$$KF_K(K,L)+LF_L(K,L)=hF(K,L) \tag{1}$$

Suppose $h=1$ (i.e., we have constant returns to scale) and there is perfect competition. Then factor prices are given by the corresponding marginal products, and equation (1) says that when capital and labor are paid their equilibrium prices, total output is just exhausted. If $h>1$, however (i.e., when we have increasing returns), factors cannot be paid their marginal product, because the required amount is larger than total output.

7 It can be shown that if F is homogeneous of degree 1, then its partial derivatives are homogeneous of degree 0, that is, for any $\lambda>0$, $F_K(K,L)=F_K(\lambda K,\lambda L)$, and similarly for F_L. See Chapter 4.

8 A similar model with a Cobb-Douglas technology was proposed simultaneously by Swan (1956). Hence, we sometimes speak of the Solow-Swan model.

9 The first equality in this expression follows by L'Hôpital's rule whenever $f(\ )$ is unbounded.

10 Notice that firms return undepreciated capital to the old workers after production takes place. Because the old "eat everything," the young have to start from scratch each period.

11 Technical progress can be handled in the same way as population growth. Let g be the rate of labor-augmenting technical progress, i.e., $A_{t+1}=(1+g)A_t$, and define $Z=K/AL$. Then we have

$$\frac{K_{t+1}}{A_{t+1}L_{t+1}}=\frac{L_ts_t}{(1+n)(1+g)A_tL_t}\Rightarrow(1+n)(1+g)Z_{t+1}=\frac{s[A_tw(Z_t),f'(Z_{t+1})+(1-\delta)]}{A_t}$$

where $Aw(Z)$ is the salary per worker. Notice, however, that this equation will not, in general, have a constant-Z solution. If preferences are homothetic, however, the savings function is of the form $s(y,R)=s(R)y$, and the previous expression simplifies to

$$(1+n)(1+g)Z_{t+1}=s[f'(Z_{t+1})+(1-\delta)]w(Z_t)$$

which does have a steady state.

12 It is shown in the proof that $\phi'(0)>1$ requires that $w'(0)>1$. Galor and Ryder (1989) have shown that this condition is stronger than the Inada condition $f'(0)=\infty$. Hence, the Inada condition is not sufficient to guarantee the existence of a nontrivial steady state.

13 Because Z is a function of A, which is not observable, it may instead be better to work with the growth rate of the capital stock per worker. Although data on this variable are indeed available for some countries, their quality is in general rather poor, and the available figures may not be fully comparable across countries. Hence, it may be better to use a transformation of (1) that will allow us to work directly with (more reliable) data on investment flows, rather than with capital stocks.

14 See Barro and Sala-i-Martin (1990, 1992) and Mankiw, Romer, and Weil (1992) for empirical applications of this methodology.

15 Because z is in logs, this is approximately the deviation from the steady state in percentage terms.

16 This can also be done inside the function. Thus we could replace the statement in *In[7]* by

```
sol1=NDSolve[{z'[t]==F[z[t], 0.25, 0.69, 0.03, 0.02,
0.01],z[0]==zs1/2}, z,{t, 0, 100}]
```

12

An Introduction to Dynamic Optimization

This chapter contains an introduction to dynamic optimization. In Section 1 we develop some basic elements of dynamic programming that are then used in Section 2 in an informal derivation of the maximum principle. Applications will be discussed in Chapter 13.

1. Dynamic Programming

Consider a system, economic or otherwise, whose evolution over time can be at least partially controlled by the actions of a decision-maker. At each point in time s, the state of the system can be described by a dated vector of real variables, $x_s\in\mathbb{R}^n$, which we call the *state vector.* In each period the decision-maker chooses a vector of *control* or *decision variables,* $u_s\in\mathbb{R}^m$. Together, the current state of the system and the choices of controls determine the value of the state vector for the following period according to the (possibly time-dependent) law of motion

$$x_{s+1}=m_s(x_s,u_s) \tag{1}$$

Thus, different choices of the control variables will yield different time paths of the system. It will be assumed that the decision-maker has preferences defined over such time paths that can be summarized by a time-additive *return* or *objective function*

$$W_t=\sum_{s=t}^{T-1}f_s(x_s,u_s) \tag{2}$$

For simplicity, we will take as given the planning horizon and the initial and terminal values of the state vector. Thus, we consider the problem faced by a planner who inherits at time t a predetermined state vector x_t, cares only about what happens between times t and $T\ (\le\infty)$, and is obliged to leave the state vector with value x_T at the end of the planning period. The agent can also be constrained by further restrictions on the state and control vectors, which we will write $(x_s, u_s) \in C_s$ for each s.

Given the initial state of the system, x_t, and a sequence $\mathbf{u}_{t,T-1} = \{u_s;\ s = t, t+1, \ldots, T-1\}$ of control variables, the evolution of the state vector is determined by the law of motion (1). Thus, x_t and $\mathbf{u}_{t,T-1}$ induce a sequence of states $\mathbf{x}_{t+1,T} = \{x_s;\ s = t+1, \ldots, T\}$. We will write $\mathbf{z}_{t,T} = \{\mathbf{u}_{t,T-1} \cup \mathbf{x}_{t+1,T}\}$ and say that a such sequence is *admissible* if both states and controls are feasible at all times and the terminal value of the state vector is equal to the required value, x_T. The set of all sequences $\mathbf{z}_{t,T}$ admissible from a given initial state vector x_t will be denoted by $\Phi(x_t)$, or by $\Phi(x_t, x_T)$ when we also want to make explicit the terminal constraint on the state. When we want to indicate explicitly the initial and terminal conditions on this sequence, we will write $\mathbf{z}(\mathbf{u}_{t,T-1}, x_t, x_T)$, and we will denote the portion of $\mathbf{z}_{t,T}$ between points a and b in time by $\mathbf{z}(\mathbf{u}_{t,T-1}, x_t, x_T)|_a^b$.

In this notation the decision-maker's objective function can be written

$$W(\mathbf{z}_{t,T}, t, T-1) = \sum_{s=t}^{T-1} f_s(x_s, u_s) \tag{2'}$$

Notice that $W(\ )$ is given by the sum of the instantaneous or period return functions $\{f_s\}$, where each f_s is a function only of time and the current state and control vectors and does not depend on either past or future values of x or u.

(a) The Principle of Optimality and Bellman's Equation

The problem the agent faces is that of choosing the time path of the control variables so as to maximize the objective function W_t subject to the law of motion (1) and appropriate feasibility constraints, taking as given the planning horizon (t, T) and the initial and terminal values of the state vector. We will denote by $V(\ )$ the *value function* for the planner's problem (i.e., the maximum attainable value of the objective function). Clearly, $V(\ )$ will be a function of the parameters of the maximization problem (the initial and terminal times and state vectors) and is equal to the objective function evaluated at the optimal control path and the induced state sequence, assuming they exist. Formally, the problem can be written

$$\begin{aligned} V(x_t, t; x_T, T) &= \max_{\mathbf{u}_{t,T-1}} \{W[\mathbf{z}(\mathbf{u}_{t,T-1}, x_t, x_T), t, T-1] \\ &= \sum_{s=t}^{T-1} f_s(x_s, u_s) \text{ s.t. } x_{s+1} = m_s(x_s, u_s),\ t,\ T,\ x_t, \\ &\quad \text{and } x_T \text{ given, } (x_s, u_s) \in C_s \subseteq \mathbb{R}^{n+m} \text{ for each } s\} \end{aligned} \tag{DP}$$

If T is finite (which may not be the case), (DP) can be solved by the stan-

applying the Lagrange or Kuhn-Tucker theorems). The structure of the problem, moreover, permits some important simplifications and will also allow us to deal with infinite-horizon problems (to which the standard theorems do not apply). The features that make things easier are the additive separability of the objective function and the simple structure of the law of motion – the fact that for each s, f_s (the period return function) and m_s (the law of motion) depend only on s and on the current values of the state and control variables, but not on their past or future values, and that the total return is simply the sum of the period return functions.

This property has the following implication. Let $\mathbf{z}_{t,T} = \mathbf{z}(\mathbf{u}_{t,T-1}, x_t, x_T)$ be an admissible sequence of controls and induced states between end points x_t and x_T, and let a and b be positive integers, with $t \le a < b \le T-1$. Then we can write the return function in the form

$$W(\mathbf{z}_{t,T}, t, T-1) = W\left(\mathbf{z}_{t,T}\Big|_t^{a-1}, t, a-1\right) + W\left(\mathbf{z}_{t,T}\Big|_a^{b-1}, a, b-1\right) + W\left(\mathbf{z}_{t,T}\Big|_b^{T-1}, b, T-1\right)$$

That is, the total payoff associated with a state–control sequence over the whole planning horizon is simply the sum of the payoffs associated with different portions of the sequence over the corresponding subperiods. Using this additivity property, it is easy to establish the following result, which gives an important property of the optimal solution of (DP).

Theorem 1.1. The principle of optimality. Let $z^*_{t,T} = \mathbf{z}(\mathbf{u}^*_{t,T-1}, x_t, x_T) = \{u^*_s, x^*_{s+1}\}$ *be the optimal solution of (DP) between given end points* (x_t, t) *and* (x_T, T). *Given arbitrary points* a *and* b, *with* $t \le a < b \le T-1$, *let* x^*_a *and* x^*_b *be the corresponding terms of the optimal state sequence* $\{x^*_s\}$. *Then the optimal solution to*

$$\begin{aligned} V(x^*_a, a; x^*_b, b) = \max_{u_{a,b-1}} \Big\{ & W(z_{a,b-1}, a, b-1) = \sum_{s=a}^{b-1} f_s(x_s, u_s) \\ & s.t.\ x_{s+1} = m_s(x_s, u_s),\ a,\ b,\ x^*_a, \text{ and } x^*_b \text{ given,} \\ & (x_s, u_s) \in C_s \subseteq \mathbb{R}^{n+m} \text{ for each } s \Big\} \end{aligned} \tag{DP.ab}$$

is given by $z^*_{t,T}|_a^{b-1}$.

Roughly speaking, the theorem says that each portion of the optimal plan is optimal on its own right. More precisely, any portion of an optimal trajectory is an optimal trajectory for an appropriate subproblem of (DP) in which we constrain the end-point values of the state vector to be equal to the corresponding terms of the optimal state sequence for the whole

Proof. We proceed by contradiction. Let $\Phi(x^*_a, x^*_b)$ be the set of feasible trajectories $\mathbf{z}_{a,b-1}$ between end points (x^*_a, a) and (x^*_b, b). This set is not empty, as it contains at least the relevant portion of the optimal sequence for the whole problem, $\mathbf{z}^*_{t,T-1}|_a^b$, which exists by assumption. Now suppose that $\mathbf{z}^*_{t,T-1}|_a^{b-1}$ is not optimal for the subperiod from a to b. Then there exists a feasible sequence between these end points, $\mathbf{z}'_{a,b-1}$, such that $W(\mathbf{z}'_{a,b-1}) > W(\mathbf{z}^*_{t,T-1}|_a^{b-1})$. By the time-additivity of the objective function,

$$W\left(\mathbf{z}^*_{t,T}\Big|_t^{a-1}\right) + W(\mathbf{z}'_{a,b-1}) + W\left(\mathbf{z}^*_{t,T}\Big|_b^{T-1}\right) > W\left(\mathbf{z}^*_{t,T}\Big|_t^{T-1}\right)$$

Hence, we have found a sequence $\mathbf{z}^*_{t,T-1}|_t^{a-1} \cup \mathbf{z}'_{a,b-1} \cup \mathbf{z}^*_{t,T-1}|_b^{T-1}$ that yields a higher return than $\mathbf{z}^*_{t,T-1}$. Moreover, because this sequence is feasible by construction, we have reached a contradiction: $\mathbf{z}^*_{t,T-1}$ cannot be an optimal solution for the "whole" problem. □

Problem 1.2. A violation of the principle of optimality. Consider an agent who lives three periods and maximizes a utility function of the form

$$V_1 = U_1 + \alpha U_2 + \beta U_3$$

where utility in period i, U_i, is a function of current and (expected) future consumption, that is,

$$U_1(c_1, c_2, c_3) = \ln(c_1 c_2 c_3), \quad U_2(c_2, c_3) = \ln(c_2 c_3), \quad \text{and} \quad U_3(c_3) = \ln c_3$$

and the budget constraint is of the form

$$A_{t+1} = A_t - c_t \qquad (A_1 \text{ given, and } A_4 = 0)$$

where A is wealth.

Notice that the return function is additive, but not separable over periods, as the period-1 utility, for example, depends on (expected) consumption at times 2 and 3. Hence, the assumptions of Theorem 1.1 do not hold, and, as we will see, the principle of optimality fails.

(i) Compute the optimal consumption plan from the perspective of time 1, $c^1 = (c^1_1, c^1_2, c^1_3)$.

(ii) Next, consider what happens as the agent begins to implement this plan. At time 1, he consumes c^1_1, receives utility U_1, and has leftover wealth $A_2 = A_1 - c^1_1$. He then faces the problem of maximizing utility over the remainder of his life,

$$\max V_2 = \alpha U_2 + \beta U_3$$

subject to $c_2 + c_3 = A_2$. Compute the new optimal plan, $c^2 = (c^2_2, c^2_3)$, and compare it with the last portion of c^1. Has the consumer changed his mind? How and why? Does the Bellman equation

The principle of optimality has an important implication, sometimes called *time consistency:* Suppose we compute the optimal path from the beginning of the planning period and start moving along it. After a while, we stop and recalculate the optimal solution from the current time and state. The principle of optimality tells us that the solution of this new problem will be the remainder of the original optimal plan. Hence, the decision-maker will not be tempted to "change his mind."

This property allows us to approach the problem sequentially, leaving for tomorrow decisions about future controls, thus breaking up the original dynamic problem into a sequence of static subproblems. To make this precise, consider one particular decomposition of the problem, that into (i) today's choice of controls and (ii) all the rest of the plan. By the additivity of the objective function, we can write

$$\begin{aligned} V(x_t, t; x_T, T) &= \max_{\mathbf{u}_{t,T-1}} W[\mathbf{z}(\mathbf{u}_{t,T-1}, x_t, x_T), t, T-1] \\ &= \max_{u_t,\, \mathbf{u}_{t+1,T-1}} \{f_t(u_t, x_t) + W[\mathbf{z}(\mathbf{u}_{t+1,T-1}, x_{t+1}, x_T), t+1, T-1]\} \end{aligned}$$

where the maximization is subject to the usual constraints and, in particular, $x_{t+1} = m_t(x_t, u_t)$. The structure of the problem allows us to approach the choice of the current (u_t) and future $(\mathbf{u}_{t+1,T-1})$ controls sequentially. Notice that states and controls dated $t+1$ or higher do not affect the current return, given by $f_t(u_t, x_t)$, and that the current state and control vectors (x_t, u_t) affect future returns only through their effects on tomorrow's state, x_{t+1}. Thus, we can solve the problem in two steps: Given any choice of the current control, tomorrow we will face the problem of choosing $\mathbf{u}_{t+1,T-1}$ so as to maximize $W[\mathbf{z}(\mathbf{u}_{t+1,T-1}, x_{t+1}, x_t), t+1, T-1]$, taking as given the state x_{t+1} resulting from today's decision – a problem identical with today's except for the initial state and time. Having solved this problem, today's decision reduces to choosing u_t, taking into account both its direct contribution to the current return and its indirect contribution to future payoffs through its effect on tomorrow's state. The principle of optimality assures us that this stepwise or sequential maximization process will yield the same result as simultaneous determination of the whole control path. Thus, we can write

$$\begin{aligned} V(x_t, t; x_T, T) &= \max_{u_t,\, \mathbf{u}_{t+1,T-1}} \{f_t(u_t, x_t) + W[\mathbf{z}(\mathbf{u}_{t+1,T-1}, x_{t+1}, x_T), t+1, T-1]\} \\ &= \max_{u_t} \Big\{ f_t(u_t, x_t) + \max_{\mathbf{u}_{t+1,T-1}} \{W[\mathbf{z}(\mathbf{u}_{t+1,T-1}, x_{t+1}, x_T), t+1, T-1]\} \\ &\qquad \text{s.t. } x_{t+1} = m_t(x_t,\ u_t) \Big\} \end{aligned}$$

Finally, observe that the payoff resulting from the inside maximization is

$$\max_{u_t}\left\{f_t(u_t,x_t)+\max_{\mathbf{u}_{t+1,T-1}}\{W[\mathbf{z}(\mathbf{u}_{t+1,T-1},x_{t+1},x_T),t+1,T-1]\}\text{ s.t. }x_{t+1}=m_t(x_t,u_t)\right\}$$
$$=\max_{u_t}\{f_t(u_t,x_t)+V(x_{t+1},t+1;x_T,T)\text{ s.t. }x_{t+1}=m_t(x_t,u_t)\}$$

and we arrive at *Bellman's equation*,

$$V(x_t,t;x_T,T)=\max_{u_t}\{f_t(u_t,x_t)+V(x_{t+1},t+1;x_T,T)\text{ s.t. }x_{t+1}=m_t(x_t,u_t)\}\quad\text{(BE)}$$

This expression formally characterizes the optimal choice of the current control vector as the solution of a static optimization problem in which the future consequences of current actions are summarized by incorporating tomorrow's value function into today's objective function. The solution to the static maximization problem in (BE) yields a *policy function* that gives the optimal value of the current control, u_t^*, as a function $g_t(x_t)$ of time and the current state. Tomorrow's state is then given by $x_{t+1}=m_t[g_t(x_t),x_t]$, and a solution to a similar problem (with x_{t+1} now given) then yields tomorrow's optimal control. (Notice that time enters both the value and policy functions as a separate argument, reflecting the fact that periods may differ in factors other than the state vector.)

The recursive relation given by (BE) is useful in that it allows us to conceptually transform a dynamic choice problem into a sequence of static problems we already know how to handle, at least in principle. But notice that the maximization in Bellman's equation is not really a standard problem in at least one sense: The value function $V(\ )$ appears both inside and outside the maximization operator (although with different arguments) and therefore is not a known function. In fact, (BE) is a functional equation – an equation in the unknown function $V(\ )$. Hence, the reformulation of the original problem does not really solve it, nor put it in a form we can solve directly. The Bellman equation, however, does provide the basis for an alternative approach to the problem that will indeed lead to an operational solution method. In the sections that follow we will consider two cases: finite-horizon problems, and infinite-horizon problems with some additional restrictions.

(i) *Solution of Finite-Horizon Problems through Backward Induction*

Dynamic programming problems over a finite planning horizon do not present any conceptual difficulties. The value and policy functions can be obtained by starting from the end and working backward. The optimal control sequence can then be computed by applying the sequence of policy functions, $g_1,\ldots,g_{T-1}$, to the initial state vector.

One period before the end of the planning period the problem reduces to choosing the last control, taking as given the terminal value of the state. Omitting some of the

$$V(x_{T-1},T-1)=\max_{u_{T-1}}\{f_{T-1}(u_{T-1},x_{T-1})\text{ s.t. }m_{T-1}(x_{T-1},u_{T-1})=x_T\text{ given}\}$$

Notice that at this stage there is no unknown value function inside the max operator; hence, $V(x_{T-1},T-1)$ is well defined by the foregoing expression, and for a fully specified problem its computation is, in principle, straightforward. On the other hand, the value of x_{T-1} is not known at this point, but this does not matter, for we are interested in the whole function $V(\cdot,T-1)$, rather than its value for a specific state vector.

This procedure will also work for a class of problems more general than those we have considered thus far. In particular, we can abandon the assumption of a predetermined terminal state vector and let the agent choose x_T taking into account its contribution to his payoff, given by a *scrap* or *salvage* value function $S(x_T)$.[1] In this case, the last-stage maximization becomes

$$V(x_{T-1},T-1)=\max_{u_{T-1}}\{f_{T-1}(u_{T-1},x_{T-1})+S(x_T)\text{ s.t. }x_T=m_{T-1}(x_{T-1},u_{T-1})\}$$

In any case, the solution of the last-period problem yields a policy function that gives the optimal value of the last control as a function of the state at the beginning of the period: $u_{T-1}^*=g_{T-1}(x_{T-1})$. As for the value function, the value of the argument is not known at this stage, but what we want is the function itself.

Given $V(x_{T-1},T-1)$, we can go back one step and compute the value function for the previous period,

$$V(x_{T-2},T-2)=\max_{u_{T-2}}\{f_{T-2}(u_{T-2},x_{T-2})+V(x_{T-1},T-1)$$
$$\text{s.t. }x_{T-1}=m_{T-2}(x_{T-2},u_{T-2})\}$$

obtaining also the corresponding policy function, $u_{T-2}^*=g_{T-2}(x_{T-2})$. Proceeding in this manner, we eventually reach the initial period and solve

$$V(x_t,t)=\max_{u_t}\{f_t(u_t,x_t)+V(x_{t+1},t+1;x_T,T)\text{ s.t. }x_{t+1}=m_t(x_t,u_t)\}$$

to obtain the value function for the original problem and the first policy function, $g_t(\cdot)$. At this point, the initial value of the state, x_t, is a given quantity, and the whole sequence of policy functions $\{g_s;\ s=t,t+1,\ldots,T-1\}$ is also known. Hence, we can recover the optimal sequence of instruments, given by $u_s^*=g_s(x_s)$, and the induced sequence of states, $x_{s+1}=m_s(x_s,u_s^*)$.

(ii) *Discounting and Stationarity*

It should be clear that the foregoing solution algorithm cannot be used when the planning horizon is infinite, for there is no terminal date from which to problems can be dealt with and in many cases are easier to solve than finite-horizon problems. In this section we introduce some notation and impose some additional structure on the dynamic programming problem before briefly discussing a particular class of infinite-horizon problems that will be analyzed in greater detail later.

Discounting. In many situations, payoffs accruing at different points in time are valued differently by the decision-maker. Typically, those that are realized further into the future are valued less than those that accrue immediately. Although our earlier specification of a time-dependent period return function $f_s(x_s, u_s)$ implicitly allows for this possibility, it will be convenient to bring it out explicitly by introducing a sequence of period-specific weights. In particular, we will assume that the period return function at time s is of the form $f_s(x_s, u_s) = \alpha_s F_s(x_s, u_s)$, where the discount factor α_s is a nonnegative real number, and consider an agent who faces a problem of the form

$$V(x_0, 0) = \max_{\mathbf{u}_{0,T-1}} \sum_{s=0}^{T-1} \alpha_s F_s(x_s, u_s)$$

subject to the usual constraints. We will interpret $F_s(\;)$ as the payoff that accrues at time s, valued from the perspective of time s itself, and $f_s(\;) = \alpha_s F_s(\;)$ as the same payoff "discounted back" to the beginning of the planning period at time zero. Thus, multiplication of the current payoff $F_s(\;)$ by α_s brings it back to time-zero units, and division of the discounted payoff $f_s(\;)$ by the same factor brings it forward to time-s units. Because first-period returns need no discounting, we set α_0 equal to 1.

As time passes and the agent gets to period t, he faces the subproblem of maximizing the remainder of the objective function,

$$V(x_t, t) = \max_{\mathbf{u}_{t,T-1}} \sum_{s=t}^{T-1} \alpha_s F_s(x_s, u_s)$$

Notice that the value function in this expression gives the maximum attainable payoff evaluated from the perspective of time zero, because each period return is multiplied by the corresponding discount factor. When maximizing over the subperiod starting at t, however, it is often more convenient to make "current" valuations (as of time t). Thus, we define the current value function by

$$V^c(x_t, t) = \frac{V(x_t, t)}{\alpha_t} = \max_{\mathbf{u}_{t,T-1}} \sum_{s=t}^{T-1} \frac{\alpha_s}{\alpha_t} F_s(x_s, u_s)$$

As usual, successive subproblems are linked by the Bell

$$V(x_t, t) = \max_{u_t} \{\alpha_t F_t(x_t, u_t) + V(x_{t+1}, t+1)\}$$

To rewrite this equation in terms of current values, we divide both sides of this expression by α_t, obtaining

$$V^c(x_t, t) = \max_{u_t} \{F_t(x_t, u_t) + \beta_t V^c(x_{t+1}, t+1)\}$$

where the one-period discount factor, $\beta_t = \alpha_{t+1}/\alpha_t$, discounts values from $t + 1$ to t (multiplying by α_{t+1} brings them back to zero, dividing by α_t takes them back up to t). The interpretation of this expression is almost exactly the same as that of the undiscounted version of the Bellman equation: Given tomorrow's state, x_{t+1}, $V^c(x_{t+1}, t + 1)$ gives the maximum attainable payoff in "tomorrow's utility units." To bring it back to "today's units," we multiply $V^c(\;)$ by β_t. The optimal policy is then to choose u_t so as to maximize the sum of today's period return and the discounted value of tomorrow's current value function.

Infinite Horizon, Stationary Problem. In many problems of interest it can be assumed that the period return function, the law of motion, the one-period discount factor, and the feasible set C to which states and controls must belong are all time-invariant, that is,

$$F_s = F, \qquad m_s = m, \qquad C_s = C, \qquad \beta_s = \beta \qquad (\forall\, s)$$

This assumption allows some further simplifications of the problem. Notice that with a constant β, we have $\alpha_{s+1} = \beta\alpha_s$. This equation, together with the assumption that $\alpha_0 = 1$, implies that the discount factor must be of the form $\alpha_s = \beta^s$. Thus, the subproblem starting at time t can be written

$$V^c(x_t, t) = \max_{u_{t,T-1}} \sum_{s=t}^{T-1} \mathbf{b}^{s-t} F(x_s, u_s)$$

In the finite-horizon case, t is still a separate argument of the current value function, as subproblems that start at different dates differ from each other not only in the initial value of the state vector but also in the time remaining until the end of the planning period. If the planning horizon is infinite, however, this is no longer the case, and all subproblems are identical. Thus, for infinite-horizon stationary problems, the current value function is a function of the initial state alone, $V^c(x_t)$, and the Bellman equation becomes

$$V^c(x_t) = \max_{u_t} \{F(x_t, u_t) + \beta V^c(x_{t+1})\}$$

It follows that the policy function, $u_t^* = g(x_t)$, is also time-invariant. This is an important simplification, because we now have to find only one such func-

As noted earlier, the backward-induction algorithm cannot be used to solve infinite-horizon problems. The following observation, however, provides the basis for a way to deal with such problems, as we will see later. Given a function $v(\ )$ from $\mathbb{R}^n$ to $\mathbb{R}$, we can define an operator T mapping the space of such functions into itself:

$$Tv(x) = \max_u \{F(x,u) + \beta v(y) \text{ s.t. } y = m(x,u),\ (x,u) \in C\}$$

The Bellman equation can then be written $V^c = TV^c$. Hence, a function V solves Bellman's equation if and only if it is a fixed point of the operator T. Under certain assumptions, the contraction mapping theorem can be used to establish the existence and uniqueness of a solution to Bellman's equation and to determine some properties of interest of such a function.

(iii) Uncertainty

Dynamic programming is particularly useful when dealing with problems that involve uncertainty in a dynamic setting. Provided we ignore some technical problems, the previous discussion can be easily extended to deal with stochastic problems.

Imagine that instead of a deterministic law of motion we have a stochastic law: x_t and u_t no longer determine the value of x_{t+1}, but only its probability distribution, described by a distribution function of the form $G(x_{t+1}; x_t, u_t)$, where

$$G(y; x_t, u_t) = \text{pr}(x_{t+1} \le y \mid x_t, u_t)$$

Agents now maximize expected utility. At time t, they choose u_t, not knowing for certain the value of next period's state. Whatever x_{t+1} turns out to be, they will optimize from tomorrow on, obtaining a value of $V^c(x_{t+1}, t+1)$. From today's perspective, then, u_t must be chosen so as to maximize the sum of the current return and the discounted value of the expectation of $V^c(x_{t+1}, t+1)$, computed using $G(\ )$. Hence, the Bellman equation becomes

$$V^c(x_t, t) = \max_{u_t} \{F_t(u_t, x_t) + \beta_t \int V^c(x_{t+1}, t+1)\, dG(x_{t+1}; x_t, u_t)\}$$

(b) Some Results for Stationary Discounted Problems

In this section we will analyze in greater detail a class of infinite-horizon problems. Given a predetermined state vector x_t, a decision-maker faces the problem of maximizing the objective function

∞

with $\beta \in (0, 1)$, over the set of feasible sequences $\mathbf{z}_{t,\infty} = \{u_s, x_{s+1}\} \in \Phi(x_t)$, where $x_{s+1} = m(x_s, u_s)$. We will assume that the series W_t converges (although possibly to plus or minus infinity) for all feasible sequences $\mathbf{z}_{t,\infty}$ and that the feasibility constraints are of the form

$$u_s \in \Gamma(x_s)$$

where Γ is a correspondence mapping points in $\mathbb{R}^n$ into sets in $\mathbb{R}^m$. The problem faced by the agent can then be written

$$V^c(x_t) = \max_{\mathbf{u}_{t,\infty}} \left\{ \sum_{s=t}^{\infty} \beta^{s-t} F(x_s, u_s) \right. \\ \left. \text{s.t. } x_{s+1} = m(x_s, u_s),\ u_s \in \Gamma(x_s),\ x_t \text{ given} \right\} \qquad \text{(DP.}\infty\text{)}$$

and the current value function $V^c(x_t)$ gives the maximum attainable value of the objective function whenever the problem has a solution. We know from our previous discussion that if the value function does exist, then it satisfies the Bellman equation:

$$V^c(x) = \max_{u \in \Gamma(x)} \{F(x,u) + \beta V^c(y) \text{ s.t. } y = m(x,u)\} \qquad \text{(BE)}$$

The converse of this statement, however, is not necessarily true. The Bellman equation may have several solutions, and only one of them can be the value function for the programming problem. Hence, we need to establish conditions under which we can be sure that a given solution of (BE) is the value function we seek.

Theorem 1.3. Let the function $v:\mathbb{R}^n \longrightarrow \mathbb{R}$ *solve the Bellman equation (BE) and satisfy the boundedness condition*

$$\lim_{n\to\infty} \beta^n v(x_n) = 0 \qquad (0)$$

for any sequence $\{x_n\}$ *feasible from the initial state* x_t. *Suppose, moreover, that there exists a sequence* $z^*_{t,\infty} = x_t \cup \{u^*_s, x^*_{s+1}\}$, *where* u^*_s *solves*

$$v(x_s) = \max_{u_s \in \Gamma(x_s)} \{F(x_s, u_s) + \beta v[m(x_s, u_s)]\} \qquad (BE.s)$$

for each s *and* $x^*_{s+1} = m(x^*_s, u^*_s)$. *Then* v *is the current value function for the programming problem, and* $\mathbf{z}^*_{t,\infty}$ *solves (DP.∞)*.

Proof. To show that $v(\ )$ is the value function for the programming problem, we need to show that for any given initial state x_t, That is, $v(x_t)$ is an upper bound for the value of the problem over the set of feasible sequences, and there is a feasible sequence that attains this value.

Let $\mathbf{z}_{t,\infty} = x_t \cup \{u_s, x_{s+1}; s \geq t\}$ be an arbitrary sequence feasible from x_t. Then, by (BE.s),

$$
\begin{aligned}
v(x_t) = \max_{u_t \in \Gamma(x_t)} \{F(x_t, u_t) + \beta v(x_{t+1})\} &\geq F(x_t, u_t) + \beta v(x_{t+1}) \\
&\geq F(x_t, u_t) + \beta[F(x_{t+1}, u_{t+1}) + \beta v(x_{t+2})] \\
&\geq \ldots \geq \sum_{s=t}^{t+n} \beta^{s-t} F(x_s, u_s) + \beta^{n+1} v(x_{t+n+1}) \qquad (3)
\end{aligned}
$$

Taking the limit of this expression as $n \to \infty$, and using the boundedness condition (0),

$$v(x_t) \geq W_t(\mathbf{z}_{t,\infty}) + \lim_{n \to \infty} \beta^{n+1} v(x_{n+1}) = W_t(\mathbf{z}_{t,\infty})$$

for any feasible sequence $\mathbf{z}_{t,\infty}$. Hence, $v(x_t)$ is an upper bound for the value of the problem. Moreover, the sequence $\mathbf{z}^*_{t,\infty} = x_t \cup \{u^*_s, x^*_{s+1}\}$ of solutions to (BE.s) attains this value. Notice that by definition,

$$v(x^*_s) = \max_{u_s \in \Gamma(x_s)} \{F(x^*_s, u_s) + \beta v[m(x^*_s, u_s)]\} = F(x^*_s, u^*_s) + \beta v(x^*_{s+1})$$

Hence, all the weak inequalities in (3) hold as equalities, and we conclude that

$$v(x^*_t) = W_t(\mathbf{z}^*_{t,\infty})$$

which proves the theorem. □

Theorem 1.3 says that if we can find a bounded solution to the Bellman equation, the original problem reduces to a sequence of static maximizations. There is, however, no assurance that such a solution will exist in all cases. Our next task is to identify conditions under which the Bellman equation has a unique bounded solution. The discussion relies heavily on the reader's familiarity with the concepts of a complete metric space and the contraction mapping theorem (for a review of this material, see Section 7 of Chapter 2).

We define the operator T mapping real-valued functions into real-valued functions by

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

Then the Bellman equation can be written in the form

Thus, we see that finding a solution to the Bellman equation is equivalent to finding a fixed point of the operator T. If we can show that under appropriate assumptions, T is a contraction mapping a complete metric space into itself, we can invoke the contraction mapping theorem to establish the existence and uniqueness of an appropriate solution to (BE).

We recall from Chapter 2 (see Theorem 7.12) that the space $C(X)$ of bounded, continuous real-valued functions defined on a set X in $\mathbb{R}^n$ is a complete metric space under the sup norm, defined by

$$\|f\|_s = \sup\{|f(x)|;\ x \in X\}$$

Next we will check that under certain continuity and boundedness restrictions on the objective function, the law of motion, and the constraint correspondence, the operator T maps $C(X)$ into itself (i.e., T maps continuous bounded functions into continuous bounded functions) and that T is a contraction. By the contraction mapping theorem, it follows that (BE′) has a unique bounded solution in $C(X)$ that, by Theorem 1.3, is the value function we are seeking.

In what follows, we will make the following assumption.

Assumption 1.4. Continuity. The period return function F is bounded and continuous, the law of motion m is continuous, the constraint correspondence Γ is continuous,[2] and the set $\Gamma(x)$ is nonempty and compact for each x.

Under these conditions we can establish the following result.

Theorem 1.5. Suppose that Assumption 1.4 holds. Then T *is an operator mapping continuous bounded functions into continuous bounded functions. Moreover* $T: C(X) \longrightarrow C(X)$ *is a contraction and therefore has a unique fixed point* V *in* C(X). *This* V *is the value function for the corresponding dynamic programming problem.*

Moreover, under Assumption 1.4, the solution function for the maximization in (BE) is the policy correspondence g( ) *for the programming problem, giving the set of optimal values of the control* u *as a function of the state, and* g( ) *is nonempty and uhc.*

Proof

- Let $v \in C(X)$. Under our assumptions, the maximization problem that defines the operator T,

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

is, for each x, that of maximizing a continuous function on a compact set. Hence, both v and F are bounded, Tv is also bounded; and because F and v are continuous and the constraint correspondence is continuous and compact-valued, the theorem of the maximum (Theorem 2.1 in Chapter 7) guarantees the continuity of Tv. Hence, T maps $C(X)$ into itself. Moreover, by the theorem of the maximum, the solution mapping for this maximization (i.e., the policy function $g(x)$), is a nonempty and uhc correspondence.

- To establish that T is a contraction, we make use of Blackwell's sufficient conditions (see Theorem 7.19 in Chapter 2). We have to show that T satisfies

  (i) monotonicity: $\forall f, g \in C(X), f(x) \le g(x)\ \forall X \Rightarrow Tf(x) \le Tg(x)$, and
  (ii) discounting: $\exists \beta \in (0, 1)$ s.th. $\forall f \in C(X), x \in X, a \ge 0$: $T[f(x) + a] \le Tf(x) + \beta a$.

  First, suppose that $w(y) \le v(y)$ for all y in X. Then for each (x, u), $w[m(x, u)] \le v[m(x, u)]$, and therefore

  $$Tv(x) = \max_{u \in \Gamma(x)}\{F(x,u) + \beta v[m(x,u)]\} \ge \max_{u \in \Gamma(x)}\{F(x,u) + \beta w[m(x,u)]\} = Tw(x)$$

  Thus, T is monotone. Next, note that for any positive constant a, we have

  $$\begin{aligned} T[v(x)+a] &= \max_{u\in\Gamma(x)}\{F(x,u)+\beta\{v[m(x,u)]+a\}\} \\ &= \max_{u\in\Gamma(x)}\{F(x,u)+\beta v[m(x,u)]\}+\beta a = Tv(x)+\beta a \end{aligned}$$

  Hence, T discounts. Because it satisfies both of Blackwell's conditions, T is a contraction.
- Because T is a contraction on a complete metric space, it follows directly from the contraction mapping theorem (Theorem 7.15 in Chapter 2) that it has a unique fixed point V.
- By Theorem 1.3, the bounded continuous function V is the value function for the corresponding dynamic programming problem. Moreover, the solution mapping for the maximum problem in the Bellman equation is the optimal policy correspondence. □

It also follows from the contraction mapping theorem that if, starting with an arbitrary continuous and bounded function V_o, we define a sequence $\{V_n\}$ of functions by

$$V_{n-1} = TV_n$$

this sequence converges to the value function V. This fact can sometimes be used to find the value function.

Knowing when the Bellman equation has a unique bounded solution (i.e., when the value function is well defined) is an important first step, but one that is of little practical help. To go further we need to establish conditions under which V will have certain desirable properties.

In the remainder of this section we will use the foregoing results relating the value function with the bounded solution of the Bellman equation to

concave, and the policy correspondence is a continuous function. For this, we will rely on the following result: Recall that if (X, d) is a complete metric space and Y is a closed subset of X, then (Y, d) is also a complete metric space (Theorem 7.9 in Chapter 2). Now suppose that $T: X \longrightarrow X$ is a contraction in X and, moreover, that T maps Y into itself. Then T is also a contraction in Y, and it follows that the unique fixed point of T on X must be in Y. A slight twist on this result yields the following theorem.

Theorem 1.6. Let (X, d) *be a complete metric space, and let* $T: X \longrightarrow X$ *be a contraction with fixed point* $v \in X$. *Further, let* Y *be a closed subset of* X, *and assume that* T *maps points in* Y *into some subset* Z *of* Y *(i.e.,* $T: Y \longrightarrow Z$*). Then the unique fixed point* v *of* T *in* X *will be in* Z.

Problem 1.7. Prove Theorem 1.6.

We will show that the set $ND(X)$ of nondecreasing bounded and continuous functions is a closed subset of $C(X)$ and that the operator T in the Bellman equation maps nondecreasing functions into strictly increasing functions. It follows by Theorem 1.6 that the value function V must be strictly increasing. A similar argument will allow us to establish strict concavity.

Lemma 1.8. Consider the normed vector space $[C(X), \|\cdot\|_s]$, *where* $C(X)$ *is the set of bounded continuous functions* $f: \mathbb{R}^n \supseteq X \longrightarrow \mathbb{R}$, *with the sup norm* $\|f\|_s = \sup\{|f(x)|;\ x \in X\}$. *Let* $ND(X)$ *be the set of nondecreasing bounded and continuous functions on* X. *Then* $ND(X)$ *is a closed subset of* $C(X)$.

Recall that a function $f: X \longrightarrow \mathbb{R}$ is said to be nondecreasing if

$$\forall\, x_0, x_1 \in X,\ x_1 > x_0 \Rightarrow f(x_1) \ge f(x_0)$$

and strictly increasing if

$$\forall\, x_0, x_1 \in X,\ x_1 > x_0 \Rightarrow f(x_1) > f(x_0)$$

Proof. Let $\{f_n\}$ be a sequence of nondecreasing continuous functions convergent (in the sup norm and hence pointwise) to a function f (which is bounded and continuous, by the completeness of $C(X)$). To establish that $ND(X)$ is a closed subset of $C(X)$, it suffices to show that f is nondecreasing. Let x_0 and x_1 be arbitrary points in X such that $x_1 > x_0$, and consider the sequence of real numbers $\{f_n(x_1) - f_n(x_0)\}$. Because $\{f_n\} \to f$, $\{f_n(x_1) - f_n(x_0)\}$ converges to $f(x_1) - f(x_0)$, and because f_n is nondecreasing, $f_n(x_1) - f_n(x_0) \ge 0$

is nonnegative. This establishes that the limit function $f(\ )$ is also nondecreasing. □

Assumption 1.9. Monotonicity. Assume that $F(\ )$ is strictly increasing in x, $m(\ )$ is increasing in x, and the constraint correspondence $\Gamma(\ )$ is increasing in the sense that

$$x_1 \geq x_0 \Rightarrow \Gamma(x_1) \supseteq \Gamma(x_0)$$

Lemma 1.10. Let $T: C(X) \longrightarrow C(X)$ *be the operator defined by*

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

and assume that Assumption 1.9 (monotonicity) holds. Then T *maps nondecreasing functions into strictly increasing functions.*

Problem 1.11. Prove Lemma 1.10.

Combining these two lemmas with Theorem 1.6, the following result is immediate.

Theorem 1.12. Suppose that Assumptions 1.4 and 1.9 (continuity and monotonicity) hold. Then the value function V *is strictly increasing in the state* x.

To summarize, we know that under the continuity assumption the Bellman equation has a unique continuous and bounded solution V that is the value function for the corresponding programming problem. This function can be characterized as a fixed point of an appropriately defined operator $T: C(X) \longrightarrow C(X)$. We have shown that the set of nondecreasing bounded and continuous functions $ND(X)$ is a closed subset of $C(X)$ and that under Assumption 1.9, T maps nondecreasing functions into strictly increasing functions. It follows that the value function must be strictly increasing. Intuitively, our assumptions ensure that an "increase" in the state is strictly desirable because it strictly increases the current return and does not reduce future opportunities.

We will now develop a very similar argument to show that under certain conditions, V is strictly concave. Recall that a function f is said to be (weakly) concave if

$$\forall\, x_0, x_1 \in X,\ \lambda \in [0,1],\ (1-\lambda)f(x_0) + \lambda f(x_1) \leq f[(1-\lambda)x_0 + \lambda x_1]$$

$$\forall\, x_0, x_1 \in X,\ \lambda \in (0,1),\ (1-\lambda)f(x_0) + \lambda f(x_1) < f[(1-\lambda)x_0 + \lambda x_1]$$

Lemma 1.13. Consider the normed vector space $[C(X), \|\cdot\|_s]$, *where* $\|\cdot\|_s$ *is the sup norm, and assume* X *is a convex set. The set of (weakly) concave functions in* C(X) *is a closed subset of* C(X).

Problem 1.14. Prove Lemma 1.13.

Assumption 1.15. Concavity. Assume that F is strictly concave, m is concave, for each x the constraint set $\Gamma(x)$ is convex, and the constraint correspondence Γ is convex in the sense that

$$\forall\, x_0, x_1 \in X,\ \lambda \in [0,1],\ u_0 \in \Gamma(x_0), u_1 \in \Gamma(x_1)$$
$$\Rightarrow (1-\lambda)u_0 + \lambda u_1 \in \Gamma[(1-\lambda)x_0 + \lambda x_1]$$

Lemma 1.16. Let $T: C(X) \longrightarrow C(X)$ *be the operator defined by*

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

and assume that the concavity and monotonicity assumptions hold. Then T *maps weakly concave functions into strictly concave functions.*

Problem 1.17. Prove Lemma 1.16.

Using Lemmas 1.13 and 1.16 and Theorem 1.6, it follows that under the continuity, monotonicity, and concavity assumptions, the value function V is strictly concave.

Theorem 1.18. Suppose the continuity, monotonicity, and concavity assumptions hold. Then the value function V *is strictly concave and strictly increasing, and the policy correspondence* g( ) *is a continuous function.*

Proof. The first part of the theorem is immediate. Moreover, we know by the maximum theorem and Theorem 1.5 that the optimal policy correspondence is uhc. Because any single-valued uhc correspondence is a continuous function (see Section 11 of Chapter 2), we need only establish that the solution u^* to the maximization in the Bellman equation is unique, but this follows immediately by the strict concavity of F, the concavity of $m(\ )$, and the concavity and monotonicity of V, all of which ensure that the objective

Differentiability of the Value Function. The maximization in the Bellman equation is a static optimization problem that looks like an ordinary Lagrange or Kuhn-Tucker problem. Hence, one is tempted to write the Lagrangean function and differentiate it with respect to u to obtain a set of first-order conditions and then proceed in the usual way (by applying the implicit-function theorem or differentiating implicitly the first-order conditions) to establish the comparative-statics properties of the optimal policy function. This approach, however, presupposes that all the functions involved are twice differentiable, an assumption that generally is not valid.

The basic problem arises because the value function for the problem, $V(\ )$, appears also inside the maximization operator. Whereas we are free to make whatever assumptions we want about $m(\ )$ and $F(\ )$, the differentiability of $V(\ )$ must be established rather than directly assumed.[3] It can be shown that $V(\ )$ will be (once or twice) differentiable for a certain class of problems, but not in general.[4] As a result, the standard approach to studying the comparative-statics properties of maximization systems is not generally available for the case of dynamic programming problems.

2. Optimal Control

We now switch from discrete time to continuous time and develop the basic elements of optimal control theory. A central result of this section is a set of necessary conditions for an optimum in a certain class of dynamic optimization problems, the so-called maximum principle of Pontryagin. We will derive the maximum principle from a dynamic programming formulation. Roughly, we start with a discrete-time problem, apply the dynamic programming techniques discussed earlier, and consider what happens in the limit as the length of the period goes to zero.

The continuous-time analogue of the problem studied in the preceding section can be written

$$V^c(x_0, 0) = \max_{u(t),0\le t\le T} \Big\{ W_0\big(u(t)|_{t=0}^{T}, x(t)|_{t=0}^{T}\big) = \int_0^T \alpha(t) F[u(t), x(t), t]\, dt$$

$$+ \alpha(T) S[x(T)] \text{ s.t. } x(0) = x_0 \text{ given,}$$

$$\text{and } \dot{x}(t) = m[u(t), x(t), t] \Big\} \qquad \text{(P.0)}$$

where, as before, x is the state vector, and u the vector of control variables. The salvage or scrap function $S(\ )$ is used to allow for the possibility that we may place some value on the state at the end of the planning horizon T (which may or may not be finite). We will assume that the discount factor corresponding to period t is of the form

$$\alpha(t) = \exp\Big(-\int_0^t \rho(s)\, ds\Big)$$

which reduces to the more familiar $e^{-\rho t}$ whenever the discount rate ρ is constant over time.[5] The notation $x(t)$ indicates that the state is a function of time. For convenience, we will often replace this functional notation by the subscript notation x_t to indicate dependence on time, or omit the t's when they are not particularly needed. We will often treat x and u as if they were single variables, but the reader should keep in mind that they are vectors.

The problem is similar to the one analyzed in Section 1 except that the planner now must choose a control trajectory (i.e., a function of time, $u(t)$, defined for $t \in [0, T]$, that describes the values of the instruments at each point in time), rather than a control sequence $\{u_t\}_{t=0}^{T-1}$. Given a control path $u^0(t)|_{t=0}^{T}$, the corresponding trajectory of the state vector, $x^0(t)|_{t=0}^{T}$, is determined by the law of motion, $\dot{x}_t = m(u, x, t)$, and the initial condition $x(0) = x_0$. Evaluating W_0, we obtain the value of the given trajectories, $W_0(u^0(t)|_{t=0}^{T}, x^0(t)|_{t=0}^{T})$. Our goal is to characterize the time paths of u and x that will yield the largest possible value for the objective function. This will be achieved by transforming the dynamic maximization problem (P.0) into a combination of two more familiar problems: a static maximization at each point in time, and a system of ordinary differential equations.

(a) The Maximum Principle

We begin with an intuitive discussion of the logic of the maximum principle. At each point in time t the planner finds herself with some predetermined value of the state x_t and must choose a control vector u_t that will determine both the immediate payoff $F_t(\ )$ and the rate of change of the state variables $\dot{x}_t$.[6] Current decisions, therefore, have two effects on total value: an immediate one through $F_t(\ )$, and an indirect one through the induced change in x. Clearly, a control chosen to maximize just the current return is unlikely to be optimal. We need some way to take into account the effects of current decisions on future opportunities. Intuitively, the maximum principle achieves this by attaching a price to the *stocks* of state variables.

The idea is to introduce a modified objective function that will add to the immediate return the value of the change in the state vector due to current decisions. To this end, we introduce a new set of variables q_t, one for each component of the state vector. These variables, known as multipliers or *costate variables*, can be interpreted as the prices associated with the state variables. The modified objective function, known as the *current-value Hamiltonian*, is then defined as

13

Some Applications of Dynamic Optimization

In this chapter we will review some applications of dynamic optimization to economics. In Section 1 we develop two models of search to illustrate the use of dynamic programming in a stochastic setting. Section 2 analyzes the decision problem faced by a social planner who maximizes the utility of an infinitely-lived representative agent in a one-good neoclassical economy. In Section 3 we study the optimal investment policy of a competitive firm when the installation of capital is costly. Finally, in Section 4 we develop the Cass-Koopmans model of a dynamic competitive economy and use it to analyze the welfare cost of factor taxes. Section 5 concludes with a series of problems.

1. Search Models

Search theory provides a simple and yet interesting application of dynamic programming to economics. In the basic search model, wage offers drawn from a given distribution arrive at fixed or random intervals, and an agent simply decides whether to accept one of them and become employed or reject them and continue searching for a better opportunity. We have, then, a very simple problem in stochastic dynamic programming: The control is simply a take-it-or-leave-it decision, and the distribution of the state variables (the offers) is time-invariant and does not depend on either the state or the control.

The first part of this section introduces the basic "microeconomic" model of job search. In addition to its interest as an application of dynamic programming, this model provides a useful counterpoint to the neoclassical model of a competitive labor market. In the latter model, transactions are assumed to take place instantaneously and at no cost, and wages are set so that the market clears continuously. Hence, there is no room for unemployment. In the search model, on the other hand, it may be optimal for an agent to remain temporarily unemployed in order to wait for a better opportunity

than those available today. Hence, the search model provides a useful framework for analyzing how rational agents will respond to changes in the level or duration of unemployment benefits, the abundance and riskiness of employment offers, and many other questions that can hardly be addressed within the neoclassical model.

The search model, however, does not necessarily require a departure from the spirit of the neoclassical model. Notice, in particular, that the unemployment that naturally arises in any search model is frictional in nature and essentially voluntary. Hence, the explicit modeling of the process of job search may well yield nothing more than a model with a natural rate of unemployment. On the other hand, it is relatively easy to incorporate additional features into a search model that add a strong Keynesian flavor to it. If we are willing to assume that an increase in the level of aggregate activity makes it easier for potential trading partners to locate each other, we have a participation externality that generates inefficiency and the possibility of multiple equilibria, thus opening the door for public intervention to improve things. A "macro" model with these features will be developed in the second part of the section.

(a) The Basic Model of Job Search

Consider an infinitely-lived, risk-neutral worker who maximizes (the expectation of) the discounted value of lifetime income,

$$E\left\{\sum_{t=0}^{\infty}\beta^t y_t\right\}$$

where income at time t, y_t, is equal to the wage rate (x) for employed workers and to a government-provided benefit (b) for the unemployed. Unemployed workers also receive one employment offer each period. All jobs are permanent and pay the same wage each period. Wages, however, may differ across jobs. Hence, x is a (nonnegative) random variable that we assume to be drawn from a time-invariant distribution described by a cumulative distribution function (cdf) $F(\ )$, where $F(w) = \mathrm{pr}(x \leq w)$.

A worker who has just received an offer has two options: One is to accept the job and work forever at the specified wage x;[1] the other is to reject the offer and wait for a better one to arrive. We will denote the value of the first option (accepting and being employed at wage x) by $W_a(x)$, and that of the second (rejecting the offer and remaining unemployed) by W_r. Clearly, $W_a(x)$, the present value of lifetime earnings on a job paying salary x, is an increasing function of x given by

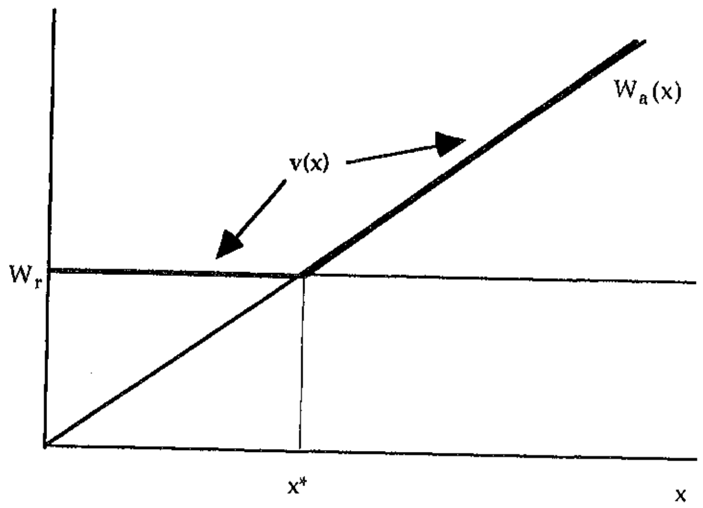

Figure 13.1. Value function and reservation wage for the search problem.

On the other hand, W_r is not a function of x: The expected present value of lifetime earnings for an unemployed worker is independent of the wage offer he has just rejected.

A rational worker will choose the action that will yield the larger value. Thus, the expected value of lifetime income for an agent who has just received an offer x is given by the value function

$$v(x) = \max[W_a(x,), W_r] \qquad (2)$$

and he accepts the offer if and only if $W_a(x) \geq W_r$ (i.e., if the value of being employed at the offered wage exceeds the value of being unemployed). As illustrated in Figure 13.1, the optimal decision strategy takes the form of a reservation-wage rule. Because $W_a(x)$ is increasing in the salary, and W_r is independent of it, a job will be accepted if and only if it pays a wage that is higher than some critical value x^*. This critical or *reservation wage* is defined as the value of x that makes the agent indifferent between taking the job and remaining unemployed, that is, x^* solves $W_a(x^*) = W_r$.

It remains, of course, to determine the reservation wage x^* or, equivalently, the value of being unemployed, W_r. As a first step, consider the situation of a worker who is currently unemployed (i.e., who has just rejected an offer): His income today is the unemployment benefit b; tomorrow he will receive a new offer, x, and will accept it or reject it depending on whether or not its value exceeds W_r. Hence, his current value one period hence (from tomorrow's perspective) will be given by $v(x) = \max[W_a(x), W_r]$. As of today, however, the realization of x is not known, so we can only work with the

we have to discount it by one period. Formally, then, the value of being unemployed is defined recursively by

$$W_r = b + \beta E\{\max[W_a(x), W_r]\} \qquad (3)$$

We can now characterize the reservation salary. By definition, x^* is the value of x that makes the agent indifferent between accepting and rejecting the offer. Hence, x^* satisfies

$$W_a(x^*) \equiv \frac{x^*}{1-\beta} = W_r$$

and therefore

$$x^* = (1-\beta)W_r \qquad (4)$$

Substituting (3) into (4),

$$x^* = W_r - \beta W_r = b + \beta E\{\max[W_a(x), W_r]\} - \beta W_r$$

Bringing the (constant) last term into the expectation and the max operator, we obtain

$$x^* = b + \beta E\{\max[W_a(x) - W_r, 0]\} \qquad (5)$$

an equation that can be solved for x^*. This expression can be simplified as follows. We begin by writing out the expectation,

$$x^* = b + \beta \int_0^\infty \max[W_a(x) - W_r, 0]\, dF(x) \qquad (6)$$

and observing that the resulting integral can be broken up into two parts:

$$\int_0^\infty \max[W_a - W_r, 0]\, dF = \int_0^{x^*} \max[W_a - W_r, 0]\, dF + \int_{x^*}^\infty \max[W_a - W_r, 0]\, dF$$

Notice that over the first interval of integration we have $x \leq x^*$, implying that $W_a(x) \leq W_r$; thus, $\max[W_a(x) - W_r, 0] = 0$ for $x \in (0, x^*]$, and the first integral vanishes. For $x \in (x^*, \infty)$, on the other hand, we have $W_a(x) \geq W_r$, implying $\max[W_a(x) - W_r, 0] = W_a(x) - W_r$. Hence, (6) reduces to

$$x^* = b + \beta \int_{x^*}^\infty [W_a(x) - W_r]\, dF(x)$$

Finally, recalling that $W_r = x^*/(1-\beta)$ and $W_a(x) = x/(1-\beta)$, we arrive at the *fundamental reservation-wage equation*,

$$x^* = b + \frac{\beta}{1-\beta} \int_{x^*}^\infty (x - x^*)\, dF(x) \qquad (R)$$

which implicitly defines the reservation wage x^* as a function of the parameters of the model and the distribution of wage offers. This equation can be used to study the comparative statics of the reservation wage, as we will

Continuous-Time and Stochastic-Offer Arrivals

One of the crucial determinants of how selective a worker can afford to be in regard to wage offers is the availability of job opportunities. The model in the preceding section, which assumes that the worker receives an offer every period, ignores this aspect of the problem. We will now relax this restrictive assumption and extend the model to incorporate a measure of the "scarcity" of work opportunities through a parameter that reflects the rate of arrival of job offers. We will also illustrate how to go from discrete time to continuous time – a formulation that, although less intuitive when it comes to the derivation of the valuation equations, turns out to be more convenient in many cases.

We will make two changes with respect to the earlier model. The first will be to parameterize the length of the period. We will assume that all periods have the same duration h and reinterpret the wage and the unemployment benefit as rates per unit of time. Thus, an unemployed worker's income during a period is now bh, and an employed worker earns xh. We will also assume that the one-period discount factor is a function $\beta(h)$ of the length of the period. To go from discrete time to continuous time, we will take limits as the length of the period goes to zero.

Second, we will now model the arrival of wage offers as a stochastic process. We will assume that an unemployed worker has probability λh of receiving an offer during the current period. In the limit, as h goes to zero, offer arrivals follow a Poisson process with parameter λ, which can be interpreted as the instantaneous probability of receiving an offer.

The solution procedure is similar to that used earlier. The value of accepting a job that pays salary x per unit of time is given by

$$W_a(x) = \sum_{t=0}^{\infty} \beta(h)^t xh = \frac{xh}{1-\beta(h)} \tag{1}$$

and the value of rejecting it, W_r, is still independent of x. The reservation wage x^* is the salary that makes the agent indifferent between accepting and rejecting employment and therefore satisfies

$$W_a(x^*) = W_r$$
$$\Rightarrow x^* = \frac{1-\beta(h)}{h} W_r \tag{2}$$

To characterize W_r, consider the prospects of an unemployed worker, which are now slightly more complicated by the fact that he no longer knows when the next offer will arrive. During the current period, his only

In the second case, his value next period will again be W_r. In the first case, his payoff next period will be given by $v(x) = \max[W_a(x), W_r]$, but because the realization of x is not known today, we have to compute the expected return. Finally, all values accruing tomorrow must be discounted by one period. Hence, the expected value of being currently unemployed is given by

$$W_r = bh + \beta(h)\{\lambda h E \max[W_a(x), W_r] + (1-\lambda h)W_r\} \tag{3}$$

The next step is to manipulate this expression so that we can substitute it into the right-hand side of (2). Subtracting $\beta(h)W_r$ from both sides of (3),

$$[1-\beta(h)]W_r = bh + \beta(h)E\{\lambda h \max[W_a(x), W_r] - \lambda h W_r\}$$
$$= bh + \beta(h)\lambda h E\{E \max[W_a(x) - W_r, 0]\}$$

and dividing by h,

$$\frac{1-\beta(h)}{h} W_r = b + \beta(h)\lambda E\{\max[W_a(x) - W_r, 0]\} \tag{4}$$

Substituting (4) into (3) and simplifying, we could obtain a reservation-wage equation very similar to the one in the preceding section. Instead, let us go to continuous time. For this, let the discount factor be of the form $\beta(h) = e^{-\rho h}$. Then we have (using L'Hôpital's rule in the second expression)

$$\lim_{h\to\infty} \beta(h) = 1 \quad \text{and} \quad \lim_{h\to 0} \frac{1-\beta(h)}{h} = \rho \tag{5}$$

Taking limits as $h \to 0$, (1) yields $W_a(x) = x/\rho$, (2) becomes

$$x^* = \rho W_r \tag{2'}$$

and (4) implies[2]

$$\rho W_r = b + \lambda E\{\max[W_a(x) - W_r, 0]\} \tag{4'}$$

Substituting (4′) into (2′) and proceeding as in the preceding section, we obtain the reservation-wage equation:

$$x^* = \rho W_r = b + \lambda \int_0^{\infty} \max[W_a(x) - W_r, 0]\, dF(x)$$

Now, if $x < x^*$, the agent rejects the offer, that is, $W_a(x) < W_r$, and therefore $\max[W_a(x) - W_r, 0] = 0$. On the other hand, if $x > x^*$, then $W_a(x) > W_r$, and therefore $\max[W_a(x) - W_r, 0] = W_a(x) - W_r$. Hence, we can break up the domain of integration into two parts, $(0, x^*)$ and (x^*, ∞), and observing that the integral over the first interval vanishes, we have

Finally, substituting $W_r = x^*/(\rho)$ and $W_a(x) = x/(\rho)$ in this expression, we obtain the fundamental reservation-wage equation:

$$x^* = b + \frac{\lambda}{\rho}\int_{x^*}^{\infty}(x - x^*)\, dF(x) \tag{R}$$

This equation has an intuitive interpretation. Rearrange it to get

$$x^* - b = \frac{\lambda}{\rho}\int_{x^*}^{\infty}(x - x^*)\, dF(x)$$

Then the left-hand side measures the immediate opportunity cost of rejecting an offer, and the right-hand side gives the present value of the expected gain from continued search. The reservation wage, by definition, equates the two quantities.

It is straightforward to do comparative statics using this expression. Write

$$H(x^*; b, \lambda, \rho) = x^* - b - \frac{\lambda}{\rho}\int_{x^*}^{\infty}(x - x^*)\, dF(x) = 0$$

and compute the partial derivatives of $H(\;)$:[3]

$$H_{x^*} = 1 - \frac{\lambda}{\rho}\left(\int_{x^*}^{\infty}(-1)\, dF(x) - (x^* - x^*)F'(x^*)\right)$$

$$= 1 + \frac{\lambda}{\rho}\int_{x^*}^{\infty} dF(x) = 1 + \frac{\lambda}{\rho}[1 - F(x^*)] > 0$$

$$H_b = -1 < 0$$

$$H_\lambda = -\frac{1}{\rho}\int_{x^*}^{\infty}(x - x^*)\, dF(x) < 0$$

$$H_\rho = \frac{\lambda}{\rho^2}\int_{x^*}^{\infty}(x - x^*)\, dF(x) > 0$$

By the implicit-function theorem,

$$\frac{\partial x^*}{\partial b} = -\frac{H_b}{H_{x^*}} > 0, \quad \frac{\partial x^*}{\partial \lambda} = -\frac{H_\lambda}{H_{x^*}} > 0, \quad \text{and} \quad \frac{\partial x^*}{\partial \rho} = -\frac{H_\rho}{H_{x^*}} < 0$$

That is, an increase in the unemployment benefit leads to an increase in the reservation salary, as workers can now afford to wait longer for a better offer (an increase in b reduces the opportunity cost of rejecting any offer). An increase in λ means that jobs become less scarce, and it has a similar effect (the expected cost of rejecting an offer is now lower because the expected delay until a new one arrives is shorter). Finally, an increase in ρ means that future benefits are discounted at a higher rate (agents are less patient); because the expected benefits of continued search will accrue in

(b) A Search-Based Macro Model

Standard neoclassical models implicitly rely on the Walrasian auctioneer to perform two crucial tasks. One is setting prices so that markets will clear continuously. The second can be called trade coordination: The auctioneer is assumed to provide clearing services that will make it unnecessary for the parties to a transaction to physically locate each other, thus simplifying the task of matching desired quantities. In short, these models assume that the allocation of resources is a costless and frictionless process. One implication of this assumption, if we take it literally, is that there is no room for involuntary unemployment. Extensions of the neoclassical model can generate fluctuations in employment levels as agents adjust their labor supply in response to price or productivity shocks, but the labor market must clear continuously, like any other market.

Search models do away with the trade-coordination function of the auctioneer and explicitly model the fact that many transactions must take place between individuals who must first find each other. Trade thus becomes a costly and time-consuming process. Applied to labor markets, this kind of model leads to the emergence of frictional unemployment, for agents will be inactive during some of the time that they wait for an acceptable job.

Moreover, this view of the process of resource allocation naturally suggests an important externality associated with the exchange technology: It seems likely that the greater the number of people who want to trade at any given time, the easier it will be for each of them to locate a suitable partner. Loosely speaking, because an increase in the level of economic activity makes it easier for the parties to an exchange to find each other, individual decisions have external effects over the opportunities available to other agents. One result of this phenomenon is that the equilibrium will not be Pareto-optimal, as agents will fail to take into account the external effects of their actions. Another implication is the possibility of multiple equilibria, as either pessimistic or optimistic expectations tend to become self-fulfilling. Thus, there is a role for government policy, both in correcting for externalities and in helping the economy select a good equilibrium. Policy may be useful as a device for improving coordination between agents in a way the market cannot achieve because of the presence of external effects.

The search model has served as a framework for some contributions to a literature which shows that macro models with "Keynesian" properties can be built from solid micro foundations. The remainder of this section develops one such model, due to Diamond (1982) and Diamond and Fudenberg (1989), in which an agent must first search for production opportunities and optimally low level of economic activity may arise as a result of the difficulty of coordinating exchange in an economy with many agents.

Diamond's Search Model

Imagine a tropical island inhabited by infinitely-lived natives who walk around the beaches looking for coconut trees (production opportunities). Having found a tree, an agent must decide whether or not to climb it. If he does, he comes down with a coconut, but he is not finished yet: An ancient taboo forbids the consumption of one's own coconuts. Hence, the agent must find another native with whom to trade coconuts (one for one) before eating.[4] Having done this, he continues to search for additional production opportunities.

All trees have exactly one piece of fruit, but they may differ in height (production cost). Consumption of a coconut yields utility y. Production costs (the disutility of climbing) are proportional to the height of the tree, which is a nonnegative random variable, c, bounded below by $\underline{c}$ and drawn from a known distribution with cdf $G(\ )$. That is, $G(x) = \text{pr}(c \le x)$, and $G(\underline{c}) = 0$. Agents maximize the expected value of discounted lifetime utility,

$$V = E\sum_{i=0}^{\infty} e^{-\rho t_i} U_{t_i}, \quad \text{where} \quad U_{t_i} = y_{t_i} - c_{t_i}$$

Notice that although time is continuous, production and consumption take place at discrete intervals. At a given time t_i, the agent may be engaged in production (climbing a tree), in which case his instantaneous utility is $-c$, in eating (with utility y), or in doing neither, in which case his instantaneous utility is zero.

An agent who is not engaged in production or consumption may be in either of two states. We will say that he is unemployed if he is looking for a production opportunity and that he is employed if he is carrying a coconut and is looking for someone with whom to trade. The arrivals of production opportunities and trading partners follow Poisson processes, with parameters that are taken as given by each individual agent. We will denote by a the instantaneous probability of finding a tree, and by $b(e)$ the instantaneous probability of finding a trading partner.

A crucial assumption of the model is that b is an increasing function of the aggregate employment rate e. That is, the larger the number of people who are walking around with coconuts in their hands, the easier it will be for them to bump into each other. We will assume that

$$b(0) = 0, \quad b'(e) > 0, \quad \text{and} \quad b''(e) < 0$$

of activity will be suboptimally low. As we will see, the externality is also at the root of the possibility of multiple equilibria, for it makes both optimism and pessimism potentially self-fulfilling. For example, if most agents believe that trading will be easy, they will have an incentive to climb even relatively high trees. If they all do, then finding a trading partner will indeed be easy, thus validating ex post their initial optimism.

Production Decisions. The only decision that an agent has to take in the model is whether or not to climb a tree he has just run into. As in the job-search model, the decision rule takes the form of a reservation level: Agents will accept all of those production opportunities whose cost is smaller than some critical level c^* (i.e., natives will climb all sufficiently low trees).

To characterize the reservation cost, we will proceed as before, beginning with a discrete-time version of the model and then taking limits as the duration of the period, h, goes to zero. In what follows, then, the relevant transition probabilities will be ah and bh for one period, and the one-period discount factor will be given by $\beta(h) = e^{-\rho h}$.

Denote by $W_e(e)$ the expected lifetime utility of an employed worker when the employment rate is equal to e, and by $W_u(e)$ the value of being unemployed given e,[5] and consider the situation of an employed worker at time t. With probability $b(e_t)h$ he will find a trading partner during the current period, consume his coconut (earning utility y), and then become unemployed. With probability $1 - b(e_t)h$ he will be unable to consume and will remain employed. Thus, his expected payoff is given by

$$W_e(e_t) = bh[y + \beta(h)W_u(e_{t+h})] + (1 - bh)\beta(h)W_e(e_{t+h})$$

where we have taken into account the fact that from this period to the next (which starts at $t + h$) the employment rate may change, altering the expected values of both employed and unemployed agents. Subtracting $\beta(h)W_e(e_t)$ from both sides of the foregoing expression and dividing both sides by h, we obtain

$$\begin{aligned}
[1 - \beta(h)]W_e(e_t) &= bhy + \beta(h)[bhW_u(e_{t+h}) + (1 - bh)W_e(e_{t+h}) - W_e(e_t)] \\
&= bhy + \beta(h)[bh[W_u(e_{t+h})] - W_e(e_{t+h})] + [W_u(e_{t+h}) - W_e(e_t)] \\
\Rightarrow \frac{1 - \beta(h)}{h} W_e(e_t) &= by + \beta(h)\left(b[W_u(e_{t+h}) - W_e(e_{t+h})] + \frac{W_e(e_{t+h}) - W_e(e_t)}{h}\right)
\end{aligned}$$

Taking the limit on both sides of this expression as h goes to zero, expectations may not be sufficient to fully close the model: Even if agents know the structure of the model and can compute the equilibrium paths, there is uncertainty concerning the actual path of the economy, for agents cannot know for sure which equilibrium will be selected.

2. Optimal Growth in Discrete Time

Consider an economy populated by a constant number of identical infinitely-lived agents. There is a single good that can be consumed directly or used as capital in production. The preferences of a representative individual are described by a utility function of the form

$$\sum_{t=0}^{\infty} \beta^t U(c_t) \tag{1}$$

where $\beta \in (0, 1)$ is the rate of time discount, a measure of the agent's "impatience," c_t is consumption at time t, and the period utility function $U(\ )$ is a strictly increasing and strictly concave C^2 function. All agents are endowed with one unit of labor time each period.

Production of the single good requires both labor (L) and capital (K). The production technology is described by a strictly concave production function,

$$Y = F(K, L)$$

where we interpret Y as gross output (i.e., new production plus undepreciated capital).[8] We assume that $F(\ )$ is C^2 and is strictly increasing and exhibits constant returns to scale (i.e., is homogeneous of degree 1). Thus, if both inputs are changed by the same factor λ, output changes also by a factor of λ, and we have

$$F(\lambda K, \lambda L) = \lambda F(K, L) \tag{2}$$

This property of the production function allows a convenient normalization. In (2), let $\lambda = 1/L$, and note that

$$F(K/L,\ L/L) = (1/L)F(K, L) \Rightarrow F(K, L) = LF(K/L, 1)$$

If we write k for the per-capita capital stock (K/L) and define the per-capita production function by

$$f(k) \equiv F(k, 1) \tag{3}$$

we can write total output as

$$y = f(k) \tag{4}$$

Imagine that this economy is regulated by a benevolent, all-powerful social planner who makes production, consumption, and investment decisions so as to maximize the lifetime utility of the representative individual. The planner chooses a sequence $\{c_t, k_{t+1}\}_{t=0}^{\infty}$ of consumption levels and capital stocks so as to maximize the utility function (1), taking as given the production technology, and subject to a resource-availability constraint. Working in per-capita terms, the initial capital stock k_0 is given, and at each point in time, consumption and investment must satisfy the constraint

$$f(k_t) = c_t + k_{t+1} \tag{5}$$

That is, current output per capita, including undepreciated capital, $f(k_t)$, can be either consumed today or used for tomorrow's production.

At any given point in time t, the initial capital stock k_t describes completely the state of the system and determines the economy's consumption possibilities for the current period and all future time. Given k_t, the planner's immediate concern is to choose current consumption. Alternatively, because $k_{t+1} + c_t$ must add up to current output, we can think of the planner as choosing an investment level k_{t+1}. Hence, the planner's problem can be written

$$V(k_0) = \max_{\{k_{t+1}\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \beta^t U[f(k_t) - k_{t+1}] \text{ s.t. } 0 \le k_{t+1} \le f(k_t),\ k_0 \text{ given} \right\} \tag{P}$$

The constraint says that next period's capital stock cannot be negative and cannot exceed current gross output. To rule out corner solutions, we will assume that both the production function and the period utility function satisfy the following conditions:

$$f(0) = 0, \quad f'(0) = \infty, \quad f'(\infty) = 0, \quad U'(0) = \infty, \quad \text{and} \quad U'(\infty) = 0 \tag{6}$$

Following our discussion in Chapter 12, the (current) value function for the planner's problem satisfies the Bellman equation,

$$V(k_t) = \max_{k_{t+1}} \{U[f(k_t) - k_{t+1}] + \beta V(k_{t+1}) \text{ s.t } 0 \le k_{t+1} \le f(k_t)\} \tag{BE.P}$$

Under our assumptions regarding preferences and technology, all but one of the conditions that would guarantee the existence and uniqueness of a bounded, continuous, strictly increasing and strictly concave solution to (BE.P) are satisfied. In particular, recall that Theorem 1.5 in Chapter 12 required the period return function to be bounded. In the current context, however, the period utility function $U(\ )$ and the production function

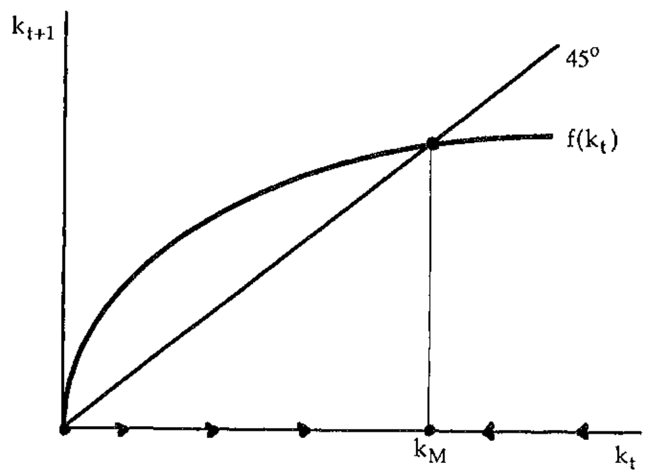

Figure 13.6.

Imagine, for a moment, that consumption is zero in all periods. Then the evolution of the capital stock is described by the difference equation

$$k_{t+1} = f(k_t) \tag{7}$$

It is easy to show (see the discussion of the Solow model in Chapter 11) that under our assumptions, the phase diagram for this equation is as shown in Figure 13.6, with a unique and globally stable steady state, k_M. Hence, even if all output is invested each period, there is a maximum sustainable per-capita capital stock. We can therefore restrict ourselves to values of k in the interval $[0, k_M]$. Because $U[f(k)]$ is certainly bounded in this set, we can apply Theorems 1.5 and 1.18 in Chapter 12 to obtain the following result.

Proposition 2.1. The Bellman equation (BE.P) has a unique continuous and bounded solution V. *This function is the value function for the planner's problem (P) and is strictly increasing and strictly concave. Moreover, the policy correspondence* g( ) *giving next period's optimal capital stock as a function of today's state* k_t *is a well-defined and continuous function.*

Given this result, we can establish some important properties of the policy function by studying the maximization inside the Bellman equation. We begin by using Theorem 2.15 in Chapter 6 to show that $V(\,)$ is differentiable. This will allow us to use the first-order condition for the maximization in (BE.P) to characterize the optimal investment decision.

Proof. Fix some k_t^0 in $(0, k_M)$, and let k_{t+1}^0 be a solution of the problem

$$V(k_t^0) = \max_{k_{t+1}}\{U[f(k_t^0) - k_{t+1}] + \beta V(k_{t+1}) \text{ s.t. } 0 \le k_{t+1} \le f(k_t^0)\} \quad \text{(BE.P)}$$

Next, define the function

$$W(k_t) = U[f(k_t) - k_{t+1}^0] + \beta V(k_{t+1}^0)$$

for k_t within some ε-ball with center at k_t^0, $B_\varepsilon(k_t^0)$. Under assumption (6), k_{t+1}^0 will be an interior solution of this problem, that is, $0 < k_{t+1}^0 < f(k_t^0)$. By the continuity of f, ε can be chosen small enough that $f(k_t) > k_{t+1}^0$ for all $k_t \in B_\varepsilon(k_t^0)$, that is, so that k_{t+1}^0 is still feasible for all $k_t \in B_\varepsilon(k_t^0)$. On the other hand, k_{t+1}^0 is not necessarily optimal for an arbitrary k_t in $B_\varepsilon(k_t^0)$. Hence,

$$W(k_t) = U[f(k_t) - k_{t+1}^0] + \beta V(k_{t+1}^0) \le \max_{k_{t+1}}\{U[f(k_t) - k_{t+1}] + \beta V(k_{t+1})\} = V(k_t)$$

for all $k_t \in B_\varepsilon(k_t^0)$, and

$$W(k_t^0) = V(k_t^0)$$

because k_{t+1}^0 is optimal for k_t^0. Moreover, $W(\,)$ is a differentiable function of k_t, because $U(\,)$ and $f(\,)$ are differentiable, and $V(k_{t+1}^0)$ is just a constant. Hence, by Theorem 2.15 in Chapter 6, $V(\,)$ is differentiable at k_t^0, and

$$V'(k_t^0) = W'(k_t^0) = U'[f(k_t^0) - k_{t+1}^0]f'(k_t^0) \qquad \square$$

Because $V(\,)$ is differentiable, an interior solution of the maximization inside the Bellman equation is characterized by the first-order condition

$$U'[f(k_t) - k_{t+1}] = \beta V'(k_{t+1}) \tag{8}$$

which implicitly defines the policy function

$$k_{t+1}^* = g(k_t)$$

Without additional restrictions there will be no guarantee that V will be twice differentiable. Hence, we cannot differentiate (8) again to establish the comparative-statics properties of the function $g(\,)$. As we will see, however, equation (8) and the concavity of the value function provide sufficient information to establish some important properties of the policy function and the optimal sequence of capital stocks.

In some cases it will be useful to rewrite (8) in an alternative way. By Proposition 2.2, applied at time $t+1$, we have that

$$V'(k_{t+1}) = U'[f(k_{t+1}) - k_{t+2}]f'(k_{t+1}) = U'(c_{t+1})f'(k_{t+1}) \tag{9}$$

$$U'(c_t) = \beta U'(c_{t+1}) f'(k_{t+1}) \tag{10'}$$

To interpret this equation, consider reducing period-t consumption by one unit in order to invest it and increase consumption at $t+1$. On the one hand, there is a utility loss of $U'(c_t)$ in period t. On the other, an additional unit of investment will allow consumption to be higher by $f'(k_{t+1})$ units next period, yielding a utility gain of $U'(c_{t+1})f'(k_{t+1})$. Because this utility gain comes one period later, however, we must discount it by β. The Euler equation says that along an optimal path, today's loss and tomorrow's gain must be equal, for otherwise a feasible rearrangement of the consumption/investment plan would increase its total value, implying that the original plan could not have been optimal. Hence, along an optimal trajectory, the planner must be indifferent, at the margin, between using an additional unit of output for current consumption or for investment.

There are now two different ways to proceed. One is to work directly with the first-order condition (8); the other is to analyze the two-equation system formed by the Euler equation (10′) and the constraint (5), reinterpreted as the law of motion for the capital stock,

$$k_{t+1} = f(k_t) - c_t \tag{5'}$$

We will work through the first approach and let the reader explore the second approach through a series of problems.

(a) Properties of the Policy Function and the Optimal Capital Sequence

Given the policy function $g(\ )$, the optimal time path for the capital stock is the solution of the difference equation $k_{t+1} = g(k_t)$. We know that the optimal sequence, $\{k_t^*\}$, must satisfy the first-order condition (8) and the Euler equation (10) and that the value function $V(\ )$ is strictly concave and increasing. In this section, we will use this information to establish some properties of $g(\ )$ and $\{k_t^*\}$.

We begin by characterizing the steady state of the system. Setting $k_t = k_{t+1} = k_{t+2} \equiv k$ in the Euler equation (10), we obtain

$$U'[f(k)-k] = \beta U'[f(k)-k]f'(k)$$
$$\Rightarrow f'(k) = 1/\beta \tag{11}$$

an equation that implicitly defines the steady-state capital stock $\bar{k}$ as a function of the discount rate β.[9] Because $f(\ )$ is strictly concave, the marginal product of capital, $f'(k)$, is a strictly decreasing function of the capital stock, implying that equation (11) has at most one solution. The assumptions that

$\bar{k}$. Moreover, we have $\bar{k} \le k_M$, as $\bar{k}$ cannot be larger than the maximum sustainable capital stock described earlier.

Next, we show that the policy function $g(\ )$ is an increasing function of k_t. This result is then used to establish that the optimal sequence of capital stocks $\{k_t^*\}_{t=0}^{\infty}$ is monotonic and converges asymptotically to the steady state for any given initial stock $k_0 > 0$.

Proposition 2.3. The policy function $k_{t+1}^* = g(k_t)$ *is increasing in* k_t.

Proof. By contradiction. Suppose $g(\ )$ is not increasing everywhere. Then there exist capital stocks k' and k'' such that $k'' > k'$ and

$$g(k'') < g(k') \tag{1}$$

Because $V(\ )$ is concave, moreover, $V'(\ )$ is decreasing, and (1) implies

$$V'[g(k'')] > V'[g(k')] \tag{2}$$

By the first-order condition

$$U'[f(k_t) - k_{t+1}] = \beta V'(k_{t+1}) \tag{8}$$

inequality (2) implies

$$U'[f(k'') - g(k'')] > U'[f(k') - g(k')]$$

Now, Because $U(\ )$ is strictly concave by assumption, the foregoing expression implies that

$$f(k'') - g(k'') < f(k') - g(k') \Rightarrow g(k'') - g(k') > f(k'') - f(k') > 0$$

where the last inequality holds because $f(\ )$ is increasing. But then $g(k'') > g(k')$, which contradicts (1). □

Proposition 2.4. The optimal capital sequence $\{k_t^*\}$, *defined recursively by* $k_{t+1}^* = g(k_t^*)$, *with* k_0 *given, is monotonic.*

Proof. Suppose $k_1^* > k_0$. Because $g(\ )$ is increasing, we have

$$k_2^* = g(k_1^*) \ge g(k_0) = k_1^*$$

which implies, in turn,

$$k_3^* = g(k_2^*) \ge g(k_1^*) = k_2^*$$

and so forth. Similarly, if $k_1^* < k_0$, then

$$k_2^* = g(k_1^*) \le g(k_0) = k_1^*$$

Proposition 2.5. If the initial capital stock k_0 *is above the steady state* $\bar{k}$, *then* $\{k_t^*\}$ *decreases monotonically; if* $k_0 < \bar{k}$ *then* $\{k_{t+1}^*\}$ *increases monotonically.*

Proof. Because $V(\ )$ is strictly concave, $V'(\ )$ is strictly decreasing. Hence

$$k'' > k' \Rightarrow V'(k'') < V'(k') \tag{1}$$

Consider two successive capital stocks, k_t^* and k_{t+1}^*, where $k_{t+1}^* = g(k_t^*)$. By (1), $k_t^* - k_{t+1}^*$ and $V'(k_t^*) - V'(k_{t+1}^*)$ will have opposite signs, that is,

$$(k_t^* - k_{t+1}^*)[V'(k_t^*) - V'(k_{t+1}^*)] \leq 0 \qquad (= 0 \text{ at the steady state}) \tag{2}$$

By equations (8) and (9), we have

$$(8) \Rightarrow V'(k_{t+1}^*) = (1/\beta)U'[f(k_t^*) - k_{t+1}^*]$$
$$(9) \Rightarrow V'(k_t^*) = U'[f(k_t^*) - k_{t+1}^*]f'(k_t^*)$$

Substituting these expressions in (2),

$$(k_t^* - k_{t+1}^*)\{U'[f(k_t^*) - k_{t+1}^*]f'(k_t^*) - (1/\beta)U'[f(k_t^*) - k_{t+1}^*]\} \leq 0$$

and, dividing by $U'(\ ) > 0$,

$$(k_t^* - k_{t+1}^*)[f'(k_t^*) - (1/\beta)] \leq 0 \tag{3}$$

Recall that at the steady state, $f'(\bar{k}) = 1/\beta$, and $f'(\ )$ is decreasing, by the concavity of $f(\ )$. Hence,

- if $k_t^* < \bar{k}$, we have $f'(k_t^*) > (1/\beta)$, and (3) implies that $k_t^* \leq k_{t+1}^*$, that is, $\{k_t^*\}$ is increasing, and
- if $k_t^* > \bar{k}$, we have $f'(k_t^*) < (1/\beta)$, and (3) implies that $k_t^* \geq k_{t+1}^*$, that is, $\{k_t^*\}$ is decreasing. □

Proposition 2.6. The optimal capital sequence $\{k_t^*\}$ *converges (monotonically) to the steady-state capital stock* $\bar{k}$ *for any initial* $k_0 > 0$.

Proof. Note that $\{k_t^*\}$ is monotonic and bounded (above by k_M, below by zero or, alternatively, by k_0 and $\bar{k}$). Because every monotonic bounded sequence converges, $\{k_t^*\}$ has a limit that we will call k^*. By the continuity of the policy function $g(\ )$, k^* must be a fixed point of $g(\ )$, for

$$k^* = \lim_{t\to\infty} k_{t+1}^* = \lim_{t\to\infty} g(k_t) = g\left(\lim_{t\to\infty} k_t\right) = g(k^*)$$

Hence, k^* is a steady state. Because there is a unique steady state $\bar{k}$, we conclude that $\{k_t^*\} \to \bar{k}$. □

(b) The Euler Equation and Dynamics

value function and the first-order condition for the maximization in the Bellman equation, we have established that the optimal capital sequence $\{k_t^*\}$ converges monotonically to the steady state of the system. We can then use the constraint again to infer the optimal path of consumption over time. We now illustrate a second and probably more instructive approach to analyzing the dynamics of the optimal-growth model. The basic idea is to treat the system formed by the Euler equation and the transition law for the capital stock,

$$U'(c_t) = \beta U'(c_{t+1})f'(k_{t+1}) \tag{11}$$
$$k_{t+1} = f(k_t) - c_t \tag{12}$$

as an ordinary system of difference equations and study its dynamics in the standard way. Thus, we first solve for the steady state; then we construct a phase diagram and compute the eigenvalues of the Jacobian matrix at the steady state to check for stability.

Setting $c_t = c_{t+1} \equiv c$ and $k_t = k_{t+1} \equiv k$ in (11) and (12), we get

$$(11) \Rightarrow U'(c) = \beta U'(c)f'(k) \Rightarrow \beta f'(k) = 1 \tag{13}$$
$$(12) \Rightarrow c = f(k) - k \tag{14}$$

As we have seen, equation (13) has a unique solution $\bar{k}$. Given $\bar{k}$, equation (14) can be solved for steady-state consumption $\bar{c}$.

The system (11)–(12) is not quite in the "standard form." In particular, we would like to have each variable (k_{t+1} and c_{t+1}) as a function of the lagged values k_t and c_t. To this end, we solve (12) for k_{t+1}, substitute the result into (11), and apply the implicit-function theorem to the resulting equation to obtain a function $\phi(\ )$ giving c_{t+1} as a function of k_t and c_t. This yields the system

$$k_{t+1} = f(k_t) - c_t \equiv \varphi(k_t, c_t) \tag{15}$$
$$U'(c_t) = \beta U'(c_{t+1})f'[f(k_t) - c_t] \Leftrightarrow c_{t+1} = \phi(k_t, c_t) \tag{16}$$

Problem 2.7. Apply the implicit-function theorem to compute the partial derivatives of the function $\phi(k_t, c_t)$ defined implicitly by equation (16), and determine their sign.

Problem 2.8. Setting $c_t = c_{t+1} = c$ and $k_t = k_{t+1} = k$ in (15) and (16), draw the phase lines $\Delta k_t = 0$ and $\Delta c_t = 0$. To complete the phase diagram, determine the directions of motion along the c and k axes in each of the four regions into which the state plane (c, x) is divided by the phase lines.

eigenvalues of the Jacobian matrix for the system are positive real numbers lying on opposite sides of 1.

The phase diagram we have constructed shows the orbits of the system (15)–(16), but only one of these trajectories corresponds to the solution of the original planning problem. These two equations can be thought of as the first-order conditions for an optimum, but they are not sufficient to fully characterize the optimal path.

Out of all the solutions of (15)–(16), we want to identify the one that corresponds to the solution of the programming problem. To select one particular solution, we need two boundary conditions to pin down one point in the phase plane through which the system will have to go. The initial value of the capital stock should be taken as given; this yields one initial condition, $k(t = 0) = k_0$, a given constant, specifying that the system starts out from some point on a vertical line through k_0 in the phase plane. On the other hand, there is no natural initial condition for the free variable c, so we need another way to identify the optimal path.

It turns out that the optimal consumption/investment plan is the one described by the saddle-path trajectory. An intuitive way to see this is by examining the phase diagram for the system after adding to it a feasibility bound requiring that consumption not exceed current output, that is, $c \leq f(k)$. Inspection of this figure suggests that all trajectories other than the saddle path eventually run into either the k axis or the feasibility bound, where present consumption exhausts output, leaving nothing for next period. In either case, consumption becomes zero and remains so thereafter. It is clear that such paths cannot be optimal, leaving us with only the saddle path.

A more formal way to identify the optimal path is through a so-called *transversality condition*. In some sense, the problem is the same as in a static maximization problem: The first-order conditions (the Euler equations here) identify possible candidates for a maximum, but they are also satisfied by points that are not maxima. To find an optimum, we need an additional criterion, some sort of second-order condition relating to the concavity of the objective function at the candidate point. The transversality condition plays a similar role in the present context, and as we will see, the sufficiency proof relies heavily on the concavity of the objective function.

An alternative way to think of the transversality condition is as a terminal condition for the system of difference equations. Consider first a finite-horizon version of the planning problem we are studying. In that case k_{T+1} is the capital stock to be left "at the end of time"; it is clear that the optimal thing to do is to leave nothing, so $k_{T+1} = 0$, providing us with a second boundary condition to identify the particular solution of (15)–(16) that solves the

transversality condition can be interpreted in a somewhat similar way, as the requirement that as $t \to \infty$ the suitably discounted value of the capital stock should go to zero. Intuitively, we want to prevent the planner from accumulating too much capital at the expense of deferring consumption forever.

Proposition 2.10. Transversality condition. Let $s^* = k_0 \cup \{c_t^*, k_{t+1}^*\}_{t=0}^{\infty}$ *be a solution sequence of the system (15)–(16). If this sequence satisfies the transversality condition*

$$\lim_{T\to\infty} \beta^T U'(c_T) f'(k_T) k_T = 0 \qquad (T)$$

then it solves the planner's problem.

Proof. Let $s^* = k_0 \cup \{c_t^*, k_{t+1}^*\}$ be a sequence satisfying the conditions of the proposition, and $s = k_0 \cup \{c_t, k_{t+1}\}$ an arbitrary feasible sequence. To establish that s^* is optimal, we show that

$$d = W_0(s^*) - W_0(s) = \sum_{t=0}^{\infty} \beta^t [U(c_t^*) - U(c_t)] \geq 0$$

That is, the total "utility value" of the candidate sequence s^* is at least as large as that of any feasible sequence.

To show this, it will be convenient to solve the resource constraint for c,

$$k_{t+1} = f(k_t) - c_t \Rightarrow c_t = f(k_t) - k_{t+1}$$

and write the period utility function as

$$U(k_t, k_{t+1}) = U[f(k_t) - k_{t+1}]$$

It is easy to show that the function $U(k_t, k_{t+1})$ is concave. Moreover, we have

$$U_1(k_t, k_{t+1}) = U'[f(k_t) - k_{t+1}] f'(k_t) > 0 \qquad (1)$$

$$U_2(k_t, k_{t+1}) = -U'[f(k_t) - k_{t+1}] < 0 \qquad (2)$$

In this notation the Euler equation can be written

$$U'[f(k_t) - k_{t+1}] = \beta U'[f(k_{t+1}) - k_{t+2}] f'(k_{t+1})$$
$$\Rightarrow U_2(k_t, k_{t+1}) + \beta U_1(k_{t+1}, k_{t+2}) = 0 \qquad (3)$$

Next, write d in the form

$$d = W_0(s^*) - W_0(s) = \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \{U(k_t^*, k_{t+1}^*) - U(k_t, k_{t+1})\}$$

and observe that, by the concavity of $U(k_t, k_{t+1})$,

$$U(k_t, k_{t+1}) \leq U(k_t^*, k_{t+1}^*) + U_1(k_t^*, k_{t+1}^*)(k_t - k_t^*) + U_2(k_t^*, k_{t+1}^*)(k_{t+1} - k_{t+1}^*)$$

$$U(k_t^*, k_{t+1}^*) - U(k_t, k_{t+1}) \geq U_1(k_t^*, k_{t+1}^*)(k_t^* - k_t) + U_2(k_t^*, k_{t+1}^*)(k_{t+1}^* - k_{t+1}) \quad (4)$$

Hence, we have

$$d = \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \{U(k_t^*, k_{t+1}^*) - U(k_t, k_{t+1})\}$$
$$\geq \lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \{U_1(k_t^*, k_{t+1}^*)(k_t^* - k_t) + U_2(k_t^*, k_{t+1}^*)(k_{t+1}^* - k_{t+1})\}$$
$$= \beta^0 \{U_1(k_0^*, k_1^*)(k_0^* - k_0) + U_2(k_0^*, k_1^*)(k_1^* - k_1)\}$$
$$+ \beta \{U_1(k_1^*, k_2^*)(k_1^* - k_1) + U_2(k_1^*, k_2^*)(k_2^* - k_2)\} + \ldots +$$
$$\beta^t \{U_1(k_t^*, k_{t+1}^*)(k_t^* - k_t) + U_2(k_t^*, k_{t+1}^*)(k_{t+1}^* - k_{t+1})\}$$
$$+ \beta^{t+1} \{U_1(k_{t+1}^*, k_{t+2}^*)(k_{t+1}^* - k_{t+1}) + U_2(k_{t+1}^*, k_{t+2}^*)(k_{t+2}^* - k_{t+2})\} + \ldots$$

Observe that the initial capital stock is given and thus is the same in both s and s^*; hence $k_0^* - k_0 = 0$, and the first term in the sum vanishes. The remaining terms can be rearranged to give

$$d = \lim_{T\to\infty} \left\{ \sum_{t=1}^{T} \beta^{t-1} \{[U_2(k_{t-1}^*, k_t^*) + \beta U_1(k_t^*, k_{t+1}^*)](k_t^* - k_t)\} + \beta^T U_2(k_T^*, k_{T+1}^*)(k_{T+1}^* - k_{T+1}) \right\} \quad (5)$$

Next, recall that s^* is assumed to satisfy the Euler equation

$$U_2(k_t^*, k_{t+1}^*) + \beta U_1(k_{t+1}^*, k_{t+2}^*) = 0 \quad (3)$$

Hence the terms in the summation vanish, and we have

$$d = W_0(s^*) - W_0(s) \geq \lim_{T\to\infty} \beta^T U_2(k_T^*, k_{T+1}^*)(k_{T+1}^* - k_{T+1})$$

Moreover, we have $k_{T+1} \geq 0$, by the feasibility constraint, and $U_2(\ ) < 0$, by (2). Hence, the product $U_2(k_T^*, k_{T+1}^*)(-k_{T+1})$ is positive, and we have, using the Euler equation,

$$d = W_0(s^*) - W_0(s) \geq \lim_{T\to\infty} \beta^T U_2(k_T^*, k_{T+1}^*) k_{t+1}^*$$
$$= \lim_{T\to\infty} \beta^{T+1} U_1(k_{T+1}^*, k_{T+2}^*) k_{T+1}^* = -\lim_{T\to\infty} \beta^{T+1} U'(c_{T+1}^*) f'(k_{T+1}^*) k_{T+1}^* = 0$$

where the next-to-last equality follows from (1), and the last limit is zero, by the transversality condition (T).

In conclusion, we have shown that

$$d = W_0(s^*) - W_0(s) \geq 0$$

Because the sequence s^* that satisfies both the Euler equation and the transversality condition must yield a greater value than any other feasible sequence, it must be optimal.

Problem 2.11. Show that the function $U(k_t, k_{t+1})$ defined in the proof of Proposition 2.10 is concave.

To conclude, it is easy to verify that the saddle-path solution satisfies the transversality condition (T) and is therefore optimal. For this solution, both c_t and k_t converge to finite values c and k. Hence, k, $U'(c)$, and $f'(k)$ are just finite constants, and

$$\lim_{T\to\infty} \beta^T U'(c) f'(k) k = 0$$

because $\beta \in (0, 1)$. Along explosive paths, however, either k or c will become zero. In that case, $f'(k) \to \infty$ or $U'(c) \to \infty$, so (T) may not hold.

3. Investment with Installation Costs

In the standard static model the firm is assumed to maximize current profits, defined as the difference between output and contemporaneous factor payments. Letting K and L denote labor and capital inputs, and w and R the wage and the rental rate of capital in units of output, the firm's problem can be written

$$\max_{K,L} F(K, L) - wL - RK \quad (1)$$

The solution functions for this problem are factor demands giving optimal input levels as functions of factor prices:

$$K^* = K(w, r) \quad \text{and} \quad L^* = L(w, r)$$

This formulation assumes that the firm can rent inputs in "spot markets" and put them to work immediately and at no cost. This clearly unrealistic assumption may lead, at best, to a theory of the determination of the optimal capital stock, but it has no implications (or very naive ones) for the optimal investment policy.

In practice, capital is typically purchased, rather than rented, and its installation may involve considerable delays and adjustment costs. Thus, a firm's stock of "installed capital" becomes a sluggish state variable, and investment decisions must be made taking into account their effect on the entire time path of profits, rather than on a period-by-period basis.

The first part of this section analyzes the optimal investment policy for a competitive firm when the installation of capital is costly. In the second part, we go from partial to general equilibrium and study the time paths of investment and share prices in a small open economy and their responses to the sequence; second, we verify that this function is bounded and continuous; third, we show that $\{f_n\} \to f$ in the sup norm.

- Given a Cauchy sequence of bounded continuous functions $\{f_n\}$, take some x in X and consider the sequence of real numbers $\{f_n(x)\}$. Note that given any positive integers m and n, we have

$$|f_m(x) - f_n(x)| \leq \sup\{|f_m(y) - f_n(y)|;\ y \in X\} \equiv \|f_m - f_n\|_s$$

Because $\{f_n\}$ is a Cauchy sequence, by choosing m and n high enough we can make $|f_m(x) - f_n(x)|$ arbitrarily small for any x. Hence, $\{f_n(x)\}$ is a Cauchy sequence of real numbers for any x, and because $\mathbb{R}$ is complete with the usual metric, $\{f_n(x)\}$ converges to some (finite) real limit, say $f(x)$.

We can therefore construct a function f that assigns to each x in X the limit $f(x)$ of the sequence of real numbers $\{f_n(x)\}$. This function, which is bounded by construction, will be our candidate for the limit of the sequence of functions $\{f_n\}$.

- To establish the continuity of f, fix an arbitrary point x in X and some $\varepsilon > 0$. Because $\{f_n\} \to f$ in the sup norm, there exists a positive integer N_1 such that $\|f - f_n\|_s < \varepsilon/3$ for all $n > N_1$. Hence,

$$|f_n(x) - f(x)| \leq \sup_y |f(y) - f_n(y)| \equiv \|f - f_n\|_s < \varepsilon/3 \tag{1}$$

for any x and all $n > N_1$. Moreover, because f_n is continuous, there is some $\delta_1 > 0$ such that for the given x,

$$|f_n(x) - f_n(y)| < \varepsilon/3 \quad \text{for all } y \text{ such that } \|x - y\|_E < \delta_1 \tag{2}$$

where $\|\cdot\|_E$ is the Euclidean norm in $\mathbb{R}^n$. Using (1), (2), and the triangle inequality, the continuity of f at x follows: For any $y \in B_{\delta_1}(x)$, and choosing $n > N_1$, we have

$$\begin{aligned}|f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq \|f - f_n\|_s + |f_n(x) - f_n(y)| + \|f - f_n\|_s < \varepsilon\end{aligned}$$

- Finally, we will show that $\|f - f_n\|_s \to 0$ as $n \to \infty$. Fix some $\varepsilon > 0$ and note that because $\{f_n\}$ is Cauchy, there is some N_2 such that

$$\|f_n - f_m\|_s < \varepsilon/2 \quad \text{for all } m, n > N_2 \tag{3}$$

By (3) and the triangle inequality, given any x in X, we have

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

for all $m, n > N_2$. Moreover, because $\{f_m(x)\} \to f(x)$, we can choose m (separately for each x if need be) so that $|f_m(x) - f(x)| < \varepsilon/2$. Hence, N_2 is such that given any $n > N_2$,

Thus, for n sufficiently high, ε is an upper bound for $\{|f_n(x) - f(x)|;\ x \in X\}$, and because $\|f_n - f\|_s$ is the smallest such upper bound, we conclude that $\|f_n - f\|_s \leq \varepsilon$ for all $n > N_2$, that is, $\{f_n\} \to f$. □

(b) Operators and the Contraction Mapping Theorem

A function $T: X \longrightarrow X$ from a metric space to itself is cometimes called an *operator*. We say that an operator is a contraction if its application to any two points of X brings them closer to each other. More formally, we have the following definition:

Definition 7.13. Contraction. Let (X, d) be a metric space, and $T: X \longrightarrow X$ an operator in it. We say that T is a contraction of modulus β if for some $\beta \in (0, 1)$ we have this: $\forall\, x, y \in X,\ d(Tx, Ty) \leq \beta d(x, y)$. The notation Tx is sometimes used instead of $T(x)$.

Theorem 7.14. Every contraction is a continuous mapping.

Proof. Let T be a contraction on (X, d). We want to show that

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ s.th. } d(x, y) < \delta \Rightarrow d(Tx, Ty) < \varepsilon$$

As T is a contraction, we have that for all $x, y \in X$ and some $\beta \in (0, 1)$,

$$d(Tx, Ty) \leq \beta d(x, y)$$

Given some ε, choose δ so that $\delta \leq \varepsilon/\beta$; then the definition of continuity is satisfied, because

$$d(Tx, Ty) \leq \beta d(x, y) < \beta\delta \leq \varepsilon$$ □

Example 7.15. Let $f: [a, b] \longrightarrow [a, b]$ be a continuous function with positive slope always smaller than 1. Then f is a contraction, because $(f(y) - f(x))/(y - x) \leq \beta < 1$. Figure 2.11 suggests that no matter how we draw it, f must cut the 45° line, that is, it must have at least one fixed point z such that $f(z) = z$.

Take any point x_0 in $[a, b]$ and define a sequence $\{x_n(x_0)\}$ recursively by

$$x_1 = f(x_0),\ x_2 = f(x_1), \ldots, x_{n+1} = f(x_n)$$

Graphically, the sequence is constructed as follows: Given the initial value x_0, we use the graph of the function to find the value of x_1; then we

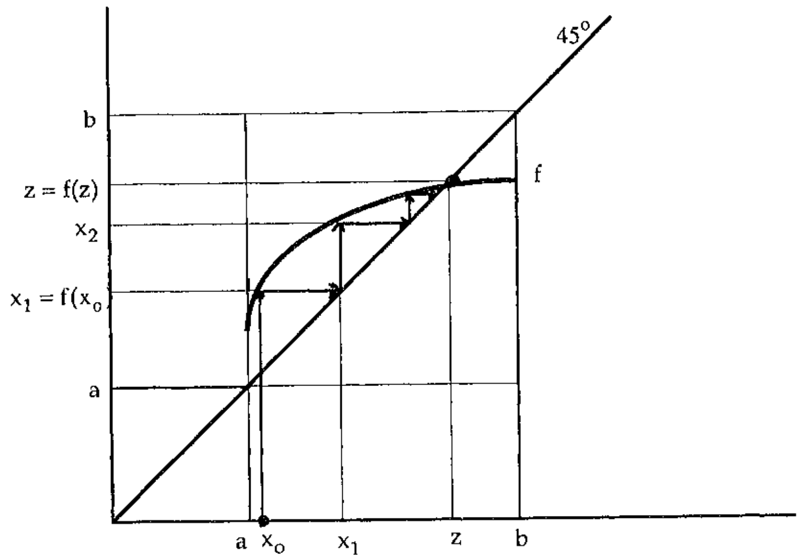

Figure 2.11. A contraction mapping.

the graph of f to find x_2, and so on. The figure suggests that no matter where we choose the initial point x_0 in $[a, b]$, the sequence converges to the fixed point z. □

The following theorem says that this result can be generalized to any contraction defined on a complete metric space.

Theorem 7.16. Contraction mapping theorem. Let (X, d) *be a complete metric space, and* $T: X \longrightarrow X$ *a contraction with modulus* $\beta < 1$. *Then*

(i) T *has precisely one fixed point* x^* *in* X *(i.e.,* $\exists! \, x^* \in X$ *s.th.* $Tx^* = x^*$*), and*
(ii) the sequence $\{x_n(x_0)\}$*, defined by*

$$x_1 = Tx_0, x_2 = Tx_1, \ldots, x_{n+1} = Tx_n$$

converges to x^* *for any starting point* x_0 *in* X.

Proof

- *Existence*: Take an arbitrary point x_0 in X and define the sequence $\{x_n(x_0)\}$ by

$$x_1 = Tx_0, x_2 = Tx_1, \ldots, x_{n+1} = Tx_n$$

We will first show that this sequence is Cauchy. Then, given that (X, d) is a complete metric space, the sequence converges to a point x^* in X. We will then show

By using the definition of contraction repeatedly, we see that the distance between two successive terms of the sequence $\{x_n(x_0)\}$ is bounded and decreasing in n:

$$\begin{aligned} d(x_{n+1}, x_n) &= d(Tx_n, Tx_{n-1}) \leq \beta d(x_n, x_{n-1}) \\ &= \beta d(Tx_{n-1}, Tx_{n-2}) \leq \beta^2 d(x_{n-1}, x_{n-2}) \\ &\leq \ldots \leq \beta^n d(x_1, x_0) \end{aligned} \tag{1}$$

Next, consider the distance between two arbitrary terms of the sequence, x_m and x_n, with $m < n$. Using the triangle inequality,

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=m}^{n-1} d(x_{i+1}, x_i) \qquad \text{[by (1)]} \\ &\leq \sum_{i=m}^{n-1} \beta^i d(x_1, x_0) = \beta^m d(x_1, x_0) \sum_{i=0}^{n-m-1} \beta^i \\ &\leq \beta^m d(x_1, x_0) \sum_{i=0}^{\infty} \beta^i = \frac{\beta^m}{1-\beta} d(x_1, x_0) \end{aligned} \tag{2}$$

Because $\beta < 1$, $\beta^m/(1-\beta) \to 0$ as $m \to \infty$. It follows that, given an arbitrary $\varepsilon > 0$, we can choose m and n sufficiently large that $d(x_m, x_n) < \varepsilon$; hence, $\{x_n(x_0)\}$ is Cauchy for any x_0, and given that (X, d) is complete by assumption, every such sequence will have a limit in X. Take one such point and call it x^*.

Next, we show that x^* is a fixed point of T. Being a contraction, T is continuous. Hence we can "take the limit out of the function" and write

$$T(x^*) = T\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x^*$$

- *Uniqueness*: Nothing we have said so far implies uniqueness. It remains to show that x^* is independent of the choice of the initial point x_0 or, equivalently, that there is only one fixed point of T. We will prove that if T has two fixed points, they must be equal.

Assume that x' and x'' are both fixed points of T (i.e., $Tx' = x'$ and $Tx'' = x''$). Because T is a contraction, we have, for some $\beta \in (0, 1)$,

$$d(x', x'') = d(Tx', Tx'') \leq \beta d(x', x'')$$

Because $\beta < 1$, this can hold only if $d(x', x'') = 0$ (i.e., if $x' = x''$), for otherwise we would arrive at

$$d(x', x'') < d(x', x'')$$

a contradication. □

The following exercise generalizes this result. It is not necessary that T itself be a contraction; it is enough that its nth iteration (T^n) be a contraction for T to have precisely one fixed point. T^n is defined recursively by

Problem 7.17. Let (X, d) be a complete metric space, and $T: X \longrightarrow X$ a function whose nth iteration T^n is a contraction. Show that T has a unique fixed point.

The contraction mapping theorem is a very useful result. It can be used to prove the existence and uniqueness of solutions to several types of equations, including differential equations and some functional equations that arise in connection with dynamic-optimization problems. Moreover, the second part of the theorem suggests a method (the *method of successive approximations*) for calculating solutions to equations that can be written in the form $Tx = x$, where T is a contraction: Beginning with a convenient trial solution, we construct a sequence $\{x_n\}$ recursively with $x_{n+1} = Tx_n$. If we can find the limit of the sequence, we will also have found the solution to the equation. Otherwise, we can approximate the solution to any desired degree of accuracy by computing sufficiently many terms of the sequence.[9]

The following theorem says, loosely speaking, that if a continuity condition holds, we can do comparative statics with fixed points of contractions.

Theorem 7.18. Continuous dependence of the fixed point on parameters. Let (X, d) *and* (Ω, ρ) *be two metric spaces, and* $T(x, \alpha)$ *a function* $X \times \Omega \longrightarrow X$. *If* (X, d) *is complete, if* f *is continuous in* α, *and if for each* $\alpha \in \Omega$ *the function* T_α, *defined by* $T_\alpha(x) = T(x, \alpha)$ *for each* $x \in X$, *is a contraction, then the solution function* $z: \Omega \longrightarrow X$, *with* $x^* = z(\alpha)$, *which gives the fixed point as a function of the parameters, is continuous.*

Proof. Consider a convergent sequence of parameter values, $\{\alpha_n\} \to \alpha$. To establish the continuity of z, it is sufficient to show that

$$d[z(\alpha_n), z(\alpha)] \to 0 \quad \text{as } \{\alpha_n\} \to \alpha \tag{1}$$

By definition, the function z satisfies the identity $T_\alpha z(\alpha) \equiv z(\alpha)$ for any α. Using this expression in (1), we have

$$\begin{aligned} d[z(\alpha_n), z(\alpha)] &= d[T_{\alpha_n} z(\alpha_n), T_\alpha z(\alpha)] \qquad \text{(by the triangle inequality)} \\ &\le d[T_{\alpha_n} z(\alpha_n), T_{\alpha_n} z(\alpha)] + d[T_{\alpha_n} z(\alpha), T_\alpha z(\alpha)] \\ &\le \beta d[z(\alpha_n), z(\alpha)] + d[T_{\alpha_n} z(\alpha), T_\alpha z(\alpha)] \end{aligned}$$

where the second inequality uses the assumption that T_{α_n} is a contraction, with modulus $\beta_n \le \beta \in (0, 1)$. Thus,

from where

$$d[z(\alpha_n), z(\alpha)] \le \frac{1}{1-\beta} d[T_{\alpha_n} z(\alpha), T_\alpha z(\alpha)]$$

Now, T is continuous in α, so the right-hand side of this expression goes to zero as $\{\alpha_n\} \to \alpha$. Hence, (1) holds, and $z(\ )$ is continuous. □

Recall that given a complete metric space (X, d) and a closed subset C of X, (C, d) is also a complete metric space (Theorem 7.9). Now suppose that $T: X \longrightarrow X$ is a contraction and maps C into itself (i.e., if $x \in C$, then $Tx \in C$). In that case, T is a contraction on C, and the unique fixed point of T in X must lie in C. Sometimes this observation allows us to establish certain properties of a fixed point by applying the contraction mapping theorem twice – first in a "large" space X to establish existence, and then again in a closed subset of X in order to show that the fixed point has certain properties. For example, if (X, d) is the space of continuous real and bounded functions with the sup norm (see Section 1), then the subset of X formed by nondecreasing functions is closed. Hence, if a contraction T in (X, d) maps nondecreasing functions into nondecreasing functions, the fixed point of T will be a nondecreasing function.

It is not always easy to determine whether or not a given function is a contraction. The following theorem, due to Blackwell, gives sufficient conditions for an operator in a useful function space to be a contraction. The advantage of this result is that in some economic applications, Blackwell's conditions are very easy to verify.

Theorem 7.19. Blackwell's sufficient conditions for a contraction. Let $B(\mathbb{R}^n, \mathbb{R})$ *be the set of bounded functions* $f: \mathbb{R}^n \longrightarrow \mathbb{R}$, *with the sup norm. If an operator* $T: B(\mathbb{R}^n, \mathbb{R}) \longrightarrow B(\mathbb{R}^n, \mathbb{R})$ *satisfies the two conditions*

(i) *monotonicity:* $\forall\, f, g \in B(\mathbb{R}^n, \mathbb{R})$, $f(x) \le g(x)\ \forall\, x \Rightarrow Tf(x) \le Tg(x)\ \forall x$,
(ii) *discounting:* $\exists\, \beta \in (0, 1)$ *s.th.* $\forall\, f \in B(\mathbb{R}^n, \mathbb{R})$, $x \in \mathbb{R}^n$, *and* $\alpha \ge 0$, *we have* $T[f(x) + \alpha] \le T[f(x)] + \beta\alpha$

then T is a contraction.

Proof. For any $f, g \in B(\mathbb{R}^n, \mathbb{R})$, we have

$$f = g + (f - g) \le g + \|f - g\|$$

By assumptions (i) and (ii),

Interchanging the roles of f and g, we obtain, by the same logic,

$$Tg \leq T(f + \|g - f\|) \leq Tf + \text{ß}\|g - f\| \Rightarrow Tf - Tg \geq -\text{ß}\|f - g\|$$

Combining the two inequalities, we obtain the desired result:

$$\|Tf - Tg\| \leq \text{ß}\|f - g\| \qquad \square$$

8. Compactness and the Extreme-Value Theorem

Let f be a real-valued function defined on some set in a metric space. A problem that frequently arises is that of finding the element of A that will maximize or minimize f. In order to guarantee that such a point exists, certain restrictions have to be placed on both the function and the set. For example, we have seen that if f is a function from a set of real numbers A to $\mathbb{R}$, a sufficient condition for the existence of a maximum is that f be continuous and A be a closed and bounded interval. One of the purposes of this section is to extend this result on continuous functions to more general sets. This brings us to the study of compactness.

(a) Compactness and Some Characterizations

To introduce the notion of compactness, we need some terminology.

Definition 8.1. Cover and open cover. A collection of sets $\mathbb{U} = \{U_i; i \in I\}$ in a metric space (X, d) is a cover of the set A if A is contained in its union, that is, if $A \subseteq \cup_{i \in I} U_i$. If all the sets U_i are open, the collection $\mathbb{U}$ is said to be an open cover of A.

Definition 8.2. Compact set. A set A in a metric space is compact if every open cover of A has a finite subcover. That is, A is compact if given *any* open cover $\mathbb{U} = \{U_i; i \in I\}$ of it, we can find a finite subset of $\mathbb{U}$, $\{U_1, \ldots, U_n\}$, that still covers A.

Notice that the definition does *not* say that a set is compact if it has a finite open cover. In fact, every set in a metric space (X, d) has a finite open cover, for the universal set X is open and covers any set in the space.

Example 8.3. $(0, 1)$ is not compact. The collection of open intervals $(1/n, 1)$ for $n \geq 2$ is an open cover of $(0, 1)$ because given any x in $(0, 1)$, there exists an integer n such that $n > 1/x$, and hence $x \in (1/n, 1)$. Thus,

where $N = \max_{1 \leq i \leq k} n_i$, and given any N there is some strictly positive real number x with $x < 1/N$. $\square$

A necessary prerequisite for the existence of a maximum of a function over a set is that the function be bounded on the set. To motivate the foregoing definition (i.e., to try to understand why sets with such strange properties may be useful), consider how we might go about extending the result given in Theorem 6.20 on the boundedness of a continuous function defined over an interval $[a, b]$ to a larger class of sets in an arbitrary metric space.

We begin by observing that a continuous function is locally bounded. Let $f: A \subseteq X \longrightarrow \mathbb{R}$ be continuous, and consider an arbitrary point a in A. Then, by the definition of continuity (with $\varepsilon = 1$), there exists a positive real number $\delta(a)$ (which depends both on the point chosen and on the particular function f we are working with) such that $|f(x) - f(a)| < 1$ for all $x \in B_{\delta(a)}(a)$. Hence, f is bounded in $B_{\delta(a)}(a)$ by $K_a = |f(a)| + 1$.

Now consider what happens when we try to extend this local boundedness property to the whole set A. The question is whether or not the continuity of f is sufficient to guarantee the existence of a bound K that will work for all x in A (for the given function). It is tempting to try to define K as the maximum of the K_a's over all points a in A, but that will not work in general, for there may be infinitely many such K_a's, and the set of such numbers may not have an upper bound. Notice, however, that the collection of open balls $\{B_{\delta(a)}(a)\}$ for all $a \in A$ is an open cover of A. If A is a compact set, there is a finite collection of such balls, $\{B_{\delta(a_1)}(a_1), \ldots, B_{\delta(a_n)}(a_n)\}$, that contains all points of A. In this case, the maximum of the (finite) set formed by the corresponding local bounds $\{K_{a_1}, \ldots, K_{a_n}\}$ is well defined and provides a global bound for the function on the set.

In conclusion, compactness allows us to replace an arbitrary open cover with a finite one. In some cases this is enough of a substitute for finiteness as to allow us to extend to infinite sets some properties that hold trivially in finite ones.

It is not always easy to work directly with the definition of compactness. In the remainder of this section we will develop some characterizations of compactness that frequently are more useful than our original definition. The first of these, known as sequential compactness, is valid in metric spaces, but not necessarily in more general topological spaces.

Definition 8.4. Sequential compactness. A set A in a metric space is sequen-

We will now show that compactness and sequential compactness (which is essentially the Bolzano-Weierstrass property) are equivalent in metric spaces. The first half of the equivalence is easily established.

Theorem 8.5. A compact set in a metric space is sequentially compact.

Proof. We will prove the contrapositive statement (a set A in a metric space that is not sequentially compact cannot be compact) by constructing an open cover of A with no finite subcover. If A is not sequentially compact, there is a sequence $\{x_n\}$ of points of A with the property that none of its subsequences converges to a point in A. Hence, no point of A is the limit of a subsequence of $\{x_n\}$, and it follows that for every x in A there exists an open ball $B_{\varepsilon(x)}(x)$ that contains only a finite number of elements of $\{x_n\}$. The family $\mathbb{B} = \{B_{\varepsilon(x)}(x); x \in A\}$ is an open cover of A. However, no finite subfamily of $\mathbb{B}$ can cover $\{x_n\}$ (and therefore A), for any such family will contain only a finite number of terms of $\{x_n\}$. Hence, A is not compact. □

The converse result takes a bit more work. We begin with some definitions.

Definition 8.6. ε-net and totally bounded set. Given some $\varepsilon > 0$ and a set A in a metric space (X, d), an ε-net for A is a set of points E in X such that

$$A \subseteq \cup_{x \in E} B_\varepsilon(x)$$

A set A in (X, d) is totally bounded if it has a finite ε-net for any $\varepsilon > 0$.

That is, a set is totally bounded if it can be covered by a finite number of balls of arbitrarily small radius. Clearly, a totally bounded set is necessarily bounded, but the converse need not be true.

Definition 8.7. Lebesgue number for an open cover. Let A be a set in a metric space, and let $\mathbb{U}$ be an open cover of A. We say that a fixed real number $\varepsilon > 0$ is a Lebesgue number for $\mathbb{U}$ if for every x in A there exists a set $U(x)$ in $\mathbb{U}$ such that $B_\varepsilon(x) \subseteq U(x)$.

Hence, if $\mathbb{U}$ has a Lebesgue number, we can "replace" it with an open cover formed by balls of constant radius, which is often more convenient. Notice that if this "ball cover" has a finite subcover, so does the original one.

Example 8.8. Notice that an open cover may not have a Lebesgue number. As in the previous example, put $A = (0, 1)$ and consider the open cover

$\mathbb{U} = \{(1/n, 1); n \geq 2\}$. For any given $\varepsilon > 0$, choose $x < \varepsilon$; then $B_\varepsilon(x) = (0, x + \varepsilon)$ is not contained in $(1/n, 1)$ for any n.

Theorem 8.9. A sequentially compact set in a metric space is totally bounded.

Proof. We will show that if a set A is not totally bounded, then it cannot be sequentially compact – that is, if for some $\varepsilon > 0$ there is no finite ε-net for A, we can construct a sequence $\{x_n\}$ in A with no convergent subsequence.

Take any x_1 in A, and let $U_1 = B_\varepsilon(x_1)$. By assumption, B_1 does not cover A, so there is some $x_2 \in A$, with $x_2 \notin U_1$. Let $U_2 = B_\varepsilon(x_2)$; then $\{U_1, U_2\}$ is still not a cover of A, and therefore there is some $x_3 \in A$, with $x_3 \notin U_1 \cup U_2$. Put $U_3 = B_\varepsilon(x_3), \ldots$, and so forth. By continuing in this fashion, we can construct a sequence $\{x_n\}$ with the property that $d(x_n, x_m) \geq \varepsilon$ for all n and m, as each new term of the sequence is chosen outside all the ε-balls centered at the previous terms. Clearly, this sequence has no Cauchy subsequences and therefore no convergent subsequences either. □

Theorem 8.10. Any open cover of a sequentially compact set in a metric space has a Lebesgue number.

Proof. Let A be a set in a metric space (X, d) with an open cover $\mathbb{U}$. If $\mathbb{U}$ has no Lebesgue number, then for every $\varepsilon > 0$ there exists some point x in A such that no set U in $\mathbb{U}$ contains $B_\varepsilon(x)$. In particular, for each integer n, we can find some point x_n in A such that $B_{1/n}(x_n)$ is not contained in any $U \in \mathbb{U}$. We will show that if A is sequentially compact, no sequence in A can have this property. Hence, given sequential compactness of A, a Lebesgue number must exist for any open cover of it (or else we have a contradiction).

By the sequential compactness of A, any sequence $\{x_n\}$ of points in A contains a convergent subsequence $\{x_{n_k}\}$ with limit $x \in A$. Because $\mathbb{U}$ covers A, $x \in U_0$ for some $U_0 \in \mathbb{U}$, and because U_0 is open, there exists some integer m such that $B_{2/m}(x) \subseteq U_0$. We will show that $B_{1/n}(x_n) \subseteq U_0$ for some terms in the sequence by exploiting the fact that we can bring $\{x_{n_k}\}$ arbitrarily close to x and make $B_{1/n_k}(x_{n_k})$ arbitrarily small.

By the convergence of $\{x_{n_k}\}$ to x, there is some N such that

$$x_{n_k} \in B_{1/m}(x) \quad \text{for all } n_k > N$$

Choose $n_k > \max\{N, m\}$, and observe that for any point y in $B_{1/n_k}(x_{n_k})$ we have

$$d(y,x) \le d(y,x_{n_k}) + d(x_{n_k},x) < \frac{1}{n_k} + \frac{1}{m} < \frac{1}{m} + \frac{1}{m} = \frac{2}{m}$$

Hence, for n_k sufficiently high, we have $y \in B_{2/m}(x)$, but then

$$B_{1/n_k}(x_{n_k}) \subseteq B_{2/m}(x) \subseteq U_0$$

contradicting the nonexistence of a Lebesgue number. □

We can now prove that sequential compactness implies compactness in a metric space.

Theorem 8.11. Any sequentially compact set in a metric space is compact.

Proof. Let $\mathbb{U}$ be an arbitrary open cover of a sequentially compact set A in a metric space. By Theorem 8.10, $\mathbb{U}$ has a Lebesgue number ε, and by Theorem 8.9 there exists a finite ε-net (for the same ε) $\{x_1, \ldots, x_n\}$ for A. For each $i = 1, \ldots, n$ there is some $U_i \in \mathbb{U}$ such that $B_\varepsilon(x_i) \subseteq U_i$, by the definition of Lebesgue number. Because $A \subseteq \cup_{i=1}^n B_\varepsilon(x_i) \subseteq \cup_{i=1}^n U_i$, $\mathbb{U}$ has a finite subcover $\{U_1, \ldots, U_n\}$. □

We will now provide an alternative characterization of compactness in terms of a property of families of closed sets.

Definition 8.12. The finite-intersection property. A nonempty family of sets $\mathbb{A} = \{A_i; i \in I\}$ has the finite-intersection property if every (nonempty) finite subfamily of $\mathbb{A}$ has a nonempty intersection.

Theorem 8.13. A set C in a metric (or topological) space (X, d) is compact if and only if every family of closed subsets of X that has the finite-intersection property has a nonempty intersection.

Proof

- Suppose C is compact. To show that any family of closed subsets that has the finite-intersection property has a nonempty intersection, we will prove the following equivalent (contrapositive) statement: Let $\mathbb{A} = \{A_i; i \in I\}$ be a family of closed subsets of C with the property that $\cap\mathbb{A} = \cap_{i\in I} A_i = \varnothing$; then there exists *some* finite subfamily of $\mathbb{A}$ with an empty intersection – that is, there exists some finite set $J \subseteq I$ such that $\cap_{i\in J} A_i = \varnothing$.

  For each i, let $U_i = \sim A_i$ be the complement of the closed set A_i. Then each U_i is an open set, and we can write, using De Morgan's laws (Theorem 1.2 in Chapter 1)

  Hence, $\{U_i; i \in I\}$ is an open cover of C. Because C is compact, $\{U_i; i \in I\}$ contains a finite subcover of C. That is, there exists a finite set $J \subseteq I$ such that

$$C \subseteq \cup_{i\in J} U_i$$

  which implies that

$$\cap_{i\in J} A_i = \sim(\cup_{i\in J} U_i) \subseteq \sim C \qquad (1)$$

  On the other hand, because each A_i is a subset of C, so is their intersection; hence, we have

$$\cap_{i\in J} A_i \subseteq C \qquad (2)$$

  Combining (1) and (2), we conclude that $\cap_{i\in J} A_i = \varnothing$, which establishes the desired result.
- For the converse, assume that C has the property that if the intersection of any family of closed subsets of C is empty, then the intersection of some finite subfamily of them is empty (we are using the contrapositive again). Let $\mathbb{U} = \{U_i; i \in I\}$ be an arbitrary open cover of C, so that

$$C \subseteq \cup_{i\in I} U_i$$

  and observe that this implies that

$$\sim(\cup_{i\in I} U_i) \subseteq \sim C \qquad (1)$$

  Next, let

$$A_i = C \cap (\sim U_i)$$

  for each i. Using (1) and De Morgan's laws, we have

$$\cap_{i\in I} A_i = \cap_{i\in I}(C \cap (\sim U_i)) = C \cap (\cap_{i\in I}(\sim U_i)) = C \cap (\sim(\cup_{i\in I} U_i)) \subseteq C \cap (\sim C) = \varnothing$$

  Hence, $\mathbb{A} = \{A_i; i \in I\}$ is a family of closed subsets of C whose intersection is empty. By assumption, there exists some finite subfamily of $\mathbb{A}$ with an empty intersection; that is, there exists some finite set $J \subseteq I$ such that $\cap_{i\in J} A_i = \varnothing$, and it follows that

$$\cap_{i\in J} A_i = \cap_{i\in J}(C \cap (\sim U_i)) = C \cap (\cap_{i\in J}(\sim U_i)) = C \cap (\sim(\cup_{i\in J} U_i)) = \varnothing$$

  This implies that C is contained in $\cup_{i\in J} U_i$. Hence, $\{U_i; i \in J\}$ is a finite subcover of $\{U_i; i \in I\}$, and we conclude that C is compact.

(b) Relationships with Other Topological Properties

In metric spaces, compactness is closely related to other topological properties, namely, closedness, completeness, and boundedness. In this section we spell out some of the interconnections among these properties.

Theorem 8.14. Any closed subset of a compact space is compact.

is closed, its complement C^c is open, and $\{U_i; i \in I\} \cup C^c$ is an open cover of X. As X is compact, this cover has a finite subcover $\{U_1, \ldots, U_n\} \cup C^c$. Then $\{U_1, \ldots, U_n\}$ is a finite subcover of C, which is therefore compact. □

Theorem 8.15. A compact set in a metric space is closed.

(This result may not hold in more general topological spaces.)

Proof. Let A be a compact set in a metric space (X, d). We will prove that A is closed by showing that it contains all its limit points. Let x_L be an arbitrary limit point of A; by Theorem 4.11 there exists a sequence $\{x_n\}$ of points of A with limit x_L. By the (sequential) compactness of A, $\{x_n\}$ has a convergent subsequence with limit in A. By the uniqueness of the limit (see Problem 2.5), x_L is the limit of the subsequence and must therefore lie in A. □

Theorem 8.16. A set in a metric space is compact if and only if it is complete and totally bounded.

Proof. We have already seen that a compact set is totally bounded. The proof that compactness implies completeness is left as an exercise. We now prove the converse implication (i.e., that a complete and totally bounded set in a metric space is compact).

Let C be complete and totally bounded. To establish (sequential) compactness, we need to show that any sequence $\{x_n\}$ in C has a subsequence converging to a point in C. And because we are assuming completeness, it is enough to show that given any sequence in C, we can produce a Cauchy subsequence, for completeness will then guarantee convergence.

Let $\{x_n\}$ be an arbitrary sequence in C. Because C is totally bounded, it can be covered by a finite number of balls of radius 1 (a 1-net). Among these balls, there must be one, say B_1, that contains infinitely many terms of the sequence. These infinitely many points of the original sequence form a new sequence that we call $\{x_n^1\}$. Next, we can cover B_1 with a finite number of balls of radius 1/2, and among these balls there must be one, say B_2, such that $B_1 \cap B_2$ contains an infinite number of points of $\{x_n^1\}$, forming a new sequence $\{x_n^2\}$. Continuing in this fashion, we obtain a sequence $\{B_i\}$ of balls with radius $1/i$ such that $B_1 \cap B_2 \cap \ldots \cap B_i$ contains infinitely many terms of the original sequence, yielding a new sequence $\{x_n^i\}$.

Consider now a "cross-sequence" $\{x_k^k\}$ formed by taking one element of each of these sequences (i.e., the kth term of $\{x_k^k\}$ is taken from $\{x_n^k\}$). We

$$x_k^k \in B_1 \cap B_2 \cap \ldots \cap B_k \quad \text{for each } k$$

Hence, given any positive integers p and q, with $p < q$, the terms x_p^p and x_q^q of $\{x_k^k\}$ are contained in the ball B_p (of radius $1/p$), and therefore

$$d(x_p^p, x_q^q) < 2/p$$

Hence, the subsequence $\{x_k^k\}$ is Cauchy: By taking p high enough, we can force all remaining terms of the sequence to fit inside a ball of arbitrarily small radius. By completeness, $\{x_k^k\}$ converges to a point in C. Hence, we have shown that an arbitrary sequence in C must contain a convergent subsequence with limit in C, thus establishing the sequential compactness of the set. □

Problem 8.17. Show that a compact set in a metric space is complete.

Problem 8.18. Let A be a compact set, and let $\{A_n\}$ be a "decreasing sequence" of nonempty closed subsets of A such that $A_{n+1} \subseteq A_n$. Show that $\cup_{n=1}^{\infty} A_n$ is not empty.

From Theorems 8.9 and 8.15, we know that a compact set in a metric space is closed and bounded. The following result tells us that the converse is true for sets of real numbers, thereby establishing an important characterization of compact sets in $\mathbb{R}$ as those that are closed and bounded.

Theorem 8.19. Heine-Borel. Any closed and bounded set of real numbers is compact.

Proof. Note that any bounded set of real numbers must be contained in a closed interval $[a, b]$ with finite end points. Because we know that any closed subset of a compact set is compact, we need only show that $[a, b]$ is compact. By the Bolzano-Weierstrass theorem, any sequence contained in this (bounded) set contains a convergent subsequence, and because $[a, b]$ is closed, the subsequence converges to a point in the interval (Theorem 4.13), establishing sequential compactness (see Problem 3.4). □

This result can be easily extended to any finite-dimensional Euclidean space.

Theorem 8.20. Any closed and bounded subset of $\mathbb{R}^m$ is compact.

Proof. Let A be a closed and bounded set in $\mathbb{R}^m$. Then there exists some number M such that $\|x\|_E \leq M$ for all x in A. Hence, A is contained in the cube of side M given by

$$C_m = I \times I \times \ldots \times I, \qquad \text{where } I = [-M, M]$$

As in the previous theorem, it is enough to show that C_m is compact, for the closedness of A then guarantees its compactness.

To simplify notation, let $m = 2$ (i.e., we will be working in the plane $\mathbb{R}^2$), and consider $C_2 = I \times I = [-M, M] \times [-M, M]$ and an arbitrary sequence $\{x_n\}$ in this set, with $x_n = (x_n^1, x_n^2)$. Observe that $\{x_n^1\}$ and $\{x_n^2\}$ are bounded sequences of real numbers contained in the compact set $I = [-M, M]$. By the Heine-Borel theorem, $\{x_n^1\}$ has a subsequence $\{x_{n_k}^1\}$ convergent to a limit x^1 in I, and the corresponding subsequence of $\{x_n^2\}$, $\{x_{n_k}^2\}$, has a convergent subsequence $\{x_{n_{kj}}^2\}$ with limit x^2 in I. Putting $x_{n_{kj}} = (x_{n_{kj}}^1, x_{n_{kj}}^2)$, it is clear that (by the equivalence between convergence in E^2 and coordinate-wise convergence in $\mathbb{R}$)

$$\{x_{n_{k_j}}\} \to (x^1, x^2) \in I \times I$$

that is, $\{x_n\}$ has a convergent subsequence with limit in C_2, which establishes the sequential compactness of C_2 and therefore of any closed and bounded set in the plane. The argument can be easily extended to any finite-dimensional Euclidean space. More generally, it can be shown that a finite product of compact sets is compact (in the sup metric).

(c) Continuous Functions on Compact Sets

Theorem 8.21. Let (X, d) *and* (Y, ρ) *be metric spaces, and* f:X ⟶ Y *a continuous function. If* C *is a compact set in* (X, d)*, its image* f(C) *is compact in* (Y, ρ).

Proof. Let $\{y_n\}$ be an arbitrary sequence in $f(C)$, and consider a companion sequence formed by points x_n in C such that $f(x_n) = y_n$. By the sequential compactness of C, $\{x_n\}$ has a convergent subsequence, say $\{x_{n_k}\}$, with limit x in C. Then, by the continuity of f,

$$\lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} f(x_{n_k}) = f\left(\lim_{k\to\infty} x_{n_k}\right) = f(x) \in f(C)$$

Hence, $\{y_n\}$ has a subsequence $\{y_{n_k}\}$ that converges to a limit in $f(C)$. This establishes the sequential compactness of $f(C)$. □

In the case of a real-valued function, the theorem says that the continuous image of a compact set is a compact interval or a collection of them. Because any such set of real numbers contains both its supremum and its infimum, we have the following important corollary:

Theorem 8.22. Extreme value (Weierstrass). Let C *be a compact set in a metric space, and* f: C ⟶ ℝ *a continuous function. Then* f *is bounded in* C *and attains both its maximum and its minimum in the set. That is, there exist points* x_M *and* x_m *in* C *such that*

$$f(x_M) = \sup f(C) \quad \text{and} \quad f(x_m) = \inf f(C)$$

Proof. We will prove the existence of a maximum. By the previous theorem, $f(C)$ is a compact set of real numbers and therefore is closed and bounded. Let β be its supremum. Then β is a limit point of $f(C)$. (Why? $\beta - 1/n$ is not an upper bound for $f(C)$). Because $f(C)$ is closed, it follows that β is contained in it, that is, there exists some point x_M in C such that $\beta = f(x_M)$. □

Problem 8.23. Give an alternative proof for Theorem 8.21 using directly the definition of compactness. (Let $\{U_i; i \in I\}$ be an open cover of $f(C)$.)

Theorem 8.24. Let (X, d) *and* (Y, ρ) *be metric spaces, with* f:X ⟶ Y *a continuous function, and* C *a compact set in* (X, d)*. Then* f *is uniformly continuous on* C.

Proof. Let $\varepsilon > 0$ be given. Because f is continuous, for each point x in C we can find a positive number $\delta(x)$ such that

$$d(x, y) < \delta(x) \Rightarrow \rho[f(x), f(y)] < \varepsilon/2 \tag{1}$$

For each $x \in C$, let $B(x)$ be the set of all points y in C for which $d(x, y) < \delta(x)/2$. The collection of all such $B(x)$'s (one for each point in C) is an open cover of C, and because C is compact, there is a finite collection of points in C, say $\{x_1, \ldots, x_n\}$, such that

$$C \subseteq B(x_1) \cup \ldots \cup B(x_n) \tag{2}$$

Put

$$\delta = \frac{\min\{\delta(x_1), \ldots, \delta(x_n)\}}{2}$$

and observe that $\delta > 0$ because this is a finite collection of positive numbers (this is why we need compactness, it guarantees that we can find a finite subcover; note that the infimum of an infinite collection of positive numbers may be zero).

Let x and y be points in C such that $d(x, y) < \delta$. By (2), there is some point x_m such that $x \in B(x_m)$, and hence

$$d(x, x_m) < \frac{\delta(x_m)}{2} \tag{3}$$

Moreover,

$$d(y, x_m) \le d(y, x) + d(x, x_m) < \delta + \frac{\delta(x_m)}{2} \le \delta(x_m) \tag{4}$$

Hence, both x and y are sufficiently close to x_m that we can use (1) to conclude that

$$\rho[f(y), f(x)] \le \rho[f(y), f(x_m)] + \rho[f(x_m), f(x)] < \varepsilon \qquad \square$$

A similar argument will yield the following result.

Theorem 8.25. Show that if a function is locally Lipschitz on a compact set, then it is Lipschitz on the set (see Definition 6.18).

Problem 8.26. Compactness of the product space. Let (X, d_1) and (Y, d_2) be metric spaces, and consider the product space $(Z = X \times Y, d_\pi)$, with the *product metric* d_π defined by

$$d_\pi(z, z') = d_\pi[(x, y), (x', y')] = \sqrt{[d_1(x, x')]^2 + [d_2(y, y')]^2} \tag{1}$$

Show that the product space $(Z = X \times Y, d_\pi)$ is compact if and only if both (X, d_1) and (Y, d_2) are compact.

9. Connected Sets

A set is said to be connected if it consists of a single piece (i.e., if it is not made up of two or more "separate components"). The following definition makes this idea more precise.

Definition 9.1. Separated and connected sets. Two sets A and B in a metric space are said to be separated if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty (i.e., if neither set has a point lying in the closure of the other). A set C in a metric space is said to be connected if it is not the union of two nonempty separated sets.

Notice that the condition for two sets to be separated is stronger than disjointedness but weaker than the requirement that the distance between them be strictly positive. Thus, the intervals $(-1, 0]$ and $(0, 1)$ are disjoint but not separated, because 0 lies in one interval and in the closure of the other.

The intervals $(-1, 0)$ and $(0, 1)$, however, are separated, but the distance between them is zero.

Connected sets on the real line have a particularly simple structure. As shown in our next result, the connected sets in $\mathbb{R}$ are precisely the intervals.

Theorem 9.2. A set S of real numbers is connected if and only if it is an interval.

Proof. Recall that a set I of real numbers is an interval if whenever x and y are in I, any real number z, with $x < z < y$, also lies in I (Problem 6.14 in Chapter 1).

- We first show that a set of real numbers that is not an interval is not connected. Let S be such a set. Then there exist real numbers x and y in S and $z \notin S$ such that $x < z < y$, and we can write S as the union of two components, as follows:

$$S = S_1 \cup S_2 \equiv [S \cap (-\infty, z)] \cup [S \cap (z, \infty)]$$

  Notice that neither of these sets is empty, because S_1 contains at least x, and S_2 contains at least y. Moreover, S_1 and S_2 are separated, because $S_1 \subseteq (-\infty, z)$ and $S_2 \subseteq (z, \infty)$, and these intervals are separated (neither of them contains the only common boundary point, z). Hence S is not connected.
- To show that every interval is connected, we show that a nonconnected set cannot be an interval. Let E be a nonconnected set of real numbers. Then there exist nonempty separated sets A and B such that $A \cup B = E$. Pick $a \in A$ and $b \in B$, and assume (relabeling the sets if necessary) that $a < b$, as in Figure 2.12. To establish that E is not an interval, we will show that there is some real number $x \notin E$ with $a < x < b$.

  We define

$$x = \sup\{A \cap [a, b]\}$$

  Then (see Problem 4.15) we have $x \in \overline{A}$ and (because A and B are separated) $x \notin B$. Moreover, we have $a \le x < b$. There are now two possibilities. If $x \notin A$, then we have found the desired number, for then $a < x < b$ and $x \notin E$. If $x \in A$, on the other hand, we have $x \notin \overline{B}$ (because A and B are separated), and it follows that x lies in the open set $\mathbb{R} \sim \overline{B}$. Hence, we can find some other point x' in this set (and therefore not in B) such that $a \le x < x' < b$. This establishes the desired result.

Figure 2.12 $\square$