Chapter 2

CONTROL OF BILINEAR SYSTEMS

Bilinear systems are an important subclass of nonlinear systems with numerous applications in engineering, biology, and economics (Mohler, 1973; Espana and Landau, 1978; Brockett, 1979; Baillieul, 1978, 1998). Most papers consider time-invariant continuous bilinear systems with linear feedback. In almost all references stability problems are studied using a sufficient condition for the existence of a feedback control. These studies usually do not deal with stability problems with a large class of inputs as considered here. The linear-quadratic optimal control problem is apparently one of the simplest and most thoroughly studied. It is not surprising, therefore, that the nonlinear analogues of this problem have long since attracted the attention of control scientists (Andreev, 1982; Aganovic, 1995; Bloch and Crouch, 1996; Bloch, 1998; Jurdjevic, 1998; Agrachev and Sachkov, 2004). The symplectic structure of the linear-quadratic problem was also studied in Faibusovich (1988a, 1991), and has naturally led to the question of nonlinear generalizations.

In this chapter we consider optimal control problems for nonlinear systems. In Section 1 we demonstrate how to reduce this problem to the equivalent problem for bilinear systems. Then we investigate an optimal control problem for nonlinear systems through necessary conditions of optimality. We apply soliton methods and Pontryagin's principle to the optimal control problem. We give theoretical justification for the equations in the Lax form, which can be constructed from the Hamiltonian system arising through the expression of the necessary optimality conditions. The Lax forms we consider, extend and unify similar results obtained previously in the publications on optimal control of bilinear systems. The theoretical justification was done by V. Yatsenko (1984),

P.M. Pardalos, V. Yatsenko, *Optimization and Control of Bilinear Systems*, doi: 10.1007/978-0-387-73669-3,

who applied the soliton method to optimal control problems. However, the approach is general and the specific Lie algebraic structure of bilinear systems is overlooked. This explains why here we obtain two equations in the Lax form instead of one. Moreover, we show that it is not difficult to extend the approach to nonlinear systems. Note that the proposed procedure is global and, in particular, it is applicable to controllable systems on manifolds. This degree of generality is dictated by potential applications (mechanical systems, robotics, multiagent systems, nonlinear circuits, control of quantum-mechanical processes) and by the fact that a number of geometrical problems that have recently been attracting considerable attention fit into the proposed framework. We also examine the selection of a performance criterion and present a number of examples. Sections 2 and 3 deal with stabilization problems for controlled physical objects. A variety of physical and biological systems are well modeled by coupled bilinear equations. In most cases such systems are capable of displaying several types of dynamical behavior: limit cycles, bistability, excitation, or chaos.

1. Optimal Control of Bilinear Systems

1.1 Optimal Control Problem

In this section, we consider the following bilinear control systems:

$$\begin{gathered} \dot{x}(t) = f_0(x) + F(x)u(t), \qquad u(t) \in \Omega, \\ z(t) = h(x) + Q(x)v(t), \qquad v(t) \in \Gamma, \\ z(t) \in \mathbb{R}^r, \qquad x(t) \in \mathbb{R}^n, \end{gathered} \tag{2.1}$$

where x is a state vector; z is an output signal; $u(t)$ and $v(t)$ are control functions of the time of dimensions m, q, respectively; $f_0(x)$ and $h(x)$ are given vector-valued functions of the $x(t)$ of dimensions n, p, respectively; and F and Q are matrix-valued functions of $x(t)$ of appropriate sizes.

At the initial time $t = t_0$, the initial condition for the system (2.1) is specified,

$$x(t_0) = x_0, \tag{2.2}$$

where x_0 is a given n-dimensional vector.

We consider the following objective function,

$$\eta = \int_0^T \sum_{i,j=1}^{m} q_{ij} u_i u_j dt, \tag{2.3}$$

where $\tilde{Q} = (q_{ij})$ is a symmetric positively determined matrix.

We seek the control $u(t)$ that yields the minimal value to the function (2.3)

$$\min \eta = \min \int_0^T \sum_{i,j=1}^m q_{ij} u_i u_j dt, \tag{2.4}$$

provided that conditions (2.1) and (2.3) hold.

1.2 Reduction of Control Problem to Equivalent Problem for Bilinear Systems

In this section we consider reduction of the nonlinear system (2.1) to a dynamically equivalent form, which allows us to simplify the solution of the corresponding optimal control problem. This simplification is achieved by the analytical solution of the optimal control problem for a dynamically equivalent bilinear system. It is shown here that the optimal control satisfies the Euler–Lagrange equation, the solutions of which are expressed analytically through the Θ-functions of Riemann surfaces. In addition, bilinear systems allow the use of known mathematical system theory results in order to investigate the system features of controlled processes (Jurdjevic, 1997; Aganović and Gajic, 1995).

Bilinearization of a nonlinear system. Write the first equation of system (2.1) as

$$\dot{x}(t) = f_0(x) + \sum_{i=1}^m u_i(t) f_i(x), \qquad x(0) = x_0, \tag{2.5}$$

where f_i is the ith column of a matrix F. Instead of considering (2.5), we study the following bilinear system,

$$\dot{y}(t) = \left(A_0 + \sum_{i=1}^m u_i(t) A_i \right) y, \qquad y(0) = y_0, \tag{2.6}$$

where $y = (y_1, \ldots, y_p)$ is a state vector; $A_0, \ldots, A_m$ are constant $p \times p$ matrices; $u(t) = (u_1(t), \ldots, u_m(t))$ is a restricted measurable control. To construct the system (2.6) the approach described in Krener (1973) is used.

Let $P^0(\tilde{A})$ be the set of functions $\{f_i : i = 0, \ldots, m\}$; $P^j(A) = P^{j-1}(\tilde{A}) \cup \{[f_i, c] : i = 0, \ldots, m,\ c \in P^{j-1}(\tilde{A})\}$ for $j \leq 1$. The completed system of $\tilde{A}$ is $P(\tilde{A}) = \bigcup_{j>0} P^j(\tilde{A})$, and we define $P(\tilde{A})_x = \{c(x) : c \in P(\tilde{A}) \subseteq \mathbb{R}^m\}$, ($[\cdot,\cdot]$ denotes the Riemann surface Lie bracket of vector fields); and $P(\tilde{B})$ is a corresponding function set for system (2.5). Consider the systems (2.5) and (2.6). Let M and N be submanifolds that carry (2.5) and (2.6) at x_0 and y_0. Then there exists such a linear map l: $\operatorname{Span} P(\tilde{A})_{x_0} \to \operatorname{Span} P(\tilde{B})_{y_0}$, such that $l(f_i(x_0)) = A_i\ (y_0)$ for $i = 0, \ldots, m, l([f_{i_1} \ldots, [f_{i_{h-1}}, f_{i_h}] \ldots](x_0)) = [A_{i_1}, \ldots, [A_{i_{h-1}}, A_{i_h}] \ldots](y_0)$ for $h \leq 2$, $1 \leq i_j \leq m$ if and only if there exist neighborhoods U and V and an analytical map $\lambda : U \to V$ such that λ carries (2.5) and (2.6) for the same control $u(t)$ and $x(t) \in U$ for $|t| < \varepsilon$. Furthermore l is a linear diffeomorphism if and only if λ is a local diffeomorphism.

Instead of considering (2.5), we study the following matrix bilinear system (Krener, 1975),

$$\dot{Y}(t) = \left(A_0 + \sum_{i=1}^{m} u_i(t) A_i\right) Y(t), \quad Y(0) = I, \quad u(t) \in \Omega, \tag{2.7}$$

where $Y(t) \in Gl(p, \mathbb{R})$; $Gl(p, \mathbb{R})$ is a group of reverse $p \times p$-matrices.

Use the minimal algebra $W(M)$, containing a totality of vector fields $\{f_0, \ldots, f_k\}$, to construct the system (2.7). The Lie algebra $W(M)$ is infinitely dimensional in general, therefore instead of the algebra $W(M)$, build up the finite-dimensional algebra $\hat{\Theta}$ using the ν-order Lie bracket:

$$[f_{i_1} \cdots [f_{i_{\nu-1}}, f_{i_\nu}] \cdots], f_0, \ldots, f_h \in \hat{\Theta}. \tag{2.8}$$

In this case, ν is stated by the accuracy of approximation of the solutions to system (2.5) by the solutions to system (2.6): the greater ν is, the higher an approximation accuracy is.

If $f_0, \ldots, f_h$ are the linearly independent elements of the algebra $\hat{\Theta}$ over the field $\mathbb{R}$, then, when the switching operation is used (up to the order ν), it is possible to obtain new elements of the algebra $\hat{\Theta}$. Assuming every Lie bracket of an order higher than ν, equal to zero, the finite-dimensional Lie algebra is derived. Let its dimensionality be equal to s. According to the Ado theorem (Ado, 1947), this algebra is isomorphic to the subalgebra $\hat{g}$ of the algebra R. Bring the $s \times s$-dimensional matrix A_i into correspondence with each f_i, and the ν-order bilinear system is obtained. The constructive building procedure, according to which the matrix bilinear system (2.7) is made up, uses the adjacent representation of the algebra $\hat{\Theta}$. If $d_1, \ldots, d_s$ is the basis of $\hat{\Theta}$, then, for each $\hat{\Theta}$, assume $ad(c)$ equal to the matrix $B = [B_{ij}]$, defined by the relation

$$[c, d_i] = \sum_{i=1}^{m} B_{ij} d_i. \tag{2.9}$$

The adjoint representation of the $(\nu + 1)$-order canonical algebra is isomorphic with respect to the canonical algebra of the same order, and it may be used to construct an equivalent system.

Build up a mapping λ, representing a local diffeomorphism of system state spaces and converting the solutions of the system (2.5) into system (2.6) under the same controls. If the linear mapping l exists, $M = \mathbb{R}^n$, $N = \mathbb{R}^p$, $c_1(x_0), \ldots, c_n(x_0)$, then there is the basis of $\mathbb{R}^n$, and $d_1(y_0), \ldots, d_n(y_0)$ are the corresponding vectors from $p(\tilde{B})y_0$. Determine the mapping $(t, x) \to \alpha_i(t)x$ as a family of integral curves for $c_i(x)$, $i = 1, \ldots, n$, that is, such that $(d/dt)\alpha_i(t)x = c_i(\alpha_i(t)x)$, $\alpha_i(0) = x$. The mapping $(t, y) \to \beta_i(t)y$ is determined by analogy from the equation

$$(d/dt)\beta_i(t)y = d_i(\beta_i(t)y), \quad \beta_i(0)y = y, \quad i = 1, \ldots, n. \tag{2.10}$$

Introduce the variable $s = (s_1, \ldots, s_n)$ and determine the mappings g_1: $s \to x$ and g_2: $s \to y$, assuming $g_1(s)\alpha_n(s_n) \cdots \alpha_2(s_2)\alpha_1(s_1)x_0$ and $g_2(s) = \beta_n(s_n) \cdots \beta_2(s_2)\beta_1(s_1)y_0$. Then, $(\partial g_1/\partial s_i)(0) = c_i(x_0)$ and g_1 has the inverse mapping g_1^{-1}:$x \to s$ for the points $x \in U$. The mapping λ is determined by the relation $\lambda = g_2 \circ g_1^{-1}$.

Global bilinearization of nonlinear systems. Let L be the differential operator defined by

$$L(g(x)) = \sum_{i=1}^{n} f_{(i)}(x) \frac{\partial}{\partial x_i}(g(x)) = g_x(x) f(x) \tag{2.11}$$

and define the set of functions

$$S = \{h(x), Lh(x), \ldots\} \cup \left[\left\{ \bigcup_{i=1}^{\hat{q}} Q_i(x), LQ_i(x), \ldots \right\} \right], \tag{2.12}$$

where Q_i is the ith column of the matrix Q; $g(x)$ is a differentiable scalar function; $f_{(i)}$ is the ith component of the vector-function $f(x)$.

If there exist integers M_i, $i = 0, \ldots, \hat{q}$, such that for $k = 1, \ldots, \hat{q}$ and all states $x(t)$, $t \in T$,

$$\begin{aligned} L^{M_0} h(x(t)) = &\sum_{i=0}^{M_0 - 1} A_0(0, 0, i+1) L^i h(x(t)) \\ &+ \sum_{j=1}^{\hat{q}} \sum_{i=0}^{M_{\hat{q}} - 1} A_0(0, j, i+1) L^i Q_i(x(t)), \end{aligned} \tag{2.13}$$

$$\ldots \tag{2.14}$$

$$L^{M_k}Q_k(x(t)) = \sum_{i=0}^{M_0-1} A(k,0,i+1)L^i h(x(t)) \tag{2.15}$$

$$+\sum_{j=1}^{q}\sum_{i=0}^{M_q-1} A(k,j,i+1)L^i Q_j(x(t)), \tag{2.16}$$

where $A_0(i,j,k)$ is a constant $(p \times p)$-matrix, and every column of $(L^i h(x))_x G(x)$ and $(L^i Q_j(x))_x G(x)$, $i = 0,\ldots,M_{j-1}$, $j = 1,\ldots,\hat{q}$ lies on S; that is, for $k = 1,\ldots,m$ and all states $x(t)$, $t \in T$, the kth column of $(L^i h(x))_x G(x)$ is equal to the expression

$$\sum_{l=1}^{M_0} A_j(k,i,0,j)L^{j-1}h(x) + \sum_{j=1}^{\hat{q}}\sum_{l=1}^{M_{\hat{q}}} A(k,i,j,l,)L^{l-1}Q_j(x), \tag{2.17}$$

and the kth column $(L^i Q_j(x))_x G(x)$ is equal to the vector

$$\sum_{l=1}^{M_0} A_j(k,i,0,l)L^{l-1}h(x) + \sum_{r=1}^{\hat{q}}\sum_{l=1}^{M_{\hat{q}}} A_j(k,i,r,l)L^{l-1}Q_r(x), \tag{2.18}$$

and for some positive integers M_i, $i = 0,\ldots,\hat{q}$,

$$\begin{aligned}&\operatorname{rank}[C, A_0'C', \ldots, (A_0')^{M_0-1}C', D_1', A_0'D_1', \ldots, (A_0')^{M_1-1}D_1', \ldots\\ &\qquad D_{\hat{q}}'A_0'D_{\hat{q}}', \ldots, (A_0')^{M_{\hat{q}}-1}D_{\hat{q}}'] = \dim A;\end{aligned} \tag{2.19}$$

then (Lo, 1975) there is the bilinear system,

$$\dot{y}(t) = \left(A_0 + \sum_{i=1}^{m} A_i u_i(t)\right) y(t), \tag{2.20}$$

$$z(t) = \left(C + \sum_{j=1}^{\hat{q}} D_j v_j(t)\right) y(t). \tag{2.21}$$

Here A_0, A_i, C, D_j are constant matrices of appropriate sizes.

1.3 Optimal Control of Bilinear Systems

An optimal control for bilinear systems is considered here to describe the basic idea behind our new approach. It is shown that the optimal control problem can be reduced to the Euler equations. The optimal control problem solution is reduced to an integration of the Euler equations admitting a Lax representation. A soliton method (Zakharov et al., 1980) is proposed for the integration of the Euler equation.

Euler–Lagrange equation. Suppose that G is a matrix Lie group with corresponding Lie algebra g. Consider the dynamically equivalent bilinear system defined on G by

$$(d/dt)Y(t) = \left(A_0 + \sum_{i=1}^{m} u_i(t)A_i\right) Y(t), \tag{2.22}$$

where $A_0, A_1, \ldots, A_m$ are constant $(p \times p)$-matrices; $Y(t)$ is a varying $(p\times p)$-matrix; $u(t) = (u_1(t), \ldots, u_m(t))$ is a control, that is, a measurable function belonging to an input set Ω.

Given an admissible class Ω of control functions, we wish to find $u_1(t), \ldots, u_m(t)$ in Ω that steer the state of (2.22) from $I \in G$ to $Y_1 \in G$ in T units of time in such a way as to minimize the cost functional

$$\eta = \int_0^T \sum_{i,j=1}^{m} q_{ij}u_iu_j dt, \tag{2.23}$$

where $\tilde{Q} = (q_{ij})$ is a symmetric and positive-definite matrix.

We assume that the $\{A_1, \ldots, A_m\}$ in (2.22) Span g. In the light of the assumption set out in this section, this makes $\{A_1, \ldots, A_m\}$ a basis for g.

Theorem 2.1. *Let R be a nonsingular matrix either symmetric or skew-symmetric such that $R^2 = \pm I$. Suppose that*

$$\begin{aligned} g &= \{C \in gl(n, \mathbb{R}) : C^tR + RC = 0\}, \\ Y_1 &\in C = \{\exp g\}_G, \quad T > 0. \end{aligned} \tag{2.24}$$

Then there exists an optimal control matrix

$$U^0(t) = \sum_{i=1}^{m} u_i^0(t)A_i, \tag{2.25}$$

that steers (2.22) from I at $t = 0$ to Y_1 at $t = T$ such that (2.23) is minimized.

The optimal control matrix satisfies the differential equation (Bailleul, 1978)

$$\begin{aligned} &(d/dt)\left(\sum_{i=1}^{m} u_i(t)A_i\right) \\ &= \tilde{Q}^{-1}\left(\left[\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right),\ A_0^t + \sum_{i=1}^{m} u_i(t)A_i^t\right]\right), \end{aligned} \tag{2.26}$$

where $[\cdot,\cdot]$ denotes the Lie bracket of the matrices $\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right)$ and $A_0^t + \sum_{i=1}^{m} u_i(t)A_i^t$.

We can now rewrite equation (2.26) as follows,

$$(d/dt)\left[\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right)\right] = \left[\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i,\, A^t + \sum_{i=1}^{m} u_i(t)A_i^t\right)\right]. \tag{2.27}$$

Using the Lie bracket, this equation can be written in the form:

$$\dot{M} = [M, \Omega], \tag{2.28}$$

where

$$\Omega = A^t + \sum_{i=1}^{m} u_i(t)A_i^t, \quad M = \tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right). \tag{2.29}$$

We call equation (2.28) *the Euler–Lagrange equation* for our optimization problem.

1.4 On the Solution of the Euler–Lagrange Equation

If $A_0^t + \sum_{i=1}^{m} u_i(t)A_i^t \in \tilde{g}$, where $\tilde{g}$ is an algebra of real skew-symmetric $(n \times n)$-matrices with the ordinary operation of commutation, and S: $\tilde{g} \to \tilde{g}$ is a linear operator in g,

$$S\left(A_0^t + \sum_{i=1}^{m} u_i(t)A_i^t\right) = \tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right), \tag{2.30}$$

then the Euler–Arnold equation is

$$d/dt\left(\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right)\right) = \left[\tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right),\, A_0^t + \sum_{i=1}^{m} u_i(t)A_i^t\right]. \tag{2.31}$$

It follows from (2.31) that is, the eigenvalues of the matrix $M = \tilde{Q}\left(\sum_{i=1}^{m} u_i(t)A_i\right)$ are preserved in time; that is, the traces of powers in equation (2.30), the power traces of M, are integrals of motion. In every invariant manifold, distinguished by these conditions, (2.30) is a Hamiltonian system.

An interesting example of Euler's equations in the group g is provided by the equation of free rotation of an n-dimension rigid body. In this case, $S\Omega = J \bullet \Omega + \Omega \bullet J$, where J is a symmetric positive-definite matrix (inertia tensor), which can always be regarded as diagonal, and (2.31) can be rewritten as

$$J\dot{\Omega} = \dot{\Omega}J = [J, \Omega_2]. \tag{2.32}$$

Equation (2.32) with arbitrary n was first considered by Mishchenko (1970) who discovered a series of nontrivial quadratic integrals.

By Liouville's theorem, there are sufficient Mishchenko integrals in the case $n = 4$ for proving the complete integrability of Euler's equations of a four-dimensional rigid body. For any n, equation (2.32) has $N(n)$ single-valued integrals of motion, and its general solution is expressible in terms of Θ-functions of Riemann surfaces (Manakov, 1976; Zakharov et al., 1980).

Example 2.1. Consider the bilinear system

$$(d/dt)X(t) = \begin{bmatrix} 0 & u_3(t) & -u_2(t) \\ -u_3(t) & 0 & u_1(t) \\ u_2(t) & -u_1(t) & 0 \end{bmatrix} X(t) \tag{2.33}$$

and the performance criterion $\eta = \int_0^T \left[q_1u_1(t)^2 + q_2u_2(t)^2 + q_3u_3(t)^2\right] dt$, $q > 0$.

From Theorem 2.1, we find that optimal controls steering this system between fixed endpoints satisfy

$$q_1(du_1/dt) = (q_2 - q_3)u_2u_3, \quad q_2(du_2/dt) = (q_3 - q_1)u_1u_3,$$
$$q_3(du_3/dt) = (q_1 - q_2)u_1u_2. \tag{2.34}$$

Interpreting the u_i as angular velocities and q_i as moments of inertia about principal axes, the optimization problem corresponds to the problem in classical mechanics of finding the equations of motion of a rotating solid body in the absence of external torques (Baillieul, 1978). Here, η is the action, and equation (2.34) are Euler equations and Theorem 2.2 (Baillieul, 1978) show that kinetic energy and angular momentum are conserved.

Unfortunately, if we do not assume that the A_is span the Lie algebra, g is no longer a necessary condition for optimal control. Nevertheless, techniques exist that allow us to develop the requisite necessary condition, even when the A_is do not span. One approach is the maximum principle of this section. Alternatively, we can use a limiting argument coupled with (2.26).

A more direct approach to the optimization problem is to involve the high-order maximum principle developed by Krener (1977).

Theorem 2.2. *Suppose that h_0 and h_0^t have the orthogonal direct sum decomposition as*

$$h_0 = k_0 \bigoplus \cdots \bigoplus k_{r-1} \bigoplus k_{r+1} \bigoplus \cdots \bigoplus k_{s-1},$$

and let $r = 1$ in (2.31).

Suppose, moreover, that $\tilde{Q} : h_0 \to h_0$ has the form assumed above. Then, the solution to (2.26) may be written as

$$\sum_{i=1}^{m} u_i^0(t)A_i = U_0(t) + \sum_{i=1}^{s-1} U_i(t)$$
$$= \exp(A_0 t) \cdot C_1 \cdot \ \exp(-A_0 t)$$
$$+ \exp(A_0 t) \cdot \exp(C_1 t) \cdot \exp(C_2 - C_1)t \cdot U_0(0) \cdot \exp(-C_1 t) \cdot \exp(-A_0 t),$$
$$\sum_{i=1}^{\nu} \cdot v_i(t)A_i = U_r(t) + U_s(t)$$
$$= \exp(A_0 t) \cdot \exp(C_1 t) \cdot (U_r(0) + U_s(0)) \cdot \exp(-C_1 t) \cdot \exp(-A_0 t),$$

where

$$C_1 = \sum_{r+1}^{s-1} U_i(t),$$

$$C_2 = \sum_{r+1}^{s-1} (\lambda_{r-1} - \lambda_j)/\lambda_{r-1} U_i(0) - \lambda_{r-1}^{-1}(U_r(0) + U_s(0)).$$

Remark 2.1. In the paper by Faibusovich (1988) an application of the method of collective Hamiltonians to a class of optimal control problems can be found. A Hamiltonian system of the maximum principle is reduced to a system of differential equations on the dual to an optimal control Lie algebra of a problem (LOC) endowed with a Lie–Berezin–Kirillov Poisson structure. It enables as to construct exactly solvable cases using some techniques developed for completely integrable systems.

2. Stability of Bilinear Systems

There has been tremendous interest in nonlinear stabilization problems in the recent years, as evidenced by several numerical research articles (Ionescu and Monopoli, 1975; Dayawansa, 1998; Hanba and Yoshihiko, 2001). The main contributing factors that have been realized are modern sensors and advanced superconducting devices, and the like, which cannot be analyzed by using linear techniques alone. More advanced techniques are necessary in order to meet the design challenges. These lead to the generalization of well-known bilinear control theories such as stabilization of active physical systems and generalization of adaptive linear systems to the bilinear settings. There exist two distinctively different approaches that are actually equivalent to feedback control: the first one is dynamic programming and the other is the regular approach (Schättler, 1998). In dynamic programming the value function is calculated as a solution of the Hamilton–Jacobi–Bellman equation. Regular synthesis is a generalization of the classical method of characteristics for first-order partial differential equations to the Hamiltonian–Jacobi–Bellman equation and hence another way to realize dynamic programming. This section describes feedback algorithms for continuous and discrete systems.

2.1 Normed Vector Space

Let $\mathbb{R}^n$ denote an n-dimensional vector space and the norm of a vector $x = (x_1, \ldots, x_n)^T$ on $\mathbb{R}^n$ be denoted by

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}.$$

If A is an $n \times m$ matrix over $\mathbb{R}$, then the norm of A is defined by

$$\|A\| = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2}.$$

If f is a linear function from $\mathbb{R}^n$ to $\mathbb{R}$ then the norm of f is defined by

$$\|f\| = \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|}.$$

Let $L^p([t_0, \infty))$ denote the set of the measurable functions $g : [t_0, \infty) \to \mathbb{R}^n$, such that

$$\|g\|_p = \left(\int_{t_0}^{\infty} \|g(t)\|^p dt \right)^{1/p} < \infty \quad \text{if} \ \ 1 \leq p < \infty$$

and

$$\|g\|_\infty = \text{ess} \sup_{t \in (t_0, \infty)} \|g(t)\| < \infty \quad \text{if} \ \ p = \infty.$$

Let $G(t)$ be an $m \times n$ matrix. Note that $G(t)$ is bounded on $L^p([t_0, \infty))$ if

$$\|G(t)\|_p \triangleq \left(\int_{t_0}^{\infty} \|G(t)\|^p dt \right)^{1/p} < \infty, \quad p \neq \infty.$$

Suppose f is a continuous function (defined from $\mathbb{R}^n \to \mathbb{R}$), $f(0) = 0$ and satisfies the following hypothesis

$$\int_0^T \frac{\omega(t)}{t} dt \leq K_f T, \tag{2.35}$$

where

$$\omega(t) = \sup_{\|x\| \leq t} |f(x)|.$$

The inequality (2.35) is satisfied if either the function is linear or satisfies the Lipschitz condition and $f(0) = 0$; that is, $f(x) \leq K\|x\|$, for some constant K.

The following lemma is useful for both continuous and discrete systems.

Lemma 2.1. *Let f be a measurable function defined from $\mathbb{R}^n$ to $\mathbb{R}$ and satisfying (2.35). Suppose*

$$u(t) = f(y(t)), \quad y(t) = H(t)x(t),$$

then

$$|u(t)| \leq 4K_f \|H(t)\| \|x(t)\|.$$

Proof: Let us write

$$|u(t)| = |f(H(t)x(t))| \leq \sum_{j=0}^{\infty} \|Z\| \leq 2^{-j} \|H(t)\| \, \|x(t)\|^{\|f(Z)\|}$$
$$= \sum_{j=0}^{\infty} \omega(2^{-j} \|H(t)\| \, \|x(t)\|)$$
$$- 2 \sum_{j=0}^{\infty} (\omega(2^{-j} \|H(t)\| \, \|x(t)\|)/2^{-j+1})(2^{-j+1} - 2^{-j}). \tag{2.36}$$

It is clear that $\omega(t)$ is a bounded increasing function. Therefore, the term in the first pair of parentheses is not larger than the minimum value of the function,

$$\omega(s\|H(t)\| \, \|x(t)\|)/s$$

on the interval

$$s \in [2^{-j}, 2^{-j+1}].$$

So, the last sum is less than

$$2\int_0^2 \frac{\omega(s\|H(t)\|\,\|x(t)\|)}{s}ds = 2\int_0^{2s\|H(t)\|\,\|x(t)\|}\frac{\omega(s)}{s}ds. \tag{2.37}$$

From (2.35), we conclude that

$$\|u(t)\| \leq 4K_f\|H(t)\|\,\|x(t)\|.$$

Remark 2.2. Here we should remark that K_f defined in (2.35) depends on the function f; for example, $K_f = \|f\|$ if f is a linear function, $K_f = K_u$ if f satisfies the Lipschitz condition: $||f(x_2) - f(x_1)|| \leq K_u|x_2 - x_1|$, and $f(0) = 0$. Therefore, for these two kinds of special functions, linear functions, or functions in the Lipschitz class, we have a better estimation,

$$|u(t)| \leq \|f\|\,\|H(t)\|\|x(t)\|.$$

It means that we can use $\|f\|$ and K_u to substitute K_f in the following theorems when f is a linear function and K_u is a function in the Lipschitz class, respectively.

2.2 Continuous Bilinear Systems

Let us define the continuous bilinear systems (BLS) as follows,

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)x(t)u(t) + c(t)u(t), \\ y(t) &= H(t)x(t), \quad u(t) = f(y(t)), \quad x_0 = x(t_0), \end{aligned} \tag{2.38}$$

where $x(t)$, $y(t)$ are n-dimensional vectors, A, B, H are $n \times n$ matrices, c is an $n \times 1$ matrix, and the entries of A, B, H, c are continuous functions. We assume that the solutions $x(t)$ are continuously differentiable functions, and the solutions are uniquely defined for the initial state x_0.

For BS (2.38), the following Theorem 2.3 may be derived from Lemma 2.1.

Theorem 2.3. *Let $x(t)$ denote a solution of the system (2.38). Suppose there exists a matrix D defined on time t, such that $A = DD^{-1}$. Let*

$$\Lambda \triangleq \int_0^\infty 4K_f\|H\|\,\|D^{-1}\|\,\|D\|\,\|c\|dt$$

and

$$G \triangleq \int_0^\infty 4K_f\|H\|\,\|D^{-1}\|\,\|D\|^2\|B\|dt.$$

If

$$1 - \Lambda e^{\Lambda} > 0,$$

then there exists a $\delta > 0$*, such that*

$$\|x(t)\| \le \frac{\|D^{-1}(t_0)X_0\|\,\|D(t)\|}{1 - (\Lambda + \|D^{-1}(t_0)X_0\|G)e^{\Lambda}} \tag{2.39}$$

for $\|X_0\| \le \delta$*.*

Remark 2.3. From (2.39), it is not difficult to show that the equilibrium at the origin of the system is stable if $\|D(t)\| \in L^\infty([t_0, \infty))$, and asymptotically stable if $\lim_{t\to\infty} \|D(t)\| = 0$.

The proof is somewhat lengthy and given by Mohler (1973).

Example 2.2. Consider the following bilinear system

$$\begin{bmatrix} \dot{X}_1(t) \\ \dot{X}_2(t) \end{bmatrix} = \begin{bmatrix} -2 & -1 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \exp(0.5t) & \exp(0.3t) \\ 0 & 0.5t \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} u(t)$$

$$+ \begin{bmatrix} (t+5)^{-3}\exp(-t) \\ \exp(-2t) \end{bmatrix} u(t), \tag{2.40}$$

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 0.5(t+1)^{-1} & -0.6 \\ 0.7 & 0.2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \tag{2.41}$$

where

$$u(t) = f(y(t)), \quad f(x) = |x| \sin\frac{1}{|x|}. \tag{2.42}$$

For the application of Theorem 2.3, let D be a fundamental matrix of (2.40):

$$D(t) = \begin{bmatrix} \exp(-2t) & \exp(-3t) \\ 0 & \exp(-3t) \end{bmatrix}.$$

Let $t_0 = 0$. In the Theorem 2.3, $\|H\| = 1.08$, $\|D\| \le \sqrt{3}\exp(-2t)$, $\|D^{-1}\| \le \sqrt{3}\exp 3t$, $\|C\| \le \sqrt{2}(t+5)^{-3}\exp(-t)$, $\|B\| \le \sqrt{3}\exp(0.5t)$, $K_f = 1$: we get $\Lambda = 0.366$, $G = 77.76$. Let $x = [x_1, x_2]^T$. Then we have,

$$\|x(t)\| \le \frac{\sqrt{3}\sqrt{(x_1+x_2)^2 + x_2^2\exp(2-t)}}{1 - 1.567 \cdot (0.366\sqrt{3G}\sqrt{(x_1+x_2)^2 + x_2^2}}.$$

Hence, the equilibrium at the origin for the system, (2.40)–(2.42), is uniformly stable and asymptotically stable.

2.3 Discrete Bilinear Systems

Consider the general form of a Discrete bilinear system with output feedback as follows,

$$x(t+1) = A(t)x(t) + \sum_{i=1}^{m} B_i(t)x(t)u_i(t) + C(t)u(t), \tag{2.43}$$

$$y(t) = H(t)x(t), \tag{2.44}$$

$$u(t) = (u_1(t), \ldots, u_m(t))^T = f(y(t)), \tag{2.45}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, $p \leq n$, $u \in \mathbb{R}^m$. $A(t)$, $B_i(t)$, $i = 1, \ldots, m$ are $n \times n$ matrices, $C(t)$ is an $n \times m$ matrix, and $H(t)$ is a $p \times n$ matrix, $f : \mathbb{R}^p \to \mathbb{R}^m$.

The following lemma is useful for the stability theory of discrete bilinear systems.

Lemma 2.2. *In the general bilinear system (2.43)–(2.45), assume that there exists $\alpha > 0$, a polynomial $h(\cdot)$ that does not include terms of degree ≤ 3, and positive coefficients, such that the following inequalities hold,*

$$\|x(t+1)\| \leq \alpha_1 \|x(t)\|^2 + h(\|x(t)\|) \tag{2.46}$$

or

$$\|x(t+1)\|^2 \leq \alpha_1 \|x(t)\|^2 + h(\|x(t)\|). \tag{2.47}$$

Then the zero state for the system (2.43)–(2.45) is uniformly stable and asymptotically stable, if $\alpha_1 < 1$.

Proof: (2.47) can be rewritten as

$$\|x(t+1)\|^2 \leq \alpha_1 \|x(t)\|^2 + g(\|x(t)\|)\|x(t)\|^2,$$

where polynomial $g(\cdot)$ has a degree > 1 and positive coefficients. If we take $t = 0$, $\|x(0)\| < \delta$ then

$$\|x(1)\|^2 \leq [\alpha_1 + g(\delta)]\delta^2 = \delta^2 \beta,$$

where

$$\beta \triangleq \alpha_1 + g(\delta).$$

For every $\epsilon > 0$, one can take small enough δ such that $\beta < 1$ and $0 < \delta < \epsilon$. If $\alpha_1 < 1$, then

$$\|x(t)\|^2 \leq [\alpha_1 + g(\delta\beta^{1/2})]\delta^2\beta \leq \delta^2\beta^2.$$

Without difficulty, by mathematical induction, one can show that

$$\|x(t)\| \leq \delta\beta^{t/2}. \tag{2.48}$$

This implies that the zero state for the system (2.43)–(2.45) is uniformly and asymptotically stable if $\beta \leq 1$ or $\beta < 1$, respectively.

Now we first consider the simple form of the bilinear system with output feedback

$$x(t+1) = A(t)x(t) + B(t)x(t)u(t), \tag{2.49}$$
$$y(t) = H(t)x(t), \tag{2.50}$$
$$u(t) = f(y(t)), \tag{2.51}$$

where $A(t)$, $B(t)$, $H(t)$ are $n \times n$ matrices, x and y are n-vectors, and $u(t)$ is a scalar input.

Let

$$\lambda_1 \triangleq \sup_{t\geq 0} \lambda_{\max}[A^T(t)A(t)], \tag{2.52}$$

$$\lambda_2 \triangleq \sup_{t\geq 0} \lambda_{\max}[B^T(t)B(t)], \tag{2.53}$$

$$\lambda_3 \triangleq \sup_{t\geq 0} \max |\lambda[B^T(t)A(t) + A^T(t)B(t)]|. \tag{2.54}$$

Theorem 2.4. *If $f : \mathbb{R}^n \to \mathbb{R}$, $H(t)$ is uniformly bounded on Z_+, $\lambda_1 < 1$, and $\lambda_2 < \infty$, $\lambda_3 < \infty$, then the zero state for the system (2.49)–(2.51) is uniformly and asymptotically stable.*

Now we consider the more general system (2.43)–(2.45) with multiple output feedback.

Let

$$\lambda_2 \triangleq \sup_{t\geq 0} \max_{1\leq i,j\leq m} \{\max |\lambda[B_i^T(t)B_j(t)]|\}, \tag{2.55}$$

$$\lambda_3 \triangleq \sup_{t\geq 0} \max_{1\leq i\leq m} \{\max |\lambda[B_i^T(t)A(t) + A^T(t)B_i(t)]|\}. \tag{2.56}$$

Theorem 2.5. *Let us suppose that $C(t)$ and $H(t)$ are uniformly bounded on Z^+,*

$$\sqrt{\lambda_1} + K_1 F_H F_c < 1,$$

where K_1 is a constant that may depend on $f(\cdot)$. Then the zero state of the system (2.43)–(2.45) is uniformly and asymptotically stable.

Example 2.3. Consider the dynamical system (Mohler, 1973)

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-R_a/L_a & 0 & 0\\ 0 & 0 & 1\\ 0 & 0 & -D/J\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

$$+\begin{bmatrix}0 & 0 & -K_a^*/L_a 0 \;\; 0 \;\; 0\\ K_y/J & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} u_1 + \begin{bmatrix}1/L_a\\ 0\\ 0\end{bmatrix} v, \tag{2.57}$$

$$\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}, \tag{2.58}$$

$$\dot{x} = Ax + Bxu_1 + cv, \tag{2.59}$$

$$y = Hx, \tag{2.60}$$

where $x_1 = i_a$, $x_2 = \theta$, $x_3 = \omega$, $u_1 = i_a$, $v = v_a$, J is the moment of intertia, D is the viscous damping ratio, R_a is the armature resistance, L_a is the applied armature inductance, K_y, K_a are motor characteristics, K_a is the motor const, i_a is the armature current, i_e is the field current, v_a is the armature voltage, ω is the angular velocity, and θ is the angular position.

The motor control problem is to choose the functions f_1, f_2 such that the obtained system with feedback is stable.

Equations (2.57) and (2.58) can be discretized by use of a first-order Euler expansion to give

$$x(t+1) = x(t) + TAtx(t) + TBx(t)u_1(t) + Tcv(t), \tag{2.61}$$

$$y(t) = Hx(t), \tag{2.62}$$

where T is the sampling interval. Equation (2.61) can be rewritten as

$$x(t+1) = A^*x(t) + B^*x(t)u_1(t) + c^*v(t), \tag{2.63}$$

where

$$A^* = I + TA = \begin{bmatrix}1 - TR_a/L_a & 0 & 0\\ 0 & 1 & T\\ 0 & 0 & 1 - TD/J\end{bmatrix},$$

$$B^* = TB = \begin{bmatrix}0 & 0 & -K_y^*T/L_a\\ 0 & 0 & 0\\ K_yT/J & 0 & 0\end{bmatrix},$$

$$c* = Tc = \begin{bmatrix}T/L_a\\ 0\\ 0\end{bmatrix}. \tag{2.64}$$

Here,

$$\lambda_1(A^*) = 1 - TR_a/L_a, \quad \lambda_2(A^*) = 1, \quad \lambda_3(A^*) = 1 - TD/J.$$

Let $u = [u_1 v]^T = Sy(t)$, where S is a constant matrix

$$S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}.$$

Thus, the corresponding feedback of system (2.61), (2.62) is

$$x(t+1) = A^* x(t) + B^* x(t) u_1(t) + c^*(t), \tag{2.65}$$

$$y(t) = Hx(t), \quad u(t) = Sy(t). \tag{2.66}$$

It can be shown that

$$\begin{aligned} \|x(t+1)\|^2 &\leq \lambda_1 \|x(t)\|^2 + \lambda_3 \|x(t)\|^2 |u_1(t)| + \lambda_2 \|x(t)\|^2 |u_1^2(t)| \\ &\leq \lambda_1 \|x(t)\|^2 + \lambda_3 \|S\| \, \|H\| \, \|x(t)\|^3 + \lambda_2 \|S\|^2 \|x(t)\|^4. \end{aligned} \tag{2.67}$$

Then, by Lemma 2.2, we conclude that the zero state of (2.65), (2.66) is uniformly and asymptotically stable when $\lambda_1 < 1$.

The principle of choosing s_{ij} is to reduce the closed-loop eigenvalues of $A^{**T}A^{**}$ where $A^{**} = A^* + CSH$. For convenience, we choose $s_{22} = 0$, and choose s_{12} such that

$$|1 - R_a/L_a + c_1 s_{21}| < 1,$$
$$-1 < 1 - R_a/L_a c_1 s_{21} < 1$$

and

$$s_{21} \in \left(\frac{-2}{c_1} + \frac{R_a}{c_1 L_a}, \frac{R_a}{c_1 L_a} \right),$$

Simulations have been made for this example.

3. Adaptive Control of Bilinear Systems

3.1 Control of Fixed Points

We apply the adaptive control to bilinear systems in $n = 2$ and 3 dimensions with a single parameter that is varied. All systems have a stable region where asymptotically the motion goes to a fixed point attractor, as well as regimes with a limit cycle (following a Hopf bifurcation) or more complicated periodic and aperiodic behavior.

In studying control we have the following situations (Isidory, 1995). Given an initial value of the parameter, the system evolves — following equation (2.1) — to its stable steady state. At time $t = t_0$, say, the system is perturbed: here this implies an instantaneous change in the parameter value. Subsequently, the system evolves under control dynamics (i.e., equations (2.1) and (2.3)), and returns to the original steady-state. Clearly, a convenient quantifier of this process is the time of recovery τ, which depends both on the perturbation as well as on the stiffness of control.

Application to single parameter perturbation. The first system we study is a textbook example with one nontrivial degree of freedom, described by the following equations,

$$\dot{r} = u - r^3, \quad \dot{\theta} = \omega, \quad u = \alpha r. \tag{2.68}$$

The sign of α determines the dynamics: when $\alpha < 0$ the system evolves to a fixed point ($r = 0$), and when $\alpha > 0$ there is a supercritical Hopf bifurcation (or soft excitation) and the system evolves to a (circular) limit cycle of radius $r_c = \alpha^{1/2}$. The control dynamics is determined by the error signal

$$\dot{\alpha} = \epsilon(r - \langle r \rangle), \tag{2.69}$$

where $\langle r \rangle$ is the steady-state value of r. To pull the system back to the fixed point, $\langle r \rangle$ in equation (2.69) must be 0.

The above system allows analytic treatment. Initially, $\alpha = \alpha_{in} < 0$ and the dynamics is attracted to $r = 0$. When r is suddenly perturbed to a positive value (at $t = 0$), we study how the system relaxes back to the fixed point. It is convenient to rewrite equation (2.68) as a Riccati equation,

$$\frac{dr}{d\alpha} = \frac{1}{\epsilon}(\alpha - r^2). \tag{2.70}$$

Further transformation to an equation of second order,

$$\frac{d^2 u(\alpha)}{d\alpha^2} + \frac{\alpha u}{\epsilon^2} = 0, \tag{2.71}$$

allows $r(\alpha)$ to be written in terms of the two linearly independent solutions u_1 and u_2 of equation (2.71),

$$\frac{r}{\alpha}(C u_1 + u_2) + C\frac{du_1}{d\alpha} + \frac{du_2}{d\alpha} = 0. \tag{2.72}$$

For positive α, and using the variable $z = \alpha/\epsilon^{2/3}$, equation (2.71) can be expressed as

$$\frac{d^2 u(z)}{dz^2} + zu = 0, \tag{2.73}$$

which has the solutions $u_1 = Ai(z)$, $u_2 = Bi(z)$. Substituting this in equation (2.72) we get

$$r = -\epsilon^{1/3} \frac{CAi(z) + Bi(z)}{CAi(z) + Bi(z)}. \tag{2.74}$$

For small ϵ (i.e., large z), we can approximate the Airy functions to obtain (in terms of $\zeta = \frac{2}{3}z^{3/2}$)

$$r(\alpha) = -\epsilon^{1/3} \frac{C\left(-\frac{1}{2}\pi^{-1/2}z^{1/4}e^{-\zeta}\right) + \pi^{-1/2}z^{1/4}e^{\zeta}}{C\left(\frac{1}{2}\pi^{-1/2}z^{-1/4}e^{-\zeta}\right) + \pi^{-1/2}z^{-1/4}e^{\zeta}} \approx -\alpha^{1/2}. \tag{2.75}$$

Then solving

$$\dot{\alpha} = \epsilon r = -\epsilon \alpha^{1/2}, \tag{2.76}$$

we get

$$\alpha^{1/2} = -\frac{1}{2}\epsilon t + \alpha_0^{1/2}, \tag{2.77}$$

where α_0 is the value of α at $t = 0$. Thus

$$r(t) = -\frac{1}{2}\epsilon t + \alpha_0^{1/2}, \tag{2.78}$$

and the recovery time τ is then given approximately by

$$\tau = 2\alpha_0^{1/2}/\epsilon. \tag{2.79}$$

Although the above analysis is valid for small ϵ, in practice we find that equation (2.78) describes the recovery behavior over a wider range. This system (unlike the logistic or other unimodal maps examined by HL) has globally attracting steady states and so there is no perturbation, however large, from which the system does not recover in finite time. The recovery time τ is always close to the estimate provided by equation (2.78).

Nonlinear feedback control. We next analyze the system for which the evolution equations are

$$\dot{x} = \alpha + uy - \beta x - x, \tag{2.80}$$

$$\dot{y} = \beta x - uy, \quad u = x^2. \tag{2.81}$$

The nature of the dynamics is determined by the parameter β. When the asymptotic motion is attracted to a fixed point, the steady-state values of the BS can easily be seen to be $y_s = \beta_{in}/\alpha_{in}$ and $x_s = \alpha_{in}$. When β is perturbed, the equation for control dynamics becomes

$$\dot{\beta} = -\epsilon(y - y_s). \tag{2.82}$$

Linear stability analysis of equation (2.82) yields the following conditions for equilibrium,

$$\beta < \alpha_{in}^2 + 1 \tag{2.83}$$

and

$$\alpha_{in}^4 + \alpha_{in}^3(\epsilon - \beta_{in}) + \alpha_{in}^2(1 - \epsilon\beta_{in}) > 0. \tag{2.84}$$

Without loss of generality we set $\alpha_{in} = 1$ and equation (2.83) then gives the condition for attraction to the fixed point to be $\beta_{in} < 2$; for $\beta_{in} > 2$ the system is attracted to a limit cycle. Furthermore, equation (2.84) gives a stability window determined by

$$1 + (1 - \beta_{in})(\epsilon + 1) > 0. \tag{2.85}$$

Note that equation (2.85) sets a restriction on the value of ϵ so here the range of control stiffness is limited by stability considerations.

Examination of the dependence of τ on ϵ reveals a novel feature that is not observed in the one-dimensional case. For small ϵ, $\tau \sim 1/\epsilon$. But τ does not decrease monotonically with ϵ and beyond an optimum value of ϵ, τ actually starts increasing. A rough argument accounting for the linear relation between τ and $1/\epsilon$ in the small ϵ range makes use of the observation that $y \approx_{in} /x$ (for small ϵ). Substitution in equation (2.80) gives $x(t) = 1 - \text{const} \cdot \epsilon^{-\text{t}}$, and equation (2.81) becomes

$$\dot{\beta} \approx = -\epsilon(\beta/x - \beta_{in}). \tag{2.86}$$

Assuming small ϵ and a small perturbation, we then have

$$\beta \approx \beta_{in} - Ce^{\epsilon t} \approx \beta_{in} - \text{const}\,(1 - \epsilon \text{t}), \tag{2.87}$$

suggesting that recovery time $\tau \propto 1/\epsilon$.

A biochemical network. The third example we study is a complex dynamical system (Loskutov and Mikhailov, 1990) which describes various biochemical processes responsible for the coherent behavior observed in spatiotemporal organization. The equations contain positive and negative feedback loops that are typically thought to occur in a variety of processes within living cells. This system, which gives rise to a variety of behavioral patterns, is

$$\dot{X}_1 = \frac{a_1}{a_2 + X_3^n} - uX_1, \tag{2.88}$$

$$\dot{X}_2 = a_3X_1 - \phi(X_2, X_3), \tag{2.89}$$

$$\dot{X}_3 = a_4\phi(X_2, X_3) - qX_3, \tag{2.90}$$

where

$$\phi(X_2, X_3) = \frac{TX_2(1+X_2)(1+X_3)^2}{L+(1+X_2)^2(1+X_3)^2}. \tag{2.91}$$

Choosing parameters (a_1, a_2, a_3, a_4, L, T and n) suitably, we have a system whose dynamics can be varied by tuning the parameters K and q. For instance, for $q = 0.5$, we get a limit cycle when $u = 0.001$, complex oscillations when $L = 0.003$, chaos when $u = 0.004$, complex oscillations and period doublings when $0.005 < u < 0.02$, a limit cycle again when $0.03 < u < 0.5$, and a steady-state when $u = 1$. For control we let u evolve as

$$\dot{u} = -\epsilon(X_1 - \langle X_1 \rangle). \tag{2.92}$$

Equation (2.92) is very effective in returning the system back to the original steady-state ($u = 1$) when u is perturbed into any of the other above-mentioned regimes, including the chaotic region.

A discrete dissipative system. We finally apply the control to a two-dimensional discrete system given by (Loskutov and Mikhailov, 1990)

$$X_{n+1} = 1 - \frac{\alpha X_n^2}{1+X_n^4} - uY_n, \tag{2.93}$$

$$Y_{n+1} = X_n, \tag{2.94}$$

which is similar to the Hennon map with the additional feature that the global asymptotic dynamics is on an attractor for $u < 1$. When α is varied, this map gives rise to the entire repertoire of behavior seen in unimodal chaotic maps. For regulation of the steady state the control dynamics given by the equation

$$\alpha_{n+1} = \alpha_n - \epsilon(X_n - X_s) \tag{2.95}$$

is very effective. For small ϵ the recovery time τ is proportional to $1/\epsilon$, but beyond $\epsilon = \epsilon_{\text{opt}}$, τ increases with ϵ, similar to what is observed for the Brusselator reaction-diffusion model.

The above dissipative systems with more than one degree of freedom can show novel behavior quite distinct from the one-dimensional case. For example, there is a maximum strength of shock for every value of ϵ beyond which the system does not recover. Below $\epsilon = \epsilon_c$ recovery is possible for shocks of all strengths (provided that the shock does not throw the perturbed value of α outside the allowed range). For $\epsilon > \epsilon_c$ the system fails to recover from all shocks however small in magnitude, so that the $\delta_{\max}$ versus characteristic is a step function rather than the linearly decreasing function.

Applications to multiparameter systems. Typically, a dynamical system has several parameters that govern the overall behavior. In order to regulate such systems, the control has to be applied for each relevant parameter.

A representative one-dimensional map with two parameters (which is of interest in population dynamics) is given by

$$X_{n+1} = u(1 + X_n)^{-\beta}, \quad u = \alpha X_n. \tag{2.96}$$

For specific α and β the map has a globally stable equilibrium state: when the parameters α and β are varied the map yields a rich variety of dynamical behavior. To regulate the steady state of the system we control both parameters in an obvious manner through

$$\alpha_{n+1} = \alpha_n + \epsilon(X_n - X_s), \tag{2.97}$$

$$\beta_{n+1} = \beta_n + \epsilon(X_n - X_s), \tag{2.98}$$

where ϵ is the control stiffness.

Similarly, for a two-dimensional discrete map of two driven coupled oscillators given by

$$x_{n+1} = u_1(1 - x_n) + \beta(y_n - x_n), \quad u_1 = \alpha x_n \tag{2.99}$$

$$y_{n+1} = u_2(1 - y_n) + \beta(x_n - y_n), \quad u_2 = \alpha y_n \tag{2.100}$$

the same control dynamics is effective. In both cases the recovery time remains linearly dependent on ϵ.

In the control dynamics implemented here, we have always chosen an error indicator that utilizes a single variable, $X_i - X_s$. In higher-dimensional systems there is an ambiguity regarding the choice of the state variable X, to be used in equation (2.1). Empirically, we observe that in most systems any of the variables can effect control, because the equilibrium condition will lead all other variables to their steady-state values when any one of them is forced to reach a steady-state. However, there are exceptions: for instance, in the Brusselator, using coordinate y in equation (2.81) results in control whereas using x does not. This is because the steady-state value of x ($x_s = \alpha$) does not constrain β to be the desired value, whereas the steady state of y ($y_s = \alpha\beta$) does. On the other hand, in the three-dimensional system given by equations (2.88)–(2.90), all three variables can effect control. X_1 works most efficiently, however, as the magnitude of X_1 (and hence the error signal) is small, leading to a more stable control dynamics.

One method of removing the above-mentioned ambiguity is by employing AND logic in the control, that is, by requiring that all variables reach their steady-state values, X_i^s, $i = 1, 2, \ldots, N$. The equation for control dynamics then becomes

$$\dot{\mu} = \epsilon \sum_{i=1}^{N} (X_i - X_i^2). \tag{2.101}$$

In the examples we have studied, either equation (2.101) works equally efficiently.

3.2 Control of Limit Cycles

For the simple oscillator described in equation (2.1), limit cycles can be adaptively controlled. In this case defining the error signal is quite unambiguous as every limit cycle is uniquely characterized by its radius r_c. The difference between the actual radius and the radius of the limit cycle to be controlled can be used effectively for regulatory feedback. This is done, for example, by setting $\langle r \rangle$ in equation (2.68) equal to r_c. When perturbed onto a different limit cycle (radius not equal to r_c) or into the fixed point region ($\alpha > 0$), the system rapidly relaxes to the original limit cycle. For small $\epsilon\,\tau$, T is inversely proportional to ϵ, but for large ϵ we observe a different phenomenon: τ oscillates about a saturation value that is roughly constant for all values of perturbation.

When r_c is small and the system is perturbed to a much larger radius the control dynamics for small ϵ is determined by a set of equations similar to equation (2.69) (as $r > r_c$): we can therefore expect the same linear trend. When the system is kicked to the fixed point region $(r/t_0, 0)$ the control dynamics is approximated by

$$\dot{\alpha} = -\epsilon r_c, \tag{2.102}$$

so that (for small ϵ)

$$\alpha(t) = -\epsilon r_c t + \alpha_0, \tag{2.103}$$

from which the inverse dependence of τ on ϵ follows. For large ϵ this is not valid.

More generally, we can extend the above procedure to control cycles in discrete systems. What is required is an error indicator that encodes as much information about the cycle as is necessary for its unique characterization. An error signal depending on $X_{n+2} - X_n$ suffices in bringing the system back to some period 2-cycle, but not to a specific 2-cycle: $X_{n+2} - X_n = 0$ for all period 2-cycles, and so cannot guide the control dynamics onto any particular cycle. To regulate specific cycles we require an error that is unique for every 2-cycle: one possibility is an error proportional to $\mid X_{n+1} - X_n \mid - \mid X_1^c - X_2^c \mid$, where X_1^c and X_2^c are the values of the iterates of X in the 2-cycle we want to control.

We implement this for the logistic map

$$X_{n+1} = uX_n(1 - X_n). \tag{2.104}$$

The control dynamics follows from

$$u_{n+2} = \alpha_n - \epsilon(\mid X_{n+1} - X_n \mid - \mid X_1^c - X_2^c \mid). \tag{2.105}$$

It is clear that this control mechanism very effectively returns the system to the desired 2-cycle. The recovery time varies inversely as ϵ. There is also a maximum strength of shock $\delta_{\max}$ (depending on ϵ), beyond which the system fails to recover: this $\delta_{\max}$ versus ϵ curve shows a step function pattern.

Similar error indicators can, in principle, be constructed for higher-order periodic orbits although the technique diminishes in utility with increase in period. This is a problem of practicality, because higher-order cycles typically have narrow windows of stability; as a consequence the control dynamics become very unstable.

For discrete dynamical systems, however, more effective algorithms can be devised. One, which employs a logical OR structure in the error indicator, is

$$\dot{\alpha} = \epsilon \prod_{i=1}^{k} (X - X_i^c), \tag{2.106}$$

where X_i^c, $i = 1, 2, \ldots, k$ is the stable period k-orbit to be controlled. Because it implies that the desired state is either $X = X_1^c$ or $X = X_2^c$ or $\ldots$ $X = X_k^c$ this adaptive algorithm works at every iteration step. For higher-order periodic orbits, this latter method is far superior to that embodied in equation (2.106). For controlling the 2-cycle, for instance, the control equation analogous to equation (2.106) is

$$\alpha_{n+1} = \alpha_n - \epsilon(X_n - X_1^c)(X_n - X_2^c). \tag{2.107}$$

Even with this latter form, the inverse dependence of τ on ϵ is unchanged.

3.3 Variations in the Control Dynamics

Apart from sudden perturbations in the system environment that lead to parameters changing value drastically (the primary case studied above) there are additional noise effects that can occur. In particular, it is interesting to consider the effect of random background noise on the control algorithm. In an effort to explore this question, we study the discrete map equations (2.104) and (2.96), with additional Gaussian noise. The control dynamics remain unchanged. The variance σ of this noise clearly determines the control behavior: for small σ, recovery times with and without noise are virtually identical, and beyond a value $\sigma = \sigma_{\max}$, recovery is not possible. However, most important, this control procedure is remarkably robust for $\sigma < \sigma_{\max}$, and the recovery time continues to remain inversely proportional to the stiffness.

A question of some importance is whether the control algorithm is sensitive to the specific form of the control dynamics, namely the choice for $g(X - X_s)$ in equation (2.5). In realistic systems, the control dynamics that can be incorporated may be of a arbitrary functional form, arising, for example, from physicochemical or engineering design considerations specific to the system. It is thus necessary to determine whether the linear recovery we observe in the examples above is an artifact of using a linear control function, and also whether such an adaptive control is more generally applicable with different functions $g(y)$. In order to explore the features of control with different (nonlinear) functions, we have varied the control function, using $g = y^2\, y^{1/2} \sin y\, 1 - e^2$ and $y(1 - y)$. Results are shown: for all functional forms recovery times remain inversely proportional to control stiffness for small ϵ. This strongly suggests that linear recovery may be a universal feature of the adaptive control algorithm.

From our study of higher-dimensional systems of varying complexity, it appears that it is possible to provide efficient regulation of the steady state of nonlinear systems through adaptive control mechanisms. The procedure studied herein utilizes an error signal proportional to the difference between the goal output and the actual output of the state variables and should be contrasted with mechanisms (Haken, 1978) using an error signal proportional to a similar difference in the parameter value. In the latter case the control will, of course, bring the parameter back to its original value, but this does not ensure that the system will regain its specific original dynamical state. An instance where this distinction is important is in systems undergoing a subcritical Hopf bifurcation or exhibiting bistability; at a given parameter value, different initial conditions lead to different dynamics. In such a case, the present adaptive control ensures that both the original parameter value and the original dynamics will be recovered.

From numerical experiments studying the dependence of controllability on the stiffness and on the strength of perturbation, we find a number of interesting phenomena quite distinct from that seen in HL, but typical of most real systems. For multiparameter systems, a simple extension of the adaptive mechanism suffices in regulating the system; furthermore it can also be adapted to regulate periodic behavior such as limit cycles.

The HL procedure is robust both to the existence of background noise, and to the variation of the form of the control function. Most interestingly, recovery times are always inversely proportional to the stiffness of control, for a small stiffness, which may be a universal feature of such adaptive control.

Biological situations where control is believed to play a crucial role include, for instance, the maintainence of homeostasis (the relative constancy of the internal environment with respect to variables such as blood pressure, pH, blood sugar, electrolytes, and osmolarity). Clinical experiments on animals show, for example, that following a quick mild hemorrhage (a sudden perturbation in arterial pressure) the blood pressure is restored to equilibrium values within a few seconds. The control of fixed points thus has a potential utility in such physicochemical contexts. Cycles are also central to a variety of biophysical and biochemical processes. Variations in these—for example, by the replacement of periodic by aperiodic behavior, or the emergence of new periodic cycles—is often associated with disease (Haken, 1978, 1988). The control of cycles has applicability in the regulation of biologically significant oscillatory phenomena.

In summary, our study confirms that adaptive control provides a simple, powerful, and robust tool for regulating multidimensional systems capable of complicated behavior. The concepts developed through the study of model systems can serve as a paradigm for understanding more complex regulatory mechanisms widespread in nature. These may also be of use in helping formulate efficient and robust design principles.

4. Notes and Sources

Much of the material on bilinear time-optimal control problems is standard in the literature on mathematical systems theory. The present exposition essentially follows Brockett (1979) and Bailleul (1978), except for the additional focus on the soliton approach, which is not specifically recognized in this literature. The soliton approach based on the Lax form is essentially taken from the study by Yatsenko (1984). The stabilization problem illustrates the classic origin of the Lyapunov method used by Gutman (1981), Gounaridis and Kalouptsidis (1986), Longchamp (1980), Quinn (1980), Ryan and Buckingham (1983), and Slemrod (1978).

Springer

http://www.springer.com/978-0-387-73668-6

Optimization and Control of Bilinear Systems
Theory, Algorithms, and Applications
Pardalos, P.; Yatsenko, V.A.
2008, XXVI, 370 p. 40 illus., Hardcover
ISBN: 978-0-387-73668-6