_0 \in \mathbb{X}, d_{\mathbb{X}}(T^{kN}(x_0), x) \leq \beta^k d_{\mathbb{X}}(x_0, x), k = 1, 2, \cdots$.*

*Proof.* We show that a unique fixed point $v$ of $T^N$ is also the unique fixed point of $T$. We have that

$$d_{\mathbb{X}}(T(v), v) = d_{\mathbb{X}}[T(T^N(v)), T^N(v)] = d_{\mathbb{X}}[T^N(T(v)), T^N(v)] \leq \beta d_{\mathbb{X}}(T(v), v).$$

Since $\beta \in (0, 1)$, this implies that $d(T(v), v) = 0$ and $v$ is a fixed point of $T$. Uniqueness follows since any fixed point of $T$ is also a fixed point of $T^N$. The proof for part 2 is the same as that of Theorem 9 and is thus omitted. $\boxdot$

The theorem below establishes the main result of this section. We understand contraction mappings as functions that have desirable properties, especially regarding fixed points. The following theorem establishes the conditions that are sufficient for a function mapping to be a contraction.

**Theorem 10.** *(Blackwell's sufficient conditions for a contraction) Let $\mathbb{X} \subseteq \mathrm{I\!R}^k$, and let $B(\mathbb{X})$ be the space of bounded functions $f : \mathbb{X} \to \mathrm{I\!R}$, with respect to the* sup *norm. Let $T : B(\mathbb{X}) \to B(\mathbb{X})$ be an operator satisfying*

1. *(monotonicity) If $f, g \in B(\mathbb{X})$ and $f(x) \leq g(x)$, $\forall\ x \in \mathbb{X}$, then $(Tf)(x) \leq (Tg)(x)$, $\forall x \in \mathbb{X}$.*
2. *(discounting) there exists some $\beta \in (0,1)$ such that*

$$T(f(x) + a) \leq (Tf)(x) + \beta a,\ \forall f \in B(\mathbb{X}), a \geq 0, x \in \mathbb{X}.$$

*Proof.* If $f(x) \leq g(x)$, $\forall\ x \in \mathbb{X}$, we write $f \leq g$. For any $f, g \in B(\mathbb{X}), f \leq g + ||f - g||$. Then properties 1 and 2 imply

$$Tf \leq T(g + ||f-g||) \leq Tg + \beta||f - g||.$$

Reversing the roles of $f, g$ gives the same logic. So $Tg \leq Tf + \beta||f - g||$. Combining the two inequalities we have that, $||Tg - Tf|| \leq \beta||f - g||$. $\boxdot$

## 3.2 Theorem of the Maximum

In this section, we apply the concepts of hemi-continuity of correspondences and establish the result that allows the study dynamic programming. The following theorem was first proposed by Pontryagin et al. (1964) and establishes that, under a set of conditions, dynamic problems are well-defined.

**Theorem 11.** *(Theorem of the Maximum) Let $\mathbb{X} \subset \mathrm{I\!R}^k$ and $\mathbb{Y} \subset \mathrm{I\!R}^l$, let $f : \mathbb{X} \times \mathbb{Y} \to \mathrm{I\!R}$ be a continuous function, and let $\Gamma : \mathbb{X} \to \mathbb{Y}$ be a compact-valued and continuous correspondence. Then, the function $h : \mathbb{X} \to \mathrm{I\!R}$ defined as*

$$h(x) = \max_{y \in \Gamma(x)} f(x, y)$$

*is continuous, and the correspondence $G : \mathbb{X} \to \mathbb{Y}$ defined as*

$$G(x) = \{y \in \Gamma(x) : f(x, y) = h(x)\}$$

*is nonempty, compact-valued, and u.h.c.*

*Proof.* Let $x \in \mathbb{X}$. The set $\Gamma(x)$ is nonempty and compact, and $f(x, \cdot)$ is continuous; hence it attains its maximum by Wierstrauss's Theorem and the set of maximizers $G(x)$ is nonempty. Moreover, since $G(x) \subseteq \Gamma(x)$ and by the compactness of $\Gamma(x)$ we have that $G(x)$ is bounded. Suppose $y_n \to y$, and $y_n \in G(x)$, $\forall\ n$. Since $\Gamma(x)$ is closed, $y \in \Gamma(x)$. Also, since $h(x) = f(x, y_n); \forall\ n$, and $f$ is continuous, it follows that $f(x, y) = h(x)$. Hence $y \in G(x)$; so $G(x)$ is closed. Thus, $G(x)$ is nonempty and compact for each $x$.

Next, we show that $G(x)$ is u.h.c. Fix $x$, and let $\{x_n\}$ be any sequence converging to $x$. Choose $y_n \in G(x_n)$, $\forall\ n$. Since $\Gamma$ is u.h.c. , there exists a subsequence $\{y_n\}$ converging to $y \in \Gamma(x)$. Let $z \in \Gamma(x)$. Since $\Gamma$ is u.h.c., there exists a subsequence $\{y_{n_k}\}$ converging to $y \in \Gamma(x)$. Hence $G$ is u.h.c.

Finally, we show that $h$ is continuous. Fix $x$, and let $\{x_n\}$ be any sequence converging to $x$. Choose $y_n \in \Gamma(x_n)$, $\forall\ n$. Let $\bar{h} = \limsup h(x_n)$ and $\underline{h} = \liminf h(x_n)$. Then there exists a subsequence $\{x_{n_k}\}$ such that $\bar{h} = \lim f(x_{n_k}, y_{n_k})$. But since $G$ is u.h.c., there exists a subsequence of $\{y_{n_k}\}$, call it $\{y'_j\}$, converging to $y \in G(x)$. Hence $\bar{h} = \lim f(x_j, y'_j) = f(x, y) = h(x)$. An analogous result establishes that $h(x) = \underline{h}$. Hence $\{h(x_n)\}$ converges to its limit of $h(x)$. ⊡

The following two results study how the results of the Theorem of the Maximum change when stronger constraints are imposed on $f$ and $\Gamma$.

**Lemma 1.** *Let $\mathbb{X} \subseteq \mathbb{R}^k$ and $\mathbb{Y} \subseteq \mathbb{R}^l$. Assume that the correspondence $\Gamma : \mathbb{X} \to \mathbb{Y}$ is nonempty, compact, convex valued, and continuous, and let $A$ be the graph $\Gamma$. Assume that the function $f : A \to \mathbb{R}$ is continuous and that $f(x, \cdot)$ is strictly concave, for each $x \in \mathbb{X}$. Define $g : \mathbb{X} \to \mathbb{Y}$ by*

$$g(x) = \arg\max_{y \in \Gamma(x)} f(x, y).$$

*Then for $\varepsilon > 0$and $x \in \mathbb{X}$, there exists $\delta_x > 0 \ni$*

$$y \in \Gamma(x) \text{ and } |f(x, g(x)) - f(x, y)| < \delta_x \text{ implies } ||g(x) - y|| < \varepsilon.$$

*If $\mathbb{X}$ is compact, then $\delta > 0$ can be chosen independently of $x$.*

*Proof.* Note that under the stated assumptions $g$ is a well-defined, continuous, and single valued. We first prove the claim for the case where $\mathbb{X}$ is compact. For each $\varepsilon > 0$ , define

$$A_\varepsilon = \{(x, y) \in A : ||g(x) - y|| \geq \varepsilon\}.$$

If $A_\varepsilon = \emptyset$, $\forall\ \varepsilon > 0$, then $\Gamma$ is single valued and the result is trivial. Otherwise there exists $\hat{\varepsilon} > 0$ sufficiently small such that for all $0 < \varepsilon < \hat{\varepsilon}$, the set $A_\varepsilon$ is nonempty and compact. For any such $\varepsilon$ define,

$$\delta = \min_{(x,y)\in A_\varepsilon} |f(x,g(x)) - f(x,y)|.$$

Since the function being minimized is continuous and $A_\varepsilon$ is compact, the minimum is attained. Moreover, since $[x, g(x)] \notin A_\varepsilon \ \forall\, x \in \mathbb{X}$, it follows that $\delta > 0$. Then,

$$y \in \Gamma(x) \ \text{ and } \ ||g(x) - y|| \geq \varepsilon \ \text{ implies } \ |f(x,g(x)) - f(x,y)| \geq \delta.$$

If $\mathbb{X}$ is not compact, the argument above can be applied separately for each fixed $x \in \mathbb{X}$. $\boxdot$

**Theorem 12.** *Let $\mathbb{X}, \mathbb{Y}, \Gamma, A$ be defined as in the Lemma above. Let $\{f_n\}$ be a sequence of continuous functions on $A$; assume that for each $n$ and each $x \in \mathbb{X}$, $f_n(x,\cdot)$ is strictly concave in its second argument. Assume that $f$ has the same properties and that $f_n \to f$ uniformly (in the* sup *norm). Define the functions $g_n$ and $g$ by*

$$\begin{aligned} g_n(x) &= \arg\max_{y\in\Gamma(x)} f_n(x,y), n = 1, 2 \cdots, \text{and} \\ g(x) &= \arg\max_{y\in\Gamma(x)} f(x,y). \end{aligned}$$

*Then, $g_n \to g$ pointwise.*

*Proof.* First note that since $g_n(x)$ is the unique maximizer of $f_n(x,\cdot)$ on $\Gamma(x)$, and $g(x)$ is the unique maximizer of $f(x,\cdot)$ on $\Gamma(x)$, it follows that

$$\begin{aligned} 0 &\leq f(x,g(x)) - f(x,g_n(x)) \\ &\leq f(x,g(x)) - f_n(x,g(x)) + f_n(x,g_n(x)) - f(x,g_n(x)) \\ &\leq 2||f - f_n||, \ \ \forall\, x \in \mathbb{X}. \end{aligned}$$

Since $f_n \to f$ uniformly, it follows immediately that for any $\delta > 0$, there exists $M_\delta \geq 1$ such that

$$0 \leq f(x,g(x)) - f(x,g_n(x)) \leq 2||f - f_n|| < \delta, \ \forall\, x \in \mathbb{X}, \forall\, n \geq M_\delta.$$

To show that $g_n \to g$ pointwise, we must establish that for each $\varepsilon > 0$ and $x \in \mathbb{X}$, there exists $N_x \geq 1$ such that

$$||g(x) - g_n(x)|| < \varepsilon, \forall\, n \geq N_x.$$

By the lemma above, it suffices to show that for any $\delta_x > 0$ and $x \in \mathbb{X}$ there exists

$N_x \geq 1$ such that

$$|f(x, g(x)) - f(x, g_n(x))| < \delta_x, \ \forall n \geq N_x.$$

It now follows that any $N_x \geq M_{\delta_x}$ has the required property. $\boxdot$

## 3.3 Dynamic Programming

The dynamic programs to be studied in this section and thereafter are represented in two different manners. The sequence problem (SP) is

$$\begin{aligned} &\max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\ &\text{s.t. } x_{t+1} \in \Gamma(x_t), t = 0, 1, 2, \cdots, \\ &x_0 \in \mathbb{X} \text{ given.} \end{aligned}$$

The alternative functional equation,(FE) , $v$ is

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)], \forall\, x \in \mathbb{X}.$$

Our focus is to characterize the properties of $F, \mathbb{X}, \mathbb{Y}, \Gamma$, and $A$ that are necessary and sufficient for the solution of SP to satisfy FE and conversely. This is desirable since there exist methods to solve the FE that might elucidate solutions to particular SPs. This result might seem surprising when one takes into account how each problem relates the states at each time period. The SP relates the states in term of sequences. At every point in time the SP might be solved with a different policy. On the other hand, the FE examines the problem one stage at a time. Moreover, the choice in the FE follows the same policy every stage.

**Definition 15.** *(Feasible Plans) A sequence $\{x_t\}_{t=0}^{\infty}$ in $\mathbb{X}$ is called a* plan. *Given $x_0 \in \mathbb{X}$, define*

$$\Pi(x_0) = \{\{x_t\}_{t=0}^{\infty} : x_{t+1} \in \Gamma(x_t), t = 0, 1, 2, \cdots\}$$

*as the set of* feasible plans from $x_0$.

The definition above provides a framework to relate the solutions of the FE to that of the SP. An element $\tilde{x} \in \Pi(x_0)$ is an infinite sequence, so that $\Pi(x_0)$ is a collection of sequences as defined by $\Gamma$. We proceed by establishing assumptions that drive our results. In Chapter 4, we reconcile these assumptions with those imposed by the neo-classical theory of economic growth on the economic agents being modelled.

**Assumption 1.** $\Gamma(x) \neq \emptyset, \forall\, x \in \mathbb{X}$.

**Assumption 2.** *For all $x_0 \in \mathbb{X}$ and $\tilde{x} \in \Pi(x_0)$,*

$$\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \in \mathbb{R}^*.$$

The following definition relates the return obtained at each period through the return function $F$ and the discount rate $\beta$ into the concept of utility.

**Definition 16.** *(Utility Function) For each $n = 0, 1, \cdots$, define $u_n : \Pi(x_0) \to \mathbb{R}$ by*

$$u_n(\tilde{x}) = \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}),$$

*so that $u_n(\tilde{x})$ is the partial sum of discounted returns through the $n^{th}$ period horizon. If assumption 2 holds, the infinite sum exists and we define*

$$u(\tilde{x}) = \lim_{n\to\infty} u_n(\tilde{x}).$$

**Definition 17.** *(Supremum Function) If assumptions 1 and 2 hold, then $\Pi(x_0) \neq \emptyset \forall x_0 \in \mathbb{X}$, and $F$ is well defined $\forall\, \tilde{x} \in \Pi(x_0)$. We define the* supremum function *as*

$$v^*(x_0) = \sup_{\tilde{x}\in\Pi(x_0)} u(\tilde{x}).$$

Being that $v^*$ is the supremum for the SP with a known $x_0$, it follows that:

1. If $|v^*(x_0)| < \infty$,
$$v^*(x_0) \geq u(\tilde{x}) \forall\, \tilde{x} \in \Pi(x_0), \tag{3.1}$$
and for any $\varepsilon > 0$,
$$v^*(x_0) \leq u(\tilde{x}) + \varepsilon \text{ for some } \tilde{x} \in \Pi(x_0). \tag{3.2}$$

2. If $v^*(x_0) = \infty$, then there exists a sequence of feasible plans $\{\tilde{x}_n\}$ such that
$$\lim_{n\to\infty} u(\tilde{x}_n) = \infty.$$

3. If $v^*(x_0) = -\infty$, then

$$u(\tilde{x}) = -\infty, \ \forall \ \tilde{x} \in \Pi(x_0).$$

The goal of this section is to find connections between the functions $v^*$ and $v$. We have that if assumptions 1 and 2 hold, $v^*$ is defined uniquely. Theorem 14 presents conditions guaranteeing $v^* = v$. In order to arrive at this result we first establish that the unique solution of the SP problem, $v^*$, satisfies the FE. This is established if the following three conditions hold:

1. If $|v^*(x_0)| < \infty$,

$$v^*(x_0) \geq F(x_0, y) + \beta v^*(y), \forall \ y \in \Gamma(x_0), \tag{3.3}$$

   and for any $\varepsilon > 0$,

$$v^*(x_0) \leq F(x_0, y) + \beta v^*(y) + \varepsilon \text{ for some } y \in \Gamma(x_0). \tag{3.4}$$

2. If $v^*(x_0) = \infty$, then there exists a sequence $\{y_n\} \subseteq \Gamma(x_0)$ such that

$$\lim_{n \to \infty} [F(x_0, y) + \beta v^*(y_n)] = \infty. \tag{3.5}$$

3. If $v^*(x_0) = -\infty$, then

$$F(x_0, y) + \beta v^*(y_k) = -\infty \ \forall y \in \Gamma(x_0). \tag{3.6}$$

The properties of $v^*$ that follow as a supremum-valued function and those outlined above have a striking resemblance. In fact, the only difference is that the conditions to be shown consider a single period and postpone the evaluation of the remaining states by relying on $v^*$ starting from the next period. The following Lemma is useful in establishing the fact that the solution to the SP satisfies the FE.

**Lemma 2.** *Let* $\mathbb{X}, \Gamma, F$*, and* $\beta$ *satisfy Assumption 2. Then, for any* $x_0 \in \mathbb{X}$ *and any* $(x_0, x_1, x_2, \cdots) = \tilde{x} \in \Pi(x_0)$,

$$u(\tilde{x}) = F(x_0, x_1) + u(\tilde{x}')$$

*where* $\tilde{x}' = (x_1, x_2, \cdots)$.

*Proof.* Under Assumption 2, for any $x_0 \in \mathbb{X}$ and any $\tilde{x} \in \Pi(x_0)$,

$$\begin{aligned} u(\tilde{x}) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \\ &= F(x_0, x_1) + \beta \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_{t+1}, x_{t+2}) \\ &= F(x_0, x_1) + \beta u(\tilde{x}). \end{aligned}$$

⊡

**Theorem 13.** *Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy Assumptions 1 and 2. Then $v^*$ satisfies (FE).*

*Proof.* If $\beta = 0$, the result is trivial. Suppose that $\beta > 0$ and choose $x_0 \in \mathbb{X}$. Suppose $v^*(x_0)$ is finite. Then (3.1) and (3.2) above hold. Let $x_1 \in \Gamma(x_0)$ and $\varepsilon > 0$. By (3.2), there exists $\tilde{x}' = (x_1, x_2, \cdots) \in \Pi(x_1)$ such that $u(\tilde{x}') \geq v^*(x_1) - \varepsilon$. Note that $\tilde{x} \in \Pi(x_0)$. From the 3.1 and the Lemma above, we have that

$$v^*(x_0) \geq U(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}') \geq F(x_0, x_1) + \beta v^*(x_1) - \beta\varepsilon.$$

Since $\varepsilon$ was chosen arbitrarily, condition (3.3) is established.

Now let $X_0 \in \mathbb{X}$, and $\varepsilon > 0$. From (3.2) and the Lemma above, we can choose $\tilde{x} = (x_0, x_1, \cdots) \in \Pi(x_0)$, so that

$$v^*(x_0) \leq u(\tilde{x}) + \varepsilon = F(x_0, x_1) + \beta u(\tilde{x}') + \varepsilon,$$

where $\tilde{x}' = (x_1, x_2, \cdots)$. It then follows from (3.1) that

$$v^*(x_0) \leq F(x_0) + \beta v^*(x_1) + \varepsilon,$$

since $x_1 \in \Gamma(x_0)$, this established (3.4).

If $v^*(x_0) = +\infty$, then there exists a sequence there exists a sequence $\{\tilde{x}_n\} \subseteq \Pi(x_0)$ such that $\lim_{n\to\infty} u(\tilde{x}_n) = +\infty$. Since $x_{1,n} \in \Gamma(x_0)$, $\forall n$, and

$$u(\tilde{x}_n) = F(x_0, x_{1,n}) + \beta u(\tilde{x}'_n) \leq F(x_0, x_{1,n}) + \beta v^*(x_{1,n}),\ \forall\, n,$$

it follows that (3.5) is established.

If $v^*(x_0) = -\infty$, then

$$u(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}') = -\infty, \forall\ (x_0, x_1, \cdots) = \tilde{x} \in \Pi(x_0),$$

where $\tilde{x}' = (x_1, x_2, \cdots)$. Since $F$ is real-valued, it follows that

$$u(\tilde{x}') = -\infty, \forall\, x_1 \in \Gamma(x_0), \forall\, \tilde{x}' \in \Pi(x_0).$$

Hence $v^*(x_0) = -\infty,\ \forall\ x_1 \in \Gamma(x_0)$. Since $F$ is real-valued and $\beta > 0$, the condition in (3.6) follows immediately. $\boxdot$

Insight regarding the result of the previous theorem is exposed when considering the sets of solution of the SP and the set of solutions of the FE. The previous theorem sates that the set of solutions to the FE contains the set of solutions of the SP. The following theorem provides a partial converse. That is, under an additional restriction, the set of solutions to the SP contains the set of solutions to the FE.

**Theorem 14.** *Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy Assumptions 1 and 2. If $v$ is a solution to (FE) and satisfies*

$$\lim_{n\to\infty} \beta^n v(x_n) = 0, \forall\, \tilde{x} \in \Pi(x_0), \forall\, x_0 \in \mathbb{X}, \tag{3.7}$$

*then $v = v^*$.*

*Proof.* If $v(x_0)$ is finite, then condition (3.3) and (3.4) hold. We show this implies (3.1) and (3.2) also hold. From (3.3), we have that for all $\tilde{x} \in \Pi(x_0)$,

$$\begin{aligned} v(x_0) &\geq F(x_0, x_1) + \beta v(x_1) \\ &\geq F(x_0, x_1) + F(x_1, x_2) + \beta^2 v(x_2) \\ &\vdots \\ &\geq u_n(\tilde{x}) + \beta^{n+1} v(n_{n+1}),\ n = 1, 2, \cdots \end{aligned}$$

Taking the limit at $\varepsilon \to \infty$ and using (3.7) above, we get that (3.1) holds.

Next, fix $\varepsilon > 0$ and choose $\{\delta_t\}_{t=1}^{\infty} \subset \mathbb{R}_+$ such that $\sum_{t=1}^{\infty} \beta^{t-1}\delta_t \leq \frac{\varepsilon}{2}$. Since (3.2) holds, choose $x_1 \in \Gamma(x_0), x_2 \in \Gamma(x_1), \cdots$ so that

$$v(x_t) \leq F(x_t, x_{t1}) + \beta v(x_{t+1}) + \delta_{t+1},\ t = 1, 2, \cdots$$

Then $\tilde{x} \in \Pi(x_0)$, and

$$\begin{aligned} v(x_0) &\leq \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) + \beta^{n+1} v(x_{n+1}) + (\delta_1 + \cdots + \beta^n \delta_{n+1}) \\ &\leq u_n(\tilde{x}) + \beta^{n+1} v(x_{n+1}) + \frac{\varepsilon}{2},\ n = 1, 2, \cdots \end{aligned}$$

Therefore, (3.7) implies that for $n$ sufficiently large, $v(x_0) \leq u_n(\tilde{x}) + \varepsilon$. Since $\varepsilon > 0$ was chosen arbitrarily , we have that (3.2) holds.

If (3.7) holds, then (3.6) implies that $v$ cannot take the value $-\infty$. If $v(x_0) = +\infty$, choose $n \geq 0$ and $(x_0, x_1, \cdots)$ such that $x_t \in \Gamma(x_{t-1})$ and $v(x_t) = +\infty$ for $t = 1, 2, \cdots, n$, and $v(x_{n+1}) \leq +\infty$ for all $x_{n+1} \in \Gamma(x_n)$. Clearly (3.7) implies that $n$ is finite. Fix $A > 0$. Since $v(x_n) = +\infty$, (3.5) implies that we can choose $x_{n+1,A} \in \Gamma(x_n)$ such that

$$F(x_n, x_{n+1,A}) + \beta v(x_{n+1,A}) \geq \beta^{-n} \left[ A + 1 - \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) \right].$$

Then choose $\tilde{x}_{n+1,A} \in \Pi(x_{n+1,A})$ such that $u(\tilde{x}_{n+1,A}) \geq v(x_{n+1,A}) - \beta^{-(n+1)}$. Since $v(x_{n+1,A})$ is finite, the argument above shows that this is possible. Then $\tilde{x}_A = (x_0, x_1, \cdots, x_n, \tilde{x}_{n+1,A}) \in \Pi(x_0)$ and

$$u(\tilde{x}_A) = \sum_{t=0}^{n-1} \beta^t F(x_t, x_{t+1}) + \beta^n F(x_n, x_{n+1,A}) + \beta^{n+1} u(\tilde{x}_{n+1,A}) \geq A.$$

Since $A > 0$ was chosen arbitrarily, it follows that $v^*(x_0) = +\infty$. ⊡

Theorem 15 relates the solutions to FE to those of SP by combining the approach in both problems. The theorem statement essentially prolongs the use of the FE approach one period and uses the SP to maximize the first period.

**Theorem 15.** *Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy assumptions 1 and 2. Let $\tilde{x}^* \in \Pi(x_0)$ be $\ni$ it attains the supremum in the SP for a given $x_0 \in \mathbb{X}$. Then,*

$$v^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta v_{t+1}^*, t = 0, 1, 2, \cdots. \tag{3.8}$$

*Proof.* Since $\tilde{x}^*$ attains the supremum,

$$\begin{aligned} v^*(x_t^*) &= u(\tilde{x}^*) = F(x_0, x_1^*) + \beta u(\tilde{x}^{*\prime}) \\ &\geq u(\tilde{x}) = F(x_0, x_1) + \beta u(\tilde{x}'),\ \forall\, \tilde{x} \in \Pi(x_0). \end{aligned} \tag{3.9}$$

In particular, the inequality holds for all plans with $x_t = x_t^*$. Since $(x_1^*, x_2, x_3, \cdots) \in \Pi(x_t^*)$ implies that $(x_0, x_1^*, x_2, x_3, \cdots) \in \Pi(x_0)$, it follows that

$$u(\tilde{x}^{*\prime}) \geq u(\tilde{x}'),\ \forall \tilde{x} \in \Pi(x_1^*).$$

hence $u(\tilde{x}^{*\prime}) = v(x_1^*)$. Substituting this into (3.9), gives (3.8) for $x_0$. Continuing by induction establishes (3.8) for all $t$. ⊡

The following theorem establishes a boundedness condition on the sequence $x_t$ and shows that any sequence satisfying this condition is an optimal plan.

**Theorem 16.** *Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy Assumptions 1 and 2. Let $\tilde{x}^* \in \Pi(x_0)$ be $\ni$ satisfies equation (3.8), and with*

$$\limsup_{t\to\infty} \beta^t v^*(x_t^*) \leq 0. \tag{3.10}$$

*Then $\tilde{x}^*$ attains the supremum in SP for a given $x_0$.*

*Proof.* Suppose that $\tilde{x}^* \in \Pi(x_0)$ satisfies (3.8) and (3.10). Then, it follows by induction on (3.8) that

$$v^*(x_0) = u_n(\tilde{x}^*) + \beta^{n+1} v^*(x_{n+1}^*),\ n = 1, 2, \cdots$$

Then, using (3.10), we find that $v^*(x_0) \leq u(\tilde{x}^*)$. Since $\tilde{x}^* \in \Pi(x_0)$, the reverse inequality holds, establishing the result . ⊡

As was mentioned above, the FE problems weighs the choice at every period using the same policy. In order to characterize this more formally, we introduce the concept of a *policy correspondence.* We study a special class of policy correspondences, that is those that are members of the set $C(\mathbb{X})$ of bounded and continuous functions.

**Definition 18.** *(Policy Correspondence) Given a solution $v \in C(\mathbb{X})$ to*

$$v(x) = \max_{y\in\Gamma(x)} [F(x, y) + \beta v(y)],$$

*we can define the policy correspondence $G : \mathbb{X} \to \mathbb{X}$ by*

$$G(x) = \{y \in \Gamma(x) : v(x) = F(x, y) + \beta v(y)\}.$$

Our focus now shifts to the study of policy correspondences. It is desirable to find conditions under which the policy correspondences have the properties necessary for the Theorem of the Maximum to describe the optimal solutions. Since policy correspondences are defined in terms of the return function $F$ and the correspondence $\Gamma$, $G$ obtains the desired properties by placing the following restrictions on these.

**Assumption 3.** *$\mathbb{X}$ is a convex subset of $\mathbb{R}^k$, and the correspondence $\Gamma : \mathbb{X} \to \mathbb{X}$ is nonempty, compact-valued, and continuous.*

**Assumption 4.** *The real valued function of the graph of $f$, $F : A \to \mathbb{R}$, is bounded and continuous, and $0 < \beta < 1$.*

It is of importance to note the relationship between the return function and the discount rate $\beta$. Although these are defined independently, their use in Theorem 17 further assumes they satisfy (3.7). In that context, the sequence $\beta^n$ is required to decrease at a sufficient pace to fulfill the condition established. The claim that Assumptions 3 and 4 modify the policy correspondence $G$ in exactly the appropriate way is substantiated in Theorem 17.

**Theorem 17.** *Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy Assumptions 3 and 4, and let $C(\mathbb{X})$ be the space of bounded continuous functions $f : \mathbb{X} \to \mathbb{R}$, with the* sup *norm. Then the unique operator $T$ maps $C(\mathbb{X})$ into itself, $T : C(\mathbb{X}) \to C(\mathbb{X})$; $T$ has a unique fixed point $v \in C(\mathbb{X})$; and for all $v_0 \in C(\mathbb{X})$,*

$$||T^n(v_0) - v|| \leq \beta^n ||v_0 - v||, n = 0, 1, 2, \cdots .$$

*Moreover, given $v$, the optimal policy correspondence $G : \mathbb{X} \to \mathbb{X}$ defined in 11 is compact-valued and u.h.c.*

*Proof.* Under Assumption 3 and 4, for each $f \in C(\mathbb{X})$ and $x \in \mathbb{X}$, the problem in (3.2) is to maximize the continuous function $[F(x, \cdot) + \beta f(\cdot)]$ over the compact set $\Gamma(x)$. Hence the maximum is attained. Since both $F$, and $f$ are bounded, clearly $Tf$ is also bounded; and since $F$ and $f$ are continuous, and $\Gamma$ is compact-valued and continuous, it follows from the Theorem of the Maximum that $Tf$ is continuous. Hence $T : C(\mathbb{X}) \to C(\mathbb{X})$.

It is them immediate that $T$ satisfies the condition's of Blackwell's sufficiency conditions for a contraction. Since $C(\mathbb{X})$ is a Banach space, it follows from the contraction mapping theorem, that $T$ has a unique fixed point $v \in C(\mathbb{X})$, and so (3.3) holds. The stated properties of $G$ then follow from the Theorem of the Maximum. $\boxdot$

In order to characterize the value function $v$ and the policy correspondence $G$, we require further assumptions on the behavior of $\Gamma$ and $F$. The following assumptions establish properties of $F$ and $\Gamma$ that allow us to use previous results about fixed points and contraction mappings.

**Assumption 5.** *For each* $y, F(\cdot, y)$ *is strictly increasing in each of its first* $k$ *arguments.*

**Assumption 6.** $\Gamma$ *is monotone in the sense that* $x \leq x'$ *implies* $\Gamma(x) \subseteq \Gamma(x')$.

To prove Theorem 18, we seek to show that $\Gamma$ is u.h.c.. Assumption 6 provides the required property to show this. If one imagines a sequence $\{x_n\}$ converging to its limit point, it follows that the sequence $\{\Gamma(x_n)\}$ is nested. This imagery helps visualize the over approximation of the target set as required by u.h.c. correspondences. Assumption 5 alone does not provide the required information about the behavior of $F$. However, in combination with Assumptions 4 and 2, we have enough information to present the following theorem.

**Theorem 18.** *Let* $\mathbb{X}, \Gamma, F,$ *and* $\beta$ *satisfy Assumptions 3–6, and let* $v$ *satisfy*

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)].$$

*Then* $v$ *is strictly increasing.*

*Proof.* Let $C'(\mathbb{X}) \subset C(\mathbb{X})$ be the set of bounded, continuous, and nondecreasing functions on $\mathbb{X}$ , and let $C''(\mathbb{X}) \subset C'(\mathbb{X})$ be the set of strictly increasing functions. Since $C'(\mathbb{X})$ is a closed subset of the complete metric space $C(\mathbb{X})$, it then follows that if $T[C'(\mathbb{X})] \subseteq T[C''(\mathbb{X})]$. Assumptions 5 and 6 ensure this is so. $\boxdot$

**Assumption 7.** $F$ *is strictly concave; that is,*

$$F[\theta(x, y) + (1 - \theta)(x', y')] \geq \theta F(x, y) + (1 - \theta)F(x', y'),$$

$\forall$ $(x, y), (x', y') \in A$, *and all* $\theta \in (0, 1)$, *and the inequality is strict if* $x \neq x'$.

**Assumption 8.** $\Gamma$ *is convex in the sense that for any* $0 \leq \theta \leq 1$, *and* $x, x' \in \mathbb{X}$,

$$\begin{gathered} y \in \Gamma(x) \quad \text{and} \quad y' \in \Gamma(x') \quad \text{implies} \\ \theta y + (1 - \theta)y' \in \Gamma[\theta x + (1 - \theta)x']. \end{gathered}$$

It is worth noting that the convexity assumption stated in Assumption 8 is established within the same period. That is, it is assumed that $x$ and $x'$ belong to a given $\Gamma(x)$ in the same time period. Convexity is not established inter-temporally.

**Theorem 19.** *Let $\mathbb{X}, \Gamma, F$ and $\beta$ satisfy Assumptions 3–4 and 7–8; let $v$ satisfy*

$$v(x) = \max_{y \in \Gamma(x)} [F(x,y) + \beta v(y)],$$

*and let $G$ satisfy*

$$G(x) = \{y \in \Gamma(x) : v(x) = F(x,y) + \beta v(y)\}. \tag{3.11}$$

*Then, $v$ is strictly concave and $G$ is a continuous, single-valued function.*

*Proof.* Let $C'(\mathbb{X}) \subset C(\mathbb{X})$ be the set of bounded, continuous, weakly concave functions on $\mathbb{X}$ , and let $C''(\mathbb{X}) \subset C'(\mathbb{X})$ be the set of strictly concave functions. Since $C'(\mathbb{X})$ is a closed subset of a complete metric space $C(\mathbb{X})$, by Theorem 17 and Corollary 1 to the Contraction Mapping Theorem, it is sufficient to show that $T[C'(\mathbb{X})] \subseteq C''(\mathbb{X})$.

To verify that this is so, let $f \in C'(\mathbb{X})$ and let

$$x_0 \neq x_1, \theta \in (0,1), \text{ and } x_\theta = \theta x_0 + (1-\theta)x_1.$$

Let $y_i \in \Gamma(x_i)$ attain $(Tf)x_i$, for $i = 0, 1$. It follows that

$$\begin{aligned} (Tf) &\geq F(x_\theta, y_\theta) + \beta f(y_\theta) \\ &\geq \theta[F(x_0, y_0) + \beta f(y_0)] + (1-\theta)F(x_1, y_1) + \beta f(y_1) \\ &= \theta(Tf)x_0 + (1-\theta)(Tf)x_1, \end{aligned}$$

where the first line uses (3.2) and the fact that $y_\theta \in \Gamma(x_\theta)$; the second uses the hypothesis that $f$ is concave and the concavity restriction of $F$ in Assumption 7; and the last follows from the way $y_0, y_1$ were selected. Since $x_0$ and $x_1$ were arbitrary, it follows that $Tf$ is strictly concave, and since $f$ was arbitrary, it follows that $T[C'(\mathbb{X})] \subseteq C''(\mathbb{X})$.

Hence, the unique fixed point $v$ is strictly concave. Since $F$ is also concave (Assumption 7) and, for each $x \in \mathbb{X}, \Gamma(x)$ is concave (Assumption 8), it follows that the maximum in (3.2) is attained. ⊡

Theorems 19 and 20 characterize $v$ by using the fact that the operator $T$ preserves certain properties (Stokey et al., 1989). In many cases, it is difficult to solve for the policy correspondences. In these cases, it is desirable to use approximations of the policy function instead. Theorem 20 outlines the conditions under which this is possible.

**Theorem 20.** *(Convergence of the Policy Functions) Let $\mathbb{X}, \Gamma, F$, and $\beta$ satisfy Assumptions 3–4 and 7-8, and let $v$ and $g$ satisfy*

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)], \text{ and } g(x) = \{y \in \Gamma(x) : v(x) = F(x, y) + \beta v(y)\}.$$

*Let $C'(\mathbb{X})$ be the set of bounded, continuous, concave functions $f : \mathbb{X} \to \mathbb{R}$, and let $v_0 \in C'(\mathbb{X})$. Let $\{(v_n, g_n)\}$ be defined by*

$$\begin{aligned} v_{n+1} &= Tv_n, \quad n = 0, 1, 2, \cdots, \text{and} \\ g_n(x) &= \arg\max_{y \in \Gamma(x)} [F(x, y) + \beta v_n(y)], \quad n = 0, 1, 2, \cdots \end{aligned}$$

*Then $g_n \to g$ pointwise. If $\mathbb{X}$ is compact, then the convergence is uniform.*

*Proof.* Let $C''(\mathbb{X}) \subset C'(\mathbb{X})$ be the set of strictly concave functions $f : \mathbb{X} \to \mathbb{R}$. As shown in Theorem 19, $v \in C''(\mathbb{X})$. Moreover, as shown in the proof of that Theorem, $T[C'(\mathbb{X})] \subset C''(\mathbb{X})$. Since $v_0 \in C'(\mathbb{X})$, it follows that every function $v_n, n = 1, 2, \cdots$, is strictly concave. Define the functions $\{f_n\}$ and $f$ by

$$\begin{aligned} f_n(x, y) &= F(x, y) + \beta v_n(y), \ n = 1, 2, \cdots, \text{ and} \\ f(x, y) &= F(x, y) + \beta v(y). \end{aligned}$$

Since $F$ satisfies Assumption 7, it follows that f and each function $f_n, n = 1, 2, \cdots$, is strictly concave. Hence Theorem 19 applies and the desired results are proved. ⊡

**Assumption 9.** *$F$ is differentiable on the interior of $A$.*

It is a logical consequence of Assumption 9 that the function $v$ in the FE is differentiable. The proof of this result assumes knowledge of results in the field of convex analysis that are considered to be overly esoteric for the target audience, and is thus omitted. Readers interested in a formal development of this result would benefit from the treatment of sub-gradients in Rockafellar (1970) and the statement of the result in Stokey et al. (1989). In the final section of this chapter we rely on this result to establish the classical variational result first introduced by Euler and then proven sound by Lagrange (McShane, 1989).

## 3.4 Euler Equations

In this section, we present a solution mechanism to recursive problems. The calculus of variations approach to dynamic problems such as the brachiostochrone and isoparametric problems motivated mathematicians to establish the results presented in this section. The isoparametric problem, first solved by Lagrange, provided the insight to establish the sufficiency of the first-order conditions proposed by Euler (Ferguson, 2004). Although, the version of the results presented here follow the SP approach to dynamic problems, the same results may be obtained following the FE approach.

**Definition 19.** *(Euler Equations) Let $F$ satisfy assumptions 3-5,7, and 9; let $F_x$ denote the $l$-vector of partial derivatives $(F_1, \cdots, F_l)$in its first $l$ arguments, $F_y$ denote the vector $(F_{l+1}, \cdots, F_{2l})$. Since $F$ is continuously differentiable and strictly concave, if $x^*_{t+1}$ is in the interior of $\Gamma(x^*_t)$ for all $t$, the first-order conditions for*

$$\max_y [F(x^*_t, y) + \beta F(y, x^*_{t+2})]$$
$$\text{s.t. } y \in \Gamma(x^*_t) \text{ and } x^*_{t+2} \in \Gamma(y)$$

*are*

$$0 = F_y(x^*_t, x^*_{t+1}) + \beta F_x(x^*_{t+1}, x^*_{t+2}), \quad t = 0, 1, 2, \cdots .$$

The Euler equations described above comprise a "system of $l$ second-order differential equations in the vector of state variables $x_t$" Stokey et al. (1989). In order to solve for the optimal solution, we require an $l$-vector of boundary conditions. The condition we refer to as the *transversality condition* provides these conditions.

**Definition 20.** *(Transversality Condition) The transversality condition states that*

$$\lim_{t\to\infty} \beta^t F_x(x_t*, x^*_{t+1}) \cdot x^*_t = 0.$$

We are now ready to establish the final result of this guide, that Definitions 19 and 20 are sufficient conditions to obtain the maximizing sequence of the SP.

**Theorem 21.** *(Sufficiency of the Euler and transversality conditions) Let $\mathbb{X} \subseteq \mathbb{R}^l_+$, and let $F$ satisfy Assumptions 3-5,7, and 9. Then the sequence $\{x^*_{t+1}\}^\infty_{t=0}$ with $x^*_{t+1} \in \Gamma(x^*_t)\backslash\Gamma(x^*_t)'$, $t = 0, 1, 2, \cdots$, is optimal for the SP problem, given $x_0$, if it satisfies the Euler and transversality conditions stated in Definitions 19 and 20.*

*Proof.* Let $x_0$ be given; let $\{x^*_t\} \in \Pi(x_0)$ satisfy the conditions imposed in Definitions 19 and 20; and let $\{x_t\} \in \Pi(x_0)$ be any feasible sequence. It is sufficient to show that the difference, call it $D$, between the objective function in SP evaluated at $\{x_t^*\}$ and at $\{x_t\}$ is nonnegative.

Since $F$ is continuous, concave, and differentiable as according to Assumptions 4, 7 and 9,

$$\begin{aligned} D &= \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t [F(x_t^*, x_{t+1}^*) + F(x_t, x_{t+1})] \\ &\geq \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t [F_y(x_t^*, x_{t+1}^*) \cdot (x_t^* - x_t) - F_x(x_t^*, x_{t+1}^*) \cdot (x_{t+1}^* - x_{t+1})], \end{aligned}$$

where the second inequality follows from the fact that the linear approximation of a concave function lies above the function itself. By the definition of $\Pi(x_0)$, we have that $x_0^* - x_0 = 0$. Then, rearranging terms gives

$$\begin{aligned} D \geq \lim_{T\to\infty} \Bigg\{ &\sum_{t=0}^{T-1} \beta^t [F_y(x_t^*, x_{t+1}^*) + \beta F_x(x_{t+1}^*, x_{t+2}^*)] \cdot (x_{t+1}^* - x_{t+1}) \\ &+ \beta^T F_y(x_T^*, x_{T+1}^*) \cdot (x_T^* - x_T) \Bigg\}. \end{aligned}$$

Since $\{x_t^*\}$ satisfies Definition 19, the summation of the first $T-1$ terms is zero. Therefore, substituting the equation in Definition 19 into the last term as well and using Definition 20 yields

$$\begin{aligned} D &\geq -\lim_{T\to\infty} \beta^T F_x(x_T^*, x_{T+1}^*) \cdot (x_T^* - x_T) \\ &\geq -\lim_{T\to\infty} \beta^T F_x(x_T^*, x_{T+1}^*) \cdot x_T^*, \end{aligned}$$

where the last line uses the fact that $F_x \geq 0$ by Assumption 4, and $x_t \geq 0, \forall\, t$. It then follows from Definition 20 that $D \geq 0$, establishing the desired result. $\boxdot$

Definitions 19 and 20 establish a straightforward procedure to solve recursive problems. Moreover, Theorem 21 proves their sufficiency so that the optimal sequence is obtained. The objective in Chapter 3 is fulfilled with the development of the rigorous set of results that prove the existence of a maximum, relate the solutions of the SP and the FE, and outline a solution mechanism. Chapter 4 considers the application of dynamic programming to the problem of Economic Growth and provides economic intuition behind the restrictions placed on $\Gamma, F, \mathbb{X}, \mathbb{Y}$ and $v$ in Chapter 3.

# Chapter 4

# Deterministic Economic Growth

In the final chapter of this guide we seek to elucidate some intuition about the results of the previous chapter. It is easy to get lost in the technicality of the field and forget that we are developing these methods not only for the sake of their beauty but also to apply them to economic theory. In this chapter, we present the one-sector model of optimal economic growth. The SP problem is presented in an abstract manner to indicate that any function satisfying the conditions outlined could be modelled in this way.

The neo-classical theory of economic growth associates behavioral assumptions of economic agents, that is households and firms, with many of the properties we give the utility and production functions in this example. We owe the formulation of this problem to Stokey et al. (1989), and the solutions to the author of this guide.

Consider the abstract notion of an economy with one good at every time period, the amount of which is $y_t$. The good is produced with the technology described by the gross production function $G(k_t, n_t) = y_t$, where $k_t$ represents capital and $n_t$ represents labor at time $t$. Neo-classical growth theory has that the gross production function $G$ has the following properties. First, since it relates current inputs to outputs it is defined over the graph of the correspondence $\Gamma$. That is, $G : A \to \mathbb{R}_+$. Moreover, $G$ is continuously differentiable, strictly increasing, homogeneous of degree one, and strictly quasi-concave. A function is homogenous of degree one if $G(tk, tn) = t^1 G(k, n)$. Moreover, a function is quasi-concave if the upper contour set of the function is concave. Further, $G$ has the properties that

$$G(0, n) = 0, \quad G_k(k, n) > 0, \quad G_n(k, n) > 0, \ \forall\ k, n > 0$$
$$\lim_{k \to 0} \ G_k(k, 1) = \infty, \quad \lim_{k \to \infty} G_k(k, 1) = 0,$$

where $G_k$ represents the partial derivative of the function with respect to its argument $k$. The two limits above imply conditions on the marginal rate of productivity.

At each period, households face the choice of allocating the output between investment, denoted $i_t$, and consumption, denoted $c_t$, so that we may relate

$$G(k_t, n_t) = y_t \geq c_t + i_t. \tag{4.1}$$

The budget constraint in equation (4.1) states that a household cannot consume or invest more of the good than what is available at a given time period. Capital is assumed to depreciate at rate $\delta \in [0, 1]$ every period. So that at every time period the following equation must hold

$$k_{t+1} = (1 - \delta)k_t + i_t. \tag{4.2}$$

Combining the expressions in equations 4.1 and 4.2, we get that

$$G(k_t, n_t) \geq c_t + k_{t+1} - (1 - \delta)k_t. \tag{4.3}$$

In this model, we do not consider the household's choice of providing labor and assume $n_t = 1 \; \forall \; t$. In order to express production in terms of net depreciation, and given that $n_t$ has been normalized to unity, we define

$$g(k_t) = G(k_t, 1) + (1 - \delta)k_t, \tag{4.4}$$

and note that $g(k_t) \geq c_t + k_{t+1}$. We assume non-satiability, which is defined to mean that we get equality in the previous equation. We now study the properties of the function $g(k_t)$. From the definition of $g$ as the sum of two continuously differentiable and strictly increasing functions we have that it inherits these properties. We also have that the sum of a strictly concave function and a linear function maintains the property of strict concavity. Moreover, we have that

$$\begin{aligned}
g(0) &= G(0, 1) = 0, \\
g'(k) &= G_k(k, 1) + (1 - \delta) > 0, \\
\lim_{k \to 0} g'(k) &= \lim_{k \to 0} [G_k(k, 1) + (1 - \delta)] = \infty, \text{and} \\
\lim_{k \to \infty} g'(k) &= \lim_{k \to \infty} [G_k(k, 1) + (1 - \delta)] = (1 - \delta),
\end{aligned}$$

so that it preserves the properties of $G$ modifying the last one by a horizontal shift.

The household's role is to find a sequence of consumptions that maximizes its utility. That is, it faces the SP described by

$$\max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t), \tag{4.5}$$

where $\beta \in (0,1)$ is the rate of time preference or discounting factor, and $U$ is the utility function of the household. The assumptions placed on $U$ are that it is a bounded function from $\mathbb{R}_+$ into $\mathbb{R}$ that is continuously differentiable, strictly increasing, strictly concave, and $\lim_{c \to 0} U'(c) = \infty$.

It is implicit in the formulation of this problem that household utility is additively separable through time and is time-invariant. That is, the utility function does not express an intrinsic trade off between consumption between periods. This trade-off is introduced through the state variable of capital. The utility function is time invariant since $U(c_t) = U(c_{t+1})$ if and only if $c_t = c_{t+1}$. These are standard assumptions in the literature of neo-classical economic theory, however, they should be kept in mind when interpreting the results.

Using our previous derivation of the net production function we recast the maximization problem in terms of choosing a sequence of capital that maximizes the household's utility in this way:

$$\begin{aligned} &\max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(g(k_t) - k_{t+1}) \\ \text{s.t.} \quad &k_{t+1} \in \Gamma(k_t), t = 0, 1, 2, \cdots, \\ &k_0 \in \mathbb{X} \text{ given.} \end{aligned}$$

We proceed to show that the assumptions made in Chapter 3 to develop the desired results follow from the behavioral assumptions imposed by neo-classical economics on the agents. That is, we wish to show that the functional restrictions imposed in the previous Chapter are consistent with economic theory and that the methods of dynamic programming developed so far are a result of the need to model economic behavior. Before we proceed, we summarize for conciseness the properties of the utility function $U$ and the net production function $g$.

| | | | |
|---|---|---|---|
| U1. | $0 < \beta < 1$. | T1. | $g$ is continuous. |
| U2. | $U$ is bounded. | T2. | $g(0) = 0$. |
| U3. | $U$ is strictly increasing. | T3. | $g$ is strictly increasing. |
| U4. | $U$ is strictly concave. | T4. | $g$ is weakly concave. |
| U5. | $U$ is continuously differentiable. | T5. | $g$ is continuously differentiable. |

We now proceed to evaluate the set of assumptions made in previous chapters.

Henceforth, we refer to the function $F$ as the inter-temporal function used in previous chapters. In this example, the function $F(k_t, k_{t+1}) = U(g(k_t) - k_{t+1})$ and the correspondence $\Gamma(k_t) = [0, g(k_{t-1})]$.

**Assumption 1.** $\Gamma(k) \neq \emptyset, \forall\, k \in \mathbb{X}$.
Since $\Gamma(k_t)$ was defined to be $\Gamma(k_t) = [0, g(k_{t-1})]$, and $g(0) = 0$ we have that $0 \in \Gamma(k_t)\ \forall\, t$ and hence $\Gamma(k_t)$ is always nonempty.

**Assumption 2.** *For all* $k_0 \in \mathbb{X}$ *and* $\tilde{k} \in \Pi(k_0)$,

$$\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t U[g(k_t) - k_{t+1}] \in \mathbb{R}^*.$$

Since $U$ is bounded, we have that $\exists\, M \in \mathbb{N} \ni |U(\cdot)| \leq M$.
So, $\lim\limits_{n\to\infty} \sum\limits_{t=0}^{n} \beta^t U[g(k_t) - k_{t+1}] \leq M \lim\limits_{n\to\infty} \sum\limits_{t=0}^{n} \beta^t = \frac{M}{1-\beta}$.

**Assumption 3.** $\mathbb{X}$ *is a convex subset of* $\mathbb{R}^k$*, and the correspondence* $\Gamma : \mathbb{X} \to \mathbb{X}$ *is nonempty, compact-valued, and continuous.*

**Assumption 4.** *The function* $U[g(k_t) - k_{t+1}] : A \to \mathbb{R}$ *is bounded and continuous, and* $0 < \beta < 1$.
We showed $U$ is bounded in Assumption 2. We have that both $U$ and $g(k_t) - k_{t+1}$ are continuous functions. Since the composition of continuous functions is continuous as well we have the second result. The third follows from hypothesis.

**Assumption 5.** *For each* $k_{t+1}, U[g(k_t) - k_{t+1}]$ *is strictly increasing in* $k_t$ *and* $k_{t+1}$.
From T3 and U3 we have that the function is strictly increasing.

**Assumption 6.** $\Gamma$ *is monotone in the sense that* $k \leq k'$ *implies* $\Gamma(k) \subseteq \Gamma(k')$.
It follows from $\Gamma(k) = [0, g(k)]$, that $k \leq k'$ implies

$$\Gamma(k) = [0, g(k)] \subseteq [0, g(k')] = \Gamma(k').$$

**Assumption 7.** $G$ *is strictly concave; that is,*

$$G[\theta(k_t, k_{t-1}) + (1-\theta)(k_t', k_{t+1}')] \geq \theta G(k_t, k_{t+1}) + (1-\theta) G(k_t', k_{t+1}'),$$

$\forall\ (k_t, k_{t+1}), (k_t', k_{t+1}') \in A$, *and all* $\theta \in (0,1)$, *and the inequality is strict if* $k_t \neq k_t'$.
This property is a direct consequence of U4.

**Assumption 8.** $\Gamma$ *is convex.* This property follow from the characterization of $\Gamma$ in term of intervals in this problem.

The application of dynamic programming to economic modelling has been shown to be consistent with the assumptions of economic theory. The problem presented in this chapter has been shown to suffice the conditions developed in Chapter 3, and is thus well defined. Using Assumption 9 and the methods presented in the last section of Chapter 3, one can solve the problem for specific functional characterizations of utility and production.

As was mentioned, several extensions including stability dynamics, stochastic models, and numerical solutions make the application of dynamic programming useful to problems in several disciplines. We hope the reader might continue the study of this field as much as the author has enjoyed this project and will continue the study of this field in the Economics Ph.D. program at the University of California at Berkeley.

# Appendix A

# Notation Index

| | |
|---|---|
| $\boxdot$ | *quod erat demonstrandum*(QED) |
| $x \in \mathbb{X}$ | membership |
| $\mathbb{X} \subset \mathbb{Y}, \mathbb{X} \subseteq \mathbb{Y}$ | strict subset, subset |
| $\bigcap, \bigcup$ | intersection, union |
| $\emptyset$ | empty set |
| $\setminus$ | set deletion |
| $\mathbb{X}^c$ | complement |
| $\mathbb{X}'$ | derived set |
| $\bar{\mathbb{X}}$ | closure |
| $\mathbb{X} \times \mathbb{Y}$ | cartesian product |
| $\mathbb{N}$ | natural numbers |
| $\mathbb{R}, \mathbb{R}^*$ | real numbers, extended real numbers |
| $\mathbb{R}^k$ | $k^{th}$ dimension Euclidean space |
| $\mathbb{R}^k_+$ | non-negative subspace of $\mathbb{R}^k$ |
| $\mathbb{R}^k_{++}$ | positive subspace of $\mathbb{R}^k$ |
| $d_{\mathbb{X}}(x, y)$ | distance function in the metric space $(\mathbb{X}, d)$ |
| $(\mathbb{X}, d_{\mathbb{X}})$ | metric space |
| $N_\delta(p)$ | $\delta$-neighborhood around the point $p$ |
| $\lvert \cdot \rvert$ | absolute value |
| $\lvert\lvert \cdot \rvert\rvert$ | norm |
| $B(\mathbb{X})$ | space of bounded functions on $\mathbb{X}$ |
| $C(\mathbb{X})$ | space of bounded and continuous functions on $\mathbb{X}$ |
| $f^n(\mathbb{X})$ | iterates of the function $f$ on $\mathbb{X}$ |
| $T^n(B(\mathbb{X}))$ | iterates of the function $T$ on the function space $B(\mathbb{X})$ |
| $\Pi(x_0)$ | the set of feasible plans from $x_0$ |

# Bibliography


Aghion, P. and R. Griffith (2005). *Competition and Growth : Reconciling Theory and Evidence*. The MIT Press.

Barro, R. and X. Sala-i-Martin (2003). *Economic Growth* (Second ed.). Boston, MA: The MIT Press.

Bellman, R. (1957). *Dynamic Programming*. Princeton, New Jersey: Princeton University Press.

Bryant, V. (1985). *Metric Spaces*. Cambridge, UK: Cambridge University Press.

Chiang, A. C. (1984). *Fundamental Methods of Mathematical Economics* (Third ed.). McGraw-Hill.

Diáz-Giménez, J. (2001). *Computational Methods for the Study of Dynamic Economies*, Chapter 1–Linear quadratic approximations: An introduction, pp. 13–30. USA: Oxford University Press.

Ferguson, J. (2004). A brief survey of the history of the calculus of variations and its applications. University of Victoria HO/0402357, Arxiv Preprint Math. available at http://arxiv.org/pdf/math/0402357.

Fischer, S. and O. Blanchard (1989). *Lectures on Macroeconomics*. Boston, MA: The MIT Press.

Fomin, S. V. and I. M. Gelfand (2000). *Calculus of Variations*. Dover Publications.

Goldstein, R. (2005). *Incompleteness: the proof and paradox of Kurt Gödel*. New York: Norton & Company.

Holder, A. (2005). *An Introduction to Mathematical Optimization* (First ed.). San Antonio, TX, 2005.

Hoy, M., J. Livernois, C. McKenna, R. Rees, and T. Stengos (2001). *Mathematics for Economics* (Second ed.). Boston, MA: The MIT Press.

Lewis Carroll (1999). *The Annotated Alice: The Definitive Edition*. W. W. Norton & Company.

Ljungqvist, L. and T. Sargent (2004). *Recursive Macroeconomic Theory* (Second ed.). Boston, MA: The MIT Press.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. USA: Oxford University Press.

McShane, E. (1989, September). The calculus of variations from the beginning through optimal control theory. *SIAM Journal on Control and Optimization 27*(5), 915–940.

Niehans, J. (1994). *A History of Economic Theory*. Baltimore, MA: The Johns Hopkins University Press.

Nobel, F. (2004). The bank of sweden prize in economic sciences in memory of alfred nobel 2004. [web page] `http://nobelprize.org/economics/laureates/2004/index.html`. [Accessed on April 26, 2006].

Pontryagin, L. S. et al. (1964). *The Mathematical Theory of Optimal Processes*. New York: The Macmillan Company.

Rockafellar, T. (1970). *Convex analysis*. Princeton, N.J.: Princeton University Press.

Romer, D. (2002, June). It's fourth down and what does the bellman equation say? a dynamic programming analysis of football strategy. NBER Working Papers 9024, National Bureau of Economic Research, Inc. available at http://ideas.repec.org/p/nbr/nberwo/9024.html.

Royden, H. L. (1988). *Real Analysis* (Third ed.). Princeton, New Jersey: Prentice Hall.

Rudin, W. (1976). *Principles of Mathematical Analysis* (Third ed.). McGraw-Hill.

Sala-i-Martin, X. (2000). *Apuntes de Crecimiento Economico* (Second ed.). Barcelona: Antoni Bosch.

Shone, R. (2002). *Economic Dynamics : Phase Diagrams and their Economic Application* (Second ed.). Cambridge, UK: Cambridge University Press.

Spiegel, H. W. (1983). *The Growth of Economic Tought* (Third ed.). Durham, NC: Duke University Press.

Stokey, N., R. Lucas, and E. Prescott (1989). *Recursive Methods in Economic Dynamics*. Boston, MA: Harvard University Press.

Vohra, R. (2005). *Advanced Mathematical Economics*. New York: Routledge.

Yong, J. and X. Y. Zhou (1999). *Stochastic Controls : Hamiltonian Systems and HJB Equations*. Springer.

Young, L. C. (2000). *Lecture on the Calculus of Variations and Optimal Control Theory*. Providence, Rhode Island: AMSChelsea Publishing.