SIX

Necessary and Sufficient Conditions for a General Class of Control Problems

Until this point, our development of the necessary and sufficient conditions of optimal control theory did not allow for constraints that depended on the control variables *and* the state variables. The goal of this chapter is to state two theorems giving necessary conditions for a class of control problems that contain such constraints, which we refer to as *mixed constraints* because of the presence of the state and control variables in them, as well as to prove two theorems giving sufficient conditions pertaining to this class of problems. We will use the necessary conditions to explicitly solve for the optimal paths of a version of the capital accumulating model of the firm without adjustment costs and irreversible investment. The theorems presented here are the most general we shall present and use in the book.

The problem under consideration in this chapter is to find a piecewise continuous control vector function $\mathbf{u}(\cdot) \stackrel{\text{def}}{=} (u_1(\cdot), u_2(\cdot), \ldots, u_M(\cdot))$ and its associated piecewise smooth state vector function $\mathbf{x}(\cdot) \stackrel{\text{def}}{=} (x_1(\cdot), x_2(\cdot), \ldots, x_N(\cdot))$, defined on the fixed time interval $[t_0, t_1]$ that will solve the ensuing constrained fixed endpoints optimal control problem:

$$\max_{\mathbf{u}(\cdot)} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \stackrel{\text{def}}{=} \int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t))\, dt$$

$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \ \ \mathbf{x}(t_0) = \mathbf{x}_0, \ \ \mathbf{x}(t_1) = \mathbf{x}_1, \tag{1}$$

$$h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0, \quad k = 1, 2, \ldots, K',$$

$$h^k(t, \mathbf{x}(t), \mathbf{u}(t)) = 0, \quad k = K'+1, K'+2, \ldots, K,$$

where $\mathbf{g}(\cdot) \stackrel{\text{def}}{=} (g^1(\cdot), g^2(\cdot), \ldots, g^N(\cdot))$, $\dot{\mathbf{x}}(\cdot) \stackrel{\text{def}}{=} (\dot{x}_1(\cdot), \dot{x}_2(\cdot), \ldots, \dot{x}_N(\cdot))$, and $\mathbf{h}(\cdot) \stackrel{\text{def}}{=} (h^1(\cdot), h^2(\cdot), \ldots, h^K(\cdot))$. It is important to observe that the K constraints in problem (1) are equivalent to the requirement that $\mathbf{u}(t) \in U(t, \mathbf{x}(t))$, where the control set

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

$U(t, \mathbf{x}(t))$ is now more generally defined as

$$U(t, \mathbf{x}(t)) \stackrel{\text{def}}{=} \{\mathbf{u}(\cdot) : h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0, \quad k = 1, \ldots, K',$$
$$h^k(t, \mathbf{x}(t), \mathbf{u}(t)) = 0, \quad k = K' + 1, \ldots, K\}. \tag{2}$$

Notice that in Eq. (2) we have used notation for the control set (or region) that makes explicit its dependence on the state variables. To simplify notation and make it less cumbersome, we will often use the set inclusion $\mathbf{u}(t) \in U(t, \mathbf{x}(t))$ to signify the values of the control functions that satisfy the system of K' inequality constraints and $K - K'$ equality constraints in problem (1). In passing, note that we will consider the various transversality conditions that arise for perturbations of problem (1) four chapters hence, *after* we establish the dynamic envelope theorem.

Because we are studying a more general class of optimal control problems than we did earlier, we must amend our definition of an admissible pair. To this end, we have the following definition, which should be contrasted with that given earlier in Definitions 2.1 and 4.1.

Definition 6.1: We call $(\mathbf{x}(t), \mathbf{u}(t))$ an *admissible pair* if $\mathbf{u}(\cdot)$ is any piecewise continuous control vector function such that $\mathbf{u}(t) \in U(t, \mathbf{x}(t)) \forall\, t \in [t_0, t_1]$ and $\mathbf{x}(\cdot)$ is a piecewise smooth state vector function such that $\dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t))$, $\mathbf{x}(t_0) = \mathbf{x}_0$, and $\mathbf{x}(t_1) = \mathbf{x}_1$.

As remarked in Chapter 4 when we dealt with inequality constraints of the simpler form $h^k(t, \mathbf{u}(t)) \geq 0,\ k = 1, 2, \ldots, K$, there is essentially no loss in generality in defining the control region by a system of inequality and equality constraints, for as a practical matter, the control set is almost always specified in such a manner. Accordingly, we must therefore introduce a constraint qualification on the constraint functions $h^k(\cdot)$, $k = 1, 2, \ldots, K$, just as we did in Chapter 4. Before doing so, first recall that by Definition 4.2, if X is a set, then card(X), the cardinal number of X, is the number of elements in X. We are now in a position to state the constraint qualification of interest to us.

Rank Constraint Qualification: Define $\iota(t, \mathbf{z}(t), \mathbf{v}(t)) \stackrel{\text{def}}{=} \{k : h^k(t, \mathbf{z}(t), \mathbf{v}(t)) = 0,$ $k = 1, 2, \ldots, K\}$ as the index set of the binding constraints along the optimal path. For every $t \in [t_0, t_1]$, if $\iota(t, \mathbf{z}(t), \mathbf{v}(t)) \neq \emptyset$, that is, $\iota(t, \mathbf{z}(t), \mathbf{v}(t))$ is nonempty, then the card$(\iota(t, \mathbf{z}(t), \mathbf{v}(t))) \times M$ Jacobian matrix

$$\left[\frac{\partial h^k}{\partial u_m}(t, \mathbf{z}(t), \mathbf{v}(t))\right]_{\substack{k \in \iota(t, \mathbf{z}(t), \mathbf{v}(t)) \\ m = 1, 2, \ldots, M}} \tag{3}$$

has a rank equal to card$(\iota(t, \mathbf{z}(t), \mathbf{v}(t)))$. That is, the rank of the above Jacobian matrix is equal to the number of its rows, the maximum rank it can have.

For example, if k_B of the constraints bind at a given point t in the planning horizon, where we note that $k_B \geq K - K'$ because there are $K - K'$ equality constraints

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

in problem (1), then $\text{card}(\iota(t, \mathbf{z}(t), \mathbf{v}(t))) = k_B$. Hence the above Jacobian matrix will be of order $k_B \times M$ and the constraint qualification will be satisfied if the rank of the Jacobian equals k_B. This rank condition on the Jacobian has two important consequences, both of which were noted in Chapter 4. First, it implies that at least one control variable must be present in each of the binding constraints. To see this, assume that the constraint qualification holds but that one of the k_B binding constraints, say, the first (without loss of generality), does not have any control variables in it. In this case, the first row of the Jacobian matrix would be identically zero, thereby implying that the rows of the Jacobian are linearly dependent. This, in turn, implies that the Jacobian is of rank less than k_B, thus violating the constraint qualification. Second, the constraint qualification implies that the number of binding constraints k_B cannot be greater than the number of control variables M. To see this, assume that the constraint qualification holds but that $k_B > M$. Then there are more rows (k_B) in the Jacobian than there are columns (M). But the rank of a matrix cannot exceed the minimum of the number of its rows or columns. This, in turn, implies that the rank of the Jacobian is $\min(k_B, M) = M$, which is less than k_B, thereby violating the constraint qualification. Note, however, that the constraint qualification permits the number of constraints (K) to exceed the number of control variables (M) in view of the fact that some of the inequality constraints may not bind.

In developing the necessary conditions in Chapters 2 and 4, we found it convenient to define a function called the Hamiltonian, and we will similarly find it convenient to do so here. In the present case, there are N state variables and N ordinary differential equations describing their rates of changes with respect to time. Consequently, we associate with each of the differential equations a costate function $\lambda_n(\cdot)$, $n = 1, 2, \ldots, N$, and define the Hamiltonian for problem (1) by

$$\begin{aligned} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) &\stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}) \\ &\stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}), \end{aligned} \tag{4}$$

where $\boldsymbol{\lambda} \stackrel{\text{def}}{=} (\lambda_1, \lambda_2, \ldots, \lambda_N)$ and $\boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}) \stackrel{\text{def}}{=} \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u})$ is the scalar or inner product of the vectors $\boldsymbol{\lambda}$ and $\mathbf{g}(t, \mathbf{x}, \mathbf{u})$. As you may recall from Theorem 4.2, the Maximum Principle requires that the optimal control vector maximize the Hamiltonian subject to the constraints that the control variables lie in the control set. Given that the control set is defined by a system of inequality and equality constraints in problem (1), the Maximum Principle dictates that a nonlinear programming problem be solved, to wit, $\max_{\mathbf{u}\in U(t,\mathbf{x})} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$, in order to find the optimal control vector, something that can be done, in principle, with Theorem 18.5 of Simon and Blume (1994), that is, the Karush-Kuhn-Tucker theorem. This is the same conclusion we arrived at in Chapter 4. Moreover, this observation connotes that a Lagrangian function be formed in order to solve the constrained optimization problem $\max_{\mathbf{u}\in U(t,\mathbf{x})} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$. To this end, let us define the Lagrangian function $L(\cdot)$

corresponding to the nonlinear programming problem $\max_{\mathbf{u}\in U(t,\mathbf{x})} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$ dictated by the Maximum Principle, by associating with each constraint function $h^k(\cdot)$ a Lagrange multiplier function $\mu_k(\cdot)$, $k = 1, 2, \ldots, K$, to get

$$\begin{aligned} L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &\stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}) + \sum_{k=1}^{K} \mu_k h^k(t, \mathbf{x}, \mathbf{u}) \\ &\stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u}) \\ &\stackrel{\text{def}}{=} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u}), \end{aligned} \tag{5}$$

where $\boldsymbol{\mu} \stackrel{\text{def}}{=} (\mu_1, \mu_2, \ldots, \mu_K)$. Note that it is precisely because we have defined the control set in terms of a system of inequality and equality constraints in problem (1) that we are led to such a formulation. With the constraint qualification now dealt with, we may turn to the necessary conditions for problem (1).

In addition to the standard assumptions on the functions $f(\cdot)$ and $\mathbf{g}(\cdot)$ given in Chapter 1, we now make the following *additional* ones:

(A.1) $\partial f(\cdot)/\partial u_m \in C^{(0)}$ with respect to the $1 + N + M$ variables $(t, \mathbf{x}, \mathbf{u})$, for $m = 1, 2, \ldots, M$.
(A.2) $\partial g^n(\cdot)/\partial u_m \in C^{(0)}$ with respect to the $1 + N + M$ variables $(t, \mathbf{x}, \mathbf{u})$ for $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$.
(A.3) $h^k(\cdot) \in C^{(0)}$ with respect to the $1 + N + M$ variables $(t, \mathbf{x}, \mathbf{u})$ for $k = 1, 2, \ldots, K$.
(A.4) $\partial h^k(\cdot)/\partial x_n \in C^{(0)}$ with respect to the $1 + N + M$ variables $(t, \mathbf{x}, \mathbf{u})$ for $k = 1, 2, \ldots, K$ and $n = 1, 2, \ldots, N$.
(A.5) $\partial h^k(\cdot)/\partial u_m \in C^{(0)}$ with respect to the $1 + N + M$ variables $(t, \mathbf{x}, \mathbf{u})$ for $k = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, M$.

We may now state the necessary conditions pertaining to problem (1).

Theorem 6.1 (Necessary Conditions, Mixed Constraints): *Let* $(\mathbf{z}(t), \mathbf{v}(t))$ *be an admissible pair for problem (1), and assume that the rank constraint qualification is satisfied. Then if* $(\mathbf{z}(t),\mathbf{v}(t))$ *yields the absolute maximum of* $J[\cdot]$, *it is necessary that there exist a piecewise smooth vector-valued function* $\boldsymbol{\lambda}(\cdot) \stackrel{\text{def}}{=} (\lambda_1(\cdot), \lambda_2(\cdot), \ldots, \lambda_N(\cdot))$ *and a piecewise continuous vector-valued Lagrange multiplier function* $\boldsymbol{\mu}(\cdot) \stackrel{\text{def}}{=} (\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_K(\cdot))$, *such that for all* $t \in [t_0, t_1]$,

$$\mathbf{v}(t) = \arg\max_{\mathbf{u}}\{H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t))\ s.t.\ \mathbf{u} \in U(t, \mathbf{z}(t))\},$$

that is, if

$$M(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)) \stackrel{\text{def}}{=} \max_{\mathbf{u}}\{H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t))\ s.t.\ \mathbf{u} \in U(t, \mathbf{z}(t))\},$$

then

$$M(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)) \equiv H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)),$$

or equivalently

$$H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)) \geq H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t))\, \forall\, \mathbf{u} \in U(t, \mathbf{z}(t)),$$

where

$$\begin{aligned} U(t, \mathbf{x}(t)) \stackrel{\text{def}}{=} \{\mathbf{u}(\cdot) : h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0, \quad k = 1, \ldots, K', \\ h^k(t, \mathbf{x}(t), \mathbf{u}(t)) = 0, \quad k = K' + 1, \ldots, K\} \end{aligned}$$

is the control set. Because the rank constraint qualification is assumed to hold, the above necessary condition implies that

$$\frac{\partial L}{\partial u_m}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad m = 1, 2, \ldots, M,$$

$$\frac{\partial L}{\partial \mu_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \geq 0,\ \mu_k(t) \geq 0,$$

$$\mu_k(t)\frac{\partial L}{\partial \mu_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad k = 1, 2, \ldots, K',$$

$$\frac{\partial L}{\partial u_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad k = K' + 1, K' + 2, \ldots, K,$$

where

$$\begin{aligned} L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &\stackrel{\text{def}}{=} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u}) \\ &= f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}) + \sum_{k=1}^{K} \mu_k h^k(t, \mathbf{x}, \mathbf{u}) \end{aligned}$$

is the Lagrangian function. Furthermore, except for the points of discontinuities of $\mathbf{v}(t)$*,*

$$\dot{z}_n(t) = \frac{\partial L}{\partial \lambda_n}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = g^n(t, \mathbf{z}(t), \mathbf{v}(t)), \quad n = 1, 2, \ldots, N,$$

$$\dot{\lambda}_n(t) = -\frac{\partial L}{\partial x_n}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)), \quad n = 1, 2, \ldots, N,$$

where the above notation means that the functions are first differentiated with respect to the particular variable and then evaluated at $(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))$.

Let us remark on several important consequences of Theorem 6.1, all of which are analogous to those noted for Theorem 4.4. First, in general, the Lagrange multiplier function $\boldsymbol{\mu}(\cdot) \stackrel{\text{def}}{=} (\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_K(\cdot))$ is piecewise continuous on $[t_0, t_1]$. It turns out, however, that $\boldsymbol{\mu}(\cdot)$ is continuous whenever the optimal control function $\mathbf{v}(\cdot)$ is continuous. Thus the discontinuities in the Lagrange multipliers can only occur at points where the optimal control is discontinuous. One implication of Theorem 6.1 is that the Lagrangian function evaluated along the optimal solution is a continuous function of t, that is, $L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))$ is a continuous function of t. Another implication, identical in flavor to that derived in Mental Exercises 2.31 and 4.13 for simpler control problems, is that the *total derivative* of $L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))$ with respect to t, scilicet,

$$\dot{L}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \stackrel{\text{def}}{=} \frac{d}{dt} L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)),$$

is equal to the *partial derivative* of $L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))$ with respect to t, namely,

$$L_t(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \stackrel{\text{def}}{=} \frac{\partial}{\partial t} L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)),$$

at all continuity points of the optimal control function $\mathbf{v}(\cdot)$, and assuming that $f(\cdot)$, $\mathbf{g}(\cdot)$, and $\mathbf{h}(\cdot)$ are $C^{(1)}$ functions of $(t, \mathbf{x}, \mathbf{u})$. Other results similar to those derived in Mental Exercises 2.31 and 4.13 also hold, as you will discover in a mental exercise. Finally, we should note that the necessary conditions involving inequalities are complementary slackness conditions exactly analogous to those encountered in linear and nonlinear programming.

Let us now indicate that the necessary conditions of Theorem 6.1 contain enough conditions to determine candidates for optimality. It should be clear from the method of Lagrange and the Karush-Kuhn-Tucker theorem that the necessary conditions

$$\frac{\partial L}{\partial u_m}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0, \quad m = 1, 2, \ldots, M,$$

$$\frac{\partial L}{\partial \mu_k}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \geq 0,\ \mu_k \geq 0,\ \mu_k \frac{\partial L}{\partial \mu_k}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0, \quad k = 1, 2, \ldots, K',$$

$$\frac{\partial L}{\partial \mu_k}(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0, \quad k = K' + 1, K' + 2, \ldots, K,$$

in principle determine the control variables and the Lagrange multipliers as functions of time, the state variables, the costate variables, and any parameters of the problem, say, $\mathbf{u} = \hat{\mathbf{u}}(t, \mathbf{x}, \boldsymbol{\lambda})$ and $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}(t, \mathbf{x}, \boldsymbol{\lambda})$. Substituting these solutions into the canonical equations then gives the system of differential equations

$$\dot{x}_n = \frac{\partial L}{\partial \lambda_n}(t, \mathbf{x}, \hat{\mathbf{u}}(t, \mathbf{x}, \boldsymbol{\lambda}), \boldsymbol{\lambda}, \hat{\boldsymbol{\mu}}(t, \mathbf{x}, \boldsymbol{\lambda})) = g^n(t, \mathbf{x}, \hat{\mathbf{u}}(t, \mathbf{x}, \boldsymbol{\lambda})), \quad n = 1, 2, \ldots, N,$$

$$\dot{\lambda}_n = -\frac{\partial L}{\partial x_n}(t, \mathbf{x}, \hat{\mathbf{u}}(t, \mathbf{x}, \boldsymbol{\lambda}), \boldsymbol{\lambda}, \hat{\boldsymbol{\mu}}(t, \mathbf{x}, \boldsymbol{\lambda})), \quad n = 1, 2, \ldots, N.$$

The solution of this system gives rise to $2N$ constants of integration, which can in principle be determined by using the $2N$ boundary conditions $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{x}(t_1) = \mathbf{x}_1$, thereby yielding the solution $(\mathbf{z}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1), \boldsymbol{\lambda}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1))$. This solution can then be substituted back into $\hat{\mathbf{u}}(t, \mathbf{x}, \boldsymbol{\lambda})$ and $\hat{\boldsymbol{\mu}}(t, \mathbf{x}, \boldsymbol{\lambda})$ to determine their optimal solution, say,

$$\mathbf{v}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1) \stackrel{\text{def}}{=} \hat{\mathbf{u}}(t, \mathbf{z}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1), \boldsymbol{\lambda}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1)),$$

$$\boldsymbol{\mu}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1) \stackrel{\text{def}}{=} \hat{\boldsymbol{\mu}}(t, \mathbf{z}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1), \boldsymbol{\lambda}(t; t_0, \mathbf{x}_0, t_1, \mathbf{x}_1)).$$

Thus, in principle, Theorem 6.1 contains enough conditions to find the optimal solution of problem (1), as was to be demonstrated. In passing, recall that in Chapter 4, a rather thorough set of remarks concerning the role that the rank constraint qualification plays in solving an optimal control problem was presented. It is advisable to revisit those remarks at this time if one's memory of them is a bit vague.

Let's now pause and consider an example of a capital accumulating model of the firm that has a bounded investment rate and irreversibility, and solve it with Theorem 6.1.

Example 6.1: This example examines the dynamic behavior of a price-taking capital accumulating firm, but unlike earlier investigations of it, the firm does *not* face any adjustment costs and investment is irreversible. The firm produces a single output y with the production function $f(\cdot)$ with values $f(K(t)) \stackrel{\text{def}}{=} K(t)$, where $K(t)$ is the capital stock of the firm at time t, and is the only input into the production of the good. The good is sold in a competitive market at the constant price $p > 0$. The firm may purchase the capital good at the rate $I(t)$ at the competitive and constant price $c > 0$. For simplicity, we assume that the firm does not discount its profit flow, which is given by $\pi(K(t), I(t); c, p) \stackrel{\text{def}}{=} pK(t) - cI(t)$. Also, for simplicity, we assume that capital does not depreciate, so that the capital stock accumulates according to the differential equation $\dot{K}(t) = I(t)$. The initial capital stock $K(0) = K_0 > 0$ is given, as is the planning horizon $[0, T]$, but the terminal stock of capital $K(T) = K_T$ is chosen by the firm. The irreversibility of the investment rate is captured by a nonnegativity constraint on it, that is, $I(t) \geq 0$. We also assume that the firm cannot borrow to finance its purchases of the capital stock, thereby implying that all investment expenditures must come from the revenue generated from the sale of the good produced. As a result, another constraint on the firm's behavior is the nonnegativity of its profit flow at each date in the planning horizon, that is, $pK(t) - cI(t) \geq 0$. Putting all of this information together, the control problem to be solved by the firm in order to determine its optimal investment rate is

$$\max_{I(\cdot), K_T} \int_0^T [pK(t) - cI(t)]\, dt$$

$$\text{s.t.} \quad \dot{K}(t) = I(t),\ K(0) = K_0,\ K(T) = K_T,$$

$$h^1(t, K(t), I(t)) \stackrel{\text{def}}{=} I(t) \geq 0,$$

$$h^2(t, K(t), I(t)) \stackrel{\text{def}}{=} pK(t) - cI(t) \geq 0.$$

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

Because the problem is linear in the investment rate, we anticipate that the solution could be piecewise continuous in t. Finally, we assume that $pT > c$. This asserts that total revenue generated from one unit of capital over the entire planning horizon exceeds the cost of purchasing it. As one might expect, without this assumption, the firm would not find investment profitable and thus would essentially cease to operate. To simplify notation, let us define $\boldsymbol{\beta} \overset{\text{def}}{=} (c, p, K_0, T)$ as the parameter vector.

Define the Hamiltonian as $H(K, I, \lambda; c, p) \overset{\text{def}}{=} pK - cI + \lambda I$ and the Lagrangian as

$$L(K, I, \lambda, \mu_1, \mu_2; c, p) \overset{\text{def}}{=} H(K, I, \lambda; c, p) + \mu_1 I + \mu_2[pK - cI].$$

The necessary conditions as dictated by Theorem 6.1 are

$$L_I(K, I, \lambda, \mu_1, \mu_2; c, p) = -c + \lambda + \mu_1 - c\mu_2 = 0, \tag{6}$$

$$L_{\mu_1}(K, I, \lambda, \mu_1, \mu_2; c, p) = I \geq 0, \mu_1 \geq 0, \mu_1 L_{\mu_1}(K, I, \lambda, \mu_1, \mu_2; c, p) = 0, \tag{7}$$

$$L_{\mu_2}(K, I, \lambda, \mu_1, \mu_2; c, p) = pK - cI \geq 0, \mu_2 \geq 0,$$
$$\mu_2 L_{\mu_2}(K, I, \lambda, \mu_1, \mu_2; c, p) = 0, \tag{8}$$

$$\dot{\lambda} = -L_K(K, I, \lambda, \mu_1, \mu_2; c, p) = -p - p\mu_2, \lambda(T) = 0, \tag{9}$$

$$\dot{K} = L_\lambda(K, I, \lambda, \mu_1, \mu_2; c, p) = I, K(0) = K_0. \tag{10}$$

Note that because K_T is a decision variable, we have employed the terminal transversality condition $\lambda(T) = 0$. Before trying to solve the control problem, let's first establish two preliminary results that will be of value in solving it.

First, seeing as $\mu_2 \geq 0$ from Eq. (8) and $p > 0$ by assumption, inspection of Eq. (9) shows that $\dot{\lambda}(t) < 0 \, \forall \, t \in [0, T]$. Moreover, because $\lambda(T) = 0$ from Eq. (9), the prior conclusion implies that $\lambda(t) > 0 \, \forall \, t \in [0, T)$. Second, all admissible values of the capital stock are positive for the entire planning period, that is, $K(t) > 0 \, \forall \, t \in [0, T]$. This follows from the facts that $K(0) = K_0 > 0$, $I(t) \geq 0 \, \forall \, t \in [0, T]$, and $\dot{K} = I$. You are asked to provide a more formal proof of this result in a mental exercise.

We can now show that the rank constraint qualification is satisfied for all admissible pairs. To do so, we first establish that, at most, only one of the two constraints can bind at any time $t \in [0, T]$ for all admissible pairs. If the first constraint $h^1(t, K(t), I(t)) \overset{\text{def}}{=} I(t) \geq 0$ binds at time $\bar{t} \in [0, T]$, then $I(\bar{t}) = 0$. But if $I(\bar{t}) = 0$, then $h^2(\bar{t}, K(\bar{t}), I(\bar{t})) \overset{\text{def}}{=} pK(\bar{t}) - cI(\bar{t}) = pK(\bar{t}) > 0$ because $K(t) > 0 \, \forall \, t \in [0, T]$; hence the second constraint isn't binding when the first one is. By the same token, if the second constraint $h^2(t, K(t), I(t)) \overset{\text{def}}{=} pK(t) - cI(t) \geq 0$ binds at some time $\tilde{t} \in [0, T]$, then $I(\tilde{t}) = pc^{-1}K(\tilde{t}) > 0$ because $K(t) > 0 \, \forall \, t \in [0, T]$. This therefore implies that $h^1(\tilde{t}, K(\tilde{t}), I(\tilde{t})) \overset{\text{def}}{=} I(\tilde{t}) > 0$; thus the first constraint isn't binding when the second one is. As a result, at most one constraint binds at any time $t \in [0, T]$, just as we intended to show. Consequently, the index set $\iota(t, K(t), I(t)) \overset{\text{def}}{=} \{k : h^k(t, K(t), I(t)) = 0, \; k = 1, 2\}$ contains at most one element for all admissible function pairs $(K(\cdot), I(\cdot))$, that is to say, $\operatorname{card}(\iota(t, K(t), I(t))) \leq 1$ for all admissible function pairs $(K(\cdot), I(\cdot))$.

Now if the first constraint binds, then $\operatorname{card}(\iota(t, K(t), I(t))) = 1$ and the relevant Jacobian matrix is given by

$$\left[\frac{\partial h^k}{\partial u_m}(t, \mathbf{x}(t), \mathbf{u}(t))\right]_{\substack{k \in \iota(t, \mathbf{x}(t), \mathbf{u}(t)) \\ m=1,2,\ldots,M}} = \frac{\partial h^1}{\partial I}(t, K(t), I(t)) = 1,$$

which has rank one for all admissible values of the capital stock and investment rate. Similarly, if instead the second constraint binds, then $\operatorname{card}(\iota(t, K(t), I(t))) = 1$ and the relevant Jacobian matrix is

$$\left[\frac{\partial h^k}{\partial u_m}(t, \mathbf{x}(t), \mathbf{u}(t))\right]_{\substack{k \in \iota(t, \mathbf{x}(t), \mathbf{u}(t)) \\ m=1,2,\ldots,M}} = \frac{\partial h^2}{\partial I}(t, K(t), I(t)) = -c,$$

which also has rank one for all admissible values of the capital stock and investment rate. We may conclude, therefore, that the rank constraint qualification is satisfied for all admissible values of the capital stock and investment rate because the rank of the relevant Jacobian matrix is equal to $\operatorname{card}(\iota(t, K(t), I(t)))$ for all admissible values of the capital stock and investment rate, just as claimed.

Let's now establish that an interior solution, that is, one in which *neither* of the constraints are binding, can only occur for an instant, if at all. In other words, we will show that an interior solution cannot hold for an interval of time. To that end, assume that an interior solution exists for a finite interval of time. This implies, via Eqs. (7) and (8), that $\mu_1 = 0$ and $\mu_2 = 0$. In turn, Eq. (6) reduces to $\lambda(t) = c$, a positive constant. This, however, cannot hold for an interval of time because $\dot{\lambda}(t) < 0\, \forall\, t \in [0, T]$, as noted above. Thus $\lambda(t) = c$ can only hold for an instant, if at all. Consequently, an interior solution of the problem, if one exists, can only occur for an instant of time.

Let's now make a conjecture about the form (or structure) of the solution to the necessary conditions (6) through (10), and then seek to verify that it solves them. To come up with a plausible conjecture, first recall that we showed an interior solution for the investment rate can only hold for an instant. Hence, we conjecture that the solution for the investment rate is of the bang-bang flavor. Next, inspection of the objective functional yields the conclusion that a larger capital stock, ceteris paribus, furnishes higher profit flow and hence wealth. With all else the same, therefore, the firm would seemingly like to build up its capital stock as fast as possible inasmuch as this would appear to generate the greatest wealth. These observations lead to the following conjecture about the investment rate:

$$I(t) = \begin{cases} pc^{-1}K(t) & \forall\, t \in [0, s] \\ 0 & \forall\, t \in (s, T], \end{cases} \tag{11}$$

where $s \in (0, T)$ is the switching time, which is to be determined.

To begin verification of the conjecture in Eq. (11), we assume that $\mu_2(t) > 0\,\forall\, t \in [0, s]$. From Eq. (8), specifically $\mu_2 L_{\mu_2}(K, I, \lambda, \mu_1, \mu_2; c, p) = \mu_2[pK - cI] = 0$, it follows that $I = pc^{-1}K$, which is the first part of the conjecture about the investment rate given in Eq. (11). Because $K(t) > 0\,\forall\, t \in [0, T]$, as noted above, $I = pc^{-1}K > 0$ too, which implies by way of Eq. (7) that $\mu_1(t) = 0\,\forall\, t \in [0, s]$. Using $I = pc^{-1}K$, the state equation (10) becomes $\dot{K} = pc^{-1}K$. Integrating the state equation using the integrating factor $\exp[-pc^{-1}t]$ and the initial condition $K(0) = K_0$ produces $K_1^*(t; \boldsymbol{\beta}) = K_0 e^{pc^{-1}t} > 0$ as the specific solution for the capital stock, where the subscript 1 indicates that this solution corresponds to the first or initial time interval $[0, s]$. Substituting $K_1^*(t; \boldsymbol{\beta}) = K_0 e^{pc^{-1}t}$ into $I = pc^{-1}K$ yields the specific solution for the investment rate, namely, $I_1^*(t; \boldsymbol{\beta}) = pc^{-1}K_0 e^{pc^{-1}t} > 0$. To determine the shadow value of the capital stock, substitute $\mu_1(t) = 0\,\forall\, t \in [0, s]$ into Eq. (6) to find that $1 + \mu_2 = \lambda c^{-1}$. Substituting this result into the costate equation (9) gives $\dot{\lambda} = -pc^{-1}\lambda$, which in light of the solution of the state equation integrates to $\lambda_1(t; \boldsymbol{\beta}) = \bar{\lambda} e^{-pc^{-1}t}$, where $\bar{\lambda}$ is a constant of integration and (again) the subscript 1 indicates that this solution corresponds to the time interval $[0, s]$. Finally, substituting $\lambda_1(t; \boldsymbol{\beta}) = \bar{\lambda} e^{-pc^{-1}t}$ into $1 + \mu_2 = \lambda c^{-1}$ yields the time path of the second Lagrange multiplier $\mu_2(t; \boldsymbol{\beta}) = c^{-1}\bar{\lambda} e^{-pc^{-1}t} - 1$. This completes the determination of the solution of the necessary conditions for the time interval $[0, s]$. Note, however, that we have not yet determined the constant $\bar{\lambda}$ or the switching time s.

To determine the solution for the second or latter time interval $(s, T]$, we assume that $\mu_1(t) > 0\,\forall\, t \in (s, T]$. Using this assumption and the fact that $\mu_1 I = 0$ from Eq. (7) implies that $I_2^*(t; \boldsymbol{\beta}) = 0\,\forall\, t \in (s, T]$, where the subscript 2 indicates that this solution corresponds to the second time interval $(s, T]$. This conclusion verifies the second part of the conjecture about the investment rate given in Eq. (11). This solution implies that the state equation (10) takes the form $\dot{K} = 0$, which has a constant solution, say, $K_2^*(t; \boldsymbol{\beta}) = \bar{K}$, where $\bar{K}$ is a constant of integration to be determined. Because $K(t) > 0\,\forall\, t \in [0, T]$ and $I_2^*(t; \boldsymbol{\beta}) = 0\,\forall\, t \in (s, T]$, we have $\mu_2 pK = 0$ from Eq. (8), which implies that $\mu_2(t) = 0\,\forall\, t \in (s, T]$. This, in turn, implies that the costate equation (9) reduces to $\dot{\lambda} = -p$. Integrating this differential equation with the aid of the transversality condition $\lambda(T) = 0$ gives $\lambda_2(t; \boldsymbol{\beta}) = p[T - t]$ as the specific solution. Finally, because $\mu_2(t) = 0\,\forall\, t \in (s, T]$ the necessary condition (6) reduces to $-c + \lambda + \mu_1 = 0$, thereby yielding $\mu_1(t; \boldsymbol{\beta}) = c - p[T - t]$ as the solution for the first Lagrange multiplier. This completes the determination of the solution of the necessary conditions for the second time interval $(s, T]$. Note, however, that we have not yet determined the constant $\bar{K}$.

To complete the verification of the conjecture, we must determine the two constants of integration $\bar{K}$ and $\bar{\lambda}$, as well as the switching time s. Let's begin with the switching time s. To that end, recall that an interior solution holds only for an instant. Given that the conjectured solution (11) is bang-bang, the only time at which an interior solution holds is at the switching time s, precisely when the investment rate jumps from its upper bound to its lower bound. Thus at time $t = s$, an interior solution is equivalent to $\mu_1(s) = \mu_2(s) = 0$ by Eqs. (7) and (8). Using this observation in Eq. (6) implies that $\lambda_2(s; \boldsymbol{\beta}) = p[T - s] = c$, which when solved for the switching time s gives $s = s^*(\boldsymbol{\beta}) \stackrel{\text{def}}{=} T - cp^{-1}$. Note that the assumption $pT > c$ implies, and is implied by, a switching time in the open interval $(0, T)$. In other words, the assumption $pT > c$ is necessary and sufficient for a bang-bang solution to the control problem. To find the value of $\bar{\lambda}$, remember that Theorem 6.1 asserts that the costate function is a piecewise smooth and thus continuous function of t. This means that at $t = s$, the value of the costate variable from the interval $[0, s]$ must equal the value of the costate variable from the interval $(s, T]$. That is, $\lambda_1(s; \boldsymbol{\beta}) = \lambda_2(s; \boldsymbol{\beta})$, or $\bar{\lambda}e^{-pc^{-1}s} = p[T - s]$. Substituting $s = s^*(\boldsymbol{\beta}) \stackrel{\text{def}}{=} T - cp^{-1}$ in the latter equation and solving for $\bar{\lambda}$ gives $\bar{\lambda} = ce^{pc^{-1}T-1}$. Finally, because the state function is piecewise smooth and hence continuous by the definition of admissibility, we similarly have that $K_1^*(s; \boldsymbol{\beta}) = K_2^*(s; \boldsymbol{\beta})$, or $\bar{K} = K_0 e^{pc^{-1}s}$. Substituting $s = s^*(\boldsymbol{\beta}) \stackrel{\text{def}}{=} T - cp^{-1}$ in the latter equation yields $\bar{K} = K_0 e^{pc^{-1}T-1}$.

Having determined all the constants of integration and the switching time, we restate our solution of the necessary conditions for ease of reference:

$$K^*(t; \boldsymbol{\beta}) = \begin{cases} K_0 e^{pc^{-1}t} & \forall\, t \in [0, s] \\ K_0 e^{pc^{-1}T-1} & \forall\, t \in (s, T], \end{cases} \tag{12}$$

$$I^*(t; \boldsymbol{\beta}) = \begin{cases} pc^{-1}K_0 e^{pc^{-1}t} & \forall\, t \in [0, s] \\ 0 & \forall\, t \in (s, T], \end{cases} \tag{13}$$

$$\lambda(t; \boldsymbol{\beta}) = \begin{cases} ce^{pc^{-1}[T-t]-1} & \forall\, t \in [0, s] \\ p[T - t] & \forall\, t \in (s, T], \end{cases} \tag{14}$$

$$\mu_1(t; \boldsymbol{\beta}) = \begin{cases} 0 & \forall\, t \in [0, s] \\ c - p[T - t] & \forall\, t \in (s, T], \end{cases} \tag{15}$$

$$\mu_2(t; \boldsymbol{\beta}) = \begin{cases} e^{pc^{-1}[T-t]-1} - 1 & \forall\, t \in [0, s] \\ 0 & \forall\, t \in (s, T], \end{cases} \tag{16}$$

where $s = s^*(\boldsymbol{\beta}) \stackrel{\text{def}}{=} T - cp^{-1}$ is the switching time. Observe that the solutions for the Lagrange multipliers are continuous functions of t, even though Theorem 6.1 permits them to be but piecewise continuous functions of t. We leave formal verification of this fact for a mental exercise.

To close out this example, we derive and then discuss the comparative dynamics of an increase in the output price p. Differentiating Eqs. (12) through (14) yields

$$\frac{\partial K^*(t; \boldsymbol{\beta})}{\partial p} = \begin{cases} c^{-1}tK_0 e^{pc^{-1}t} > 0 & \forall\, t \in [0, s] \\ c^{-1}TK_0 e^{pc^{-1}T-1} > 0 & \forall\, t \in (s, T], \end{cases} \tag{17}$$

$$\frac{\partial I^*(t;\boldsymbol{\beta})}{\partial p} = \begin{cases} c^{-1}K_0 e^{pc^{-1}t}[1 + pc^{-1}t] > 0 & \forall\, t \in [0, s] \\ 0 & \forall\, t \in (s, T], \end{cases} \tag{18}$$

$$\frac{\partial \lambda(t;\boldsymbol{\beta})}{\partial p} = \begin{cases} [T - t]e^{pc^{-1}[T-t]-1} > 0 & \forall\, t \in [0, s] \\ [T - t] \geq 0 & \forall\, t \in (s, T]. \end{cases} \tag{19}$$

These computations show that an increase in the price of the firm's product results in the investment rate, capital stock, and shadow value of the capital stock being higher throughout the planning horizon. The higher capital stock is a result of the higher maximum rate of investment in the interval $[0, s]$ and the lengthening of that interval implied by the delay in the switching time, that is, $\partial s^*(\boldsymbol{\beta})/\partial p = cp^{-2} > 0$. Given that the final good has gone up in value, so too has the shadow value of the capital stock, since capital is the only input used to produce the good.

In wrapping up this example, observe that the comparative dynamics at time $t = s^*(\boldsymbol{\beta})$ are a little bit more complicated. The reason is that there are two effects resulting from an increase in the output price, namely, the *explicit* or *direct effect* given by Eqs. (17) through (19), and an *indirect effect* given by the chain rule. For example, for the capital stock, the indirect effect is given by

$$\dot{K}^*(s^*(\boldsymbol{\beta}); \boldsymbol{\beta})\frac{\partial s^*(\boldsymbol{\beta})}{\partial p} = \left[pc^{-1}K_0 e^{pc^{-1}s^*(\boldsymbol{\beta})}\right] cp^{-2} > 0.$$

In other words, the *total effect* of an increase in the output price on the capital stock at time $t = s^*(\boldsymbol{\beta})$ is given by the chain rule as

$$\frac{\partial K^*(s^*(\boldsymbol{\beta}); \boldsymbol{\beta})}{\partial p} + \dot{K}^*(s^*(\boldsymbol{\beta}); \boldsymbol{\beta})\frac{\partial s^*(\boldsymbol{\beta})}{\partial p} = c^{-1}s^*(\boldsymbol{\beta})K_0 e^{pc^{-1}s^*(\boldsymbol{\beta})} + \left[pc^{-1}K_0 e^{pc^{-1}s^*(\boldsymbol{\beta})}\right] cp^{-2} > 0.$$

We leave the remaining comparative dynamics for a mental exercise.

Let's now state and prove a generalization of the Mangasarian sufficiency theorem we have seen in simpler forms in earlier chapters.

Theorem 6.2 (Mangasarian Sufficient Conditions, Mixed Constraints): *Let* $(\mathbf{z}(t), \mathbf{v}(t))$ *be an admissible pair for problem (1). Suppose that* $(\mathbf{z}(t), \mathbf{v}(t))$ *satisfies the necessary conditions of Theorem 6.1 for problem (1) with costate vector* $\boldsymbol{\lambda}(t)$ *and Lagrange multiplier vector* $\boldsymbol{\mu}(t)$*, and let* $L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \overset{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u})$ *be the value of the Lagrangian function. If* $L(\cdot)$ *is a concave function of* $(\mathbf{x}, \mathbf{u})\, \forall\, t \in [t_0, t_1]$ *over an open convex set containing all the admissible values of* $(\mathbf{x}(\cdot), \mathbf{u}(\cdot))$ *when the costate vector is* $\boldsymbol{\lambda}(t)$ *and Lagrange multiplier vector is* $\boldsymbol{\mu}(t)$*, then* $\mathbf{v}(t)$ *is an optimal control and* $(\mathbf{z}(t), \mathbf{v}(t))$ *yields the global maximum of* $J[\cdot]$*. If* $L(\cdot)$ *is a strictly concave function under the same conditions, then* $(\mathbf{z}(t), \mathbf{v}(t))$ *yields the unique global maximum of* $J[\cdot]$*.*

Proof: Let $(\mathbf{x}(t), \mathbf{u}(t))$ be any admissible pair. By hypothesis, $L(\cdot)$ is a $C^{(1)}$ concave function of $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$. It therefore follows from Theorem 21.3 in Simon and Blume (1994) that

$$\begin{aligned} L(t, \mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) &\leq L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \\ &\quad + L_{\mathbf{x}}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))\,[\mathbf{x}(t) - \mathbf{z}(t)] \\ &\quad + L_{\mathbf{u}}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))\,[\mathbf{u}(t) - \mathbf{v}(t)], \end{aligned} \tag{20}$$

for every $t \in [t_0, t_1]$. Using the fact that $L_{\mathbf{u}}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \equiv \mathbf{0}'_M$ by Theorem 6.1, and then integrating both sides of the resulting reduced inequality over the interval $[t_0, t_1]$ using the definitions of $L(\cdot)$ and $J[\cdot]$, yields

$$\begin{aligned} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \leq J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)] &+ \int_{t_0}^{t_1} \boldsymbol{\lambda}(t)'\,[\mathbf{g}(t, \mathbf{z}(t), \mathbf{v}(t)) - \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t))]\,dt \\ &+ \int_{t_0}^{t_1} \boldsymbol{\mu}(t)'\,[\mathbf{h}(t, \mathbf{z}(t), \mathbf{v}(t)) - \mathbf{h}(t, \mathbf{x}(t), \mathbf{u}(t))]\,dt \\ &+ \int_{t_0}^{t_1} L_{\mathbf{x}}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))\,[\mathbf{x}(t) - \mathbf{z}(t)]\,dt. \end{aligned} \tag{21}$$

By admissibility, $\dot{\mathbf{z}}(t) \equiv \mathbf{g}(t, \mathbf{z}(t), \mathbf{v}(t))$ and $\dot{\mathbf{x}}(t) \equiv \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t))$ for every $t \in [t_0, t_1]$, whereas Theorem 6.1 implies that $\dot{\boldsymbol{\lambda}}(t)' \equiv -L_{\mathbf{x}}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t))$ for every $t \in [t_0, t_1]$. Substituting these three results in Eq. (21) gives

$$\begin{aligned} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \leq J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)] &+ \int_{t_0}^{t_1} [\boldsymbol{\lambda}(t)'[\dot{\mathbf{z}}(t) - \dot{\mathbf{x}}(t)] + \dot{\boldsymbol{\lambda}}(t)'[\mathbf{z}(t) - \mathbf{x}(t)]]\,dt \\ &+ \int_{t_0}^{t_1} \boldsymbol{\mu}(t)'\,[\mathbf{h}(t, \mathbf{z}(t), \mathbf{v}(t)) - \mathbf{h}(t, \mathbf{x}(t), \mathbf{u}(t))]\,dt. \end{aligned} \tag{22}$$

Moreover, Theorem 6.1 also implies that (i) $\mu_k(t)h^k(t, \mathbf{z}(t), \mathbf{v}(t)) \equiv 0$ for $k = 1, 2, \ldots, K$ because $\mu_k(t)h^k(t, \mathbf{z}(t), \mathbf{v}(t)) \equiv 0$ for $k = 1, 2, \ldots, K'$ and $h^k(t, \mathbf{z}(t), \mathbf{v}(t)) \equiv 0$ for $k = K'+1, K'+2, \ldots, K$, (ii) $\mu_k(t)h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0$ for $k = 1, 2, \ldots, K'$ by virtue of $\mu_k(t) \geq 0$ and $h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0$ for $k = 1, 2, \ldots, K'$, and (iii) $\mu_k(t)h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \equiv 0$ for $k = K'+1, K'+2, \ldots, K$ on account of $h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \equiv 0$ for $k = K'+1, K'+2, \ldots, K$. These three implications of

Theorem 6.1 therefore imply that

$$\int_{t_0}^{t_1} \boldsymbol{\mu}(t)' \left[\mathbf{h}(t, \mathbf{z}(t), \mathbf{v}(t)) - \mathbf{h}(t, \mathbf{x}(t), \mathbf{u}(t))\right] dt \leq 0. \tag{23}$$

Using the inequality in Eq. (23) permits Eq. (22) to be rewritten in the reduced form

$$J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \leq J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)] + \int_{t_0}^{t_1} [\boldsymbol{\lambda}(t)'[\dot{\mathbf{z}}(t) - \dot{\mathbf{x}}(t)] + \dot{\boldsymbol{\lambda}}(t)'[\mathbf{z}(t) - \mathbf{x}(t)]]\, dt. \tag{24}$$

To wrap up the proof, simply note that

$$\begin{aligned}\frac{d}{dt}[\boldsymbol{\lambda}(t)'[\mathbf{z}(t) - \mathbf{x}(t)]] &= \boldsymbol{\lambda}(t)' [\dot{\mathbf{z}}(t) - \dot{\mathbf{x}}(t)] + [\mathbf{z}(t) - \mathbf{x}(t)]' \dot{\boldsymbol{\lambda}}(t) \\ &= \boldsymbol{\lambda}(t)' [\dot{\mathbf{z}}(t) - \dot{\mathbf{x}}(t)] + \dot{\boldsymbol{\lambda}}(t)'[\mathbf{z}(t) - \mathbf{x}(t)],\end{aligned}$$

and substitute this result into Eq. (24) to get

$$\begin{aligned}J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] &\leq J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)] + \int_{t_0}^{t_1} \frac{d}{dt}[\boldsymbol{\lambda}(t)'[\mathbf{z}(t) - \mathbf{x}(t)]]\, dt \\ &= J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)] + \boldsymbol{\lambda}(t)' [\mathbf{z}(t) - \mathbf{x}(t)]\Big|_{t=t_0}^{t=t_1} \\ &= J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)],\end{aligned}$$

because by admissibility, we have $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{z}(t_0) = \mathbf{x}_0$, $\mathbf{x}(t_1) = \mathbf{x}_1$, and $\mathbf{z}(t_1) = \mathbf{x}_1$. We have thus shown that $J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \leq J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)]$ for all admissible functions $(\mathbf{x}(\cdot), \mathbf{u}(\cdot))$, just as we wished to. If $L(\cdot)$ is a strictly concave function of $(\mathbf{x}, \mathbf{u})\, \forall\, t \in [t_0, t_1]$, then the inequality in Eq. (20) becomes strict if either $\mathbf{x}(t) \neq \mathbf{z}(t)$ or $\mathbf{u}(t) \neq \mathbf{v}(t)$ for some $t \in [t_0, t_1]$. Then $J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] < J[\mathbf{z}(\cdot), \mathbf{v}(\cdot)]$ follows. This shows that any admissible pair of functions $(\mathbf{x}(\cdot), \mathbf{u}(\cdot))$ that are not identically equal to $(\mathbf{z}(\cdot), \mathbf{v}(\cdot))$ are suboptimal. Q.E.D.

An important feature of this sufficiency theorem is that the rank constraint qualification is *not* required, in sharp contrast to the necessary conditions of Theorem 6.1. It is also noteworthy that this theorem can be strengthened to give an Arrow-type sufficiency result. We will pursue this very shortly.

The ensuing lemma gives sufficient conditions for the Lagrangian of problem (1) to be a concave function of $(\mathbf{x}, \mathbf{u})\, \forall\, t \in [t_0, t_1]$. Its proof is left for a mental exercise.

Lemma 6.1: *The Lagrangian function $L(\cdot)$ defined by*

$$L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u})$$

for problem (1) is a concave function of $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$ *if the following conditions hold:*

(i) $f(\cdot)$ *is concave in* $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$,
(ii) $\lambda_n g^n(\cdot)$, $n = 1, 2, \ldots, N$, *is concave in* $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$,
(iii) $\mu_k h^k(\cdot)$, $k = 1, 2, \ldots, K'$, *is concave in* $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$,
(iv) $\mu_k h^k(\cdot)$, $k = K' + 1, K' + 2, \ldots, K$, *is concave in* $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$.

A few remarks on the lemma are in order. First, it should be clear that condition (ii) holds if $g^n(\cdot)$ is concave in $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$ and $\lambda_n(t) \geq 0$, $n = 1, 2, \ldots, N$, or if $g^n(\cdot)$ is convex in $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$ and $\lambda_n(t) \leq 0$, $n = 1, 2, \ldots, N$. Second, no such remark is required for condition (iii) because $\mu_k(t) \geq 0$ for $k = 1, 2, \ldots, K'$. Third, condition (iv) holds if $h^k(\cdot)$ is concave in $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$ and $\mu_k(t) \geq 0$, $k = K' + 1, K' + 2, \ldots, K$, or if $h^k(\cdot)$ is convex in $(\mathbf{x}, \mathbf{u})\,\forall\, t \in [t_0, t_1]$ and $\mu_k(t) \leq 0$, $k = K' + 1, K' + 2, \ldots, K$.

Let's pause briefly so that we can see Theorem 6.2 in action. Specifically, we will use it to determine if the solution of the necessary conditions in Example 6.1 is a solution to the posed control problem.

Example 6.2: Recall that the Lagrangian for the control problem of Example 6.1 is given by

$$L(K, I, \lambda, \mu_1, \mu_2; c, p) \stackrel{\text{def}}{=} pK - cI + \lambda I + \mu_1 I + \mu_2[pK - cI].$$

The Hessian matrix of the Lagrangian with respect to the capital stock and investment rate is

$$\begin{bmatrix} L_{II} & L_{IK} \\ L_{KI} & L_{KK} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which is negative semidefinite. Hence $L(\cdot)$ is a concave function of (K, I) for all $t \in [0, T]$. Thus the solution of the necessary conditions is a solution of the posed control problem. The same conclusion is also arrived at by simply observing that $L(\cdot)$ is a linear and thus a concave function of (K, I) for all $t \in [0, T]$. Note also that because $L(\cdot)$ is a linear function of (K, I), assumptions (A.1) through (A.5) are satisfied.

Just as we did in Chapter 3, we now provide an alternative and equivalent statement of the necessary conditions of Theorem 6.1 in terms of the maximized Hamiltonian. This permits us to introduce a generalized version of the Arrow-type sufficiency theorem alluded to above. First, we introduce the definition of the maximized Hamiltonian for the general control problem under consideration in this chapter.

Definition 6.2: For the control problem (1), the *maximized Hamiltonian* $M(\cdot)$ is defined as

$$M(t,\mathbf{x},\boldsymbol{\lambda}) \stackrel{\text{def}}{=} \max_{\mathbf{u}}\{H(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda}) \text{ s.t. } u \in U(t,\mathbf{x})\}, \tag{25}$$

where $U(t,\mathbf{x}) \stackrel{\text{def}}{=} \{\mathbf{u} : h^k(t,\mathbf{x},\mathbf{u}) \geq 0,\ k = 1,\ldots,K',\ h^k(t,\mathbf{x},\mathbf{u}) = 0,\ k = K'+1,\ldots,K\}$ is the control set and $H(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda}) \stackrel{\text{def}}{=} f(t,\mathbf{x},\mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t,\mathbf{x},\mathbf{u})$ is the Hamiltonian.

Given that problem (25) is a static constrained optimization problem, the necessary conditions implied by the existence of a maximizing control are a subset of those from Theorem 6.1, to wit,

$$\frac{\partial L}{\partial u_m}(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda},\boldsymbol{\mu}) = 0, \quad m = 1,2,\ldots,M,$$

$$\frac{\partial L}{\partial \mu_k}(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda},\boldsymbol{\mu}) \geq 0,\ \mu_k \geq 0,\ \mu_k\frac{\partial L}{\partial \mu_k}(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda},\boldsymbol{\mu}) = 0, \quad k = 1,2,\ldots,K',$$

$$\frac{\partial L}{\partial \mu_k}(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda},\boldsymbol{\mu}) = 0, \quad k = K'+1, K'+2,\ldots,K,$$

where $L(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda},\boldsymbol{\mu}) \stackrel{\text{def}}{=} f(t,\mathbf{x},\mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t,\mathbf{x},\mathbf{u}) + \boldsymbol{\mu}'\mathbf{h}(t,\mathbf{x},\mathbf{u})$ is the Lagrangian for problem (25) and $\boldsymbol{\mu} \in \Re^K$ is the vector of Lagrange multipliers for the mixed constraints. As remarked earlier, these, in principle, define $\mathbf{u}$ and $\boldsymbol{\mu}$ as functions of time, the state variables, the costate variables, and any parameters of the problem, say, $\mathbf{u} = \hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda})$ and $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}(t,\mathbf{x},\boldsymbol{\lambda})$. We acknowledge that $\mathbf{u} = \hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda})$ is the maximizing value of the control variable by using the notation

$$\hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda}) \stackrel{\text{def}}{=} \arg\max_{\mathbf{u}}\{H(t,\mathbf{x},\mathbf{u},\boldsymbol{\lambda}) \text{ s.t. } u \in U(t,\mathbf{x})\}.$$

Substituting $\mathbf{u} = \hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda})$ into the Hamiltonian $H(\cdot)$ yields the value of the maximized Hamiltonian $M(\cdot)$, that is,

$$M(t,\mathbf{x},\boldsymbol{\lambda}) \equiv H(t,\mathbf{x},\hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda}),\boldsymbol{\lambda}) = f(t,\mathbf{x},\hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda})) + \boldsymbol{\lambda}'\ \mathbf{g}(t,\mathbf{x},\hat{\mathbf{u}}(t,\mathbf{x},\boldsymbol{\lambda})). \tag{26}$$

Equation (26) demonstrates how one would go about constructing the maximized Hamiltonian in practice. Given Definition 6.2, we are now in a position to restate Theorem 6.1 in terms of the maximized Hamiltonian $M(\cdot)$.

Theorem 6.3 (Necessary Conditions, Mixed Constraints): *Let* $(\mathbf{z}(t),\mathbf{v}(t))$ *be an admissible pair for problem (1), and assume that the rank constraint qualification is satisfied. Then if* $(\mathbf{z}(t),\mathbf{v}(t))$ *yields the absolute maximum of* $J[\cdot]$, *it is necessary that there exist a piecewise smooth vector-valued function* $\boldsymbol{\lambda}(\cdot) \stackrel{\text{def}}{=} (\lambda_1(\cdot),\lambda_2(\cdot),\ldots,\lambda_N(\cdot))$ *and a piecewise continuous vector-valued Lagrange*

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

multiplier function $\boldsymbol{\mu}(\cdot) \stackrel{\text{def}}{=} (\mu_1(\cdot), \mu_2(\cdot), \ldots, \mu_K(\cdot))$*, such that for all* $t \in [t_0, t_1]$*,*

$$\mathbf{v}(t) = \arg\max_{\mathbf{u}} \left\{ H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t)) \; s.t. \; \mathbf{u} \in U(t, \mathbf{z}(t)) \right\},$$

that is, if

$$M(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)) \stackrel{\text{def}}{=} \max_{\mathbf{u}} \left\{ H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t)) \; s.t. \; \mathbf{u} \in U(t, \mathbf{z}(t)) \right\},$$

then

$$M(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)) \equiv H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)),$$

or equivalently

$$H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)) \geq H(t, \mathbf{z}(t), \mathbf{u}, \boldsymbol{\lambda}(t)) \, \forall \, \mathbf{u} \in U(t, \mathbf{z}(t)),$$

where

$$U(t, \mathbf{x}(t)) \stackrel{\text{def}}{=} \{\mathbf{u}(\cdot) : h^k(t, \mathbf{x}(t)\mathbf{u}(t)) \geq 0, \; k = 1, \ldots, K', \; h^k(t, \mathbf{x}(t), \mathbf{u}(t)) = 0, \\ k = K' + 1, \ldots, K\}$$

is the control set and $\mathbf{v}(t) \stackrel{\text{def}}{=} \hat{\mathbf{u}}(t, \mathbf{z}(t), \boldsymbol{\lambda}(t))$*. Because the rank constraint qualification is assumed to hold, the above necessary condition implies that*

$$\frac{\partial L}{\partial u_m}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad m = 1, 2, \ldots, M,$$

$$\frac{\partial L}{\partial \mu_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) \geq 0, \; \mu_k(t) \geq 0,$$

$$\mu_k(t)\frac{\partial L}{\partial \mu_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad k = 1, 2, \ldots, K',$$

$$\frac{\partial L}{\partial \mu_k}(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = 0, \quad k = K' + 1, K' + 2, \ldots, K,$$

where

$$\begin{aligned} L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &\stackrel{\text{def}}{=} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{x}, \mathbf{u}) \\ &= f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}) + \sum_{k=1}^{K} \mu_k h^k(t, \mathbf{x}, \mathbf{u}) \end{aligned}$$

is the Lagrangian function. Furthermore, except for the points of discontinuities of $\mathbf{v}(t)$*,*

$$\dot{z}_n(t) = \frac{\partial M}{\partial \lambda_n}(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)) = g^n(t, \mathbf{z}(t), \mathbf{v}(t)), \quad n = 1, 2, \ldots, N,$$

$$\dot{\lambda}_n(t) = -\frac{\partial M}{\partial x_n}(t, \mathbf{z}(t), \boldsymbol{\lambda}(t)), \quad n = 1, 2, \ldots, N,$$

where the above notation means that the functions are first differentiated with respect to the particular variable and then evaluated at $(t, \mathbf{z}(t), \boldsymbol{\lambda}(t))$.

Notice that the real difference between Theorems 6.1 and 6.3 lies in the statement of the canonical equations. That the two versions are equivalent is a simple matter to demonstrate via the envelope theorem and is therefore left for a mental exercise. Given Theorem 6.3, we can now state the Arrow-type sufficiency theorem. The proof follows that of Theorem 6.2 and so is left for a mental exercise too.

Theorem 6.4 (Arrow Sufficiency Theorem, Mixed Constraints): *Let* $(\mathbf{z}(t), \mathbf{v}(t))$ *be an admissible pair for problem (1). Suppose that* $(\mathbf{z}(t), \mathbf{v}(t))$ *satisfy the necessary conditions of Theorem 6.3 for problem (1) with costate vector* $\boldsymbol{\lambda}(t)$ *and Lagrange multiplier vector* $\boldsymbol{\mu}(t)$*, and let* $M(t, \mathbf{x}, \boldsymbol{\lambda}) \stackrel{\text{def}}{=} \max_{\mathbf{u}}\{H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda})$ *s.t.* $\mathbf{u} \in U(t, \mathbf{x})\}$ *be the value of the maximized Hamiltonian function. If* $M(\cdot)$ *is a concave function of* $\mathbf{x}\, \forall\, t \in [t_0, t_1]$ *over an open convex set containing all the admissible values of* $\mathbf{x}(\cdot)$ *when the costate vector is* $\boldsymbol{\lambda}(t)$*, then* $\mathbf{v}(t)$ *is an optimal control and* $(\mathbf{z}(t), \mathbf{v}(t))$ *yields the global maximum of* $J[\cdot]$*. If* $M(\cdot)$ *is a strictly concave function of* $\mathbf{x}\, \forall\, t \in [t_0, t_1]$ *under the same conditions, then* $(\mathbf{z}(t), \mathbf{v}(t))$ *yields the unique global maximum of* $J[\cdot]$ *and* $\mathbf{z}(t)$ *is unique, but* $\mathbf{v}(t)$ *is not necessarily unique.*

The Arrow sufficiency theorem replaces the assumption of concavity of the Lagrangian $L(\cdot)$ in $(\mathbf{x}, \mathbf{u})$ from the Mangasarian theorem, with the assumption that the maximized Hamiltonian $M(\cdot)$ is concave in $\mathbf{x}$. Note, however, that checking the curvature properties of a derived function such as $M(\cdot)$ can be more difficult than checking the curvature properties of $L(\cdot)$. By Theorem 3.5, we know that Mangasarian's sufficiency theorem is a *special case* of Arrow's sufficiency theorem. This means that even if $L(\cdot)$ is not concave in $(\mathbf{x}, \mathbf{u})$ it is still possible that $M(\cdot)$ is concave in $\mathbf{x}$, so that Arrow's theorem applies to a larger class of problems.

In Chapter 9, we take up the study of the envelope theorem for optimal control problems. This will (i) permit us to achieve a thorough and complete economic interpretation of the costate vector, (ii) pave the way for the introduction of the primal-dual method of comparative dynamics, and (iii) establish, with relative ease, the numerous transversality conditions that are an integral part of the necessary conditions when the planning horizon and endpoints are decision variables.

Before doing so, however, the next two chapters introduce a class of control problems, scilicet, isoperimetric problems, that are an important class of optimal control problems in several fields of economics. These two chapters may be skipped on a first reading, without loss of continuity.

MENTAL EXERCISES

6.1 In Example 6.1, prove that
  (a) admissible $K(t) > 0\, \forall\, t \in [0, T]$, and
  (b) the Lagrange multipliers are continuous functions of t.

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

6.2 Derive and economically interpret the comparative dynamics of an increase in c and T in Example 6.1. Do not concern yourself with the Lagrange multipliers.

6.3 Prove Lemma 6.1.

6.4 Prove that Theorem 6.3 is equivalent to Theorem 6.1.

6.5 Prove Theorem 6.4.

6.6 Consider the optimal control problem

$$\min_{u(\cdot),x_T} \int_0^T [x(t) + u(t)]\, dt$$

$$\text{s.t.} \quad \dot{x}(t) = -u(t),\ x(0) = 1,\ x(T) = x_T,$$

$$h^1(t, x(t), u(t)) \overset{\text{def}}{=} u(t) \geq 0,$$

$$h^2(t, x(t), u(t)) \overset{\text{def}}{=} x(t) - u(t) \geq 0,$$

where $T > 1$.

(a) Show that all admissible solutions satisfy $x(t) > 0\, \forall\, t \in [0, T]$.
(b) Show that the rank constraint qualification is satisfied for all admissible pairs.
(c) Find a solution to the necessary conditions.
(d) Is the solution of the necessary conditions a solution of the control problem? Show your work and explain.

6.7 Consider the optimal control problem

$$\max_{u(\cdot),x_T} \int_0^T [x(t) - u(t)] dt$$

$$\text{s.t.} \quad \dot{x}(t) = u(t), x(0) = x_0, x(T) = x_T,$$

$$h^1(t, x(t), u(t)) \overset{\text{def}}{=} u(t) \geq 0,$$

$$h^2(t, x(t), u(t)) \overset{\text{def}}{=} x(t) - u(t) \geq 0.$$

(a) Show that all admissible solutions satisfy $x(t) > 0\, \forall\, t \in [0, T]$.
(b) Show that the rank constraint qualification is satisfied for all admissible pairs.
(c) Find a solution to the necessary conditions.
(d) Is the solution of the necessary conditions a solution of the control problem? Show your work and explain.

6.8 Show that the *isoperimetric* optimal control problem

$$\max_{\mathbf{u}(\cdot)} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \overset{\text{def}}{=} \int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t))\, dt$$

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \ \mathbf{x}(t_0) = \mathbf{x}_0, \ \mathbf{x}(t_1) = \mathbf{x}_1,$$

$$\int_{t_0}^{t_1} G^k(t, \mathbf{x}(t), \mathbf{u}(t))\, dt = \gamma_k, \quad k = 1, 2, \ldots, K,$$

can be rewritten as an equivalent mixed constraint optimal control problem. **Hint:** Define a new vector of state variables.

6.9 Solve the optimal control problem

$$\max_{u(\cdot), x_{1/2}} \int_{a}^{1/2} [-(u(t))^2 - x(t)]\, dt$$

$$\text{s.t.} \quad \dot{x}(t) = -u(t), \ x(a) = 7/4, \ x(1/2) = x_{1/2},$$

$$h^1(t, x(t), u(t)) \stackrel{\text{def}}{=} x(t) - u(t) \geq 0,$$

where $a = -1 - \ln 3$.

6.10 Solve the optimal control problem

$$\max_{u_1(\cdot), u_2(\cdot), x_1} \int_{0}^{1} [u_2(t) - x(t)]\, dt$$

$$\text{s.t.} \quad \dot{x}(t) = u_1(t), \ x(0) = 1/8, \ x(1) = x_1,$$

$$h^1(t, x(t), u_1(t), u_2(t)) \stackrel{\text{def}}{=} u_1(t) \geq 0,$$

$$h^2(t, x(t), u_1(t), u_2(t)) \stackrel{\text{def}}{=} 1 - u_1(t) \geq 0,$$

$$h^3(t, x(t), u_1(t), u_2(t)) \stackrel{\text{def}}{=} x(t) - (u_2(t))^2 \geq 0.$$

Note that there are two control variables and one state variable in this problem.

6.11 Consider the optimal control problem

$$\max_{\mathbf{u}(\cdot)} J[\mathbf{x}(\cdot), \mathbf{u}(\cdot)] \stackrel{\text{def}}{=} \int_{t_0}^{t_1} f(t, \mathbf{x}(t), \mathbf{u}(t))\, dt$$

$$\text{s.t.} \quad \dot{\mathbf{x}}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)), \ \mathbf{x}(t_0) = \mathbf{x}_0, \ \mathbf{x}(t_1) = \mathbf{x}_1,$$

$$h^k(t, \mathbf{x}(t), \mathbf{u}(t)) \geq 0, \quad k = 1, 2, \ldots, K',$$

$$h^k(t, \mathbf{x}(t), \mathbf{u}(t)) = 0, \quad k = K' + 1, K' + 2, \ldots, K.$$

Let $(\mathbf{z}(t), \mathbf{v}(t))$ be the optimal pair, $\boldsymbol{\lambda}(t)$ the corresponding value of the costate vector, and $\boldsymbol{\mu}(t)$ the corresponding value of the Lagrange multiplier vector. Define the Hamiltonian as

$$H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) \stackrel{\text{def}}{=} f(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda}'\mathbf{g}(t, \mathbf{x}, \mathbf{u}) = f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}),$$

and the Lagrangian as

$$\begin{aligned} L(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &\stackrel{\text{def}}{=} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) + \boldsymbol{\mu}'\mathbf{h}(t, \mathbf{u}) \\ &= f(t, \mathbf{x}, \mathbf{u}) + \sum_{n=1}^{N} \lambda_n g^n(t, \mathbf{x}, \mathbf{u}) + \sum_{k=1}^{K} \mu_k h^k(t, \mathbf{x}, \mathbf{u}). \end{aligned}$$

Assume that $f(\cdot) \in C^{(1)}$, $\mathbf{g}(\cdot) \in C^{(1)}$, and $\mathbf{h}(\cdot) \in C^{(1)}$ in $(t, \mathbf{x}, \mathbf{u})$.

(a) Prove that

$$\frac{d}{dt} L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)) = \frac{\partial}{\partial t} L(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t), \boldsymbol{\mu}(t)).$$

(b) Prove that if the optimal control problem is autonomous, that is, the independent variable t doesn't enter the functions $f(\cdot)$, $\mathbf{g}(\cdot)$, or $\mathbf{h}(\cdot)$ explicitly, that is, $f_t(t, \mathbf{x}, \mathbf{u}) \equiv 0$, $\mathbf{g}_t(t, \mathbf{x}, \mathbf{u}) \equiv \mathbf{0}_N$, and $\mathbf{h}_t(t, \mathbf{u}) \equiv \mathbf{0}_K$, then $L(\cdot)$ is constant along the optimal path.

(c) Prove that

$$\frac{d}{dt} H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)) = \frac{\partial}{\partial t} H(t, \mathbf{z}(t), \mathbf{v}(t), \boldsymbol{\lambda}(t)).$$

Note that this result is *not* the same as that in Mental Exercise 2.31.

(d) Prove that if the optimal control problem is autonomous, then $H(\cdot)$ is constant along the optimal path too.

6.12 *Rational Procrastination*. Here's a typical situation faced by students at universities all across the world. A research paper is assigned by the professor of a class the first day of the term, say, $t = 0$, and is due at the end of the term, say, $T > 0$ hours later. At University of California, Davis, this would mean that the research paper is due in $24 \times 7 \times 10 = 1{,}680$ hours, or 10 weeks, from the present. The total effort required by the typical student to complete the paper is known to be $\varepsilon \in (0, T)$ hours. Define $e(t) \in [0, 1]$ as the proportion of each hour that the student devotes to working on the paper (i.e., research effort), and define $\ell(t) \in [0, 1]$ as the proportion of each hour that the student devotes to leisure activities. It is assumed that the student will complete the term paper by the required date, thereby implying the isoperimetric constraint

$$\int_0^T e(t)dt = \varepsilon.$$

Because each hour is made up entirely of leisure time and research effort, we also have the equality constraint that $e(t) + \ell(t) = 1$ for all $t \in [0, T]$. The instantaneous preferences of the typical student are defined over leisure time, a good, and denoted by $U(\ell)$, where $U(\cdot) \in C^{(2)}$, $U'(\ell) > 0$, and $U''(\ell) < 0$ for all $\ell(t) \in (0, 1)$. We assume that $e(t) \in (0, 1)$ and $\ell(t) \in (0, 1)$ for all $t \in [0, T]$ in an optimal plan, thereby ruling out these constraints from binding. These two assumptions will simplify the analysis considerably. The student is asserted to maximize the present discounted value of utility over the term, subject to

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press

completing the research paper. Hence the optimal control problem faced by the typical student can be stated as

$$V(\boldsymbol{\beta}) \stackrel{\text{def}}{=} \max_{e(\cdot),\ell(\cdot)} \int_0^T U(\ell(t))\, e^{-rt}\, dt$$

$$\text{s.t.} \quad \int_0^T e(t)dt = \varepsilon,$$

$$e(t) + \ell(t) = 1,$$

where $r > 0$ is the student's intertemporal rate of time preference and $\boldsymbol{\beta} \stackrel{\text{def}}{=} (\varepsilon, r, T) \in \Re^3_{++}$. Assume that the pair of curves $(e^*(t;\boldsymbol{\beta}), \ell^*(t;\boldsymbol{\beta}))$ is a solution to the necessary conditions of the optimal control problem, with corresponding costate variable $\lambda(t;\boldsymbol{\beta})$ and Lagrange multiplier $\mu(t;\boldsymbol{\beta})$. Note that the above optimal control problem is *not* in standard form because the state equation is absent.

(a) Convert the above *isoperimetric* optimal control problem to one in standard form by defining an appropriate state variable, say, $x(t)$, and calling it cumulative effort. Also include the optimal value function in your problem statement. Please do not substitute the equality constraint $e(t) + \ell(t) = 1$ out of the problem.

(b) Write the constraint $e(t) + \ell(t) = 1$ as $1 - e - \ell$ in forming the Lagrangian function, and then derive the necessary conditions for this problem.

(c) Prove that the triplet $(x^*(t;\boldsymbol{\beta}), e^*(t;\boldsymbol{\beta}), \ell^*(t;\boldsymbol{\beta}))$ is a solution of the optimal control problem in standard form.

(d) Prove that $\lambda(t;\boldsymbol{\beta}) = \mu(t;\boldsymbol{\beta})\, \forall\, t \in [0, T]$ and that $\lambda(t;\boldsymbol{\beta}) > 0\, \forall\, t \in [0, T]$. Provide an economic interpretation of these two results. Does the latter result make sense? Explain.

(e) Prove that $\dot{\ell}^*(t;\boldsymbol{\beta}) < 0\, \forall\, t \in [0, T]$ and $\dot{e}^*(t;\boldsymbol{\beta}) > 0\, \forall\, t \in [0, T]$. Provide an economic interpretation. Is this the rational procrastination result alluded to in the problem? Why or why not?

6.13 *Fattening up the Fish on the Farm*. You are the sole owner of a fish farm and wish to minimize the present discounted value of the fish feeding costs over the fixed planning horizon $[0, T]$. Let $w(t)$ be the weight of the stock of fish at time t, $w(0) = w_0 > 0$ be the fixed initial weight of the fish stock when the fattening plan begins, and $w(T) = w_T > 0$ be the required (and fixed) terminal weight of the fish stock when the fattening operation comes to a close. Assume that $w_T > w_0$, so that the terminal weight of the fish stock is larger than its initial weight. Thus, given the initial weight of the fish stock of $w(0) = w_0 > 0$, you want to minimize the present discounted value of the feeding costs of the stock so as to bring the fish up to the required weight $w(T) = w_T > 0$ over the period $[0, T]$. Let $u(t)$ be the feeding rate of the stock of fish at time t and $r > 0$ be the discount rate that you use in discounting all future cash flows. The instantaneous feeding cost function is given by $C(u) \stackrel{\text{def}}{=} \frac{1}{2}u^2$. The rate of change in the weight of the fish stock is proportional to the feeding rate; hence the differential equation governing the weight dynamics is $\dot{w} = \alpha^{-1}u$, where $\alpha > 0$ is the inverse of the marginal product of the feeding rate. For notational clarity, define $\boldsymbol{\beta} \stackrel{\text{def}}{=} (\alpha, r, w_0, T, w_T)$ as the parameter vector of the problem.

(a) Set up this problem as an optimal control problem, and define the optimal value function mathematically, say, $C^*(\cdot)$. Provide an economic interpretation of the optimal value function.

(b) Derive the solution of the necessary conditions, say, $(w^*(t; \boldsymbol{\beta}), u^*(t; \boldsymbol{\beta}))$, and let $\lambda(t; \boldsymbol{\beta})$ be the corresponding time path of the costate variable.

(c) Prove that the pair $(w^*(t; \boldsymbol{\beta}), u^*(t; \boldsymbol{\beta}))$ is the solution to the control problem. Is it the unique solution? Clearly explain your reasoning.

(d) Prove that

$$\frac{\partial u^*}{\partial \alpha}(t; \boldsymbol{\beta}) > 0 \, \forall \, t \in [0, T] \text{ and } \frac{\partial w^*}{\partial \alpha}(t; \boldsymbol{\beta}) \equiv 0 \, \forall \, t \in [0, T].$$

Provide an *economic* (not literal) interpretation of this comparative dynamics result.

(e) Prove that

$$\left. \frac{\partial u^*}{\partial r}(t; \boldsymbol{\beta}) \right|_{t=0} < 0 \text{ and } \left. \frac{\partial u^*}{\partial r}(t; \boldsymbol{\beta}) \right|_{t=T} > 0.$$

Provide an economic interpretation.

(f) Prove that

$$\frac{\partial u^*}{\partial w_T}(t; \boldsymbol{\beta}) > 0 \, \forall \, t \in [0, T].$$

Provide an economic interpretation.

(g) Derive an explicit expression for the optimal value function $C^*(\cdot)$. Now consider another stage to the optimization problem. In this stage, you are to choose the terminal weight of the fish stock w_T so as to maximize the present discounted value of profit that results from selling the fish at the market-determined price of $p > 0$ per unit of weight.

(h) Set up this new optimization problem for finding the optimal value of w_T.

(i) Solve the first-order necessary condition of this optimization problem for $w_T = W(\boldsymbol{\gamma})$, where $\boldsymbol{\gamma} \stackrel{\text{def}}{=} (\alpha, p, r, w_0, T)$. Show that $W(\boldsymbol{\gamma}) > w_0$.

(j) Prove that the solution $w_T = W(\boldsymbol{\gamma})$ to the first-order necessary condition is the unique solution to the optimization problem.

(k) Prove that $\partial W(\boldsymbol{\gamma})/\partial p > 0$ and provide an economic interpretation.

(l) Prove that $\partial W(\boldsymbol{\gamma})/\partial \alpha > 0$ and provide an economic interpretation.

(m) Define the optimal feeding rate evaluated at the optimal terminal weight by $u^{**}(t; \boldsymbol{\gamma}) \stackrel{\text{def}}{=} u^*(t; w_0, T, W(\boldsymbol{\gamma}), \alpha, r)$. Derive an explicit formula for

$u^{**}(t;\gamma)$. Prove that

$$\frac{\partial u^{**}}{\partial p}(t;\gamma) > 0 \,\forall\, t \in [0,T]$$

by differentiating the explicit formula for $u^{**}(t;\gamma)$ as well as by using the above definition. Provide an economic interpretation.

(n) Prove that

$$\frac{\partial u^{**}}{\partial \alpha}(t;\gamma) < 0 \,\forall\, t \in [0,T].$$

Provide an economic interpretation. Moreover, contrast this result with the one obtained in part (d). You should reconsider the definition in part (m) for a sound answer.

FURTHER READING

Readers looking for a discussion of necessary and sufficient conditions for optimal control problems with pure state constraints, that is, constraints that depend only on the state variable and possibly t, are encouraged to consult Kamien and Schwartz (1991), Léonard and Van Long (1992), Seierstad and Sydsæter (1977), and, of course, the textbook by Seierstad and Sydsæter (1987). Optimal control problems with pure state constraints introduce several nontrivial complications, and as such, are best studied *after* one has internalized the necessary and sufficient conditions used in this book. The survey article by Hartl, Sethi, and Vickson (1995) is a comprehensive piece that presents existence, necessary, and sufficient conditions for optimal control problems with state constraints. A straightforward existence theorem is given in Steinberg and Stalford (1973). Tomiyama (1985) and Tomiyama and Rossana (1989) derive necessary conditions for two-stage optimal control problems.

REFERENCES

Hartl, R.F., Sethi, S.P., and Vickson, R.G. (1995), "A Survey of the Maximum Principles for Optimal Control Problems with State Constraints," *SIAM Review*, 37, 181–218.

Kamien, M.I. and Schwartz, N.L. (1991, 2nd Ed.), *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management* (New York: Elsevier Science Publishing Co., Inc.).

Léonard, D. and Van Long, N. (1992), *Optimal Control Theory and Static Optimization in Economics* (New York: Cambridge University Press).

Seierstad, A. and Sydsæter, K. (1977), "Sufficient Conditions in Optimal Control Theory," *International Economic Review*, 18, 367–391.

Seierstad, A. and Sydsæter, K. (1987), *Optimal Control Theory with Economic Applications* (New York: Elsevier Science Publishing Co., Inc.).

Simon, C.P. and Blume, L. (1994), *Mathematics for Economists* (New York: W.W. Norton & Company, Inc.).

Steinberg, A.M. and Stalford, H.L. (1973), "On Existence of Optimal Controls," *Journal of Optimization Theory and Applications*, 11, 266–273.

Tomiyama, K. (1985), "Two-Stage Optimal Control Problems and Optimality Conditions," *Journal of Economic Dynamics and Control*, 9, 317–337.

Tomiyama, K. and Rossana, R.J. (1989), "Two-Stage Optimal Control Problems with an Explicit Switch Point Dependence: Optimality Criteria and an Example of Delivery Lags and Investment," *Journal of Economic Dynamics and Control*, 13, 319–337.

https://doi.org/10.1017/CBO9780511806827.008 Published online by Cambridge University Press