

**Dynamic Optimization in Continuous-Time Economic Models  
(A Guide for the Perplexed)**

Maurice Obstfeld\*

*University of California at Berkeley*

First Draft: April 1992

\*I thank the National Science Foundation for research support.

## I. Introduction

The assumption that economic activity takes place continuously is a convenient abstraction in many applications. In others, such as the study of financial-market equilibrium, the assumption of continuous trading corresponds closely to reality. Regardless of motivation, continuous-time modeling allows application of a powerful mathematical tool, the theory of optimal dynamic control.

The basic idea of optimal control theory is easy to grasp-- indeed it follows from elementary principles similar to those that underlie standard static optimization problems. The purpose of these notes is twofold. First, I present intuitive derivations of the first-order necessary conditions that characterize the solutions of basic continuous-time optimization problems. Second, I show why very similar conditions apply in deterministic and stochastic environments alike. <sup>1</sup>

A simple unified treatment of continuous-time deterministic and stochastic optimization requires some restrictions on the form that economic uncertainty takes. The stochastic models I discuss below will assume that uncertainty evolves continuously according to a type of process known as an Itô <sup>^</sup> (or Gaussian

---

<sup>1</sup>When the optimization is done over a finite time horizon, the usual second-order sufficient conditions generalize immediately. (These second-order conditions will be valid in all problems examined here.) When the horizon is infinite, however, some additional "terminal" conditions are needed to ensure optimality. I make only passing reference to these conditions below, even though I always assume (for simplicity) that horizons are infinite. Detailed treatment of such technical questions can be found in some of the later references.

diffusion) process. Once mainly the province of finance theorists, Itô processes have recently been applied to interesting and otherwise intractable problems in other areas of economics, for example, exchange-rate dynamics, the theory of the firm, and endogenous growth theory. Below, I therefore include a brief and heuristic introduction to continuous-time stochastic processes, including the one fundamental tool needed for this type of analysis, Itô's chain rule for stochastic differentials. Don't be intimidated: the intuition behind Itô's Lemma is not hard to grasp, and the mileage one gets out of it thereafter truly is amazing.

## II. Deterministic Optimization in Continuous Time

The basic problem to be examined takes the form: Maximize

$$(1) \quad \int_0^{\infty} e^{-\delta t} U[c(t), k(t)] dt$$

subject to

$$(2) \quad \dot{k}(t) = G[c(t), k(t)], \quad k(0) \text{ given,}$$

where  $U(c, k)$  is a strictly concave function and  $G(c, k)$  is concave. In practice there may be some additional inequality constraints on  $c$  and/or  $k$ ; for example, if  $c$  stands for consumption,  $c$  must be nonnegative. While I will not deal in any detail with such constraints, they are straightforward to

incorporate.<sup>2</sup> In general,  $c$  and  $k$  can be vectors, but I will concentrate on the notationally simpler scalar case. I call  $c$  the *control* variable for the optimization problem and  $k$  the *state* variable. You should think of the control variable as a flow--for example, consumption per unit time--and the state variable as a stock--for example, the stock of capital, measured in units of consumption.

The problem set out above has a special structure that we can exploit in describing a solution. In the above problem, planning starts at time  $t = 0$ . Since no exogenous variables enter (1) or (2), the maximized value of (1) depends *only* on  $k(0)$ , the predetermined initial value of the state variable. In other words, the problem is *stationary*, i.e., it does not change in form with the passage of time.<sup>3</sup> Let's denote this maximized value by  $J[k(0)]$ , and call  $J(k)$  the *value function* for the problem. If  $\{c^*(t)\}_{t=0}^{\infty}$  stands for the associated optimal path of the control and  $\{k^*(t)\}_{t=0}^{\infty}$  for that of the state,<sup>4</sup> then by definition,

---

<sup>2</sup>The best reference work on economic applications of optimal control is still Kenneth J. Arrow and Mordecai Kurz, *Public Investment, the Rate of Return, and Optimal Fiscal Policy* (Baltimore: Johns Hopkins University Press, 1970).

<sup>3</sup>Nonstationary problems often can be handled by methods analogous to those discussed below, but they require additional notation to keep track of the exogenous factors that are changing.

<sup>4</sup>According to (2), these are related by

$$k^*(t) = \int_0^t G[c^*(s), k^*(s)] ds + k(0).$$

$$J[k(0)] = \int_0^{\infty} e^{-\delta t} U[c^*(t), k^*(t)] dt.$$

The nice structure of this problem relates to the following property. Suppose that the optimal plan has been followed until a time  $T > 0$ , so that  $k(T)$  is equal to the value  $k^*(T)$  given in the last footnote. Imagine a *new* decision maker who maximizes the discounted flow of utility from time  $t = T$  onward,

$$(3) \quad \int_T^{\infty} e^{-\delta(t-T)} U[c(t), k(t)] dt,$$

subject to (2), but with the initial value of  $k$  given by  $k(T) = k^*(T)$ . *Then the optimal program determined by this new decision maker will coincide with the continuation, from time  $T$  onward, of the optimal program determined at time 0, given  $k(0)$ .* You should construct a proof of this fundamental result, which is intimately related to the notion of "dynamic consistency."

You should also convince yourself of a key implication of this result, that

$$(4) \quad J[k(0)] = \int_0^T e^{-\delta t} U[c^*(t), k^*(t)] dt + e^{-\delta T} J[k^*(T)],$$

where  $J[k^*(T)]$  denotes the maximized value of (3) given that  $k(T) = k^*(T)$  and (2) is respected. Equation (4) implies that we can think of our original,  $t = 0$ , problem as the finite-horizon problem of maximizing

$$\int_0^T e^{-\delta t} U[c(t), k(t)] dt + e^{-\delta T} J[k(T)]$$

subject to the constraint that (2) holds for  $0 \leq t \leq T$ . Of course, in practice it may not be so easy to determine the correct functional form for  $J(k)$ , as we shall see below!

Nonetheless, this way of formulating our problem--which is known as *Bellman's principle of dynamic programming*--leads directly to a characterization of the optimum. Because this characterization is derived most conveniently by starting in discrete time, I first set up a discrete-time analogue of our basic maximization problem and then proceed to the limit of continuous time.

Let's imagine that time is carved up into discrete intervals of length  $h$ . A decision on the control variable  $c$ , which is a flow, sets  $c$  at some fixed level per unit time over an entire period of duration  $h$ . Furthermore, we assume that changes in  $k$ , rather than accruing continuously with time, are "credited" only at the very end of a period, like monthly interest on a bank account. We thus consider the problem: Maximize

$$(5) \quad \sum_{t=0}^{\infty} e^{-\delta t h} U[c(t), k(t)] h$$

subject to

$$(6) \quad k(t+h) - k(t) = hG[c(t),k(t)], \quad k(0) \text{ given.}$$

Above,  $c(t)$  is the fixed rate of consumption over period  $t$  while  $k(t)$  is the given level of  $k$  that prevails from the very end of period  $t - 1$  until the very end of  $t$ . In (5) [resp. (6)] I have multiplied  $U(c,k)$  [resp.  $G(c,k)$ ] by  $h$  because the cumulative flow of utility [resp. change in  $k$ ] over a period is the product of a fixed instantaneous rate of flow [resp. rate of change] and the period's length.

Bellman's principle provides a simple approach to the preceding problem. It states that the problem's value function is given by

$$(7) \quad J[k(t)] = \max_{c(t)} \left\{ U[c(t),k(t)]h + e^{-\delta h} J[k(t+h)] \right\},$$

subject to (6), for any initial  $k(t)$ . It implies, in particular, that optimal  $c^*(t)$  must be chosen to maximize the term in braces. By taking functional relationship (7) to the limit as  $h \rightarrow 0$ , we will find a way to characterize the continuous-time optimum.<sup>5</sup>

We will make four changes in (7) to get it into useful form. First, subtract  $J[k(t)]$  from both sides. Second, impose the

---

<sup>5</sup>All of this presupposes that a well-defined value function exists--something which in general requires justification. (See the extended example in this section for a concrete case.) I have also not proven that the value function, when it does exist, is *differentiable*. We know that it will be for the type of problem under study here, so I'll feel free to use the value function's first derivative whenever I need it below. With somewhat less justification, I'll also use its second and third derivatives.

constraint (6) by substituting for  $k(t+h)$ ,  $k(t) + hG(c(t),k(t))$ . Third, replace  $e^{-\delta h}$  by its power-series representation,  $1 - \delta h + (\delta^2 h^2)/2 - (\delta^3 h^3)/6 + \dots$ . Finally, divide the whole thing by  $h$ . The result is

$$(8) \quad 0 = \max_c \left[ U(c,k) - [\delta - (\delta^2 h/2) + \dots] J[k + hG(c,k)] + \{J[k + hG(c,k)] - J(k)\}/h \right],$$

where implicitly all variables are dated  $t$ . Notice that

$$\frac{J[k + hG(c,k)] - J(k)}{h} = \frac{\{J[k + hG(c,k)] - J(k)\}G(c,k)}{G(c,k)h}.$$

It follows that as  $h \rightarrow 0$ , the left-hand side above approaches  $J'(k)G(c,k)$ . Accordingly, we have proved the following

**PROPOSITION II.1.** *At each moment, the control  $c^*$  optimal for maximizing (1) subject to (2) satisfies the Bellman equation*

$$(9) \quad 0 = U(c^*,k) + J'(k)G(c^*,k) - \delta J(k) \\ = \max_c \left\{ U(c,k) + J'(k)G(c,k) - \delta J(k) \right\}.$$

This is a very simple and elegant formula. What is its interpretation? As an intermediate step in interpreting (9), define the *Hamiltonian* for this maximization problem as



$$(10) \quad \mathcal{H}(c,k) \equiv U(c,k) + J'(k)G(c,k).$$

In this model, the intertemporal tradeoff involves a choice between higher current  $c$  and higher future  $k$ . If  $c$  is consumption and  $k$  wealth, for example, the model is one in which the utility from consuming now must continuously be traded off against the utility value of savings. The Hamiltonian  $\mathcal{H}(c,k)$  can be thought of as a measure of the flow value, in current utility terms, of the consumption-savings combination implied by the consumption choice  $c$ , given the predetermined value of  $k$ . The Hamiltonian solves the problem of "pricing" saving in terms of current utility by multiplying the flow of saving,  $G(c,k) = \dot{k}$ , by  $J'(k)$ , the effect of an increment to wealth on total lifetime utility. A corollary of this observation is that  $J'(k)$  has a natural interpretation as the shadow price (or marginal current utility) of wealth. More generally, leaving our particular example aside,  $J'(k)$  is the shadow price one should associate with the state variable  $k$ .

This brings us back to the Bellman equation, equation (9). Let  $c^*$  be the value of  $c$  that maximizes  $\mathcal{H}(c,k)$ , given  $k$ , which is arbitrarily predetermined and therefore might not have been chosen optimally.<sup>6</sup> Then (9) states that

$$(11) \quad \mathcal{H}(c^*,k) = \max_c \{\mathcal{H}(c,k)\} = \delta J(k).$$

---

<sup>6</sup>It is important to understand clearly that at a given point in time  $t$ ,  $k(t)$  is not an object of choice (which is why we call it a state variable). Variable  $c(t)$  can be chosen freely at time  $t$  (which is why it is called a control variable), but its level influences the *change* in  $k(t)$  over the next infinitesimal time interval,  $k(t + dt) - k(t)$ , not the current value  $k(t)$ .

In words, the maximized Hamiltonian is a fraction  $\delta$  of an optimal plan's total lifetime value. Equivalently, the instantaneous value flow from following an optimal plan divided by the plan's total value--i.e., the plan's rate of return--must equal the rate of time preference,  $\delta$ . Notice that if we were to measure the current payout of the plan by  $U(c^*,k)$  alone, we would err by not taking proper account of the value of the current increase in  $k$ . This would be like leaving investment out of our measure of GNP! The Hamiltonian solves this accounting problem by valuing the increment to  $k$  using the shadow price  $J'(k)$ .

To understand the implications of (9) for optimal consumption we must go ahead and perform the maximization that it specifies (which amounts to maximizing the Hamiltonian). As a by-product, we obtain the Pontryagin necessary conditions for optimal control.

Maximizing the term in braces in (9) over  $c$ , we get <sup>7</sup>

$$(12) \quad U_c(c^*,k) = -G_c(c^*,k)J'(k).$$

The reason this condition is necessary is easy to grasp. At each moment the decision maker can decide to "consume" a bit more, but at some cost in terms of the value of current "savings." A unit of additional consumption yields a marginal payoff of  $U_c(c^*,k)$ , but at the same time, savings change by  $G_c(c^*,k)$ . The utility

---

<sup>7</sup>I assume interior solutions throughout.

value of a marginal *fall* in savings thus is  $-G_c(c^*,k)J'(k)$ ; and if the planner is indeed at an optimum, it must be that this marginal cost just equals the marginal current utility benefit from higher consumption. In other words, unless (12) holds, there will be an incentive to change  $c$  from  $c^*$ , meaning that  $c^*$  cannot be optimal.

Let's get some further insight by exploiting again the recursive structure of the problem. It is easy to see from (12) that for any date  $t$  the optimal level of the control,  $c^*(t)$ , depends only on the inherited state  $k(t)$  (regardless of whether  $k(t)$  was chosen optimally in the past). Let's write this functional relationship between optimal  $c$  and  $k$  as  $c^* = c(k)$ , and assume that  $c(k)$  is differentiable. (For example, if  $c$  is consumption and  $k$  total wealth,  $c(k)$  will be the household's consumption function.) Functions like  $c(k)$  will be called *optimal policy functions*, or more simply, *policy functions*. Because  $c(k)$  is defined as the solution to (9), it automatically satisfies

$$0 = U[c(k),k] + J'(k)G[c(k),k] - \delta J(k).$$

Equation (12) informs us as to the optimal relation between  $c$  and  $k$  at a point in time. To learn about the implied optimal behavior of consumption over time, let's differentiate the preceding equation with respect to  $k$ :

$$0 = [U_c(c^*, k) + J'(k)G_c(c^*, k)]c'(k) + U_k(c^*, k) \\ + [G_k(c^*, k) - \delta]J'(k) + J''(k)G(c^*, k).$$

The expression above, far from being a hopeless quagmire, is actually just what we're looking for. Notice first that the left-hand term multiplying  $c'(k)$  drops out entirely thanks to (12): another example of the envelope theorem. This leaves us with the rest,

$$(13) \quad U_k(c^*, k) + J'(k)[G_k(c^*, k) - \delta] + J''(k)G(c^*, k) = 0.$$

Even the preceding simplified expression probably isn't totally reassuring. Do not despair, however. A familiar economic interpretation is again fortunately available.

We saw earlier that  $J'(k)$  could be usefully thought of as the shadow price of the state variable  $k$ . If we think of  $k$  as an asset stock (capital, foreign bonds, whatever), this shadow price corresponds to an asset price. Furthermore, we know that under perfect foresight, asset prices adjust so as to equate the asset's total instantaneous rate of return to some required or benchmark rate of return, which in the present context can only be the time-preference rate,  $\delta$ . As an aid to clear thinking, let's introduce a new variable,  $\lambda$ , to represent the shadow price  $J'(k)$  of the asset  $k$ :

$$\lambda \equiv J'(k).$$

Our next step will be to rewrite (13) in terms of  $\lambda$ . The key observation allowing us to do this concerns the last term on the right-hand side of (13),  $J''(k)G(c^*,k)$ . The chain rule of calculus implies that

$$J''(k)G(c^*,k) = \frac{dJ'(k)}{dk} \times \frac{dk}{dt} = \frac{d\lambda}{dk} \times \frac{dk}{dt} = \frac{d\lambda}{dt} = \dot{\lambda};$$

and with this fact in hand, it is only a matter of substitution to express (13) in the form

$$(14) \quad \frac{U_k + \lambda G_k + \dot{\lambda}}{\lambda} = \delta.$$

This is just the asset-pricing equation promised in the last paragraph.

Can you see why this last assertion is true? To understand it, let's decompose the total return to holding a unit of stock  $k$  into "dividends" and "capital gains." The "dividend" is the sum of two parts, the direct effect of an extra unit of  $k$  on utility,  $U_k$ , and its effect on the rate of increase of  $k$ ,  $\lambda G_k$ . (We must multiply  $G_k$  by the shadow price  $\lambda$  in order to express the *physical* effect of  $k$  on  $\dot{k}$  in the same terms as  $U_k$ , that is, in terms of utility.) The "capital gain" is just the increase in the price of  $k$ ,  $\dot{\lambda}$ . The sum of dividend and capital gain, divided by the asset price  $\lambda$ , is just the rate of return on  $k$ , which, by

(14) must equal  $\delta$  along an optimal path.

---

### Example

Let's step back for a moment from this abstract setting to consolidate what we've learned through an example. Consider the standard problem of a consumer who maximizes  $\int_0^{\infty} e^{-\delta t} U[c(t)] dt$  subject to  $\dot{k} = f(k) - c$  (where  $c$  is consumption,  $k$  capital, and  $f(k)$  the production function). Now  $U_k = 0$ ,  $G(c, k) = f(k) - c$ ,  $G_c = -1$ , and  $G_k = f'(k)$ . In this setting, (14) becomes the statement that the rate of time preference should equal the marginal product of capital plus the rate of accrual of utility capital gains,

$$\delta = f'(k) + \frac{\dot{\lambda}}{\lambda}.$$

Condition (12) becomes  $U'(c) = \lambda$ . Since this last equality implies that  $\dot{\lambda} = U''(c)\dot{c}$ , we can express the optimal dynamics of  $c$  and  $k$  as a nonlinear differential-equation system:

$$(15) \quad \dot{c} = -\frac{U'(c)}{U''(c)} [f'(k) - \delta], \quad \dot{k} = f(k) - c.$$

You can see the phase diagram for this system in figure 1. (Be sure you can derive it yourself! The diagram assumes that  $\lim_{k \rightarrow \infty} f'(k) = 0$ , so that a steady-state capital stock exists.) The diagram makes clear that, given  $k$ , any positive initial  $c$

initiates a path along which the two preceding differential equations for  $c$  and  $k$  are respected. *But not all of these paths are optimal, since the differential equations specify conditions that are merely necessary, but not sufficient, for optimality.*

Indeed, only one path will be optimal in general: we can write the associated policy function as  $c^* = c(k)$  (it is graphed in figure 1). For given  $k$ , paths with initial consumption levels exceeding  $c(k)$  imply that  $k$  becomes negative after a finite time interval. Since a negative capital stock is nonsensical, such paths are not even feasible, let alone optimal. Paths with initial consumption levels below  $c(k)$  imply that  $k$  gets to be too large, in the sense that the individual could raise lifetime utility by eating some capital and never replacing it. These "overaccumulation" paths violate a sort of terminal condition stating that the present value of the capital stock should converge to zero along an optimal path. I shall not take the time to discuss such terminal conditions here.

If we take

$$U(c) = \frac{c^{1-(1/\varepsilon)} - 1}{1 - (1/\varepsilon)}, \quad f(k) = rk,$$

where  $\varepsilon$  and  $r$  are positive constants. we can actually find an algebraic formula for the policy function  $c(k)$ .

Let's conjecture that optimal consumption is proportional to wealth, that is, that  $c(k) = \eta k$  for some constant  $\eta$  to be

determined. If this conjecture is right, the capital stock  $k$  will follow  $\dot{k} = (r - \eta)k$ , or, equivalently,

$$\frac{\dot{k}}{k} = r - \eta.$$

This expression gives us the key clue for finding  $\eta$ . If  $c = \eta k$ , as we've guessed, then also

$$\frac{\dot{c}}{c} = r - \eta.$$

But necessary condition (15) requires that  $\frac{\dot{c}}{c} = \varepsilon(r - \delta)$ , which contradicts the last equation unless

$$(16) \quad \eta = (1 - \varepsilon)r + \varepsilon\delta.$$

Thus,  $c(k) = [(1 - \varepsilon)r + \varepsilon\delta]k$  is the optimal policy function. In the case of log utility ( $\varepsilon = 1$ ), we simply have  $\eta = \delta$ . We get the same simple result if it so happens that  $r$  and  $\delta$  are equal.

Equation (16) has a nice interpretation. In Milton Friedman's permanent-income model, where  $\delta = r$ , people consume the annuity value of wealth, so that  $\eta = r = \delta$ . This rule results in a *level* consumption path. When  $\delta \neq r$ , however, the optimal consumption path will be tilted, with consumption rising over time if  $r > \delta$  and falling over time if  $r < \delta$ . By writing



(16) as

$$\eta = r - \varepsilon(r - \delta)$$

we can see these two effects at work. Why is the deviation from the Friedman permanent-income path proportional to  $\varepsilon$ ? Recall that  $\varepsilon$ , the elasticity of intertemporal substitution, measures an individual's willingness to substitute consumption today for consumption in the future. If  $\varepsilon$  is high and  $r > \delta$ , for example, people will be quite willing to forgo present consumption to take advantage of the relatively high rate of return to saving; and the larger is  $\varepsilon$ , *certeris paribus*, the lower will be  $\eta$ . Alert readers will have noticed a major problem with all this. If  $r > \delta$  and  $\varepsilon$  is sufficiently large,  $\eta$ , and hence "optimal" consumption, will be *negative*. How can this be? Where has our analysis gone wrong?

The answer is that when  $\eta < 0$ , *no optimum consumption plan exists!* After all, nothing we've done demonstrates existence: our arguments merely indicate some properties that an optimum, *if* one exists, will need to have.

No optimal consumption path exists when  $\eta < 0$  for the following reason. Because optimal consumption growth necessarily satisfies  $\dot{c}/c = \varepsilon(r - \delta)$ , and  $\varepsilon(r - \delta) > r$  in this case, optimal consumption would have to grow at a rate exceeding the rate of return on capital,  $r$ . Since capital growth obeys  $\dot{k}/k = r - (c/k)$ , however, and  $c \geq 0$ , the growth rate of capital, and hence

that of output, is at most  $r$ . With consumption positive and growing at 3 percent per year, say, but with capital growing at a lower yearly rate, consumption would eventually grow to be greater than total output--an impossibility in a closed economy. So the proposed path for consumption is not feasible. This means that no feasible path--other than the obviously suboptimal path with  $c(t) = 0$ , for all  $t$ --satisfies the necessary conditions for optimality. Hence, no feasible path is optimal: no optimal path exists.

Let's take our analysis a step further to see how the value function  $J(k)$  looks. Observe first that at any time  $t$ ,  $k(t) = k(0)e^{(r-\eta)t} = k(0)e^{\varepsilon(r-\delta)t}$ , where  $k(0)$  is the starting capital stock and  $\eta$  is given by (16). Evidently, the value function at  $t = 0$  is just

$$\begin{aligned}
 J[k(0)] &= \left[1 - \frac{1}{\varepsilon}\right]^{-1} \left\{ \int_0^{\infty} e^{-\delta t} [\eta k(t)]^{1-(1/\varepsilon)} dt - \frac{1}{\delta} \right\} \\
 &= \left[1 - \frac{1}{\varepsilon}\right]^{-1} \left\{ \int_0^{\infty} e^{-\delta t} [\eta k(0) e^{\varepsilon(r-\delta)t}]^{1-(1/\varepsilon)} dt - \frac{1}{\delta} \right\} \\
 &= \left[1 - \frac{1}{\varepsilon}\right]^{-1} \left\{ \frac{[\eta k(0)]^{1-(1/\varepsilon)}}{\delta - (\varepsilon - 1)(r - \delta)} - \frac{1}{\delta} \right\} .
 \end{aligned}$$

So the value function  $J(k)$  has the same general form as the utility function, but with  $k$  in place of  $c$ . This is not the last

time we'll encounter this property. Alert readers will again notice that to carry out the final step of the last calculation, I had to assume that the integral in braces above is convergent, that is, that  $\delta - (\varepsilon - 1)(r - \delta) > 0$ . Notice, however, that  $\delta - (\varepsilon - 1)(r - \delta) = r - \varepsilon(r - \delta) = \eta$ , so the calculation is valid provided an optimal consumption program exists. If one doesn't, the value function clearly doesn't exist either: we can't specify the maximized value of a function that doesn't attain a maximum. This counterexample should serve as a warning against blithely assuming that all problems have well-defined solutions and value functions.

---

Return now to the theoretical development. We have seen how to solve continuous-time deterministic problems using Bellman's method of dynamic programming, which is based on the value function  $J(k)$ . We have also seen how to interpret the derivative of the value function,  $J'(k)$ , as a sort of shadow asset price, denoted by  $\lambda$ . The last order of business is to show that we have actually proved a simple form of Pontryagin's *Maximum Principle*:<sup>8</sup>

**PROPOSITION II.2.** (Maximum Principle) *Let  $c^*(t)$  solve the problem of maximizing (1) subject to (2). Then there exist variables  $\lambda(t)$ --called costate variables--such that the Hamiltonian*

---

<sup>8</sup>First derived in L.S. Pontryagin et al., *The Mathematical Theory of Optimal Processes* (New York and London: Interscience Publishers, 1962).

$$\mathcal{H}[c, k(t), \lambda(t)] \equiv U[c, k(t)] + \lambda(t)G[c, k(t)]$$

is maximized at  $c = c^*(t)$  given  $\lambda(t)$  and  $k(t)$ ; that is,

$$(17) \quad \frac{\partial \mathcal{H}}{\partial c}(c^*, k, \lambda) = U_c(c^*, k) + \lambda G_c(c^*, k) = 0$$

at all times (assuming, as always, an interior solution). Furthermore, the costate variable obeys the differential equation

$$(18) \quad \dot{\lambda} = \lambda \delta - \frac{\partial \mathcal{H}}{\partial k}(c^*, k, \lambda) = \lambda \delta - [U_k(c^*, k) + \lambda G_k(c^*, k)]$$

for  $\dot{k} = G(c^*, k)$  and  $k(0)$  given. <sup>9</sup>

---

<sup>9</sup>You should note that if we integrate differential-equation (18), we get the *general* solution

$$\lambda(t) = \int_t^{\infty} e^{-\delta(s-t)} \frac{\partial \mathcal{H}}{\partial k}[c^*(s), k(s), \lambda(s)] ds + Ae^{\delta t},$$

where  $A$  is an arbitrary constant. [To check this claim, just differentiate the foregoing expression with respect to  $t$ : if the integral in the expression is  $I(t)$ , we find that  $\dot{\lambda} = \delta I - (\partial \mathcal{H} / \partial k) + \delta Ae^{\delta t} = \delta \lambda - (\partial \mathcal{H} / \partial k)$ .] I referred in the prior example to an additional terminal condition requiring the present value of the capital stock to converge to zero along an optimal path. Since  $\lambda(t)$  is the price of capital at time  $t$ , this terminal condition usually requires that  $\lim_{t \rightarrow \infty} e^{-\delta t} \lambda(t) = 0$ , or that  $A = 0$  in the

You can verify that if we identify the costate variable with the derivative of the value function,  $J'(k)$ , the Pontryagin necessary conditions are satisfied by our earlier dynamic-programming solution. In particular, (17) coincides with (12) and (18) coincides with (14). So we have shown, in a simple stationary setting, why the Maximum Principle "works." The principle is actually more broadly applicable than you might guess from the foregoing discussion--it easily handles nonstationary environments, side constraints, etc. And it has a stochastic analogue, to which I now turn. 10

---

solution above. The *particular* solution that remains equates the shadow price of a unit of capital to the discounted stream of its shadow "marginal products," where the latter are measured by partial derivatives of the flow of value,  $\mathcal{H}$ , with respect to  $k$ .  
10For more details and complications on the deterministic Maximum Principle, see Arrow and Kurz, *op. cit.*