3 Mathematical Preliminaries

In Chapter 2 the optimal growth problem

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

$$\text{s.t.} \quad c_t + k_{t+1} \leq f(k_t),$$

$$c_t, k_{t+1} \geq 0, \quad t = 0, 1, \ldots,$$

$$\text{given } k_0,$$

was seen to lead to the functional equation

(1) $$v(k) = \max_{c,y} [U(c) + \beta v(y)]$$

$$\text{s.t.} \quad c + y \leq f(k),$$

$$c, y \geq 0.$$

The purpose of this chapter and the next is to show precisely the relationship between these two problems and others like them and to develop the mathematical methods that have proved useful in studying the latter. In Section 2.1 we argued in an informal way that the solutions to the two problems should be closely connected, and this argument will be made rigorous later. In the rest of this introduction we consider alternative methods for finding solutions to (1), outline the one to be pursued, and describe the mathematical issues it raises. In the remaining sections of the chapter we deal with these issues in turn. We draw upon this material extensively in Chapter 4, where functional equations like (1) are analyzed.

In (1) the functions U and f are given—they take specific forms known to us—and the value function v is unknown. Our task is to prove the existence and uniqueness of a function v satisfying (1) and to deduce its properties, given those of U and f. The classical (nineteenth-century) approach to this problem was the *method of successive approximations*, and it works in the following very commonsensical way. Begin by taking an initial guess that a specific function, call it v_0, satisfies (1). Then define a new function, v_1, by

$$v_1(k) = \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v_0(y)\}. \tag{2}$$

If it should happen that $v_1(k) = v_0(k)$, for all $k \ge 0$, then clearly v_0 is a solution to (1). Lucky guessing (cf. Exercise 2.3) is one way to establish the existence of a function satisfying (1), but it is notoriously unreliable. The method of successive approximations proceeds in a more systematic way.

Suppose, as is usually the case, that $v_1 \ne v_0$. Then use v_1 as a new guess and define the sequence of functions $\{v_n\}$ recursively by

$$v_{n+1}(k) = \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v_n(y)\}, \quad n = 0, 1, 2, \ldots. \tag{3}$$

The hope behind this iterative process is that as n increases, the successive approximations v_n get closer to a function v that actually satisfies (1). That is, the hope is that the limit of the sequence $\{v_n\}$ is a solution v. Moreover, if it can be shown that $\lim_{n\to\infty} v_n$ is the same for any initial guess v_0, then it will follow that this limit is the only function satisfying (1). (Why?)

Is there any reason to hope for success in this analytical strategy? Recall that our reason for being interested in (1) is to use it to locate the optimal capital accumulation policy for a one-sector economy. Suppose we begin by choosing any feasible capital accumulation policy, that is, any function g_0 satisfying $0 \le g_0(k) \le f(k)$, all $k \ge 0$. [An example is the policy of saving a constant fraction of income: $g_0(k) = \theta f(k)$, where $0 < \theta < 1$.] The lifetime utility yielded by this policy, as a function of the initial capital stock k_0, is

$$w_0(k_0) = \sum_{t=0}^{\infty} \beta^t U[f(k_t) - g_0(k_t)],$$

where

$$k_{t+1} = g_0(k_t),\ t = 0, 1, 2, \ldots.$$

The following exercise develops a result about (g_0, w_0) that is used later.

Exercise 3.1 Show that

$$w_0(k) = U[f(k) - g_0(k)] + \beta w_0[g_0(k)], \quad \text{all } k \ge 0.$$

If the utility from the policy g_0 is used as the initial guess for a value function—that is, if $v_0 = w_0$—then (2) is the problem facing a planner who can choose capital accumulation optimally for one period but must follow the policy g_0 in all subsequent periods. Thus $v_1(k)$ is the level of lifetime utility attained, and the maximizing value of y—call it $g_1(k)$—is the optimal level for end-of-period capital. Both v_1 and g_1 are functions of beginning-of-period capital k.

Notice that since $g_0(k)$ is a feasible choice in the first period, the planner will do no worse than he would by following the policy g_0 from the beginning, and in general he will be able to do better. That is, for any feasible policy g_0 and associated initial value function v_0,

$$\begin{aligned} v_1(k) &= \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v_0(y)\} \\ &\ge \{U[f(k) - g_0(k)] + \beta v_0[g_0(k)]\} \\ &= v_0(k), \end{aligned} \tag{4}$$

where the last line follows from Exercise 3.1.

Now suppose the planner has the option of choosing capital accumulation optimally for two periods but must follow the policy g_0 thereafter. If y is his choice for end-of-period capital in the first period, then from the second period on the best he can do is to choose $g_1(y)$ for end-of-period 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 $X = \{x \in \mathbf{R}^2 : x = \alpha z$, 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*

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

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*

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

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 $\mathbf{R}$ 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 $f: X \to \mathbf{R}$ 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 $f^i = d^i f(x)/dx^i$. 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 $T: S \to S$ 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 $T: S \to S$ 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 $T: S \to S$ 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 $T^0x = x$, and $T^nx = T(T^{n-1}x)$, $n = 1, 2, \ldots$. 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,

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

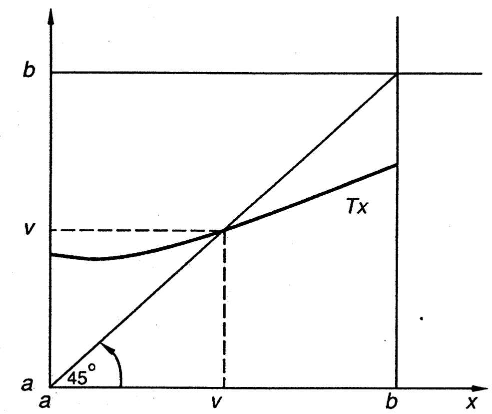

Figure 3.1

Continuing by induction, we get

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

Hence, for any $m > n$,

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

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 $v_n \to v \in S$.

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* $f(x) \le g(x)$, *for all* $x \in X$, *implies* $(Tf)(x) \le (Tg)(x)$, *for all* $x \in X$;

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

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

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

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

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

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

Combining these two inequalities, we find that $\|Tf - Tg\| \le \beta\|f - g\|$, 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

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

If $v(y) \le w(y)$ 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

$$T(v + a)(k) = \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta[v(y) + a]\}$$

$$= \max_{0 \le y \le f(k)} \{U[f(k) - y] + \beta v(y)\} + \beta a$$

$$= (Tv)(k) + \beta a.$$

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

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

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

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

is well defined, as is the nonempty set

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

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

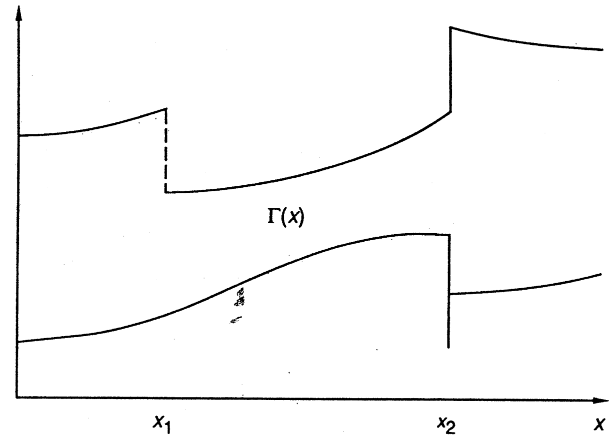

Figure 3.2

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 $\phi: X \to Y$ and $\psi: X \to Y$ 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 $\phi: X \to Y$ and $\psi: X \to Y$ 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 $\phi: X \to Y$ and $\psi: Y \to Z$ 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 $\Gamma_i: X \to Y_i$, $i = 1, \ldots, k$, be compact-valued and u.h.c. Show that $\Gamma: X \to Y = Y_1 \times \ldots \times Y_k$ 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 $\Gamma: X \to Y$ 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 $\Gamma: \mathbf{R}^k \to \mathbf{R}^{l+m}$, 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 l.h.c., then so is ϕ.

c. Let $\phi: X \to Y$ and $\psi: X \to Y$ 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 $\phi: X \to Y$ and $\psi: X \to Y$ 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 $\text{int } \phi(x) \cap \text{int } \psi(x) \neq \emptyset$, then Γ is l.h.c. at x.

e. Show that if $\phi: X \to Y$ and $\psi: Y \to Z$ 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 $\Gamma_i: X \to Y_i$, $i = 1, \ldots, k$, be l.h.c. Show that $\Gamma: X \to Y = Y_1 \times \ldots \times Y_k$ 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 $\Gamma: \mathbf{R}_+ \to \mathbf{R}_+$ 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 $\Gamma: \mathbf{R}^l_+ \to \mathbf{R}_+$ 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 $\Gamma: \mathbf{R}^l_+ \times \mathbf{R}^m \to \mathbf{R}^l_+$ by

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

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 $\Gamma: X \to Y$, 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 $\Gamma: X \to Y$ 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: $B(x, \varepsilon) = \{x' \in X: \|x - x'\| \leq \varepsilon\}$.

THEOREM 3.5 *Let $\Gamma: X \to Y$ be a nonempty-valued correspondence, and let A be the graph of Γ. Suppose that A is convex; 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 $x \in \hat{X}$; and that for every $x \in X$, there exists some $\varepsilon > 0$ such that the set $B(x, \varepsilon) \cap X$ is closed and convex. Then Γ is l.h.c.*

Proof. Choose $\hat{x} \in X$; $\hat{y} \in \Gamma(\hat{x})$; and $\{x_n\} \subset X$ with $x_n \to \hat{x}$. Choose $\varepsilon > 0$ such that the set $\hat{X} = B(\hat{x}, \varepsilon) \cap X$ is closed and convex. 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

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

Since D is a bounded set and $x_n \to \hat{x}$, it follows that $\alpha_n \to 0$. 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. ∎

We are now ready to answer the questions we posed at the beginning of this section: 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, which will repeatedly be applied later.

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 and 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

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

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

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

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

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

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.