J Optim Theory Appl (2011) 149: 26–46
DOI 10.1007/s10957-010-9773-3

Equations for the Missing Boundary Values in the Hamiltonian Formulation of Optimal Control Problems

Vicente Costanza · Pablo S. Rivadeneira · Ruben D. Spies

Published online: 6 January 2011

Abstract Partial differential equations for the unknown final state and initial costate arising in the Hamiltonian formulation of regular optimal control problems with a quadratic final penalty are found. It is shown that the missing boundary conditions for Hamilton's canonical ordinary differential equations satisfy a system of first-order quasilinear vector partial differential equations (PDEs), when the functional dependence of the H-optimal control in phase-space variables is explicitly known. Their solutions are computed in the context of nonlinear systems with $\mathbb{R}^n$-valued states. No special restrictions are imposed on the form of the Lagrangian cost term. Having calculated the initial values of the costates, the optimal control can then be constructed from on-line integration of the corresponding $2n$-dimensional Hamilton ordinary differential equations (ODEs). The off-line procedure requires finding two auxiliary $n \times n$ matrices that generalize those appearing in the solution of the differential Riccati equation (DRE) associated with the linear-quadratic regulator (LQR) problem. In all equations, the independent variables are the finite time-horizon duration T and the final-penalty matrix coefficient S, so their solutions give information on a whole two-parameter family of control problems, which can be used for design purposes. The mathematical treatment takes advantage from the symplectic structure of the Hamiltonian formalism, which allows one to reformulate Bellman's conjectures concerning the "invariant-embedding" methodology for two-point boundary-value problems. Results for LQR problems are tested against solutions of the associ-

Communicated by P.M. Pardalos.

V. Costanza (✉) · P.S. Rivadeneira
Instituto de Desarrollo Tecnológico para la Industria Química INTEC, CONICET-UNL, Güemes 3450, S3000GLN Santa Fe, Argentina
e-mail: tsinoli@santafe-conicet.gov.ar

R.D. Spies
Instituto de Matemática Aplicada del Litoral, IMAL, CONICET-UNL, Güemes 3450, S3000GLN Santa Fe, Argentina

ated differential Riccati equation, and the attributes of the two approaches are illustrated and discussed. Also, nonlinear problems are numerically solved and compared against those obtained by using shooting techniques.

Keywords Optimal control · Hamilton equations · First order PDEs · Boundary-value problems · Riccati equations

1 Introduction

Hamilton's canonical equations (HCEs) play a central role in Mechanics, since their equivalence with the principle of least action and the variational approach leading to the Euler-Lagrange equation was established [1]. After the foundational work of Pontryagin [2], HCEs have been also at the core of modern Optimal Control Theory. When the problem concerning an n-dimensional control system and an additive cost objective is regular [3], i.e. when the Hamiltonian function can be uniquely optimized by a control value u^0 depending on the remaining variables (t, x, λ) (where x denotes the state vector and λ the Hamiltonian costate vector), then HCEs appear as a set of $2n$ ordinary differential equations, whose solutions are the optimal state-costate time trajectories.

For the finite-horizon optimization set-up, with a free final state, the cost penalty $K(x)$ imposed on the final deviation from equilibrium generates a two-point boundary-value situation. This is often a rather difficult numerical problem to solve. In the linear-quadratic regulator (LQR) case with $K(x) = x'Sx$, there exist well-known methods (see for instance Bernhard [4] and Sontag [5]) to transform the boundary-value into a final-value problem, related to the differential Riccati equation (DRE).

For the infinite-horizon bilinear-quadratic regulator and the change of set-point servo problems, there exists a recent attempt to find the missing initial condition for the costate variable, based on a state-dependent (generalized) algebraic Riccati equation (GARE) with solution $P_\infty(x)$, which allows one to integrate the HCEs online with the underlying control process [6]. The same approach in finite time leads to a first-order PDE (GDRE, see [7, 8]) for a time-dependent generalized Riccati matrix $P_T(t, x)$, whose solution gives immediately the missing initial costate $\lambda(0) = 2P_T(0, x_0)x_0$ and exhibits, for $S = 0$, a limiting behavior [3] similar to that of linear systems with the same cost, i.e.

$$\lim_{T\to\infty} P_T(0, x) = P_\infty(x), \tag{1}$$

where T is the duration of each optimization process (or the time-horizon of the problem).

To our knowledge, the general nonlinear case has only been tackled in the one-dimensional case [9, 10], where the problem reduces to solving two first-order uncoupled PDEs for the missing boundary values $x(T)$ and $\lambda(0)$. In this article, we derive PDEs for the multi-dimensional case and their solutions are illustrated using standard software. This approach provides a promising alternative to multiple shooting methods and nonlinear–programming-type ones [11, 12], which are classical for boundary value problems.

Shooting methods are not guaranteed to converge in sensitive (mildly chaotic) problems. Even when they do converge, the computational effort involved in each individual problem is similar to that of integrating the PDEs used here for a two-parameter family of such problems. In the present approach, the main numerical work (solving the PDEs) is intended to be processed off-line in order to alleviate the on-line construction of the optimal control. This is a clear advantage compared to model-predictive (or nonlinear-programming) control, which needs an indeterminate amount of on-line computation to find the (suboptimal) control at each sample time, mainly due to the discretization of the control-values space. On the other hand, the method presented here is not, at present, adapted to manage arbitrary constraints on control and state values, which are handled effectively by control-parametrization schemes.

Another convenient feature of the PDEs method is that their solutions provide the initial value $P(0)$ of the Riccati matrix associated to the linearization of the original Hamiltonian system, which is used for constructing the compensator and the Kalman filter gains in two-degrees-of-freedom control strategies [13].

The elaboration process has been mostly in the line of the early invariant-embedding ideas introduced by Bellman [14] and completely disjoint from Riccati equations. Here these ideas are retaken and reformulated in the light of the symplectic structure inherent to the Hamiltonian formalism.

Besides the possibility of integrating the HCEs on-line with the real plant (and constructing the optimal control for the model in real time), the approach presented in this paper may also be useful when studying input-output $\mathcal{L}$-stability of control systems. In fact, the trajectory cost

$$\int_0^T \left[\|y(t)\|^2 - \gamma^2 \|u(t)\|^2 \right] dt \tag{2}$$

may be analyzed, as shall be shown, for variable gain γ, even for nonlinear observation functions

$$y(t) = h(x(t), u(t)) \tag{3}$$

and nonlinear dynamics [15, 16], provided the resulting Hamiltonian is regular. We shall say that the Hamiltonian is "regular" iff the H-minimal control function $u^0(x, \lambda)$ is explicitly known (not only its existence but also its closed form), and it is continuous in all of its variables. Therefore, this approach suggests a novel framework, where the "performance *versus* stability" balance can be adequately explored.

After the relevant mathematical objects associated with the finite-horizon control problem are presented in Sect. 2, the problem is embedded into a (T, S) family in Sect. 3. Afterwards, the main PDEs for the missing boundary values are substantiated in Sect. 4. Applications to linear and bilinear systems are then developed and illustrated in Sect. 5. A few numerical examples are devised in Sect. 6, and finally in Sect. 7 the step by step implementation of the method is explained, its advantages and difficulties are discussed, the conclusions are summarized and the perspectives are advanced. An Appendix is added to outline the extension from $S = sI$ to S non-negative definite and symmetric.

2 The Hamiltonian Formalism

Initialized control systems of the form

$$\dot{x} = f(x, u), \qquad x(0) = x_0 \tag{4}$$

will be considered, where the state x moves in some region $\mathcal{O}_{x_0}$ of $\mathbb{R}^n$, containing x_0 and, without any loss of generality, the admissible control strategies are the real, piecewise continuous functions of time. The function $f : \mathcal{O}_{x_0} \times \mathbb{R} \to \mathbb{R}^n$ is assumed to be smooth enough as to guarantee existence and uniqueness of solutions to the dynamics (4) in the range of interest. The finite-horizon optimization context will imply here that a cost functional of the form

$$\mathcal{J}(T, 0, x_0, u(\cdot)) = \int_0^T L(x(\tau), u(\tau))d\tau + x'(T)Sx(T) \tag{5}$$

has to be minimized on the set of admissible control trajectories, where $T < \infty$, L is a non-negative smooth function called "the *Lagrangian*" of the problem, and S is a positive semi-definite symmetric matrix called "the *final penalty* coefficient". The "*value function*" $\mathcal{V}$ can always be defined for such a problem, namely

$$\mathcal{V}(t, x) \triangleq \inf_{u(\cdot)} \mathcal{J}(T, t, x, u(\cdot)), \quad t \in [0, T], \; x \in \mathcal{O}_x, \tag{6}$$

and, if the problem has a unique solution, then this solution is called "the *optimal control* strategy" u^*,

$$u^*(\cdot) \triangleq \arg \inf_{u(\cdot)} \mathcal{J}(T, 0, x_0, u(\cdot)), \tag{7}$$

which in turn will generate "the *optimal state* trajectory"

$$x^*(\cdot) \triangleq \text{solution to (4)} \quad \text{with } u(\cdot) = u^*(\cdot). \tag{8}$$

The *Hamiltonian* of such a problem is defined as usual by

$$H(x, \lambda, u) \triangleq L(x, u) + \lambda' f(x, u), \tag{9}$$

where λ is called the "costate", $\lambda \in \mathbb{R}^n$, (x, λ) ranging in $2n$-dimensional "phase-space". Since H is assumed to be regular, there exists a unique *H-optimal control*

$$u^0(x, \lambda) \triangleq \arg \min_u H(x, \lambda, u). \tag{10}$$

The *control Hamiltonian*,

$$\mathcal{H}^0(x, \lambda) \triangleq H(x, \lambda, u^0(x, \lambda)), \tag{11}$$

gives rise to the HCEs (see Pontryagin et al. [2] for general problems and Sontag [5], p. 406 for the free final state case)

$$\dot{x} = \left(\frac{\partial \mathcal{H}^0}{\partial \lambda}\right)' \triangleq \mathcal{F}(x, \lambda); \qquad x(0) = x_0, \tag{12}$$

$$\dot{\lambda} = -\left(\frac{\partial \mathcal{H}^0}{\partial x}\right)' \triangleq -\mathcal{G}(x,\lambda); \qquad \lambda(T) = 2Sx(T), \tag{13}$$

that is a $2n$-dimensional ODE for a (Hamiltonian) vector field $\mathcal{X}$,

$$\begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} \mathcal{F}(x,\lambda) \\ -\mathcal{G}(x,\lambda) \end{pmatrix} \triangleq \mathcal{X}(x,\lambda). \tag{14}$$

Under the hypotheses being considered, solutions to (12), (13) result in the optimal state and costate trajectories (denoted $x^*(t)$ and $\lambda^*(t)$ respectively), which are also related to the value-function by

$$\lambda^*(t) = \left(\frac{\partial \mathcal{V}}{\partial x}(t, x^*(t))\right)', \quad t \in [0, T], \tag{15}$$

and the optimal control can be expressed as (see Sontag [5], p. 361)

$$u^*(t) = u^0(x^*(t), \lambda^*(t)). \tag{16}$$

It is useful to recall that the control Hamiltonian is constant along the optimal trajectories since

$$\frac{d}{dt}\mathcal{H}^0(x^*(t), \lambda^*(t)) = \left(\frac{\partial \mathcal{H}^0}{\partial x}\right) \cdot \mathcal{F} + \left(\frac{\partial \mathcal{H}^0}{\partial \lambda}\right) \cdot [-\mathcal{G}] = 0. \tag{17}$$

3 Embedding the Problem into a (T, S)-Family

The following notation for the missing boundary values will be used in this section

$$\rho(T, S) \triangleq x^*(T), \tag{18}$$

$$\sigma(T, S) \triangleq \lambda^*(0). \tag{19}$$

By assuming that the Hamiltonian vector field be at least of class C^1, the existence and uniqueness of a flow

$$\phi : \mathbb{R} \times \mathcal{O}_x \times \mathcal{O}_\lambda \to \mathbb{R}^{2n} \tag{20}$$

(vectors in $\mathbb{R}^{2n}$ will be written in column form), satisfying

$$D_1\phi(t, x, \lambda) = \begin{pmatrix} \mathcal{F}(\phi(t,x,\lambda)) \\ -\mathcal{G}(\phi(t,x,\lambda)) \end{pmatrix} = \mathcal{X}(\phi(t,x,\lambda)), \tag{21}$$

$$\phi(0, x, \lambda) = \begin{pmatrix} x \\ \lambda \end{pmatrix}, \quad \forall \begin{pmatrix} x \\ \lambda \end{pmatrix} \in \mathcal{O}_x \times \mathcal{O}_\lambda, \tag{22}$$

is guaranteed. Here $\mathcal{O}_x, \mathcal{O}_\lambda$ are appropriate regions of $\mathbb{R}^n$ and, in order to avoid confusion, "D_1" is used instead of "$\frac{\partial}{\partial t}$", and, in general, "$D_k$" stands for partial derivative with respect to the kth variable.

By denoting with ϕ^t the t-advance transformations ($\phi^t(x,\lambda) \triangleq \phi(t,x,\lambda)$) associated with this flow, the following identities become clear

$$\begin{pmatrix} \rho \\ 2S\rho \end{pmatrix} = \phi^T(x_0,\sigma) \triangleq \phi(T,x_0,\sigma) = \begin{pmatrix} \phi_1(T,x_0,\sigma) \\ \phi_2(T,x_0,\sigma) \end{pmatrix}, \tag{23}$$

where ϕ_1, ϕ_2 denote the "components" of the flow over the state and costate subspaces, respectively. Thus, the first component of (21) reads

$$\begin{aligned} D_1\phi_1(T,x_0,\sigma(T,S)) &= \mathcal{F}(\phi(T,x_0,\sigma(T,S))) = \mathcal{F}\left((x^*(T),\lambda^*(T))'\right) \\ &= \mathcal{F}\left((\rho(T,S), 2S\rho(T,S))'\right) \triangleq F(\rho,S). \end{aligned} \tag{24}$$

Similarly, from the second component of (21) we obtain

$$-D_1\phi_2(T,x_0,\sigma(T,S)) = \mathcal{G}(\rho(T,S), 2S\rho(T,S)) \triangleq G(\rho,S). \tag{25}$$

The phase-space derivative of the T-advance function associated with the Hamiltonian flow will be needed later, so a special notation is introduced for it and its natural partition, namely:

$$\begin{aligned} V \triangleq D\phi^T(x_0,\sigma) &= \begin{pmatrix} V_1 & V_2 \\ V_3 & V_4 \end{pmatrix} \\ &= \begin{pmatrix} D_1\phi_1^T & D_2\phi_1^T \\ D_1\phi_2^T & D_2\phi_2^T \end{pmatrix}(x_0,\sigma) = \begin{pmatrix} \phi_{1_x}^T & \phi_{1_\lambda}^T \\ \phi_{2_x}^T & \phi_{2_\lambda}^T \end{pmatrix}, \end{aligned} \tag{26}$$

where $\phi_{1_x}^T \triangleq \frac{\partial \phi_1^T}{\partial x}$ and similarly for $\phi_{2_x}^T, \phi_{1_\lambda}^T, \phi_{2_\lambda}^T$. Existence and uniqueness of solutions imply that the inverse of V exists and verifies

$$U \triangleq V^{-1} = D\phi^{-T}(\rho, 2S\rho). \tag{27}$$

It is well known that Hamiltonian vector fields have flows with the following important symplectic properties (see Jacobson [17], p. 378 and also Katok and Hasselblatt [18], p. 220)

$$V_1'V_4 - V_3'V_2 = I = V_4'V_1 - V_2'V_3, \tag{28}$$

$$V_1'V_3 - V_3'V_1 = 0 = V_2'V_4 - V_4'V_2, \tag{29}$$

where the following notation was adopted for $n \times n$ submatrices: $V_i' \triangleq (V_i)'$, resulting from partitioning the $2n \times 2n$ matrix V.

Taking into account (28), (29), the inverse of V can be calculated in terms of the submatrices V_i, namely

$$V^{-1} = U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix} = \begin{pmatrix} V_4' & -V_2' \\ -V_3' & V_1' \end{pmatrix}. \tag{30}$$

Now, by differentiating with respect to T the first component of (23), we obtain

$$D_1\rho(T,S) = D_1\phi_1(T,x_0,\sigma(T,S)) + D_3\phi_1(T,x_0,\sigma(T,S))D_1\sigma(T,S), \tag{31}$$

which we write (with $\phi^T_{1_\lambda} = \frac{\partial\phi^T_1}{\partial\lambda} = D_3\phi^T_1 = D_2\phi^T_1$) simply as

$$\rho_T(T,S) = F(\rho,S) + V_2\sigma_T(T,S). \tag{32}$$

Similarly, by differentiating (23) with respect to T, we obtain

$$2S\rho_T(T,S) = -G(\rho,S) + V_4\sigma_T(T,S). \tag{33}$$

Using the symplectic properties of the Hamiltonian vector field, (32), (33) yield

$$V'_4(\rho_T - F) = V'_4V_2\sigma_T = V'_2V_4\sigma_T = V'_2(2S\rho_T + G). \tag{34}$$

Remark From now on the matrix $S \in \mathbb{R}^{n\times n}$ will be of the form $\mathrm{diag}(s,\dots,s) = sI$, where s will be a real non-negative number. The extension to a general symmetric non-negative S will be discussed in the Appendix.

By repeating the procedure for the s-derivatives of the first and second components of (23), analogous equations are obtained, namely

$$\rho_s(T,s) = \phi^T_{1_\lambda}\sigma_s(T,s) = V_2\sigma_s(T,s), \tag{35}$$

$$2(\rho(T,s) + s\rho_s(T,s)) = \phi^T_{2_\lambda}\sigma_s(T,s) = V_4\sigma_s(T,s), \tag{36}$$

from which, applying once again the symplectic properties of the flow, there follows that

$$V'_4\rho_s = V'_4V_2\sigma_s = V'_2V_4\sigma_s = V'_2(2\rho + 2s\rho_s). \tag{37}$$

Now, using $V' = \left(\begin{smallmatrix} V'_1 & V'_3 \\ V'_2 & V'_4 \end{smallmatrix}\right)$ and (34), (37), the following expression results

$$\begin{aligned} &V'\begin{pmatrix} 2s\rho_T + G & 2\rho + 2s\rho_s \\ F - \rho_T & -\rho_s \end{pmatrix} \\ &\quad = \begin{pmatrix} V'_1(2s\rho_T + G) + V'_3(F - \rho_T) & V'_1(2\rho + 2s\rho_s) - V'_3\rho_s \\ 0 & 0 \end{pmatrix}, \end{aligned} \tag{38}$$

and combining with (32), (33) and (35), (36), it follows that the left-hand side is also equal to

$$\cdots = \begin{pmatrix} V'_1V_4\sigma_T - V'_3V_2\sigma_T & V'_1V_4\sigma_s - V'_3V_2\sigma_s \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sigma_T & \sigma_s \\ 0 & 0 \end{pmatrix}, \tag{39}$$

which means that there are really only two independent first-order vector PDEs for ρ and σ, namely

$$V'_1(2s\rho_T + G) + V'_3(F - \rho_T) = \sigma_T, \tag{40}$$

$$V_1'(2\rho + 2s\rho_s) - V_3'\rho_s = \sigma_s. \tag{41}$$

Notice that only V_1 and V_3 are involved, although a pair of equivalent PDEs can be obtained involving V_2 and V_4. In the one-dimensional case these equations can be decoupled, obtaining [9]

$$\rho\rho_T - \left(sF + \frac{G}{2}\right)\rho_s = \rho F, \tag{42}$$

$$\rho\sigma_T - \left(sF + \frac{G}{2}\right)\sigma_s = 0, \tag{43}$$

but for $n > 1$ a more careful and involved treatment is needed, as shown in the next section.

4 The Variational Equation

It will be convenient to assign a name to the combined variable $v \triangleq \binom{x}{\lambda}$. Equation (21) can then be written as

$$D_1\phi(t, v) = \mathcal{X}(\phi(t, v)), \tag{44}$$

and differentiating both sides with respect to v, (i.e. $D_2 = \frac{\partial}{\partial v}$,) and then interchanging the order of differentiation, the "*variational equation*" (see for instance Hirsh and Smale [19], p. 299) is obtained, namely

$$D_1[D_2\phi(t, v)] = D\mathcal{X}(\phi(t, v)) \cdot D_2\phi(t, v), \tag{45}$$

or, by a slight abuse of notation, for a fixed $v \in \mathcal{O}_x \times \mathcal{O}_\lambda$,

$$\frac{d}{dt}\widetilde{V}(t) = \mathcal{A}(t)\widetilde{V}(t), \tag{46}$$

with the initial condition

$$\widetilde{V}(0) = I, \tag{47}$$

where $\widetilde{V}(t) \triangleq D\phi^t$ (the symbol V is reserved for $\widetilde{V}(T)$), $\mathcal{A}(t) \triangleq D\mathcal{X} \circ \phi^t$. Actually, this means that $\widetilde{V}(T) = V = \Phi(T, 0)$ (the fundamental solution of (46)), which verifies (see Sontag [5], p. 488)

$$\frac{\partial\Phi(T, 0)}{\partial T} = \mathcal{A}(T)\Phi(T, 0), \tag{48}$$

i.e. the following identity is established for $v = (x_0, \sigma)'$ (see (26))

$$V_T \triangleq \frac{\partial V}{\partial T}(T, x_0, \sigma) = D\mathcal{X}(\rho, 2s\rho) \cdot V(T, x_0, \sigma), \tag{49}$$

or, in short, reserving the symbol $\mathcal{A}$ for $\mathcal{A}(T)$,

$$V_T = \mathcal{A}V. \tag{50}$$

Now, previous treatments of the LQR problem in Hamiltonian form [4, 5], employ two auxiliary matrices, which in the present context can be defined as:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \triangleq U \begin{pmatrix} I \\ 2sI \end{pmatrix} = \begin{pmatrix} V_4' - 2V_2's \\ -V_3' + 2V_1's \end{pmatrix}. \tag{51}$$

Differentiating these matrices with respect to T and s, using (30), (50), (51), and partitioning as before $\mathcal{A} = \left(\begin{smallmatrix} \mathcal{A}_1 & \mathcal{A}_2 \\ \mathcal{A}_3 & \mathcal{A}_4 \end{smallmatrix}\right)$, it follows, up to first order in s, that

$$\begin{aligned} \begin{pmatrix} \alpha_T \\ \beta_T \end{pmatrix} &= \begin{pmatrix} (V_4')_T - 2(V_2')_T s \\ -(V_3')_T + 2(V_1')_T s \end{pmatrix} \\ &= \begin{pmatrix} V_2'\mathcal{A}_3' + V_4'\mathcal{A}_4' - 2(V_2'\mathcal{A}_1' + V_4'\mathcal{A}_2')s \\ -(V_1'\mathcal{A}_3' + V_3'\mathcal{A}_4') + 2(V_1'\mathcal{A}_1' + V_3'\mathcal{A}_2')s \end{pmatrix}, \end{aligned} \tag{52}$$

$$\begin{pmatrix} \alpha_s \\ \beta_s \end{pmatrix} = \begin{pmatrix} -2V_2' \\ 2V_1' \end{pmatrix}. \tag{53}$$

Then the following approximate solutions of (50) are obtained:

$$\begin{aligned} &V_1' = \frac{1}{2}\beta_s, \qquad V_2' = -\frac{1}{2}\alpha_s, \\ &V_3' = 2V_1's - \beta = \beta_s s - \beta, \qquad V_4' = \alpha + 2V_2's = \alpha - \alpha_s s. \end{aligned} \tag{54}$$

Also,

$$\begin{pmatrix} \alpha_T \\ \beta_T \end{pmatrix} = \begin{pmatrix} \alpha_s\mathcal{M} - \alpha\mathcal{N} \\ \beta_s\mathcal{M} - \beta\mathcal{N} \end{pmatrix}, \tag{55}$$

where the new matrices $\mathcal{M}, \mathcal{N}$ take the form

$$\mathcal{M} \triangleq \frac{1}{2}(2\mathcal{A}_1's - \mathcal{A}_3') + s(2\mathcal{A}_2's - \mathcal{A}_4'), \tag{56}$$

$$\mathcal{N} \triangleq 2\mathcal{A}_2's - \mathcal{A}_4' \quad \Longrightarrow \quad \mathcal{M} = \frac{1}{2}(2\mathcal{A}_1's - \mathcal{A}_3') + s\mathcal{N}. \tag{57}$$

Since for a process of zero duration, $\phi^T = \phi^0 = \text{id}$ (the identity function), then in such case $V = D\phi^T = I = U$, and therefore, the main matrix PDEs in (55) are subject to the initial conditions

$$\alpha(0, s) = I, \tag{58}$$

$$\beta(0, s) = 2sI. \tag{59}$$

Now, (56), (57) still include the unknown final state ρ inside the $\mathcal{A}_i's$ so the (matrix) PDEs in (55) cannot be solved alone. But, having found expressions for the partitions of V in terms of the auxiliary matrices α, β and their derivatives, (vector) (40), (41) turn into

$$\begin{pmatrix} \sigma_T \\ \sigma_s \end{pmatrix} = \begin{pmatrix} \beta_s(sF + \frac{G}{2}) - \beta(F - \rho_T) \\ \beta_s \rho + \beta \rho_s \end{pmatrix} \tag{60}$$

which become solvable, at least in principle, when coupled to the matrix PDEs for α, β, and subject to initial conditions

$$\rho(0, s) = x_0, \tag{61}$$

$$\sigma(0, s) = 2sx_0. \tag{62}$$

In short, the problem requires us to solve, in parallel, two matrix first-order PDEs for (α, β), and another two vector first-order PDEs for (ρ, σ), all meeting appropriate initial conditions. If instead of V_1 and V_3 the remaining submatrices V_2 and V_4 were chosen, then (34), (37) also take a condensed form, namely

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_s(sF + \frac{G}{2}) - \alpha(F - \rho_T) \\ \alpha_s \rho + \alpha \rho_s \end{pmatrix}, \tag{63}$$

but σ can not be recovered directly from them.

Concerning the existence and uniqueness of solutions to the coupled system of (55), (60), there exist only local results (see Folland [20], p. 51), although the field of vector and matrix PDE integration is in active development (see for instance Zenchuk and Santini [21]).

5 Applications to Linear and Nonlinear Systems

5.1 The LQR Problem Revisited

The Hamiltonian form of the LQR problem with dynamics $f = Ax + Bu$ and Lagrangian $L = x'Qx + u'Ru$ reads

$$\dot{v} = \begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} A & -\frac{1}{2}W \\ -2Q & -A' \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \mathbf{H}v, \tag{64}$$

where $W \triangleq BR^{-1}B'$. Therefore in this case the HCEs become a linear, time-invariant dynamical system or vector field $\mathcal{X}(v) = \mathbf{H}v$, whose flow verifies

$$\phi^T(v) = e^{\mathbf{H}T}v, \tag{65}$$

and consequently

$$V = D\phi^T = \phi^T = e^{\mathbf{H}T}, \tag{66}$$

$$U = V^{-1} = e^{-\mathbf{H}T}. \tag{67}$$

Then, the following identities are easily obtained

$$V_T = \mathbf{H}e^{\mathbf{H}T}, \tag{68}$$

$$D\mathcal{X} = \mathbf{H} = \mathcal{A} \quad \text{(time-invariant)}, \tag{69}$$

$$\mathcal{A}' = \begin{pmatrix} A' & -2Q \\ -\frac{1}{2}W & -A \end{pmatrix} = \begin{pmatrix} \mathcal{A}'_1 & \mathcal{A}'_3 \\ \mathcal{A}'_2 & \mathcal{A}'_4 \end{pmatrix}, \tag{70}$$

$$\begin{aligned} \mathcal{N} &= A - Ws, \\ \mathcal{M} &= \frac{1}{2}(2A's + 2Q) + s\mathcal{N} = A's + sA + Q - sWs. \end{aligned} \tag{71}$$

Therefore, (55) for α, β can be integrated alone, since they do not depend on ρ, σ. Actually, from

$$\begin{pmatrix} x_0 \\ \sigma \end{pmatrix} = e^{-\mathbf{H}T} \begin{pmatrix} \rho \\ 2s\rho \end{pmatrix} = U \begin{pmatrix} I \\ 2sI \end{pmatrix} \rho = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \rho, \tag{72}$$

it follows that no further equations are needed for ρ, σ. Since α is always invertible (see Sontag [5], p. 371), then the missing boundary values result from

$$\rho = \alpha^{-1} x_0, \tag{73}$$

$$\sigma = \beta\rho. \tag{74}$$

A natural consistency check would consist in verifying that the expressions in (73), (74) are the solutions to the vector PDEs in (60). From (64)

$$F = \mathcal{F}(\rho, 2s\rho) = A\rho - Ws\rho = \mathcal{N}\rho, \tag{75}$$

$$G = \mathcal{G}(\rho, 2s\rho) = 2Q\rho + 2A's\rho, \tag{76}$$

$$sF + \frac{G}{2} = \left[s(A - Ws) + Q + A's\right]\rho = \mathcal{M}\rho, \tag{77}$$

and from (74)

$$\sigma_T = \beta_T\rho + \beta\rho_T, \tag{78}$$

$$\sigma_s = \beta_s\rho + \beta\rho_s, \tag{79}$$

so the second row of (60) is trivially verified, and the validity of its first row follows from

$$\sigma_T = \beta_T\rho + \beta\rho_T = (\beta_s\mathcal{M} - \beta\mathcal{N})\rho + \beta\rho_T = \beta\rho_T - \beta F - \beta_s\left(sF + \frac{G}{2}\right) = \sigma_T. \tag{80}$$

From (15) for the LQR case, the initial costate has here the form

$$\sigma(T, s) = \lambda^{T,s}(0) = \left(\frac{\partial \mathcal{V}}{\partial x}(0, x^*(0))\right)' = 2P(0)x_0; \tag{81}$$

and therefore, from (74), for each (T,s)-problem the Riccati matrix $P(t)$ should also verify

$$P(0)=\frac{1}{2}\beta(T,s)\,[\alpha(T,s)]^{-1}\,. \tag{82}$$

The method based on PDEs for missing boundary values avoids solving the DRE for each particular (T,s)-problem. It also avoids storing, necessarily as an approximation, the Riccati matrix $P(t)$ for the values of $t\in[0,T]$ chosen by the numerical integrator, possibly different from the time instants for which the control $u(t)$ is constructed.

Actually, it follows immediately from (73), (74), (81), (82), and from the fact that for this case $u^0(x,\lambda)=-\frac{1}{2}R^{-1}B'\lambda$, that a feedback form can be obtained for the optimal control at intermediate times $t\in[0,T]$, namely

$$u^*(t)=-\frac{1}{2}R^{-1}B'\beta(T-t,s)\,[\alpha(T-t,s)]^{-1}\,x(t), \tag{83}$$

whenever $x(t)\in\mathcal{O}_x$. As a side-result, an alternative formula for the Riccati matrix is obtained:

$$P(t)=\frac{1}{2}\beta(T-t,s)\,[\alpha(T-t,s)]^{-1}\,,\quad \forall t\in[0,T]. \tag{84}$$

5.2 Bilinear Systems

The bilinear-quadratic case (with $x\in\mathbb{R}^n$, $u\in\mathbb{R}$) will be used to illustrate the application of the previous results to nonlinear systems. The dynamics and trajectory cost will be, respectively,

$$f(x,u)=Ax+(b+Nx)u, \tag{85}$$

$$L(x,u)=x'Qx+ru^2. \tag{86}$$

The H-optimal control is readily obtained [5, 8]

$$u^0(x,\lambda)=-\frac{1}{2r}\lambda'(b+Nx), \tag{87}$$

and then the PDE (55), (60) result similar to the LQR example, but with the following matrices

$$\bar{W}=\frac{1}{r}(b+N\rho)(b+N\rho)', \tag{88}$$

$$\hat{A}=A-\frac{1}{r}[\rho's(b+N\rho)N+(b+N\rho)\rho'sN], \tag{89}$$

$$\tilde{Q}=Q+\frac{1}{r}N's\rho\rho'sN. \tag{90}$$

In short, the relevant objects read in this case

$$\mathcal{N}=2\mathcal{A}_2's-\mathcal{A}_4'=\hat{A}-\bar{W}s, \tag{91}$$

$$\mathcal{M} = \mathcal{A}_1' s - \frac{1}{2}\mathcal{A}_3' + s\mathcal{N} = \hat{A}' s + s\hat{A} + \tilde{Q} - s\bar{W}s, \tag{92}$$

$$F = \left[A - \frac{1}{r}(b + N\rho)(b + N\rho)' s \right] \rho, \tag{93}$$

$$G = 2\left[Q + A' s - \frac{\rho' s(b + N\rho)}{r} N' s \right] \rho, \tag{94}$$

$$sF + \frac{G}{2} = \left\{ A' s + sA + Q - \frac{1}{r}[s(b + N\rho)(b + N\rho)' + \rho' s(b + N\rho)N'] \right\} s\rho. \tag{95}$$

To illustrate these abstract outcomes, calculations were made with a bilinear system having states in $\mathbb{R}^2$ and subject to quadratic cost; in particular a problem with the matrices (A, B, N, Q, R)

$$A = \begin{pmatrix} -2 & 0 \\ 3 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad N = \begin{pmatrix} 2.5 & 0.5 \\ 0 & 4 \end{pmatrix},$$

$$Q = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 2 \end{pmatrix}, \qquad R = 1, \quad \text{and} \quad (T, s) \in [0, 1] \times [0, 1].$$

The "*NDSolve*" command from *Mathematica* was used to obtain the numerical results (as in all PDE integrations), which for the bilinear example are shown in Figs. 1 and 2, where the influence of the nonlinearity translates into a stronger stabilization effect than for LQR (control affects the dynamics with an extra negative term Nxu^0).

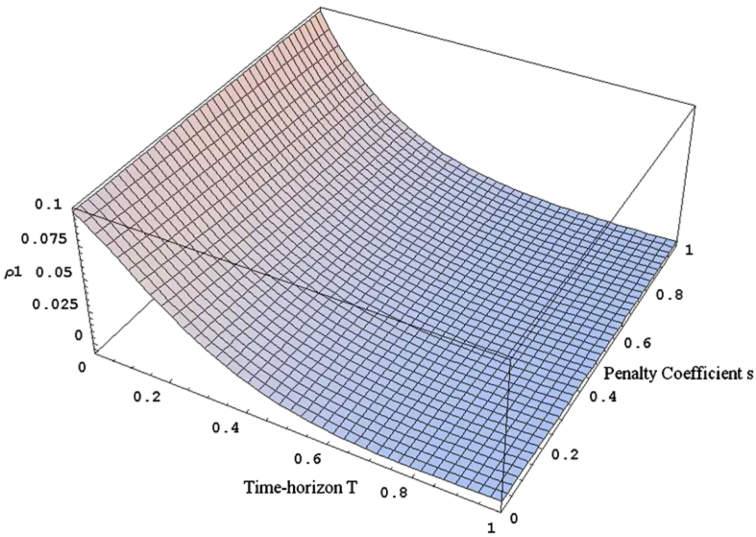

Fig. 1 First component $\rho_1(T, s)$ of final state for the bilinear-quadratic problem

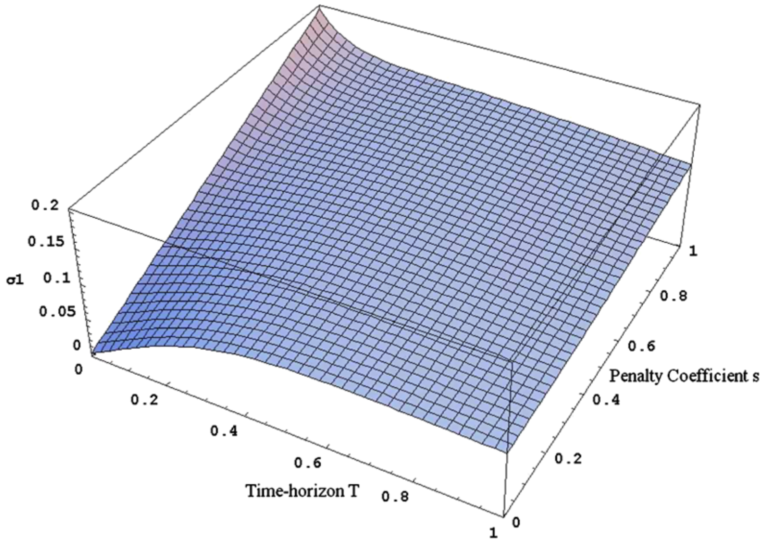

Fig. 2 First component $\sigma_1(T,s)$ of initial costate for the bilinear-quadratic problem

It was also numerically verified that

$$\mathcal{H}^0_{T,s}(x_0,\sigma) = \mathcal{H}^0_{T,s}(\rho, 2s\rho) \tag{96}$$

for all chosen (T,s) points (see (17)) up to minor tolerance.

6 Comparison of the New PDE Formulation and Shooting Methods

6.1 A "Chaotic" Nonlinear Example

Consider the following nonlinear system (see [10, 22], and the references therein)

$$\dot{x} = -x^3 + u; \qquad x(0) = x_0 = 1, \tag{97}$$

coupled to the quadratic cost

$$\mathcal{J}(u) = \int_0^T \Big[Q(x(t) - \bar{x}_c)^2 + Ru^2(t) \Big] dt + s\,(x(T) - \bar{x}_\ell)^2, \tag{98}$$

where $\bar{x} = \bar{x}_c = \bar{x}_\ell = 1.5$, and $Q = R = 1$. The Hamilton Canonical Equations (HCEs) result

$$\begin{cases} \dot{x} = -x^3 - \frac{\lambda}{2}; & x(0) = 1, \\ \dot{\lambda} = -2(x - \bar{x}) + 3\lambda x^2; & \lambda(T) = 2s(x(T) - \bar{x}). \end{cases} \tag{99}$$

By solving the new PDEs [10, 13] for the missing boundary values ρ, σ given by (42), (43), where

$$F = -\rho^3 - s(\rho - \bar{x}), \qquad G = 2(\rho - \bar{x}) - 6s(\rho - \bar{x})\rho^2, \tag{100}$$

the following results (inside the two-parameter family of solutions) were obtained:

$$\rho(5, 60) = 1.4899, \qquad \sigma(5, 60) = -0.572050296. \tag{101}$$

As a first checking of these results, the HCEs can be numerically integrated with initial conditions [10]

$$x(0) = x_0, \qquad \lambda(0) = \sigma(5, 60). \tag{102}$$

The trajectory accurately reaches the expected value $x(T) \approx \rho(5, 60)$ [10].

A second test is the constancy of the Hamiltonian along the optimal trajectory. The Hamiltonian (103) along the optimal trajectory

$$\mathcal{H}^0(x^*(t), \lambda^*(t)) = \left[x^*(t) - 1.5\right]^2 - \lambda^*(t)\left[x^*(t)\right]^3 - \frac{[\lambda^*(t)]^2}{4} \tag{103}$$

should be constant, which was confirmed numerically with high accuracy. In particular it must be the same at the initial and final times, as checked below

$$\mathcal{H}^0(x^*(0), \lambda^*(0)) = \mathcal{H}^0(1.00000, -0.572050) \approx 0.74024, \tag{104}$$

$$\mathcal{H}^0(x^*(5), \lambda^*(5)) = \mathcal{H}^0(1.4899, -10.92079) \approx 0.74025. \tag{105}$$

Now, several attempts at applying straightforward shooting (the "*ode45*" routine from *Matlab* was used for integrating the HCEs) to this problem were unsuccessful, i.e.:

(i) Starting with $x(0) = x_0$, $\lambda(0) = -0.573 \approx \sigma(5, 60) = -0.572050296$, and running the HCEs in the direction of increasing time, the resulting trajectories strongly differ from the expected behavior. So, a better performance was sought before shooting in the reverse direction, but

(ii) the trajectories still diverge for very close approximations to the initial costate, for instance for $\lambda(0) = -0.57206$.

(iii) A backward integration was assayed from the final conditions

$$x(T) = \rho(5, 60) + \varepsilon, \qquad \lambda(T) = 2s\,(x(T) - \bar{x}_\ell)$$

with $\varepsilon = 0.000005$, which diverges too.

The conclusion is that the shooting method was plainly inapplicable to this "chaotic" example, whereas the PDEs gave consistent and accurate results.

Table 1 Boundary values

$T=0.1$, $S=3$	$t=0$	$t=T$
$x(t)$	1	0.219331
$\lambda(t)$	0.896592	1.31599

6.2 A Stable Problem

This is a standard problem with linear dynamics

$$\dot{x} = -5x + u; \qquad x(0) = x_0 = 1, \tag{106}$$

and quadratic cost

$$\mathcal{J}(u) = \int_0^T \left[Qx^2(t) + Ru^2(t) \right] dt + Sx^2(T), \tag{107}$$

with $Q = R = 1$, $T = 0.1$, $S = 3$. The corresponding HCEs are

$$\begin{cases} \dot{x} = -x - \frac{9}{2}\lambda; & x(0) = 1, \\ \dot{\lambda} = -2x + 5\lambda x; & \lambda(T) = 6x(T). \end{cases} \tag{108}$$

The missing boundary values recovered from the appropriate PDEs, namely

$$\rho_T - (2AS - S^2W + Q)\rho_S = (A - WS)\rho, \tag{109}$$

$$\sigma_T - (2AS - S^2W + Q)\sigma_S = 0, \tag{110}$$

$$A = -5, \qquad B = 1, \qquad W = B^2/R = 1 \tag{111}$$

are given in Table 1.

Table 2 shows the results of applying an "intelligent" shooting strategy (boundary values are slightly perturbed from those given by the PDEs' method, although they are directly guessed in usual applications). The strategy is easily followed from the numerical values given in the table. The method recovers all significant digits after 16 iterations, consuming 0.078 CPU time units. The PDEs took 0.063 CPU time units and solved the problem for the whole family of pairs $\{(T, S), 0 < T \leq 2, 0 < S \leq 5\}$.

Even for simple examples, and using good guesses to start shooting strategies, the results are here surpassed by the PDEs method. It should be remarked that for large values of T the shooting method becomes unstable due to the mixed-sign eigenvalues inherent to the dynamics of Hamiltonian systems. This inconvenience is not present in the PDEs method.

7 Conclusions

The solutions to the PDEs, established in the previous sections, transform the classical boundary-value problem, posed for Hamilton equations in $2n$ dimensions, into an

Table 2 Shooting iterations for a fixed optimization horizon of duration 0.1

Iteration #	$x(0)$	$\lambda(0)$	$x(T)$	$\lambda(T)$
1 →	1	0.9	0.217709	1.32179
2 ←	0.996302	0.893277	0.21852	1.31112
3 →	1	0.894934	0.22012	1.31316
4 ←	1.0018	0.898205	0.219726	1.31835
5 →	1	0.897398	0.218948	1.31736
6 ←	0.999125	0.895808	0.219139	1.31484
7 →	1	0.896201	0.219518	1.31532
8 ←	1.00043	0.896974	0.219425	1.31655
9 →	1	0.896783	0.219241	1.31631
10 ←	0.999793	0.896407	0.219286	1.31572
11 →	1	0.896530	0.219361	1.31588
12 ←	1.00007	0.896652	0.219346	1.31608
13 →	1	0.896607	0.219324	1.31601
14 ←	0.999984	0.896577	0.219328	1.31597
15 →	1	0.896588	0.219333	1.31598
16 ←	1	0.896596	0.219332	1.31599

initial-value situation when the Hamiltonian is regular. This, in turn, allows us to numerically integrate the original HCEs on-line with the control process, and to generate the manipulated variable $u^*(t) = u^0(x^*(t), \lambda^*(t))$, for any desired time instant t, from the state and costate values provided by this integration, since the H-optimal control function $u^0(\cdot, \cdot)$ is known. The on-line accessibility to an accurate value for the optimal state is valuable in practical situations, since physical states of nonlinear control systems are hardly available at all times. Mismatches and perturbations can be handled via compensation schemes, for which the PDEs' solutions give useful results, for instance the initial value of the Riccati matrix. It should be warned however that the assumption of regularity precludes the present approach from systematically handling restrictions on state and/or control values. For such non-regular problems, the Pontryagin principle [2] is commonly accepted to be the proper theoretical setup. Numerically, problems with restrictions are better treated by control-parametrization methods [11].

The procedure for treating a nonlinear regular optimal control problem with a finite horizon and a quadratic final penalty, developed and tested in previous sections, can be summarized as follows:

(i) From the H-optimal control $u^0(x, \lambda)$, (10), and the control Hamiltonian $\mathcal{H}^0(x, \lambda)$, (11), obtain the $\mathcal{F}$, $\mathcal{G}$ expressions of the vector field $\mathcal{X}(x, \lambda) = \binom{\mathcal{F}(x,\lambda)}{-\mathcal{G}(x,\lambda)}$.

(ii) By replacing x by ρ, and λ by $2s\rho$, obtain the formal expressions of $F(\rho, s)$ and $G(\rho, s)$, (24) and (25).

(iii) Obtain the matrix of the derivative of vector field $\mathcal{X}$, i.e. $D\mathcal{X}(x, \lambda)$, and again, replacing (x, λ) by $(\rho, 2s\rho)$, obtain the expression of the matrix $\mathcal{A}(T) =$

$D\mathcal{X}(\rho, 2s\rho)$, which is partitioned into $n \times n$ submatrices as $\begin{pmatrix} \mathcal{A}_1 & \mathcal{A}_3 \\ \mathcal{A}_2 & \mathcal{A}_4 \end{pmatrix}$. Then, the matrices $\mathcal{M}$ and $\mathcal{N}$ can be easily constructed from (56) and (57).

(iv) The PDEs for α, β, ρ, σ, namely (55) and (60), together with their initial conditions (58), (59), (61), (62) are then solved. If used in Mathematica, this amounts to just a line with the command "*NDSolve*"; if Matlab is preferred, then the "*pdepe*" instruction can be used instead. More sophisticated software is under construction, for instance in [21], and [23].

(v) The desired $\lambda(0) = \sigma(T, s)$ can then be extracted from the numerical solution. The value of $x(T) = \rho(T, s)$ also becomes available, and gives a measure of how close the final state would come to the target for the chosen parameter pair (T, s).

(vi) Finally, the HCEs (12), (13) can be implemented online, starting with initial conditions $x(0) = x_0$, $\lambda(0) = \sigma(T, s)$.

Controlling through HCEs has obvious advantages over optimal control methods based on iterative approximations of the Hamilton-Jacobi-Bellman PDE, as in the case of power series expansion approaches previously devised for the bilinear-quadratic problem. Also, the numerical effort involved in solving the new quasi-linear PDEs is, in principle, smaller than that required by classical multiple shooting methods. Hence, the integration of the system of quasilinear PDEs proposed in this article is worth the effort, specially acknowledging that standard mathematically oriented software can render accurate solutions, as it has been illustrated by the examples.

The PDEs' method solves a whole family of (T, s) problems, avoiding additional off-line calculations and burdensome storing of information for each particular situation, as in methods of the DRE or GDRE type. The examples worked out in previous sections have been solved using a few instructions from standard mathematically oriented software. No special programming was needed. However, further improvements in their numerical solution can in principle be implemented by treating the quasilinear first-order PDEs as a set of ODEs with initial conditions (see for instance Folland [20], Chap. 1). In any case, the equations proposed in the paper provide an "analytic" intent of solution to the optimal control problem. This is a theoretical improvement over "approximation" approaches, like the usual shooting methods. The main disadvantage of running the HCEs online to generate the state and costate variables needed for control, is their intrinsic instability, due to the pairing in positive and negative eigenvalues of their linear part. As a result, their integration is usually sensitive to variations in initial conditions. Some methods have been devised to partially cope with this inconvenience (see for instance [10, 22]).

Solutions providing the missing boundary values can be used to construct the optimal control in feedback form for the LQR problem as illustrated. The extension to general nonlinear feedback is currently under investigation.

Another useful result coming from the PDEs that we derived is the initial value $P(0)$ (82) of the generalized Riccati matrix $P(t)$ associated with the time-invariant "*equations on variations*" of the HCEs, which govern the compensation (feedback) stage in "*two-degree-of-freedom*" (2DOF) strategies [13]. This initial value allows to integrate on-line the Differential Riccati Equation (DRE) of this stage, avoiding the classical inconveniences arising from solving the DRE off-line, storing the solution in memory, and interpolating it at the required sampling times.

Using the equations on variations for constructing the compensation control allows designing an optimal filter for I/O perturbations. The matrix $P(t)$ is part of the gain of the filter, which can then be run on-line, together with all the remaining parts of the control scheme [13, 24]. The filter actually works as an observer. This is of particular importance in chemical operations since part or all the physical states of the plant can not often be measured with the frequency required by continuous-time-model control, or when the states are not measurable at all.

Having the values of ρ, σ for a wide range of T, s parameter values may be helpful at the design stage. The values of T, s can be reconsidered by the designer when acknowledging the final values of the state $\rho(T, s)$ that will be obtained under present conditions (see [25] and the references therein); and if a change in the parameter values is decided, then it will not be necessary to perform additional calculations to manage the new situation. Besides, the value of $\sigma(T, s)$ is an accurate measure of the "marginal cost" of the process, i.e. it measures how much the optimal cost would change under perturbations, which can also influence the decision on adopting the final T, s values.

Regarding potential applications, the approach presented in this paper may also be useful in studying the input-output $\mathcal{L}$-stability of nonlinear control systems, as announced in the introduction. Therefore, these PDEs seem to provide a novel environment where to explore the balance "performance versus stability".

Other aspects of this approach deserve research. For instance, the curves $\sigma(., s)$ are potentially a safeguard against Hamiltonian systems' instabilities (their linearizations have eigenvalues with positive real parts, because those associated with λ are opposite to those corresponding to x). Therefore, it will probably add to robustness to construct the control by imposing a bound on costates $\lambda(.)$, for instance impeding the costates to trespass the reverse $\sigma(., s)$ curve starting from $\sigma(\bar{T}, s)$ when a finite horizon of duration $\bar{T}$ is being optimized.

Appendix: Extension to General Symmetric Positive Semi-Definite Final Penalty Matrices S

For the sake of completeness the general case of a positive semi-definite symmetric final penalty matrix S is discussed below. First, for the diagonal case, i.e. for $S = \text{diag}(s_1, s_2, \ldots, s_n)$, where $s_i \geq 0$, $i = 1, \ldots, n$, the following identity can be easily obtained after differentiating the first equality in (51):

$$\begin{pmatrix} \alpha_{s_i} \\ \beta_{s_i} \end{pmatrix} = \begin{pmatrix} \frac{\partial \alpha}{\partial s_i}(T, s_1, s_2, \ldots, s_n) \\ \frac{\partial \beta}{\partial s_i}(T, s_1, s_2, \ldots, s_n) \end{pmatrix} = U \begin{pmatrix} 0_n \\ 2I_n^i \end{pmatrix}, \tag{112}$$

where 0_n is the $n \times n$ zero matrix and $I_n^i = \text{diag}(0, \ldots, 0, 1, 0, \ldots, 0)$, with the "1" in the ith position (the sub-index n will be dropped in the sequel). Then, by defining

$$\alpha_S \triangleq \sum_{i=1}^{n} \alpha_{s_i}, \qquad \beta_S \triangleq \sum_{i=1}^{n} \beta_{s_i}, \tag{113}$$

the next expression follows

$$\begin{pmatrix}\alpha_S\\ \beta_S\end{pmatrix} = U\begin{pmatrix}0\\ 2I\end{pmatrix}. \tag{114}$$

Then, (52)–(53) are all valid under the new meaning of α_S, β_S, especially (55), namely

$$\begin{pmatrix}\alpha_T\\ \beta_T\end{pmatrix} = \begin{pmatrix}\alpha_S\mathcal{M} - \alpha\mathcal{N}\\ \beta_S\mathcal{M} - \beta\mathcal{N}\end{pmatrix}. \tag{115}$$

Equation (115) should be integrated simultaneously with the first rows of both (60), (63) in order to avoid s_i-derivatives of ρ and σ.

Now, for a general (symmetric, positive semi-definite) matrix S it is well known that S can be diagonalized via an orthogonal matrix E (which depends on S), i.e.

$$\widetilde{S} \triangleq ESE' = \mathrm{diag}(s_1, \dots, s_n), \quad s_i \geq 0,\ i = 1, \dots, n.$$

Then, after defining

$$\begin{pmatrix}\widetilde{\alpha}(T,S)\\ \widetilde{\beta}(T,S)\end{pmatrix} \triangleq U\begin{pmatrix}I_n\\ 2\widetilde{S}\end{pmatrix} = \begin{pmatrix}V_4' - 2V_2'\widetilde{S}\\ -V_3' + 2V_1'\widetilde{S}\end{pmatrix}, \tag{116}$$

and maintaining the original definition (51) for α and β, and using definitions (113) and (114), the following relations can be easily derived:

$$\begin{pmatrix}\widetilde{\alpha}_{\widetilde{S}}\\ \widetilde{\beta}_{\widetilde{S}}\end{pmatrix} = U\begin{pmatrix}0\\ 2I\end{pmatrix} = \begin{pmatrix}-2V_2'\\ 2V_1'\end{pmatrix}, \tag{117}$$

$$\begin{aligned}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} &= U\begin{pmatrix}I\\ 2S\end{pmatrix} = \begin{pmatrix}V_4' - 2V_2'S\\ -V_3' + 2V_1'S\end{pmatrix} = \begin{pmatrix}\widetilde{\alpha} - 2V_2'(S - \widetilde{S})\\ \widetilde{\beta} + 2V_1'(S - \widetilde{S})\end{pmatrix}\\ &= \begin{pmatrix}\widetilde{\alpha} + \widetilde{\alpha}_{\widetilde{S}}(S - \widetilde{S})\\ \widetilde{\beta} + \widetilde{\beta}_{\widetilde{S}}(S - \widetilde{S})\end{pmatrix}.\end{aligned} \tag{118}$$

Therefore, (115) and the first rows of both (60), (63) can be integrated to obtain $\widetilde{\alpha}$, $\widetilde{\beta}$, $\widetilde{\alpha}_{\widetilde{S}}$, $\widetilde{\beta}_{\widetilde{S}}$, $\tilde{\rho}$ and $\tilde{\sigma}$. By replacing in (118), the values of α and β can be calculated, leading to approximate values for $\rho(T,S)$ and $\sigma(T,S)$, namely

$$\begin{pmatrix}\rho\\ \sigma\end{pmatrix} = \begin{pmatrix}\alpha^{-1}\tilde{\alpha}\tilde{\rho}\\ \beta\alpha^{-1}\tilde{\alpha}\tilde{\rho}\end{pmatrix}. \tag{119}$$

Some illustrations of non-scalar S-cases are presented in [26].

References

1. Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, New York (1978)

2. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mischenko, E.F.: The Mathematical Theory of Optimal Processes. Wiley, New York (1962)
3. Kalman, R.E., Falb, P.L., Arbib, M.A.: Topics in Mathematical System Theory. McGraw-Hill, New York (1969)
4. Bernhard, P.: Introducción a la teoría de control óptimo. Cuaderno Nro. 4, Instituto de Matemática "Beppo Levi", Rosario, Argentina (1972)
5. Sontag, E.D.: Mathematical Control Theory. Springer, New York (1998)
6. Costanza, V., Neuman, C.E.: Optim. Control Appl. Methods **27**, 41–60 (2006)
7. Costanza, V.: Lat. Am. Appl. Res. **35**(4), 327–335 (2005)
8. Pardalos, P.M., Yatsenko, V.: Optimization and Control of Bilinear Systems. Springer, Berlin (2008)
9. Costanza, V.: Optim. Control Appl. Methods **29**, 225–242 (2008)
10. Costanza, V., Rivadeneira, P.S.: Automatica **44**(9), 2427–2434 (2008)
11. Goh, C.J., Teo, K.L.: Automatica **24**(1), 3–18 (1998)
12. Chinchuluun, A., Pardalos, P., Enkhbat, R., Tseveendorj, I.: Optimization and Control: Theory and Applications. Springer Optimization and Its Applications, vol. 39. Springer, New York (2010)
13. Costanza, V.: Proceedings of the XIII Workshop on Information Processing and Control, Rosario-Argentina (2009)
14. Bellman, R., Kalaba, R.: Q. Appl. Math. **XXI**, 166–168 (1963)
15. Isidori, A.: Nonlinear Control Systems II. Springer, Berlin (1999)
16. Khalil, H.K.: Nonlinear Systems, 3rd edn. Prentice-Hall, Upper Saddle River (2002)
17. Jacobson, N.: Basic Algebra I. Freeman, San Francisco (1974)
18. Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. Cambridge University Press, Cambridge (1999)
19. Hirsch, M.W., Smale, S.: Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press, San Diego (1974)
20. Folland, G.B.: Introduction to Partial Differential Equations, 2nd edn. Princeton University Press, Princeton (1995)
21. Zenchuk, A.I., Santini, P.M.: Dressing method based on homogeneous Fredholm equation: quasilinear PDEs in multidimensions. arXiv:nlin/0701031 (2007)
22. Rao, A.V., Mease, K.D.: Automatica **35**, 633–642 (1999)
23. Costanza, V., Troparevsky, M.I., Rivadeneira, P.S.: Proceedings of the XIII Workshop on Information Processing and Control, Rosario-Argentina (2009)
24. Costanza, V.: Optim. Control Appl. Methods **28**, 209–228 (2007)
25. Ierapetritou, M.G., Pistikopoulos, E.N.: Comput. Chem. Eng. **19**, 627–632 (1995)
26. Costanza, V., Rivadeneira, P.S.: Proceedings of the XXI Congreso Argentino de Control Automático, Buenos Aires-Argentina, paper A001, Sept 2008