capital and enjoy total utility $v_1(y)$. His problem in the first period is thus $\max[U(c) + \beta v_1(y)]$, subject to the constraints in (1). The maximized value of this objective function was defined, in (3), as $v_2(k)$. Hence it follows from (4) that

$$\begin{aligned} v_2(k) &= \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v_1(y)\} \\ &\ge \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v_0(y)\} \\ &= v_1(k). \end{aligned}$$

Continuing in this way, one establishes by induction that $v_{n+1}(k) \ge v_n(k)$, all k, $n = 0, 1, 2, \ldots$. The successive approximations defined in (3) are improvements, reflecting the fact that planning flexibility over longer and longer finite horizons offers new options without taking any other options away. Consequently it seems reasonable to suppose that the sequence of functions $\{v_n\}$ defined in (3) might converge to a solution v to (1). That is, the method of successive approximations seems to be a reasonable way to locate and characterize solutions.

This method can be described in a somewhat different and much more convenient language. As we showed in the discussion above, for any function w: $\mathbf{R}_+ \to \mathbf{R}$, we can define a new function—call it Tw: $\mathbf{R}_+ \to \mathbf{R}$—by

$$(5) \qquad (Tw)(k) = \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta w(y)\}.$$

When we use this notation, the method of successive approximations amounts to choosing a function v_0 and studying the sequence $\{v_n\}$ defined by $v_{n+1} = Tv_n$, $n = 0, 1, 2, \ldots$. The goal then is to show that this sequence converges and that the limit function v satisfies (1). Alternatively, we can simply view the operator T as a mapping from some set C of functions into itself: T: $C \to C$. In this notation solving (1) is equivalent to locating a *fixed point* of the mapping T, that is, a function $v \in C$ satisfying $v = Tv$, and the method of successive approximations is viewed as a way to construct this fixed point.

To study operators T like the one defined in (5), we need to draw on several basic mathematical results. To show that T maps an appropriate space C of functions into itself, we must decide what spaces of functions are suitable for carrying out our analysis. In general we want to limit attention to continuous functions. This choice raises the issue of whether, given a continuous function w, the function Tw defined by (5) is also continuous. Finally, we need a fixed-point theorem that applies to operators like T on the space C we have selected. The rest of the chapter deals with these issues.

In Section 3.1 we review the basic facts about metric spaces and normed vector spaces and define the space C that will be used repeatedly later. In Section 3.2 we prove the Contraction Mapping Theorem, a fixed-point theorem of vast usefulness. In Section 3.3 we review the main facts we will need about functions, like Tw above, that are defined by maximization problems.

3.1 Metric Spaces and Normed Vector Spaces

The preceding section motivates the study of certain functional equations as a means of finding solutions to problems posed in terms of infinite sequences. To pursue the study of these problems, as we will in Chapter 4, we need to talk about infinite sequences $\{x_t\}_{t=0}^{\infty}$ of states, about candidates for the value function v, and about the convergence of sequences of various sorts. To do this, we will find it convenient to think of both infinite sequences and certain classes of functions as elements of infinite-dimensional normed vector spaces. Accordingly, we begin here with the definitions of vector spaces, metric spaces, and normed vector spaces. We then discuss the notions of convergence and Cauchy convergence, and define the notion of completeness for a metric space. Theorem 3.1 then establishes that the space of bounded, continuous, real-valued functions on a set $X \subseteq \mathbf{R}^l$ is complete.

We begin with the definition of a vector space.

DEFINITION *A (real)* ***vector space*** *X is a set of elements (vectors) together with two operations, addition and scalar multiplication. For any two vectors $x, y \in X$, addition gives a vector $x + y \in X$; and for any vector $x \in X$ and any real number $\alpha \in \mathbf{R}$, scalar multiplication gives a vector $\alpha x \in X$. These operations obey the usual algebraic laws; that is, for all x, y, $z \in X$, and $\alpha, \beta \in \mathbf{R}$:*

a. $x + y = y + x$;

b. $(x + y) + z = x + (y + z)$;

c. $\alpha(x + y) = \alpha x + \alpha y$;
d. $(\alpha + \beta)x = \alpha x + \beta x$; *and*
e. $(\alpha\beta)x = \alpha(\beta x)$.

Moreover, there is a zero vector $\theta \in X$ *that has the following properties:*
f. $x + \theta = x$; *and*
g. $0x = \theta$.

Finally,
h. $1x = x$.

The adjective "real" simply indicates that scalar multiplication is defined taking the real numbers, not elements of the complex plane or some other set, as scalars. All of the vector spaces used in this book are real, and the adjective will not be repeated. Important features of a vector space are that it has a "zero" element and that it is closed under addition and scalar multiplication. Vector spaces are also called *linear spaces.*

Exercise 3.2 Show that the following are vector spaces:
a. any finite-dimensional Euclidean space $\mathbf{R}^l$;
b. the set $X = \{x \in \mathbf{R}^2: x = \alpha z, \text{ some } \alpha \in \mathbf{R}\}$, where $z \in \mathbf{R}^2$;
c. the set X consisting of all infinite sequences $(x_0, x_1, x_2, \ldots)$, where $x_i \in \mathbf{R}$, all i;
d. the set of all continuous functions on the interval $[a, b]$.

Show that the following are not vector spaces:
e. the unit circle in $\mathbf{R}^2$;
f. the set of all integers, $I = \{\ldots, -1, 0, +1, \ldots\}$;
g. the set of all nonnegative functions on $[a, b]$.

To discuss convergence in a vector space or in any other space, we need to have the notion of distance. The notion of distance in Euclidean space is generalized in the abstract notion of a *metric,* a function defined on any two elements in a set the value of which has an interpretation as the distance between them.

DEFINITION *A **metric space** is a set S, together with a metric (distance function)* $\rho: S \times S \to \mathbf{R}$, *such that for all* $x, y, z \in S$:
a. $\rho(x, y) \geq 0$, *with equality if and only if* $x = y$;
b. $\rho(x, y) = \rho(y, x)$; *and*
c. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

The definition of a metric thus abstracts the four basic properties of Euclidean distance: the distance between distinct points is strictly positive; the distance from a point to itself is zero; distance is symmetric; and the triangle inequality holds.

Exercise 3.3 Show that the following are metric spaces.
a. Let S be the set of integers, with $\rho(x, y) = |x - y|$.
b. Let S be the set of integers, with $\rho(x, y) = 0$ if $x = y$, 1 if $x \neq y$.
c. Let S be the set of all continuous, strictly increasing functions on $[a, b]$, with $\rho(x, y) = \max_{a \leq t \leq b} |x(t) - y(t)|$.
d. Let S be the set of all continuous, strictly increasing functions on $[a, b]$, with $\rho(x, y) = \int_a^b |x(t) - y(t)|dt$.
e. Let S be the set of all rational numbers, with $\rho(x, y) = |x - y|$.
f. Let $S = \mathbf{R}$, with $\rho(x, y) = f(|x - y|)$, where $f: \mathbf{R}_+ \to \mathbf{R}_+$ is continuous, strictly increasing, and strictly concave, with $f(0) = 0$.

For vector spaces, metrics are usually defined in such a way that the distance between any two points is equal to the distance of their difference from the zero point. That is, since for any points x and y in a vector space S, the point $x - y$ is also in S, the metric on a vector space is usually defined in such a way that $\rho(x, y) = \rho(x - y, \theta)$. To define such a metric, we need the concept of a norm.

DEFINITION *A **normed vector space** is a vector space S, together with a norm* $\|\cdot\|: S \to \mathbf{R}$, *such that for all* $x, y \in S$ *and* $\alpha \in \mathbf{R}$:
a. $\|x\| \geq 0$, *with equality if and only if* $x = \theta$;
b. $\|\alpha x\| = |\alpha| \cdot \|x\|$; *and*
c. $\|x + y\| \leq \|x\| + \|y\|$ *(the triangle inequality).*

Exercise 3.4 Show that the following are normed vector spaces.
a. Let $S = \mathbf{R}^l$, with $\|x\| = [\sum_{i=1}^l x_i^2]^{1/2}$ (Euclidean space).
b. Let $S = \mathbf{R}^l$, with $\|x\| = \max_i |x_i|$.
c. Let $S = \mathbf{R}^l$, with $\|x\| = \sum_{i=1}^l |x_i|$.
d. Let S be the set of all bounded infinite sequences $(x_1, x_2, \ldots)$, $x_k \in \mathbf{R}$, all k, with $\|x\| = \sup_k |x_k|$. (This space is called l_∞.)
e. Let S be the set of all continuous functions on $[a, b]$, with $\|x\| = \sup_{a \leq t \leq b} |x(t)|$. (This space is called $C[a, b]$.)
f. Let S be the set of all continuous functions on $[a, b]$, with $\|x\| = \int_a^b |x(t)|dt$.

It is standard to view any normed vector space $(S, \|\cdot\|)$ as a metric space, where the metric is taken to be $\rho(x, y) = \|x - y\|$, all $x, y \in S$.

The notion of convergence of a sequence of real numbers carries over without change to any metric space.

DEFINITION *A sequence $\{x_n\}_{n=0}^{\infty}$ in S* ***converges*** *to $x \in S$, if for each $\varepsilon > 0$, there exists N_ε such that*

$$(1) \qquad \rho(x_n, x) < \varepsilon, \quad \text{all } n \geq N_\varepsilon.$$

Thus a sequence $\{x_n\}$ in a metric space (S, ρ) converges to $x \in S$ if and only if the sequence of distances $\{\rho(x_n, x)\}$, a sequence in $\mathbf{R}_+$, converges to zero. In this case we write $x_n \to x$.

Verifying convergence directly involves having a "candidate" for the limit point x so that the inequality (1) can be checked. When a candidate is not immediately available, the following alternative criterion is often useful.

DEFINITION *A sequence $\{x_n\}_{n=0}^{\infty}$ in S is a* ***Cauchy sequence*** *(satisfies the* ***Cauchy criterion****) if for each $\varepsilon > 0$, there exists N_ε such that*

$$(2) \qquad \rho(x_n, x_m) < \varepsilon, \quad \text{all } n, m \geq N_\varepsilon.$$

Thus a sequence is Cauchy if the points get closer and closer to each other. The following exercise illustrates some basic facts about convergence and the Cauchy criterion.

Exercise 3.5 a. Show that if $x_n \to x$ and $x_n \to y$, then $x = y$. That is, if $\{x_n\}$ has a limit, then that limit is unique.

b. Show that if a sequence $\{x_n\}$ is convergent, then it satisfies the Cauchy criterion.

c. Show that if a sequence $\{x_n\}$ satisfies the Cauchy criterion, then it is bounded.

d. Show that $x_n \to x$ if and only if every subsequence of $\{x_n\}$ converges to x.

The advantage of the Cauchy criterion is that, in contrast to (1), (2) can be checked with knowledge of $\{x_n\}$ only. For the Cauchy criterion to be useful, however, we must work with spaces where it implies the existence of a limit point.

DEFINITION *A metric space (S, ρ) is* ***complete*** *if every Cauchy sequence in S converges to an element in S.*

In complete metric spaces, then, verifying that a sequence satisfies the Cauchy criterion is a way of verifying the existence of a limit point in S.

Verifying the completeness of particular spaces can take some work. We take as given the following

FACT *The set of real numbers* **R** *with the metric $\rho(x, y) = |x - y|$ is a complete metric space.*

Exercise 3.6 a. Show that the metric spaces in Exercises 3.3a,b and 3.4a–e are complete and that those in Exercises 3.3c–e and 3.4f are not. Show that the space in 3.3c is complete if "strictly increasing" is replaced with "nondecreasing."

b. Show that if (S, ρ) is a complete metric space and S' is a closed subset of S, then (S', ρ) is a complete metric space.

A complete normed vector space is called a ***Banach space.***

The next example is no more difficult than some of those in Exercise 3.6, but since it is important in what follows and illustrates clearly each of the steps involved in verifying completeness, we present the proof here.

THEOREM 3.1 *Let $X \subseteq \mathbf{R}^l$, and let $C(X)$ be the set of bounded continuous functions $f: X \to \mathbf{R}$ with the sup norm, $\|f\| = \sup_{x \in X} |f(x)|$. Then $C(X)$ is a complete normed vector space. (Note that if X is compact then every continuous function is bounded. Otherwise the restriction to bounded functions must be added.)*

Proof. That $C(X)$ is a normed vector space follows from Exercise 3.4e. Hence it suffices to show that if $\{f_n\}$ is a Cauchy sequence, there exists $f \in C(X)$ such that

$$\text{for any } \varepsilon > 0 \text{ there exists } N_\varepsilon \text{ such that } \|f_n - f\| \leq \varepsilon, \quad \text{all } n \geq N_\varepsilon.$$

Three steps are involved: to find a "candidate" function f; to show that $\{f_n\}$ converges to f in the sup norm; and to show that $f \in C(X)$ (that f is bounded and continuous). Each step involves its own entirely distinct logic.

Fix $x \in X$; then the sequence of real numbers $\{f_n(x)\}$ satisfies

$$|f_n(x) - f_m(x)| \leq \sup_{y \in X} |f_n(y) - f_m(y)| = \|f_n - f_m\|.$$

Therefore it satisfies the Cauchy criterion; and by the completeness of the real numbers, it converges to a limit point—call it $f(x)$. The limiting values define a function $f\colon X \to \mathbf{R}$ that we take to be our candidate.

Next we must show that $\|f_n - f\| \to 0$ as $n \to \infty$. Let $\varepsilon > 0$ be given and choose N_ε so that $n, m \geq N_\varepsilon$ implies $\|f_n - f_m\| \leq \varepsilon/2$. Since $\{f_n\}$ satisfies the Cauchy criterion, this can be done. Now for any fixed $x \in X$ and all $m \geq n \geq N_\varepsilon$,

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq \|f_n - f_m\| + |f_m(x) - f(x)| \\ &\leq \varepsilon/2 + |f_m(x) - f(x)|. \end{aligned}$$

Since $\{f_m(x)\}$ converges to $f(x)$, we can choose m separately for each fixed $x \in X$ so that $|f_m(x) - f(x)| \leq \varepsilon/2$. Since the choice of x was arbitrary, it follows that $\|f_n - f\| \leq \varepsilon$, all $n \geq N_\varepsilon$. Since $\varepsilon > 0$ was arbitrary, the desired result then follows.

Finally, we must show that f is bounded and continuous. Boundedness is obvious. To prove that f is continuous, we must show that for every $\varepsilon > 0$ and every $x \in X$, there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon \text{ if } \|x - y\|_E < \delta,$$

where $\|\cdot\|_E$ is the Euclidean norm on $\mathbf{R}^l$. Let ε and x be given. Choose k so that $\|f - f_k\| < \varepsilon/3$; since $f_n \to f$ (in the sup norm), such a choice is possible. Then choose δ so that

$$\|x - y\|_E < \delta \text{ implies } |f_k(x) - f_k(y)| < \varepsilon/3.$$

Since f_k is continuous, such a choice is possible. Then

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &\leq 2\|f - f_k\| + |f_k(x) - f_k(y)| \\ &< \varepsilon. \quad \blacksquare \end{aligned}$$

Although we have organized these component arguments into a theorem about a function space, each should be familiar to students of calculus. Convergence in the sup norm is simply uniform convergence. The proof above is then just an amalgam of the standard proofs that a sequence of functions that satisfies the Cauchy criterion uniformly converges uniformly and that uniform convergence "preserves continuity."

Exercise 3.7 a. Let $C^1[a, b]$ be the set of all continuously differentiable functions on $[a, b] = X \subset \mathbf{R}$, with the norm $\|f\| = \sup_{x \in X}\{|f(x)| + |f'(x)|\}$. Show that $C^1[a, b]$ is a Banach space. [*Hint.* Notice that

$$\sup_{x \in X}|f(x)| + \sup_{x \in X}|f'(x)| \geq \|f\| \geq \max\{\sup_{x \in X}|f(x)|, \sup_{x \in X}|f'(x)|\}.]$$

b. Show that this set of functions with the norm $\|f\| = \sup_{x \in X}|f(x)|$ is not complete. That is, give an example of a sequence of functions that is Cauchy in the given norm that does not converge to a function in the set. Is this sequence Cauchy in the norm of part (a)?

c. Let $C^k[a, b]$ be the set of all k times continuously differentiable functions on $[a, b] = X \subset \mathbf{R}$, with the norm $\|f\| = \sum_{i=0}^{k} \alpha_i \max_{x \in X}|f^i(x)|$, where $f^i = d^if(x)/dx^i$. Show that this space is complete if and only if $\alpha_i > 0$, $i = 0, 1, \ldots, k$.

3.2 The Contraction Mapping Theorem

In this section we prove two main results. The first is the Contraction Mapping Theorem, an extremely simple and powerful fixed point theorem. The second is a set of sufficient conditions, due to Blackwell, for establishing that certain operators are contraction mappings. The latter are useful in a wide variety of economic applications and will be drawn upon extensively in the next chapter.

We begin with the following definition.

DEFINITION *Let (S, ρ) be a metric space and $T: S \rightarrow S$ be a function mapping S into itself. T is a* **contraction mapping** *(with* **modulus** *β) if for some $\beta \in (0, 1)$, $\rho(Tx, Ty) \leq \beta\rho(x, y)$, for all $x, y \in S$.*

Perhaps the most familiar examples of contraction mappings are those on a closed interval $S = [a, b]$, with $\rho(x, y) = |x - y|$. Then $T: S \rightarrow S$ is a contraction if for some $\beta \in (0, 1)$.

$$\frac{|Tx - Ty|}{|x - y|} \leq \beta < 1, \quad \text{all } x, y \in S \text{ with } x \neq y.$$

That is, T is a contraction mapping if it is a function with slope uniformly less than one in absolute value.

Exercise 3.8 Show that if T is a contraction on S, then T is uniformly continuous on S.

The ***fixed points*** of T, the elements of S satisfying $Tx = x$, are the intersections of Tx with the 45° line, as shown in Figure 3.1. Hence it is clear that any contraction on this space has a unique fixed point. This conclusion is much more general.

THEOREM 3.2 (*Contraction Mapping Theorem*) *If (S, ρ) is a complete metric space and $T: S \rightarrow S$ is a contraction mapping with modulus β, then*

a. T has exactly one fixed point v in S, and

b. for any $v_0 \in S$, $\rho(T^n v_0, v) \leq \beta^n \rho(v_0, v)$, $n = 0, 1, 2, \ldots .$

Proof. To prove (a), we must find a candidate for v, show that it satisfies $Tv = v$, and show that no other element $\hat{v} \in S$ does.

Define the iterates of T, the mappings $\{T^n\}$, by $T^0 x = x$, and $T^n x = T(T^{n-1}x)$, $n = 1, 2, \ldots$. Choose $v_0 \in S$, and define $\{v_n\}_{n=0}^{\infty}$ by $v_{n+1} = Tv_n$, so that $v_n = T^n v_0$. By the contraction property of T,

$$\rho(v_2, v_1) = \rho(Tv_1, Tv_0) \leq \beta\rho(v_1, v_0).$$

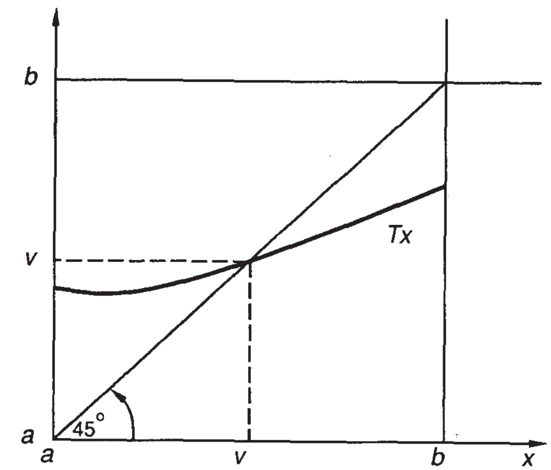

Figure 3.1

Continuing by induction, we get

$$\text{(1)} \qquad \rho(v_{n+1}, v_n) \leq \beta^n \rho(v_1, v_0), \quad n = 1, 2, \ldots .$$

Hence, for any $m > n$,

$$\begin{aligned} \rho(v_m, v_n) &\leq \rho(v_m, v_{m-1}) + \cdots + \rho(v_{n+2}, v_{n+1}) + \rho(v_{n+1}, v_n) \\ &\leq [\beta^{m-1} + \cdots + \beta^{n+1} + \beta^n]\, \rho(v_1, v_0) \\ &= \beta^n[\beta^{m-n-1} + \cdots + \beta + 1\, \rho(v_1, v_0) \\ \text{(2)} \qquad &\leq \frac{\beta^n}{1 - \beta}\, \rho(v_1, v_0), \end{aligned}$$

where the first line uses the triangle inequality and the second follows from (1). It is clear from (2) that $\{v_n\}$ is a Cauchy sequence. Since S is complete, it follows that $v_n \rightarrow v \in S$.

To show that $Tv = v$, note that for all n and all $v_0 \in S$,

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

We have demonstrated that both terms in the last expression converge to zero as $n \to \infty$; hence $\rho(Tv, v) = 0$, or $Tv = v$.

Finally, we must show that there is no other function $\hat{v} \in S$ satisfying $T\hat{v} = \hat{v}$. Suppose to the contrary that $\hat{v} \neq v$ is another solution. Then

$$0 < a = \rho(\hat{v}, v) = \rho(T\hat{v}, Tv) \leq \beta\rho(\hat{v}, v) = \beta a,$$

which cannot hold, since $\beta < 1$. This proves part (a).

To prove part (b), observe that for any $n \geq 1$

$$\rho(T^n v_0, v) = \rho[T(T^{n-1}v_0), Tv] \leq \beta\rho(T^{n-1}v_0, v),$$

so that (b) follows by induction. ■

Recall from Exercise 3.6b that if (S, ρ) is a complete metric space and S' is a closed subset of S, then (S', ρ) is also a complete metric space. Now suppose that $T: S \to S$ is a contraction mapping, and suppose further that T maps S' into itself, $T(S') \subseteq S'$ (where $T(S')$ denotes the image of S' under T). Then T is also a contraction mapping on S'. Hence the unique fixed point of T on S lies in S'. This observation is often useful for establishing qualitative properties of a fixed point. Specifically, in some situations we will want to apply the Contraction Mapping Theorem twice: once on a large space to establish uniqueness, and again on a smaller space to characterize the fixed point more precisely.

The following corollary formalizes this argument.

COROLLARY 1 *Let (S, ρ) be a complete metric space, and let $T: S \to S$ be a contraction mapping with fixed point $v \in S$. If S' is a closed subset of S and $T(S') \subseteq S'$, then $v \in S'$. If in addition $T(S') \subseteq S'' \subseteq S'$, then $v \in S''$.*

Proof. Choose $v_0 \in S'$, and note that $\{T^n v_0\}$ is a sequence in S' converging to v. Since S' is closed, it follows that $v \in S'$. If in addition $T(S') \subseteq S''$, then it follows that $v = Tv \in S''$. ■

Part (b) of the Contraction Mapping Theorem bounds the distance $\rho(T^n v_0, v)$ between the nth approximation and the fixed point in terms of the distance $\rho(v_0, v)$ between the initial approximation and the fixed point. However, if v is not known (as is the case if one is computing v), then neither is the magnitude of the bound. Exercise 3.9 gives a computationally useful inequality.

Exercise 3.9 Let (S, ρ), T, and v be as given above, let β be the modulus of T, and let $v_0 \in S$. Show that

$$\rho(T^n v_0, v) \leq \frac{1}{1-\beta}\rho(T^n v_0, T^{n+1}v_0).$$

The following result is a useful generalization of the Contraction Mapping Theorem.

COROLLARY 2 *(N-Stage Contraction Theorem) Let (S, ρ) be a complete metric space, let $T: S \to S$, and suppose that for some integer N, $T^N: S \to S$ is a contraction mapping with modulus β. Then*

a. T has exactly one fixed point in S, and

b. for any $v_0 \in S$, $\rho(T^{kN}v_0, v) \leq \beta^k\rho(v_0, v)$, $k = 0, 1, 2, \ldots$.

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

$$\rho(Tv, v) = \rho[T(T^N v), T^N v] = \rho[T^N(Tv), T^N v] \leq \beta\rho(Tv, v).$$

Since $\beta \in (0, 1)$, this implies that $\rho(Tv, v) = 0$, so v is a fixed point of T. To establish uniqueness, note that any fixed point of T is also a fixed point of T^N. Part (b) is established using the same argument as in the proof of Theorem 3.2. ■

The next exercise shows how the Contraction Mapping Theorem is used to prove existence and uniqueness of a solution to a differential equation.

Exercise 3.10 Consider the differential equation and boundary condition $dx(s)/ds = f[x(s)]$, all $s \geq 0$, with $x(0) = c \in \mathbf{R}$. Assume that $f: \mathbf{R} \to \mathbf{R}$ is continuous, and for some $B > 0$ satisfies the Lipschitz condition $|f(a) - f(b)| \leq B|a - b|$, all $a, b \in \mathbf{R}$. For any $t > 0$, consider $C[0, t]$, the space of bounded continuous functions on $[0, t]$, with the sup norm. Recall from Theorem 3.1 that this space is complete.

a. Show that the operator T defined by

$$(Tv)(s) = c + \int_0^s f[v(z)]dz, \quad 0 \leq s \leq t,$$

maps $C[0, t]$ into itself. That is, show that if v is bounded and continuous on $[0, t]$, then so is Tv.

b. Show that for some $\tau > 0$, T is a contraction on $C[0, \tau]$.

c. Show that the unique fixed point of T on $C[0, \tau]$ is a differentiable function, and hence that it is the unique solution on $[0, \tau]$ to the given differential equation.

Another useful route to verifying that certain operators are contractions is due to Blackwell.

THEOREM 3.3 *(Blackwell's sufficient conditions for a contraction) Let* $X \subseteq \mathbf{R}^l$, and let $B(X)$ *be a space of bounded functions* $f: X \to \mathbf{R}$, *with the sup norm. Let* $T: B(X) \to B(X)$ *be an operator satisfying*

a. *(monotonicity)* $f, g \in B(X)$ *and* $f(x) \leq g(x)$, *for all* $x \in X$, *implies* $(Tf)(x) \leq (Tg)(x)$, *for all* $x \in X$;

b. *(discounting) there exists some* $\beta \in (0, 1)$ *such that*

$$[T(f + a)](x) \leq (Tf)(x) + \beta a, \quad \text{all } f \in B(X),\ a \geq 0,\ x \in X.$$

[Here $(f + a)(x)$ *is the function defined by* $(f + a)(x) = f(x) + a$.*] Then* T *is a contraction with modulus* β.

Proof. If $f(x) \leq g(x)$ for all $x \in X$, we write $f \leq g$. For any $f, g \in B(X)$, $f \leq g + \|f - g\|$. Then properties (a) and (b) imply that

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

Reversing the roles of f and g gives by the same logic

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

Combining these two inequalities, we find that $\|Tf - Tg\| \leq \beta\|f - g\|$, as was to be shown. ■

In many economic applications the two hypotheses of Blackwell's theorem can be verified at a glance. For example, in the one-sector optimal growth problem, an operator T was defined by

$$(Tv)(k) = \max_{0 \leq y \leq f(k)} \{U[f(k) - y] + \beta v(y)\}.$$

If $v(y) \leq w(y)$ for all values of y, then the objective function for which Tw is the maximized value is uniformly higher than the function for which Tv is the maximized value; so the monotonicity hypothesis (a) is obvious. The discounting hypothesis (b) is equally easy, since

$$\begin{aligned} T(v + a)(k) &= \max_{0 \leq y \leq f(k)} \{U[f(k) - y] + \beta[v(y) + a]\} \\ &= \max_{0 \leq y \leq f(k)} \{U[f(k) - y] + \beta v(y)\} + \beta a \\ &= (Tv)(k) + \beta a. \end{aligned}$$

Blackwell's result will play a key role in our analysis of dynamic programs.

3.3 The Theorem of the Maximum

We will want to apply the Contraction Mapping Theorem to analyze dynamic programming problems that are much more general than the examples that have been discussed to this point. If x is the beginning-of-period state variable, an element of $X \subseteq \mathbf{R}^l$, and $y \in X$ is the end-of-period state to be chosen, we would like to let the current period return $F(x, y)$ and the set of feasible y values, given x, be specified as generally as possible. On the other hand, we want the operator T defined by

$$(Tv)(x) = \sup_y [F(x, y) + \beta v(y)]$$

$$\text{s.t. } y \text{ feasible given } x,$$

to take the space $C(X)$ of bounded continuous functions of the state into itself. We would also like to be able to characterize the set of maximizing values of y, given x.

To describe the feasible set, we use the idea of a ***correspondence*** from a set X into a set Y: a relation that assigns a set $\Gamma(x) \subseteq Y$ to each $x \in X$. In the case of interest here, $Y = X$. Hence we seek restrictions on the correspondence $\Gamma: X \to X$ describing the feasibility constraints and on the return function F, which together ensure that if $v \in C(X)$ and $(Tv)(x) =$

$\sup_{y \in \Gamma(x)}[F(x, y) + \beta v(y)]$ then $Tv \in C(X)$. Moreover, we wish to determine the implied properties of the correspondence $G(x)$ containing the maximizing values of y for each x. The main result in this section is the Theorem of the Maximum, which accomplishes both tasks.

Let $X \subseteq \mathbf{R}^l$; let $Y \subseteq \mathbf{R}^m$; let $f: X \times Y \to \mathbf{R}$ be a (single-valued) function; and let $\Gamma: X \to Y$ be a (nonempty, possibly multivalued) correspondence. Our interest is in problems of the form $\sup_{y \in \Gamma(x)} f(x, y)$. If for each x, $f(x, \cdot)$ is continuous in y and the set $\Gamma(x)$ is nonempty and compact, then for each x the maximum is attained. In this case the function

$$(1) \qquad h(x) = \max_{y \in \Gamma(x)} f(x, y)$$

is well defined, as is the nonempty set

$$(2) \qquad G(x) = \{y \in \Gamma(x): f(x, y) = h(x)\}$$

of y values that attain the maximum. In this section further restrictions on f and Γ will be added, to ensure that the function h and the set G vary in a continuous way with x.

There are several notions of continuity for correspondences, and each can be characterized in a variety of ways. For our purposes it is convenient to use definitions stated in terms of sequences.

DEFINITION *A correspondence $\Gamma: X \to Y$ is* ***lower hemi-continuous*** *(l.h.c.) at x if $\Gamma(x)$ is nonempty and if, for every $y \in \Gamma(x)$ and every sequence $x_n \to x$, there exists $N \geq 1$ and a sequence $\{y_n\}_{n=N}^{\infty}$ such that $y_n \to y$ and $y_n \in \Gamma(x_n)$, all $n \geq N$. [If $\Gamma(x')$ is nonempty for all $x' \in X$, then it is always possible to take $N = 1$.]*

DEFINITION *A compact-valued correspondence $\Gamma: X \to Y$ 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 $\{y_n\}$ such that $y_n \in \Gamma(x_n)$, all n, there exists a convergent subsequence of $\{y_n\}$ whose limit point y is in $\Gamma(x)$.*

Figure 3.2 displays a correspondence that is l.h.c. but not u.h.c. at x_1; is u.h.c. but not l.h.c. at x_2; and is both u.h.c. and l.h.c. at all other points. Note that our definition of u.h.c. applies only to correspondences that are compact-valued. Since all of the correspondences we will be dealing

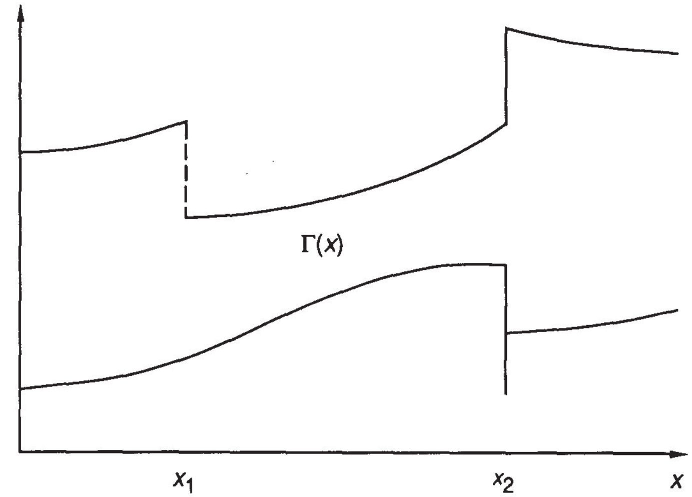

Figure 3.2

with satisfy this requirement, the restriction will not be binding. (A definition of u.h.c. for all correspondences is available, but it is stated in terms of images of open sets. For our purposes this definition is much less convenient, and its wider scope is never useful.)

DEFINITION *A correspondence $\Gamma: X \to Y$ is* ***continuous*** *at $x \in X$ if it is both u.h.c. and l.h.c. at x.*

A correspondence $\Gamma: X \to Y$ is called l.h.c., u.h.c., or continuous if it has that property at every point $x \in X$. The following exercises highlight some important facts about upper and lower hemi-continuity. Note that if $\Gamma: X \to Y$, then for any set $\hat{X} \subset X$, we define

$$\Gamma(\hat{X}) = \{y \in Y: y \in \Gamma(x), \text{ for some } x \in \hat{X}\}.$$

Exercise 3.11 a. Show that if Γ is single-valued and u.h.c., then it is continuous.

b. Let $\Gamma: \mathbf{R}^k \to \mathbf{R}^{l+m}$, and define $\phi: \mathbf{R}^k \to \mathbf{R}^l$ by

$$\phi(x) = \{y_1 \in \mathbf{R}^l: (y_1, y_2) \in \Gamma(x) \text{ for some } y_2 \in \mathbf{R}^m\}.$$

Show that if Γ is compact-valued and u.h.c., then so is ϕ.

c. Let ϕ: $X \to Y$ and ψ: $X \to Y$ be compact-valued and u.h.c., and define $\Gamma = \phi \cup \psi$ by

$$\Gamma(x) = \{y \in Y: y \in \phi(x) \cup \psi(x)\}, \quad \text{all } x \in X.$$

Show that Γ is compact-valued and u.h.c.

d. Let ϕ: $X \to Y$ and ψ: $X \to Y$ be compact-valued and u.h.c., and suppose that

$$\Gamma(x) = \{y \in Y: y \in \phi(x) \cap \psi(x)\} \neq \emptyset, \quad \text{all } x \in X.$$

Show that Γ is compact-valued and u.h.c.

e. Show that if ϕ: $X \to Y$ and ψ: $Y \to Z$ are compact-valued and u.h.c., then the correspondence $\psi \circ \phi = \Gamma$: $X \to Z$ defined by

$$\Gamma(x) = \{z \in Z: z \in \psi(y), \text{ for some } y \in \phi(x)\}$$

is also compact-valued and u.h.c.

f. Let Γ_i: $X \to Y_i$, $i = 1, \ldots, k$, be compact-valued and u.h.c. Show that Γ: $X \to Y = Y_1 \times \ldots \times Y_k$ defined by

$$\Gamma(x) = \{y \in Y: y = (y_1, \ldots, y_k), \text{ where } y_i \in \Gamma_i(x), i = 1, \ldots, k\},$$

is also compact-valued and u.h.c.

g. Show that if Γ: $X \to Y$ is compact-valued and u.h.c., then for any compact set $K \subseteq X$, the set $\Gamma(K) \subseteq Y$ is also compact. [*Hint.* To show that $\Gamma(K)$ is bounded, suppose the contrary. Let $\{y_n\}$ be a divergent sequence in $\Gamma(K)$, and choose $\{x_n\}$ such that $y_n \in \Gamma(x_n)$, all n.]

Exercise 3.12 a. Show that if Γ is single-valued and l.h.c., then it is continuous.

b. Let Γ: $\mathbf{R}^k \to \mathbf{R}^{l+m}$, and define ϕ: $\mathbf{R}^k \to \mathbf{R}^l$ by

$$\phi(x) = \{y_1 \in \mathbf{R}^l: (y_1, y_2) \in \Gamma(x), \text{ for some } y_2 \in \mathbf{R}^m\}.$$

Show that if Γ is l.h.c., then so is ϕ.

c. Let ϕ: $X \to Y$ and ψ: $X \to Y$ be l.h.c., and define $\Gamma = \phi \cup \psi$ by

$$\Gamma(x) = \{y \in Y: y \in \phi(x) \cup \psi(x)\}, \quad \text{all } x \in X.$$

Show that Γ is l.h.c.

d. Let ϕ: $X \to Y$ and ψ: $X \to Y$ be l.h.c., and suppose that

$$\Gamma(x) = \{y \in Y: y \in \phi(x) \cap \psi(x)\} \neq \emptyset, \quad \text{all } x \in X.$$

Show by example that Γ need not be l.h.c. Show that if ϕ and ψ are both convex-valued, and if int $\phi(x) \cap$ int $\psi(x) \neq \emptyset$, then Γ is l.h.c. at x.

e. Show that if ϕ: $X \to Y$ and ψ: $Y \to Z$ are l.h.c., then the correspondence $\psi \circ \phi = \Gamma$: $X \to Z$ defined by

$$\Gamma(x) = \{z \in Z: z \in \psi(y), \text{ for some } y \in \phi(x)\}$$

is also l.h.c.

f. Let Γ_i: $X \to Y_i$, $i = 1, \ldots, k$, be l.h.c. Show that Γ: $X \to Y = Y_1 \times \ldots \times Y_k$ defined by

$$\Gamma(x) = \{y \in Y: y = (y_1, \ldots, y_k), \text{ where } y_i \in \Gamma_i(x), i = 1, \ldots, k\}$$

is l.h.c.

The next two exercises show some of the relationships between constraints stated in terms of inequalities involving continuous functions and those stated in terms of continuous correspondences. These relationships are extremely important for many problems in economics, where constraints are often stated in terms of production functions, budget constraints, and so on.

Exercise 3.13 a. Let Γ: $\mathbf{R}_+ \to \mathbf{R}_+$ be defined by $\Gamma(x) = [0, x]$. Show that Γ is continuous.

b. Let f: $\mathbf{R}^l_+ \to \mathbf{R}_+$ be a continuous function, and define the correspondence Γ: $\mathbf{R}^l_+ \to \mathbf{R}_+$ by $\Gamma(x) = [0, f(x)]$. Show that Γ is continuous.

c. Let f_i: $\mathbf{R}^l_+ \times \mathbf{R}^m \to \mathbf{R}_+$, $i = 1, \ldots, l$, be continuous functions. Define Γ: $\mathbf{R}^l_+ \times \mathbf{R}^m \to \mathbf{R}^l_+$ by

$$\Gamma(x, z) = \left\{y \in \mathbf{R}^l_+: 0 \le y_i \le f_i(x^i, z), i = 1, \ldots, l; \text{ and } \sum_{i=1}^{l} x^i \le x\right\}.$$

Show that Γ is continuous.

Exercise 3.14 a. Let $H(x, y)$: $\mathbf{R}^l_+ \times \mathbf{R}^m_+ \to \mathbf{R}$ be continuous, strictly increasing in its first l arguments, strictly decreasing in its last m arguments, with $H(0, 0) = 0$. Define Γ: $\mathbf{R}^l \to \mathbf{R}^m$ by $\Gamma(x) = \{y \in \mathbf{R}^m: H(x, y) \geq 0\}$. Show that if $\Gamma(x)$ is compact-valued, then Γ is continuous at x.

b. Let $H(x, y)$: $\mathbf{R}^l \times \mathbf{R}^m \to \mathbf{R}$ be continuous and concave, and define Γ as in part (a). Show that if $\Gamma(x)$ is compact-valued and there exists some $\hat{y} \in \Gamma(x)$ such that $H(x, \hat{y}) > 0$, then Γ is continuous at x.

c. Define H: $\mathbf{R} \times \mathbf{R} \to \mathbf{R}$ by $H(x, y) = 1 - \max\{|x|, |y|\}$, and define $\Gamma(x)$ as in part (a). Where does Γ fail to be l.h.c.?

When trying to establish properties of a correspondence Γ: $X \to Y$, it is sometimes useful to deal with its ***graph***, the set

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

The next two results provide conditions on A that are sufficient to ensure the upper and lower hemi-continuity respectively of Γ.

THEOREM 3.4 *Let Γ: $X \to Y$ be a nonempty-valued correspondence, and let A be the graph of Γ. Suppose that A is closed, and that for any bounded set $\hat{X} \subseteq X$, the set $\Gamma(\hat{X})$ is bounded. Then Γ is compact-valued and u.h.c.*

Proof. For each $x \in X$, $\Gamma(x)$ is closed (since A is closed) and is bounded (by hypothesis). Hence Γ is compact-valued.

Let $\hat{x} \in X$, and let $\{x_n\} \subseteq X$ with $x_n \to \hat{x}$. Since Γ is nonempty-valued, we can choose $y_n \in \Gamma(x_n)$, all n. Since $x_n \to \hat{x}$, there is a bounded set $\hat{X} \subset X$ containing $\{x_n\}$ and $\hat{x}$. Then by hypothesis $\Gamma(\hat{X})$ is bounded. Hence $\{y_n\} \subset \Gamma(\hat{X})$ has a convergent subsequence, call it $\{y_{n_k}\}$; let $\hat{y}$ be the limit point of this subsequence. Then $\{(x_{n_k}, y_{n_k})\}$ is a sequence 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 Γ is u.h.c. at $\hat{x}$. Since $\hat{x}$ was arbitrary, this establishes the desired result. ∎

To see why the hypothesis of boundedness is required in Theorem 3.4, consider the correspondence Γ: $\mathbf{R}_+ \to \mathbf{R}_+$ defined by

$$\Gamma(0) = 0, \quad \text{and} \quad \Gamma(x) = \{0, 1/x\}, \quad \text{all } x > 0.$$

The graph of Γ is closed, but Γ is not u.h.c. at $x = 0$.

The next exercise is a kind of converse to Theorem 3.4.

Exercise 3.15 Let Γ: $X \to Y$ be a compact-valued u.h.c. correspondence with graph A. Show that if X is compact then A is compact.

The next theorem deals with lower hemi-continuity. For any $x \in \mathbf{R}^l$ and any $\varepsilon > 0$, let $B(x, \varepsilon)$ denote the closed ball of radius ε about x.

THEOREM 3.5 *Let Γ: $X \to Y$ be a nonempty-valued correspondence, and let A be the graph of Γ. Suppose that A is convex; and that for any bounded set $\hat{X} \subseteq X$, there is a bounded set $\hat{Y} \subseteq Y$ such that $\Gamma(x) \cap \hat{Y} \neq \emptyset$, all $x \in \hat{X}$. Then Γ is l.h.c. at every interior point of X.*

Proof. Choose $\hat{x} \in \mathring{X}$; $\hat{y} \in \Gamma(\hat{x})$; and $\{x_n\} \subset X$ with $x_n \to \hat{x}$. Choose $\varepsilon > 0$ such that the set $\hat{X} = B(\hat{x}, \varepsilon) \subseteq X$. Note that for some $N \geq 1$, $x_n \in \hat{X}$, all $n \geq N$; without loss of generality we take $N = 1$.

Let D denote the boundary of the set $\hat{X}$. 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 $x_n \to \hat{x}$, so $\alpha_n \to 0$. Choose $\hat{Y}$ such that $\Gamma(x) \cap \hat{Y} \neq \emptyset$, all $x \in \hat{X}$. Then for each n, choose $\hat{y}_n \in \Gamma(d_n) \cap \hat{Y}$, and define

$$y_n = \alpha_n \hat{y}_n + (1 - \alpha_n)\hat{y}, \quad \text{all } n.$$

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

To see why $\hat{x}$ must be an interior point, consider the case where X is a disk and A is an inverted cone that is slanted so the tip is directly above the boundary of X. Let $\hat{x}$ be the point below the tip of the cone, and take a sequence $\{x_n\}$ along the boundary of the disk. Then each set $\Gamma(x_n)$ is a single point, but $\Gamma(\hat{x})$ is an interval.

We are now ready to answer the questions: Under what conditions do the function $h(x)$ defined by the maximization problem in (1) and the associated set of maximizing y values $G(x)$ defined in (2) vary continuously with x? An answer is provided in the following theorem.

THEOREM 3.6 *(Theorem of the Maximum) Let $X \subseteq \mathbf{R}^l$ and $Y \subseteq \mathbf{R}^m$, let $f: X \times Y \to \mathbf{R}$ be a continuous function, and let $\Gamma: X \to Y$ be a compact-valued and continuous correspondence. Then the function $h: X \to \mathbf{R}$ defined in (1) is continuous, and the correspondence $G: X \to Y$ defined in (2) is nonempty, compact-valued, and u.h.c.*

Proof. Fix $x \in X$. The set $\Gamma(x)$ is nonempty and compact, and $f(x, \cdot)$ is continuous; hence the maximum in (1) is attained, and the set $G(x)$ of maximizers is nonempty. Moreover, since $G(x) \subseteq \Gamma(x)$ and $\Gamma(x)$ is compact, it follows that $G(x)$ is bounded. Suppose $y_n \to y$, and $y_n \in G(x)$, all n. Since $\Gamma(x)$ is closed, $y \in \Gamma(x)$. Also, since $h(x) = f(x, y_n)$, all 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 will show that $G(x)$ is u.h.c. Fix x, and let $\{x_n\}$ be any sequence converging to x. Choose $y_n \in G(x_n)$, all n. Since Γ is u.h.c., there exists a subsequence $\{y_{n_k}\}$ converging to $y \in \Gamma(x)$. Let $z \in \Gamma(x)$. Since Γ is l.h.c., there exists a sequence $z_{n_k} \to z$, with $z_{n_k} \in \Gamma(x_{n_k})$, all k. Since $f(x_{n_k}, y_{n_k}) \geq f(x_{n_k}, z_{n_k})$, all k, and f is continuous, it follows that $f(x, y) \geq f(x, z)$. Since this holds for any $z \in \Gamma(x)$, it follows that $y \in G(x)$. Hence G is u.h.c.

Finally, we will show that h is continuous. Fix x, and let $\{x_n\}$ be any sequence converging to x. Choose $y_n \in G(x_n)$, all n. Let $\bar{h} = \lim\sup h(x_n)$ and $\underline{h} = \lim\inf h(x_n)$. 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 argument establishes that $h(x) = \underline{h}$. Hence $\{h(x_n)\}$ converges, and its limit is $h(x)$. ■

The following exercise illustrates through concrete examples what this theorem does and does not say.

Exercise 3.16 a. Let $X = \mathbf{R}$, and let $\Gamma(x) = Y = [-1, +1]$, all $x \in X$. Define $f: X \times Y \to \mathbf{R}$ by $f(x, y) = xy^2$. Graph $G(x)$; show that $G(x)$ is u.h.c. but not l.h.c. at $x = 0$.

b. Let $x = \mathbf{R}$, and let $\Gamma(x) = [0, 4]$, all $x \in X$. Define

$$f(x, y) = \max\{2 - (y - 1)^2, x + 1 - (y - 2)^2\}.$$

Graph $G(x)$ and show that it is u.h.c. Exactly where does it fail to be l.h.c.?

c. Let $X = \mathbf{R}_+$, $\Gamma(x) = \{y \in \mathbf{R}: -x \leq y \leq x\}$, and $f(x, y) = \cos(y)$. Graph $G(x)$ and show that it is u.h.c. Exactly where does it fail to be l.h.c.?

Suppose that in addition to the hypotheses of the Theorem of the Maximum the correspondence Γ is convex-valued and the function f is strictly concave in y. Then G is single-valued, and by Exercise 3.11a it is a continuous function—call it g. The next two results establish properties of g. Lemma 3.7 shows that if $f(x, y)$ is close to the maximized value $f[x, g(x)]$, then y is close to $g(x)$. Theorem 3.8 draws on this result to show that if $\{f_n\}$ is a sequence of continuous functions, each strictly concave in y, converging uniformly to f, then the sequence of maximizing functions $\{g_n\}$ converges pointwise to g. The latter convergence is uniform if X is compact.

LEMMA 3.7 *Let $X \subseteq \mathbf{R}^l$ and $Y \subseteq \mathbf{R}^m$. Assume that the correspondence $\Gamma: X \to Y$ is nonempty, compact- and convex-valued, and continuous, and let A be the graph of Γ. Assume that the function $f: A \to \mathbf{R}$ is continuous and that $f(x, \cdot)$ is strictly concave, for each $x \in X$. Define the function $g: X \to Y$* by

$$g(x) = \underset{y \in \Gamma(x)}{\operatorname{argmax}} f(x, y).$$

Then for each $\varepsilon > 0$ and $x \in X$, there exists $\delta_x > 0$ such that

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

If 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 (single-valued) function. We first prove the claim for the case where X is compact. Note that in this case A is a compact set by Exercise 3.15. For each $\varepsilon > 0$, define

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

If $A_\varepsilon = \emptyset$, all $\varepsilon > 0$, then Γ 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_ε is nonempty and compact. For any such ε, let

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

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

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

as was to be shown.

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

THEOREM 3.8 *Let X, Y, Γ, and A be as defined in Lemma 3.7. Let $\{f_n\}$ be a sequence of continuous (real-valued) functions on A; assume that for each n and each $x \in X$, $f_n(x, \cdot)$ 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*

$$g_n(x) = \operatorname*{argmax}_{y \in \Gamma(x)} f_n(x, y), \quad n = 1, 2, \ldots, \text{ and}$$

$$g(x) = \operatorname*{argmax}_{y \in \Gamma(x)} f(x, y).$$

Then $g_n \to g$ pointwise. If X is compact, $g_n \to g$ uniformly.

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

$$\begin{aligned} 0 &\leq f[x, g(x)] - f[x, g_n(x)] \\ &\leq f[x, g(x)] - f_n[x, g(x)] + f_n[x, g_n(x)] - f[x, g_n(x)] \\ &\leq 2\|f - f_n\|, \quad \text{all } x \in X. \end{aligned}$$

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

$$0 \leq f[x, g(x)] - f[x, g_n(x)] \leq 2\|f - f_n\| < \delta, \quad \text{all } x \in X, \text{ all } n \geq M_\delta. \tag{3}$$

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

$$\|g(x) - g_n(x)\| < \varepsilon, \quad \text{all } n \geq N_x. \tag{4}$$

By Lemma 3.7, it suffices to show that for any $\delta_x > 0$ and $x \in X$ there exists $N_x \geq 1$ such that

$$|f[x, g(x)] - f[x, g_n(x)]| < \delta_x, \quad \text{all } n \geq N_x. \tag{5}$$

From (3), it follows that any $N_x \geq M_{\delta_x}$ has the required property.

Suppose X is compact. To establish that $g_n \to g$ uniformly, we must show that for each $\varepsilon > 0$ there exists $N \geq 1$ such that (4) holds for all $x \in X$. By Lemma 3.7, it suffices to show that for any $\delta > 0$, there exists $N \geq 1$, such that (5) holds for all $x \in X$. From (3) it follows that any $N \geq M_\delta$ has the required property. ∎

3.4 Bibliographic Notes

For a more detailed discussion of metric spaces, see Kolmogorov and Fomin (1970, chap. 2) or Royden (1968, chap. 7). Good discussions of normed vector spaces can be found in Kolmogorov and Fomin (1970, chap. 4) and Luenberger (1969, chap. 2), both of which also treat the Contraction Mapping Theorem. Blackwell's sufficient condition is Theorem 5 in Blackwell (1965). The Theorem of the Maximum dates from Berge (1963, chap. 6), and can also be found in Hildenbrand (1974, pt. I.B). Both of these also contain excellent treatments of upper and lower hemi-continuity. We are grateful to David Levine and Michael Jansson for pointing out that the argument in Theorem 3.5 applies only on the interior of the set.

4 Dynamic Programming under Certainty

Posed in terms of infinite sequences, the problems we are interested in are of the form

$$\text{(SP)} \qquad \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})$$

$$\text{s.t.} \quad x_{t+1} \in \Gamma(x_t), \quad t = 0, 1, 2, \ldots,$$

$$x_0 \in X \text{ given.}$$

Corresponding to any such problem, we have a functional equation of the form

$$\text{(FE)} \qquad v(x) = \sup_{y \in \Gamma(x)} [F(x, y) + \beta v(y)], \quad \text{all } x \in X.$$

In this chapter we establish the relationship between solutions to these two problems and develop methods for analyzing the latter.

Exercise 4.1 a. Show that the one-sector growth model discussed at the beginning of Chapter 3 can be expressed as in (SP).

b. Show that the many-sector growth model

$$\sup_{\{(c_t, k_{t+1})\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t)$$

$$\text{s.t.} \quad (k_{t+1} + c_t, k_t) \in Y, \quad t = 0, 1, 2, \ldots,$$

$$\text{given } k_0 \in \mathbf{R}^l_+,$$

where $Y \subset \mathbf{R}^{2l}_+$ is a fixed production set, can also be written this way.

As we hinted in the last chapter and will show in this one, some very powerful—and relatively simple—mathematical tools can be used to study the functional equation (FE). To take advantage of these, however, we must show that solutions to (FE) correspond to solutions to the sequence problem (SP). In Section 4.1 we rigorously establish the connections between solutions to these two problems, connections that Richard Bellman called the *Principle of Optimality*. Section 4.2 then develops the main results of the chapter: existence, uniqueness, and characterization theorems for solutions to (FE) under the assumption that the return function F is bounded. The case where F displays constant returns to scale is treated in Section 4.3, and the case where F is an arbitrary unbounded return function in Section 4.4. Section 4.5 treats the relationship between the dynamic programming approach to optimization over time and the classical (variational) approach. Section 4.6 contains references for further discussion of some of the mathematical and economic ideas. In Chapter 5 we illustrate how the methods developed in Sections 4.2–4.4 can be applied to a wide variety of economic problems.

4.1 The Principle of Optimality

In this section we study the relationship between solutions to the problems (SP) and (FE). (Note that "sup" has been used instead of "max" in both, so that we can ignore—for the moment—the question of whether the optimum is attained.) The general idea, of course, is that the solution v to (FE), evaluated at x_0, gives the value of the supremum in (SP) when the initial state is x_0 and that a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ attains the supremum in (SP) if and only if it satisfies

$$\text{(1)} \qquad v(x_t) = F(x_t, x_{t+1}) + \beta v(x_{t+1}), \quad t = 0, 1, 2, \ldots.$$

Richard Bellman called these ideas the Principle of Optimality. Intuitive as it is, the Principle requires proof. Spelling out precisely the conditions under which it holds is our task in this section.

The main results are Theorem 4.2, establishing that the supremum function v^* for the sequence problem (SP) satisfies the functional equation (FE), and Theorem 4.3, establishing a partial converse. The "partial" nature of the converse arises from the fact that a boundedness condition must be imposed. Theorems 4.4 and 4.5 then deal with the characterization of optimal policies. Theorem 4.4 shows that if $\{x_{t+1}\}_{t=0}^{\infty}$ is a sequence attaining the supremum in (SP), then it satisfies (1) for $v = v^*$. Conversely, Theorem 4.5 establishes that any sequence $\{x_{t+1}\}_{t=0}^{\infty}$ that satisfies (1) for $v = v^*$, and also satisfies a boundedness condition, attains the supremum in (SP). The four theorems taken together thus establish conditions under which solutions to (SP) and to (FE) coincide exactly, and optimal policies are those that satisfy (1).

To begin we must establish some notation. Let X be the set of possible values for the state variable x. In this section we will not need to impose any restrictions on the set X. It may be a subset of a Euclidean space, a set of functions, a set of probability distributions, or any other set. Let $\Gamma: X \to X$ be the correspondence describing the feasibility constraints. That is, for each $x \in X$, $\Gamma(x)$ is the set of feasible values for the state variable next period if the current state is x. Let A be the graph of Γ:

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

Let the real-valued function $F: A \to \mathbf{R}$ be the one-period return function, and let $\beta \geq 0$ be the (stationary) discount factor. Thus the "givens" for the problem are X, Γ, F, and β.

First we must establish conditions under which the problem (SP) is well defined. That is, we must find conditions under which the feasible set is nonempty and the objective function is well defined for every point in the feasible set.

Call any sequence $\{x_t\}_{t=0}^{\infty}$ in X a *plan*. Given $x_0 \in X$, let

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

be the set of plans that are *feasible from* x_0. That is, $\Pi(x_0)$ is the set of all sequences $\{x_t\}$ satisfying the constraints in (SP). Let $\underline{x} = (x_0, x_1, \ldots)$ denote a typical element of $\Pi(x_0)$. The following assumption ensures that $\Pi(x_0)$ is nonempty, for all $x_0 \in X$.

ASSUMPTION 4.1 *$\Gamma(x)$ is nonempty, for all $x \in X$.*

The only additional restriction on X, Γ, F, and β we will need in this section is a requirement that all feasible plans can be evaluated using the objective function F and the discount rate β.

ASSUMPTION 4.2 *For all $x_0 \in X$ and $\underline{x} \in \Pi(x_0)$, $\lim_{n\to\infty}\sum_{t=0}^{n}\beta^t F(x_t, x_{t+1})$ exists (although it may be plus or minus infinity).*

There are a variety of ways of ensuring that Assumption 4.2 holds. Clearly it is satisfied if the function F is bounded and $0 < \beta < 1$. Alternatively, for any $(x, y) \in A$, let

$$F^+(x, y) = \max\{0, F(x, y)\} \quad \text{and} \quad F^-(x, y) = \max\{0, -F(x, y)\}.$$

Then Assumption 4.2 holds if for each $x_0 \in X$ and $\underline{x} \in \Pi(x_0)$, either

$$\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F^+(x_t, x_{t+1}) < +\infty, \quad \text{or}$$

$$\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F^-(x_t, x_{t+1}) < +\infty, \quad \text{or both.}$$

Thus a sufficient condition for Assumptions 4.1–4.2 is that F be bounded above or below and $0 < \beta < 1$. Another sufficient condition is that for each $x_0 \in X$ and $\underline{x} \in \Pi(x_0)$, there exist $\theta \in (0, \beta^{-1})$ and $0 < c < \infty$ such that

$$F(x_t, x_{t+1}) \leq c\theta^t, \quad \text{all } t.$$

The following exercise provides a way of verifying that the latter holds.

Exercise 4.2 a. Show that Assumption 4.2 is satisfied if $X = \mathbf{R}^l_+$; $0 < \beta < 1$; there exists $0 < \theta < 1/\beta$ such that $y \in \Gamma(x)$ implies $\|y\| \leq \theta\|x\|$; $F(0, 0) = 0$; F is increasing in its first l arguments and decreasing in its last l arguments; F is concave in its first l arguments; and $0 \in \Gamma(x)$, all x.

b. Show that Assumption 4.2 is satisfied if $X = \mathbf{R}^l_+$; $0 < \beta < 1$; there exists $0 < \theta < 1/\beta$ such that $y \in \Gamma(x)$ implies $F(y, 0) \leq \theta F(x, 0)$; F is increasing in its first l arguments and decreasing in its last l arguments; and $0 \in \Gamma(x)$, all x.

For each $n = 0, 1, \ldots,$ define $u_n: \Pi(x_0) \to \mathbf{R}$ by

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

Then $u_n(\underline{x})$ is the partial sum of the (discounted) returns in periods 0 through n from the feasible plan $\underset{\sim}{x}$. Under Assumption 4.2 we can also define u: $\Pi(x_0) \to \overline{\mathbf{R}}$ by

$$u(\underset{\sim}{x}) = \lim_{n\to\infty} u_n(\underset{\sim}{x}),$$

where $\overline{\mathbf{R}} = \mathbf{R} \cup \{+\infty, -\infty\}$ is the set of extended real numbers. Thus $u(\underset{\sim}{x})$ is the (infinite) sum of discounted returns from the feasible sequence $\underset{\sim}{x}$.

If Assumptions 4.1 and 4.2 both hold, then the set of feasible plans $\Pi(x_0)$ is nonempty for each $x_0 \in X$, and the objective function in (SP) is well defined for every plan $\underset{\sim}{x} \in \Pi(x_0)$. We can then define the *supremum function* v^*: $X \to \overline{\mathbf{R}}$ by

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

Thus $v^*(x_0)$ is the supremum in (SP). Note that it follows by definition that v^* is the unique function satisfying the following three conditions:

a. if $|v^*(x_0)| < \infty$, then

$$v^*(x_0) \geq u(\underset{\sim}{x}), \quad \text{all } \underset{\sim}{x} \in \Pi(x_0); \tag{2}$$

and for any $\varepsilon > 0$,

$$v^*(x_0) \leq u(\underset{\sim}{x}) + \varepsilon, \quad \text{some } \underset{\sim}{x} \in \Pi(x_0); \tag{3}$$

b. if $v^*(x_0) = +\infty$, then there exists a sequence $\{\underset{\sim}{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\underset{\sim}{x}^k) = +\infty$; and

c. if $v^*(x_0) = -\infty$, then $u(\underset{\sim}{x}) = -\infty$, for all $\underset{\sim}{x} \in \Pi(x_0)$.

Our interest is in the connections between the supremum function v^* and solutions v to the functional equation (FE). In interpreting the next results, it is important to remember that v^* is always uniquely defined (provided Assumptions 4.1–4.2 hold), whereas (FE) may—for all we know so far—have zero, one, or many solutions.

We will say that v^* satisfies the functional equation if three conditions hold:

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

$$v^*(x_0) \geq F(x_0, y) + \beta v^*(y), \quad \text{all } y \in \Gamma(x_0), \tag{4}$$

and for any $\varepsilon > 0$,

$$v^*(x_0) \leq F(x_0, y) + \beta v^*(y) + \varepsilon, \quad \text{some } y \in \Gamma(x_0); \tag{5}$$

b. if $v^*(x_0) = +\infty$, then there exists a sequence $\{y^k\}$ in $\Gamma(x_0)$ such that

$$\lim_{k\to\infty} [F(x_0, y^k) + \beta v^*(y^k)] = +\infty; \tag{6}$$

c. if $v^*(x_0) = -\infty$, then

$$F(x_0, y) + \beta v^*(y) = -\infty, \quad \text{all } y \in \Gamma(x_0). \tag{7}$$

Before we prove that the supremum function v^* satisfies the functional equation, it is useful to establish a preliminary result.

LEMMA 4.1 *Let X, Γ, F, and β satisfy Assumption 4.2. Then for any $x_0 \in X$ and any $(x_0, x_1, \ldots) = \underset{\sim}{x} \in \Pi(x_0)$,*

$$u(\underset{\sim}{x}) = F(x_0, x_1) + \beta u(\underset{\sim}{x}'),$$

where $\underset{\sim}{x}' = (x_1, x_2, \ldots)$.

Proof. Under Assumption 4.2, for any $x_0 \in X$ and any $\underset{\sim}{x} \in \Pi(x_0)$,

$$\begin{aligned} u(\underset{\sim}{x}) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \\ &= F(x_0, x_1) + \beta \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_{t+1}, x_{t+2}) \\ &= F(x_0, x_1) + \beta u(\underset{\sim}{x}'). \quad \blacksquare \end{aligned}$$

THEOREM 4.2 *Let X, Γ, F, and β satisfy Assumptions 4.1–4.2. Then the function v^* satisfies (FE).*

Proof. If $\beta = 0$, the result is trivial. Suppose that $\beta > 0$, and choose $x_0 \in X$.

Suppose $v^*(x_0)$ is finite. Then (2) and (3) hold, and it is sufficient to show that this implies (4) and (5) hold. To establish (4), let $x_1 \in \Gamma(x_0)$ and $\varepsilon > 0$ be given. Then by (3) there exists $\underset{\sim}{x}' = (x_1, x_2, \ldots) \in \Pi(x_1)$ such that $u(\underset{\sim}{x}') \geq v^*(x_1) - \varepsilon$. Note, too, that $\underset{\sim}{x} = (x_0, x_1, x_2, \ldots) \in \Pi(x_0)$. Hence it follows from (2) and Lemma 4.1 that

$$v^*(x_0) \ge u(\underset{\sim}{x}) = F(x_0, x_1) + \beta u(\underset{\sim}{x}') \ge F(x_0, x_1) + \beta v^*(x_1) - \beta\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, (4) follows.

To establish (5), choose $x_0 \in X$ and $\varepsilon > 0$. From (3) and Lemma 4.1, it follows that one can choose $\underset{\sim}{x} = (x_0, x_1, \ldots) \in \Pi(x_0)$, so that

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

where $\underset{\sim}{x}' = (x_1, x_2, \ldots)$. It then follows from (2) that

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

Since $x_1 \in \Gamma(x_0)$, this establishes (5).

If $v^*(x_0) = +\infty$, then there exists a sequence $\{\underset{\sim}{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\underset{\sim}{x}^k) = +\infty$. Since $x_1^k \in \Gamma(x_0)$, all k, and

$$u(\underset{\sim}{x}^k) = F(x_0, x_1^k) + \beta u(\underset{\sim}{x}'^k) \le F(x_0, x_1^k) + \beta v^*(x_1^k), \quad \text{all } k,$$

it follows that (6) holds for the sequence $\{y^k = x_1^k\}$ in $\Gamma(x_0)$.

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

$$u(\underset{\sim}{x}) = F(x_0, x_1) + \beta u(\underset{\sim}{x}') = -\infty, \quad \text{all } (x_0, x_1, x_2, \ldots) = \underset{\sim}{x} \in \Pi(x_0),$$

where $\underset{\sim}{x}' = (x_1, x_2, \ldots)$. Since F is real-valued (it does *not* take on the values $-\infty$ or $+\infty$), it follows that

$$u(\underset{\sim}{x}') = -\infty, \quad \text{all } x_1 \in \Gamma(x_0), \text{ all } \underset{\sim}{x}' \in \Pi(x_1).$$

Hence $v^*(x_1) = -\infty$, all $x_1 \in \Gamma(x_0)$. Since F is real-valued and $\beta > 0$, (7) follows immediately. ■

The next theorem provides a partial converse to Theorem 4.2. It shows that v^* is the only solution to the functional equation that satisfies a certain boundedness condition.

THEOREM 4.3 *Let X, Γ, F, and β satisfy Assumptions 4.1–4.2. If v is a solution to (FE) and satisfies*

(8) $$\lim_{n\to\infty} \beta^n v(x_n) = 0, \quad \text{all } (x_0, x_1, \ldots) \in \Pi(x_0), \text{ all } x_0 \in X,$$

then $v = v^*$.

Proof. If $v(x_0)$ is finite, then (4) and (5) hold, and it suffices to show that this implies (2) and (3) hold. Now (4) implies that for all $\underset{\sim}{x} \in \Pi(x_0)$,

$$\begin{aligned} v(x_0) &\ge F(x_0, x_1) + \beta v(x_1) \\ &\ge F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 v(x_2) \\ &\;\;\vdots \\ &\ge u_n(\underset{\sim}{x}) + \beta^{n+1} v(x_{n+1}), \quad n = 1, 2, \ldots. \end{aligned}$$

Taking the limit as $n \to \infty$ and using (8), we find that (2) holds.

Next, fix $\varepsilon > 0$ and choose $\{\delta_t\}_{t=1}^\infty$ in $\mathbf{R}_+$ such that $\sum_{t=1}^\infty \beta^{t-1}\delta_t \le \varepsilon/2$. Since (5) holds, we can choose $x_1 \in \Gamma(x_0)$, $x_2 \in \Gamma(x_1), \ldots$ so that

$$v(x_t) \le F(x_t, x_{t+1}) + \beta v(x_{t+1}) + \delta_{t+1}, \quad t = 0, 1, \ldots.$$

Then $\underset{\sim}{x} = (x_0, x_1, \ldots) \in \Pi(x_0)$, and

$$\begin{aligned} v(x_0) &\le \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) + \beta^{n+1} v(x_{n+1}) + (\delta_1 + \ldots + \beta^n \delta_{n+1}) \\ &\le u_n(\underset{\sim}{x}) + \beta^{n+1} v(x_{n+1}) + \varepsilon/2, \quad n = 1, 2, \ldots. \end{aligned}$$

Hence (8) implies that for all n sufficiently large, $v(x_0) \le u_n(\underset{\sim}{x}) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, it follows that (3) holds.

If (8) holds, then (7) implies that v cannot take the value $-\infty$. If $v(x_0) = +\infty$, choose $n \ge 0$ and $(x_0, x_1, \ldots, x_n)$ such that $x_t \in \Gamma(x_{t-1})$ and $v(x_t) = +\infty$ for $t = 0, 1, \ldots, n$, and $v(x_{n+1}) < +\infty$ for all $x_{n+1} \in \Gamma(x_n)$. Clearly (8) implies that n is finite. Fix any $A > 0$. Since $v(x_n) = +\infty$, (6) implies that we can choose $x_{n+1}^A \in \Gamma(x_n)$ such that

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

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

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

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

It is an immediate consequence of Theorem 4.3 that the functional equation (FE) has at most one solution satisfying (8).

In summary, we have established two main results about solutions to (FE). Theorem 4.2 shows that v^* satisfies (FE). The functional equation may have other solutions as well, but Theorem 4.3 shows that these extraneous solutions always violate (8). Hence a solution to (FE) that satisfies (8) is v^*. The following example is a case where (FE) has an extraneous solution in addition to v^*.

Consider a consumer whose objective function is simply discounted consumption. The consumer has initial wealth $x_0 \in X = \mathbf{R}$, and he can borrow or lend at the interest rate $\beta^{-1} - 1$, where $\beta \in (0, 1)$. There are no constraints on borrowing, so his problem is simply

$$\max_{\{(c_t, x_{t+1})\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t c_t$$

$$\text{s.t.} \quad 0 \leq c_t \leq x_t - \beta x_{t+1}, \quad t = 0, 1, \ldots,$$

$$x_0 \text{ given.}$$

Since consumption is unbounded, the supremum function is obviously $v^*(x) = +\infty$, all x. Now consider the recursive formulation of this problem. The return function is $F(x, y) = x - \beta y$, and the correspondence describing the feasible set is $\Gamma(x) = (-\infty, \beta^{-1}x]$; so the functional equation is

$$v(x) = \sup_{y \leq \beta^{-1}x} [x - \beta y + \beta v(y)].$$

The function $v^*(x) = +\infty$ satisfies this equation, as Theorem 4.2 implies, but the function $v(x) = x$ does, too. But since the sequence $x_t = \beta^{-t}x_0$, $t = 0, 1, \ldots$, is in $\Pi(x_0)$, (8) does not hold and Theorem 4.3 does not apply.

The next exercise gives two variations on Theorem 4.3 that are sometimes useful when (8) does not hold.

Exercise 4.3 Let X, Γ, F, and β satisfy Assumptions 4.1–4.2. Let v be a solution to (FE) with

$$\limsup_{n\to\infty} \beta^n v(x_n) \leq 0, \quad \text{all } x_0 \in X, \text{ all } (x_0, x_1, \ldots) \in \Pi(x_0),$$

a. Show that $v \leq v^*$.

b. Suppose in addition that for each $x_0 \in X$ and $\underset{\sim}{x} \in \Pi(x_0)$, there exists $\underset{\sim}{x}' = (x_0, x_1', x_2', \ldots) \in \Pi(x_0)$ such that $\lim_{n\to\infty} \beta^n v(x_n') = 0$ and $u(\underset{\sim}{x}') \geq u(\underset{\sim}{x})$. Show that $v = v^*$.

Our next task is to characterize feasible plans that attain the optimum, if any do. Call a feasible plan $\underset{\sim}{x} \in \Pi(x_0)$ an *optimal plan from* x_0 if it attains the supremum in (SP), that is, if $u(\underset{\sim}{x}) = v^*(x_0)$. The next two theorems deal with the relationship between optimal plans and those that satisfy the policy equation (1) for $v = v^*$. The next theorem shows that optimal plans satisfy (1).

THEOREM 4.4 *Let X, Γ, F, and β satisfy Assumption 4.1–4.2. Let $\underset{\sim}{x}^* \in \Pi(x_0)$ be a feasible plan that attains the supremum in (SP) for initial state x_0. Then*

(9) $$v^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta v^*(x_{t+1}^*), \quad t = 0, 1, 2, \ldots.$$

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

(10) $$v^*(x_0^*) = u(\underset{\sim}{x}^*) = F(x_0, x_1^*) + \beta u(\underset{\sim}{x}^{*\prime})$$
$$\geq u(\underset{\sim}{x}) = F(x_0, x_1) + \beta u(\underset{\sim}{x}'), \quad \text{all } \underset{\sim}{x} \in \Pi(x_0).$$

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

$$u(\underset{\sim}{x}^{*\prime}) \geq u(\underset{\sim}{x}'), \quad \text{all } \underset{\sim}{x} \in \Pi(x_1^*).$$

Hence $u(\underset{\sim}{x}^{*\prime}) = v(x_1^*)$. Substituting this into (10) gives (9) for $t = 0$. Continuing by induction establishes (9) for all t. ■

The next theorem provides a partial converse to Theorem 4.4. It shows that any sequence satisfying (9) and a boundedness condition is an optimal plan.

THEOREM 4.5 *Let X, Γ, F, and β satisfy Assumptions 4.1–4.2. Let $\underset{\sim}{x}^* \in \Pi(x_0)$ be a feasible plan from x_0 satisfying (9), and with*

$$\limsup_{t\to\infty} \beta^t v^*(x_t^*) \le 0. \tag{11}$$

Then $\underset{\sim}{x}^*$ attains the supremum in (SP) for initial state x_0.

Proof. Suppose that $\underset{\sim}{x}^* \in \Pi(x_0)$ satisfies (9) and (11). Then it follows by an induction on (9) that

$$v^*(x_0) = u_n(\underset{\sim}{x}^*) + \beta^{n+1} v^*(x_{n+1}^*), \quad n = 1, 2, \ldots.$$

Then using (11), we find that $v^*(x_0) \le u(\underset{\sim}{x}^*)$. Since $\underset{\sim}{x}^* \in \Pi(x_0)$, the reverse inequality holds, establishing the result. ■

The consumption example used after Theorem 4.3 can be modified to illustrate why (11) is needed. Let preferences be as specified before, so that $c_t = x_t - \beta x_{t+1} = F(x_t, x_{t+1})$, all t. However, let us prohibit indebtedness by requiring $x_t \ge 0$, all t. Then in sequence form the problem is

$$\max_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t (x_t - \beta x_{t+1})$$

$$\text{s.t. } 0 \le x_{t+1} \le \beta^{-1} x_t, \quad t = 0, 1, \ldots,$$

$$x_0 \text{ given.}$$

If we cancel all of the offsetting terms in the objective function, it follows immediately that the supremum function is $v^*(x_0) = x_0$, all $x_0 \ge 0$. It is also clear that v^* satisfies the functional equation

$$v^*(x) = \max_{y \in [0, \beta^{-1}x]} [(x - \beta y) + \beta v^*(y)], \quad \text{all } x,$$

as Theorem 4.2 implies.

Now consider plans that attain the optimum. Given any $x_0 \ge 0$, the set of feasible plans $\Pi(x_0)$ consists of the sequences

$$(x_0, 0, 0, 0, \ldots), (x_0, \beta^{-1}x_0, 0, 0, \ldots),$$

$$(x_0, \beta^{-1}x_0, \beta^{-2}x_0, 0, \ldots), \text{ etc.,}$$

and all convex combinations thereof. Hence every feasible plan satisfies (9). It is straightforward to verify that, as Theorem 4.5 implies, any plan that satisfies (11) as well yields utility $v^*(x_0) = x_0$. (Essentially, it does not matter when consumption occurs as long as it occurs in finite time.) On the other hand, the feasible plan $x_t = \beta^{-t}x_0$, $t = 0, 1, \ldots$, (in each period invest everything and consume nothing) yields discounted utility of zero, for all $x \ge 0$. For $x > 0$, however, it violates (11), so Theorem 4.5 does not apply.

We will call any nonempty correspondence $G: X \to X$, with $G(x) \subseteq \Gamma(x)$, all $x \in X$, a *policy correspondence,* since the set $G(x)$ is a feasible set of actions if the state is x. If G is single-valued, we will call it a *policy function* and denote it by a lowercase g. If a sequence $\underset{\sim}{x} = (x_0, x_1, \ldots)$ satisfies $x_{t+1} \in G(x_t)$, $t = 0, 1, 2, \ldots$, we will say that $\underset{\sim}{x}$ is *generated from* x_0 *by* G. Finally, we will define the *optimal policy correspondence* G^* by

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

Then Theorem 4.4 shows that every optimal plan $\{x_t^*\}$ is generated from G^*, and Theorem 4.5 shows that any plan $\{x_t^*\}$ generated from G^*—if, in addition, it satisfies (11)—is an optimal plan.

4.2 Bounded Returns

In this section we study functional equations of the form

$$v(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)], \tag{1}$$

under the assumption that the function F is bounded and the discount factor β is strictly less than one.

As above, let X be the set of possible values for the state variable; let Γ: $X \to X$ be the correspondence describing the feasibility constraints; let

$A = \{(x, y) \in X \times X: y \in \Gamma(x)\}$ be the graph of Γ; let $F: A \rightarrow \mathbf{R}$ be the return function; and let $\beta \geq 0$ be the discount factor. Throughout this section, we will impose the following two assumptions on X, Γ, F, and β.

ASSUMPTION **4.3** *X is a convex subset of $\mathbf{R}^l$, and the correspondence Γ: $X \rightarrow X$ is nonempty, compact-valued, and continuous.*

ASSUMPTION **4.4** *The function $F: A \rightarrow \mathbf{R}$ is bounded and continuous, and $0 < \beta < 1$.*

It is clear that under Assumptions 4.3–4.4, Assumptions 4.1–4.2 hold, so the sequence problem corresponding to (1) is well defined. Moreover, Theorems 4.2–4.5 imply that under these assumptions solutions to (1) coincide exactly—in terms of both values and optimal plans—to solutions of the sequence problem.

The requirement that X be a subset of a finite-dimensional Euclidean space could be relaxed in much of what follows, but at the expense of a substantial additional investment in terminology and notation. (Recall that the definitions of u.h.c. and l.h.c. provided in Chapter 3 applied only to correspondences from one Euclidean space to another.) However, most of the arguments in this section apply much more broadly. Also note that the assumption that X is convex is not needed for Theorems 4.6 and 4.7.

If B is a bound for $|F(x, y)|$, then the supremum function v^* satisfies $|v^*(x)| \leq B/(1 - \beta)$, all $x \in X$. In this case it is natural to seek solutions to (1) in the space $C(X)$ of bounded continuous functions $f: X \rightarrow \mathbf{R}$, with the sup norm: $\|f\| = \sup_{x \in X}|f(x)|$. Clearly, any solution to (1) in $C(X)$ satisfies the hypothesis of Theorem 4.3 and hence is the supremum function. Moreover, given a solution $v \in C(X)$ to (1), we can define the policy correspondence $G: X \rightarrow X$ by

$$G(x) = \{y \in \Gamma(x): v(x) = F(x, y) + \beta v(y)\}, \tag{2}$$

and Theorems 4.4 and 4.5 imply that for any $x_0 \in X$, a sequence $\{x_t^*\}$ attains the supremum in the sequence problem if and only if it is generated by G.

The rest of the section proceeds as follows. Define the operator T on $C(X)$ by

$$(Tf)(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta f(y)], \tag{3}$$

so (1) becomes $v = Tv$. First, if we use only the boundedness and continuity restrictions in Assumptions 4.3 and 4.4, Theorem 4.6 establishes that $T: C(X) \rightarrow C(X)$, that T has a unique fixed point in $C(X)$, and that the policy correspondence G defined in (2) is nonempty and u.h.c. Theorem 4.7 establishes that under additional monotonicity restrictions on F and Γ, v is strictly increasing. Theorem 4.8 establishes that under additional concavity restrictions on F and convexity restrictions on Γ, v is strictly concave and G is a continuous (single-valued) function. Theorem 4.9 shows that if $\{v_n\}$ is a sequence of approximations defined by $v_n = T^n v_0$, with v_0 appropriately chosen, then the sequence of associated policy functions $\{g_n\}$ converges uniformly to the optimal policy function g given by (2). Finally, Theorem 4.11 establishes that if F is continuously differentiable, then v is, too.

THEOREM **4.6** *Let X, Γ, F, and β satisfy Assumptions 4.3 and 4.4, and let $C(X)$ be the space of bounded continuous functions $f: X \rightarrow \mathbf{R}$, with the sup norm. Then the operator T maps $C(X)$ into itself, $T: C(X) \rightarrow C(X)$; T has a unique fixed point $v \in C(X)$; and for all $v_0 \in C(X)$,*

$$\|T^n v_0 - v\| \leq \beta^n \|v_0 - v\|, \quad n = 0, 1, 2, \ldots. \tag{4}$$

Moreover, given v, the optimal policy correspondence $G: X \rightarrow X$ defined by (2) is compact-valued and u.h.c.

Proof. Under Assumptions 4.3 and 4.4, for each $f \in C(X)$ and $x \in X$, the problem in (3) 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 Γ is compact-valued and continuous, it follows from the Theorem of the Maximum (Theorem 3.6) that Tf is continuous. Hence $T: C(X) \rightarrow C(X)$.

It is then immediate that T satisfies the hypotheses of Blackwell's sufficient conditions for a contraction (Theorem 3.3). Since $C(X)$ is a Banach space (Theorem 3.1), it then follows from the Contraction Mapping Theorem (Theorem 3.2), that T has a unique fixed point $v \in C(X)$, and (4) holds. The stated properties of G then follow from the Theorem of the Maximum, applied to (1). ∎

It follows immediately from Theorem 4.3 that under the hypotheses of Theorem 4.6, the unique bounded continuous function v satisfying

(1) is the supremum function for the associated sequence problem. That is, Theorems 4.3 and 4.6 together establish that under Assumptions 4.3–4.4 the supremum function is bounded and continuous. Moreover, it then follows from Theorems 4.5 and 4.6 that there exists at least one optimal plan: any plan generated by the (nonempty) correspondence G is optimal.

To characterize v and G more sharply, we need more information about F and Γ. The next two results show how Corollary 1 to the Contraction Mapping Theorem can be used to obtain more precise characterizations of v and G.

ASSUMPTION 4.5 *For each y, $F(\cdot, y)$ is strictly increasing in each of its first l arguments.*

ASSUMPTION 4.6 *Γ is monotone in the sense that $x \le x'$ implies $\Gamma(x) \subseteq \Gamma(x')$.*

THEOREM 4.7 *Let X, Γ, F, and β satisfy Assumptions 4.3–4.6, and let v be the unique solution to (1). Then v is strictly increasing.*

Proof. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, nondecreasing functions on X, and let $C''(X) \subset C'(X)$ be the set of strictly increasing functions. Since $C'(X)$ is a closed subset of the complete metric space $C(X)$, by Theorem 4.6 and Corollary 1 to the Contraction Mapping Theorem (Theorem 3.2), it is sufficient to show that $T[C'(X)] \subseteq C''(X)$. Assumptions 4.5 and 4.6 ensure that this is so. ∎

ASSUMPTION 4.7 *F is strictly concave; that is,*

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

$$\text{all } (x, y), (x', y') \in A, \quad \text{and all } \theta \in (0, 1),$$

and the inequality is strict if $x \neq x'$.

ASSUMPTION 4.8 *Γ is convex in the sense that for any $0 \le \theta \le 1$, and $x, x' \in X$,*

$$y \in \Gamma(x) \quad \text{and} \quad y' \in \Gamma(x') \quad \text{implies}$$

$$\theta y + (1 - \theta)y' \in \Gamma[\theta x + (1 - \theta)x'].$$

Assumption 4.8 implies that for each $x \in X$, the set $\Gamma(x)$ is convex and there are no "increasing returns." Note that since X is convex, Assumption 4.8 is equivalent to assuming that the graph of Γ (the set A) is convex.

THEOREM 4.8 *Let X, Γ, F, and β satisfy Assumptions 4.3–4.4 and 4.7–4.8; let v satisfy (1); and let G satisfy (2). Then v is strictly concave and G is a continuous, single-valued function.*

Proof. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, weakly concave functions on X, and let $C''(X) \subset C'(X)$ be the set of strictly concave functions. Since $C'(X)$ is a closed subset of the complete metric space $C(X)$, by Theorem 4.6 and Corollary 1 to the Contraction Mapping Theorem (Theorem 3.2), it is sufficient to show that $T[C'(X)] \subseteq C''(X)$.

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

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

Let $y_i \in \Gamma(x_i)$ attain $(Tf)(x_i)$, for $i = 0, 1$. Then by Assumption 4.8, $y_\theta = \theta y_0 + (1 - \theta)y_1 \in \Gamma(x_\theta)$. It follows that

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

where the first line uses (3) and the fact that $y_\theta \in \Gamma(x_\theta)$; the second uses the hypothesis that f is concave and the concavity restriction on F in Assumption 4.7; and the last follows from the way y_0 and y_1 were selected. Since x_0 and x_1 were arbitrary, it follows that Tf is strictly concave, and since f was arbitrary, that $T[C'(X)] \subseteq C''(X)$.

Hence the unique fixed point v is strictly concave. Since F is also concave (Assumption 4.7) and, for each $x \in X$, $\Gamma(x)$ is convex (Assumption 4.8), it follows that the maximum in (3) is attained at a unique y value. Hence G is a single-valued function. The continuity of G then follows from the fact that it is u.h.c. (Exercise 3.11). ∎

Theorems 4.7 and 4.8 characterize the value function by using the fact that the operator T preserves certain properties. Thus if v_0 has property

P and if P is preserved by T, then we can conclude that each function in the sequence $\{T^n v_0\}$ has property P. Then, if P is preserved under uniform convergence, we can conclude that v also has property P. The same general idea can be used to establish facts about the policy function g, but we need to establish the sense in which the approximate policy functions—the functions g_n that attain $T^n v_0$—converge to g. The next result draws on Theorem 3.8 to address this issue.

THEOREM 4.9 (*Convergence of the policy functions*) *Let X, Γ, F, and β satisfy Assumptions 4.3–4.4 and 4.7–4.8, and let v and g satisfy (1) and (2). Let $C'(X)$ be the set of bounded, continuous, concave functions f: $X \to \mathbf{R}$, and let $v_0 \in C'(X)$. Let $\{(v_n, g_n)\}$ be defined by*

$$v_{n+1} = Tv_n, \quad n = 0, 1, 2, \ldots, \quad \text{and}$$

$$g_n(x) = \underset{y \in \Gamma(x)}{\operatorname{argmax}} \, [F(x, y) + \beta v_n(y)], \quad n = 0, 1, 2, \ldots.$$

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

Proof. Let $C''(X) \subset C'(X)$ be the set of strictly concave functions f: $X \to \mathbf{R}$. As shown in Theorem 4.8, $v \in C''(X)$. Moreover, as shown in the proof of that theorem, $T[C'(X)] \subseteq C''(X)$. Since $v_0 \in C'(X)$, it then follows that every function v_n, $n = 1, 2, \ldots$, is strictly concave. Define the functions $\{f_n\}$ and f by

$$f_n(x, y) = F(x, y) + \beta v_n(y), \quad n = 1, 2, \ldots, \quad \text{and}$$

$$f(x, y) = F(x, y) + \beta v(y).$$

Since F satisfies Assumption 4.7, it follows that each function f_n, $n = 1, 2, \ldots$, is strictly concave, as is f. Hence Theorem 3.8 applies and the desired results are proved. ■

The next exercise deals with the case where the state space X is finite or countable, as it is in computational applications.

Exercise 4.4 Let $X = \{x_1, x_2, \ldots\}$ be a finite or countable set; let the correspondence Γ: $X \to X$ be nonempty and finite-valued; let $A = \{(x, y) \in X \times X: y \in \Gamma(x)\}$; let F: $A \to \mathbf{R}$ be a bounded function; and let $0 < \beta < 1$. Let $B(X)$ be the set of bounded functions f: $X \to \mathbf{R}$, with the sup norm. Define the operator T by (3).

a. Show that T: $B(X) \to B(X)$; that T has a unique fixed point $v \in B(X)$; that (4) holds for all $v_0 \in B(X)$; and that the optimal policy correspondence G: $X \to X$ defined by (2) is nonempty.

Let H be the set of functions h: $X \to X$ such that $h(x) \in \Gamma(x)$, all $x \in X$. For any $h \in H$, define the operator T_h on $B(X)$ by $(T_h f)(x) = F[x, h(x)] + \beta f[h(x)]$.

b. Show that for any $h \in H$, T_h: $B(X) \to B(X)$, and T_h has a unique fixed point $w \in B(X)$.

Let $h_0 \in H$ be given, and consider the following algorithm. Given h_n, let w_n be the unique fixed point of T_{h_n}. Given w_n, choose h_{n+1} so that $h_{n+1}(x) \in \operatorname{argmax}_{y \in \Gamma(x)} [F(x, y) + \beta w_n(y)]$.

c. Show that the sequence of functions $\{w_n\}$ converges to v, the unique fixed point of T. [*Hint*. Show that $w_0 \leq Tw_0 \leq w_1 \leq Tw_1 \leq \ldots$.]

An algorithm based on Exercise 4.4 involves applying the operators T_{h_n}—operators that require no maximization—repeatedly and applying T only infrequently. Since maximization is usually the expensive step in these computations, the savings can be considerable.

Once the existence of a unique solution $v \in C(X)$ to the functional equation (1) has been established, we would like to treat the maximum problem in that equation as an ordinary programming problem and use the standard methods of calculus to characterize the policy function g. For example, consider the functional equation for the one-sector growth model:

$$v(x) = \max_{0 \leq y \leq f(x)} \{U[f(x) - y] + \beta v(y)\}.$$

If we knew that v was differentiable (and that the solution to the maximum problem in (1) was always interior), then the policy function g would be given implicitly by the first-order condition

$$U'[f(x) - g(x)] - \beta v'[g(x)] = 0. \tag{5}$$

Moreover, if we knew that v was twice differentiable, the monotonicity of g could be established by differentiating (5) with respect to x and examining the resulting expression for g'. However, the legitimacy of these methods depends upon the differentiability of the functions U, f, v, and g. We are free to make whatever differentiability assumptions we choose for U and f, but the properties of v and g must be established. We turn next to what is known about this issue.

It has been shown by Benveniste and Scheinkman (1979) that under fairly general conditions the value function v is *once* differentiable. That is, (5) is valid under quite broad conditions. However, known conditions ensuring that v is *twice* differentiable (and hence that g is once differentiable) are extremely strong (see Araujo and Scheinkman 1981). Thus differentiating (5) is seldom useful as a way of establishing properties of g. However, in cases where g is monotone, it is usually possible to establish that fact by a direct argument involving a first-order condition like (5).

We begin with the theorem proved by Benveniste and Scheinkman.

THEOREM 4.10 (*Benveniste and Scheinkman*) *Let $X \subseteq \mathbf{R}^l$ be a convex set, let $V: X \to \mathbf{R}$ be concave, let $x_0 \in \operatorname{int} X$, and let D be a neighborhood of x_0. If there is a concave, differentiable function $W: D \to \mathbf{R}$, with $W(x_0) = V(x_0)$ and with $W(x) \leq V(x)$ for all $x \in D$, then V is differentiable at x_0, and*

$$V_i(x_0) = W_i(x_0), \quad i = 1, 2, \ldots, l.$$

Proof. Any subgradient p of V at x_0 must satisfy

$$p \cdot (x - x_0) \geq V(x) - V(x_0) \geq W(x) - W(x_0), \quad \text{all } x \in D,$$

where the first inequality uses the definition of a subgradient and the second uses the fact that $W(x) \leq V(x)$, with equality at x_0. Since W is differentiable at x_0, p is unique, and any concave function with a unique subgradient at an interior point x_0 is differentiable at x_0 (cf. Rockafellar 1970, Theorem 25.1, p. 242). ■

Figure 4.1 illustrates the idea behind this result.

Applying this result to dynamic programs is straightforward, given the following additional restriction.

ASSUMPTION 4.9 *F is continuously differentiable on the interior of A.*

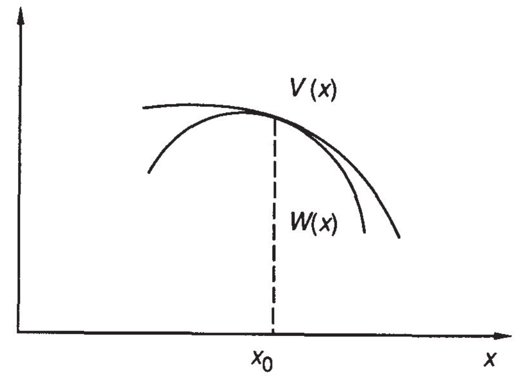

Figure 4.1

THEOREM 4.11 (*Differentiability of the value function*) *Let X, Γ, F, and β satisfy Assumptions 4.3–4.4 and 4.7–4.9, and let v and g satisfy (1) and (2). If $x_0 \in \operatorname{int} X$ and $g(x_0) \in \operatorname{int} \Gamma(x_0)$, then v is continuously differentiable at x_0, with derivatives given by*

$$v_i(x_0) = F_i[x_0, g(x_0)], \quad i = 1, 2, \ldots, l.$$

Proof. Since $g(x_0) \in \operatorname{int} \Gamma(x_0)$ and Γ is continuous, it follows that $g(x_0) \in \operatorname{int} \Gamma(x)$, for all x in some neighborhood D of x_0. Define W on D by

$$W(x) = F[x, g(x_0)] + \beta v[g(x_0)].$$

Since F is concave (Assumption 4.7) and differentiable (Assumption 4.9), it follows that W is concave and differentiable. Moreover, since $g(x_0) \in \Gamma(x)$ for all $x \in D$, it follows that

$$W(x) \leq \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)] = v(x), \quad \text{all } x \in D,$$

with equality at x_0. Hence v and W satisfy the hypotheses of Theorem 4.10, and the desired results follow immediately. ■

Note that the proof requires only that F be differentiable in its first l arguments.

With differentiability of the value function established, it is often straightforward to show that the optimal policy function g is monotone, and to bound its slope.

Exercise 4.5 Consider the first-order condition (5). Assume that U, f, and v are strictly increasing, strictly concave, and once continuously differentiable, and that $0 < g(x) < f(x)$, all x. Use (5) to show that g is strictly increasing and has slope less than the slope of f. That is,

$$0 < g(x') - g(x) < f(x') - f(x), \quad \text{if } x' > x.$$

[*Hint.* Refer to Figure 4.2.]

In specific applications it is often possible to obtain much sharper characterizations of v or of G or of both than those provided by the theorems above. It is useful to keep in mind that once the existence and uniqueness of the solution to (1) has been established, the right side of that equation can be treated as an ordinary maximization problem. Thus whatever tools can be brought to bear on that problem should be exploited. But such arguments usually rely on properties of F or of Γ or of both that are specific to the application at hand. The problems in Chapter 5 provide a variety of illustrations of specific arguments of this type.

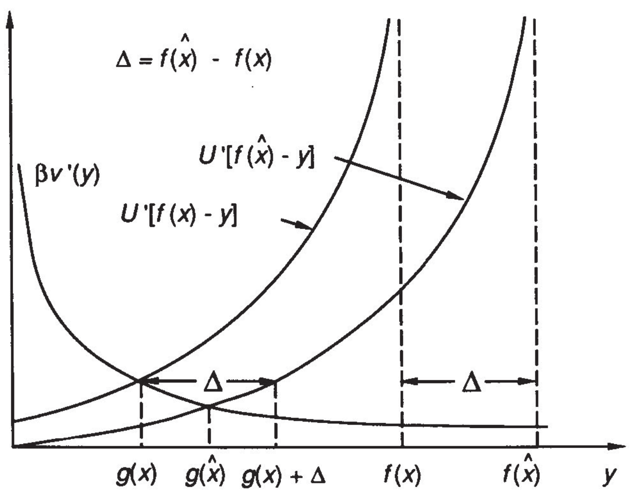

Figure 4.2

It should also be emphasized that even in cases that do not quite fit the assumptions of this section, arguments similar to the ones above can often be used. In this sense the results above should be viewed as suggestive, not (by any means) definitive. Sections 5.12 and 5.15 illustrate this point, as do many other applications in the literature. One particularly good illustration is the case of dynamic programming problems that exhibit constant returns to scale, to which we turn next.

4.3 Constant Returns to Scale

We sometimes wish to work with return functions F that are unbounded. For example, in the one-sector model of optimal growth, any utility function of the form $U(c) = (c^{1-\sigma} - 1)/(1 - \sigma)$, $\sigma \geq 1$, together with any technology of the form $f(k) = k^{\alpha}$, $0 < \alpha \leq 1$, leads to an unbounded return function. In this case and others like it, Assumption 4.4 is violated if X is taken to be all of $\mathbf{R}^l_+$. There are several ways to deal with problems of this type.

In some cases it is natural to restrict the state space to be a compact set $X \subset \mathbf{R}^l_+$. If Γ is compact-valued and continuous and if F is continuous, then with this restriction on X imposed, F is bounded on the compact set A. In these cases the arguments in Section 4.2 can be applied directly. Thus a judicious choice of the state space is very often all that is needed to apply those arguments to problems in which utility functions, profit functions, and so on, are unbounded. Illustrations of this method are given in Sections 5.1 and 5.9.

However, there are also many interesting cases where the state space cannot be so restricted. For example, no model of capital accumulation in which the technology permits sustained growth can be treated in this way. In this section and the next, we describe two ways in which the arguments in Section 4.2 can be adapted to models with unbounded returns.

This section deals with systems in which the return function and feasibility constraints both display constant returns to scale, and the constraints have the further property that feasible sequences $\{x_t\}$ cannot grow "too fast." First we show that Theorems 4.2–4.5 hold for problems of this type, so solutions to the functional equation correspond exactly to solutions of the original problem posed in terms of sequences, in terms of both values and policies. Theorems 4.12 and 4.13 then establish that the functional equation has a unique solution, and that this solution and the associated policy correspondence are homogeneous of degree one.

Throughout this section we let X be a *convex cone* in $\mathbf{R}^l$. That is, $X \subseteq \mathbf{R}^l$ is a convex set with the property that $x \in X$ implies $\alpha x \in X$, for any $\alpha \geq 0$. For example, $\mathbf{R}^l$ and $\mathbf{R}^l_+$ are both convex cones. In place of Assumptions 4.3 and 4.4, we will use the following restrictions. As in Section 4.2, let A denote the graph of Γ.

ASSUMPTION 4.10 *$X \subseteq \mathbf{R}^l$ is a convex cone. The correspondence Γ: $X \to X$ is nonempty, compact-valued, and continuous, and for any $x \in X$,*

$$y \in \Gamma(x) \text{ implies } \lambda y \in \Gamma(\lambda x), \quad \text{all } \lambda \geq 0.$$

That is, the graph of Γ is a cone. In addition, for some $\alpha \in (0, \beta^{-1})$,

$$\|y\|_E \leq \alpha\|x\|_E, \quad \text{all } x \in X \text{ and } y \in \Gamma(x),$$

(where $\|\cdot\|_E$ denotes the Euclidean norm on $\mathbf{R}^l$).

ASSUMPTION 4.11 *$\beta \in (0, 1)$; and $F: A \to \mathbf{R}$ is continuous and homogeneous of degree one, and for some $0 < B < \infty$,*

$$|F(x, y)| \leq B(\|x\|_E + \|y\|_E), \quad \text{all } (x, y) \in A.$$

Assumption 4.10 says that the correspondence Γ describing the feasibility constraints shows constant returns to scale, and it bounds the rate of growth of $\{\|x_t\|_E\}$ for feasible sequences $\{x_t\}$ by β^{-1}. Assumption 4.11 says that the return function F displays constant returns to scale, and it imposes a uniform bound on the ratio of F to the norm of its arguments.

Under Assumptions 4.10–4.11 we have the following results.

Exercise 4.6 Show that under Assumptions 4.10–4.11,

a. $\|x_t\|_E \leq \alpha^t\|x_0\|_E$, $t = 1, 2, \ldots$, all $x_0 \in X$, all $\{x_t\} \in \Pi(x_0)$;

b. Assumptions 4.1–4.2 hold; and

c. the supremum function v^*: $X \to \mathbf{R}$ defined in Section 4.1 is homogeneous of degree one, and for some $0 < c < \infty$, satisfies $|v^*(x)| \leq c\|x\|_E$, all $x \in$ X.

Part (b) of this exercise establishes that under Assumptions 4.10–4.11, Theorems 4.2 and 4.4 hold. That is, the supremum function v^* satisfies

the functional equation, and every optimal sequence $\{x_t^*\}$ (if any exist) is generated from the policy correspondence G associated with v^*. Moreover, v^* has the properties established in part (c) of the exercise.

Our next task is to choose an appropriate space of functions within which to look for solutions to the functional equation and then to define an appropriate operator on that space. In view of the results in Exercise 4.6c, it is natural to seek solutions to the functional equation within the space of functions $f: X \to \mathbf{R}$ that are continuous and homogeneous of degree one, and bounded in the sense that $|f(x)|/\|x\|_E < +\infty$, all $x \in X$. To capture the latter fact, it is useful to use the norm

$$\|f\| = \max_{\substack{\|x\|_E=1 \\ x\in X}} |f(x)|. \tag{1}$$

Let $H(X)$ be the space of functions $f: X \to \mathbf{R}$ that are continuous and homogeneous of degree one, and bounded in the norm in (1). Define the operator T on $H(X)$ by

$$(Tf)(x) = \sup_{y\in\Gamma(x)} [F(x, y) + \beta f(y)]. \tag{2}$$

Exercise 4.7 a. Show that $H(X)$, with the norm in (1), is a complete normed vector space.

b. Show that under Assumptions 4.10 and 4.11, $T: H(X) \to H(X)$.

It follows directly from Exercise 4.6a that for any $f \in H(X)$,

$$|f(x_t)| \leq \|x_t\|_E \|f\| \leq \alpha^t\|x_0\|_E \|f\|, \quad \text{all } x_0 \in X, \quad \text{all } \{x_t\} \in \Pi(x_0).$$

Since $\alpha\beta < 1$, it then follows that the hypotheses of Theorems 4.3 and 4.5 hold. Therefore v^* is the only solution in $H(X)$ to the functional equation, and every sequence $\{x_t^*\}$ generated by the associated policy correspondence G is optimal. Thus the Principle of Optimality applies to this type of constant-returns-to-scale problem.

The contraction property of the operator T can be verified by using a modification of Blackwell's sufficient conditions for a contraction (Theorem 3.3). For any function f that is homogeneous of degree one and for any $a \in \mathbf{R}$, we will in this context define the function $f + a$ by

$$(f + a)(x) = f(x) + a\|x\|,$$