Annals of Operations Research 98, 19–44, 2000

Approximate Optimal Control and Stability of Nonlinear Finite- and Infinite-Dimensional Systems

S.P. BANKS and K. DINESH s.banks@sheffield.ac.uk
Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

Abstract. We consider first nonlinear systems of the form

$$\dot{x} = A(x)x + B(x)u,$$

together with a standard quadratic cost functional and replace the system by a sequence of time-varying approximations for which the optimal control problem can be solved explicitly. We then show that the sequence converges. Although it may not converge to a global optimal control of the nonlinear system, we also consider a similar approximation sequence for the equation given by the necessary conditions of the maximum principle and we shall see that the first method gives solutions very close to the optimal solution in many cases. We shall also extend the results to parabolic PDEs which can be written in the above form on some Hilbert space.

Keywords: optimal control, nonlinear systems, parabolic systems

1. Introduction

The optimal control of general nonlinear systems of the form

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

with the quadratic cost function

$$\min J = \frac{1}{2}x^{\mathrm{T}}(t_f)Fx(t_f) + \frac{1}{2}\int_0^{t_f} \left(x^{\mathrm{T}}Qx + u^{\mathrm{T}}Ru\right) \mathrm{d}t$$

can be solved, in principle, by the use of Lie series and infinite-dimensional bilinear systems theory [3–5]. However, the solution is complex and difficult to implement. For this reason we consider "pseudo-linear" systems of the form

$$\dot{x} = A(x)x + B(x)u$$

together with a quadratic cost and introduce a sequence of (time-varying) linear-quadratic approximations to the nonlinear problem. It should be noted that the method proposed here can easily be generalised to the case of nonlinear cost functions of the form

$$J = \frac{1}{2}x^{\mathrm{T}}(t_f)Fx(t_f) + \frac{1}{2}\int_0^{t_f} \left(x^{\mathrm{T}}Q(x)x + u^{\mathrm{T}}R(x)u\right) \mathrm{d}t.$$

See also [6] for a related “freezing technique”.

This method has been applied to bilinear and more general nonlinear systems in [1,9]. Here we study the method in complete generality and show that it is very effective in that it is easy to apply and converges rapidly. Other robust techniques for nonlinear systems (e.g., [18]) are more difficult to apply since they are not based on simple linear approximations. In fact the methods based on viscosity solutions of the Hamilton–Jacobi–Isaacs equation are powerful, but apply only to certain types of nonlinearity and are very complex [20]. Of course, there are many linearisation techniques which are only of local usefulness (see [15]). Moreover, our method applies directly to distributed systems with very little extra effort. It will also apply to many other aspects of nonlinear systems theory; for example stability and chaotic motion see [10].

In section 2 we shall introduce the approximating sequence and in section 3 we shall prove convergence of the sequence in an appropriate space. Some examples will be given in section 4. In section 5 we shall consider the global optimal solution and in section 6 the method is extended to parabolic PDEs, while in section 7 we show that the method applies to stability theory and exemplify this by studying a large-scale nonlinear distributed delay system. Further results in the theory of “pseudo-linear” systems using Lie and Clifford algebras can be found in [7,8].

The results in this paper are very general; the basic assumption made is of local Lipschitz continuity. Many of the results are proved assuming Lipschitz continuity, but as noted below they are easily extended to the local case and so apply to a very wide range of systems as will be shown by the examples.

2. The approximating sequence

Consider the pseudo-linear system

$$\dot{x} = A(x)x + B(x)u, \quad x(0) = x_0 \tag{2.1}$$

together with the quadratic cost

$$J = x^{\mathrm{T}}(t_f)Fx(t_f) + \int_0^{t_f} \left(x^{\mathrm{T}}Qx + u^{\mathrm{T}}Ru\right)\mathrm{d}t. \tag{2.2}$$

We introduce the following sequence of approximations to the problem of minimising the cost (2.2) subject to the dynamics (2.1):

$$\begin{aligned}
\dot{x}^{[0]} &= A(x_0)x^{[0]} + B(x_0)u^{[0]}, \quad x^{[0]}(0) = x_0,\\
J^{[0]} &= x^{[0]\mathrm{T}}(t_f)Fx^{[0]}(t_f) + \int_0^{t_f} \left(x^{[0]\mathrm{T}}Qx^{[0]} + u^{[0]\mathrm{T}}Ru^{[0]}\right)\mathrm{d}t
\end{aligned}$$

and for $k \geqslant 1$,

$$\begin{aligned}
\dot{x}^{[k]} &= A\big(x^{[k-1]}(t)\big)x^{[k]} + B\big(x^{[k-1]}(t)\big)u^{[k]}, \quad x^{[k]}(0) = x_0,\\
J^{[k]} &= x^{[k]\mathrm{T}}(t_f)Fx^{[k]}(t_f) + \int_0^{t_f} \left(x^{[k]\mathrm{T}}Qx^{[k]} + u^{[k]\mathrm{T}}Ru^{[k]}\right)\mathrm{d}t.
\end{aligned} \tag{2.3}$$

Since each approximating problem in (2.3) is linear (time-varying), quadratic we can write the optimal control in the form

$$u^{[k]} = -R^{-1}B^{\mathrm{T}}\big(x^{[k-1]}(t)\big)P^{[k]}x^{[k]}(t),$$

where $P^{[k]}$ is the solution of the usual Riccati equation

$$\begin{aligned}\dot{P}^{[k]}(t) &= -Q - P^{[k]}A\big(x^{[k-1]}(t)\big) - A\big(x^{[k-1]}(t)\big)^{\mathrm{T}}P^{[k]}\\ &\quad + P^{[k]}B\big(x^{[k-1]}(t)\big)R^{-1}B^{\mathrm{T}}\big(x^{[k-1]}(t)\big)P^{[k]},\\ P^{[k]}(t_f) &= F\end{aligned} \tag{2.4}$$

and the kth dynamical system becomes

$$\dot{x}^{[k]} = A\big(x^{[k-1]}(t)\big)x^{[k]} - B\big(x^{[k-1]}(t)\big)R^{-1}B^{\mathrm{T}}\big(x^{[k-1]}(t)\big)P^{[k]}x^{[k]}(t). \tag{2.5}$$

3. Proof of convergence

In this section we shall prove that the sequence of coupled equations (2.4) and (2.5) converges under certain conditions on $A(x)$ and $B(x)$ and for small enough horizon time t_f (or small enough initial condition x_0). (Throughout this paper, we shall be using Euclidean norms.) To do this first note that if $\Phi^{[i-1]}(t, t_0)$ denotes the transition matrix generated by $A(x^{[i-1]}(t))$ then (by [12]) we have

$$\big\|\Phi^{[i-1]}(t, t_0)\big\| \leqslant \exp\left[\int_{t_0}^{t} \mu\Big(A\big(x^{[i-1]}(\tau)\big)\Big)\,\mathrm{d}\tau\right],$$

where $\mu(A)$ is the logarithmic norm of A. We next require an estimate for $\Phi^{[i-1]} - \Phi^{[i-2]}$, and we shall make the following assumptions:

(A1) $\mu(A(x)) \leqslant \mu_0 \quad \forall x \in \mathbb{R}^n$,

(A2) $\|A(x) - A(y)\| \leqslant \alpha\|x - y\| \quad \forall x, y \in \mathbb{R}^n$,

(A3) $\|C(x) - C(y)\| \leqslant \beta\|x - y\| \quad \forall x, y \in \mathbb{R}^n$,

(A4) $\|C(x)\| \leqslant \gamma \quad \forall x \in \mathbb{R}^n$.

Lemma 3.1. Suppose that $\mu(A(x)) \leqslant \mu$ for some constant μ and for all x and that

$$\big\|A(x) - A(y)\big\| \leqslant \alpha\|x - y\|, \quad \forall x, y \in \mathbb{R}^n$$

for some constant α.

Then

$$\big\|\Phi^{[i-1]}(t, t_0) - \Phi^{[i-2]}(t, t_0)\big\| \leqslant \alpha\, \mathrm{e}^{\mu(t-t_0)}(t - t_0) \sup_{s\in[t_0,t]} \big\|x^{[i-1]}(s) - x^{[i-2]}(s)\big\|.$$

Proof. $\Phi^{[i-1]}$, $\Phi^{[i-2]}$ are solutions of the respective equations

$$\dot{z} = A\big(x^{[i-1]}(t)\big)z, \quad z(t_0) = I,$$
$$\dot{w} = A\big(x^{[i-2]}(t)\big)w, \quad w(t_0) = I.$$

Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}(z - w) = A\big(x^{[i-1]}(t)\big)(z - w) + \big[A\big(x^{[i-1]}(t)\big) - A\big(x^{[i-2]}(t)\big)\big]w$$

and so

$$z - w = \int_{t_0}^{t} \Phi^{[i-1]}(t, s)\big[A\big(x^{[i-1]}(s)\big) - A\big(x^{[i-2]}(s)\big)\big]w(s)\,\mathrm{d}s,$$

i.e.,

$$\begin{aligned}\|z - w\| &\leqslant \int_{t_0}^{t} \exp\bigg(\int_{s}^{t} \mu\big(A\big(x^{[i-1]}(\tau)\big)\big)\,\mathrm{d}\tau\bigg) \exp\bigg(\int_{t_0}^{s} \mu\big(A\big(x^{[i-2]}(\tau)\big)\big)\,\mathrm{d}\tau\bigg)\\ &\qquad \times \alpha\big\|x^{[i-1]}(s) - x^{[i-2]}(s)\big\|\,\mathrm{d}s\\ &\leqslant \exp\big(\mu(t - t_0)\big)\alpha(t - t_0) \sup_{s\in[t_0,t]} \big\|x^{[i-1]}(s) - x^{[i-2]}(s)\big\|. \qquad \square\end{aligned}$$

We first consider the more general equation

$$\begin{aligned}&\dot{x}^{[0]}(t) = A(x_0)x^{[0]}(t) + C(x_0)x^{[0]}(t), \quad x^{[0]}(0) = x_0,\\ &\dot{x}^{[i]}(t) = A\big(x^{[i-1]}(t)\big)x^{[i]}(t) + C\big(x^{[i-1]}(t)\big)x^{[i]}(t), \quad x^{[i]}(0) = x_0, \quad i \geqslant 1.\end{aligned} \tag{3.1}$$

Then we have

$$\begin{aligned}x^{[i]}(t) - x^{[i-1]}(t) = {}&\big(\Phi^{[i-1]}(t, 0) - \Phi^{[i-2]}(t, 0)\big)x_0\\ &+ \int_{0}^{t} \Phi^{[i-1]}(t, s)C\big(x^{[i-1]}(s)\big)\big(x^{[i]}(s) - x^{[i-1]}(s)\big)\,\mathrm{d}s\\ &+ \int_{0}^{t} \Phi^{[i-1]}(t, s)\big(C\big(x^{[i-1]}(s)\big) - C\big(x^{[i-2]}(s)\big)\big)x^{[i-1]}(s)\,\mathrm{d}s\\ &+ \int_{0}^{t} \big(\Phi^{[i-1]}(t, s) - \Phi^{[i-2]}(t, s)\big)C\big(x^{[i-2]}(s)\big)x^{[i-1]}(s)\,\mathrm{d}s\end{aligned}$$

and so, under the conditions (A1)–(A4) we have

$$\begin{aligned}\xi^{[i]}(t) \leqslant \alpha t\, e^{\mu_0 t}\xi^{[i-1]}(t)\|x_0\| &+ \int_0^t e^{\mu_0(t-s)}\gamma\xi^{[i]}(s)\,ds\\ &+ \|x_0\|\int_0^t e^{\mu_0(t-s)}\beta\xi^{[i-1]}(s)e^{(\mu_0+\gamma)s}\,ds\\ &+ \|x_0\|\int_0^t \alpha(t-s)e^{\mu_0(t-s)}\xi^{[i-1]}(t)\gamma\, e^{(\mu_0+\gamma)s}\,ds,\end{aligned}$$

where

$$\xi^{[i]}(t) = \sup_{\tau\in[0,t]}\left\|x^{[i]}(t) - x^{[i-1]}(t)\right\|$$

since

$$\left\|x^{[i]}(t)\right\| \leqslant e^{(\mu_0+\gamma)t}\|x_0\|$$

from (3.1). Hence,

$$\begin{aligned}&\left(1-\gamma\int_0^t e^{\mu_0(t-s)}\,ds\right)\xi^{[i]}(t) \leqslant \alpha t\, e^{\mu_0 t}\xi^{[i-1]}(t)\|x_0\|\\ &\quad+ \|x_0\|\xi^{[i-1]}(t)\left\{\int_0^t e^{\mu_0(t-s)}\beta\, e^{(\mu_0+\gamma)s}\,ds + \int_0^t \alpha(t-s)e^{\mu_0(t-s)}\gamma\, e^{(\mu_0+\gamma)s}\,ds\right\}\end{aligned}$$

and so

$$\xi^{[i]}(t) \leqslant \lambda(t)\xi^{[i-1]}(t),$$

where

$$\lambda(t) = \|x_0\|\left[\alpha t\, e^{\mu_0 t} + \left(\frac{\beta}{\gamma}+\alpha\right)\left(e^{(\mu_0+\gamma)t} - e^{\mu_0 t}\right)\right]\Big/\left(1-\gamma\int_0^t e^{\mu_0(t-s)}\,ds\right) \tag{3.2}$$

and so, if $|\lambda(t)| < 1$ for $t \in [0, T]$ we have

$$x^{[i]}(t) \to x(t) \quad \text{on } C\big([t_0, T], \mathbb{R}^n\big).$$

Moreover, it can easily be seen that

$$\sup_{t\in[0,T]}\left\|x(t)\right\| \leqslant \frac{1}{1-\nu}\xi^{[2]}(t) + \sup_{t\in[0,T]}\left\|x^{[1]}(t)\right\|, \tag{3.3}$$

where

$$\nu = \sup_{t\in[0,T]}\left|\lambda(t)\right|.$$

Lemma 3.2. If $\bar{\mu} = \mu(A(x_0))$, then, under the assumptions (A1)–(A4) we have

$$\left\|x^{[1]}(t)\right\| \leqslant e^{(\bar{\mu}+\gamma)t}\|x_0\|$$

and

$$\left\|x^{[2]}(t) - x^{[1]}(t)\right\| \leqslant \left(\mathrm{e}^{(\mu_0+\gamma)t} + \mathrm{e}^{(\bar{\mu}+\gamma)t}\right)\|x_0\| \leqslant 2\mathrm{e}^{(\mu_0+\gamma)t}\|x_0\|$$

(since $\bar{\mu} \leqslant \mu_0$).

Proof. This follows from Gronwall's lemma and the inequality

$$\left\|x^{[2]}(t)\right\| \leqslant \mathrm{e}^{(\mu_0+\gamma)t}\|x_0\|. \qquad \square$$

Corollary 3.1. Under the assumptions (A1)–(A4) we have

$$\sup_{t\in[0,T]}\left\|x(t)\right\| \leqslant \left(\frac{2}{1-\nu}+1\right)\mathrm{e}^{(\mu_0+\gamma)t}\|x_0\|.$$

Proof. This follows from the above results. $\square$

In the case of the controlled sequence (2.5) we have

$$C(x) = -B(x)R^{-1}B^{\mathrm{T}}(x)P(x)$$

and so

$$\left\|C(x)\right\| \leqslant \left\|B(x)\right\|^2\left\|R^{-1}\right\|\left\|P(x)\right\|.$$

We can get a bound on $\|P^{[i]}(x)\|$ directly from (2.4) (and a bound on $\|P^{[i]}(x) - P^{[i]}(y)\|$ in a similar way). In fact, if we put $\rho = \|P^{[i]}(x)\|$, then ρ satisfies an inequality of the form

$$\dot{\rho} \leqslant c_1 + c_2\rho + c_3\rho^2, \quad \rho(t_f) = \|F\|$$

for some constants $c_1, c_2, c_3 > 0$. Hence for some $\tau \in [0, t_f)$, we have $\rho(t) < \|F\| + \varepsilon$ for all $t \in [\tau, t_f)$ for any $\varepsilon > 0$. Hence so on $[\tau, t_f]$,

$$\dot{\rho} \leqslant c_1 + \left(c_2 + c_3\left(\|F\| + \varepsilon\right)\right)\rho$$

so that

$$\rho \leqslant \mathrm{e}^{(c_2+c_3(\|F\|+\varepsilon))(t_f-t)}\|F\| + c_1\int_t^{t_f}\mathrm{e}^{(c_2+c_3(\|F\|+\varepsilon))(t_f-s)}\,\mathrm{d}s. \tag{3.4}$$

Hence, for small enough t_f (or small enough $c_2,\ c_3$) it follows that $\|P^{[i]}(x)\|$ is bounded on $[0, t_f]$. Hence we have

Theorem 3.1. If t_f (or x_0) is small enough, then under the conditions (A1)–(A2) and

(A3)′ $\|B(x) - B(y)\| \leqslant \beta\|x - y\| \quad \forall x, y \in \mathbb{R}^n$;

(A4)′ $\|B(x)\| \leqslant \gamma \quad \forall x \in \mathbb{R}^n$

approximations $x^{[i]}$ and $u^{[i]}$ converge (in $C(0, t_f; \mathbb{R}^n)$) to functions $x(t),\ u(t)$ which minimise (2.2) over the set of feedback controls of the form $-B(x)R^{-1}B^{\mathrm{T}}(x)P(x)$. $\square$

Remark 3.1. We can replace assumptions (A1)–(A4) by local versions, so the results are not as restrictive as might appear.

Remark 3.2. An actual upper bound on t_f is somewhat involved, but can be shown to be given by the requirement that $\sup_{t\in[0,t_f]}(\lambda(t)) < 1$ where λ is given by (3.2) and by the condition that the right-hand side of (3.4) is less than $\|F\| + \varepsilon$ on $[0, t_f]$. Hence, we require

$$\mathrm{e}^{\nu t_f}\|F\| + \frac{c_1}{\nu}\big(\mathrm{e}^{\nu t_f} - 1\big) < \|F\| + \varepsilon,$$

where

$$\nu = \big(c_2 + c_3(\|F\| + \varepsilon)\big)t_f$$

and $c_1,\ c_2,\ c_3$ are easily found from the Riccati equation.

4. Example

In this section we shall apply the theory to the inverted pendulum, given by the equations

$$\begin{aligned}
\dot{x}_1 &= x_2,\\
\dot{x}_2 &= \frac{mrx_4^2\sin x_3}{M + m\sin^2 x_3} - \frac{mg\sin 2x_3}{2(M + m\sin^2 x_3)} + \frac{u}{M + m\sin^2 x_3},\\
\dot{x}_3 &= x_4,\\
\dot{x}_4 &= \frac{(M+m)g\sin x_3}{r(M + m\sin^2 x_3)} - \frac{mx_4^2\sin x_3}{2(M + m\sin^2 x_3)} - \frac{u\cos x_3}{r(M + m\sin^2 x_3)}.
\end{aligned}$$

The equations are easily put in the form (1.1) and so we may apply the theory above. Figures 1, 2 show the state approximations $x^{[0]},\ x^{[1]},\ x^{[2]}$ and the corresponding controls $u^{[0]},\ u^{[1]},\ u^{[2]}$ for the case of the initial angles $20^0,\ 50^0$ respectively. The converged solution is shown with a solid line. The method is seen to converge rapidly to provide very effective control.

5. Global control

In this section we shall consider the equations for the global optimal control (in the case where such a control exists, for example, in coercive problems – of course a unique optimal control may not exist even in this case) by writing down the full necessary

(a)

(b)

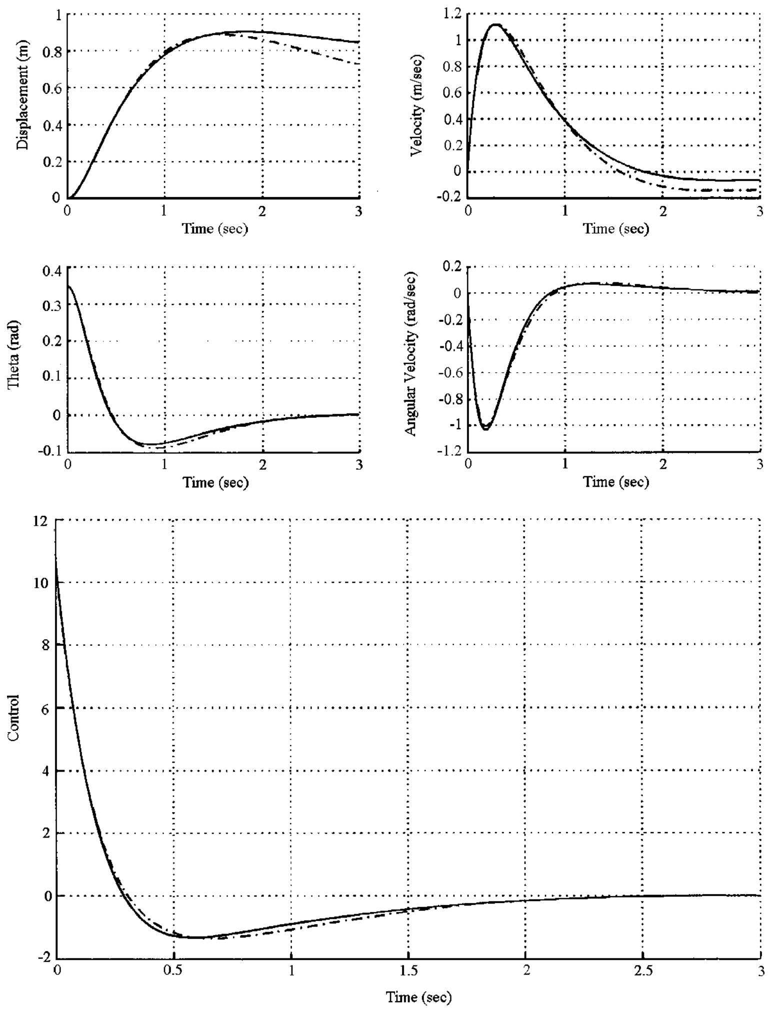

Figure 1. Response of the states of the inverted pendulum to the initial angle $\theta = 20°$, using the method (I), sequence of approximations: (a) Sequence of states $x^{[0]}(t)$, $x^{[1]}(t)$ and $x^{[2]}(t)$; (b) Sequence of control inputs $u^{[0]}(t)$, $u^{[1]}(t)$ and $u^{[2]}(t)$.

conditions for a minimum from Pontryagin's minimum principle. Thus, we consider the optimisation problem

$$\min J = \frac{1}{2} x^{\mathrm{T}}(t_f) F x(t_f) + \frac{1}{2} \int_0^{t_f} \left(x^{\mathrm{T}} Q x + u^{\mathrm{T}} R u \right) \mathrm{d}t$$

subject to the dynamics

$$\dot{x} = A(x)x + B(x)u, \quad x(0) = x_0.$$

(a)

(b)

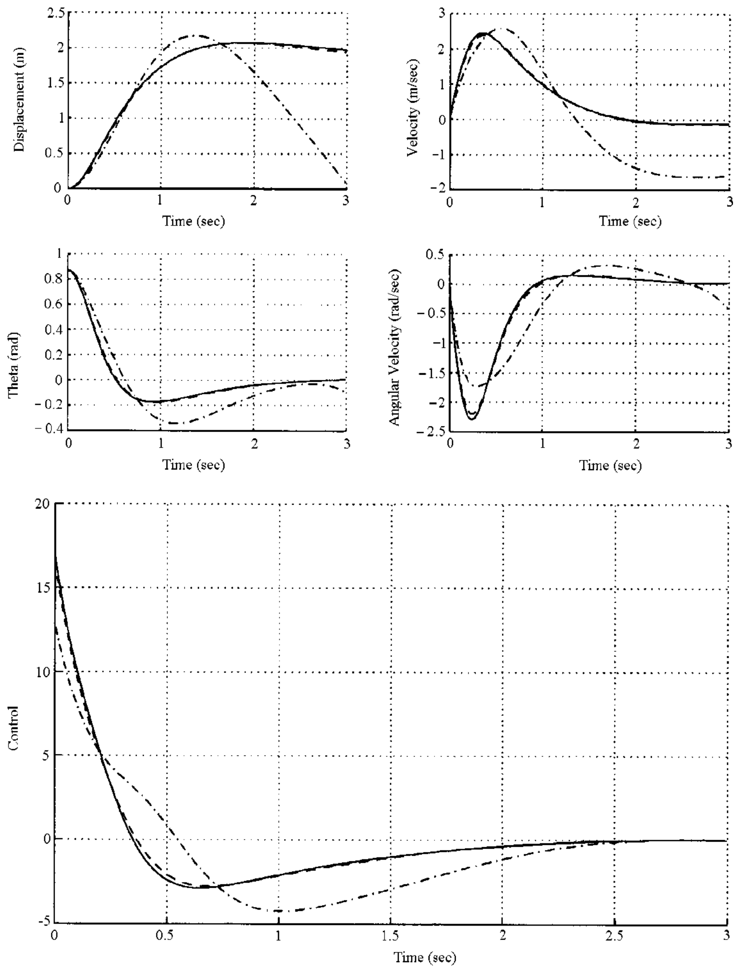

Figure 2. Response of the states of the inverted pendulum to the initial angle $\theta = 50°$, using the method (I), sequence of approximations: (a) Sequence of states $x^{[0]}(t)$, $x^{[1]}(t)$ and $x^{[2]}(t)$; (b) Sequence of control inputs $u^{[0]}(t)$, $u^{[1]}(t)$ and $u^{[2]}(t)$.

The Hamiltonian $\mathcal{H}$ for the problem is

$$\mathcal{H} = \frac{1}{2}\big(x^{\mathrm{T}}Qx + u^{\mathrm{T}}Ru\big) + \lambda^{\mathrm{T}}\big(A(x)x + B(x)u\big)$$

and so the necessary conditions for an optimum are

$$\begin{aligned}
\frac{\partial\mathcal{H}}{\partial\lambda} &= \dot{x} = A(x)x + B(x)u,\\
\frac{\partial\mathcal{H}}{\partial x} &= -\dot{\lambda} = Qx + \frac{\partial}{\partial x}\big(A(x)x\big)^{\mathrm{T}}\lambda + u^{\mathrm{T}}\frac{\partial}{\partial x}B(x)^{\mathrm{T}}\lambda,\\
\frac{\partial\mathcal{H}}{\partial u} &= 0 = Ru + B(x)^{\mathrm{T}}\lambda
\end{aligned}$$

so that

$$u = -R^{-1}B(x)^{\mathrm{T}}\lambda.$$

From these equations we obtain the $2n$-dimensional coupled two-point bvp

2 point boundary values problem

$$\begin{aligned}
\dot{x} &= A(x)x - B(x)R^{-1}B(x)^{\mathrm{T}}\lambda, \quad x(0) = x_0,\\
\dot{\lambda} &= -Qx + C(x,\lambda)\lambda, \quad \lambda(t_f) = Fx(t_f),
\end{aligned}$$

where

$$C(x,\lambda) = -\frac{\partial}{\partial x}\big(A(x)x\big)^{\mathrm{T}} - u^{\mathrm{T}}\frac{\partial}{\partial x}B(x)^{\mathrm{T}}.$$

As before, we introduce the following sequence of approximations to the above equations:

$$\begin{aligned}
\dot{x}^{[0]}(t) &= A(x_0)x^{[0]}(t) - B(x_0)R^{-1}B(x_0)^{\mathrm{T}}\lambda^{[0]}(t), \quad x^{[0]}(0) = x_0,\\
\dot{\lambda}^{[0]}(t) &= -Qx^{[0]}(t) + C(x_0, Fx_0)\lambda^{[0]}(t), \quad \lambda^{[0]}(t_f) = Fx_0
\end{aligned} \tag{5.1}$$

and

$$\begin{aligned}
&\dot{x}^{[i]}(t) = A\big(x^{[i-1]}(t)\big)x^{[i]}(t) - B\big(x^{[i-1]}(t)\big)R^{-1}B\big(x^{[i-1]}(t)\big)^{\mathrm{T}}\lambda^{[i]}(t),\\
&x^{[i]}(0) = x_0,\\
&\dot{\lambda}^{[i]}(t) = -Qx^{[i]}(t) + C\big(x^{[i-1]}(t), \lambda^{[i-1]}(t)\big)\lambda^{[i]}(t), \quad \lambda^{[i]}(t_f) = Fx^{[i-1]}(t_f).
\end{aligned} \tag{5.2}$$

If $z = [x,\lambda]^{\mathrm{T}}$, then we have the coupled two-point bvp

$$\dot{z}^{[i]}(t) = \bar{A}\big(z^{[i-1]}(t)\big)z^{[i]}(t), \quad z^{[i]}(t) = \big[x^{[i]}(t), \lambda^{[i]}(t)\big]^{\mathrm{T}},$$

where

$$\bar{A}\big(z^{[i-1]}(t)\big) = \begin{pmatrix} A(x^{[i-1]}(t)) & -B(x^{[i-1]}(t))R^{-1}B(x^{[i-1]}(t))^{\mathrm{T}} \\ -Q & C(x^{[i-1]}(t), \lambda^{[i-1]}(t)) \end{pmatrix}.$$

Since these equations are linear and time-varying it is well-known (see [19]) that, for each i, we can solve the system explicitly for the missing initial values of λ. We shall recall the procedure here, briefly. We have a $2n$-dimensional linear time-varying system of the form

$$\dot{\varsigma}(t) = \Gamma(t)\varsigma(t)$$

with given initial conditions

$$\varsigma_k(0) = c_k, \quad 1 \leqslant k \leqslant n$$

and terminal conditions

$$\varsigma_k(t_f) = d_{k-n}, \quad n+1 \leqslant k \leqslant 2n.$$

Consider the adjoint system

$$\dot{\psi}(t) = -\Gamma^{\mathrm{T}}(t)\psi(t).$$

We have

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=1}^{2n}(\varsigma_k\psi_k)=0$$

and so, integrating, we have

$$\sum_{k=1}^{2n}\varsigma_k(t_f)\psi_k(t_f)-\sum_{k=1}^{2n}\varsigma_k(0)\psi_k(0)=0.$$

Now choose the n sets of final conditions for ψ:

$$\psi_k^{(\ell)}(t_f)=\begin{cases}1 & \text{if } k=\ell,\\ 0 & \text{if } k\neq\ell\end{cases}$$

then it follows easily that the missing initial conditions for ς are given by

$$\begin{pmatrix}\varsigma_{n+1}(0)\\ \vdots\\ \varsigma_{2n}(0)\end{pmatrix}$$
$$=\begin{pmatrix}\psi_{n+1}^{(1)}(0) & \psi_{n+2}^{(1)}(0) & \cdots & \psi_{2n}^{(1)}(0)\\ \cdots & \cdots & \cdots & \cdots\\ \psi_{n+1}^{(n)}(0) & \psi_{n+2}^{(n)}(0) & & \psi_{2n}^{(n)}(0)\end{pmatrix}^{-1}\begin{pmatrix}\varsigma_{n+1}(t_f)-\sum_{i=1}^{n}\varsigma_i(0)\psi_i^{(1)}(0)\\ \vdots\\ \varsigma_{2n}(t_f)-\sum_{i=1}^{n}\varsigma_i(0)\psi_i^{(n)}(0)\end{pmatrix}. \quad (5.3)$$

Then we have

Theorem 5.1. Under the conditions (A1), (A2), (A3)$'$, (A4)$'$ of theorem 3.1, the sequence of systems given by (5.1) and (5.2) converges (in $C(0,t_f;\mathbb{R}^n)$) for sufficiently small t_f or x_0.

Proof. It is easy to see that, since we can solve the two-point bvp for each i explicitly using (5.3), the sequence of functions $\lambda^{[i]}(t)$ is uniformly bounded on $[0,t_f]$ for sufficiently small t_f or x_0. Using the techniques of the last section, we can then show that $x^{[i]}(t)$ converges in $C(0,t_f;\mathbb{R}^n)$, so that $\lambda^{[i]}(t_f)=Fx^{[i]}(t_f)$ converges. Hence the same method then proves that $\lambda^{[i]}(t)$ converges in $C(0,t_f;\mathbb{R}^n)$. □

Example. We re-examine the example in section 4 and apply the full minimum principle as discussed above. We obtain the results shown in figures 3, 4 for the corresponding angles of 20^0, 50^0. It is clear that, in this case, the two methods give very similar results. Since the first method is much easier to compute, it seems that this simple approximation method will give solutions very close to the optimal one (when such an optimal control exists), at least for small enough t_f or x_0.

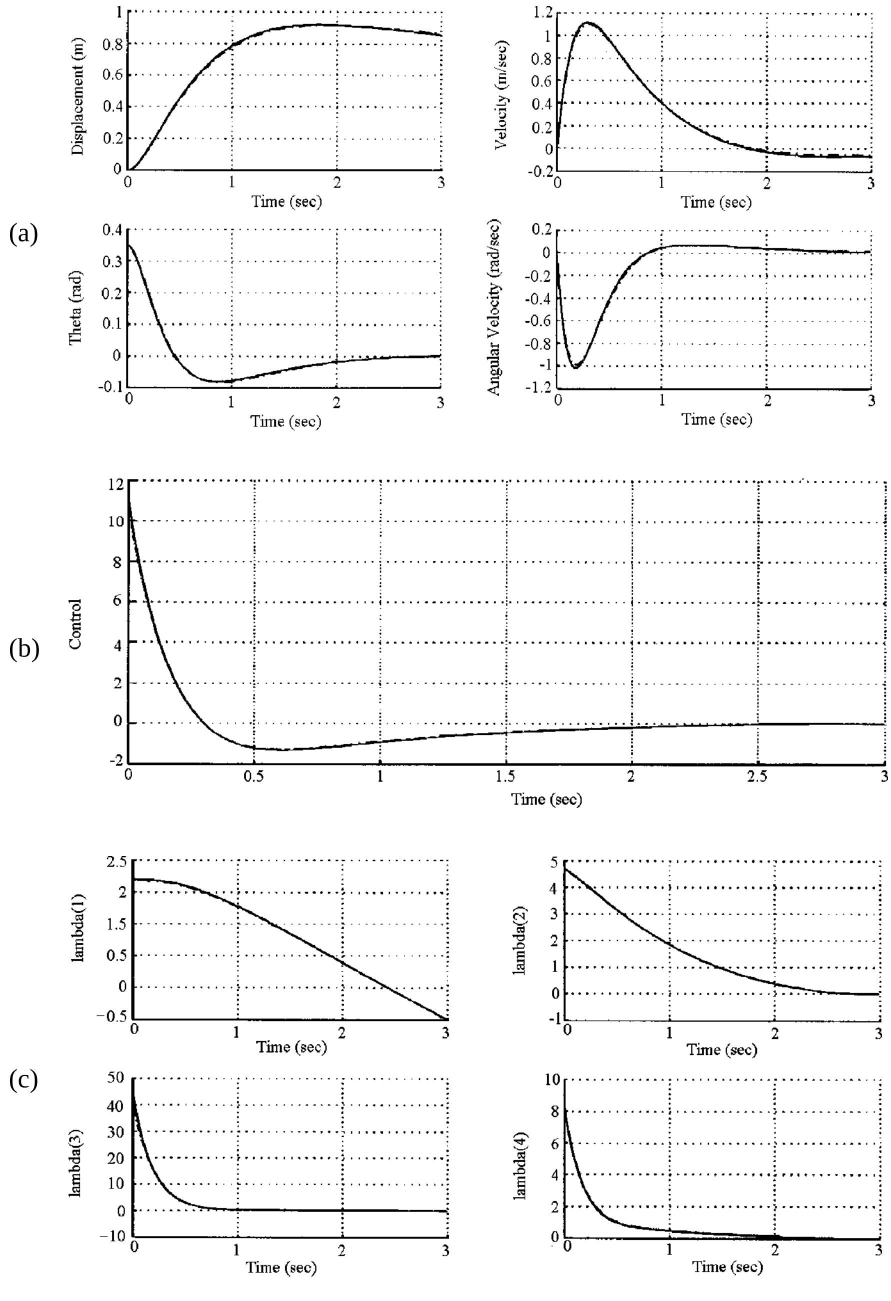

Figure 3. Response of the states of the inverted pendulum to the initial angle $\theta = 20°$, using the method (II), sequence of two point boundary value problems: (a) Sequence of states $x^{[0]}(t)$, $x^{[1]}(t)$ and $x^{[2]}(t)$; (b) Sequence of control inputs $u^{[0]}(t)$, $u^{[1]}(t)$ and $u^{[2]}(t)$; (c) Sequence of states $\lambda^{[0]}(t)$, $\lambda^{[1]}(t)$ and $\lambda^{[2]}(t)$.

(a)

(b)

(c)

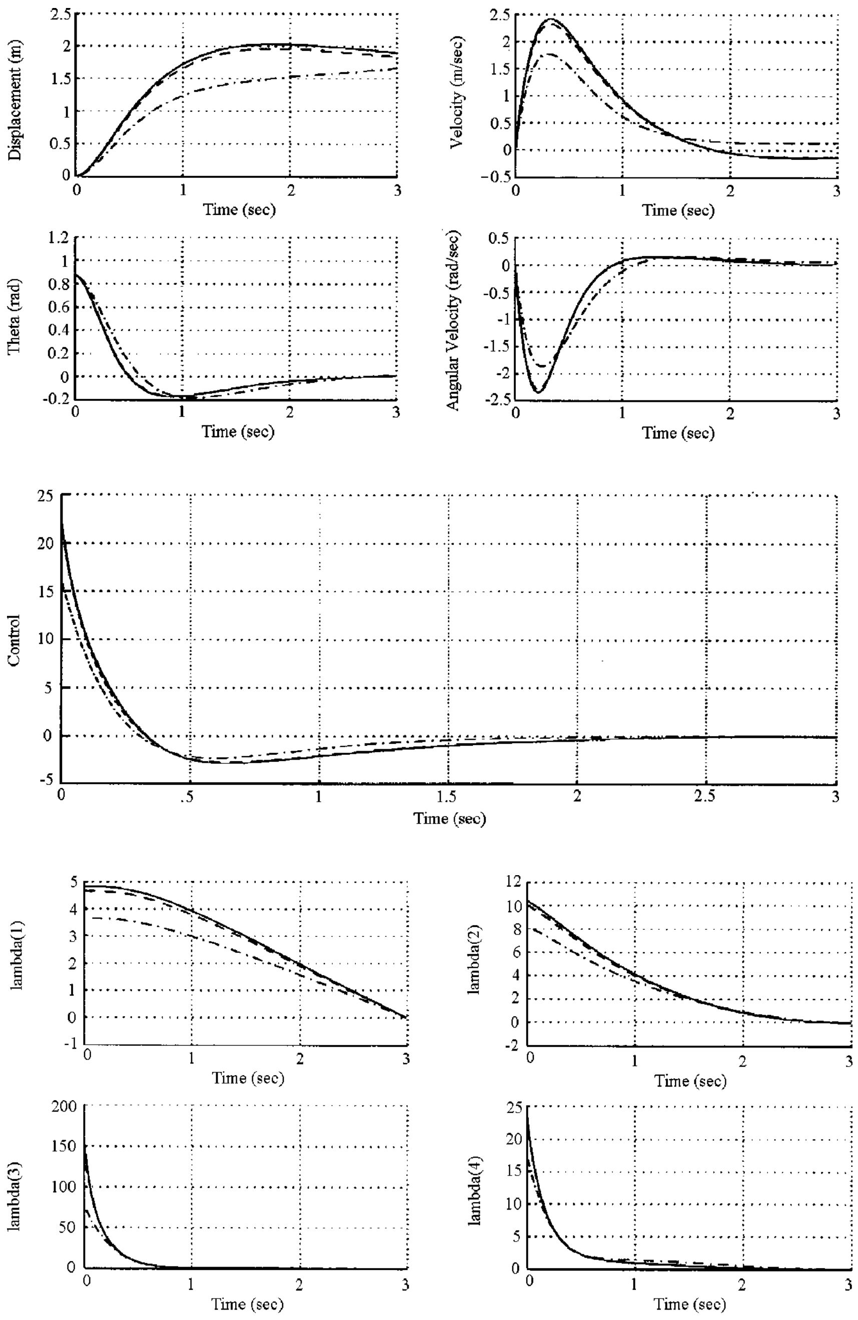

Figure 4. Response of the states of the inverted pendulum to the initial angle $\theta = 50°$, using the method (II), sequence of two point boundary value problems: (a) Sequence of states $x^{[0]}(t)$, $x^{[1]}(t)$ and $x^{[2]}(t)$; (b) Sequence of control inputs $u^{[0]}(t)$, $u^{[1]}(t)$ and $u^{[2]}(t)$; (c) Sequence of states $\lambda^{[0]}(t)$, $\lambda^{[1]}(t)$ and $\lambda^{[2]}(t)$.

6. Parabolic systems

In this section we shall extend the above results to the case of a nonlinear parabolic system of the form

$$\dot{\phi}(t) = -A\big(t, \phi(t)\big)\phi(t) + B\big(t, \phi(t)\big)u(t),$$

where A is an unbounded operator (defined on a Hilbert space H) which satisfies the conditions

(P1) For each ϕ, $\{A(t, \phi)\colon\ t \geqslant 0\}$ is a family of closed, densely defined linear operators in H with domain $D(A(t, \phi))$ independent of t and $\phi \in H$.

(P2) The resolvent $R(\lambda, A(t, \phi))$ (i.e., $(\lambda - A(t, \phi))^{-1}$) exists for all λ with $\operatorname{Re}\lambda \leqslant 0$ and

$$\big\| R\big(\lambda, A(t, \phi)\big)\big\| \leqslant C\big(1 + |\lambda|\big)^{-1}$$

for some constant C.
Thus, $(A(t, \phi))^{-1}$ exists in $L(H)$ and we put $A = A(0, 0)$.

(P3) $A(\cdot, \phi) : [0, \infty) \to L(D(A), H)$ is Hölder continuous.
Let $A^{\alpha}(t, \phi)$ denote the usual fractional power, $0 < \alpha \leqslant 1$. Then we also assume

(P4) $\|A^{-\alpha}(A(t_1, \phi) - A(t_2, \psi))\| \leqslant C_1(|t_1 - t_2|^{\mu} + \|\phi - \psi\|_{\alpha})$ for all $\phi, \psi \in D(A^{\alpha})$ for some constant C_1 and $\mu > 0$, where $\|x\|_{\alpha} = \|A^{\alpha}x\|$.

The following are well-known (see [2,14]):

(i) $D(A^{\alpha}(t, \phi)) \subseteq D(A^{\beta}(t, \phi))$, if $\alpha \geqslant \beta > 0$.

Put $H_{\alpha} = (D(A^{\alpha}), \|\cdot\|_{\alpha})$. Then H_{α} is a Hilbert space and $H_{\beta} \underset{\to}{\subseteq} H_{\alpha}$, for $0 \leqslant \alpha \leqslant \beta \leqslant 1$.

(ii) Let $U(t, \tau; \phi) \in L(H)$ denote the evolution operator generated by $A(t, \phi)$. Then

(a) $U(t, t; \phi) = I$;

(b) $U(t, s) \cdot U(s, \tau) = U(t, \tau),\ 0 \leqslant \tau \leqslant s \leqslant t$;

(c) $\|U(t, \tau; \phi)\|_{\alpha,\beta} \leqslant \frac{\text{const}}{(t-\tau)^{\gamma}}$, if $0 \leqslant \alpha \leqslant \beta \leqslant 1$ and $\beta - \alpha < \gamma < 1$;

(d) $\|U(t, \tau; \phi)\|_{\alpha,\beta} \leqslant$ const if $0 \leqslant \beta < \alpha < 1$;

(e) $\|U(t, \tau; \phi) - U(s, \tau; \phi)\|_{\beta,\alpha} \leqslant \text{const} \cdot |t - s|^{\gamma}$ for $0 \leqslant \gamma < \beta - \alpha$.

Here, $\|\cdot\|_{\alpha,\beta}$ denotes the norm of an operator from H_{α} to H_{β}. For simplicity of exposition, B will be assumed to be bounded, but the results may be extended to boundary control in the usual way (see [13]).

We shall need the following extension of the generalized Gronwall inequality for parabolic PDEs given in [17].

Lemma 6.1. Let $a, b, c \geqslant 0$, $0 < \alpha < 1$, $\alpha + \beta - 1 > 0$, $\gamma > 0$ and $\alpha + \beta + \gamma > 2$ and suppose that $t^{\gamma-1}x(t)$ is locally integrable on $[0, t_f]$ and

$$x(t) \leqslant a + b\int_0^t (t-s)^{\beta-1}s^{\gamma-1}x(s)\,\mathrm{d}s + c\int_0^t (t-s)^{\alpha-1}x(s)\,\mathrm{d}s \quad \text{a.e. on } [0, t_f].$$

Then

$$x(t) \leqslant a\bigl(1 + t_f^{\alpha}/\alpha\bigr)E_{\alpha+\beta-1,\gamma}\bigl[k\Gamma(\alpha+\beta-1)^{1/\nu}t\bigr]$$

for some constant k, where $\nu = \alpha + \beta + \gamma - 2$, $E_{\mu,\gamma}(s) = \sum_{m=0}^{\infty} c_m s^{m\nu}$ with $c_0 = 1$, $c_{m+1}/c_m = \Gamma(m\nu + \gamma)/\Gamma(m\nu + \gamma + \mu)$ for $m \geqslant 0$.

Proof. Put $p(t) = a + b\int_0^t (t-s)^{\beta-1}s^{\gamma-1}x(s)\,\mathrm{d}s$. Then

$$x(t) \leqslant p(t) + c\int_0^t (t-s)^{\alpha-1}x(s)\,\mathrm{d}s.$$

By [17, lemma 7.1.1], we have

$$x(t) \leqslant p(t) + \theta\int_0^t E'_\alpha\bigl(\theta(t-s)\bigr)p(s)\,\mathrm{d}s,$$

where

$$\theta = \bigl[c\Gamma(\alpha)\bigr]^{1/\alpha}, \quad E_\alpha(z) = \sum_{n=0}^{\infty}\frac{z^{n\alpha}}{\Gamma(n\alpha+1)}, \quad E'_\alpha = \frac{\mathrm{d}}{\mathrm{d}z}E_\alpha(z).$$

Since $E'_\alpha(z) \sim z^{\alpha-1}/\Gamma(\alpha)$ as $z \to 0^+$ we see that

$$x(t) \leqslant p(t) + C_1\int_0^t (t-s)^{\alpha-1}p(s)\,\mathrm{d}s$$

for some constant C_1. Now,

$$\begin{aligned}
&\int_0^t (t-s)^{\alpha-1}p(s)\,\mathrm{d}s\\
&\quad = a\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}s + b\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}s\int_0^s (s-\tau)^{\beta-1}\tau^{\gamma-1}x(\tau)\,\mathrm{d}\tau\\
&\quad = a\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}s + b\int_0^t \tau^{\gamma-1}x(\tau)\,\mathrm{d}\tau\int_\tau^t (t-s)^{\alpha-1}(s-\tau)^{\beta-1}\,\mathrm{d}s\\
&\quad = a\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}s + b\int_0^t (t-\tau)^{\alpha+\beta-2}\tau^{\gamma-1}x(\tau)\,\mathrm{d}\tau\int_0^1 \eta^{\beta-1}(1-\eta)^{\alpha-1}\,\mathrm{d}\eta\\
&\quad \leqslant C_2 + C_3\int_0^t (t-\tau)^{\alpha+\beta-2}\tau^{\gamma-1}x(\tau)\,\mathrm{d}\tau
\end{aligned}$$

for new constants $C_2,\ C_3$ ($C_2 = at_f^{\alpha}/\alpha,\ C_3 = \text{beta}(\alpha, \beta)$). Hence

$$\begin{aligned} x(t) &\leqslant a + b\int_0^t (t-s)^{\beta-1}s^{\gamma-1}x(s)\,\mathrm{d}s + C_2 + C_3\int_0^t (t-s)^{\alpha+\beta-2}s^{\gamma-1}x(s)\,\mathrm{d}s \\ &\leqslant (a + C_2) + C_4\int_0^t (t-s)^{\alpha+\beta-2}s^{\gamma-1}x(s)\big[1 + (t-s)^{1-\alpha}\big]\,\mathrm{d}s \\ &\leqslant a_1 + a_2\int_0^t (t-s)^{\alpha+\beta-2}s^{\gamma-1}x(s)\,\mathrm{d}s, \end{aligned}$$

where $a_1 = a(1 + t_f^{\alpha}/\alpha)$, $a_2 = C_4(1 + t_f^{1-\alpha})$. Hence, by [17, lemma 7.1.2] we have

$$x(t) \leqslant a_1 E_{\alpha+\beta-1,\gamma}\Big(\big(a_2\Gamma(\alpha+\beta-1)\big)^{1/\nu}t\Big)$$

and the result follows. □

We shall consider the optimal control problem consisting of the nonlinear equation

$$\dot{\phi}(t) + A\big(t, \phi(t)\big)\phi(t) = B\big(t, \phi(t)\big)u(t), \quad \phi(0) = \psi, \tag{6.1}$$

where, for each ϕ, $A(t, \phi(t))$ satisfies the conditions (P1) to (P4) in section 1, and the quadratic cost functional

$$J = \big\langle \phi(t_f), F\phi(t_f)\big\rangle + \int_0^{t_f} \Big(\big\langle \phi(s), M\phi(s)\big\rangle + \big\langle u(s), Ru(s)\big\rangle\Big)\,\mathrm{d}s, \tag{6.2}$$

where $M, F \in L(H)$ and $R \in L(U)$ for some Hilbert spaces $H,\ U$ and

$$\langle Ru, u\rangle \geqslant \alpha\|u\|^2 \quad \text{for some } \alpha > 0,\ \forall u \in U.$$

We shall find an approximating sequence of controls which converge for sufficiently small t_f. For similar ideas applied to robust stability see [11]. For simplicity of exposition we shall assume that, for each ϕ, $B(t, \phi(t))$ is a bounded operator. By suitable modifications such as those in [13] we can consider boundary control in a similar way.

In order to study this problem we shall consider the following sequence of approximating problems:

$$\dot{\phi}^{[i]}(t) = -A\big(t, \phi^{[i-1]}(t)\big)\phi^{[i]}(t) + B\big(t, \phi^{[i-1]}(t)\big)u^{[i]}(t), \quad \phi^{[i]}(0) = \psi \tag{6.3}$$

subject to

$$J^{[i]} = \big\langle \phi^{[i]}(t_f), F\phi^{[i]}(t_f)\big\rangle + \int_0^{t_f} \Big(\big\langle \phi^{[i]}(s), M\phi^{[i]}(s)\big\rangle + \big\langle u^{[i]}(s), Ru^{[i]}(s)\big\rangle\Big)\,\mathrm{d}s \tag{6.4}$$

for each $i \geqslant 1$, where $\phi^{[0]}$ is defined as the optimal solution of the problem

$$\dot{\phi}^{[0]}(t) = -A(t, \psi)\phi^{[0]}(t) + B(t, \psi)u^{[0]}(t), \quad \phi^{[0]}(0) = \psi$$

subject to

$$J^{[0]} = \langle \phi^{[0]}(t_f), F\phi^{[0]}(t_f) \rangle + \int_0^{t_f} \left(\langle \phi^{[0]}(s), M\phi^{[0]}(s) \rangle + \langle u^{[0]}(s), Ru^{[0]}(s) \rangle \right) \mathrm{d}s.$$

(Note that we are now denoting the inner product by $\langle x, y \rangle$ rather than $x^{\mathrm{T}} y$ in the finite-dimensional case.)

equal

As is well-known (see [13]) the solution of the problem (6.3) and (6.4) is given by

$$u^{[i]}(t) = -R^{-1} B^*\left(t, \phi^{[i-1]}(t)\right) P^{[i]}(t) \phi^{[i]}(t), \tag{6.5}$$

where $P^{[i]}(t)$ is the solution of the “weak” Riccati equation

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \langle P^{[i]}(t)h, k \rangle - \langle P^{[i]}(t)h, A\left(t, \phi^{[i-1]}(t)\right)k \rangle - \langle A\left(t, \phi^{[i-1]}(t)\right)h, P^{[i]}(t)k \rangle + \langle Mh, k \rangle \\ = \langle P^{[i]}(t) B\left(t, \phi^{[i-1]}(t)\right) R^{-1} B^*\left(t, \phi^{[i-1]}(t)\right) P^{[i]}(t)h, k \rangle \end{aligned} \tag{6.6}$$

on $[0, t_f]$, where

$$P^{[i]}(t_f) = F$$

for $h, k \in D(A(t, \phi^{[i-1]}(t)))$. It will be useful to write (6.6) in the integral form

$$\begin{aligned} P^{[i]}(t) = & \int_t^{t_f} U^{[i]*}(s, t) M U^{[i]}(s, t) \, \mathrm{d}s \\ & - \int_t^{t_f} U^{[i]*}(s, t) P^{[i]}(s) B\left(s, \phi^{[i-1]}(s)\right) R^{-1} B^*\left(s, \phi^{[i-1]}(s)\right) P^{[i]}(s) U^{[i]}(s, t) \, \mathrm{d}s, \end{aligned}$$

where $U^{[i]}(t, s)$ is the evolution operator generated by $-A(t, \phi^{[i-1]}(t))$. With the control (6.5), equation (6.3) becomes

$$\dot{\phi}^{[i]}(t) = -A\left(t, \phi^{[i-1]}(t)\right)\phi^{[i]}(t) - B\left(t, \phi^{[i-1]}(t)\right) R^{-1} B^*\left(t, \phi^{[i-1]}(t)\right) P^{[i]}(t) \phi^{[i]}(t). \tag{6.7}$$

First we use (6.7) to put a bound on $\| P^{[i]}(t) \|$, $0 \leqslant t \leqslant t_f$. In fact, we shall require a bound on $\| P^{[i]}(t) \|_{\alpha,0}$. However, since $H_\alpha \underset{\to}{\subseteq} H_0$ we have

$$\| P^{[i]}(t) \|_{\alpha,0} \leqslant C \| P^{[i]}(t) \|$$

for some constant C. Now, by (6.7) we have

$$\begin{aligned} \rho_i(t) \leqslant \| F \| + & \int_t^{t_f} \left\| U^{[i]}(s, t) \right\|^2 \| M \| \, \mathrm{d}s \\ & + \int_t^{t_f} \left\| U^{[i]}(s, t) \right\|^2 \left\| B\left(s, \phi^{[i-1]}(s)\right) \right\|^2 \| R^{-1} \| \rho_i^2(s) \, \mathrm{d}s, \end{aligned}$$

where $\rho_i(t) = \| P^{[i]}(t) \|$. From the assumptions on $A(t, \phi)$ we shall show that

$$\left\| U^{[i]}(s, t) \right\| \leqslant K \tag{6.8}$$

(independent of i) for $\|\phi^{[i-1]}(t) - \phi_0\|_{C([0,t_f],H_\alpha)} \leqslant N$, for any given N. (Of course, K depends on N.) Thus,

$$\rho_i(t) \leqslant \|F\| + t_f K^2\|M\| + K^2\alpha \int_t^{t_f} \rho_i^2(s)\,\mathrm{d}s,$$

where

$$\big\|B\big(s,\phi^{[i-1]}(s)\big)\big\|^2\big\|R^{-1}\big\| \leqslant \alpha.$$

Let $V = \|F\| + t_f K^2\|M\|$. Then for all $t \leqslant t_f$ for which $\rho_i(t) < 2V$ we have

$$\rho_i(t) \leqslant V + 2K^2\alpha V \int_t^{t_f} \rho_i(s)\,\mathrm{d}s$$

so $\rho_i(t) \leqslant V\mathrm{e}^{2K^2\alpha V(t_f-t)}$ and so if

$$\mathrm{e}^{2K^2\alpha V t_f} < 2 \tag{6.9}$$

we have the bound $\|P^{[i]}(t)\|_{\alpha,0} \leqslant 2(\|F\| + t_f K^2\|M\|)$. In order to prove (6.9) note that $U^{[i]}(t,0)$ satisfies the equation $\dot\psi = -A(t,\phi^{[i-1]}(t))\psi,\ \psi(s,s) = I$. Then

$$\dot\psi(t) = -A(0,\phi_0)\psi(t) - \big(A\big(t,\phi^{[i-1]}(t)\big) - A(0,\phi_0)\big)\psi(t)$$

and so

$$\psi(t) = U^0(t,0)\psi(0) - \int_0^t U^0(t,s)\big(A\big(s,\phi^{[i-1]}(s)\big) - A(0,\phi_0)\big)\psi(s)\,\mathrm{d}s,$$

where $U^0(t,0)$ is the evolution operator generated by $-A(0,\phi_0)$. Thus,

$$\begin{aligned}
\big\|\psi(t)\big\| \leqslant{}& \big\|U^0(t,0)\big\|\big\|\psi(0)\big\| + \int_0^t \big\|U^0(t,s)\big\|_1\big\|A^{-1}\big(A\big(s,\phi^{[i-1]}(s)\big) \\
&- A(0,\phi_0)\big)\big\|\big\|\psi(s)\big\|\,\mathrm{d}s \\
\leqslant{}& \big\|U^0(t,0)\big\|\big\|\psi(0)\big\| + \int_0^t \frac{C}{(t-s)^\gamma}\big(s^\mu + \big\|\phi^{[i-1]}(s) - \phi_0\big\|_\alpha\big)\,\mathrm{d}s \\
\leqslant{}& \big\|U^0(t,0)\big\|\big\|\psi(0)\big\| + \int_0^t \frac{C}{(t-s)^\gamma}\big(s^\mu + N\big)\,\mathrm{d}s
\end{aligned}$$

and the result follows from the generalized Gronwall inequality above. A similar argument proves (6.9) for general t and s.

Next, from (6.6),

$$\begin{aligned}
&P^{[i]}(t) - P^{[i-1]}(t) \\
&\quad= \int_t^{t_f} U^{[i]*}(s,t)\big(A^*\big(t,\phi^{[i-1]}\big) - A^*\big(t,\phi^{[i-2]}\big)\big)P^{[i-1]}(s)U^{[i]}(s,t)\,\mathrm{d}s \\
&\qquad+ \int_t^{t_f} U^{[i]*}(s,t)P^{[i-1]}(s)\big(A\big(t,\phi^{[i-1]}\big) - A\big(t,\phi^{[i-2]}\big)\big)U^{[i]}(s,t)\,\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&+\int_t^{t_f}\big(P^{[i]}(s)-P^{[i-1]}(s)\big)B\big(s,\phi^{[i-1]}\big)R^{-1}B^*\big(s,\phi^{[i-1]}\big)P^{[i]}(s)\,\mathrm{d}s\\
&+\int_t^{t_f}P^{[i-1]}(s)\big(B\big(s,\phi^{[i-1]}\big)-B\big(s,\phi^{[i-2]}\big)\big)R^{-1}B^*\big(s,\phi^{[i-1]}\big)P^{[i]}(s)\,\mathrm{d}s\\
&+\int_t^{t_f}P^{[i-1]}(s)B\big(s,\phi^{[i-2]}\big)R^{-1}\big(B^*\big(s,\phi^{[i-1]}\big)-B^*\big(s,\phi^{[i-2]}\big)\big)P^{[i]}(s)\,\mathrm{d}s\\
&+\int_t^{t_f}P^{[i-1]}(s)B\big(s,\phi^{[i-2]}\big)R^{-1}B^*\big(s,\phi^{[i-2]}\big)\big(P^{[i]}(s)-P^{[i-1]}(s)\big)\,\mathrm{d}s
\end{aligned}
$$

and a similar proof to the bound on $P^{[i]}$ shows that $\|P^{[i]}(t)-P^{[i-1]}(t)\|\leqslant\Xi$, $t\in[0,t_f]$ for some Ξ independent of i.

We now proceed to show that $\{\phi^{[i]}\}$ is a Cauchy sequence in $C([0,t_f],H_\alpha)$. First note that it is easy to see that

$$\big\|\phi^{[i]}(t)\big\|_\alpha\leqslant Ct^{\eta-1}$$

for some constant C and some $\eta\in(0,1)$. Now,

$$
\begin{aligned}
\dot\phi^{[i]}(t)-\dot\phi^{[i-1]}(t)=&-A\big(t,\phi^{[i-1]}(t)\big)\big[\phi^{[i]}(t)-\phi^{[i-1]}(t)\big]\\
&-\big(A\big(t,\phi^{[i-1]}(t)\big)-A\big(t,\phi^{[i-2]}(t)\big)\big)\phi^{[i-1]}(t)\\
&-B\big(t,\phi^{[i-1]}(t)\big)R^{-1}B^*\big(t,\phi^{[i-1]}(t)\big)P^{[i]}(t)\phi^{[i]}(t)\\
&+B\big(t,\phi^{[i-2]}(t)\big)R^{-1}B^*\big(t,\phi^{[i-2]}(t)\big)P^{[i-1]}(t)\phi^{[i-1]}(t)
\end{aligned}
$$

and so

$$
\begin{aligned}
&\big\|\phi^{[i]}(t)-\phi^{[i-1]}(t)\big\|_\alpha\\
&\quad\leqslant\int_0^t\big\|U^{[i-1]}(t,\tau)\big[A\big(\tau,\phi^{[i-1]}(\tau)\big)-A\big(\tau,\phi^{[i-2]}(\tau)\big)\big]\phi^{[i-1]}(\tau)\,\mathrm{d}\tau\big\|_\alpha\\
&\qquad+\left\|\int_0^t U^{[i-1]}(t,\tau)\big[F^{[i]}(\tau)-F^{[i-1]}(\tau)\big]\,\mathrm{d}\tau\right\|_\alpha,
\end{aligned}
$$

where

$$F^{[i]}(t)=-B\big(t,\phi^{[i-1]}(t)\big)R^{-1}B^*\big(t,\phi^{[i-1]}(t)\big)P^{[i]}(t)\phi^{[i]}(t)$$

and so

$$\big\|F^{[i]}(t)-F^{[i-1]}(t)\big\|\leqslant C_1+C_2\big\|\phi^{[i]}(t)-\phi^{[i-1]}(t)\big\|_\alpha$$

for some new constants $C_1,\ C_2$ (depending on t_f and the other bounds above). Hence,

$$
\begin{aligned}
\big\|\phi^{[i]}(t)-\phi^{[i-1]}(t)\big\|_\alpha\leqslant&\int_0^t\frac{C_3t^{\eta-1}}{(t-\tau)^\alpha}\big\|\phi^{[i-1]}(\tau)-\phi^{[i-2]}(\tau)\big\|_\alpha\,\mathrm{d}\tau\\
&+\int_0^t\frac{(C_1+C_2)}{(t-\tau)^\alpha}\big\|\phi^{[i-1]}(\tau)-\phi^{[i-2]}(\tau)\big\|_\alpha\,\mathrm{d}\tau \qquad (6.10)
\end{aligned}
$$

for some new constant C_3. By the generalized Gronwall inequality, we see that

$$\sup_{t\in[0,t_f]} \left\| \phi^{[i]}(t) - \phi^{[i-1]}(t) \right\|_\alpha \leqslant K \sup_{t\in[0,t_f]} \left\| \phi^{[i-1]}(t) - \phi^{[i-2]}(t) \right\|_\alpha,$$

where K is a constant such that $K \to 0$ as $t_f \to 0$. Hence, if t_f is small enough we find that $\{\phi^{[i]}\}$ is indeed a Cauchy sequence. The limit function $\phi(t) \in C([0, t_f], H_\alpha)$ is an "almost optimal" control in the sense that there is an arbitrarily close approximation $\bar{\phi}(t)$ which satisfies an equation "close" to (6.3) and minimizes (6.4).

As an example, let

$$-A(x, t, \phi, \nabla\phi)\phi = \sum_{i,j=1}^{n} a_{ij} D_i D_j \phi + \sum_{i=1}^{n} b_i(t, x, \phi, \nabla\phi) D_i \phi + c(t, x, \phi, \nabla\phi)\phi,$$

where $a_{ij},\ b_i,\ c$ belong to $C^{\nu/2,\nu,1,1}([0, t_f] \times \overline{\Omega} \times \mathbb{R} \times \mathbb{R}^n)$ where $0 < \nu < 1$ and

$$\sum_{i,j=1}^{n} a_{ij} \eta_i \eta_j \geqslant \mathrm{const}\|\eta\|^2$$

for all $\eta \in \mathbb{R}^n$. Also let $B(t, \phi)$ be the nonlinear integral operator

$$B(t, \phi)\phi = \int k\big(x, \phi(x)\big)\phi(x)\, \mathrm{d}\Omega,$$

where $k(\cdot, \phi) \in L^2(\Omega)$ for each ϕ and

$$\left\| k(\cdot, \phi) - k(\cdot, \psi) \right\|_{L^2} \leqslant \|\phi - \psi\|_{L^2}$$

for each $x \in \Omega$. Then standard elliptic estimates show that the conditions of the above theory hold and so the system

$$\dot{\phi}(t) = -A(x, t, \phi, \nabla\phi)\phi + \int k\big(x, \phi(x)\big)\phi(x)\, \mathrm{d}\Omega$$

has an almost optimal solution with respect to the cost functional (6.4) for sufficiently small t_f. This time t_f is one which satisfies (6.9) and is such that

$$K = \left(\sup_{t\in[0,t_f]} C_3 \frac{t_f^{\eta-1}}{(t_f - t)^\alpha} + \sup_{t\in[0,t_f]} \frac{C_1 + C_2}{(t_f - t)^\alpha} \right) < 1$$

for the constants $C_1,\ C_2,\ C_3$ in (6.10).

7. Exponential stability of large-scale nonlinear systems with multiple time delays

By way of an illustration, we shall show that the same method applies to almost any nonlinear problem. Here we shall consider the exponential stability of a large-scale nonlinear parabolic system with multiple time delays. We shall see that, by reducing the problem to a sequence of linear, time-varying ones, we may again obtain results for nonlinear systems easily from the corresponding ones for linear systems. Thus, consider the large scale nonlinear distributed systems with delay of the form

$$\begin{aligned}\dot{x}(t) &= -A\big(x(t),t\big)x(t) + B\big(x(t),t\big)u(t)\\ &\quad + \sum_{k=1}^{L} C^k\big(x(t),t\big)x\big(t-\tau_k(t)\big) + f\big(t,x(t),x\big(t-\tau_k(t)\big)\big),\\ x_i(t) &= \phi_i(t), \quad -\infty < t \leqslant t_0,\ i=1,\dots,m.\end{aligned} \tag{7.1}$$

Here, $x(t)$ is of the form $x(t) = (x_1(t),\dots,x_m(t))^{\mathrm{T}}$, where $x_i = (x_1^i,\dots,x_{n_i}^i)^{\mathrm{T}}$ is the state of the i-th isolated subsystem given by

$$\dot{x}_i(t) = -A_{ii}\big(x_i(t),t\big)x_i(t) + B_{ii}\big(x_i(t),t\big)u_i(t),$$

where $u_i = (u_1^i,\dots,u_{r_i}^i)^{\mathrm{T}}$. Stabilization will be achieved by finding a feedback control of the form

$$u_i(t) = K_{ii}\big(x_i(t),t\big)x_i(t)$$

so that the system

$$\dot{x}_i(t) = -A_{ii}\big(x_i(t),t\big)x_i(t) + B_{ii}\big(x_i(t),t\big)K_{ii}\big(x_i(t),t\big)x_i(t) \tag{7.2}$$

is stable. The resulting operator $-A_{ii}(x_i(t),t) + B_{ii}(x_i(t),t)K_{ii}(x_i(t),t)$ generates a nonlinear evolution operator which can be approximated arbitrarily closely by the systems

$$\dot{x}_i^{[k]}(t) = -A_{ii}\big(x_i^{[k-1]}(t),t\big)x_i^{[k]}(t) + B_{ii}\big(x_i^{[k-1]}(t),t\big)K_{ii}\big(x_i^{[k-1]}(t),t\big)x_i^{[k]}(t).$$

Details of the convergence of this sequence are as before, if we assume the conditions

$$\begin{aligned}&\big\|A_{ij}\big(x(t),t\big)\big\|_\alpha \leqslant a_{ij},\\ &\big\|B_{ij}\big(x(t),t\big)\big\|_\alpha \leqslant b_{ij},\\ &\big\|C_{ij}^k\big(x(t),t\big)\big\|_\alpha \leqslant c_{ij}^k,\\ &\big\|f_i\big(t,x(t),x\big(t-\tau_k(t)\big)\big)\big\| \leqslant \sum_{j=1}^{m} e_{ij}\big\|x_j(t)\big\|_\alpha + \sum_{k=1}^{K}\sum_{j=1}^{m} d_{ij}^k\big\|x_j\big(t-\tau_k(t)\big)\big\|_\alpha.\end{aligned}$$

However, as we are going to prove stability, the convergence holds for all t.

Consider the ith isolated subsystem given by

$$\dot{x}_i(t) = -A_{ii}\big(x_i(t),t\big)x_i(t) + B_{ii}\big(x_i(t),t\big)u_i(t).$$

For simplicity of exposition, we shall assume that $B_{ii}(x(t),t)$ is a bounded operator for all $x(t)$, t. However, the theory is easily extended to the case of unbounded B (i.e., boundary control) in the usual way, as stated above. We shall assume that the system

$$\dot{x}_i(t) = -A_{ii}(0,0)x_i(t) + B_{ii}(0,0)u_i(t) \tag{7.3}$$

is stabilizable with (linear) stabilizing feedback

$$u_i(t) = K_{ii}(0,0)x_i(t)$$

so that $K_{ii}(0,0) \in L(L^2(\Omega))$ and the operator $-A_{ii}(0,0) + B_{ii}(0,0)K_{ii}(0,0)$ generates an analytic semigroup with

$$\begin{aligned} \|T_{ii}(t)\| &\leqslant M\mathrm{e}^{a_{ii}t} \\ \|A_{ii}(0,0)T_{ii}(t)\| &\leqslant M\mathrm{e}^{a_{ii}t}/t, \quad t > 0, \end{aligned}$$

where $a_{ii} < 0$.

We have

$$\begin{aligned} \dot{x}_i(t) = {} & -A_{ii}(0,0)x_i(t) + B_{ii}(0,0)u_i(t) + \big(-A_{ii}\big(x_i(t),t\big) + A_{ii}(0,0)\big)x_i(t) \\ & + \big(B_{ii}\big(x_i(t),t\big) - B_{ii}(0,0)\big)u_i(t) \end{aligned}$$

so that

$$\begin{aligned} x_i(t) = {} & T_{ii}(t)x_{i0} + \int_0^t T_{ii}(t-s)\big(-A_{ii}\big(x_i(s),s\big) + A_{ii}(0,0)\big)x_i(s)\,\mathrm{d}s \\ & + \int_0^t T_{ii}(t-s)\big(B_{ii}\big(x_i(s),s\big) - B_{ii}(0,0)\big)K_{ii}(0,0)x_i(s)\,\mathrm{d}s, \end{aligned}$$

whence

$$\begin{aligned} \|x_i(t)\|_\alpha \leqslant {} & M\mathrm{e}^{a_{ii}t}\|x_{i0}\|_\alpha + \int_0^t \frac{M\mathrm{e}^{a_{ii}(t-s)}}{(t-s)^\alpha}\big(C_1 s^\mu + C_2\|x_i(s)\|_\alpha\big)\Delta\|x_i(s)\|_\alpha\,\mathrm{d}s \\ & + \int_0^t \frac{M\mathrm{e}^{a_{ii}(t-s)}}{(t-s)^\alpha}\big(C_4 s^\mu + C_5\|x_i(s)\|_\alpha\big)\Delta\|K_{ii}(0,0)\|\,\|x_i(s)\|_\alpha\,\mathrm{d}s \\ = {} & M\mathrm{e}^{a_{ii}t}\|x_{i0}\|_\alpha + M\Delta\big(C_1 + C_4\|K_{ii}(0,0)\|\big)\int_0^t \frac{\mathrm{e}^{a_{ii}(t-s)}}{(t-s)^\alpha}s^\mu\|x_i(s)\|_\alpha\,\mathrm{d}s \\ & + M\Delta\big(C_{21} + C_5\|K_{ii}(0,0)\|\big)\int_0^t \frac{\mathrm{e}^{a_{ii}(t-s)}}{(t-s)^\alpha}\|x_i(s)\|_\alpha\|x_i(s)\|_\alpha\,\mathrm{d}s \end{aligned}$$

for some new constants C_i, Δ. Suppose that $\|x_{i0}\|_\alpha < \lambda < 1$ and that $\|x_i(t)\|_\alpha < 1$ for $t \in [0,T)$ for some λ to be specified. Then, by lemma 6.1, for $t \in [0,T)$, we have

$$\begin{aligned} \|x_i(t)\|_\alpha &\leqslant M\mathrm{e}^{a_{ii}t}\|x_{i0}\|_\alpha\left(1 + C\frac{t^\alpha}{\alpha}\right)E_{\alpha+\beta-1,\gamma}\Big(\big(L\big(1+t^{1-\alpha}\big)\Gamma(\alpha+\beta-1)\big)^{1/\nu}t\Big) \\ &= M\|x_{i0}\|_\alpha v(t) \end{aligned}$$

say, where

$$L = M\Delta(C_1 + C_4)\|K_{ii}(0,0)\|\max\big(1,\mathrm{beta}(\alpha,\beta)\big)$$

and

$$v(t) = \left(1 + C\frac{t^\alpha}{\alpha}\right)E_{\alpha+\beta-1,\gamma}\Big(\big(L\big(1+t^{1-\alpha}\big)\Gamma(\alpha+\beta-1)\big)^{1/\nu}t\Big).$$

Now

$$E_{\alpha+\beta-1,\gamma}(z) \sim z^{1/2(\nu/\beta-\gamma)} \exp\left(\frac{\beta}{\nu} z^{\nu/\beta}\right),$$

and since $\nu/\beta = (\alpha+\beta+\gamma-2)/\beta$ we see that if $-a_{ii} > (\alpha+\gamma-2)/\beta + 1$ then

$$\zeta \doteq \max_{t\in[0,\infty)} v(t)$$

exists. Hence, if we take $\lambda = 1/(M\zeta)$ then $\|x_i(t)\|_\alpha \leqslant 1$ for $t = T$ and we can extend the set of t for which $\|x_i(t)\|_\alpha \leqslant 1$ to $[0,\infty)$. Hence $\|x_i(t)\|_\alpha \to 0$ as $t \to \infty$ and the system is stabilized. Summarizing, we have

Theorem 7.1. If $-a_{ii} > (\alpha+\gamma-2)/\beta + 1$ the system (7.1) can be stabilized for small enough initial conditions. Moreover, a similar proof shows that the nonlinear evolution operator $\widetilde{U}_{ii}(t,s)$ generated by $-A_{ii}(x_i(t),t)x_i(t) + B_{ii}(x_i(t),t)K_{ii}(0,0)$ can be approximated arbitrarily closely by a linear evolution operator $U_{ii}(t,s)$ given by the approximate system

$$\dot{x}_i^{[k]}(t) = -A_{ii}\big(x_i^{[k-1]}(t),t\big)x_i^{[k]}(t) + B_{ii}\big(x_i^{[k-1]}(t),t\big)K_{ii}(0,0)x_i^{[k]}(t).$$

We also have

$$\big\|U_{ii}(t,s)\big\|_\alpha \leqslant M_{ii}\mathrm{e}^{a_{ii}(t-s)}. \qquad \square$$

Remark. A second much simpler condition for the stabilization of (7.2) can be given at the expense of a much stronger condition on A_{ii} and B_{ii} namely, if there exists a uniformly bounded operator $K_{ii}(x_i(t),t)$ for each fixed $(x_i(t),t)$ such that

$$\big\langle x, \big(A_{ii}(y,\tau) + B_{ii}(y,\tau)K_{ii}(y,\tau)\big)x\big\rangle \leqslant -\varepsilon\|x\|^2$$

for each y,τ then the system (7.2) is clearly stabilized by K_{ii}.

We then obtain the main result of this section:

Theorem 7.2. If $K_{ii}(0,0) \in L(L^2(\Omega)) \cap L(D(A^\alpha))$ exists which stabilizes the system (1) and the conditions of lemma 6.1 and theorem 7.1 are satisfied and, moreover, we have

(i) $\lambda_k = \sup_{t\geqslant t_0}\{\dot{\tau}_k(t)\} < 1$;

(ii) $a_{ii} + b_{ii} < 0,\ i = 1,\dots,m$;

(iii) $M_{ii}[(a_{ij} + e_{ij} + \alpha_i b_{ij}) + \sum_{k=1}^{K} v_k(c_{ij}^k + d_{ij}^k)] < 0,\ i,j = 1,\dots,m,$

where $\alpha_i = \|K_{ii}(0,0)\|$ and $v_k = (1-\lambda_k)^{-1} > 0,\ k = 1,\dots,K$, then the zero solution of (7.1) is exponentially stable.

Proof. From the above remarks we may assume that the stabilized isolated system generates a linear evolution operator $U_{ii}(t,s)$. Then we have

$$x_i(t) = U_{ii}(t,t_0)\phi_i(t) + \int_{t_0}^{t} U_{ii}(t,s)\Bigg[\sum_{j\neq i}^{m} A_{ij}\big(x(s),s\big)x_j(s) + \sum_{j\neq i}^{m} B_{ij}\big(x(s),s\big)u_j(s) + \sum_{k=1}^{K}\sum_{j=1}^{m} C_{ij}^{k}\big(x(s),s\big)x_j\big(s-\tau_k(s)\big) + f_i\big(s,x(s),x\big(s-\tau_k(s)\big)\big)\Bigg]\mathrm{d}s.$$

Hence,

$$\begin{aligned}\|x_i(t)\|_\alpha &\leqslant M_{ii}e^{a_{ii}(t-t_0)}\|\phi_i(t)\|_\alpha + \int_{t_0}^{t} M_{ii}e^{a_{ii}(t-s)}\Bigg[\sum_{j\neq i}^{m}\|A_{ij}\big(x(s),s\big)\|_\alpha\|x_j(s)\|_\alpha \\ &\quad + \sum_{j\neq i}^{m}\|B_{ij}\big(x(s),s\big)\|_\alpha\|u_j(s)\|_\alpha \\ &\quad + \sum_{k=1}^{K}\sum_{j=1}^{m}\|C_{ij}\big(x(s),s\big)\|_\alpha\|x_j\big(s-\tau_k(s)\big)\|_\alpha \\ &\quad + \|f_i\big(s,x(s),x\big(s-\tau_k(s)\big)\big)\|_\alpha\Bigg]\mathrm{d}s \\ &\leqslant M_{ii}e^{a_{ii}(t-t_0)}\|\phi_i(t)\|_\alpha + \int_{t_0}^{t} M_{ii}e^{a_{ii}(t-s)}\Bigg[\sum_{j\neq i}^{m}(a_{ij}+e_{ij}+\alpha_i b_{ij})\|x_j(s)\|_\alpha \\ &\quad + e_{ii}\|x_i(s)\|_\alpha + \sum_{k=1}^{K}\sum_{j=1}^{m}\big(c_{ij}^{k}+d_{ij}^{k}\big)\|x_j\big(s-\tau_k(s)\big)\|_\alpha\Bigg]\mathrm{d}s.\end{aligned}$$

Now let $q_i(t)$ be defined by

$$\begin{aligned}q_i(t) &= \|\phi_i(t)\|_\alpha, \quad -\infty < t \leqslant t_0, \\ q_i(t) &= M_{ii}\mathrm{e}^{a_{ii}(t-t_0)}\|\phi_i(t)\|_\alpha + \int_{t_0}^{t} M_{ii}\mathrm{e}^{a_{ii}(t-s)}\Bigg[\sum_{j\neq i}^{m}(a_{ij}+e_{ij}+\alpha_i b_{ij})\|x_j(s)\|_\alpha \\ &\quad + e_{ii}\|x_i(s)\|_\alpha + \sum_{k=1}^{K}\sum_{j=1}^{m}\big(c_{ij}^{k}+d_{ij}^{k}\big)\|x_j\big(s-\tau_k(s)\big)\|_\alpha\Bigg]\mathrm{d}s\end{aligned}$$

for $t \geqslant t_0$.

Then we have

$$\|x_i(t)\|_\alpha \leqslant q_i(t).$$

From the definition of $q_i(t)$ we have

$$\dot{q}_i(t) = a_{ii}q_i(t) + M_{ii}\mathrm{e}^{a_{ii}(t-t_0)}\Bigg[\sum_{j\neq i}^{m}(a_{ij} + e_{ij} + \alpha_i b_{ij})\|x_j(s)\|_\alpha + e_{ii}\|x_i(s)\|_\alpha + \sum_{k=1}^{K}\sum_{j=1}^{m}\big(c_{ij}^k + d_{ij}^k\big)\big\|x_j\big(s - \tau_k(s)\big)\big\|_\alpha\Bigg]\mathrm{d}s$$

so that

$$\dot{q}_i(t) \leqslant M_{ii}\sum_{j=i}^{m}(a_{ij} + e_{ij} + \alpha_i b_{ij}) - \alpha_i b_{ij} q_i(t) + \sum_{k=1}^{K}\sum_{j=1}^{m}\big(c_{ij}^k + d_{ij}^k\big)q_i\big(t - \tau_k(t)\big).$$

Define a Lyapunov (-like) function

$$V_i\big(t, q_i(t)\big) = q_i(t) + \sum_{k=1}^{K} v_k \int_{t-\tau_k(t)}^{t}\sum_{j=1}^{m}\big(c_{ij}^k + d_{ij}^k\big)q_j(s)\,\mathrm{d}s$$

and

$$V\big(t, q(t)\big) = \sum_{i=1}^{m}\mu_i V_i\big(t, q_i(t)\big),$$

where $\mu_i > 0$. Elementary computation shows that

$$\dot{V}\big(t, q(t)\big) \leqslant -\beta\sum_{j=1}^{m}q_j(t),$$

where

$$\beta = -\max_{1\leqslant j\leqslant m}\sum_{i=1}^{m}\mu_i\left[M_{ii}\left\{(a_{ij} + e_{ij} + \alpha_i b_{ij}) + \sum_{k=1}^{K}v_k\big(c_{ij}^k + d_{ij}^k\big)\right\}\right] > 0.$$

The result now follows from Krasovskii's theorem [16]. □

8. Conclusions

In this paper we have introduced a general technique for studying nonlinear problems – that is by replacing their defining equations by sequences of time-varying systems which can be approached by classical means. The sequences are shown to converge under mild assumptions and can be applied in a great variety of situations. Thus, we have shown how to solve finite- and infinite-dimensional nonlinear-quadratic control problems (although the quadratic cost can be replaced by a much more general functional without any extra difficulty) and we have demonstrated its application to a complicated stability problem. The method clearly has many other applications – in fact, any nonlinear problem for which the corresponding problem for time-varying linear systems can be solved is amenable to this method.

References

[1] Z. Aganovic and Z. Gajic, The successive approximation procedure for finite time optimal control of bilinear systems, IEEE Trans. Automat. Control 39(9) (1994) 324–328.

[2] H. Amann, Periodic solutions of semilinear parabolic equations, in: *Nonlinear Analysis*, A collection of papers in honor of Erich Roth (Academic Press, New York, 1978) pp. 1–29.

[3] S.P. Banks and M.K. Yew, On a class of suboptimal controls for infinite-dimensional bilinear systems, Systems Control Lett. 5 (1985) 327–333.

[4] S.P. Banks, On the optimal control of nonlinear systems, Systems Control Lett. 6 (1986) 337–343.

[5] S.P. Banks, Infinite-dimensional Carleman linearisation, the Lie series and optimal control of nonlinear PDEs, Internat. J. Systems Sci. 23 (1992) 663–675.

[6] S.P. Banks and K. Mhana, Optimal control and stabilization for nonlinear systems, IMA J. Control and Inform. 9 (1992) 179–196.

[7] S.P. Banks and S.K. Al-Jurani, Lie algebras and the stability of nonlinear systems, Internat. J. Control 60 (1994) 315–329.

[8] S.P. Banks, A. Moser and D. McCaffrey, Clifford algebras, dynamical systems and periodic orbits, IMA J. Math. Control and Inform. 13 (1996) 279–298.

[9] S.P. Banks and D. McCaffrey, Optimal control of nonlinear distributed parameter systems, Proc. of 17th IMACS World Congress, Berlin (July 1997).

[10] S.P. Banks and D. McCaffrey, Lie algebras, structure of nonlinear systems and chaotic motion, Internat. J. Bifur. and Chaos 8(7) (1998) 1437–1462.

[11] S.P. Banks, A. Moser and D. McCaffrey, Robust exponential stability of evolution equations, Arch. Control Sci. 4(XL) (1995) 261–279.

[12] F. Brauer, Perturbations of nonlinear systems of differential equations, J. Math. Anal. App. 14 (1966) 198–206.

[13] R.F. Curtain and A.J. Pritchard, *Infinite-Dimensional Linear Systems Theory* (Springer, 1978).

[14] A. Friedman, *Partial Differential Equations* (Holt, New York, 1969).

[15] W.L. Garrard and J.M. Jordan, Design of nonlinear automatic flight control systems, Automatica 13 (1977) 497–505.

[16] J.K. Hale, *Theory of Functional Differential Equations* (Springer, New York, 1977).

[17] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math. 840 (Springer, 1981).

[18] S.K. Nguang, Robust nonlinear H_∞-output feedback control, IEEE Trans. Automat. Control 41(7) (1996) 1003–1007.

[19] S.M. Roberts and J.S. Shipman, *Two-Point Boundary Value Problems: Shooting Methods*, Modern Analytic and Computational Methods in Science and Mathematics 31, ed. R. Bellman (Elsevier, New York, 1972).

[20] P. Soravia, H_∞ control of nonlinear systems: Differential games and viscosity solutions, SIAM J. Control Optim. 34(3) (1996) 1071–1097.