Online Appendix A

Deterministic Optimal Control

As far as the laws of mathematics refer to reality,
they are not certain;
and as far as they are certain,
they do not refer to reality.
—Albert Einstein (1879–1955), quoted by J.R. Newman
in *The World of Mathematics*

$$m = L/c^2.$$
—Albert Einstein, the original form
of his famous energy-mass relation $E = mc^2$,
where L is the Lagrangian, sometimes a form of energy
and the cost part of the Hamiltonian
in deterministic control theory

It probably comes as a surprise to many Americans that the Wright brothers, Orville and Wilbur, did not invent flying, but they developed the *first free, controlled and sustained powered flight by man* as reviewed in Repperger's historical perspective on their technical challenges [233]. Indeed, control is embedded in many modern appliances working silently in computers, motor vehicles and other useful appliances. Beyond engineering design there are natural control systems, like the remarkable human brain working with other components of the central nervous system [172]. Basar [21] lists 25 seminal papers on control and Bernstein [29] reviews control history through feedback control. The state and future directions of control of dynamical systems were summarized in the 1988 Fleming panel report [90] and more recently in the 2003 Murray panel report [91].

This chapter provides summary background as a review to provide a basis for examining the difference between deterministic optimal control and stochastic optimal control, treated in Chapter 6. Summarized with commentary are Hamilton's equations, the maximum principle and dynamic programming formulation. A

special and useful canonical model, the linear quadratic (LQ) model, is presented.

A.1   Hamilton's Equations: Hamiltonian and Lagrange Multiplier Formulation of Deterministic Optimal Control

For deterministic control problems [164, 44], many can be cast as systems of ordinary differential equations so there are many standard numerical methods that can be used for the solution. For example, if $\mathbf{X}(t)$ is the state n_x-vector on the state space $\mathcal{X}$ in continuous time t and $\mathbf{U}(t)$ is the control n_u-vector on the control space $\mathcal{U}$, then the differential equation for the deterministic system dynamics is

$$\frac{d\mathbf{X}}{dt}(t) = \mathbf{f}(\mathbf{X}(t), \mathbf{U}(t), t), \quad \mathbf{X}(t_0) = \mathbf{x}_0. \tag{A.1}$$

Here, $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ is called the **plant function** and may be nonlinear. The **cost objective** functional or **performance index** is to achieve the minimal cumulative **running** or **instantaneous costs** $C(\mathbf{x}, \mathbf{u}, t)$ on (t_0, f_f) plus **terminal cost** function $S(\mathbf{x}, t)$, that is,

$$V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}_0, t_0) = \int_{t_0}^{t_f} C\left(\mathbf{X}(t), \mathbf{U}(t), t\right) dt + S\left(\mathbf{X}(t_f), t_f\right). \tag{A.2}$$

Often in deterministic control theory and the calculus of variations, the cost function is also called the **Lagrangian**, i.e., $L(\mathbf{x}, \mathbf{u}, t) = C(\mathbf{x}, \mathbf{u}, t)$, from analogy with classical mechanics. The notation $V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}_0, t_0)$ means that the cost is a functional of the state and control trajectory functions $V[\mathbf{X}, \mathbf{U}, t_f]$, i.e., a function of functions, but also is a function of the values of the initial data $(\mathbf{x}_0, t_0)$, i.e., a function dependence in the ordinary sense. This fairly general functional form with running and terminal costs is called the **Bolza form** of the objective functional. However, the notation $C(\mathbf{x}, \mathbf{u}, t)$ will be used for the instantaneous component of the objective even when it is not a cost and the overall objective is maximization rather than minimization, e.g., the maximization of profit.

Here, the value of the minimum total costs with respect to the control space $\mathcal{U}$ will be considered,

$$v^*(\mathbf{x}_0, t_0) = \min_{\mathbf{U} \in \mathcal{U}} \left[V[\mathbf{X}, \mathbf{U}, t_f](x_0, t_0) \right], \tag{A.3}$$

unless otherwise specified, subject to the initial value problem for the controlled dynamics in (A.1). There is very little difference between the global minimum and the global maximum problem; the smallest value is found in the former and the largest value in the latter. The search in both cases is over all **critical points**, which consist of the set of all **regular points** or local optima, which here are points where the control derivative or gradient is zero, **boundary points** of the control domain and **singular points** or other irregular points. If the control space $\mathcal{U}$ is the whole space $\mathbb{R}^{n_u}$, the control problem is said to be **unconstrained**, or in absence

of constraints, the problem is mainly searching for regular points, assuming there are no singular points, so

$$v^*(\mathbf{x}_0, t_0) = v^{(\mathrm{reg})}(\mathbf{x}_0, t_0) = \min_{\mathbf{U} \in \mathbb{R}^{n_u}} \left[V[\mathbf{X}, \mathbf{U}, t_f](x_0, t_0) \right]. \qquad (A.4)$$

In the Hamiltonian formulation [164], the Bolza form of optimization objective is replaced by a running cost optimal objective extended to include the state dynamics and the new optimization objective function is called the **Hamiltonian**:

$$\mathcal{H}(\mathbf{X}(t), \mathbf{U}(t), \boldsymbol{\lambda}(t), t) \equiv C(\mathbf{X}(t), \mathbf{U}(t), t) + \boldsymbol{\lambda}^T(t)\mathbf{f}(\mathbf{X}(t), \mathbf{U}(t), t), \qquad (A.5)$$

where $\boldsymbol{\lambda}(t)$ is the n_x-vector **Lagrange multiplier**, also called the **adjoint state** or **costate** or auxiliary vector. The Lagrange multiplier provides the objective extension for including the state dynamics. The symbol $\boldsymbol{\lambda}$ should not be confused with the Poisson rate usd in stochastic jump modeling, since the jump rate does not appear in deterministic problems, but both deterministic and stochastic uses are standard notations in the appropriate context.

Theorem A.1. *Gradient Necessary Conditions for a Regular Control Optimum – Interior Point Optimum Principle.*
*Let the Hamiltonian $\mathcal{H}$ have continuous first order derivatives in the state, costate and control vectors, $\{\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}\}$. Then the **necessary conditions** for an **interior point optimum** (maximum or minimum) of the Hamiltonian $\mathcal{H}$ at the optimal set of three vectors, $\{\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t)\}$, marked with an asterisk $(*)$, are called **Hamilton's equations:***

$$\frac{d\mathbf{X}^*}{dt}(t) = \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}}\right)^* \equiv \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}}\right)(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) = \mathbf{f}(\mathbf{X}^*(t), \mathbf{U}^*(t), t), \quad (A.6)$$

$$-\frac{d\boldsymbol{\lambda}^*}{dt}(t) = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^* \equiv \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) = \left(\frac{\partial C}{\partial \mathbf{x}} + \frac{\partial \mathbf{f}^T}{\partial \mathbf{x}}\boldsymbol{\lambda}\right)^*, \quad (A.7)$$

$$\mathbf{0} = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{u}}\right)^* \equiv \left(\frac{\partial \mathcal{H}}{\partial \mathbf{u}}\right)(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) = \left(\frac{\partial C}{\partial \mathbf{u}} + \frac{\partial \mathbf{f}^T}{\partial \mathbf{u}}\boldsymbol{\lambda}\right)^*, \quad (A.8)$$

where from the critical condition (A.8), *the optimal control is the regular control, i.e.,*

$$\mathbf{U}^*(t) = \mathbf{U}^{(\mathrm{reg})}(t),$$

*at a **regular or interior point** and $\mathbf{U}^{(\mathrm{reg})}(t)$ is called a **regular control**, so critical condition* (A.8) *does not necessarily apply to boundary points or singular points of the control, but certainly does apply to the case of unconstrained control. The associated final conditions are listed in Table A.1 below.*

Proof. The proof is a standard optimization proof in the calculus of variations [40, 15, 164, 44] and is a significant generalization of the usual *first derivative optima*

Table A.1. *Some final conditions for deterministic optimal control.*

	$\mathbf{X}(t_f) = \mathbf{x}_f$ Fixed	$\mathbf{X}(t_f)$ Free & t_f-Independent
t_f Fixed	$\mathbf{x}_f^* = \mathbf{X}^*(t_f) = \mathbf{x}_f$ at $t = t_f$	$\boldsymbol{\lambda}_f^* = \boldsymbol{\lambda}^*(t_f) = \nabla_{\mathbf{x}}[S](\mathbf{x}_f^*, t_f)$ at $t = t_f$
t_f Free	$\mathbf{x}_f^{**} = \mathbf{X}^*(t_f^*) = \mathbf{x}_f$ $(\mathcal{H} + S_t)_f^{**} = 0$ at $t = t_f^*$	$\boldsymbol{\lambda}_f^{**} = \boldsymbol{\lambda}^*(t_f^*) = \nabla_{\mathbf{x}}[S](\mathbf{x}_f^{**}, t_f^*)$ $(\mathcal{H} + S_t)_f^{**} = 0$ at $t = t_f^*$

Notation: $\mathbf{x}_f^* \equiv \mathbf{X}^*(t_f)$, $\mathbf{u}_f^* \equiv \mathbf{U}^*(t_f)$, $\boldsymbol{\lambda}_f^* \equiv \boldsymbol{\lambda}^*(t_f)$ *and* $\mathcal{H}_f^* \equiv \mathcal{H}(\mathbf{x}_f^*, \mathbf{u}_f^*, \boldsymbol{\lambda}_f^*, t_f)$ *in the case of fixed final time* t_f, *while* $\mathbf{x}_f^{**} \equiv \mathbf{X}^*(t_f^*)$, $\mathbf{u}_f^{**} \equiv \mathbf{U}^*(t_f^*)$, $\boldsymbol{\lambda}_f^{**} \equiv \boldsymbol{\lambda}^*(t_f^*)$ *and* $\mathcal{H}_f^{**} \equiv \mathcal{H}(\mathbf{x}_f^{**}, \mathbf{u}_f^{**}, \boldsymbol{\lambda}_f^{**}, t_f^*)$ *in the case of free final time with optimal final time* t_f^*.

test. Our formal justification is a brief formulation after Kirk's description [164], but in our notation.

Note that the gradient

$$\left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^* \equiv \nabla_{\mathbf{x}}[\mathcal{H}](\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) = \left[\frac{\partial \mathcal{H}}{\partial x_i}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t)\right]_{n_x \times 1},$$

so is the $\mathbf{x}$-gradient and a column n_x-vector like $\mathbf{X}$ itself here (elsewhere row vector gradients may be used, e.g., [44]) The gradients of C and $\mathbf{f}$ have corresponding dimensions. The triple set (A.6), (A.7), (A.8) of equations form a set of three vector ordinary differential equations for the optimal trajectory under the optimal control $\mathbf{U}^*(t)$. The first equation (A.6) merely reaffirms the specified state dynamical system (A.1) and that the inclusion with the Lagrange multiplier $\boldsymbol{\lambda}^*(t)$ is proper. The prefix minus on the time derivative of the Lagrange multiplier in (A.7) indicates that it is a backward-time ODE, in contrast to the forward-time state ODE (A.6).

For the calculus of variations, the objective (A.2) is extended in two ways. First, the terminal cost is absorbed in the integral of running costs using the fundamental theorem of calculus,

$$S(\mathbf{X}(t_f), t_f) = S(\mathbf{x}_0, t_0) + \int_{t_0}^{t_f} \frac{dS}{dt}(\mathbf{X}(t), t)dt$$
$$= S(\mathbf{x}_0, t_0) + \int_{t_0}^{t_f} \left(\frac{\partial S}{\partial t}(\mathbf{X}(t), t) + \dot{\mathbf{X}}^\top(t)\frac{\partial S}{\partial \mathbf{x}}(\mathbf{X}(t), t)\right) dt,$$

noting that the initial condition $S(\mathbf{x}_0, t_0)$ is fixed and so can be ignored in the optimization, but the final time t_f will be allowed to be free rather than fixed.

Second, the negative of the state derivative, $-\dot{\mathbf{X}}(t)$, is included in the Lagrange coefficient of the Hamiltonian. Thus, the extended or augmented objective is

$$V^+[\mathbf{Z}, \dot{\mathbf{X}}, t_f] \equiv \int_{t_0}^{t_f} C^+(\mathbf{Z}(t), \dot{\mathbf{X}}(t), t)dt, \tag{A.9}$$

where for brevity an extended state vector is defined as

$$\mathbf{Z}(t) \equiv \left[\begin{array}{c} \mathbf{X}(t) \\ \mathbf{U}(t) \\ \boldsymbol{\lambda}(t) \end{array} \right] \tag{A.10}$$

and the extended cost function is

$$C^+(\mathbf{Z}(t), \dot{\mathbf{X}}(t), t) \equiv \mathcal{H}(\mathbf{Z}(t), t) + \frac{\partial S}{\partial t}(\mathbf{X}(t), t) + \dot{\mathbf{X}}^\top(t)\left(\frac{\partial S}{\partial \mathbf{x}}(\mathbf{X}(t), t) - \boldsymbol{\lambda}(t)\right). \tag{A.11}$$

The objective extension also enables the optimal treatment of the final or stopping time t_f when t_f is a free variable.

Next, the variations of the independent variables about potential optima, e.g., $\mathbf{Z}^*(t)$, are introduced,

$$\mathbf{Z}(t) \equiv \mathbf{Z}^*(t) + \delta\mathbf{Z}(t);$$
$$\dot{\mathbf{X}}(t) \equiv \dot{\mathbf{X}}^*(t) + \delta\dot{\mathbf{X}}(t);$$
$$t_f \equiv t_f^* + \delta t_f,$$

the latter permitting optimal stopping times t_f^* in addition to free final states for generality. Assuming all variations are small and neglecting higher order variations, i.e., $O(|\delta\mathbf{Z}(t)|^2)$, a preliminary form of the first variation of the extended objective

$$V^+[\mathbf{Z}, \dot{\mathbf{X}}, t_f] \simeq V^+[\mathbf{Z}^*, \dot{\mathbf{X}}^*, t_f^*] + \delta V^+[\mathbf{Z}, \dot{\mathbf{X}}, t_f]$$

is

$$\delta V^+[\mathbf{Z}, \dot{\mathbf{X}}, t_f] \simeq \int_{t_0}^{t_f^*}\left(\delta\mathbf{Z}^\top\left(\frac{\partial C^+}{\partial \mathbf{z}}\right)^* + \delta\dot{\mathbf{X}}^\top\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^*\right) dt + \delta t_f \left.(C^+)^*\right|_{t=t_f^*},$$

where the latter term derives from a forward approximation of the final integral fragment on $[t_f^*, t_f^* + \delta t_f]$ for small first variation δt_f, ignoring second variations. Also, the shorthand notation such as $(\partial C^+/\partial \mathbf{z})^* = (\partial C^+/\partial \mathbf{z})(\mathbf{Z}^*(t), \dot{\mathbf{X}}^*(t), t)$ has been used.

Since

$$\delta\mathbf{X}(t) = \delta\mathbf{X}(t_0) + \int_{t_0}^t \delta\dot{\mathbf{X}}(s) ds,$$

the variation $\delta\dot{\mathbf{X}}(t)$ is not independent of its integral $\delta\mathbf{X}(t)$, but this dependence can be removed by a primary applied mathematics technique of *integration by parts*. So, replacing the objective variation δV^+ by $\delta V^\dagger$ without $\delta\dot{\mathbf{X}}(t)$,

$$\begin{aligned} \delta V^\dagger[\mathbf{Z}, t_f] \simeq \ & \int_{t_0}^{t_f^*}\left(\delta\mathbf{Z}^\top\left(\frac{\partial C^+}{\partial \mathbf{z}}\right)^* - \delta\mathbf{X}^\top\frac{d}{dt}\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^*\right) dt \\ & + \left.\left(\delta t_f \,(C^+)^* + \delta\mathbf{X}^\top\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^*\right)\right|_{t=t_f^*}. \end{aligned}$$

However, the variation

$$\delta \mathbf{X}(t_f^*) \equiv \mathbf{X}(t_f^*) - \mathbf{X}^*(t_f^*)$$

is only the variation at $t = t_f^*$ and not the total final variation required, which is

$$\delta \widehat{\mathbf{X}}(t_f^*) \equiv \mathbf{X}(t_f^* + \delta t_f) - \mathbf{X}^*(t_f^*),$$

the difference between a final trial value at $t_f = t_f^* + \delta t_f$ and a final optimal state value at the optimal stopping time $t = t_f^*$. By using a tangent line approximation, the former can be converted to the other with sufficient first variation accuracy,

$$\delta \widehat{\mathbf{X}}(t_f^*) \simeq \mathbf{X}(t_f^*) + \dot{\mathbf{X}}(t_f^*)\delta t_f - \mathbf{X}^*(t_f^*) \simeq \delta \widehat{\mathbf{X}}(t_f^*) + \dot{\mathbf{X}}^*(t_f^*)\delta t_f,$$

where $\dot{\mathbf{X}}(t_f^*)\delta t_f \simeq \dot{\mathbf{X}}^*(t_f^*)\delta t_f$ within first variation accuracy. Hence, the proper final first variation $\delta \widehat{\mathbf{X}}(t_f^*)$ with tangent correction can be substituted for $\delta \widehat{\mathbf{X}}(t_f^*)$, yielding

$$\delta V^\dagger[\mathbf{Z}, t_f] \simeq \int_{t_0}^{t_f^*} \left(\delta \mathbf{Z}^\top \left(\frac{\partial C^+}{\partial \mathbf{z}} \right)^* - \delta \mathbf{X}^\top \frac{d}{dt} \left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}} \right)^* \right) dt \\ + \left(\delta t_f \left(C^+ - \left(\dot{\mathbf{X}} \right)^\top \left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}} \right)^* \right) + \delta \widehat{\mathbf{X}}^\top \left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}} \right)^* \right) \Bigg|_{t=t_f^*}. \tag{A.12}$$

The **fundamental theorem of the calculus of variations** [164] states that the first variation, here $\delta V^\dagger[\mathbf{Z}, t_f]$, must vanish for all admissible variations, here assuming $\delta \mathbf{Z}(t)$ is continuous, on an optimal trajectory, here $\mathbf{Z}^*(t)$. Thus,

$$\delta V^\dagger[\mathbf{Z}, t_f] = 0.$$

Further, the **fundamental lemma of the calculus of variations** [164] states that given a continuous function $\mathcal{F}_i(t)$ and

$$\int_{t_0}^{t_f} \delta X_i(t)\mathcal{F}_i(t)dt = 0$$

for every continuous trajectory $\delta X_i(t)$ on $[t_0, t_f]$, then

$$\mathcal{F}_i(t) = 0$$

on $[t_0, t_f]$. For multidimensional trajectories and independent component variations $\delta X_i(t)$ for $i = 1{:}n_x$, the result holds for all components.

Using the definition of the extended cost C^+ in (A.11), extended state $\mathbf{Z}$ in (A.10) and the Hamiltonian (A.5) with the first variation $\delta V^\dagger[\mathbf{Z}, t_f]$ in (A.12), we have

• Coefficient of $\delta \boldsymbol{\lambda}^\top(t) \implies$

$$\left(\frac{\partial C^+}{\partial \boldsymbol{\lambda}} \right)^* = \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} \right)^* - \dot{\mathbf{X}}^*(t) = \mathbf{0} \implies$$

$$\dot{\mathbf{X}}^*(t) = \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{\lambda}} \right)^* = \mathbf{f}(\mathbf{X}^*(t), \mathbf{U}^*(t), t) \text{ on } t_0 < t \le t_f.$$

- Coefficient of $\delta\mathbf{X}^\top(t) \implies$

$$\left(\frac{\partial C^+}{\partial \mathbf{x}}\right)^* - \frac{d}{dt}\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^* = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^* + \dot{\boldsymbol{\lambda}}^*(t) = \mathbf{0} \implies$$

$$\dot{\boldsymbol{\lambda}}^*(t) = -\left(\frac{\partial \mathcal{H}}{\partial \mathbf{x}}\right)^* = -\left(\frac{\partial C}{\partial \mathbf{x}} + \frac{\partial \mathbf{f}^T}{\partial \mathbf{x}}\boldsymbol{\lambda}\right)^*, \text{ on } t_0 \le t < t_f.$$

- Coefficient of $\delta\mathbf{U}^\top(t) \implies$

$$\left(\frac{\partial C^+}{\partial \mathbf{u}}\right)^* = \left(\frac{\partial \mathcal{H}}{\partial \mathbf{u}}\right)^* = \left(\frac{\partial C}{\partial \mathbf{u}} + \frac{\partial \mathbf{f}^T}{\partial \mathbf{u}}\boldsymbol{\lambda}\right)^* = \mathbf{0}, \text{ on } t_0 \le t < t_f.$$

Cautionary Remark: *This critical point result is valid only for isolated, interior critical optima, so it would not be valid for the case that $\mathcal{H}$ is linear in $\mathbf{U}$ or a singular case. However, the equations for $\dot{\mathbf{X}}^*(t)$ and $\dot{\boldsymbol{\lambda}}^*(t)$ remain valid.*

- Coefficient of $\delta t_f \implies$
  If t_f is fixed, then $\delta t_f \equiv 0$ and no information can be implied about the coefficient, else if t_f is free and if $\delta t_f \ne 0$ is otherwise arbitrary, then

$$\left((C^+)^* - \left(\dot{\mathbf{X}}^*\right)^\top\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^*\right)\Bigg|_{t=t_f^*} = \left(\mathcal{H}^* + \left(\frac{\partial S}{\partial t}\right)^*\right)\Bigg|_{t=t_f^*} = 0.$$

- Coefficient of $\delta\widehat{\mathbf{X}}^\top(t_f^*) \implies$
  If $\mathbf{X}(t_f)$ is fixed and t_f fixed, then $\delta\widehat{\mathbf{X}}^\top(t_f^*) \equiv 0$ and no information can be implied about the coefficient, else if $\mathbf{X}(t_f)$ is free and t_f is fixed, then $\delta\widehat{\mathbf{X}}^\top(t_f^*) \ne 0$ and

$$\left(\frac{\partial C^+}{\partial \dot{\mathbf{x}}}\right)^*\Bigg|_{t=t_f^*} = \left(\left(\frac{\partial S}{\partial \mathbf{x}}\right)^* - \boldsymbol{\lambda}^*\right)\Bigg|_{t=t_f} = \mathbf{0} \implies$$

$$\boldsymbol{\lambda}^*(t_f) = \frac{\partial S}{\partial \mathbf{x}}(\mathbf{X}^*(t_f), t_f),$$

or else if both $\mathbf{X}(t_f)$ and t_f are free, then the combined conditions are

$$\boldsymbol{\lambda}_f^{**} \equiv \boldsymbol{\lambda}^*(t_f^*) = \frac{\partial S}{\partial \mathbf{x}}(\mathbf{X}^*(t_f^*), t_f^*),$$

$$\left(\mathcal{H} + \frac{\partial S}{\partial t}\right)^{**} \equiv \left(\mathcal{H}^* + \left(\frac{\partial S}{\partial t}\right)^*\right)\Bigg|_{t=t_f^*} = 0,$$

the double asterisk notion denoting the optimal stopping time on the optimal path.

The first three items complete the proof of the theorem, while the last two items complete the justifications of the final conditions listed in Table A.1. □

The state vector $\mathbf{X}^*(t)$ satisfies specified initial conditions $\mathbf{X}^*(t_0) = \mathbf{x}_0^*$ at t_0. The **final conditions** for the state $\mathbf{X}^*(t_f)$ and **costate** or **adjoint state** $\boldsymbol{\lambda}^*(t_f)$, if any, depend on the application and a fairly complete set is tabulated in Kirk [164], Bryson and Ho [44], and Athans and Falb [15]. The final conditions depend on whether the final time t_f is fixed (specified) or free (unspecified) and whether the corresponding final state vector $\mathbf{x}_f = \mathbf{X}(t_f)$ is fixed or free. A partial list of some of the conditions is given in Table A.1.

See the classical sources of Athans and Falb [15], Kirk [164], and Bryson and Ho [44] for additional final conditions such as moving boundaries $\boldsymbol{\Gamma}(\mathbf{X}(t)) = \mathbf{0}$ or $\boldsymbol{\Gamma}(\mathbf{X}(t), t) = \mathbf{0}$ and other variants that enter into the final conditions. For other variants with more economic interpretations, the bioeconomics book by Clark [57] is very readable and useful. Other condition variants include a multitude of mixed and hybrid cases that are vector component combinations of the purely fixed and free vector cases presented in Table A.1. Some of these final conditions arise as natural boundary conditions because they cannot be independently specified but follow from the structure of the optimal control problem by the method of **calculus of variations** [40, 15, 164, 44].

The final conditions for the free terminal time and free terminal state case

$$\boldsymbol{\lambda}_f^{**} = \boldsymbol{\lambda}^*(t_f^*) = \nabla_{\mathbf{x}}[S](\mathbf{x}_f^{**}, t_f^*), \tag{A.13}$$

$$0 = \mathcal{H}(\mathbf{x}_f^{**}, \mathbf{u}_f^{**}, \boldsymbol{\lambda}_f^{**}, t_f^*) + S_t(\mathbf{x}_f^{**}, t_f^*) \tag{A.14}$$

in Table A.1 are a good example of the results from the calculus of variations. The equation (A.13) is the final or **transversality** condition for the optimal Lagrange multipier that usually accompanies the stationary point Euler–Lagrange equations (A.7) for the optimal multiplier and (A.8) for the optimal control [44]. The Euler–Lagrange equations along with the dynamic constraint equation and initial condition (A.1) satisfy a two-point boundary value problem, also called a **final-initial** value problem.

Theorem A.2. *Legendre–Clebsch Sufficient Conditions for Regular Control Optimum.*
*If the Hamiltonian $\mathcal{H}$ (A.5) has continuous second order derivatives in the control vector $\mathbf{u}$, then the **sufficient condition** for a **regular point maximum** is that the Hessian matrix must be **negative definite**, i.e., $\mathcal{H}$ is **concave** at the regular point,*

$$\mathcal{H}_{\mathbf{uu}}^* = \nabla_{\mathbf{u}} \left[\nabla_{\mathbf{u}}^\top [\mathcal{H}] \right] (\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) < \mathbf{0} \tag{A.15}$$

*and the **sufficient condition** for a **regular control minimum** is that the Hessian matrix must be **positive definite**, i.e., $\mathcal{H}$ is **convex** at the regular control,*

$$\mathcal{H}_{\mathbf{uu}}^* = \nabla_{\mathbf{u}} \left[\nabla_{\mathbf{u}}^\top [\mathcal{H}] \right] (\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) > \mathbf{0}. \tag{A.16}$$

These sufficient conditions are called the (strengthened) **Legendre–Clebsch** *conditions.*

The proof is a standard optimization proof in multivariate calculus [263, 221, 44] and is a general form of the so-called *second derivative optimum test.*

If the Legendre–Clebsch conditions do not hold, then extra conditions usually are needed. For example, if $\mathcal{H}$ is linear in the control $\mathbf{u}$, then the control problem may be singular [24] and more basic optimization principles may be needed.

See the next section for how to handle some of these exceptions to **regular control** or **normal control** with the critical, stationary condition with respect to the control $\mathbf{u}$ here, using basic optimization principles in terms of a maximum or minimum principle.

Example A.3. *Regular Control Problem.*
This problem is a simplified fragment of a financial portfolio application. Let the dynamics be linear in the positive scalar state $X(t) > 0$, denoting the measure of the wealth at time t, but bilinear in the control-state, such that

$$\dot{X}(t) \equiv \frac{dX}{dt}(t) = (\mu_0 - U(t))X(t), \quad X(0) = x_0 > 0, \quad 0 \le t \le t_f, \quad \text{(A.17)}$$

where μ_0 is a fixed mean production rate of the wealth and $U(t)$ is the control variable that is a measure of the rate of consumption of the wealth at time t. The consumption is constrained to be nonnegative and bounded above,

$$U^{(\min)} = 0 \le U(t) \le U^{(\max)}. \quad \text{(A.18)}$$

The objective is to maximize the cumulative utility of instantaneous consumption where the utility is a **power law**

$$C(x, u, t) = u^\gamma/\gamma \quad \text{(A.19)}$$

for positive powers $\gamma > 0$, but in the following analysis we will exclude the linear case $\gamma = 1$ to keep this a regular or normal control problem. In addition, let there be terminal wealth utility using the same power law,

$$S(x, t) = x^\gamma/\gamma. \quad \text{(A.20)}$$

Thus, this is a Bolza problem described above, but here the maximum utility is sought rather than the minimum cost. The difference between solving a maximum versus a minimum problem is trivial, as can be seen from the Legendre–Clebsch sufficient conditions, (A.15) and (A.16), with only a difference in the sign of the inequality.

Solution. *The Hamiltonian is then*

$$\mathcal{H}(x, u, \lambda, t) = u^\gamma/\gamma + \lambda(\mu_0 - u)x. \quad \text{(A.21)}$$

Hamilton's equations for a **regular control** *solution are*

$$\dot{X}^*(t) = +\mathcal{H}_\lambda^* = (\mu_0 - U^{(\mathrm{reg})}(t))X^*(t), \tag{A.22}$$

$$\dot{\lambda}^*(t) = -\mathcal{H}_\mathbf{x}^* = -(\mu_0 - U^{(\mathrm{reg})}(t))\lambda^*(t), \tag{A.23}$$

$$0 = \mathcal{H}_u^* = (U^{(\mathrm{reg})})^{\gamma-1}(t) - \lambda^*(t)X^*(t); \tag{A.24}$$

the latter equation yields the regular control,

$$U^{(\mathrm{reg})}(t) = (\lambda^*(t)X^*(t))^{1/(\gamma-1)} \tag{A.25}$$

provided that $\gamma \neq 1$, as promised, i.e., excluding what is called the **neutral risk** *case. Since the control is a regular control, then, strictly speaking, $X^*(t) = X^{(\mathrm{reg})}(t)$ and $\lambda^*(t) = \lambda^{(\mathrm{reg})}(t)$.*

Before designating the regular control as the optimal control, the Legendre–Clebsch second derivative sufficient conditions are examined:

$$\mathcal{H}_{uu} = (\gamma - 1)u^{\gamma-2}; \tag{A.26}$$

it is seen from the Legendre–Clebsch sufficient condition for a maximum that $\mathcal{H}$ is **concave** *or $(\mathcal{H}_{uu})^{(\mathrm{reg})} < 0$, and this condition is only satisfied for $\gamma < 1$, the "low"* **risk adverse** *case. Hence, $U^*(t) = U^{(\mathrm{reg})}$.*

However, for $\gamma > 1$ and **risk-seeking** *utility, the regular control (A.25) yields a minimum since $\mathcal{H}$ is* **convex** *or $(\mathcal{H}_{uu})^{(\mathrm{reg})} > 0$, but it would not be rational to get a minimum utility. If maximizing the utility is needed when $\gamma > 1$, then the control constraints must be used. See Exercise 6 for how to obtain the proper maximum utility solution when $\gamma > 1$.*

The first two Hamilton's equations, though seemingly complicated, can be solved by dividing both sides of the equations and examining them in the phase plane without the time dependence,

$$\frac{dX^*}{d\lambda^*} = -\frac{X^*}{\lambda^*}, \tag{A.27}$$

which is just the product rule of differentiation, $d(X^*\lambda^*)/dt = 0$, slightly rearranged and the solution is

$$X^*\lambda^* = K, \tag{A.28}$$

where K is a constant of integration. Consequently, our optimal control is the regular control and must be a constant as well,

$$U^*(t) = U^{(\mathrm{reg})} = K^{1/(\gamma-1)} \equiv K_0, \tag{A.29}$$

provided $0 \leq U^{(\mathrm{reg})} \leq U^{(\mathrm{max})}$. Constant control means that the state and costate equations here are equations of simple exponential growth, so

$$X^*(t) = x_0 e^{(\mu_0 - K_0)t}, \tag{A.30}$$

$$\lambda^*(t) = \lambda^*(t_f)e^{-(\mu_0 - K_0)(t-t_f)}, \tag{A.31}$$

where the constant K_0 and the final adjoint value $\lambda^*(t_f) = \lambda_f^*$ need to be determined. By the transversality condition in Table A.1 for t_f fixed and $X^*(t_f) = x_f^*$ unspecified,

$$\lambda_f^* = S_x(x_f^*, t_f) = (x_f^*)^{\gamma-1} = \left(x_0 e^{(\mu_0 - K_0)t_f}\right)^{\gamma-1}, \qquad (A.32)$$

using the derivative of the terminal utility $S(x, t)$ in (A.20) and the state solution $X^*(t)$ in (A.30). Finally, the definitions of K in (A.28) and K_0 in (A.29) yield a nonlinear equation for the control constant $U^*(t) = K_0$ using (A.28–(A.32)),

$$K_0 = K^{\frac{1}{\gamma-1}} = (x_f^* \lambda_f^*)^{\frac{1}{\gamma-1}} = (x_f^*)^{\frac{\gamma}{\gamma-1}} = \left(x_0 e^{(\mu_0 - K_0)t_f}\right)^{\frac{\gamma}{\gamma-1}}, \qquad (A.33)$$

in terms of the specified X_0, μ_0 and $\gamma < 1$.

We are assuming that the control constraint $U^{(\max)}$ is sufficiently larger than K_0, so that the control remains regular. Control constraint violations, bang control and linear or singular control are treated in the next section.

The Hamiltonian when $\gamma = 0.5$ is displayed along with some sample optimal wealth state $X^*(t)$ and costate $\lambda^*(t)$ solutions in Figure A.1 such that the Hamiltonian is in Subfigure A.1(a) while the optimal solution for maximum utility is in Subfigure A.1(b) using the Online Appendix C code C.23 called *RegCtrlExample6p1a.m*. The terminal wealth at the terminal time $t_f = 1.0$ starting from $x_0 = 10.0$ is $S = 1.038$ for $\gamma = 0.5$. The mean production rate was $\mu_0 = 0.10$ or *10%* in absence of consumption. The MATLAB modification of Brent's zero finding algorithm *fzero* [88] is used to find the control constant $U^*(t) = K_0$ whose approximate value is 3.715 when $\gamma = 0.5$ to accuracy of order 10^{-15} in satisfying (A.33).
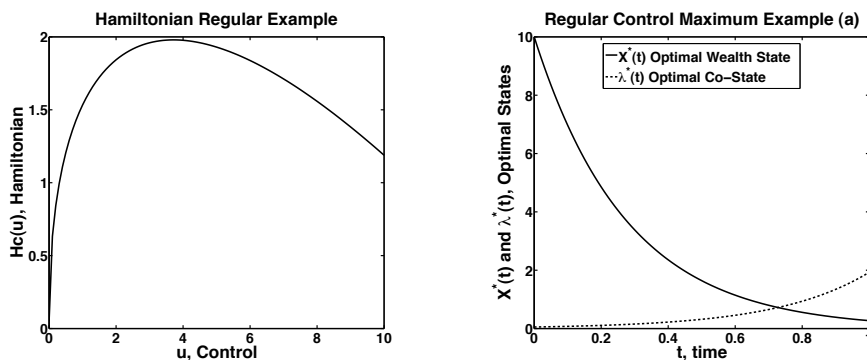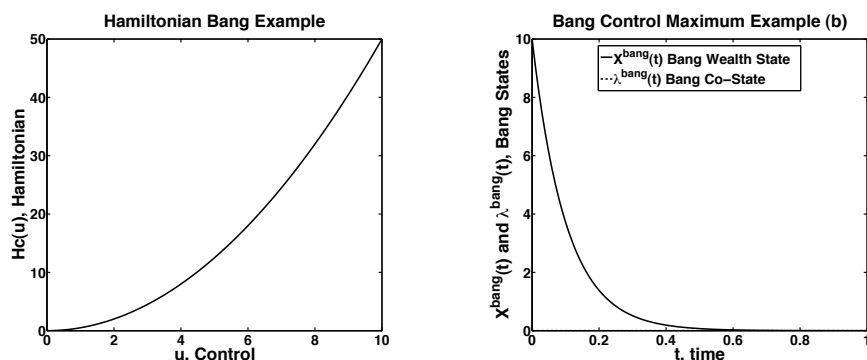

(a) Hamiltonian for regular maximum utility example for power $\gamma = 0.5$.

(b) Optimal paths for regular maximum utility example for power $\gamma = 0.5$.

Figure A.1. *Hamitonian and optimal solutions for regular control problem example from* (A.30) *for* $X^*(t)$ *and* (A.31) *for* $\lambda^*(t)$. *Note that the* $\gamma = 0.5$ *power utility is only for illustration purposes.*

For completeness and to provide a contrasting illustration with a nonregular, bang control case for a power utility with $\gamma = 2.0$, the Hamiltonian and optimal paths

are displayed in Subfigures A.2(a)–A.2(b), *respectively, using the same code C.23 called* `RegCtrlExample6p1a.m`. *The control constant* $U^*(t)$ *has an approximate value of 10.0 when* $\gamma = 2.0$. *The terminal wealth is* $S = 5.02e\text{-}4$ *at the terminal time* $t_f = 1.0$ *starting from* $x_0 = 10.0$ *for* $\gamma = 2.0$. *See Exercise* 6 *for obtaining a proper maximum utility solution when* $\gamma > 1$.

(a) Hamiltonian for end point maximum utility example for power $\gamma = 2.0$.

(b) Optimal paths for end point maximum utility example for power $\gamma = 2.0$.

Figure A.2. *Hamiltonian and optimal solutions for bang control problem example from* (A.30) *for* $X^*(t)$ *and* (A.31) *for* $\lambda^*(t)$. *Note that the* $\gamma = 2.0$ *power utility is only for illustration purposes.*

Remark A.4. *Many control problems are not this easy, since they may require much more analysis, especially in multiple dimensions, and often numerical approximation is needed. For more information on optimal finance portfolios with consumption, see Section* 10.4 *in Chapter* 10 *on financial applications.*

A.1.1 Deterministic Computation and Computational Complexity

Except for simple or analytical homework problems, usually numerical discretization and iterations are required until the solution $(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t))$ converges to some prescribed accuracy. If there are n_t discrete time nodes, $T_k = t_0 + (k-1)\Delta T$ for $k = 1 : N_t$ with $\Delta T = (t_f - t_0)/(N_t - 1)$, then the n_x dimensional state vector $\mathbf{X}^*(t)$ is discretized into $\mathbf{X}^*(T_k) = \mathbf{X}_k = [X_{i,k}]_{n_x \times N_t}$ or $n_x \cdot N_t$ discrete variables. For the three-vector solution the **computational complexity** or the order of the **computational cost** [111] is

$$CC(n_x, n_t) = O(3n_x \cdot N_t) \tag{A.34}$$

per iteration, i.e., bilinear in the dimension and number of time nodes, a very manageable computational problem, even for today's powerful personal computers.

In addition, MATLAB [210] has a good number of control Toolboxes for handling problems. There are also several good online tutorials available, such as Tilbury and Messner's [268, 205] *Control Tutorials for MATLAB and Simulink.*

Some early surveys on computational methods for optimal control problems are by Larson [182], Dyer and McReynolds [77], and Polak [227].

A.2   Optimum Principles: The Basic Principles Approach

For many problems, as discussed in Section B.15 of Online Appendix B of preliminaries, the unconstrained or regular control conditions expressed by Hamilton's equations (A.6), (A.7), (A.8) are in general inadequate. The inadequacy arises in problems for which the optima are not located at interior points but are located at the boundaries of the state and control domains, such as when the domains have bounded constraints in addition to dynamical constraints like (A.1). One exceptional case is the linear control problem. Another exception is when the optima are at interior points at which the derivatives in Hamilton's equations cease to exist, or any of the multitude of combinations of these exceptions depending on all or a subset of the components of the variables involved.

Basic Optimum Principle. *Hence, for general optimization theory and its application, it is essential to return to basic optimization principles, that the **global minimum is the smallest** or that the **global maximum is the biggest**.*

Example A.5. ***Simple static Example of State-Dependent Control with Quadratic Costs and Control Constraints.***
Consider the static quadratic cost function with scalar control u and state x,

$$\mathcal{H}(x,u) = C(x,u) = 2 + x + \frac{1}{2}x^2 - xu + \frac{1}{2}u^2 = 2 + x + \frac{1}{2}(u-x)^2 \quad \text{(A.35)}$$

with control constraints

$$-1 \leq u \leq +1, \tag{A.36}$$

but without any dynamical constraints like (A.1). The objective is to find the optimal control law and optimal cost.
Solution: *The control gradient or derivative is*

$$\frac{\partial C}{\partial u}(x,u) = -x + u,$$

*yielding the critical, stationary point with respect to the control, called a **regular control** in control theory,*

$$U^{(\mathrm{reg})}(x) = x,$$

which would be the global minimum in absence of control constraints since the second partial with respect to the control is positive, $C_{uu}(x,u) = +1 > 0$ with corresponding

regular cost

$$C^{(\mathrm{reg})}(x) \equiv C(x, u^{(\mathrm{reg})}(x)) = 2 + x$$

that is linear (affine) in the state variable.

However, this example has control constraints (A.36) which forces the correct optimal control to assume the constrained values when the regular control goes beyond those constraints, i.e.,

$$U^*(x) = \left\{ \begin{array}{cc} -1, & x \le -1 \\ x, & -1 \le x \le +1 \\ +1, & +1 \le x \end{array} \right\}. \tag{A.37}$$

*This type of optimal control could be called a **bang-regular-bang** control, where the term **bang** signifies hitting the control constraints, the control boundaries becoming active. The corresponding correct optimal cost is*

$$C^*(x) = C(x, u^*(x)) = \left\{ \begin{array}{cc} 2 + x + \frac{1}{2}(x+1)^2, & x \le -1 \\ 2 + x, & -1 \le x \le +1 \\ 2 + x + \frac{1}{2}(x-1)^2, & +1 \le x \end{array} \right\}. \tag{A.38}$$

For this example, $C^*(x)$ is continuous and continuously differentiable, but not twice continuously differentiable. However, optimal controls and optimal costs of some problems can have much worse analytic properties. The optimal solution (A.38) for this simple, somewhat artificial, static optimal control problem is illustrated in Figure A.3 with the optimal control in Subfigure A.3(a) and the optimal cost in Subfigure A.3(b), using the Online Appendix C code C.24 called $SimpleOptExample.m$. *This simple example provides motivation about why the stationary optimality condition (A.8) for the optimal control is not generally valid.*
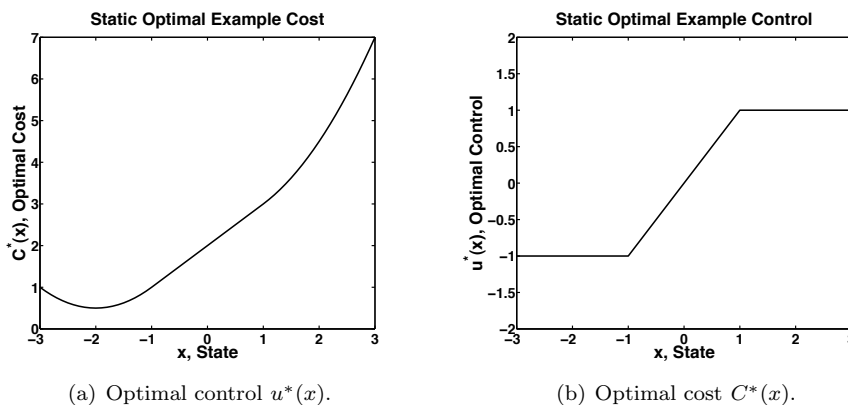

(a) Optimal control $u^*(x)$.   (b) Optimal cost $C^*(x)$.

Figure A.3. *Optimal solutions for a simple, static optimal control problem represented by (A.35) and (A.36), respectively.*

The **basic optimum principle** is just the underlying principle for optimization, but the rigorous justification is beyond the scope of this text. In control theory the optimum principle is associated with the name **Pontryagin maximum principle** [226] in the Russian literature, where the Hamiltonian is formed with an extra multiplier λ_0 to include the objective functional as the 0th dynamical constraint

$$\dot{X}_0(t) = C(X(t), U(t), t),$$

so the maximum refers to the Hamiltonian when the objective is minimum costs and λ_0 must be nonpositive. (See also (A.39) below.) Often the optimum principle is called the **minimum principle** in the English literature [164, 44, 258], particularly when dealing with minimum cost problems, though not exclusively. The difference between a maximum and a minimum principle is essentially a difference in the sign of the Hamiltonian and the fact that the conversion from a maximum objective to a minimum objective problem is quite simple:

$$\max_u[F(u)] = -\min_u[-F(u)]. \tag{A.39}$$

With regard to applications, which version of the optimum principle is used depends on the whether the optimal objective is minimum costs or maximum profit, or minimum energy or maximum energy, minimum time or maximum speed, and there are many other objective choices:

- Minimum time ($C = 1$ and $S = 0$).

- Minimum control ($C = |u|$ and $S = 0$).

- Minimum fuel ($C = |u|$, i.e., thrust measure of fuel consumption, and $S = 0$).

- Minimum energy ($C = u^2$, i.e., energy, and $S = 0$).

- Minimum net profit ($C = p_0 X - c_0$, i.e., profit less cost, and $S = 0$).

- Maximum utility of consumption ($C = \mathcal{U}(u)$, i.e., utility of consumption, and $S = \mathcal{U}(x)$, i.e., utility of portfolio wealth).

- Maximum thrust angle ($C = sin(\theta(t))$ and $S = 0$).

- Minimum distance.

- Minimum surface area.

Here, the maximum and minimum principles are only stated, but see the references at the end of the chapter for more information, such as Anderson-Moore [8], Athans-Falb [15], Bryson-Ho [44], Kirk [164], Pontryagin et al. [226] and Bell-Jacobson [24]. While the statement of the principle seems very intuitive, the rigorous proof is far from easy.

Theorem A.6. *Optimum Principles.*
The **necessary condition for a maximum** *or* **maximum principle** *is*

$$\mathcal{H}^* = \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) \geq \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}(t), \boldsymbol{\lambda}^*(t), t), \tag{A.40}$$

*but the **necessary condition for a minimum** or **minimum principle** is*

$$\mathcal{H}^* = \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) \leq \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}(t), \boldsymbol{\lambda}^*(t), t), \qquad (A.41)$$

in general replacing (A.8), *where* $\mathbf{X}^*(t)$ *and* $\boldsymbol{\lambda}^*(t)$ *are candidates for optimal state or costate, respectively. The optimal state* $\mathbf{X}^*(t)$ *must satisfy the dynamical constraint* $\dot{\mathbf{X}}^*(t) = (\mathcal{H}_{\boldsymbol{\lambda}})^*$ *(A.6) and the costate* $\boldsymbol{\lambda}^*(t)$ *must satisfy the costate equation* $\dot{\boldsymbol{\lambda}}^*(t) = -(\mathcal{H}_{\mathbf{x}})^*$ *(A.7). The optimal control* $\mathbf{U}^*(t)$ *is the argument of the corresponding maximum in* (A.40) *or minimum in* (A.41).

Remarks A.7.

- *Note that the optimal principles* (A.40) *and* (A.41) *as in the basic optimizing principles are used as a general replacement for the necessary conditions for a regular point* $\mathcal{H}_u^* = \mathbf{0}$ *(A.8), the Legendre–Clebsch second order sufficient conditions* $\mathcal{H}_{uu}^* < 0$ *(A.15) for a maximum and* (A.16) $\mathcal{H}_{uu}^* > 0$ *for a minimum. However, these first and second order derivative conditions are still valid for interior or regular points.*

- *In fact, Pontryagin et al. [226] justify briefly that the optimum principles are sufficient conditions because they are more basic conditions.*

- *If we let the control perturbation be*

$$\delta\mathbf{U}(t) \equiv \mathbf{U}(t) - \mathbf{U}^*(t), \qquad (A.42)$$

then the corresponding perturbation or variation in the Hamiltonian is

$$\Delta_u\mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) \equiv \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t) + \delta\mathbf{U}(t), \boldsymbol{\lambda}^*(t), t) \quad (A.43)$$
$$- \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t)$$

and the maximum principle can be reformulated as

$$\Delta_u\mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) \leq 0, \qquad (A.44)$$

while the minimum principle can be reformulated as

$$\Delta_u\mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) \geq 0. \qquad (A.45)$$

In the language of the calculus of variations, the optimum principles are that the first variation of the Hamiltonian is negative semidefinite for a maximum, while it is positive semidefinite for a minimum.

- *Concerning the simple static example* A.5, *the perturbation form of the minimum principle* (A.45) *can be used to justify the choice of the bang controls given in* (A.37). *The perturbation for the example is*

$$\Delta_u\mathcal{H}^* = (U^* - x)\delta U^* + \frac{1}{2}(\delta U^*)^2,$$

where only the linear term need be considered for its contribution to the non-negativity of the perturbation since the quadratic term is never negative. When there is minimal bang control, $U^* = -1$, then the perturbation δU^* must necessarily be nonnegative, otherwise the control constraints (A.36) would be violated, so for nonnegativity of the Hamiltonian perturbation the control perturbation coefficient $(-1-x)$ must also be nonnegative or that $x \leq -1$. Similarly, when there is maximal bang control, $U^* = +1$, then the perturbation has to be nonpositive, $\delta U^* \leq 0$, to avoid violating the control constraints. So $\Delta_u \mathcal{H}^* \geq 0$ (A.45) implies that the coefficient $(1-x)$ of δU^* must be nonpositive or that $x \geq +1$.

- *Similar techniques work with the application of the optimum principles to the case where the Hamiltonian is linear in the control. For example, consider the scalar, **linear control** Hamiltonian,*

$$\mathcal{H}(x, u, \lambda, t) = C_0(x,t) + C_1(x,t)u + \lambda(F_0(x,t) + F_1(x,t)u)$$

subject to control constraints

$$U^{(\min)} \leq U(t) \leq U^{(\max)}$$

and such that

$$\mathcal{H}_u(x, u, \lambda, t) = C_1(x,t) + \lambda F_1(x,t) = \mathcal{H}_u(x, 0, \lambda, t),$$

so no regular control exists. However, the perturbed Hamiltonian has the form

$$\Delta_u \mathcal{H}(X^*, U^*, \lambda^*, t) = \mathcal{H}_u(X^*, 0, \lambda^*, t)\delta U^*,$$

*so optimal control is of the **bang-bang** form, which for a minimum of $\mathcal{H}$ using $\Delta_u \mathcal{H} \geq 0$ yields the composite form,*

$$U^*(t) = \left\{ \begin{array}{ll} U^{(\min)}, & (\mathcal{H}_u)^* = C_1(X^*(t),t) + \lambda^*(t)F_1(X^*(t),t) > 0 \\ U^{(\max)}, & (\mathcal{H}_u)^* = C_1(X^*(t),t) + \lambda^*(t)F_1(X^*(t),t) < 0 \end{array} \right\}, \quad (A.46)$$

since for $(\mathcal{H}_u)^* > 0$, then $\delta U^* \geq 0$ or equivalently $U^*(t) = U^{(\min)}$. Similarly when $(\mathcal{H}_u)^* < 0$, then $\delta U^* \leq 0$ or equivalently $U^*(t) = U^{(\max)}$, but if $(\mathcal{H}_u)^* = 0$ no information on either δU^* or $U^*(t)$ can be determined.

Example A.8. *Bang-Bang Control Problem.*

Consider a simple lumped model of a leaky reservoir (after Kirk [164]) given by

$$\dot{X}(t) = -aX(t) + U(t), \quad X(0) = x_0,$$

where $X(t)$ is the depth of the reservoir, $U(t)$ is the net inflow of water at time t and $a > 0$ is the rate of leakage as well as usage. The net inflow is constrained pointwise $0 \leq U(t) \leq M$ for all $0 < t \leq t_f$ and also cumulatively by

$$\int_0^{t_f} U(t)dt = K > 0, \tag{A.47}$$

where K, M and t_f are fixed constants, such that $K \leq M \cdot t_f$ for consistency. Find the optimal control law that maximizes the cumulative depth

$$J[X] = \int_0^{t_f} X(t)dt$$

and optimal depth $X^*(t)$.

Solution. *The extra integral condition* (A.47) *presents a variation on our standard control problem, but can be treated nicely by extending the state space letting $X_1(t) = X(t)$ and $\dot{X}_2(t) = U(t)$ starting at $X_2(0) = 0$, so that $X_2(t_f) = K$ is precisely the constraint* (A.47). *Thus, the Hamiltonian is*

$$\mathcal{H}(x_1, x_2, u, \lambda_1, \lambda_2, t) = x_1 + \lambda_1(-ax_1 + u) + \lambda_2 u, \qquad (A.48)$$

where λ_1 and λ_2 are Lagrange multipliers. The Hamilton equations for the optimal state and costate solutions are

$$\dot{X}_1^*(t) = \mathcal{H}_{\lambda_1}^* = -aX_1^*(t) + U^*(t), \quad X_1^*(0) = x_0;$$
$$\dot{X}_2^*(t) = \mathcal{H}_{\lambda_2}^* = U^*(t), \quad X_2^*(0) = 0;$$
$$\dot{\lambda}_1^*(t) = -\mathcal{H}_{x_1}^* = -1 + a\lambda_1^*(t);$$
$$\dot{\lambda}_2^*(t) = -\mathcal{H}_{x_2}^* = 0.$$

Consequently, $\lambda_2^*(t) = C_2$, a constant, and $X_2^*(t_f) = K$ is fixed. Also, $\lambda_1^*(t) = C_1 \exp(at) + 1/a$ with the constant determined from the transversality condition $\lambda_1^*(t_f) = 0$ of Table
textnormalrefDetFCs with $X_1^*(t_f)$ free and no terminal cost, i.e., $S(x) \equiv 0$, so $C_1 = -\exp(-at_f)/a$ and

$$\lambda_1^*(t) = \frac{1}{a}\left(1 - e^{-a(t_f - t)}\right). \qquad (A.49)$$

Since

$$\mathcal{H}_u^* = \lambda_1^*(t) + \lambda_2^*(t) \neq 0$$

in general, the usual critical point condition will not directly produce an optimal control $U^*(t)$, but a bang-bang control will work. By applying the essential Pontryagin maximum principle (first derivative test) in the form (A.43)–(A.44) with $\delta U(t) = U(t) - U^*(t)$,

$$\Delta_u \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t), t) = (\lambda_1^*(t) + \lambda_2^*(t))(U(t) - U^*(t)) \leq 0,$$

so if $(\lambda_1^*(t) + \lambda_2^*(t)) > 0$, then $U(t) - U^*(t) \leq 0$ and $U^*(t) = \max[U(t)] = M$, but if $(\lambda_1^*(t) + \lambda_2^*(t)) < 0$, then $U(t) - U^*(t) \geq 0$ and $U^*(t) = \min[U(t)] = 0$. If $(\lambda_1^*(t) + \lambda_2^*(t)) = 0$, then $U^*(t)$ cannot be determined. Now, $U^*(t)$ cannot be zero on all of $[0, t_f]$ or be M on all of $[0, t_f]$, because both options would violate the constraint (A.47) in the strict case $K < M \cdot t_f$. In this case and noting that $\lambda_1^*(t)$ is decreasing

in time, there must be a switch time t_s on $[0, t_f]$ such that $\lambda_1^*(t_s) + \lambda_2^*(t_s) = 0$,
$C_2 = \lambda_2^*(t_s) = -\lambda_1^*(t_s) = -(1 - \exp(-a(t_f - t_s)))/a < 0$ *and*

$$X_2^*(t_f) = K = \int_0^{t_s} M dt + \int_{t_s}^{t_f} 0 dt = M t_s,$$

so $t_s = K/M$. The composite bang-bang control law is then

$$U^*(t) = \left\{ \begin{array}{ll} M, & 0 \leq t < t_s \\ 0, & t_s < t \leq t_f \end{array} \right\}, \tag{A.50}$$

and the corresponding state trajectory is given by

$$X_1^*(t) = X^*(t) = x_0 \, e^{-at} + \frac{M}{a} \left\{ \begin{array}{ll} (1 - e^{-at}), & 0 \leq t \leq t_s \\ e^{-at}(e^{+at_s} - 1), & t_s < t \leq t_f \end{array} \right\}. \tag{A.51}$$

The optimal control (A.50)*, state* (A.51) *and the switch time indicator multiplier sum* (A.49)*, $\lambda_1^*(t) + \lambda_2^*(t)$, are plotted together in Figure* A.4 *with sample numerical parameter values using Online Appendix* C *code* C.25 *called* `bangbangdetctrl05-fig1.m`*.*
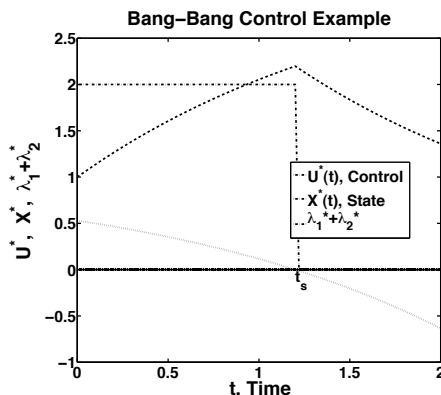

Bang–Bang Control Example

Figure A.4. *Optimal control, state and switch time multiplier sum are shown for bang-bang control example with sample parameter values $t_0 = 0$, $t_f = 2.0$, $a = 0.6$, $M = 2$, $K = 2.4$ and $x_0 = 1.0$. The computed switch time t_s is also indicated.*

Example A.9. *Singular Control Problem.*
Consider the scalar dynamical system for a natural resource with state or mass $X(t)$

$$\dot{X}(t) \equiv \frac{dX}{dt}(t) = (\mu_0 - U(t))X(t), \quad X(t_0) = x_0 > 0, \quad t_0 \leq t \leq t_f, \tag{A.52}$$

where μ_0 is the natural growth rate and $U(t)$ is the harvest rate or effort that will be taken as the control variable. Thus, (A.52) *represents exponential growth of the*

resource whose growth rate is modified by the control. Let the running "cost" for the objective functional be

$$C(x, u, t) = e^{-\delta_0 t} \max \left[p_0 x - c_0, 0 \right] u(t), \tag{A.53}$$

where $p_0 > 0$ is the fixed price per unit effort per unit mass and $c_0 > 0$ is the fixed cost per unit effort, so $p_0 X(t) - c_0$ is the net instantaneous profit at time t.

Note that only positive profit is considered to avoid the possibility of loss, so $X(t) > c_0/p_0$ needs to be maintained. Since the objective concerns profit rather that costs, the objective will be the maximization of profit and the maximum version of the optimum principle is applicable here. The factor $\delta_0 > 0$ is the fixed discount rate or time value of money, but $\delta_0 > \mu_0$ is also assumed as a result of the analysis. There is no terminal cost S. Since real applications have constraints, let the control domain be defined by

$$0 \leq U(t) \leq U^{(\mathrm{max})}, \tag{A.54}$$

where $U^{(\mathrm{max})}$ is positive but whose value is left open for the moment. Since the dynamics are linear and the initial condition is positive, the state domain will also be positive values $X(t) > 0$.

Solution. *To find the solution, the Hamiltonian is written*

$$\mathcal{H}(x, u, \lambda, t) = C(x, u, t) + \lambda \dot{X} = e^{-\delta_0 t}(p_0 x - c_0)u + \lambda(\mu_0 - u)x,$$

assuming a positive profit. Before applying basic optimization principles, we first seek critical, stationary solutions in the control dependence. The control derivative is

$$\mathcal{H}_u(x, u, \lambda, t) = e^{-\delta_0 t}(p_0 x - c_0) - \lambda x, \tag{A.55}$$

which is independent of the control u and when set to zero for stationarity yields the optimal candidate for the adjoint variable, say,

$$\widehat{\lambda}(t) = e^{-\delta_0 t} \left(p_0 - c_0/\widehat{x}(t) \right).$$

However, the other Hamilton's equations specify the potential optimal dynamics of the adjoint and state variables,

$$\dot{\widehat{\lambda}}(t) = -\widehat{\mathcal{H}}_x = -e^{-\delta_0 t} p_0 \widehat{u}(t) - \widehat{\lambda}(t)(\mu_0 - \widehat{u}(t)), \tag{A.56}$$

$$\dot{\widehat{x}}(t) = \widehat{\mathcal{H}}_\lambda = (\mu_0 - \widehat{u}(t))\widehat{x}(t). \tag{A.57}$$

*So, combining the last three equations, it is found that the control terms $\widehat{u}$ cancel out exactly. Consequently, this yields a **singular solution** for the state,*

$$X^{(\mathrm{sing})} = \widehat{x}(t) = (c_0/p_0)/(1 - \mu_0/\delta_0). \tag{A.58}$$

This singular solution leads to the requirement that $\delta_0 > \mu_0$ to maintain the profit restriction that $X(t) > c_0/p_0$. Note that the singular solution in this case is also a

constant. The solution (A.58) *is called a singular solution, rather that a regular or normal solution, since* (A.55) *does not define a stationary point or regular control and by the way the control cancels out due to the linear dependence on control. However, the singular control can be recovered from inverting the state dynamics,*

$$U^{(\mathrm{sing})} = \widehat{u}(t) = \mu_0 - \dot{X}^{(\mathrm{sing})}/X^{(\mathrm{sing})} = \mu_0.$$

For the optimal solution, the control constraints and the initial condition $X(0) = x_0$ *need to be considered.*

If $U^{(\max)} \geq \mu_0$*, then* $U^*(t) = U^{(\mathrm{sing})} = \mu_0$ *and* $X^*(t) = X^{(\mathrm{sing})}$ *on* $0 < t^* \leq t \leq T_0^{(\max)}$*, where* $T_0^{(\max)}$ *is a transition time where the initial trajectory connects to the singular trajectory at a point that is called a* **corner***. The initial trajectory must be chosen using the control bound that allows the singular path to be reached and this control trajectory could be called a* **bang control** *trajectory.*

If $X^{(\mathrm{sing})} < x_0$ *and* $U^{(\max)} > \mu_0$*, then* $U^*(t) = U^{(\max)}$ *on* $[0, T_0^{(\max)}]$*, where the maximal state trajectory starting from* x_0 *at* $t = 0$ *integrating* (A.57) *is*

$$X_0^{(\max)}(t) = x_0 \exp\left(\left(\mu_0 - U^{(\max)}\right)t\right), \quad 0 \leq t \leq T_0^{(\max)},$$

$$T_0^{(\max)} = -\frac{\ln\left(X^{(\mathrm{sing})}/x_0\right)}{\left(U^{(\max)} - \mu_0\right)} > 0.$$

If $X^{(\mathrm{sing})} > x_0$*, then* $U^*(t) = 0$ *on* $[0, T_0^{(\min)}]$ *where the minimal state trajectory starting from* x_0 *at* $t = 0$ *integrating* (A.57) *is*

$$X_0^{(\min)}(t) = x_0 e^{\mu_0 t}, \quad 0 \leq t \leq T_0^{(\min)},$$

$$T_0^{(\min)} = +\frac{\ln\left(X^{(\mathrm{sing})}/x_0\right)}{\mu_0} > 0.$$

At the final time the adjoint final or transversality condition must be used as a final value condition for the adjoint dynamics (A.56)*, which from the scalar version of the entry for fixed* t_f *and free* $X(t_f)$ *in Table* A.1 *on p.* A4 *is*

$$\lambda^*(t_f) = S_x(\mathbf{x}_f^*, t_f) = 0,$$

since there is no terminal value S *in this example. Note that this is consistent with the maximum principle using the calculus of variation and that the regular, critical relation* $\mathcal{H}_u = 0$ *cannot be used as it was for the singular path. Obviously, it is necessary to use the maximal control in* (A.56) *to reach the condition* $\lambda^*(t_f) = 0$ *from the singular path,*

$$\lambda^{(\mathrm{sing})}(t) = e^{-\delta_0 t} p_0 \mu_0/\delta_0,$$

since that leads to a positive running cost and the minimum control cannot be used to physically reach $\lambda^*(t_f) = 0$*. Letting* $\lambda_f(t) = \widehat{\lambda}(t)$ *be the solution of the adjoint*

dynamics equation (A.56) *with conditions* $\widehat{\lambda}(T) = 0$ *and connection or corner time* T_f *such that* $\lambda_f(T_f) = \widehat{\lambda}(T_f) = \lambda^{(\mathrm{sing})}(T_f)$, *then*

$$T_f = t_f + \frac{\ln\left(1 - \mu_0(\delta_0 + U^{(\max)} - \mu_0)/(\delta_0 U^{(\max)})\right)}{(\delta_0 + U^{(\max)} - \mu_0)}.$$

Given the value of T_f, *the corresponding state trajectory is*

$$X_f(t) = X^{(\mathrm{sing})} e^{-(U^{(\max)} - \mu_0)(t - T_f)}$$

on $[T_t, t_f]$.

Thus, the composite optimal control might be called **bang-singular-bang** with the form

$$U^*(t) = \left\{ \begin{array}{l} \left\{ \begin{array}{ll} U^{(\max)}, & 0 \le t \le T_0^{(\max)} \\ U^{(\mathrm{sing})} = \mu_0, & T_0^{(\max)} \le T_f \end{array} \right\}, \quad x_0 > X^{(\mathrm{sing})} \\[4mm] \left\{ \begin{array}{ll} 0, & 0 \le t \le T_0^{(\min)} \\ U^{(\mathrm{sing})} = \mu_0, & T_0^{(\min)} \le T_f \end{array} \right\}, \quad x_0 < X^{(\mathrm{sing})} \\[4mm] \qquad U^{(\max)}, \qquad\qquad\qquad T_f \le t \le t_f \end{array} \right\} \quad (A.59)$$

and a composite optimal state trajectory is

$$X^*(t) = \left\{ \begin{array}{l} \left\{ \begin{array}{ll} X_0^{(\max)}(t), & 0 \le t \le T_0^{(\max)} \\ X^{(\mathrm{sing})} & T_0^{(\max)} \le T_f \end{array} \right\}, \quad x_0 > X^{(\mathrm{sing})} \\[4mm] \left\{ \begin{array}{ll} X_0^{(\min)}(t), & 0 \le t \le T_0^{(\min)} \\ X^{(\mathrm{sing})}, & T_0^{(\min)} \le T_f \end{array} \right\}, \quad x_0 < X^{(\mathrm{sing})} \\[4mm] \qquad X_f(t), \qquad\qquad\qquad T_f \le t \le t_f \end{array} \right\}, \quad (A.60)$$

where it has been assumed for both $U^*(t)$ *in* (A.59) *and* $X^*(t)$ *in* (A.60) *that* $T_0^{(\min)} < T_f$ *or* $T_0^{(\max)} < T_f$ *so that there is a nontrivial singular path. Thus, the possibility of a pure* **bang-bang** *control is excluded, for example, when a minimal bang path* $X_0^{(\min)}(t)$ *from* x_0 *intersects the maximal bang path* $X_f(t)$ *from* x_f^* *before hitting the singular path* $X^{(\mathrm{sing})}$.

Note that this solution is for the case, $U^{(\max)} > \mu_0$. The case for $U^{(\max)} \le \mu_0$ *is left as an open problem in Exercise* 7 *for the reader, who should realize that some parameter values fail to lead to a control problem solution. One possible reason for this failure is the realistic assumption that the control is bounded and does not allow the state to jump from the initial condition to the singular path. Unbounded control that could do that is called* **impulse control**. *Impulse control could be implemented as a Dirac delta function in the differential equation and more on this matter and similar examples can be found in Clark* [57], *and Bryson and Ho* [44].

Some sample results for this singular control example are displayed in Figure A.5 using model parameters $\mu_0 = 0.08$, $\delta_0 = 0.144$, $p_0 = 5.0$, $c_0 = 12.0$, $t_0 = 0$ and $t_f = 15.0$. In Subfigure A.5(a) the optimal state trajectory starts out from $x_0 = 10.0$ at $t = 0$ using maximal bang control with $U^{(\max)} = 0.16$ moving down to reach the singular path at $X^{(\text{sing})} = 9.0$ below when $T_0^{(\max)} = 1.317$, proceeding along the singular path until reaching the singular-bang final corner when $T_f = 8.285$ and then moving down the maximal bang path using $U^{(\max)}$ until reaching the end of the time horizon at $t = t_f = 15.0$. The trajectory displayed in Subfigure A.5(b) is similar except it starts at $x_0 = 8.0$ and moves up to the singular path until reaching the singular path at $(X^{(\text{sing})}, T_0^{(\min)}) = (9.0, 1.472)$; the rest of the path is the same for this example as for the maximal initial bang trajectory. Both subfigures have been generated by the sample Appendix C code C.26 called *SingCtrlExample63.m*.
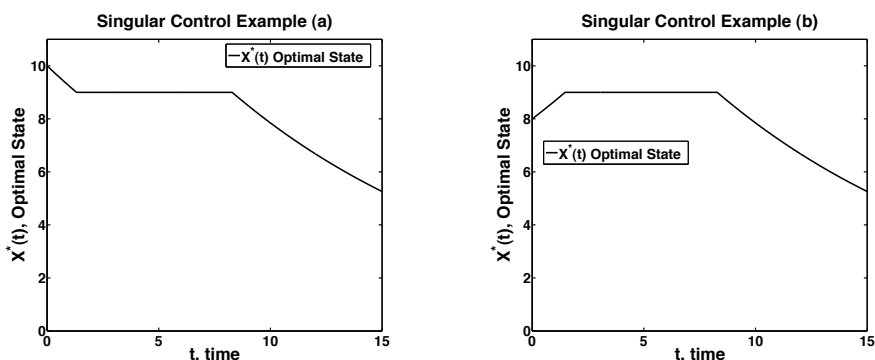
(a) Singular control optimal state $X^*(x)$ when $x_0 = 10.0$.

(b) Singular control optimal state $X^*(x)$ when $x_0 = 8.0$.

Figure A.5. *Optimal state solutions for singular control example leading to a bang-singular-bang trajectory represented by* (A.60). *Subfigure* (a) *yields a maximal bang trajectory from x_0 using $U^{(\max)}$, whereas Subfigure* (b) *yields a minimal bang trajectory from x_0 using $U^{(\min)}$.*

A.3 Linear Quadratic (LQ) Canonical Models

The linear dynamics, quadratic costs or (LQ) problem has the advantage that the regular control can be found fairly explicitly in terms of the state or the costate, thus avoiding the singular complications of linear control problems.

A.3.1 Scalar, Linear Dynamics, Quadratic Costs (LQ)

In the scalar, constant coefficient case the linear dynamics is given by

$$\dot{X}(t) = a_0 X(t) + b_0 U(t), \quad t_0 \le t \le t_f, \quad X(t_0) = x_0 \ne 0, \quad (A.61)$$

where $a_0 \neq 0$ and $b_0 \neq 0$ are assumed so the dynamics is assumed to be nontrivial. The quadratic cost objective is given by

$$V[X, U, t_f](x_0, t_0) = \int_{t_0}^{t_f} C\left(X(t), U(t), t\right) dt + S\left(X(t_f), t_f\right), \qquad (A.62)$$

with the quadratic running cost in state and control,

$$C(x, u, t) = \frac{1}{2} q_0 x^2 + \frac{1}{2} r_0 u^2, \qquad (A.63)$$

where $r_0 > 0$ for minimum costs and $q_0 \geq 0$, while the terminal quadratic cost is quadratic in the state only,

$$S(x, t) = \frac{1}{2} s_0 x^2, \qquad (A.64)$$

where $s_0 \geq 0$. It is assumed there are no bounds on the control $U(t)$ to preserve the nice canonical features of the LQ model. Otherwise the model features would have much more complexity.

Consequently, the Hamiltonian has the form

$$\mathcal{H}(x, u, t) = \frac{1}{2} q_0 x^2 + \frac{1}{2} r_0 u^2 + \lambda(a_0 x + b_0 u). \qquad (A.65)$$

Without control constraints and with quadratic control costs, the regular control policy is the optimal one, governed by the corresponding Hamilton's equations

$$\dot{X}^*(t) = +(\mathcal{H}_\lambda)^* = a_0 X^*(t) + b_0 U^*(t), \qquad (A.66)$$

$$\dot{\lambda}^*(t) = -(\mathcal{H}_x)^* = -q_0 X^*(t) - a_0 \lambda^*(t), \qquad (A.67)$$

$$0 = +(\mathcal{H}_u)^* = r_0 U^*(t) + b_0 \lambda^*(t). \qquad (A.68)$$

The Legendre–Clebsch second order minimum condition is satisfied, since

$$(\mathcal{H}_{uu})^* = r_0 > 0 \qquad (A.69)$$

by the positive definite assumption on r_0. Thus, the optimal control is

$$U^*(t) = U^{(\mathrm{reg})}(t) = -b_0 \lambda^*(t)/r_0, \qquad (A.70)$$

while using (A.70) both the state and costate optimal dynamics satisfies a linear first order matrix system of differential equations,

$$\dot{\mathbf{Z}}(t) \equiv \begin{bmatrix} \dot{X}^*(t) \\ \dot{\lambda}^*(t) \end{bmatrix} = M\mathbf{Z}(t) \equiv \begin{bmatrix} a_0 & -b_0^2/r_0 \\ -q_0 & -a_0 \end{bmatrix} \mathbf{Z}(t). \qquad (A.71)$$

The matrix differential equation (A.71) has the general eigensolution,

$$\mathbf{Z}(t) = c_1 e^{\mu_1(t-t_0)} \begin{bmatrix} 1 \\ (a_0 - \mu_1) r_0/b_0^2 \end{bmatrix} + c_2 e^{-\mu_1(t-t_0)} \begin{bmatrix} 1 \\ (a_0 + \mu_1) r_0/b_0^2 \end{bmatrix}, \qquad (A.72)$$

where c_1 and c_2 are constants of integration and

$$\mu_1 = \sqrt{a_0^2 + q_0 b_0^2 / r_0} \qquad (A.73)$$

is the principal eigenvalue of the matrix M defined in (A.71). This eigenvalue must be real by the coefficient assumptions, but $q_0 > -r_0 a_0^2 / b_0^2$ would be a sufficient condition for μ_1 to be real instead of the condition $q_0 > 0$.

The constants of integration (c_1, c_2) are determined by the initial condition

$$X^*(t_0) = x_0$$

from the first component of $\mathbf{Z}(t)$ in (A.72) and since t_f is fixed but not $X(t_f)$, the final or transversality condition in Table A.1 on p. A4 provides a second condition,

$$\lambda^*(t_f) = S_x(X^*(t_f), t_f) = s_0 X^*(t_f) \qquad (A.74)$$

from the second component of $\mathbf{Z}(t)$ in (A.72). Upon substitution of the constants of integration, the solution $(X^*(t), \lambda^*(t))$ can be found explicitly, say, by symbolic computation systems such as MAPLE or Mathematica, but it is too long and complicated to present here. However, an important property is that both $X^*(t)$ and $\lambda^*(t)$ are proportional to the initial state. The linear feedback relationship between the optimal control and the optimal state can be found from these two solutions and the linear relationship between the optimal control and the costate in (A.70) yields a **linear feedback control law**,

$$U^*(t) = K(t) X^*(t), \qquad (A.75)$$

where

$$K(t) = -(b_0/r_0)\lambda^*(t)/X^*(t), \qquad (A.76)$$

which is called the **feedback gain** coefficient and is independent of the initial state x_0 since it cancels out of the costate to state ratio. The linear feedback control law (A.75) with (A.76) is called feedback or **closed loop control** because it uses state information. However, if the control law is just time-dependent and state-independent, then the law would be called an **open loop control**.

If the plant manager is just concerned with what optimal control input is needed to achieve optimal control in the next time-step, then only the feedback gain is required, assuming the current state output $X^*(t)$ is known. This gain $K(t)$ (sometimes the control law is expressed with a minus sign, $U^*(t) = -\widehat{K}(t)X^*(t)$) can be found directly from a bilinear (quadratic) first order equation, called a **Riccati equation**,

$$\dot{K}(t) = -b_0 K^2(t) - 2a_0 K(t) + b_0 q_0 / r_0, \qquad (A.77)$$

using a numerical differential equation solver backward in time, with just knowledge of the system and cost parameters, as well as the final condition

$$K(t_f) = -b_0 s_0 / r_0 \qquad (A.78)$$

from (A.76) and (A.74).

A.3.2   Matrix, Linear Dynamics, Quadratic Costs (LQ)

In general, LQ control problems will have time-dependent matrix coefficients, and will have both multidimensional vector states and controls. Again, let $\mathbf{X}(t)$ be n_x-dimensional and $\mathbf{U}(t)$ be n_u-dimensional. With some more effort the matrix form of the LQ problem can be solved, using the symbolic tools of MAPLE and Mathematica or the numerical tools of MATLAB.

Let the matrix form of the linear (L) state dynamics be

$$\dot{\mathbf{X}}(t) = A(t)\mathbf{X}(t) + B(t)\mathbf{U}(t), \quad t_0 \leq t \leq t_f, \quad \mathbf{X}(t_0) = \mathbf{x}_0, \qquad (A.79)$$

where the coefficient matrices are $A(t) = [a_{i,j}]_{n_x \times n_x}$ and $A(t) = [b_{i,j}]_{n_x \times n_u}$, commensurate in matrix-vector multiplication. The quadratic (Q) cost objective is

$$V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}_0, t_0) = \frac{1}{2} \int_{t_0}^{t_f} \left[\mathbf{X}^\top(t)Q(t)\mathbf{X}(t) + \mathbf{U}^\top(t)R(t)\mathbf{U}(t) \right] dt \quad (A.80)$$

$$+ \frac{1}{2}\mathbf{X}^\top(t_f)S_f(t_f)\mathbf{X}(t_f),$$

where the cost coefficient matrices are all symmetric, $n_x \times n_x$ state cost coefficients $Q(t)$ and $S_f(t)$ are positive semidefinite ($Q(t) \geq 0$, $S_f(t) \geq 0$), while the $n_u \times n_u$ control cost coefficients must be positive definite, $R(t) > 0$ to ensure minimum costs. The Hamiltonian auxiliary objective is

$$\mathcal{H}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t) = \frac{1}{2}\left(\mathbf{x}^\top Q(t)\mathbf{x} + \mathbf{u}^\top R(t)\mathbf{u} \right) + \boldsymbol{\lambda}^\top \left(A(t)\mathbf{x} + B(t)\mathbf{u} \right), \quad (A.81)$$

where $\boldsymbol{\lambda} = [\lambda_i]_{n_x \times 1}$ is the auxiliary costate vector used to include the dynamical constraints to the running cost objective. In absence of control constraints and with $R(t) > 0$, the regular control is the optimal control and Hamilton's equations are

$$\dot{\mathbf{X}}^*(t) = +(\mathcal{H}_{\boldsymbol{\lambda}})^* = A(t)\mathbf{X}^*(t) + B(t)\mathbf{U}^*(t), \qquad (A.82)$$

$$\dot{\boldsymbol{\lambda}}^*(t) = -(\mathcal{H}_{\mathbf{x}})^* = -Q(t)\mathbf{X}^*(t) - A^\top(t)\boldsymbol{\lambda}^*(t), \qquad (A.83)$$

$$\mathbf{0} = (\mathcal{H}_{\mathbf{u}})^* = R(t)\mathbf{U}^*(t) + B^\top(t)\boldsymbol{\lambda}^*(t), \qquad (A.84)$$

where, by the gradient peel theorem (B.131), the transposes of $A(t)$ and $B(t)$ multiply $\boldsymbol{\lambda}^*(t)$ in (A.83) and (A.84), respectively.

Since $R(t) > 0$, i.e., $R(t)$ is positive definite and has positive $R(t)$ eigenvalues, it is invertible (B.134). Hence, the optimal control in absence of control constraints is proportional to the costate vector,

$$\mathbf{U}^*(t) = -R^{-1}(t)B^\top(t)\boldsymbol{\lambda}^*(t). \qquad (A.85)$$

As in the scalar case, we seek to show, as least formally, that the optimal control is also feedback control depending on the state vector $\mathbf{X}^*(t)$. Our approach will resemble the 2×2 scalar solution, but using $(2n_x) \times (2n_x)$ matrices partitioned

into $n_x \times n_x$ submatrices to keep the analysis compact and close to the scalar case as much as possible. Thus, our system has the form

$$\dot{\mathbf{Z}}(t) = M(t)\mathbf{Z}(t), \tag{A.86}$$

where the partitioned forms are

$$\mathbf{Z}(t) \equiv \left[\begin{array}{c} \mathbf{X}^*(t) \\ \boldsymbol{\lambda}^*(t) \end{array} \right], \tag{A.87}$$

which has dimension $(2n_x)$ and

$$M(t) \equiv \left[\begin{array}{cc} A(t) & -B(t)R^{-1}(t)B^\top(t) \\ -Q(t) & -A^\top \end{array} \right], \tag{A.88}$$

which has dimension $(2n_x)\times(2n_x)$. The multiplication of partitioned matrices works essentially the same way that multiplication of nonpartitioned matrices works.

Since the ODE system in (A.87) for $\mathbf{Z}(t)$ is linear, the usual exponential approximations works. So let a simple trial exponential solution form be

$$\mathbf{Z}(t) = Ce^{\mu t}\boldsymbol{\zeta}, \tag{A.89}$$

where C is a constant of integration, μ is a constant exponent coefficient and $\boldsymbol{\zeta}$ is a constant vector with the same $(2n_x)$ dimension as $\mathbf{Z}(t)$. Substitution into (A.87) yields the $(2n_x)$ dimensional eigenvalue problem (B.129)

$$M(t)\boldsymbol{\zeta} = \mu\boldsymbol{\zeta}, \tag{A.90}$$

so there should be $(2n_x)$ eigenvalues $[\mu_i]_{(2n_x)\times 1}$ and $(2n_x)$ associated eigenvectors

$$\boldsymbol{\zeta}_j = [\zeta_{i,j}]_{(2n_x)\times 1}$$

which are represented as columns of the matrix

$$\Psi = \left[\boldsymbol{\zeta}_j\right]_{1\times(2n_x)} \equiv [\zeta_{i,j}]_{(2n_x)\times(2n_x)}. \tag{A.91}$$

Linear superposition of these $(2n_x)$ eigensolutions yields the general solution

$$\mathbf{Z}(t) = \sum_{k=1}^{2n_x} C_k e^{\mu_k t}\boldsymbol{\zeta}_k = (\Psi.*\mathbf{E}(t))\,\mathbf{C} \equiv \widehat{\Psi}(t)\mathbf{C}, \tag{A.92}$$

where $\mathbf{E}(t) \equiv [\exp(\mu_i t)]_{(2n_x)\times 1}$ is the exponential growth vector at the eigenmode rate, the symbol pair $.*$ is the MATLAB dot-multiplication notation for element-wise multiplication (e.g., $\mathbf{x}.*\mathbf{y} = [x_i y_i]_{n_x\times n_x}$ for vector-vector multiplication or $A.*\mathbf{x} = [a_{i,j}x_j]_{n_x\times n_x}$ in matrix-vector multiplication), and

$$\widehat{\Psi}(t) = \left[\begin{array}{cc} \widehat{\Psi}_{11}(t) & \widehat{\Psi}_{12}(t) \\ \widehat{\Psi}_{21}(t) & \widehat{\Psi}_{22}(t) \end{array} \right] \equiv \Psi.*\mathbf{E}(t) = \left[\begin{array}{cc} \Psi_{11}e^{\mu_1 t} & \Psi_{12}e^{\mu_2 t} \\ \Psi_{21}e^{\mu_1 t} & \Psi_{22}e^{\mu_2 t} \end{array} \right] \tag{A.93}$$

is a convenient abbreviation for the coefficient matrix of $\mathbf{C}$, also given partitioned into 4 $n_x \times n_x$ submatrices. The constant of integration vector

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \tag{A.94}$$

is determined from the initial state condition

$$[Z_i(0)]_{n_x \times 1} = \widehat{\Psi}_{11}(0)\mathbf{C}_1 + \widehat{\Psi}_{12}(0)\mathbf{C}_2 = \mathbf{X}^*(0) = \mathbf{x}_0 \tag{A.95}$$

and the final costate or transversality condition for free $\mathbf{X}^*(t_f)$ from Table A.1 on p. A4,

$$\begin{aligned} [Z_{n+i}(t_f)]_{n_x \times 1} &= \widehat{\Psi}_{21}(t_f)\mathbf{C}_1 + \widehat{\Psi}_{22}(t_f)\mathbf{C}_2 \\ &= \boldsymbol{\lambda}^*(t_f) = \frac{1}{2}\nabla_{\mathbf{x}} \left[\mathbf{X}^\top S_f \mathbf{X} \right](t_f) = S_f(t_f)\mathbf{X}(t_f) \\ &= S_f(t_f) \left(\widehat{\Psi}_{11}(t_f)\mathbf{C}_1 + \widehat{\Psi}_{12}(t_f)\mathbf{C}_2 \right). \end{aligned} \tag{A.96}$$

So this final condition is an algebraic equation that is homogeneous in $\mathbf{C}$. Upon rearranging the initial and final conditions, (A.95) and (A.96), the complete linear algebraic problem for $\mathbf{C}$ becomes

$$\begin{aligned} G\mathbf{C} &\equiv \begin{bmatrix} \widehat{\Psi}_{11}(0) & \widehat{\Psi}_{12}(0) \\ \widehat{\Psi}_{21}(t_f) - S_f(t_f)\widehat{\Psi}_{11}(t_f) & \widehat{\Psi}_{22}(t_f) - S_f(t_f)\widehat{\Psi}_{12}(t_f) \end{bmatrix} \mathbf{C} \\ &= \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{bmatrix}. \end{aligned} \tag{A.97}$$

Assuming that the constant coefficient matrix G is invertible (this can be tested by one of the numerical or symbolic toolboxes), then the solution, using partitioning and simplification due to the homogeneity of the final condition, will formally be of the form

$$\mathbf{C} = \widehat{G}^{-1} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11}^{-1} & \widehat{G}_{12}^{-1} \\ \widehat{G}_{21}^{-1} & \widehat{G}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \widehat{G}_{11}^{-1} \\ \widehat{G}_{21}^{-1} \end{bmatrix} \mathbf{x}_0, \tag{A.98}$$

where $\widehat{G}^{-1}$ is the inverse of G, i.e., $\widehat{G}^{-1}G = I_{2n_x \times 2n_x}$. The same relation does not necessarily hold for the $n_x \times n_x$ partitioned matrices, so $\widehat{G}_{i,j}^{-1}$ is not necessarily the inverse of $G_{i,j}$. Hence, the state and costate solutions will be linear in the initial condition vector $\mathbf{x}_0$,

$$\mathbf{X}^*(t) = \left(\widehat{\Psi}_{11}(t)\widehat{G}_{11}^{-1} + \widehat{\Psi}_{12}(t)\widehat{G}_{21}^{-1} \right) \mathbf{x}_0, \tag{A.99}$$

$$\boldsymbol{\lambda}^*(t) = \left(\widehat{\Psi}_{21}(t)\widehat{G}_{11}^{-1} + \widehat{\Psi}_{22}(t)\widehat{G}_{21}^{-1} \right) \mathbf{x}_0. \tag{A.100}$$

Assuming that the coefficient matrix in (A.99) can be inverted so the backward evolution of the state is

$$\mathbf{x}_0 = \left(\widehat{\Psi}_{11}(t)\widehat{G}_{11}^{-1} + \widehat{\Psi}_{12}(t)\widehat{G}_{21}^{-1} \right)^{-1} \mathbf{X}^*(t), \qquad (A.101)$$

then the optimal control is a feedback control, i.e., linear in the state vector, and is given by

$$\mathbf{U}^*(t) = K(t)\mathbf{X}^*(t), \qquad (A.102)$$

where $K(t)$ is the gain matrix, using (A.85) with (A.99)–(A.102). The initial state thus far has been arbitrary and is

$$K(t) = -R(t)^{-1}B^\top(t) \left(\widehat{\Psi}_{21}(t)\widehat{G}_{11}^{-1} + \widehat{\Psi}_{22}(t)\widehat{G}_{21}^{-1} \right) \qquad (A.103)$$

$$\left(\widehat{\Psi}_{11}(t)\widehat{G}_{11}^{-1} + \widehat{\Psi}_{12}(t)\widehat{G}_{21}^{-1} \right)^{-1}.$$

Note that other texts may define the gain matrix differently, some using the state to costate relation, but here we take the view that the user is the plant manager, who would be interested in the relation between the optimal control and the state. See Kalman [157] for justification of (A.103). An alternative to the eigenvalue problem approach to the solution of the dynamic equations, provided that the gain matrix is the main interest, is the Riccati differential equation approach. (See Anderson and Moore [8] or Kirk [164].) Using the state to costate relation,

$$\boldsymbol{\lambda}^*(t) = J(t)\mathbf{X}^*(t), \qquad (A.104)$$

where the matrix $J(t)$ is defined so that

$$K(t) = -R^{-1}(t)B^\top J(t), \qquad (A.105)$$

to avoid having to differentiate the variable coefficients. By differentiating both sides of (A.104) with respect to t, substituting for $\dot{\boldsymbol{\lambda}}^*(t)$ from (A.83), $\dot{\mathbf{X}}^*(t)$ from (A.82), $\boldsymbol{\lambda}^*(t)$ from (A.104), and $\mathbf{U}^*(t)$ from (A.85), and setting the common coefficient of $\mathbf{X}^*(t)$ equal to zero produces the quadratic, **matrix Riccati equation**,

$$\dot{J}(t) = \left[JBR^{-1}B^\top J - JA - A^\top J - Q \right](t) \qquad (A.106)$$

with the final condition

$$J(t_f) = S_f(t_f) \qquad (A.107)$$

from the final condition $\boldsymbol{\lambda}^*(t_f) = S_f(t_f)\mathbf{X}(t_f)$ in (A.96). Hence, $J(t)$ is just an extension of the terminal cost quadratic coefficient $S_f(t)$ for $0 \le t < t_f$. This makes the Riccati problem (A.106) a final value problem rather than an initial value problem. It can be shown that $J(t)$ is symmetric from (A.106) and $S_f(t_f)$ is assumed to be symmetric, so only the upper or lower half of $J(t)$ plus its diagonal need to be

calculated. The control gain matrix $K(t)$ can be recovered using (A.105). Numerical approximation is almost always needed using methods of ordinary differential equation solvers in the numeric and symbolic computational toolboxes or elsewhere.

Once the feedback gain, either as $K(t)$ or $J(t)$, and the optimal state trajectory $\mathbf{X}^*(t)$ are obtained, the corresponding optimal control trajectory can be computed, and then the optimal total cost value $v^*(\mathbf{x}_0, t_0) = \min_U[V[\mathbf{X}, \mathbf{U}](\mathbf{x}_0, t_0)]$ can be computed from (A.3) by integrating the running cost and adding the sum to the terminal cost term.

In the case where the cost function is a full quadratic polynomial in $\mathbf{x}$ and $\mathbf{u}$, i.e., with linear (affine) cost terms, then the control has $\mathbf{X}^*(t)$-independent terms requiring another companion ordinary differential equation for $J(t)$.

A.4   Deterministic Dynamic Programming (DDP)

Dynamic programming is another approach to the optimal control problem whose aim is to obtain the feedback optimal control $\mathbf{u}^*(\mathbf{x}, t)$ and the optimal value $v^*(\mathbf{x}, t)$, rather than primarily seeking the optimal trajectory set $\{\mathbf{X}^*(t), \boldsymbol{\lambda}^*(t), \mathbf{U}^*(t)\}$ using Hamilton's equations (A.6), (A.7), (A.8). The dynamic programming approach is principally due to Bellman [25] and begins with a slightly different formulation of the Bolza problem designed for better analytical manipulation using an arbitrary initial state $\mathbf{X}(t) = \mathbf{x}$ in the state domain. The deterministic dynamical system (A.1) is reformulated as

$$\frac{d\mathbf{X}}{ds}(s) = \mathbf{f}(\mathbf{X}(s), \mathbf{U}(s), s), \quad \mathbf{X}(t) = \mathbf{x}, \tag{A.108}$$

and the objective value functional as

$$V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}, t) = \int_t^{t_f} C(\mathbf{X}(s), \mathbf{U}(s), s)\, ds + S(\mathbf{X}(t_f), t_f) \tag{A.109}$$

with total minimum costs or optimal value starting from $(\mathbf{x}, t)$

$$v^*(\mathbf{x}, t) = \min_{\mathbf{U}(t, t_f]}[V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}, t)] \tag{A.110}$$

and optimal terminal value

$$v^*(\mathbf{x}, t_f) = S(\mathbf{x}, t_f). \tag{A.111}$$

When $t = t_f$ the running cost integral vanishes leaving only the terminal cost term and since the initial state is reduced to the final state when $t = t_f$, then the minimization is no longer operative. The $\mathbf{x}$ in (A.111) thus can be arbitrary, coinciding with the fact that $\mathbf{X}(t_f)$ is unspecified in this optimal control formulation.

A.4.1   Deterministic Principle of Optimality

Dynamic programming relies crucially on a recursion for the current optimal value in terms of a future optimal value called **Bellman's principle of optimality**. The

basic concept is the assumption that the minimization operation in (A.110) can be decomposed over the control path $\mathbf{U}(s)$ for the time variable s on $(t, t_f]$, open on the left since the state $\mathbf{x}$ at time t is given, into a product over increments in time using the **minimization operator multiplicative decomposition** rule, written symbolically without arguments,

$$\min_{\boldsymbol{U}(t,t_f]} \stackrel{\text{op}}{=} \min_{\boldsymbol{U}(t,t+\Delta t]} \min_{\boldsymbol{U}(t+\Delta t,t_f]} \tag{A.112}$$

for some positive time increment Δt such that $t < t + \Delta t < t_f$ and with an analogous rule for maximization. Using this rule and the fact that an integral has a corresponding additive decomposition rule:

$$\int_t^{t_f} C(\mathbf{X}(s), \mathbf{U}(s), s) ds = \int_t^{t+\Delta t} C(\mathbf{X}(s), \mathbf{U}(s), s) ds \tag{A.113}$$
$$+ \int_{t+\Delta t}^{t_f} C(\mathbf{X}(s), \mathbf{U}(s), s) ds.$$

Application of the minimization and integration decompositions leads to

$$
\begin{aligned}
v^*(\mathbf{x}, t) &= \min_{\boldsymbol{U}(t,t+\Delta t]} \left[\int_t^{t+\Delta t} C(\mathbf{X}(s), \mathbf{U}(s), s) ds \right. \\
&\quad + \min_{\boldsymbol{U}(t+\Delta t,t_f]} \left[\int_{t+\Delta t}^{t_f} C(\mathbf{X}(s), \mathbf{U}(s), s) ds \right] + S(\mathbf{X}(t_f), t_f) \bigg] \\
&= \min_{\boldsymbol{U}(t,t+\Delta t]} \left[\int_t^{t+\Delta t} C(\mathbf{X}(s), \mathbf{U}(s), s) ds + v^*(\mathbf{X}(t+\Delta t), t+\Delta t) \right], \quad \text{(A.114)}
\end{aligned}
$$

where the optimal value $v^*(\mathbf{x}, t)$ definition (A.110, (A.109) has been reused when starting at the future state $\mathbf{X}(t + \Delta t) = \mathbf{x} + \Delta\mathbf{X}(t)$ at time $t + \Delta t$. Thus, the following form of the optimality principle has been formally derived.

Lemma A.10. *Bellman's Deterministic Principle of Optimality.*
Under the assumptions of the operator decomposition rules (A.112), (A.113),

$$v^*(\mathbf{x}, t) = \min_{\boldsymbol{U}(t,t+\Delta t]} \left[\int_t^{t+\Delta t} C(\mathbf{X}(s), \mathbf{U}(s), s) ds + v^*(\mathbf{x} + \Delta\mathbf{X}(t), t+\Delta t) \right]. \tag{A.115}$$

A.4.2 Hamilton–Jacobi–Bellman (HJB) Equation of Deterministic Dynamic Programming

In the derivation of the partial differential equation of deterministic dynamic programming or Hamilton–Jacobi–Bellman (HJB) equation, Bellman's principle of optimality is applied for small increments Δt, so Δt is replaced by the differential dt. The future state is approximated by a first order Taylor approximation,

$$\mathbf{X}(t + dt) \stackrel{\text{dt}}{=} \mathbf{X}(t) + \frac{d\mathbf{X}}{dt}(t)dt = \mathbf{x} + \frac{d\mathbf{X}}{dt}(t)dt, \tag{A.116}$$

provided the state vector $\mathbf{X}(t)$ is continuously differentiable. Consequently, the first order approximation for the optimal value $v^*(\mathbf{x}, t)$ according to the principle of optimality with $\mathbf{X}(t) = \mathbf{x}$ is

$$v^*(\mathbf{x}, t) \overset{\mathrm{dt}}{=} \min_{\boldsymbol{U}(t,t+dt]} \bigg[C(\mathbf{x}, \mathbf{U}(t), t)dt + v^*(\mathbf{x}, t) + v_t^*(\mathbf{x}, t)dt \qquad \text{(A.117)}$$
$$+ \ \nabla_{\mathbf{x}}^{\top}[v^*](\mathbf{x}, t) \cdot \mathbf{f}(\mathbf{x}, \mathbf{U}(t), t)dt \bigg],$$

provided $v^*(\mathbf{x}, t)$ is continuously differentiable in $\mathbf{x}$ and t and $C(\mathbf{x}, \mathbf{u}, t)$ is continuous so $\mathrm{o}(dt)$ can be neglected. Note that the optimal value $v^*(\mathbf{x}, t)$ appears alone on both sides of (A.117), so both of these $v^*(\mathbf{x}, t)$ terms can be cancelled. Upon letting $\mathbf{U}(t) \equiv \mathbf{u}$ and replacing the vector set $\mathbf{U}(t, t + dt]$ by $\mathbf{u}$ the PDE of deterministic dynamic programming can be summarized as the following result.

Theorem A.11.   *HJB Equation for Deterministic Dynamic Programming.*
If $v^*(\mathbf{x}, t)$ is once differentiable in $\mathbf{x}$ and once differentiable in t, while the decomposition rules (A.112), (A.113) are valid, then

$$0 = v_t^*(\mathbf{x}, t) + \min_{\mathbf{u}} \left[\mathcal{H}(\mathbf{x}, \mathbf{u}, t) \right] \equiv v_t^*(\mathbf{x}, t) + \mathcal{H}^*(\mathbf{x}, t), \qquad \text{(A.118)}$$

*where the **Hamiltonian** (technically a **pseudo-Hamiltonian**) functional is given by*

$$\mathcal{H}(\mathbf{x}, \mathbf{u}, t) \equiv C(\mathbf{x}, \mathbf{u}, t) + \nabla_{\mathbf{x}}^{\top}[v^*](\mathbf{x}, t) \cdot \mathbf{f}(\mathbf{x}, \mathbf{u}, t). \qquad \text{(A.119)}$$

The optimal control, if it exists, is given by

$$\mathbf{u}^*(\mathbf{x}, t) = \underset{\mathbf{u}}{\operatorname{argmin}} \left[\mathcal{H}(\mathbf{x}, \mathbf{u}, t) \right]. \qquad \text{(A.120)}$$

This **HJB** equation (A.118), (A.119) is no ordinary PDE, but has the following properties or attributes.

Properties A.12.

- *The HJB equation is a **functional PDE** due to the presence of the minimum operator* min.

- *The HJB equation is a **scalar valued equation**, but solution output has dimension $(nu + 1)$ consisting of the scalar optimal value function $v^* = v^*(\mathbf{x}, t)$ and the optimal control vector $\mathbf{u}^* = \mathbf{u}^*(\mathbf{x}, t)$ as well. These **dual solutions** are generally tightly coupled in functional dependence. In general, this tight coupling requires a number of iterations between v^* and $\mathbf{u}^*$ to obtain a reasonable approximation to the $(nu + 1)$-dimensional solution over the $(n_x + 1)$-dimensional space of independent variables $(\mathbf{x}, t)$. However, it should be noted that the optimal control $\mathbf{u}(\mathbf{x}, t)$ in (6.21) is also feedback optimal control if the $\mathbf{x}$ dependence is genuine.*

- *In contrast to the Hamilton's equations formulation, the dynamic programming solution does not give the state trajectory directly but the state dynamics* (A.108) *must be solved using the feedback optimal control* $u^*(\mathbf{X}(t), t)$ *using* (A.120). *If the optimal control solution is computational, which is usual except for special or canonical problems, then the state dynamic solution would also be computational.*

A.4.3 Computational Complexity for Deterministic Dynamic Programming

The state-time vector valued form of the solution set, $\{v^*(\mathbf{x}, t), \mathbf{u}^*(\mathbf{x}, t)\}$, given independent state and time variables, $\mathbf{x}$ and t, makes the dynamic programming quite different from the Hamilton's equations for optimal time-dependent vector trajectories $\{\mathbf{X}(t), \boldsymbol{\lambda}(t), \mathbf{U}(t)\}$. If time is fixed at a single discrete value $T_k = t_0 + (k-1)\Delta T$ for some k, where $k = 1 : N_t$ with $\Delta T = (t_f - t_0)/(N_t - 1)$, then the independent discretization of the n_x-dimensional state vector $\mathbf{x}$ is replaced by $\mathbf{X_j} = [X_{i,j_i}]_{n_x \times 1}$, where $\mathbf{j} = [j_i]_{n_x \times 1}$, $j_i = 1 : N_x$ for $i = 1 : n_x$ and N_x is the common number of state nodes, simply taken to be the same for each component (otherwise, N_x could be the geometric mean of n_x node counts N_i for $i = 1 : n_x$). However, $\mathbf{X_j}$ represents only one point in state space and there is a total $N_x^{n_x}$ numerical nodes or points in n_x state-dimensions. Thus, total numerical representation optimal value $v(\mathbf{x}, T_k)$ is

$$V^{(k)} = [V_{j_1, j_2, \ldots, j_{n_x}}^{(k)}]_{N_x \times N_x \times \cdots \times N_x} \tag{A.121}$$

per time-step k, so that the computational complexity is

$$CC(N_x, n_x) = O(N_x^{n_x}) = O(\exp(n_x \ln(N_x))), \tag{A.122}$$

which by the law of exponents is exponential in the dimension with an exponent coefficient depending on the logarithm of the common number of nodes N_x, symbolizing the exponential computational complexity of Bellman's curse of dimensionality. This is also the exponential order of the complexity for solving multidimensional PDEs. For the optimal control vector, the order is n_x times this order, but that does not change the exponential order dependency. The deterministic dynamic programming exponential complexity (A.122) should be compared with the deterministic Hamilton's equation formulation in (A.34) with its linear or bilinear complexity $O(3n_x \cdot N_t)$.

Further, for second order finite difference errors, the total error for one state dimension ($n_x = 1$) will be by definition

$$E_T(N_x, 1) = O(N_x^{-2}). \tag{A.123}$$

So even if the order of the complexity is fixed in state dimension $n_x > 1$, i.e., $N = N_x^{n_x}$ is a constant, then $N_x(N) = N^{1/n_x}$ and

$$E_T(N_x(N), n_x) = O\left(N^{-2/n_x}\right) \to O(1) \tag{A.124}$$

as $n_x \to +\infty$ for fixed N and accuracy, i.e., diminishing accuracy in the limit of large dimension.

There are many other computational issues but there is not enough space here to discuss them. Many of these are covered in the author's computational stochastic dynamic programming chapter [109] and more recently in [111].

A.4.4 Linear Quadratic (LQ) Problem by Deterministic Dynamic Programming

The linear quadratic problem is also a good demonstration of the method of dynamic programming as it was as an application of Hamilton's equations and the optimum principle. Using the same formulation, but modified for dynamic programming analysis to start at an arbitrary time t rather than a fixed time t_0, with the dynamics linear in both the control vector $\mathbf{U}(t)$ and the state vector $\mathbf{X}(t)$, the state dynamics is given by

$$\dot{\mathbf{X}}(s) = A(s)\mathbf{X}(s) + B(s)\mathbf{U}(s), \quad t \le s \le t_f, \quad \mathbf{X}(t) = \mathbf{x}. \qquad (A.125)$$

The objective cost functional is given by

$$V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}, t) = \frac{1}{2} \int_t^{t_f} \left[\mathbf{X}^\top(s)Q(s)\mathbf{X}(s) + \mathbf{U}^\top(s)R(s)\mathbf{U}(s) \right] ds \quad (A.126)$$

$$+ \frac{1}{2}\mathbf{X}^\top(t_f)S_f(t_f)\mathbf{X}(t_f).$$

The total minimum cost is again from (A.110)

$$v^*(\mathbf{x}, t) = \min_{\boldsymbol{U}(t, t_f]} \left[V[\mathbf{X}, \mathbf{U}, t_f](\mathbf{x}, t) \right], \qquad (A.127)$$

provided mainly that the quadratic cost matrix $R(t) > 0$, i.e., is positive definite. The HJB equation is

$$0 = v_t^*(\mathbf{x}, t) + \min_{\mathbf{u}} \left[\mathcal{H}(\mathbf{x}, \mathbf{u}, t) \right], \qquad (A.128)$$

where the **pseudo-Hamiltonian** functional simplifies to

$$\mathcal{H}(\mathbf{x}, \mathbf{u}, t) = \frac{1}{2} \left(\mathbf{x}^\top Q(t)\mathbf{x} + \mathbf{u}^\top R(t)\mathbf{u} \right) + \nabla_{\mathbf{x}}^\top [v^*](\mathbf{x}, t) \left(A(t)\mathbf{x} + B(t)\mathbf{u} \right). \quad (A.129)$$

Comparing the dynamic programming pseudo-Hamiltonian (A.119) with the standard Hamiltonian in (A.81) shows that the optimal value gradient $\nabla_{\mathbf{x}}[v^*](\mathbf{x}, t)$ (the marginal value or shadow value in economics) plays the same role as the Lagrange multiplier vector $\boldsymbol{\lambda}$ in (A.81).

Although the decomposition of the optimal value can be rigorously proven, it is sufficient for the purposes here to propose the decomposition is a quadratic form,

$$v^*(\mathbf{x}, t) = \frac{1}{2}\mathbf{x}^\top J(t)\mathbf{x}, \qquad (A.130)$$

and justify it heuristically, i.e., by showing the form (A.130) works. The quadratic coefficient $J(t)$ is a $(n_x \times n_x)$ matrix and since the quadratic form ignores the asymmetric part of the quadratic coefficient, $J(t)$ will be assumed to be symmetric. Thus, the optimal value gradient with respect to the state vector by (B.136) is

$$\nabla_{\mathbf{x}}[v^*](\mathbf{x}, t) = J(t)\mathbf{x}. \tag{A.131}$$

In the case that the cost function is a general quadratic form with linear and zeroth degree terms, then the optimal value LQ decomposition (A.130) will have the same kind of terms.

It is also assumed that there are no constraints on the control to maintain the classical LQ problem form. Thus, stationary points of the pseudo-Hamiltonian are sought,

$$\nabla_{\mathbf{u}}[\mathcal{H}](\mathbf{x}, \mathbf{u}, t) = R(t)\mathbf{u} + B^{\top}(t)J(t)\mathbf{x} = \mathbf{0}, \tag{A.132}$$

using (B.131), (B.136) and the fact that $R(t)$ is symmetric. Thus, the unconstrained optimal control is the linear feedback control

$$\mathbf{u}^*(\mathbf{x}, t) = K(t) \equiv -R^{-1}(t)B^{\top}(t)J(t)\mathbf{x}, \tag{A.133}$$

where the inverse of the quadratic cost coefficient $R(t)$ exists since $R(t)$ is positive definite and where $K(t)$ is the same **gain matrix** as in (A.103) found from the Hamilton's equation formulation. Upon substitution into the HJB equation leads to a pure quadratic form using

$$v_t^*(\mathbf{x}, t) = \mathbf{x}^{\top} J' \mathbf{x} \tag{A.134}$$

and

$$\mathcal{H}^*(\mathbf{x}, t) \equiv \mathcal{H}(\mathbf{x}, \mathbf{u}^*, t)$$

$$= \mathbf{x}^{\top}\left[-\frac{1}{2}J(t)B(t)R^{-1}(t)B^{\top}(t)J(t) + J(t)A(t) + Q(t)\right]\mathbf{x}. \tag{A.135}$$

Taking two partial derivatives or using the Hessian matrix formula (B.137) yields the **matrix Riccati equation**

$$J'(t) = -\frac{1}{2}J(t)B(t)R^{-1}(t)B^{\top}(t)J(t) - J(t)A(t) - A^{\top}(t)J(t) - Q(t) \tag{A.136}$$

subject to the same final condition as well,

$$J(t_f) = S_f(t_f). \tag{A.137}$$

For feedback optimal control and optimal value, the dynamic programming approach is a more direct approach and the algebra is more manageable than the approach through Hamilton's equations. However, the state trajectory is not produced directly by dynamic programming. The more general linear quadratic problem with jump-diffusion processes and other features are treated in Chapter 6.

A.5   Control of PDE Driven Dynamics: Distributed Parameter Systems (DPS)

Thus far, only the control of ODE-driven systems has been considered. However, many dynamical systems are governed by PDEs, such as in fluid and solid dynamics. The PDE dynamics do greatly complicate the optimal control problem and there are many cases to consider. The control of PDE-driven systems usually appears under the heading of *distributed parameter systems* (DPS) and the control is called *distributed control*, while ODE-driven systems are classified as *lumped parameter systems* in contrast. For a more thorough but very applied approach to DPS control, consult Ahmed and Teo's [4] DPS book, Gunzberger's [102] recent monograph on flow control, or the many applications in the DPS research directions proceedings [234]. See also the recent biomedical application to cancer drug delivery to the brain by Chakrabarty and Hanson [49] (briefly summarized in the biomedical application Section 11.2.2). Only one fairly general deterministic model will be presented here since the focus is on stochastic problems.

A.5.1   DPS Optimal Control Problem

Let $\mathbf{y}(\mathbf{x}, t)$ be a n_y-vector state variable in space-time, where $\mathbf{x}$ is the n_x-dimensional space vector. The state dynamics for $\mathbf{y}(\mathbf{x}, t)$ satisfy a nonlinear reaction diffusion equation with drift,

$$\frac{\partial \mathbf{y}}{\partial t}(\mathbf{x}, t) = D\nabla_x^2[\mathbf{y}](\mathbf{x}, t) + C\nabla_x[\mathbf{y}](\mathbf{x}, t) + \mathbf{B}(\mathbf{y}(\mathbf{x}, t), \mathbf{x}, t) + A\mathbf{u}(\mathbf{x}, t), \quad (A.138)$$

$\mathbf{x} \in \mathcal{D}_x$ and $t_0 < t \le t_f$, with initial condition

$$\mathbf{y}(\mathbf{x}, t_0) = \mathbf{y}_0(\mathbf{x})$$

and mixed boundary condition

$$(\alpha(\widehat{\mathbf{n}}^\top \nabla_x)[\mathbf{y}] + \beta\mathbf{y} + \boldsymbol{\gamma})(\mathbf{x}, t) = \mathbf{0}$$

for $\mathbf{x}$ on the space domain boundary $\partial\mathcal{D}_x$ while $\widehat{\mathbf{n}}(\mathbf{x}, t)$ is the outward normal to the boundary. Here $\mathbf{u}(\mathbf{x}, t)$ is the n_u-dimensional space-time control variable in a linear control-dependent term. All coefficient functions are assumed to be bounded while being commensurate in multiplication and sufficiently differentiable as needed. In particular, the diffusion tensor $D = [D_i \delta_{i,j}]_{n_y \times n_y}$ is a positive-definite diagonal matrix and the drift coefficient $C = [C_{i,k} \delta_{i,j}]_{n_y \times n_y \times n_x}$. The main *reaction* vector $\mathbf{B}(\mathbf{y}(\mathbf{x}, t), \mathbf{x}, t)$ is the only term assumed to be nonlinear since reaction terms are often naturally nonlinear. The control coefficient is $A = [A_{i,j}]_{n_y \times n_u}$ and is assumed to be constant but could depend on $(\mathbf{x}, t)$, as could C and D.

Further, let the space-time objective be in the form of the total quadratic costs,

$$V[\mathbf{y}, \mathbf{u}, t_f] = \frac{1}{2} \int_{t_0}^{t_f} \!\!dt \!\! \int_{\mathcal{D}_x} \!\!d\mathbf{x} \big(\mathbf{y}^\top Q\mathbf{y} + \mathbf{u}^\top R\mathbf{u}\big)(\mathbf{x}, t) + \frac{1}{2} \int_{\mathcal{D}_x} \!\!d\mathbf{x}\, \big(\mathbf{y}^\top S\mathbf{y}\big)(\mathbf{x}, t_f), \quad (A.139)$$

where the quadratic control coefficient R is symmetric positive-definite, while Q and S are symmetric positive-semidefinite to ensure a minimum. Equations (A.138)–(A.139) provide the underlying formulation of the DPS optimal control problem.

A.5.2 DPS Hamiltonian Extended Space Formulation

For the formulation of the equations for the optimal solutions to this control problem, the dynamic and initial-boundary constraints need to be combined into a pseudo-Hamiltonian,

$$
\begin{aligned}
\mathcal{H}(\mathbf{y}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}) = {} & V[\mathbf{y}, \mathbf{u}, t_f] \\
& + \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x}\, \boldsymbol{\lambda}^\top \big(\mathbf{y}_t - D\nabla_x^2[\mathbf{y}] - C\nabla_x[\mathbf{y}] - \mathbf{B} - A\mathbf{u}\big)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f} dt \int_{\partial\mathcal{D}_x} d\boldsymbol{\Gamma}\, \boldsymbol{\mu}^\top \big(\alpha(\widehat{\mathbf{n}}^\top \nabla_x)[\mathbf{y}] + \beta\mathbf{y} + \boldsymbol{\gamma}\big)(\mathbf{x}, t) \qquad \text{(A.140)} \\
& + \int_{\mathcal{D}_x} d\mathbf{x}\, \boldsymbol{\nu}^\top \big(\mathbf{y}(\mathbf{x}, t_0^+) - \mathbf{y}_0(\mathbf{x})\big),
\end{aligned}
$$

where $\{\boldsymbol{\lambda}(\mathbf{x}, t), \boldsymbol{\mu}(\mathbf{x}, t), \boldsymbol{\nu}(\mathbf{x})\}$ is a set of Lagrange multipliers that provide the mechanism for including the control problem constraints at the expense of extending the state-control space to higher dimension with

$$
\mathbf{z}(\mathbf{x}, t) \equiv \{\mathbf{y}(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t), \boldsymbol{\lambda}(\mathbf{x}, t), \boldsymbol{\mu}(\mathbf{x}, t), \boldsymbol{\nu}(\mathbf{x})\}
$$

denoting the extended space-control vector. Next, assuming an optimal extended state $\mathbf{z}(\mathbf{x}, t) = \mathbf{z}^*(\mathbf{x}, t)$ exists under sufficient differentiability properties of $\mathcal{H}(\mathbf{z})$, perturb about this optimal extended state as $\mathbf{z}(\mathbf{x}, t) = \mathbf{z}^*(\mathbf{x}, t) + \delta\mathbf{z}(\mathbf{x}, t)$, where $\delta\mathbf{z}(\mathbf{x}, t)$ is the *variation*, and then expand the pseudo-Hamiltonian about this variation,

$$
\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t) + \delta\mathbf{z}(\mathbf{x}, t)) = \mathcal{H}(\mathbf{z}^*(\mathbf{x}, t)) + \delta\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t), \delta\mathbf{z}(\mathbf{x}, t)) + \mathrm{O}(|\delta\mathbf{z}|^2(\mathbf{x}, t)).
$$

Neglecting quadratic order perturbation terms, including the second variation of $\mathcal{H}(\mathbf{z})$, then the first variation $\delta\mathcal{H}(\mathbf{z}^*(\mathbf{x}, t), \delta\mathbf{z}(\mathbf{x}, t))$ is found to be a linear function of the extended state perturbation $\delta\mathbf{z}(\mathbf{x}, t)$ using (A.139)–(A.140). For this perturbation, the nonlinear reaction term $\mathbf{B}(\mathbf{y}(\mathbf{x}, t), \mathbf{x}, t)$ is assumed to be more than once differentiable so that

$$
\mathbf{B}(\mathbf{y}^* + \delta\mathbf{y}, \mathbf{x}, t) = \mathbf{B}(\mathbf{y}^*, \mathbf{x}, t) + (\delta\mathbf{y}^\top \nabla_y)[\mathbf{B}](\mathbf{y}^*, \mathbf{x}, t) + \mathrm{O}(|\delta\mathbf{y}|^2),
$$

for example, twice differentiable to guarantee the quadratic order error term. For simplicity, let $B^* \equiv \mathbf{B}(\mathbf{y}^*, \mathbf{x}, t)$. Applying multiple Taylor approximations,

$$
\begin{aligned}
\delta\mathcal{H}(\mathbf{z}^*, \delta\mathbf{z}) = & \int_{t_0}^{t_f}\!\!dt\!\int_{\mathcal{D}_x}\!\!d\mathbf{x}\left((\mathbf{y}^*)^\top Q\delta\mathbf{y} + (\mathbf{u}^*)^\top R\delta\mathbf{u}\right)(\mathbf{x}, t) + \int_{\mathcal{D}_x}\!\!d\mathbf{x}\left((\mathbf{y}^*)^\top S\delta\mathbf{y}\right)(\mathbf{x}, t_f) \\
& + \int_{t_0}^{t_f}\!\!dt\!\int_{\mathcal{D}_x}\!\!d\mathbf{x}\,(\boldsymbol{\lambda}^*)^\top\!\left(\delta\mathbf{y}_t - D\nabla_x^2[\delta\mathbf{y}] - C\nabla_x[\delta\mathbf{y}] - (\delta\mathbf{y}^\top\nabla_y)[\mathbf{B}]^* - A\delta\mathbf{u}\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f}\!\!dt\!\int_{\mathcal{D}_x}\!\!d\mathbf{x}\,d\boldsymbol{\lambda}^\top\!\left(\mathbf{y}_t^* - D\nabla_x^2[\mathbf{y}^*] - C\nabla_x[\mathbf{y}^*] - \mathbf{B}^* - A\mathbf{u}^*\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f}\!\!dt\!\int_{\partial\mathcal{D}_x}\!\!d\boldsymbol{\Gamma}(\boldsymbol{\mu}^*)^\top\!\left(\alpha(\widehat{\mathbf{n}}^\top\nabla_x)[\delta\mathbf{y}] + \beta\delta\mathbf{y}\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f}\!\!dt\!\int_{\partial\mathcal{D}_x}\!\!d\boldsymbol{\Gamma}(\delta\boldsymbol{\mu})^\top\!\left(\alpha(\widehat{\mathbf{n}}^\top\nabla_x)[\mathbf{y}^*] + \beta\mathbf{y}^* + \boldsymbol{\gamma}\right)(\mathbf{x}, t) \\
& + \int_{\mathcal{D}_x}\!\!d\mathbf{x}\left((\boldsymbol{\nu}^*)^\top\delta\mathbf{y}(\mathbf{x}, t_0^+) + \delta\boldsymbol{\nu}^\top\!\left(\mathbf{y}^*(\mathbf{x}, t_0^+) - \mathbf{y}_0(\mathbf{x})\right)\right).
\end{aligned}
\tag{A.141}
$$

Obtaining the critical or optimal conditions requires the reduction of the highest order partial derivative terms, since under integration the perturbations $\delta\mathbf{y}_t(\mathbf{x}, t)$ and $\nabla_x^2[\delta\mathbf{y}]$ are not independent of lower order derivatives and the higher order derivatives can be reduced by integration by parts techniques to lower order derivatives. Thus, using integration by parts

$$
\int_{t_0}^{t_f}\!\!dt\,(\boldsymbol{\lambda}^*)^\top\delta\mathbf{y}_t(\mathbf{x}, t) = (\boldsymbol{\lambda}^*)^\top\delta\mathbf{y}\Big|_0^{t_f} - \int_{t_0}^{t_f}\!\!dt\,\delta\mathbf{y}^\top\boldsymbol{\lambda}_t^*,
$$

$$
-\int_{\mathcal{D}_x}\!\!d\mathbf{x}\,(\boldsymbol{\lambda}^*)^\top C\nabla_x[\delta\mathbf{y}] = -\int_{\partial\mathcal{D}_x}\!\!d\boldsymbol{\Gamma}\,\delta\mathbf{y}^\top\widehat{\mathbf{n}}^\top C^\top\boldsymbol{\lambda}^* + \int_{\mathcal{D}_x}\!\!d\mathbf{x}\,\delta\mathbf{y}^\top\nabla_x^\top[C^\top\boldsymbol{\lambda}^*],
$$

where $C^\top \equiv [C_{k,i}\delta_{k,j}]_{n_x \times n_y \times n_y}$ defines the transpose of a three subscript array, and finally using a double integration by parts [103],

$$
\begin{aligned}
-\int_{\mathcal{D}_x}\!\!d\mathbf{x}\,(\boldsymbol{\lambda}^*)^\top D\nabla_x^2[\delta\mathbf{y}] = & -\int_{\partial\mathcal{D}_x}\!\!d\boldsymbol{\Gamma}\left((\widehat{\mathbf{n}}^\top\nabla_x)[\delta\mathbf{y}^\top]D\boldsymbol{\lambda}^* - \delta\mathbf{y}^\top(\widehat{\mathbf{n}}^\top\nabla_x)[D\boldsymbol{\lambda}^*]\right) \\
& -\int_{\mathcal{D}_x}\!\!d\mathbf{x}\,\delta\mathbf{y}^\top\nabla_x^2[D\boldsymbol{\lambda}^*].
\end{aligned}
$$

Using these reduced forms in (A.141) and collecting terms as coefficients of like extended state perturbations produces a more useful form:

$$
\begin{aligned}
\delta\mathcal{H}(\mathbf{z}^*, \delta\mathbf{z}) = & \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x}\, (\delta\mathbf{y})^\top \left(-\boldsymbol{\lambda}_t^* - \nabla_x^2[D\boldsymbol{\lambda}^*] + \nabla_x^\top[C^\top\boldsymbol{\lambda}^*] + \nabla_y[\mathbf{B}^\top]^*\boldsymbol{\lambda}^* + Q\mathbf{y}^*\right) \\
& + \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x}\, (\delta\mathbf{u})^\top \left(R\mathbf{u}^* - A^\top\boldsymbol{\lambda}^*\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f} dt \int_{\mathcal{D}_x} d\mathbf{x}\, (\delta\boldsymbol{\lambda})^\top \left(\mathbf{y}_t^* - D\nabla_x^2[\mathbf{y}^*] - C\nabla_x[\mathbf{y}^*] - \mathbf{B}^* - A\mathbf{u}^*\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f} dt \int_{\partial\mathcal{D}_x} d\boldsymbol{\Gamma}(\widehat{\mathbf{n}}^\top\nabla_x)[\delta\mathbf{y}^\top] \left(\alpha^\top\boldsymbol{\mu}^* - D\boldsymbol{\lambda}^*\right)(\mathbf{x}, t) \qquad \text{(A.142)} \\
& + \int_{t_0}^{t_f} dt \int_{\partial\mathcal{D}_x} d\boldsymbol{\Gamma}(\delta\mathbf{y})^\top \left((\widehat{\mathbf{n}}^\top\nabla_x)[D\boldsymbol{\lambda}^*] - \widehat{\mathbf{n}}^\top C^\top\boldsymbol{\lambda}^* + \beta^\top\boldsymbol{\mu}^*\right)(\mathbf{x}, t) \\
& + \int_{t_0}^{t_f} dt \int_{\partial\mathcal{D}_x} d\boldsymbol{\Gamma}(\delta\boldsymbol{\mu})^\top \left(\alpha(\widehat{\mathbf{n}}^\top\nabla_x)[\mathbf{y}^*] + \beta\mathbf{y}^* + \boldsymbol{\gamma}\right)(\mathbf{x}, t) \\
& + \int_{\mathcal{D}_x} d\mathbf{x}\, \delta\mathbf{y}^\top (S\mathbf{y}^* + \boldsymbol{\lambda}^*)(\mathbf{x}, t_f) \\
& + \int_{\mathcal{D}_x} d\mathbf{x}\, \delta\mathbf{y}^\top (\boldsymbol{\nu}^* - \boldsymbol{\lambda}^*)(\mathbf{x}, t_0) + \int_{\mathcal{D}_x} d\mathbf{x}\, \delta\boldsymbol{\nu}^\top \left(\mathbf{y}^*(\mathbf{x}, t_0^+) - \mathbf{y}_0(\mathbf{x})\right).
\end{aligned}
$$

A.5.3 DPS Optimal State, Costate and Control PDEs

Our interest here is to present a usable formulation for those whose prime interest is obtaining concrete solutions for applications, so our approach is a formal applied mathematical one. If the interest of the reader is in existence and uniqueness properties of the solution rather than the solution itself, the reader should explore the [4, 102, 234], and the references therein for abstract notions of Hilbert spaces with related Sobolev spaces and functional derivatives. However, such abstract approaches have little utility in solving real problems.

The optimal state, costate and control trajectory dynamics follow from setting to zero the coefficients of each of the independent state, costate and control first variations of the pseudo-Hamiltonian in (A.142), as well as any relevant boundary, initial and final values which are assumed to be independent of the space-time interior values.

The optimal state equation for $\mathbf{y}^*(\mathbf{x}, t)$ follows from setting the critical coefficient of the costate variation $\delta\boldsymbol{\lambda}(\mathbf{x}, t)$ on each interior point of $\mathcal{D}_x \times (t_0, t_f)$ yielding

$$
\mathbf{y}_t^*(\mathbf{x}, t) = \left(D\nabla_x^2[\mathbf{y}^*] + C\nabla_x[\mathbf{y}^*] + \mathbf{B}^* + A\mathbf{u}^*\right)(\mathbf{x}, t) \qquad \text{(A.143)}
$$

for $\mathbf{x} \in \mathcal{D}_x$ and $t_0 < t \le t_f$, while the intial condition $\mathbf{y}^*(\mathbf{x}, t_0^+) = \mathbf{y}_0(\mathbf{x})$ follows from setting the coefficient of the initial condition costate variation $\delta\boldsymbol{\nu}(\mathbf{x})$ to zero and the boundary condition

$$
\left(\alpha(\widehat{\mathbf{n}}^\top\nabla_x)[\mathbf{y}^*] + \beta\mathbf{y}^* + \boldsymbol{\gamma}\right)(\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \partial\mathcal{D}_x, \quad t_0 < t < t_f,
$$

follows from setting the coefficient of the final condition costate variation $\delta\boldsymbol{\nu}(\mathbf{x})$ to zero. The optimal state equation (A.143), of course, has the same form as that of the original state equation (A.138), which is a forward parabolic PDE for $D > 0$.

The optimal costate equation for $\boldsymbol{\lambda}^*$ is derived by setting to zero the state variation $\delta\mathbf{y}(\mathbf{x}, t)$ coefficient to zero, so that

$$\left(\boldsymbol{\lambda}_t^* + \nabla_x^2[D\boldsymbol{\lambda}^*] - \nabla_x^\top[C^\top\boldsymbol{\lambda}^*] - \nabla_y[\mathbf{B}^\top]^*\boldsymbol{\lambda}^* - Q\mathbf{y}^*\right)(\mathbf{x}, t) = \mathbf{0} \quad (A.144)$$

for $\mathbf{x} \in \mathcal{D}_x$ and $t_0 \leq t < t_f$, noting that (A.144) is a backward parabolic PDE since the diffusion term has an opposite sign to that of the forward equation (A.143). So a final condition is needed by setting the coefficient of $\delta\mathbf{y}(\mathbf{x}, t_f)$ to zero, i.e.,

$$\left(\boldsymbol{\lambda}^* + S\mathbf{y}^*\right)(\mathbf{x}, t_f) = \mathbf{0}, \quad (A.145)$$

coupling the computed final condition of $\boldsymbol{\lambda}^*$ to the computed final value of $(-S\mathbf{y}^*)$. The boundary conditions follow from setting the coefficient of $\delta\mathbf{y}$ on the boundary to zero, so

$$\left((\widehat{\mathbf{n}}^\top\nabla_x)[D\boldsymbol{\lambda}^*] - \widehat{\mathbf{n}}^\top C^\top\boldsymbol{\lambda}^* + \beta^\top\boldsymbol{\mu}^*\right)(\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \partial\mathcal{D}_x, \quad (A.146)$$

giving rise to another complication, in that the boundary condition costate $\boldsymbol{\mu}^*(\mathbf{x}, t)$ appears. However, the coefficient of the normal gradient $(\widehat{\mathbf{n}}^\top\nabla_x)[\delta\mathbf{y}^\top]$ yields

$$\left(\alpha^\top\boldsymbol{\mu}^* - D\boldsymbol{\lambda}^*\right)(\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \partial\mathcal{D}_x,$$

which, if $\alpha^\top$ is invertible, can be used to eliminate $\boldsymbol{\mu}^*$ on the boundary. Another costate condition comes from the initial value of $\delta\mathbf{y}$ which gives

$$\boldsymbol{\nu}^*(\mathbf{x}) = \boldsymbol{\lambda}^*(\mathbf{x}, t_0),$$

where $\boldsymbol{\lambda}^*(\mathbf{x}, t_0)$ is the terminal output of the backward integration of the prime optimal costate PDE (A.144) starting from the final condition (A.145).

From the coefficient of the control variation $\delta\mathbf{u}(\mathbf{x}, t)$, the optimal control is given by

$$\left(R\mathbf{u}^* - A^\top\boldsymbol{\lambda}^*\right)(\mathbf{x}, t) = \mathbf{0}, \quad \mathbf{x} \in \mathcal{D}_x, \quad t_0 \leq t < t_f,$$

and since $R(\mathbf{x}, t)$ should be invertible due to its positive-definite property, then

$$\mathbf{u}^*(\mathbf{x}, t) = \left(R^{-1}A^\top\boldsymbol{\lambda}^*\right)(\mathbf{x}, t) \quad (A.147)$$

in absence of control constraints, else it is merely the regular optimal control $\mathbf{u}^{(\mathrm{reg})}(\mathbf{x}, t)$.

A numerical scheme developed in Chakrabarty and Hanson [49] for a biomedical application uses a forward state integration of (A.143) and backward costate integration of (A.144) with subsequent iterations until the norm of the iteration difference is sufficiently small. The forward integration step for (A.143) requires a good starting guess for the optimal control space-time distribution in addition to using the specified state initial condition. The final time approximation to $\mathbf{y}^*(\mathbf{x}, t)$ is then used as the final condition to start the costate $\boldsymbol{\lambda}^*(\mathbf{x}, t)$ backward integration of (A.144). The end approximation of the costate space-time distribution of $\boldsymbol{\lambda}^*(\mathbf{x}, t)$ is used by (A.147) to update the optimal control distribution approximation $\mathbf{u}^*(\mathbf{x}, t)$, which in turn is used in the next state forward integration.

A.6 Exercises

1. For the deterministic linear first order dynamics,

$$\dot{X}(t) = -\mu_0 X(t) + \beta_0 U(t), t > 0, \quad \text{given } X(0) = x_0 \neq 0, \quad \mu_0 > 0, \quad \beta_0 \neq 0,$$

and quadratic performance measure,

$$V[U] = \frac{r_0}{2} \int_0^{t_f} U^2(t)dt, \quad r_0 > 0,$$

find the optimal state trajectory and optimal (unconstrained) control to bring the state from the initial state to the final state x_f in t_f seconds while minimizing the functional $V[U]$ with respect to the control u, with the answer depending on the parameter set $\{x_0, x_f, t_f, \mu_0, \beta_0, r_0\}$. *Note that the final state and time are fixed.*

2. Consider another simple lumped model of a **leaky reservoir** (after Kirk [164]) given by

$$\dot{X}(t) = -aX(t) + U(t), \quad X(0) = x_0,$$

where $X(t)$ is the depth of the reservoir, $U(t)$ is the net flow of water per unit time into the reservoir at time t and $a > 0$ is the rate of leakage and usage. The net inflow is constrained pointwise $0 \leq U(t) \leq M$ for all $0 < t \leq t_f$ and also cumulatively by

$$\int_0^{t_f} U(t)dt = K > 0,$$

where K, M and t_f are fixed constants, such that $K \leq M \cdot t_f$ for consistency. Find the optimal control law $U^*(t)$ that maximizes only the final depth,

$$J[X] = bX(t_f)$$

with $b > 0$, the optimal state $X^*(t)$, optimal final depth $J[X^*]$ and the optimal Hamiltonian $\mathcal{H}^*$.

3. **Pontryagin's auxiliary necessary conditions for the Hamiltonian in the special case of no explicit dependence on time:** Assume sufficient differentiability for the Hamiltonian and that

$$\mathcal{H}^* = \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t))$$

so $\partial \mathcal{H}^*/\partial t \equiv 0$. Then show the following

(a) If the final time t_f is *fixed* and the Hamiltonian $\mathcal{H}$ does not depend explicitly on time, then the Hamiltonian must be constant when evaluated on a locally (interior) extreme trajectory, i.e.,

$$\mathcal{H}^* = \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t)) = c, \tag{A.148}$$

where c is a constant.

Explain why fixed t_f and local extremes are needed; also, explain why Example A.8 and Exercise 2 are counter-examples for the result A.148 if a certain condition is not satisfied, stating what that condition is.

(b) If the final time t_f is *free* and both the Hamiltonian $\mathcal{H}$ and the final cost function $S = S(x)$ do not depend explicitly on time, then the Hamiltonian must be zero when evaluated on an locally (interior) extreme trajectory, i.e.,

$$\mathcal{H}^* = \mathcal{H}(\mathbf{X}^*(t), \mathbf{U}^*(t), \boldsymbol{\lambda}^*(t)) = 0. \qquad (A.149)$$

4. Solve the deterministic optimal control problem with wealth state dynamics

$$dX(t) = (\mu_0 - U(t))X(t)dt$$

for $0 \le t \le t_f$, $X(0) = x_0 > 0$, μ_0 is a constant mean rate, and the wealth consumption is unconstrained $-\infty < U(t) < +\infty$. The objective is maximum cumulative utility, the running consumption is the risk-adverse utility $C(x, u, t) = 2\sqrt{u}$, and similarly the utility of final wealth is $S(x, t) = 2\sqrt{x}$.

(a) Formulate the Hamiltonian $\mathcal{H}(x, u, \lambda, t)$ and the associated Hamilton's equations.

(b) Show that the optimal Hamiltonian $\mathcal{H}^*$ is a maximum at the regular point $(X^*(t), U^*(t), \lambda^*(t), t)$, where $\lambda^*(t)$ is the optimal costate.

(c) Show that optimal trajectories satisfy $\lambda^*(t)X^*(t) = K$ and $U^*(t) = 1/K^2 \equiv K_0$, where K is a constant.

(d) Show that K_0 is satisfied by the nonlinear equation $K_0 x_0 \exp((\mu_0 - K_0)t_f) = 1$.
{*Hint: the transversality condition*

$$\lambda^*(t_f) = (\partial S/\partial x)(X^*(t_f), t_f)$$

since $X^*(t_f)$ is free and t_f is fixed.}

5. Find the maximum discounted net profit with objective function

$$C(x, u, t) = e^{-\delta_0 t}[p_0 X(t) - c_0]u, \quad S(x, t) = \sigma_0 x,$$

subject to the linear control-state dynamics,

$$\dot{X}(t) = \mu_0 X(t) - U(t), \quad X(t_0) = x_0, \quad t_0 \le t \le t_f,$$

where δ_0, p_0, c_0, $\sigma_0 < 1$, μ_0, and x_0 are fixed, positive constant parameters. Assume that $X(t) < c_0/p_0$. Find intervals in parameter space where there is a maximal control solution.

Discuss the difference between the solution to this problem and the solution to a similar problem in Example A.9 with a bilinear control-state term $U(t)X(t)$ rather than just linear in the control $U(t)$.

6. For the regular control demonstration in Example A.3 with dynamics (A.17), utility of instantaneous consumption (A.19), and terminal wealth (A.20), but with the utility power $\gamma > 1$ (for example, $\gamma = 2$), solve the Bolza problem for the proper maximum utility objective by using bang control with the bounded control constraints (A.18). Recall that the regular control solution yields a minimum rather than a maximum solution.

7. For the singular control demonstration in Example A.9 with state dynamics (A.52), cost function $C(x, u, t)$ (A.53), and control constraints (A.54), analyze the case when the maximum control $U^{(\max)}$ exceeds the mean rate μ_0, i.e., $U^{(\max)} > \mu_0$. When the parameter values permit a control solution, then find the solution; otherwise list the parameter ranges in which there fails to be a control solution.

8. Find the minimal control $U^*(t)$ for the optimal performance

$$v^*(x_1, x_2, t) = \max_U \left[V[X_1, X_2, U](x_1, x_2, t) \right]$$

of the measure

$$V[X_1, X_2, U](x_1, x_2, t) = \frac{1}{2} \int_t^{t_f} (q_1 X_1^2(s) + q_2 X_2^2(s) + rU^2(s))ds,$$

$q_1 > 0$, $q_2 > 0$, $r > 0$, subject to the dynamics

$$\dot{X}_1(t) = a_{1,1}X_1(t) + a_{1,2}X_2(t) + s_1U(t), \ \ s_1 > 0,$$

$$\dot{X}_2(t) = a_{2,1}X_1(t) + a_{2,2}X_2(t) + s_2U(t), \ \ s_2 > 0$$

and the control constraints

$$|U(t)| \le K, \ \ K > 0,$$

formally solving for $U^*(t)$ in terms of t, (x_1, x_2), first order partial derivatives $v_{x_1}^*(x_1, x_2, t)$ and $v_{x_2}^*(x_1, x_2, t)$ using dynamic programming. Do not solve the partial differential equation of dynamic programming but only substitute the composite formulae for $U^*(t)$ into it.

What changes in the solution form if the optimum is specified as a maximum rather than a minimum?

Suggested References for Further Reading

- Ahmed and Teo, 1981 [4]

- Anderson and Moore, 1990 [8]

- Athans and Falb 1966 [15]

- Bell and Jacobson 1975 [24]

- Betts, 2001 [30]

- Bliss, 1946 [40]

- Bryson and Ho, 1975 [44]

- Chakrabarty and Hanson, 2005 [49]

- Clark, 1976 [57]

- Gunzberger, 2003 [102]

- Kirk, 1970 [164]

- Nise, 2000 [220]

- Pontryagin et al., 1962 [226]

- *Research Directions in Distributed Parameter Systems*, 2003 [234]

- Stengel, 1986 [258]