Appendix 14.1 The optimal control problem and its solution using the maximum principle

NOTE: Many occurrences of f, x, u, λ and μ in this file (in equations or as whole words in text) are purposefully in **bold** in order to refer to vectors.

Optimal control theory, using the Maximum Principle, is a technique for solving constrained dynamic optimisation problems. In this appendix we aim to

- explain what is meant by a constrained dynamic optimisation problem;
- show one technique - the maximum principle - by which such a problem can be solved.

We will not give any proofs of the conditions used in the Maximum Principle. Our emphasis is on explaining the technique and showing how to use it. For the reader who wishes to go through these proofs, some recommendations for further reading are given at the end of Chapter 14. After you have finished reading this appendix, it will be useful to go through Appendices 14.2 and 14.3. Appendix 14.2 shows how the maximum principle is used to derive the optimal solution to the simple non-renewable resource depletion problem discussed in Part 1 of this chapter. Appendix 14.3 considers the optimal allocation of a renewable or non-renewable resource in the case where extraction of the resource involves costs, the model discussed in Part 2 of this chapter.

Let us begin by laying out the various elements of a constrained dynamic optimisation problem. In doing this, you will find it useful to refer to Tables 14.2 and 14.3, where we have summarised the key elements of the optimal control problem and its solution.

Table 14.2 The optimal control problem and its solution

Objective function (See Notes 1 & 5)	$\max J(\mathbf{u}) = \int_{t_0}^{t_T} L(\mathbf{x},\mathbf{u},t)dt \quad + F(\mathbf{x}(t_T),t_T)$			
System	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},\mathrm{t}) \qquad \mathbf{x}(\mathrm{t}_0) = \mathbf{x}_0$			
Terminal state	$\mathbf{x}(\mathrm{t}_T) = \mathbf{x}_T$		$\mathbf{x}(\mathrm{t}_T)$ free	
Terminal point	t_T fixed	t_T free	t_T fixed	t_T free
Hamiltonian (See Note 2)	$H_C = H(\mathbf{x},\mathbf{u},t,\lambda)$ $= \mathrm{L}(\mathbf{x},\mathbf{u},\mathrm{t}) + \boldsymbol{\lambda}\mathbf{f}(\mathbf{x},\mathbf{u},\mathrm{t})$			
Equations of motion	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},t) \qquad \dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \mathbf{x}}$			
Max H (See Note 3)	$\frac{\partial \mathrm{H}}{\partial \mathbf{u}} = 0$			
Transversality condition	$\mathbf{x}(\mathrm{t}_T) = \mathbf{x}_T$		$\boldsymbol{\lambda}(\mathrm{t}_T) = 0$ if $t_T = \infty$ or $F(\bullet) = 0$ otherwise $\boldsymbol{\lambda}(\mathrm{t}_T) = \mathrm{F}'(\cdot)$	
(See Note 4)		$H_C(t_T) + \partial F / \partial t_T = 0$		$\mathrm{H}_C(\mathrm{t}_T) + \partial \mathrm{F} / \partial \mathrm{t}_T = 0$

Notes to Table 14.2

Note 1	The term $F(\mathbf{x}(t_T), t_T)$ may not be present in the objective function, and cannot be if $t_T = \infty$.
Note 2	H_C denotes the "current value" Hamiltonian (see the notes to Table 14.3)
Note 3	The Max H condition given in the table is for the special case of an interior solution (**u** is an interior point). A more general statement of this condition (the "maximum principle") is: $\mathbf{u}(t)$ maximises H over $\mathbf{u}(t) \in \mathbf{U}$, for $t_0 \le t \le t_T$.
Note 4	The term $\partial F/\partial t_T$ does not enter the transversality condition in the final line of the table if $F(.) = 0$.
Note 5	Notation used in the table (see below):
$\mathbf{x}(t)$	the vector of state variables
$\mathbf{u}(t)$	the vector of control variables
$J(\mathbf{u})$	the objective function to be maximised, which may be augmented by a final function $F(\cdot)$
$\boldsymbol{\lambda}(t)$	the vector of co-state variables (or shadow prices)
$\mathbf{f}(\mathbf{x},\mathbf{u},t)$	the state equation functions, describing the relevant physical-economic system
t_0	the initial point in time
t_T	the terminal point in time
H	the Hamiltonian function
$F(\mathbf{x}(t_T), t_T)$	a 'final function' (the role of which is explained below)

Table 14.3 The optimal control problem with a discounting factor and its solution

Objective Function (See Note 1)	$\max J(\mathbf{u}) = \int_{t_0}^{t_T} L(\mathbf{x},\mathbf{u},t)e^{-\rho t}\, dt \quad + F(\mathbf{x}(t_T),t_T)e^{-\rho t_T}$			
System	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},t) \qquad \mathbf{x}(t_0) = \mathbf{x}_0$			
Terminal state	$\mathbf{x}(t_T) = \mathbf{x}_T$		$\mathbf{x}(t_T)$ free	
Terminal point	t_T fixed	t_T free	t_T fixed	t_T free
Present-value Hamiltonian (See Note 2)	$H_P = H(\mathbf{x},\mathbf{u},t,\boldsymbol{\lambda}) = L(\mathbf{x},\mathbf{u},t)e^{-\rho t} + \boldsymbol{\lambda}\mathbf{f}(\mathbf{x},\mathbf{u},t)$			
Current-value Hamiltonian (See Note 2)	$H_C = H_P e^{\rho t} = L(\mathbf{x},\mathbf{u},t) + \boldsymbol{\mu}\mathbf{f}(\mathbf{x},\mathbf{u},t) \quad , \quad (\boldsymbol{\mu} = \boldsymbol{\lambda}e^{\rho t})$			
Equations of motion	$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},t) \qquad \dot{\boldsymbol{\mu}} = \rho\boldsymbol{\mu} - \frac{\partial H_C}{\partial \mathbf{x}}$			
Max H_C (See Note 3)	$\frac{\partial H_C}{\partial \mathbf{u}} = 0$			
Transversality condition	$\mathbf{x}(t_T) = \mathbf{x}_T$		$\boldsymbol{\mu}(t_T) = 0$ if $t_T = \infty$ or if $F(\cdot) = 0$ otherwise $\boldsymbol{\mu}(t_T) = F'(\cdot)$	
(See Note 4)		$H_C(t_T) + \partial F / \partial t_T = 0$		$H_C(t_T) + \partial F / \partial t_T = 0$

Notes to Table 14.3

Note 1	The term $F(\mathbf{x}(t_T), t_T)e^{-\rho t_T}$ may not be present in the objective function, and cannot be if $t_T = \infty$.
Note 2	Two versions of the Hamiltonian function are given. The first (called here H_P) is known as the 'present value' Hamiltonian, as the presence of the term $e^{-\rho t}$ in the objective function (which converts magnitudes into present value terms) carries over into the Hamiltonian, H_P. The second (called here H_C) is known as the 'current-value' Hamiltonian (see Chiang [1992], page 210). Premultiplying H_P by $e^{\rho t}$ removes the discounting factor from the expression, and generates the current-value Hamiltonian, H_C . The property that H_C is expressed in current-value terms can be seen by noting that the L(**x**, **u**, t) function is not multiplied by the discounting term $e^{-\rho t}$. For mathematical convenience, it is usually better to work with the Hamiltonian in current-value terms.
Note 3	The Max H_C condition given in the table is for the special case of an interior solution (**u** is an interior set). A more general statement of this condition (the "maximum principle") is: **u**(t) maximises H_C over $\mathbf{u}(t) \in \mathbf{U}$, for $t_0 \leq t \leq t_T$.
Note 4	The term $\partial F/\partial t_T$ does not enter the transversality condition in the final line of the table if F(.) = 0.
Note 5	Notation is as defined in Table 14.2. The term ρ denotes the utility social discount rate.

1. The function to be maximised is known as the objective function, denoted J(**u**). This takes the form of an integral over a time period from an initial time t_0 to the terminal time t_T. Two points should be borne in mind about the terminal point in time, t_T:
   - In some optimisation problems t_T is fixed; in others it is free (and so becomes an endogenous variable, the value of which is solved for as part of the optimisation exercise).
   - In some optimisation problems the terminal point is a finite quantity (it is a finite number of time periods later than the initial time); in others, the terminal point is infinite ($t_T = \infty$). When a problem has an infinite terminal point in time, t_T should be regarded as free.

2. The objective function will, in general, contain as its arguments three types of variable:

- **x**(t), a *vector* of n state variables at time t;
- **u**(t), a *vector* of m control (or instrument) variables at time t;
- t, time itself.

Although the objective function may contain each of these three types of variables, it is not necessary that all be present in the objective function (as you will see from the examples worked through in Appendices 14.2 and 14.3).

3. The objective function may (but will often not) be augmented with the addition of a 'final function' (also known as a ‘salvage function’), denoted by the function F(·) in Tables 14.2 and 14.3. It will appear whenever the value of the objective function is dependent on some particular function F(·) of the levels of the state variables **x** at the terminal time, and possibly on the terminal time itself. (The applications in Chapter 14 did not involve the use of a final function.)

4. The solution to a dynamic optimal control problem will contain, among other things, the values of the state and control variables at each point in time over the period from t_0 to t_T. It is this that makes the exercise a dynamic optimisation exercise.

5. Underlying the optimal control problem will be some economic, biological or physical system (which we shall call simply 'the economic system'), describing the initial values of a set of state variables of interest, and how they evolve over time. The evolution of the state variables over time will, in general, be described by a set of differential equations (known as state equations) of the form:

   $$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$$

   where $\dot{\mathbf{x}}$ = d**x**/dt is the time derivative of **x** (the rate of change of **x** with respect to time). Note that as **x** is a vector of n state variables, there will in general be n state equations. Any solution to the optimal control problem must satisfy these state equations. This is one reason why we use the phrase ‘constrained’ dynamic optimisation problems.

6. A second way in which constraints may enter the problem is through the terminal conditions of the problem. There are two aspects here: one concerns the value of the state variables at the terminal point in time, the other concerns the terminal point in time itself.
   - First, in some problems the values that the state variables take at the terminal point in time are fixed; in others these values are free (and so are endogenously determined in the optimisation exercise).
   - Secondly, either the particular problem that we are dealing with will fix the terminal point in time, or that point will be free (and so, again, be determined endogenously in the optimisation exercise).

7. The optimisation exercise must satisfy a so-called transversality condition. The particular transversality condition that must be satisfied in any particular problem will depend upon which of the four possibilities outlined in (6) applies. (Four possibilities exist because for each of the two possibilities for the terminal values of the state variables there are two possibilities for the terminal point in time.) It follows from this that when we read Tables 14.2 and 14.3, then (ignoring the column of labels) there are four columns referring to these four possibilities. Where cells are merged and so cover more than one column, the condition shown refers to all the possibilities it covers. We shall come back to the transversality condition in a moment.

8. The control variables (or instruments) are variables whose value can be chosen by the decision maker in order to steer the evolution of the state variables over time in a desired manner.

9. In addition to the three kinds of variables we have discussed so far – time, state and control variables – a fourth type of variable enters optimal control problems. This is the vector of co-state variables **λ** (or **μ** in the case where the objective function contains a discount factor). Co-state variables are similar to the Lagrange multiplier variables one finds in static constrained optimisation exercises. But in the present context, where we are dealing with a dynamic optimisation problem over some sequence of time periods, the value taken by each co-state variable will in general vary over time, and so it is appropriate to denote **λ** (*t*) as the vector of co-state variables at time *t*.

10. The analogy of co-state variables with Lagrange multipliers carries over to their economic interpretation: the co-state variables can be interpreted as shadow prices, which denote the decision maker's marginal valuation of the state variable at each point in time (along the optimal time path of the state and control variables).

11. Finally, let us return to the transversality condition. Looking at the final rows in Tables 14.2 and 14.3 you will see four possible configurations of transversality condition. All relate to something that must be satisfied by the solution at the terminal point in time. Where the terminal value of the state variables is fixed, this will always be reflected in the transversality condition. On the other hand, where the terminal value of the state variables is free, the transversality condition will usually[9] require that the shadow price of the state variables be zero. Intuitively, this means that if we do not put any constraints on how large the stocks of the state variables must be at the terminal point in time, then they must have a zero value at that time. For if they had any positive value, it would have been optimal to deplete them further prior to the end of the planning horizon. Note also that whenever the terminal point in time is free (whether or not the state variables are fixed at the terminal point), an additional part of the transversality condition requires that the Hamiltonian have a zero value at the endogenously

---

9 Strictly speaking, the shadow price will be zero where the time horizon is of infinite length or if there is no final function in the objective function. However, where the objective function contains a final function, the shadow price must equal the first derivative of that final function with respect to the state variable. This is shown in Tables 14.2 and 14.3.

determined terminal point in time.[10] If it did not, then the terminal point could not have been an optimal one!

The general case referred to in Tables 14.2 and 14.3, and special cases

In the description we have given above of the optimal control problem, we have been considering a general case. For example, we allow there to be *n* state variables and *m* control variables. In some special cases, *m* and *n* may each be one, so there is only one state and one control variable. Also, we have written the state equation for the economic system of interest as being a function of three types of variables: time, state and control. In many particular problems, not all three types of variables will be present as arguments in the state equation. For example, in many problems, time does not enter explicitly in the state equation. A similar comment applies to the objective function: while in general it is a function of three types of variables, not all three will enter in some problems. Finally, often the objective function will not be augmented by the presence of a 'final function'.

Limitations to the optimal control technique outlined in this appendix

The statement of the optimal control problem and its solution given in this appendix is not as general as it might be. For example, the terminal condition might require that a control variable must be greater than some particular quantity (but is otherwise unconstrained). Or the problem might require that one or more state variables (such as a resource stock) be non-negative in the terminal state. Another case of obvious interest to economists is the requirement that a control variable (such as a resource depletion rate) be non-negative over the planning horizon. Details of generalisations of the optimal control model to cover these (and many other) cases can be found in Chiang (1992) or, more rigorously, Kamien and Schwartz (1991).

[10] Or, as shown in Table 14.3, the additional part of the transversality condition requires that the Hamiltonian plus the derivative of F with respect to T have a zero value at the endogenously determined terminal point in time.

The presence of a discount factor in the objective function

For some dynamic optimisation problems, the objective function to be maximised, J(**u**), will be an integral over time of some function of time, state variables and control variables. That is:

$$J(\mathbf{u}) = \int_{t_0}^{t_T} L(\mathbf{x}, \mathbf{u}, t)dt$$

However, in many dynamic optimisation problems that are of interest to economists, the objective function will be a discounted (or present-value) integral of the form:

$$J(\mathbf{u}) = \int_{t_0}^{t_T} L(\mathbf{x}, \mathbf{u}, t)e^{-\rho t}dt$$

For example, equation 14.8 in the text of this chapter is of this form. There, L is actually a utility function U(·) (which is a function of only one control variable, C). Indeed, throughout this book, the objective functions with which we deal are almost always discounted or present-value integral functions.

The solution of the optimal control problem

The nature of the solution to the optimal control problem will differ depending on whether or not the objective function contains a discounting factor. Table 14.2 states formally the optimal control problem and its solution using general notation, for the case where the objective function does not include a discount factor. Table 14.3 presents the same information for the case where the objective function is a discounted (or present-value) integral. Some (brief) explanation and discussion of how the conditions listed in Tables 14.2 and 14.3 may be used to obtain the required solution is provided below those tables. However, we strongly urge you also to read Appendices 14.2 and 14.3, so that you can get a feel for how the general results we have described here can be used in practice (and how we have used them in this chapter).

Interpreting the two tables

It will help to focus on one case: we will look at an optimal control problem with a discounting factor, an infinite time horizon (so that t_T is deemed to be free), and no restriction being placed on the values of the state variable in the terminal time period (so that $\mathbf{x}(t_T)$ is free). The relevant statement of the optimal control problem and its solution is, therefore, that given in the final column to the right in Table 14.3.

We can express the problem as

$$\max J(\mathbf{u}) = \int_{t_0}^{t_T} L(\mathbf{x}, \mathbf{u}, t)e^{-\rho t}\ dt$$

subject to

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad \text{and} \quad \mathbf{x}(t_0) = \mathbf{x}_0, \qquad \mathbf{x}(0)\ \text{given},\ \mathbf{x}(t_T)\ \text{free.}$$

To obtain the solution we first construct the current-value Hamiltonian:

$$H_C = L(\mathbf{x}, \mathbf{u}, t) + \boldsymbol{\mu}\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$$

The current-value Hamiltonian consists of two components:

- The first $L(\mathbf{x},\mathbf{u},t)$ is the function which, after being multiplied by the discounting factor and then being integrated over the relevant time horizon, enters the objective function. Note carefully by examining Table 14.3 that in the Hamiltonian the L function itself enters, not its integral. Furthermore, although the discounting factor enters the objective function, it does not enter in the current-value Hamiltonian.
- The second component that enters the Hamiltonian is the right-hand side of the state variable equations of motion, $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$, after having been premultiplied by the co-state variable vector in current-value form. Remember that in the general case there are n state variables, and so n co-state variables, one for each state equation. In order for this multiplication to be conformable, it is actually the transpose of the co-state vector $\boldsymbol{\mu}$ that premultiplies the vector of functions from the state equations.

Our next task is to find the values of the control variables **u** which maximise the current-value Hamiltonian at each point in time; it is this which gives this approach its name of 'the maximum principle'. If the Hamiltonian function H_C is non-linear and differentiable in the control variables **u**, then the problem will have an interior solution, which can be found easily. This is done by differentiating H_C with respect to **u** and setting the derivatives equal to zero. Hence in this case one of the necessary conditions for the solution will be

$$\frac{\partial H_C}{\partial \mathbf{u}} = 0$$ (a set of m equations, one for each of the m control variables).

More generally, there may be a corner solution. Obtaining this solution may be a difficult task in some circumstances, as it involves searching for the values of **u**(t) which maximises $H_C(t)$ (at all points in time) in some other way.

Bringing together all the necessary conditions for the complete solution of the optimisation problem we have:

- The maximum principle conditions (assuming an interior solution and no final function present):

  $\frac{\partial H_C}{\partial \mathbf{u}} = 0$ (a set of m equations, one for each of the m control variables).

- Those given in the row labelled 'Equations of motion' in Table 14.3, that is

  $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ (a set of n equations)

  $\dot{\boldsymbol{\mu}} = \rho\boldsymbol{\mu} - \frac{\partial H_C}{\partial \mathbf{x}}$ (a set of n equations)

- The initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$
- The transversality condition $H_C(t_T) = 0$

Solving these necessary conditions simultaneously, we can obtain the optimal time path for each of the m control variables over the (infinite) time horizon. Corresponding to this time path of the control variables are the optimised time paths of the n state variables and their associated current-value shadow prices (values of the co-state variables) along the optimal path.

It should be clear that obtaining this complete solution could be a daunting task in problems with many control and state variables. However, where the number of variables is small and the relevant functions are easy to handle, the solution can often be obtained quite simply. We demonstrate this assertion in the following two appendices.

One final point warrants mention. Tables 14.2 and 14.3 give necessary but not sufficient conditions for a maximum. In principle, to confirm that our solution is indeed a maximum, second-order conditions should be checked as well. However, in most problems of interest to economists (and in all problems investigated in this book), assumptions are made about the shapes of functions which guarantee that second-order conditions for a maximum will be satisfied, thereby obviating the need for checking second-order conditions.

Let us try to provide some intuitive content to the foregoing by considering a problem where there is just one state variable, x, and one control variable, u, where t does not enter either the objective function or the equation describing the system, no final function is present, and where $t_0 = 0$ and we have an infinite terminal point ($t_T = \infty$). Then the problem is to maximise [1]

$$\int_0^\infty L(x_t, u_t)e^{-\rho t}dt$$

subject to

$$\dot{x} = f(x_t, u_t) \text{ and } x(t_0) = x_o$$

for which the current-value Hamiltonian is

$$H_{C_t} = L(x_t, u_t) + \mu_t f(x_t, u_t) = L(x_t, u_t) + \mu_t \dot{x}_t$$

In the original problem, we are looking to maximise the integral of the discounted value of $L(x_t, u_t)$. The first term in the Hamiltonian is just $L(x_t, u_t)$, the instantaneous value of that we seek the maximum of. Recalling that co-state variables are like Lagrangian multipliers and that those are shadow prices (see Appendix 4.1), the second term in the Hamiltonian is the increase in the state variable, some stock, valued by the appropriate shadow price. So, H_{Ct} can be regarded as the value of interest at t plus the increase in the value of the stock at t. In that case, the maximum principle condition $\partial H_{Ct}/\partial u_t = 0$ makes a good deal of sense. It says, at every point in time, set the control variable so that it maximises H_{Ct}, which is value plus an increase in value. It is intuitive that such maximisation at every point in time is required for maximisation of the integral.
The equation of motion condition $\dot{x} = f(x_t, u_t)$ ensures that the optimal path is one that is feasible for the system. Aside from transversality, the remaining condition is $\dot{\mu} = \rho\mu_t - \partial H_{C_t}/\partial x_t$ which governs how the shadow, or imputed, price of the state variable must evolve over time.

This condition can be given some intuitive content by considering a model which is, mathematically, further specialised, and which has some economic content. Consider the simplest possible optimal growth model in which the only argument in the production function is capital. Then, the optimal paths for consumption and capital accumulation are given by maximising

[1] As x and u are now single variables, not vectors, we now drop the bold (vector) notation

$$\int_0^\infty U(C_t)e^{-\rho t}dt$$

subject to

$$\dot{K} = Q(K_t) - C_t$$

giving the current-value Hamiltonian

$$H_{Ct} = U(C_t) + \mu_t(Q(K_t) - C_t) = U(C_t) - \mu_t\dot{K}$$

Here the Hamiltonian is current utility plus the increase in the capital stock valued using the shadow price of capital. In Appendices 19.1 and 19.2 we shall explore this kind of Hamiltonian in relation to the question of the proper measurement of national income.

The maximum principle condition here is $\partial H_{Ct}/\partial C_t = \partial U_t/\partial C_t - \mu_t = 0$ which gives the shadow price of capital as equal to the marginal utility of consumption. Given that a marginal addition to the capital stock is at the cost of a marginal reduction in consumption, this makes sense. Here the condition governing the behaviour of the shadow price over time is

$$\dot{\mu} = \rho\mu_t - \partial H_{Ct}/\partial K_t = \rho\mu_t - \mu_t \partial Q_t/\partial K_t$$

where $\partial Q_t/\partial K_t$ is the marginal product of capital. This condition can be written with the proportionate rate of change of the shadow price on the left-hand side, as

$$\dot{\mu}/\mu_t = \rho - (\partial Q_t/\partial K_t)$$

where the right-hand side is the difference between the utility discount rate and the marginal product of capital adjusted for the marginal utility of consumption. The first term on the right-hand side reflects impatience for future consumption and the second term the pay-off to delayed consumption. According to this expression for the proportional rate of change of the shadow price of capital:

a) μ is increasing when 'impatience' is greater than 'pay-off';

b) μ is constant when 'impatience' is equal to 'pay-off';

c) μ is decreasing when 'impatience' is less than 'pay-off'.

This makes sense, given that:

a) when 'impatience' is greater than 'pay-off', the economy will be running down K;

b) when 'impatience' and 'pay-off' are equal, K will be constant;

c) when 'impatience' is less than 'pay-off', the economy will be accumulating K.

These remarks should be compared with the results in Table 14.1 where it will be seen that the calculated shadow price of capital decreases over time, while the shadow price of oil, which is becoming scarcer, increases over time.