Math. Control Signals Systems (1992) 5: 233–261

Mathematics of Control, Signals, and Systems

Nonlinear H^∞ Control Theory for Stable Plants*

Joseph A. Ball† and J. William Helton‡

Abstract. We analyze various aspects of the nonlinear time-invariant H^∞ control problem in the discrete-time setting. A recipe is presented that is shown to generate a solution of the H^∞ problem in a precise but weak sense, and which is conjectured to generate a genuine solution in very general circumstances. The recipe involves a version of the Hamilton–Jacobi–Bellman–Isaacs equation from differential game theory. An illustrative example is presented.

Key words. H^∞-control, Differential games, Nonlinear optimal control.

1. Introduction

For mathematical convenience we deal only with discrete-time problems in this paper. For $\mathscr{C}$ any real Hilbert space (usually finite-dimensional), we denote by $l^{2+}_{\mathscr{C}}$ the space of $\mathscr{C}$-valued sequences $\vec{u} = \{\vec{u}(n)\}_{n\geq 0}$ that are norm square-summable. The following basic problem of H^∞ control theory is illustrated in Fig. 1.1 (see [F]).

(C) *Given a system $\mathscr{P}$, find a feedback K that produces an internally stable system whose input–output map T_{zw}: $\vec{w} \to \vec{z}$ satisfies*

$$\|T_{zw}(\vec{w})\|_{l^{2+}_Z} \leq \|\vec{w}\|_{l^{2+}_W}.$$

Here $\mathscr{P}$ and K are assumed to be causal time-invariant input–output (IO) maps. The input signals $\vec{w}$ and $\vec{u}$ for $\mathscr{P}$ are assumed to have values in finite-dimensional input spaces W and U, while the output signals $\vec{z}$ and $\vec{y}$ are assumed to have values in output spaces Z and Y. We demand that the system be well-posed so that the closed loop IO map T_{zw} is well defined, causal, and time-invariant. We do not insist that $\mathscr{P}$ or K be linear but do require that they take 0 to 0. In the linear case, internal stability of the closed loop system can be defined in two ways (see [F]). In terms of IO maps, it means that the output signal $\vec{z}$, as well as the internal signals $\vec{y}$ and $\vec{u}$, are in l^{2+}_Z, l^{2+}_Y, and l^{2+}_U, respectively, for any choice of l^{2+}_W-input $\vec{w}$, even in the

* Date received: July 20, 1990. Date revised: September 2, 1991.

† Department of Mathematics, Virginia Tech, Blacksburg, Virginia 24061, U.S.A. Phone (703) 231-7080.

‡ Department of Mathematics, University of California at San Diego, La Jolla, California 92093, U.S.A.

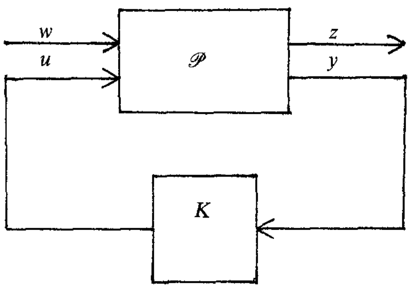

Fig. 1.1

presence of l^{2+}-perturbations of the internal signals. A state space formulation of internal stability is that the state for the composite closed-loop system in Fig. 1.1 tends to zero for any initial state if the input signal is $\vec{w} = 0$. These two definitions are equivalent in the linear case if we assume that we are working with minimal state space realizations for $\mathscr{P}$ and for K; see [S1], Section 6.3. Relations between input–output and internal stability were studied in some detail in [S2]. For the nonlinear case, we shall be working with a state space representation for $\mathscr{P}$; using a tilde to denote the state update operator, we have

$$\mathscr{P}: \begin{cases} \tilde{x} = F(x, w, u), \\ z = G_1(x, w, u), \\ y = G_2(x, w, u), \end{cases} \tag{1.1}$$

and we will seek a state space representation for K

$$\mathscr{K}: \begin{cases} \tilde{\xi} = f(\xi, y), \\ u = g(\xi, y), \end{cases} \tag{1.2}$$

so that the resulting closed-loop state space system is stable in one of the senses defined in Section 2.

In previous work [BH1–3] we presented a theory and a recipe which under certain hypotheses, the most crucial being the existence of a unique critical point for a certain energy function, led to a state space representation for a nonlinear fractional parametrization of a large class of solutions of (C). Construction of the nonlinear fractional parametrizer involved some work in addition to solving a critical point equation, namely, computation of a Morse-theoretic diffeomorphic change of variable. In practice, we may be interested in obtaining some particular solution of the problem in a simple way without doing the extra work required to obtain a nonlinear fractional parametrization of a whole class of solutions. In this paper, we show how solving a modified critical point equation leads to the construction of a particular solution (at least in a modified weak but suggestive sense) of the H^∞ control problem. Actually, we first consider a special case of the general problem, a *full information* version of the problem, where it is assumed that the sensor whose output is fed into the compensator K can measure the present value of the state x and the input w; this amounts to the special case of the general problem where the

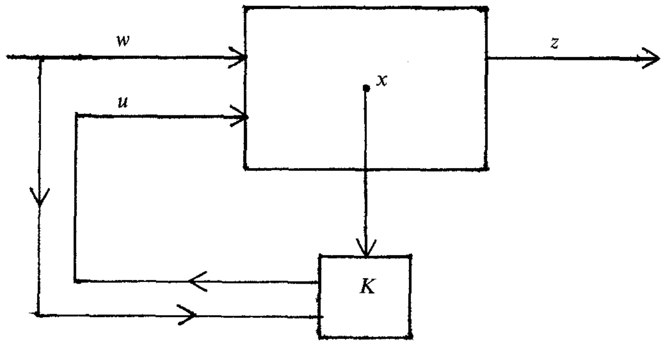

Fig. 1.2

compensator K is required to be memoryless and the second output y of $\mathscr{P}$ is assumed to be $\begin{bmatrix} x \\ w \end{bmatrix}$. This leads to the *full information control problem* (see Fig. 1.2), which plays a major role in [DGKF] for the linear case.

(FC) *Solve* (C) *for the special case where K is required to be memoryless and y is assumed to be given by* $y = \begin{bmatrix} x \\ w \end{bmatrix}$.

We next reduce the solution of the output feedback problem (C) to the (FC) problem under the assumption that we can solve for the disturbance from the measurement for each fixed value of the state; this corresponds to the 2-block case in the linear theory and includes most examples of applied interest (e.g., the mixed sensitivity problem for a square invertible plant). For a large class of linear mixed sensitivity problems, our recipe produces the usual maximum entropy or central solution.

Our goal is to give a recipe that produces a controller which solves (FC) and ultimately (C), in general, which is computationally implementable at least for systems having only "mild" nonlinearities. We shall give theorems that describe the range of validity of the recipe to a reasonable extent. Also, we work out in detail the example of a plant consisting of a linear system followed by a memoryless function M. We find that one instance where the recipe indeed leads to a solution is the case where $\|M(x)\|^2$ is strictly convex. At the other extreme, if M is a saturation nonlinearity, then serious difficulties arise.

Unlike the work in [BH1–3], in this paper we make use of the special structure for the critical point of an energy function, namely, that the critical point is a max–min point for the energy function. This leads to a simple direct proof that the full information controller constructed via our recipe leads to a passive closed-loop IO map T_{zw} as required in (FC).

Recently, a number of papers have explored the connections of linear H^∞ control with differential game theory (see [BO]). For example, [B] gives a game-theoretic interpretation of a state feedback H^∞ problem and obtains the solution as a direct application of game theory results, and [PS] provides a survey of the connection in the literature. This paper continues the development begun in [BH4] of such connections for the nonlinear setting. In the beginning, it was not at all clear which of the many possible information structures and games (as in [BO]) to associate with the conventional H^∞ control problem. For example, in [B] the author solves a state feedback H^∞ problem where the controller is assumed to have access to only the currect state and not the current disturbance; however, as was shown in [DGKF] and as we show here, the (FC) problem is the state feedback problem that serves as the stepping stone to the output feedback problem. Certain aspects of our recipe can already be found in [BO] in the context of a very general, but finite horizon, nonlinear differential game. Indeed, the equation for the unknown energy function e that we present is just a form of the classical Hamilton–Jacobi–Bellman–Isaacs equation of differential game theory, and the method of solving for the critical points of e via the Hamiltonian is also well known in differential game theory (see [BO]). The work in [BH1–3], originally motivated by a search for a nonlinear Beurling–Lax theory with applications, led to a rediscovery of some of these ideas in differential game theory. Our contribution here is to adapt the machinery from differential games to the setting of the infinite horizon nonlinear H^∞ control problem, where stability is a crucial consideration.

Since the first draft of this paper was submitted for publication, there have appeared at least two papers ([PS] and [IA]) that treat the H^∞ problem for continuous-time systems as an application of the nonlinear version of the bounded real lemma (see [W] and [HM]). This underlying idea is implicit in our recipe for the (FC) problem as well. Also, it is clear that certain formulas can be made more explicit if we assume that the state space representation of the plant $\mathscr{P}$ is affine with respect to the inputs. In [BHW] we present more details on these points for the continuous-time case.

2. Preliminaries

In this section we give some background material that will be needed for the exposition in later sections.

Much of our analysis depends on finding critical points for a smooth, real-valued function defined on a manifold. In general, if $\varphi: S \to \mathbb{R}$ is a real-valued function on a manifold S, a point s^* in S is said to be a *critical point* for φ if the gradient at s^*, $\nabla\varphi(s^*)$, vanishes

$$\nabla\varphi(s^*) = 0.$$

For us it is often more convenient to express this in terms of directional derivatives of φ at s^*:

$$D\varphi(s^*)[h] = 0 \qquad \text{for all} \quad h \in T_{s^*}S.$$

Here $D\varphi(s^*)[h]$ expresses the directional derivative of φ at s^* in the direction h, where h is a vector in the tangent space $T_{s^*}S$ of S at s^*. If s^* is a critical point for φ, we say the corresponding value $\varphi(s^*) \in \mathbb{R}$ is a *critical value* for φ.

In general, we may also take the second derivative of φ at a point s^* to get a bilinear form $(h, k) \to D^2\varphi(s^*)[h, k]$ on $T_{s^*}S \times T_{s^*}S$, called the Hessian of φ at s^*. If we assume that S is a Riemannian manifold, so that $T_{s^*}S$ has a Hilbert space inner product $\langle \ , \ \rangle$ defined on it, then with respect to some basis on $T_{s^*}S$ the bilinear form $D\varphi(s^*)$ on $T_{s^*}S \times T_{s^*}S$ is induced by a symmetric matrix A:

$$D^2\varphi(s^*)[h, k] = \langle Ah, k\rangle.$$

This matrix A we refer to simply as the Hessian matrix for φ at s^*, also denoted by $D^2\varphi(s^*)$. When S has a Cartesian product decomposition, $S = W \times U$, and $s^* = (w^*, u^*)$, then $T_{s^*}S = T_{w^*}W \times T_{u^*}U$ and we denote by $D^2_{uu}\varphi(s^*)$, $D^2_{uw}\varphi(s^*)$, $D^2_{ww}\varphi(s^*)$ the restrictions of $D^2\varphi(s^*)$ to natural coordinate subspaces of $T_{s^*}S$:

$$D^2_{ww}\varphi(s^*)[h_1, h_2] = D^2\varphi(s^*)[(h_1, 0), (h_2, 0)],$$

$$D^2_{wu}\varphi(s^*)[h_1, k_2] = D^2\varphi(s^*)[(h_1, 0), (0, k_2)],$$

$$D^2_{uu}\varphi(s^*)[k_1, k_2] = D^2\varphi(s^*)[(0, k_1), (0, k_2)],$$

for $h_1, h_2 \in T_{w^*}W$ and $k_1, k_2 \in T_{u^*}U$. These we refer to as the partial Hessians of φ at s^*.

When $S = W \times U$, it is natural to consider *max–min points*, i.e., points $s^* = (w^*, u^*)$ for which

$$\varphi(w^*, u^*) = \max_{w \in W} \min_{u \in U} \varphi(w, u). \tag{2.1}$$

By a *local max–min point*, we mean a point $s^* = (w^*, u^*)$ for which (2.1) holds at least in a local form, i.e.,

$$\varphi(w^*, u^*) = \max_{w \in W_{w^*}} \min_{u \in U_{u^*}} \varphi(w, u), \tag{2.2}$$

where W_{w^*} and U_{u^*} are neighborhoods of w^* and u^*, respectively. The following elementary result gives the precise connection between max–min points and critical points.

Lemma 2.1. *Let φ: $W \times U \to \mathbb{R}$ be a smooth function and suppose that the Hessian matrix $D^2_{uu}\varphi(w^*, u^*)$ of φ with respect to the u variable at (w^*, u^*) is invertible. If (w^*, u^*) is a local max–min point for φ, then:*

(i) *The point (w^*, u^*) is a critical point for φ.*
(ii) *The Hessian of φ with respect to u evaluated at (w^*, u^*) is positive semidefinite:*

$$D^2_{uu}\varphi(w^*, u^*)[k, k] \geq 0 \qquad \textit{for all} \quad k \in T_{u^*}U.$$

(iii) *The Schur complement is negative semidefinite:*

$$(D^2_{ww}\varphi - D^2_{wu}\varphi \cdot [D^2_{uu}\varphi]^{-1} \cdot [D^2_{uw}\varphi])(w^*, u^*)[h, h] \leq 0 \qquad \textit{for all} \quad h \in T_{w^*}W.$$

Conversely, if (i) *holds and* (ii) *and* (iii) *hold with strict inequality for $k \neq 0$, $h \neq 0$, then (w^*, u^*) is a local max–min point for φ.*

Proof. Assume first that (w^*, u^*) is a local max–min point for φ. Then, for each $w \in W_{w^*}$, there is a $u = u^*(w) \in U_{u^*}$ such that $\varphi(w, u^*(w)) = \min_{u \in U_{u^*}} \varphi(w, u)$ where by assumption $u^*(w^*) = u^*$. By the smoothness of φ we have $D_u\varphi(w, u^*(w)) = 0$. In particular, $D_u\varphi(w^*, u^*) = 0$. Since $D_{uu}\varphi(w^*, u^*)$ is invertible, by the implicit function theorem, $u^*(w)$ is uniquely determined in a sufficiently small neighborhood of u^* by the equation $D_u\varphi(w, u^*(w)) = 0$ and the function $w \to u^*(w)$ is also smooth. Since (w^*, u^*) is a max–min point for φ, we have

$$\varphi(w^*, u^*) = \max_{w \in W_{w^*}} \psi(w),$$

where $\psi(w) = \varphi(w, u^*(w))$. By the chain rule, ψ is also smooth and hence $D\psi(w^*) = 0$. On the other hand,

$$\begin{aligned} D\psi(w) &= D_w\varphi(w, u)|_{u=u^*(w)} + D_u\varphi(w, u)|_{u=u^*(w)} \cdot Du^*(w) \\ &= D_w\varphi(w, u)|_{u=u^*(w)}. \end{aligned}$$

When $w = w^*$, $u^*(w) = u^*$ and hence $D_w\varphi(w^*, u^*) = 0$. We conclude that (w^*, u^*) is a critical point of φ.

Since $\varphi(w^*, u^*) = \min_{u \in U_{u^*}} \varphi(w^*, u)$, it follows that $D_u^2\varphi(w^*, u^*)$ is positive semidefinite. Moreover, since $\varphi(w^*, u^*) = \max_{w \in W} \psi(w)$ it follows that $D^2\psi(w^*)$ is negative semidefinite. We can compute that $D^2\psi = D_{ww}^2\varphi - D_{wu}^2\varphi \cdot [D_{uu}^2\varphi]^{-1} \cdot D_{uw}^2\varphi$. This completes the proof of the direct side in Lemma 2.1.

For the converse, if (w^*, u^*) is a critical point and $D_{uu}^2\varphi(w^*, u^*)$ is positive definite, we again can solve uniquely for $u = u^*(w) \in U_{u^*}$ for each w in a neighborhood of W_{w^*} of w^* so that $D_u\varphi(w, u^*(w)) = 0$. By continuity, $D_{uu}^2\varphi(w, u^*(w))$ remains positive definite in a sufficiently small neighborhood, so $u^*(w)$ is a local minimum for $\varphi(w, \cdot)$ for each w in a sufficiently small neighborhood of w^*. By reversing the calculation in the first part of the proof and using that $D^2\psi(w^*) = (D_{ww}^2\varphi - D_{wu}^2\varphi \cdot [D_{uu}^2\varphi]^{-1} \cdot [D_{uw}^2\varphi])(w^*, u^*)$ is negative definite, we also see that $\psi(w^*) = \max_{w \in W_{w^*}} \psi(w)$ (as long as W_{w^*} is taken sufficiently small). ■

The following lemma enables us to interchange $\max_w$ and $\min_u$ in certain specialized situations. Parts of this result overlap Lemma 3.10 from [BH1].

Lemma 2.2. *Suppose that f is a smooth function defined on the Cartesian product of two manifolds $\mathscr{R}$ and $\mathscr{S}$. Suppose, in addition, that $\mathscr{R}$ and $\mathscr{S}$ have Cartesian product decompositions*

$$\mathscr{R} = \mathscr{R}_+ \times \mathscr{R}_-,$$

$$\mathscr{S} = \mathscr{S}_+ \times \mathscr{S}_-.$$

Suppose also that:

1. *f has an isolated max–min point at (r^*, s^*); in more detail, by this we mean that (r^*, s^*) has the form $((r_+^*, r_-^*), (s_+^*, s_-^*))$ and $\max_{r_+, s_+} \min_{r_-, s_-} f((r_+, r_-), (s_+, s_-)) = f((r_+^*, r_-^*), (s_+^*, s_-^*))$ where (r_+, r_-, s_+, s_-) vary over a neighborhood of $(r_+^*, r_-^*, s_+^*, s_-^*)$.*
2. *For each fixed $r = (r_+, r_-)$ in a neighborhood of (r_+^*, r_-^*) the function $\psi(s_+, s_-) \triangleq f((r_+, r_-), (s_+, s_-))$ has a unique max–min point at $\varphi(r) = (\varphi_+(r), \varphi_-(r))$:*

$$\max_{s_+} \min_{s_-} f(r, (s_+, s_-)) = f(r, (\varphi_+(r), \varphi_-(r))).$$

3. *The partial Hessian matrices*

$$D^2 f(r^*, s^*),\ D^2_{(r_-, s_-)^2} f((r_+^*, r_-^*), (s_+^*, s_-^*)),\ D^2_{s^{*2}} f(r^*, s^*),\ D^2_{s^{*2}} f((r_+^*, r_-^*), (s_+^*, s_-^*)),$$

are all invertible.

Then the function Ψ: $\mathscr{R}_+ \times \mathscr{R}_-$ *defined by*

$$\Psi(r_+, r_-) \triangleq f((r_+, r_-), (\varphi_+(r_+, r_-), \varphi_-(r_+, r_-)))$$

has a local max–min point at $r^* = (r_+^*, r_-^*)$

$$\max_{r_+} \min_{r_-} \Psi(r_+, r_-) = \Psi(r_+^*, r_-^*)$$

and $s^* = \varphi(r^*)$.

Proof. The proof involves a brute force verification, using an intricate Schur complement decomposition of the partial Hessian condition for a local max–min point given by Lemma 2.1. We give the details in the Appendix. ■

We now return to the setting of the control problem. Let us suppose that we are given IO map $\mathscr{P}$: $\vec{u} \to \vec{y}$ as in Fig. 2.1, which is modeled by state space equations

$$\tilde{x} = F(x, u), \qquad y = G(x, u). \tag{2.3}$$

Thus, if the initial state is x, the output sequence $\vec{y}$ is generated by the input sequence $\vec{u}$ according to the recursion

$$\begin{aligned} \vec{x}(n+1) &= F(\vec{x}(n), \vec{u}(n)), \qquad \vec{x}(0) = x, \\ \vec{y}(n) &= G(\vec{x}(n), \vec{u}(n)). \end{aligned} \tag{2.4}$$

We assume that the element 0 of the state space X is an equilibrium point; then $0 = F(0, 0)$, $0 = G(0, 0)$. Now suppose that we have specified an energy function e: $X \to \mathbb{R}^+$ for which $e(x) = 0$ if and only if $x = 0$. We say that the system (2.3) is *asymptotically e-controllable* if for each initial state x there is some choice of input sequence $\vec{u}_x \in l_U^{2+}$ for which the resulting sequence of states $\vec{x}$ generated by (2.4) has the property that

$$\lim_{n \to \infty} e(\vec{x}(n)) = 0. \tag{2.5}$$

Similarly, we say that (2.3) is *asymptotically e-stable* if (2.5) holds for *any* choice of input sequence $\vec{u} \in l_U^{2+}$. Next, we say that (2.3) is simply *e-stable* if at least

$$\sup_n e(\vec{x}(n)) < \infty$$

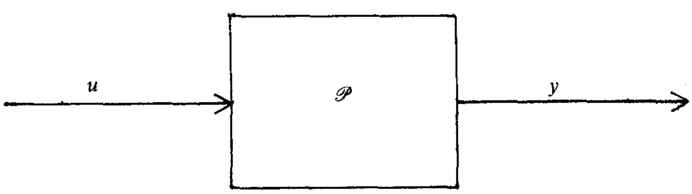

Fig. 2.1

for any choice of input sequence $\vec{u} \in l_U^{2+}$. By asymptotically controllable we mean that for each initial state x there is a choice of input sequence $\vec{u}_x \in l_U^{2+}$ for which the resulting sequence of states $\vec{x}$ itself tends to the equilibrium point 0

$$\lim_{n\to\infty} \vec{x}(n) = 0$$

in the state space. We define *asymptotically stable* and *stable* by similar modifications of the definitions of asymptotically e-stable and e-stable. The conventional notion of asymptotically stable, that $\vec{x}(n) \to 0$ for any initial state when $\vec{u} = \vec{0}$ is fed in, we shall call zero-input asymptotically stable. Note that if e is proper (inverse image of bounded sets is bounded) and if $x \neq 0$ implies $e(x) > 0$, then asymptotically e-controllable implies asymptotically controllable, and similarly for asymptotically e-stable and e-stable. Finally, we say that the system (2.3) is *input–output passive* if for any $\vec{u} \in l_U^{2+}$ we have

$$\sum_{k=0}^{n} \|\vec{y}(k)\|^2 \leq \sum_{k=0}^{n} \|\vec{u}(k)\|^2,$$

where the output sequence $\vec{y}$ is generated from (2.4) with $x = 0$. Alternatively, we may express this as

$$\|P_n\vec{y}\,\|_{l_Y^{2+}}^2 \leq \|P_n\vec{u}\,\|_{l_U^{2+}}^2,$$

where, in general,

$$(P_n\vec{x})(k) = \begin{cases} \vec{x}(k), & 0 \leq k \leq n, \\ 0, & k > n. \end{cases}$$

3. The Recipe and Main Results

3.A. *Recipe for Solving* (FC)

Now we turn to the recipe for solving (FC). We assume that we are given a state space representation as in (1.1) for the plant $\mathscr{P}$. The main unknown is a function $e\colon X \to \mathbb{R}^+$ on the state space; intuitively, we think of $e(x)$ as representing how much potential energy is in the state x, and the main equation is an energy balance inequality.

Recipe 3.1 for (FC).

1. Find a function $e\colon X \to \mathbb{R}^+$ satisfying

   (E) $$e(x) \geq \max_{w} \min_{u} \{e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2\}.$$

   With this choice of the function e define

   $$Q(x, w, u) \triangleq e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2.$$

   Denote the critical points by w_x^*, u_x^*.
2. Compute

   $$u^*(x, w) \triangleq \arg\min_{u} Q(x, w, u).$$

3. Use as the feedback law

$$u = u^*(x, w).$$

The output feedback control problem (C) can easily be reduced to the (FC) if we make some simplifying assumptions. Namely, we assume that G_2 is independent of u

$$y = G_2(x, w) \tag{3.1}$$

and, for each fixed x, that G_2 is a diffeomorphism as a function of w. Thus, there is a smooth function $G_2^I(x, \cdot)$ such that

$$y = G_2(x, G_2^I(x, y)), \qquad w = G_2^I(x, G_2(x, w)). \tag{3.2}$$

Then the same performance T_{zw} can be achieved by an output feedback with a dynamic compensator given by the following recipe. In the linear case these assumptions are generally satisfied for the 2-block case and fail to hold for the general 4-block case; we are content here to analyze a nonlinear generalization of the 2-block case.

Recipe 3.2 for (C). Assume (3.1) and (3.2).

1. Find e and u^* as in Steps 1 and 2 of Recipe 3.1.
2. Use as feedback law the dynamic compensator $\vec{u} = K(\vec{y})$ given by state space equations

$$\tilde{\xi} = F(\xi, G_2^I(\xi, y), u^*(\xi, G_2^I(\xi, y))), \qquad u = u^*(\xi, G_2^I(\xi, y)).$$

It is also possible to give continuous-time analogs of these recipes. In the continuous-time setting, the state space representation for the plant has the form

$$\mathscr{P}\colon \quad \dot{x} = F(x, w, u), \quad z = G_1(x, w, u), \quad y = G_2(x, w, u),$$

and we seek a state space representation for the compensator K of the form

$$K\colon \quad \dot{\xi} = f(\xi, y), \quad u = g(\xi, y).$$

If, in the recipes, (E) is replaced by

$$\text{(E–C)} \qquad 0 \geq \max_{w} \min_{u} \{\nabla e(x) \cdot F(x, w, u) + \|G_1(x, w, u)\|^2 - \|w\|^2\}$$

we obtain, at least formally, continuous-time analogs of all the results stated here explicitly for discrete time. We shall treat the continuous-time case in more detail in [BHW].

We now state our results concerning the validity of Recipe 3.1. The definitions of the various stability notions are given at the end of Section 2.

Theorem 3.1. *Suppose a function $e\colon X \to \mathbb{R}^+$ exists and $u^*(x, w)$ is defined as in Recipe 3.1. Then the closed-loop transfer function T_{zw}, as in Fig. 1.2 with $u = u^*(x, w)$, is input–output passive, i.e.,*

$$\|P_n T_{zw}(\vec{w})\|^2_{l_Y^{2+}} \leq \|P_n \vec{w}\|^2_{l_W^{2+}} \qquad \textit{for} \quad n = 0, 1, 2, \ldots.$$

Consequently,

$$\|T_{zw}(\vec{w})\|^2_{l^{2+}_Z} \leq \|\vec{w}\|^2_{l^{2+}_W}$$

for all $\vec{w} \in l^{2+}_W$. *Moreover, the state space system in Fig. 1.2 is e-stable. Suppose, in addition, that* $e(x) > 0$ *for* $x \neq 0$, *that* e *is proper, and that the closed-loop system is detectable, in the sense that*

$$\vec{x}(k+1) = F(\vec{x}(k), 0, u^*(\vec{x}(k), 0)), \qquad 0 = G_1(\vec{x}(k), 0, u^*(\vec{x}(k), 0))$$

$$\textit{implies} \quad \lim_{k \to 0} \vec{x}(k) = 0.$$

Then the closed-loop system is, in addition, zero-input asymptotically stable.

Proof. By assumption, $e\colon X \to \mathbb{R}^+$ satisfies (E), so

$$e(x) \geq \max_w \min_u Q(x, w, u),$$

where

$$Q(x, w, u) = e(F(x, w, u)) + \|G_1(x, w, u)\|^2_Z - \|w\|^2_W.$$

Hence, for any fixed x and w,

$$e(x) \geq \min_u Q(x, w, u) = Q(x, w, u^*(x, w)).$$

Plugging in the definition of Q gives $e(x) \geq e(F(x, w, u^*)) + \|G_1(x, w, u^*)\|^2 - \|w\|^2$, so

$$e(x) - e(F(x, w, u^*)) \geq \|G_1(x, w, u^*)\|^2 - \|w\|^2 \tag{3.3}$$

for all $x \in X$ and $w \in W$ where $u^* = u^*(x, w)$ is given as in Recipe 3.1. Now let $\vec{w} = \{\vec{w}(n)\}_{n \geq 0}$ be an input string for the closed-loop system T_{zw}. The resulting output string $\vec{z} = \{\vec{z}(n)\}_{n \geq 0}$ is determined recursively by

$$\vec{x}(n+1) = F(\vec{x}(n)), \vec{w}(n), u^*(\vec{x}(n), \vec{w}(n))), \qquad \vec{x}(0) = 0,$$

$$\vec{z}(n) = G_1(\vec{x}(n), \vec{w}(n), u^*(\vec{x}(n), \vec{w}(n))).$$

Comparing with (3.3) we see that

$$e(\vec{x}(k)) - e(\vec{x}(k+1)) \geq \|\vec{z}(k)\|^2 - \|\vec{w}(k)\|^2 \tag{3.4}$$

for all k. Summing from $k = 0$ to $k = n$ gives

$$e(\vec{x}(0)) - e(\vec{x}(n+1)) \geq \sum_{k=0}^{n} \|\vec{z}(k)\|^2 - \sum_{k=0}^{n} \|\vec{w}(k)\|^2. \tag{3.5}$$

By assumption, $e(\vec{x}(0)) = 0$ and $e(x) \geq 0$ for all x. We conclude that

$$\sum_{k=0}^{n} \|\vec{z}(k)\|^2 \leq \sum_{k=0}^{n} \|\vec{w}(k)\|^2$$

whenever $\vec{z} = T_{zw}(\vec{w})$. This shows that T_{zw} is input–output passive as asserted.

To prove e-stability of the system in Fig. 1.2, observe that (3.5) implies that

$$\begin{aligned} e(\vec{x}(0)) + \sum_{k=0}^{n} \|\vec{w}(k)\|^2 &\geq \sum_{k=0}^{n} \|\vec{z}(k)\|^2 + e(\vec{x}(n+1)) \\ &\geq e(\vec{x}(n+1)). \end{aligned}$$

Hence $\vec{w} \in l_W^{2+}$ implies that $\sup_n e(\vec{x}(n)) < \infty$, as required.

The remaining assertion of the theorem follows by a standard type of Lyapunov argument (see [PS] and [IA] for similar type arguments). If we specialize the argument above to the case $\vec{w}(k) = 0$, from (3.4) we see that $e(\vec{x}(k))$ is decreasing in k and hence $\lim_{k\to\infty} e(\vec{x}(k)) = e_*$ exists. If e is proper, from the boundedness of $e(\vec{x}(k))$, we conclude that $\vec{x}(k)$ is bounded, and hence that there is a ω-limit set Ω associated with the orbit $\{\vec{x}(k)\}$ which is invariant under the closed-loop dynamics (with $\vec{w}(k) = 0$) and in which e has the constant value e^*. From (3.4) again we conclude that $G_1(x, 0, u^*(x, 0)) = 0$ for each x in Ω. Now, from the detectability assumption, it follows that $\lim_{k\to\infty} \vec{x}(k) = 0$ if $\vec{x}(0) \in \Omega$ and hence $e_* = 0$. Thus $\lim_{k\to\infty} e(\vec{x}(k)) = 0$ for the original orbit (with an arbitrary value of $\vec{x}(0)$) as well. Finally, since e is proper and $x \neq 0$ implies $e(x) > 0$, we conclude that $\lim_{k\to\infty} \vec{x}(k) = 0$ as well. ■

3.B. The Input–Output Approach

To implement Recipe 3.1 we must be able to find a function e satisfying (E) to complete the first step. We next give a formula for a function e that automatically satisfies (E) whenever it is well defined. In the course of doing this we develop a correspondence that is important in game theory and dynamic programming (see [BO]) (an area where stability considerations for the infinite horizon context have received little attention). As an additional dividend, this correspondence enables us to derive another stability result for Recipe 3.1.

Denote by $\mathscr{F}_x^{\mathscr{P}}$ the input–output map associated with the system $\mathscr{P}$ with initial state x, i.e., if $\begin{bmatrix} \vec{w} \\ \vec{u} \end{bmatrix}$ if a sequence (also called a string) of inputs in $l_W^+ \oplus l_U^+$, then $\begin{bmatrix} \vec{z} \\ \vec{y} \end{bmatrix} = \begin{bmatrix} \mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}, \vec{u}) \\ \mathscr{F}_{x2}^{\mathscr{P}}(\vec{w}, \vec{u}) \end{bmatrix} = \mathscr{F}_x^{\mathscr{P}}(\vec{w}, \vec{u})$ is the associated sequence of outputs generated with initial state equal to x defined recursively by

$$\begin{aligned} \vec{x}(n+1) &= F(\vec{x}(n), \vec{w}(n), \vec{u}(n)), & \vec{x}(0) &= x, \\ \vec{z}(n) &= G_1(\vec{x}(n), \vec{w}(n), \vec{u}(n)), & \vec{y}(n) &= G_2(\vec{x}(n), \vec{w}(n), \vec{u}(n)). \end{aligned} \tag{3.6}$$

Let $\mathscr{D}_x \subset l_W^{2+} \oplus l_U^{2+}$ be the set of l^{2+}-sequences $\begin{bmatrix} \vec{w} \\ \vec{u} \end{bmatrix}$ for which the associated output sequence $\vec{z}$ is also norm-square-summable:

$$\mathscr{D}_x = \left\{ \begin{bmatrix} \vec{w} \\ \vec{u} \end{bmatrix} \in l_W^2 \oplus l_U^2 \colon \mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}, \vec{u}) \in l_Z^{2+} \right\}.$$

We now provide a link between critical points for two different types of energy functions. In general, if $x \in X$ and $\vec{x} \in l_X^2$, by $(x, \vec{x})$ we denote the sequence

$$(x, \vec{x})(k) = \begin{cases} x, & k = 0, \\ \vec{x}(k-1), & k \geq 1. \end{cases}$$

Proposition 3.2. *Define a function*

$$\mathscr{E}: \left\{(x, \vec{w}, \vec{u}): x \in X, \begin{bmatrix} \vec{w} \\ \vec{u} \end{bmatrix} \in \mathscr{D}_x \right\} \to \mathbb{R}$$

by

$$\mathscr{E}(x, \vec{w}, \vec{u}) = \|\mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}, \vec{u})\|_{l_Z^{2+}}^2 - \|\vec{w}\|_{l_W^{2+}}^2$$

and suppose that:

(S) *For each $x \in X$ there is a choice of isolated critical point $(\vec{w}_x^*, \vec{u}_x^*)$, depending smoothly on x, in the interior of $\mathscr{D}_x$ for $\mathscr{E}(x, \cdot, \cdot)$ which is a local max–min point for $\mathscr{E}(x, \cdot, \cdot)$:*

$$\mathscr{E}(x, \vec{w}_x^*, \vec{u}_x^*) = \max_{\vec{w} \in W_x} \min_{\vec{u} \in U_x} \mathscr{E}(x, \vec{w}, \vec{u}),$$

where W_x and U_x are neighborhoods of $\vec{w}_x^*$ and $\vec{u}_x^*$ in l_W^{2+} and l_U^{2+}, respectively. Define a function $e: X \to \mathbb{R}$ by

$$e(x) = \mathscr{E}(x, \vec{w}_x^*, \vec{u}_x^*).$$

Then e satisfies (E) *with equality locally, i.e., for each $x \in X$ the function*

$$Q(x, w, u) = e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2$$

has a critical point (w_x^*, u_x^*) which is a local max–min for $Q(x, \cdot, \cdot)$ and

$$e(x) = Q(x, w_x^*, u_x^*) = \max_{w \in Q_w} \min_{u \in Q_u} Q(x, w, u)$$

(here Q_w and Q_u are open neighborhoods of w_x^* and u_x^*, respectively). Moreover, the critical point $(\vec{w}_x^*, \vec{u}_x^*)$ for $\mathscr{E}(x, \cdot, \cdot)$ and the critical point (w_x^*, u_x^*) for $Q(x, \cdot, \cdot)$ are connected in the following way:

$$\vec{w}_x^* = (w_x^*, \vec{w}_{F(x, w_x^*, u_x^*)}^*), \tag{3.7}$$

$$\vec{u}_x^* = (u_x^*, \vec{u}_{F(x, w_x^*, u_x^*)}^*). \tag{3.8}$$

Remark. Equations (3.7) and (3.8) tell us how to get $\begin{bmatrix} w_x^* \\ u_x^* \end{bmatrix}$ from $\begin{bmatrix} \vec{w}_x^* \\ \vec{u}_x^* \end{bmatrix}$. Conversely, under the assumption of Proposition 3.2, we can recover $\begin{bmatrix} \vec{w}_x^* \\ \vec{u}_x^* \end{bmatrix}$ from the function $x \to \begin{bmatrix} w_x^* \\ u_x^* \end{bmatrix} \in W \times U$ as the solution of the recurrence equations

$$\vec{x}^*(k+1) = F(\vec{x}^*(k), \vec{w}_x^*(k), \vec{u}_x^*(k)), \qquad \vec{x}^*(0) = x,$$

$$\begin{bmatrix} \vec{w}_x^*(k) \\ \vec{u}_x^*(k) \end{bmatrix} = \begin{bmatrix} w_{\vec{x}^*(k)}^* \\ u_{\vec{x}^*(k)}^* \end{bmatrix}.$$

If we only assume that e is known and that the function $Q(x, w, u) = e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2$ has a max–min point (w_x^*, u_x^*) for each x, the above recursion generates a candidate for a max–min point for $\mathscr{E}(x, \vec{w}, \vec{u})$. A natural stability constraint for the problems here is that the resulting sequence $\begin{bmatrix} \vec{w}_x^* \\ \vec{u}_x^* \end{bmatrix}$ is in $l_W^2 \oplus l_U^2$.

Remark. The assumption (S) is not as strong as it may seem. Existence of $\vec{u}_x^*$ can be proved as follows. Since $\mathscr{F}_{x1}(x, \vec{w}, \cdot)$ is causal, we can show that it is weak-$*$ continuous. We can then get the existence of $\vec{u}_x^*$ from the weak-$*$ compactness of bounded sets and the semicontinuity of the norm function $\|\cdot\|$ with respect to the weak-$*$ topology.

With the assumption (S) we now obtain the following stability result for the full information controller constructed in the Recipe 3.1. The proof will be given in the next section.

Theorem 3.3. *Let $\mathscr{P}$ be as in* (FC), *assume that* (S) *holds, and that the full information feedback $u = u^*(x, w)$ is constructed as in Recipe* 3.1 (*where equality holds in* (E)). *Then the closed-loop system in Fig.* 1.2 *is asymptotically e-controllable for the e given by $e(x) = \mathscr{E}(x, \vec{w}_x^*, \vec{u}_x^*)$.*

Internal stability for the closed-loop system in Fig. 3.1, with compensator constructed as in Recipe 3.2, appears to be more delicate. We are able to get a precise result if we impose an additional hypothesis which leads to a feedback configuration of a model matching problem (see [F] for the linear case):

(MM) $$G_2(x, w, u) = G_2(w).$$

We have the following result.

Theorem 3.4. *Let $\mathscr{P}$ be a plant with state space equations of the form*

$$\tilde{x} = F(x, w, u), \qquad z = G_1(x, w, u), \qquad y = G_2(x, w),$$

suppose that (3.1) *and* (3.2) *hold, and assume $(x, w) \to u^*(x, w)$ is constructed as in Recipe* 3.1. *Then the compensator K, defined by state space equations as in Recipe* 3.2 *used as an output feedback in Fig.* 1.1, *induces the same IO map T_{zw} for the closed-loop system, as does the full information controller given by Recipe* 3.1. *If, in addition, $\mathscr{P}$ is in the special form* (MM) *and the closed loop full information configuration in Fig.* 1.2 *with K constructed via Recipe* 3.1 *is asymptotically e-stable, then the closed-loop output feedback configuration with compensator K as above is also asymptotically e-stable.*

We conjecture that the internal stability result in Theorem 3.4 in fact holds under the weaker conditions that the system with input variable $\begin{bmatrix} u \\ y \end{bmatrix}$ and output variable $\begin{bmatrix} z \\ w \end{bmatrix}$ given by state space equations

$$\tilde{x} = F(x, G_2^I(x, y), u), \qquad z = G_1(x, G_2^I(x, y), u), \qquad w = G_2^I(x, y),$$

be stable. This is consistent with the linear case and is the reason for the qualification "stable" in the title of this paper.

Proofs of Theorems 3.2–3.4 in this section will be discussed in the next section. We also mention that if assumption (3.1) fails, we replace assumption (3.2) with the assumption (3.1′) that the function $G_2(x, w) \triangleq G_2(x, w, u^*(x, w))$ be a diffeomorphism as a function of w for each fixed x, i.e., that there is a function $\tilde{G}_2^I(x, y)$ for which

$$y = \tilde{G}_2(x, \tilde{G}_2^I(x, y)), \qquad w = \tilde{G}_2^I(x, \tilde{G}_2(x, w)). \tag{3.2′}$$

Then replacing G_2^I by $\tilde{G}_2^I$ gives a recipe for the output feedback problem under the more general situation allowed by assumption (3.1′) and (3.2′), which is consistent with the central solution in the linear case.

4. Proofs of the Main Results

In this section we turn to the proofs of the results announced in Section 3.

Proof of Proposition 3.2. Apply Lemma 2.2 to the function

$$f((w, u), (\vec{\beta}, \vec{\alpha})) = \mathscr{E}(x, (w, \vec{\beta}), (u, \vec{\alpha})).$$

Here $w \in W$, $u \in U$, $\vec{\beta} \in l_W^{2+}$, $\alpha \in l_U^{2+}$, and

$$(w, \vec{\beta})(k) = \begin{cases} w, & k = 0, \\ \vec{\beta}(k-1), & k \geq 1. \end{cases}$$

Note that from the definition of $\mathscr{E}$ in Proposition 3.2 and the recursive definition of the IO map $\mathscr{F}_{x1}^{\mathscr{P}}$ we have

$$\mathscr{E}(x, (w, \vec{\beta}), (u, \vec{\alpha})) = \mathscr{E}(F(x, w, u), \vec{\beta}, \vec{\alpha}) + \|G_1(x, w, u)\|^2 - \|w\|^2.$$

For fixed (w, u), the associated function $\psi_{(w,u)}(\vec{\beta}, \vec{\alpha})$, as in Lemma 2.2, is given by

$$\psi_{(w,u)}(\vec{\beta}, \vec{\alpha}) = \mathscr{E}(F(x, w, u), \vec{\beta}, \vec{\alpha}) + \|G_1(x, w, u)\|^2 - \|w\|^2.$$

Note that the last two terms are independent of $\vec{\beta}$, $\vec{u}$. Thus, by definition of the function $x \to e(x)$ in Proposition 3.2, we see that

$$\begin{aligned} Q(x, w, u) &\triangleq \max_{\vec{\beta}} \min_{\vec{\alpha}} \psi_{(w,u)}(\vec{\beta}, \vec{\alpha}) \\ &= e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2, \end{aligned} \tag{4.1}$$

and that the max–min occurs at the point $(\vec{\beta}, \vec{\alpha}) = (\vec{w}^*_{F(x,w,u)}, \vec{u}^*_{F(x,w,u)})$. Thus, in the notation of Lemma 2.2,

$$\varphi(w, u) = (\vec{w}^*_{F(x,w,u)}, \vec{u}^*_{F(x,w,u)}). \tag{4.2}$$

Also, in the notation of Lemma 2.2, the function $f((r_+, r_-), (s_+, s_-))$ has a $\max_{r_+, s_+}$–$\min_{r_-, s_-}$ point at $(r^*, s^*) = (r^*, \varphi(r^*))$ while the function $\Psi(r_+, r_-)$ has a max–min point at r^*. This general observation, applied to our setting here gives us that the critical point $(\vec{w}_x^*, \vec{u}_x^*)$ for $\mathscr{E}(x, \cdot, \cdot)$, has the form

$$\vec{w}_x^* = (w_x^*, \vec{w}^*_{F(x,w,u)}), \tag{4.3}$$

$$\vec{u}_x^* = (u_x^*, \vec{u}^*_{F(x,w,u)}), \tag{4.4}$$

where $(w, u) = (w_x^*, u_x^*)$ is the local max–min point for the associated function $Q(x, w, u)$ given by (4.1). Also, in Lemma 2.2, we had

$$\begin{aligned} f((r_+^*, r_-^*), (s_+^*, s_-^*)) &= \max_{r_+, s_+} \min_{r_-, s_-} f((r_+, r_-), (s_+, s_-)) \\ &= \max_{r_+} \min_{r_-} \left\{ \max_{s_+} \min_{s_-} f((r_+, r_-), (s_+, s_-)) \right\}, \end{aligned}$$

which for our setting here becomes

$$\begin{aligned} e(x) &= \min_{(w,\vec{\beta})} \min_{(u,\vec{\alpha})} \mathscr{E}(x, (w, \vec{\beta}), (u, \vec{\alpha})) \\ &= \max_{w} \min_{u} \{ e(F(x, w, u) + \|G_1(x, w, u)\|^2 - \|w\|^2 \}. \end{aligned} \tag{4.5}$$

Since the right side of (4.5) is only a local max–min, this shows that e satisfies a local form of (E) with equality. Finally, formulas (3.7) and (3.8) follow from (4.3) and (4.4), and Proposition 3.2 follows. ■

We are now in a position to prove Theorem 3.3.

Proof of Theorem 3.3. For each $x \in X$, let $(\vec{w}_x^*, \vec{u}_x^*)$ be the max–min critical point of $\mathscr{E}(x, \cdot, \cdot)$ as in (S). We show that the input string $\vec{w}_x = \vec{w}_x^* \in l_W^{2+}$ has all the properties required in Theorem 3.3. Using $\vec{w}_x^*$ as the input string, the evolution of the state in the closed-loop system with full information control given by Recipe 3.1 is given by

$$\vec{x}(k+1) = F(\vec{x}(k), \vec{w}_x^*(k), u^*(\vec{x}(k), \vec{w}_x^*(k))), \qquad \vec{x}(0) = x.$$

For $k = 0$, we have

$$\vec{x}(1) = F(x, \vec{w}_x^*(0), u^*(x, \vec{w}_x^*(0))), \qquad \vec{x}(0) = x.$$

By (3.7) we know that $\vec{w}_x^*(0) = w_x^*$. From the definition of the function $(x, w) \to u^*(x, w)$ and by (3.8), it is easy to see that

$$u^*(x, w_x^*) = u_x^* = \vec{u}_x^*(0).$$

Thus

$$\vec{x}(1) = F(x, \vec{w}_x^*(0), \vec{u}_x^*(0)).$$

Also, from (3.7) and (3.8), we see that

$$\vec{w}^*_{\vec{x}(1)} = S^* \vec{w}_x^*, \qquad \vec{u}^*_{\vec{x}(1)} = S^* \vec{u}_x,$$

where S^* is the backwards unilateral shift operator. Hence, from the recursive definition of the IO map $\mathscr{F}_{x1}^{\mathscr{P}}: (\vec{w}, \vec{u}) \to z$ (see (3.6)), we get

$$\begin{aligned} e(\vec{x}(1)) &= \|\mathscr{F}_{\vec{x}(1),1}^{\mathscr{P}}(\vec{w}_{\vec{x}(1)}^*, \vec{u}_{\vec{x}(1)}^*)\|^2 - \|\vec{w}_{\vec{x}(1)}^*\|^2 \\ &= \|S^*\mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}_x^*, \vec{u}_x^*)\|^2 - \|S^*\vec{w}_x^*\|^2. \end{aligned}$$

Using a similar argument, by induction we can show that, in general,

$$e(\vec{x}(n)) = \|S^{*n}\mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}_x^*, \vec{u}_x^*)\|^2 - \|S^{*n}\vec{w}_x^*\|^2. \tag{4.6}$$

Since $\mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}_x^*, \vec{u}_x^*) \in l_Z^{2+}$ and $\vec{w}_x^* \in l_U^{2+}$ it follows now from (4.6) that

$$\lim_{n\to\infty} e(\vec{x}(n)) = 0,$$

as asserted in Theorem 3.3. ■

Proof of Theorem 3.4. The composite system in Fig. 1.1 with $\mathscr{P}$ and K as in the theorem is governed by state space equations

$$\begin{aligned} &\tilde{x} = F(x, w, u^*(\xi, G_2^I(\xi, y))), \qquad \tilde{\xi} = F(\xi, w, u^*(\xi, G_2^I(\xi, y))), \\ &z = G_1(x, w, u^*(\xi, G_2^I(\xi, y))), \end{aligned} \tag{4.7}$$

where

$$y = G_2(x, w).$$

If $x = \xi$, we see from the first two equations that $\tilde{x} = \tilde{\xi}$. Thus, if $\vec{x}(0) = \vec{\xi}(0) = 0$ and a sequence $\vec{w} \in l_W^+$ is fed in, it follows that $\vec{x}(k) = \vec{\xi}(k)$ for all $k = 0, 1, 2, \ldots$, and hence

$$\vec{u}(k) = u^*(\vec{\xi}(k), G_2^I(\vec{\xi}(k), \vec{y}(k))) = u^*(\vec{x}(k), G_2^I(\vec{x}(k), \vec{y}(k))) = u^*(\vec{x}(k), \vec{w}(k))$$

for all k. From this we easily see that the closed-loop IO map T_{zw} with the output feedback K is the same as with the full information controller $u = u^*(x, w)$ in Recipe 3.1.

Now, assume, in addition, that $G_2(x, w) = G_2(w)$, and that the full information scheme in Recipe 3.1 is asymptotically e-stable. Then, for *any* input string $\vec{w} \in l_W^{2+}$ and *any* initial state x, $\lim_{k\to\infty} e(\vec{x}(k)) = 0$ if $\{\vec{x}(k)\}_{k\geq 0}$ is generated by

$$\vec{x}(k+1) = F(\vec{x}(k), \vec{w}(k), u^*(\vec{x}(k), \vec{w}(k))). \tag{4.8}$$

Now, if the sequence $\vec{w}$ is fed into the system (4.7) which is in initial state $\begin{bmatrix} x \\ \xi \end{bmatrix}$, then the resulting sequences of states $\vec{x}$ and $\vec{\xi}$ are generated by (4.8) but with different initializations $\vec{x}(0) = x$ and $\vec{\xi}(0) = \xi$. From the e-stability of the full information configuration, we now see that both

$$\lim_{k\to\infty} e(\vec{x}(k)) = 0$$

and

$$\lim_{k\to\infty} e(\vec{\xi}(k)) = 0,$$

and hence the system (4.7) is e-stable in this case. ■

5. A Practical Refinement

The most difficult part of the recipe is computing the energy function $e\colon X \to \mathbb{R}^+$. Once e is computed we must compute max–min points (w_x^*, u_x^*) for the function

$$Q(x, w, u) = e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2.$$

We saw in Proposition 3.2 how these critical points may arise from max–min points $(\vec{w}_x^*, \vec{u}_x^*)$ for an energy function $\mathscr{E}$ defined on l^{2+}-strings:

$$\mathscr{E}(x, \vec{w}, \vec{u}) = \|\mathscr{F}_{x1}^{\mathscr{P}}(\vec{w}, \vec{u})\|^2 - \|\vec{w}\|^2.$$

We now show an alternative method for finding the max–min points (w_x^*, u_x^*) and $(\vec{w}_x^*, \vec{u}_x^*)$ which bypasses having to first solve the (E) equation for e. In this method, we introduce the Hamiltonian associated with the system $\mathscr{P}$ by

$$H(p, x, w, u) = p \cdot F(x, w, u) + \|G_1(x, w, u)\|^2 - \|w\|^2. \tag{5.1}$$

Here we assume that the state space X is $\mathbb{R}^n$ and p is an n-dimensional column vector. We shall show that $(\vec{w}_x^*, \vec{u}_x^*)$ can be found by finding an appropriate solution to the following coupled system of recurrence relations, called the *associated Hamiltonian system*:

$$(\vec{w}_x^*(k), \vec{u}_x^*(k)) = \underset{(w,u)}{\arg\operatorname{crit}}\, H(\vec{p}(k+1), \vec{x}^*(k), w, u), \tag{5.2}$$

$$\vec{x}^*(k+1) = F(\vec{x}^*(k), \vec{w}_x^*(k), \vec{u}_x^*(k)), \qquad \vec{x}^*(0) = x, \tag{5.3}$$

$$\vec{p}(k) = \nabla_x H(\vec{p}(k+1), \vec{x}^*(k), \vec{w}_x^*(k), \vec{u}_x^*(k)), \tag{5.4}$$

and

$$\lim_{k\to\infty} \vec{p}(k) = 0. \tag{5.5}$$

Theorem 5.1. *Suppose a smooth solution $e\colon X \to \mathbb{R}^+$ of* (E) *with equality exists. Then there exists a sequence $\vec{p} = \{\vec{p}(k)\}_{k\geq 0}$ of costate vectors and a sequence $\vec{x} = \{\vec{x}^*(k)\}_{k\geq 0}$ of state vectors such that $\{\vec{p}_x, \vec{x}^*, \vec{w}_x^*, \vec{u}_x^*\}$ satisfies the Hamiltonian system equations* (5.2), (5.3), *and* (5.4). *If $\lim_{k\to\infty} \vec{x}^*(k) = 0$ and 0 is a critical point of e, then also* (5.5) *is satisfied. Moreover, the controller $u = u^*(x, w)$ in Recipe* 3.1 *is a solution of the equation*

$$D_u F(x, w, u)^T[\vec{\tilde{p}}(0)] + 2D_u G_1(x, w, u)^T[G_1(x, w, u)] = 0, \tag{5.6}$$

where $\vec{\tilde{p}}(k)$ refers to the solution $\vec{p}(k)$ of (5.1)–(5.5) *with initial condition $\vec{x}^*(0) = F(x, w, u)$.*

Proof. Let us set

$$Q(x, w, u) = e(F(x, w, u)) + \|G_1(x, w, u)\|^2 - \|w\|^2. \tag{5.7}$$

Comparing (5.6) and (5.1) we see that

$$D_{(w,u)} H(\tilde{p}, x, w, u) = D_{(w,u)} Q(x, w, u), \tag{5.8}$$

if we set $\tilde{p} = \nabla e(F(x, w, u))$. Hence, if we set

$$(w_x^*, u_x^*) = \arg \max_w \min_u Q(x, w, u), \tag{5.9}$$

(5.7) gives

$$(w_x^*, u_x^*) = \arg \operatorname*{crit}_{(w,u)} H(\tilde{p}, x, w, u) \qquad \text{when} \quad \tilde{p} = \nabla e(F(x, w_x^*, u_x^*)). \tag{5.10}$$

Now, in general, define sequences $\vec{x}^*$, $\vec{w}_x^*$, $\vec{u}_x^*$, and $\vec{p}$ by the equations

$$(\vec{w}_x^*(k), \vec{u}_x^*(k)) = \arg \max_w \min_u Q(\vec{x}(k), w, u), \tag{5.11}$$

$$\vec{x}^*(k+1) = F(\vec{x}^*(k), \vec{w}_x^*(k), \vec{u}_x^*(k)), \qquad \vec{x}^*(0) = x, \tag{5.12}$$

$$\vec{p}(k) = \nabla e(\vec{x}^*(k)). \tag{5.13}$$

Then, by substituting (5.11)–(5.13) into (5.9) and (5.10), we verify (5.2), while (5.3) is the same as (5.12). To prove (5.4), we use the assumption that e satisfies (E) with equality

$$\begin{aligned} e(x) &= \max_w \min_u Q(x, w, u) \\ &= Q(x, w_x^*, u_x^*), \end{aligned}$$

where Q is given by (5.7). Differentiate both sides with respect to x and use that (w_x^*, u_x^*) is a critical point for $Q(x, \cdot, \cdot)$ to get

$$\begin{aligned} De(x) &= D_x Q(x, w_x^*, u_x^*) \\ &= De(F(x, w_x^*, u_x^*)) \cdot D_x F(x, w_x^*, u_x^*) + 2DG_1(x, w_x^*, u_x^*)[G_1(x, w_x^*, u_x^*)]^T \\ &\quad - 2w^T. \end{aligned}$$

Take transposes and use (5.11), (5.3), and (5.13) and the definition (5.1) of H to get (5.4). Finally, if $\lim_{k \to \infty} \vec{x}^*(k) = 0$ and 0 is assumed to be a critical point for e, we get that

$$\lim_{k \to \infty} \vec{p}(k) = \lim_{k \to \infty} \nabla e(\vec{x}^*(k)) = 0.$$

Finally, by definition, $u = u^*(x, w)$ is a critical point of $Q(x, w, u)$ with respect to the variable u. Thus $u = u^*(x, w)$ satisfies

$$D_u Q(x, w, u) = 0. \tag{5.14}$$

By substituting in (5.7) for Q and using that, in general, $\vec{p}(0) = \nabla e(x)$ if $\vec{x}^*(0) = x$, we get (5.6) from (5.14). ∎

A crucial point in the proof of Theorem 3.1 is that the values of e are nonnegative real numbers, i.e., that e have a global minimum at 0. A necessary condition for 0 to be a local minimum for e is that $D^2 e(0)$ (the Hessian of e at 0) be positive semidefinite. A sufficient condition for 0 to be a global minimum is that $D^2 e(0)$ be positive definite and that $D^2 e(x)$ be positive semidefinite for all $x \in X$. In the Hamiltonian formulation, $\vec{p}(0)$ corresponds to $\nabla e(x)$, and hence the derivative $D_x \vec{p}(0)$ of $\vec{p}(0)$, with respect to the initial condition $\vec{x}^*(0) = x$, corresponds to $D^2 e(x)$. In this way, we arrive at a necessary condition ($D_x \vec{p}(0) > 0$) and a sufficient condition

($D_x\vec{p}(0) > 0$ and $D_x^2\vec{p}(x) \geq 0$ for all x) in terms of the Hamiltonian formulation for solutions to (FC) to exist.

For examples with mild nonlinearities it is sometimes possible to solve the Hamiltonian system (5.2)–(5.5) for the critical point $(\vec{w}_x^*, \vec{u}_x^*)$; we shall illustrate this with an example in the next section. Once $(\vec{w}_x^*, \vec{u}_x^*)$ is found we can define e by

$$e(x) = \mathscr{E}(x, \vec{w}_x^*, \vec{u}_x^*).$$

The second major computational hurdle in the implementation of Recipe 3.1 is the computation of $u^*(x, w)$. In practice this should not be so difficult, since u^* arises as $\arg\min_u Q$ and minima are readily approximated numerically by the method of gradient descent. Given an initial guess $u^0(x, w)$, a next approximation to u^* is given by

$$u^{**}(x, w) = u^0(x, w) - t\nabla_u Q(x, w, u^0(x, w))$$

for a $t > 0$. The point is that we arrive at an excellent approximation to $u^*(x, w)$ in very few steps, if the initial guess is reasonably good. We expect such would be the case in real time control problems after a brief warm-up time.

6. An Example

We illustrate the theory of the preceding sections with a special class of plants $\mathscr{P}$ for which the recipe is more tractable. A similar type of example is discussed in greater detail in [BH3]; however, there the solution was not derived via the Hamiltonian system (5.1)–(5.5). We assume that the plant $\mathscr{P}$ in Fig. 1.2 has state space equations of the form

$$\text{(ML)} \qquad \begin{aligned} \tilde{x} &= F(x, w, u) = Ax + B_1 w + B_2 u, \\ z &= G_1(x, w, u) = M(C_1 x + D_{12} u), \\ y &= \begin{bmatrix} x \\ w \end{bmatrix}. \end{aligned}$$

Here $A\colon X \to X, B_1\colon W \to X, B_2\colon U \to X, C_1\colon X \to Z, D_{12}\colon U \to Z$ are linear transformations, and $M\colon Z \to Z$ is a (possibly) nonlinear mapping. The state space X, the input spaces U and W, and the output spaces Z and Y are all assumed to be finite-dimensional linear spaces. Thus, $\mathscr{P}$ is of the form of a *linear* IO map composed with a *memoryless* nonlinearity M on the output space Z. The form of the equation for y means that we will ideal with the (FC) problem and Recipe 3.1. Note that the assumption (W) below is satisfied if the memoryless term M is strictly convex.

Theorem 6.1. *Let A, B_1, B_2, C_1, D_{12}, and M be as in* (ML). *Set*

$$W(\sigma) = DM(\sigma)^T[M(\sigma)],$$

$$A^\times = A - B_2 D_{12}^{-1} C_1,$$

$P_{as}^\times$ = *the spectral projection for $A^\times$ associated with eigenvalues outside the unit disk,*

$$A_{as}^\times = A^\times \,|\, \mathrm{Im}\, P_{as}^\times, \quad B_{2as} = P_{as}^\times B_2,$$

and define $L\colon \operatorname{Im}(P_{as}^{\times})^T \to \operatorname{Im} P_{as}^{\times}$ *by*

$$L(p) = \sum_{n=0}^{\infty} (A_{as}^{\times})^{-(n+1)} B_{2as} D_{12}^{-1} \{W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T (A^{\times\,T})^{-(n+1)} p)$$

$$+ \frac{1}{2} \sum_{n=0}^{\infty} (A_{as}^{\times})^{-(n+1)} B_1 B_1^T (A_{as}^{\times\,T})^{-(n+1)} p.$$

Assume:

(U) $A^{\times}$ *has no eigenvalues on the unit circle;*

(W) W *is a diffeomorphism on* Z;

(L) L *is a diffeomorphism from* $\operatorname{Im}(P_{as}^{\times})^T$ *onto* $\operatorname{Im} P_{as}^{\times}$;

and

(A) $$\sigma(A) \subset \Delta = \{z \in \mathbb{C}\colon |z| < 1\}.$$

Then a solution $(\vec{p}(k), \vec{x}^*(k), \vec{w}_x^*(k), \vec{u}_x^*(k))$ *of the Hamiltonian system* (5.1)–(5.4) *associated with* (ML) *is given by*

$$\vec{p}(k) = (A_{as}^{\times\,T})^{-k} L^{-1}(P_{as}^{\times} x), \tag{6.1}$$

$$\vec{w}_x^* = \tfrac{1}{2} B_1^T (A_{as}^{\times\,T})^{-(k+1)} L^{-1}(P_{as}^{\times} x), \tag{6.2}$$

$$\vec{x}^*(k) = (A^{\times})^k x + \sum_{l=0}^{k-1} (A^{\times})^l [\tfrac{1}{2} B_1 B_1^T (A_{as}^{\times\,T})^{-(k-l)} L^{-1}(P_{as}^{\times} x) + B_2 D_{12}^{-1} \delta_{k+1-l}], \tag{6.3}$$

$$\vec{u}^*(k) = D_{12}^{-1}(\delta_k - C_1 \vec{x}^*(k)), \tag{6.4}$$

where

$$\delta_k = W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_{2_{as}}^T (A_{as}^{\times\,T})^{-(k+1)} L^{-1}(P_{as}^{\times} x)). \tag{6.5}$$

The associated energy function $e(x) = \mathscr{E}(x, \vec{w}_x^*, \vec{u}_x^*)$ *is given by*

$$e(x) = \sum_{k=0}^{\infty} [\|M(\delta_k)\|^2 - \tfrac{1}{4}\|B_1^T (A^{\times\,T})^{-k} L^{-1}(P_{as}^{\times} x)\|^2], \tag{6.6}$$

and the control $u = u^*(x, w)$ *given by* (5.6) *satisfies the equation*

$$B_2^T L^{-1}(P_{as}^{\times}(Ax + B_1 w + B_2 u)) + 2D_{12}^T W(C_1 x + D_{12} u) = 0.$$

Proof. Our task is to solve the Hamiltonian system of equations (5.1)–(5.5) specialized to a system of the form (ML). We first note that the Hamiltonian $H(p, x, w, u)$ given by (5.1) becomes

$$H(p, x, w, u) = p \cdot (Ax + B_1 w + B_2 u) + \|M \circ (C_1 x + D_{12} u)\|^2 - \|w\|^2. \tag{6.7}$$

Therefore, the derivative with respect to the state variable x in direction h is given by

$$\begin{aligned} D_x H(p, x, w, u)[h] &= p^T A h + 2\langle C_1^T DM(\sigma)[M(\sigma)], h\rangle \\ &= \langle A^T p, h\rangle + 2\langle C_1^T DM(\sigma)[M(\sigma)], h\rangle, \end{aligned}$$

where we have set

$$\sigma = C_1 x + D_{12} u,$$

so the gradient of H is given by

$$\nabla_x H(p, x, w, u) = A^T p + 2C_1^T DM(\sigma)^T[M(\sigma)]. \tag{6.8}$$

Thus, the recursion (5.4) for the costate vector p becomes

$$\vec{p}(k) = A^T \vec{p}(k+1) + 2C_1^T W(\vec{\sigma}^*(k)), \tag{6.9}$$

where

$$\begin{aligned}\vec{\sigma}^*(k) &= \sigma(\vec{x}^*(k), \vec{u}^*(k)) \\ &= C_1 \vec{x}^*(k) + D_{12}\vec{u}^*(k),\end{aligned} \tag{6.10}$$

and, in general, we have set

$$W(\sigma) = DM(\sigma)^T[M(\sigma)]. \tag{6.11}$$

Next, we wish to solve (5.2)

$$(\vec{w}^*(k), \vec{u}^*(k)) = \arg \operatorname*{crit}_{w,u} H(\vec{p}(k+1), \vec{x}^*(k), w, u)$$

for $\vec{w}^*(k)$ and $\vec{u}^*(k)$. To do this, we compute

$$\begin{aligned}D_w H(p, x, w, u)[h] &= p \cdot B_1 h - 2\langle w, h\rangle \\ &= \langle B_1^T p - 2w, h\rangle.\end{aligned}$$

This quantity being equal to zero in all directions h forces

$$\vec{w}^*(k) = \tfrac{1}{2} B_1^T \vec{p}(k+1). \tag{6.12}$$

Similarly, from

$$D_u H(p, x, w, u)[k] = \langle B_2^T p + 2D_{12}^T W(\sigma), k\rangle = 0,$$

in all directions k we get

$$2D_{12}^T W(\sigma) = -B_2^T p.$$

Using our assumption that D_{12} is invertible we can rewrite this as

$$W(\vec{\sigma}^*(k)) = -\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T \vec{p}(k+1). \tag{6.13}$$

If we now plug (6.13) into the recursion (6.9) for $\vec{p}(k)$ we obtain

$$\begin{aligned}\vec{p}(k) &= A^T \vec{p}(k+1) + 2C_1^T\{-\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T \vec{p}(k+1)\} \\ &= A^{\times T} \vec{p}(k+1),\end{aligned} \tag{6.14}$$

where we have set

$$A^\times = A - B_2 D_{12}^{-1} C_1. \tag{6.15}$$

A forward time recursion for $\vec{p}$ is therefore given by

$$\vec{p}(k+1) = (A^{\times T})^{-1} \vec{p}(k), \qquad \vec{p}(0) = p_0,$$

which has the solution

$$\vec{p}(k) = (A^{\times T})^{-k} p_0, \tag{6.16}$$

where $\vec{p}(0) = p_0$ is to be determined. Plugging (6.16) into the expressions (6.12) and (6.13) and using assumption (W) leads to

$$\vec{w}^*(k) = \tfrac{1}{2} B_1^T (A^{\times T})^{-k-1} p_0 \tag{6.17}$$

and

$$\vec{\sigma}^*(k) = W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T (A^{\times T})^{-k-1} p_0).$$

From $\sigma^*(k) = C_1 \vec{x}^*(k) + D_{12}\vec{u}^*(k)$ we can solve this last expression for $\vec{u}^*(k)$ to get

$$\vec{u}^*(k) = -D_{12}^{-1} C_1 \vec{x}^*(k) + D_{12}^{-1} W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T (A^{\times T})^{-k-1} p_0). \tag{6.18}$$

We next impose the condition that $\vec{u}^* \in l_U^{2+}$, $\vec{w}^* \in l_W^{2+}$; it turns out that these conditions uniquely determine p_0. From $\vec{w}^* \in l_W^{2+}$ we see that $p_0 \in \operatorname{Im}(P_{as}^x)^T$, where P_{as}^x in the Riesz spectral projection of $A^\times$ associated with eigenvalues of modulus larger than 1 (here we use assumption (U). To analyze the meaning of $\vec{u}^* \in l_U^{2+}$ we introduce the Fourier (or $Z-$) transform; in general, for a sequence $\vec{x} \in l_X^{2+}$ denote by $\hat{x}$ the H_X^2-function given by

$$\hat{x}(z) = \sum_{k=0}^{\infty} \vec{x}(k) z^k.$$

From the state space equations

$$\vec{x}^*(k+1) = A\vec{x}(k) + B_1 \vec{w}^*(k) + B_2 \vec{u}^*(k), \qquad \vec{x}^*(0) = x,$$

and

$$\vec{w}^*(k) = \tfrac{1}{2} B_1^T (A^{\times T})^{-k-1} p_0,$$

we get

$$\hat{x}(z) = (I - zA)^{-1} x + \frac{z}{2}(I - zA)^{-1} B_1 B_1^T (I - z(A^{\times T})^{-1})^{-1} (A^{\times T})^{-1} p_0$$
$$+ z(I - zA)^{-1} B_2 \hat{u}(z).$$

Plugging this into the Fourier transformed version of (6.18) gives

$$\hat{u}^*(z) = -D_{12}^{-1} C_1 \left\{ (I - zA)^{-1} x + \frac{z}{2}(I - zA)^{-1} B_1 B_1^T (I - z(A^{\times T})^{-1})^{-1} (A^{\times T})^{-1} p_0 \right.$$
$$\left. + z(I - zA)^{-1} B_2 \hat{u}^*(z) \right\}$$
$$+ D_{12}^{-1} \sum_{k=0}^{\infty} W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2 (A^{\times T})^{-k-1} p_0).$$

Solve for $\hat{u}^*(z)$ to get

$$F(z)\hat{u}^*(z) = -C_1(I - zA)^{-1} x + \frac{z}{2} C_1 (I - zA)^{-1} B_1 B_1^T (I - z(A^{\times T})^{-1})^{-1} (A^{\times T})^{-1} p_0$$
$$+ \sum_{k=0}^{\infty} W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2 (A^{\times T})^{-k-1} p_0), \tag{6.19}$$

where we have set

$$F(z) = D_{12} + zC_1(I - zA)^{-1}B_2. \tag{6.20}$$

Note that then $F(z)^{-1}$ is given by

$$F(z)^{-1} = D_{12}^{-1} - zD_{12}^{-1}C_1(I - zA^\times)^{-1}B_2D_{12}^{-1},$$

where $A^\times$ is an in (6.15). Note that assumption (A) and the condition $p_0 \in \operatorname{Im}(P_{as}^\times)^T$ imply that the right-hand side of (6.19) is in H_Z^2; hence, by a computation as in Theorem 1.2 and 1.3 from [BR], we can show from (6.19) that $\hat{u}^* \in H_Y^2$ is equivalent to

$$\sum_{n=0}^{\infty} (A_{as}^\times)^{-n-1}B_{2as}D_{12}^{-1}\cdot\Big\{W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1}B_2^T(A^{\times\,T})^{-(n+1)}p_0) - \sum_{k=1}^{n} \tfrac{1}{2}C_1A^{k-1}B_1B_1^T(A^{\times T})^{-(n+1-k)}p_0 - C_1A^n x\Big\} = 0, \tag{6.21}$$

where we have set $B_{as} = P_{as}^\times B_2$. Define $L\colon \operatorname{Im}(P_{as}^\times)^T \to \operatorname{Im} P_{as}^\times$ by

$$L(p) = \sum_{n=0}^{\infty} (A_{as}^\times)^{-(n+1)}B_{2as}D_{12}^{-1}\cdot\Big[W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1}B_2^T(A^{\times\,T})^{-(n+1)}p) - \sum_{k=1}^{n} \tfrac{1}{2}C_1A^{k-1}B_1B_1^T(A_{as}^{\times\,T})^{-(n+1-k)}p\Big]. \tag{6.22}$$

With a little algebra we can show that L, defined as in (6.22), is identical to L having the more symmetric form as given in the statement of the theorem; we postpone the verification of this until after the proof of Theorem 6.1. With L defined as in (6.22) we see that condition (6.21) can be expressed as

$$\begin{aligned} L(p_0) &= \sum_{p=0}^{\infty} (A_{as}^\times)^{-(n+1)}B_{2as}D_{12}^{-1}C_1A^n x \\ &\triangleq \Gamma x. \end{aligned} \tag{6.23}$$

Note that Γ is a solution of the Stein equation

$$\Gamma - (A_{as}^\times)^{-1}\Gamma A = (A_{as}^\times)^{-1}B_{2as}D_{12}^{-1}C_1, \tag{6.24}$$

and since both $\sigma(A)$ and $\sigma(A_{as}^\times)$ are contained in Δ, Γ is uniquely determined as a solution of (6.24). On the other hand, it is easy to verify that $P_{as}^\times\colon X \to \operatorname{Im} P_{as}^\times$ solves (6.24); we conclude that $\Gamma = P_{as}^\times$ and hence (6.23) takes the form

$$L(p_0) = P_{as}^\times x.$$

Using assumption (L) we finally solve for p_0:

$$p_0 = L^{-1}(P_{as}^\times x). \tag{6.25}$$

Verification of the statements in Theorem 6.1 is now a routine matter.

The proof of theorem 6.1 is thus complete once we verify

Lemma 6.2. *The mapping* $L\colon \operatorname{Im}(P_{as}^\times)^T \to \operatorname{Im} P_{as}^\times$ *given by* (6.22) *is identical to the mapping* L *given in the statement of Theorem* 6.1.

Proof. Note that L, as given in (6.22), can be expressed as $L = L_1 + L_2$ where

$$L_1(p) = \sum_{n=0}^{\infty} (A_{as}^{\times})^{-(n+1)} B_{2as} D_{12}^{-1} W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1} B_2^T (A^{\times T})^{-(n+1)} p),$$

and

$$L_2(p) = -\frac{1}{2} \sum_{n=1}^{\infty} \sum_{k=1}^{n} (A_{as}^{\times})^{-(n+1)} B_{2as} D_{12}^{-1} C_1 A^{k-1} B_1 B_1^T (A_{as}^{\times T})^{-(n+1-k)} p. \tag{6.26}$$

The content of the lemma is the identity

$$L_2(p) = \frac{1}{2} \sum_{n=0}^{\infty} (A_{as}^{\times})^{-(n+1)} B_1 B_1^T (A_{as}^{\times T})^{-(n+1)} p.$$

To simplify L_2, the first step is to use the identity

$$-B_2 D_{12}^{-1} C_1 = A^{\times} - A.$$

Substitution of this into (6.26) gives

$$\begin{aligned} L_2 &= \frac{1}{2} \sum_{n=1}^{\infty} \sum_{k=1}^{n} (A_{as}^{\times})^{-(n+1)} [A^{\times} - A] A^{k-1} B_1 B_1^T (A_{as}^{\times T})^{-(n+1-k)} \\ &= \frac{1}{2} \sum_{n=1}^{\infty} \sum_{k=1}^{n} [Y_{k,n} - Z_{k,n}], \end{aligned} \tag{6.27}$$

where

$$Y_{k,n} = (A_{as}^{\times})^{-n} A^{k-1} B_1 B_1^T (A_{as}^{\times T})^{-(n+1-k)}$$

and

$$Z_{k,n} = (A_{as}^{\times})^{-(n+1)} A^{k} B_1 B_1^T (A_{as}^{\times T})^{-(n+1-k)}. \tag{6.28}$$

Note that $Y_{k+1,n+1} = Z_{k,n}$ and hence the series in (6.27) telescopes. Explicitly, we have

$$\begin{aligned} 2L_2 &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} Y_{k+1,n+1} - \sum_{n=1}^{\infty} \sum_{k=1}^{n} Z_{k,n} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} Z_{k,n} - \sum_{n=1}^{\infty} \sum_{k=1}^{n} Z_{k,n} = \sum_{n=1}^{\infty} Z_{0,n} + Z_{0,0} = \sum_{n=0}^{\infty} Z_{0,n}. \end{aligned}$$

Recalling (6.28), we get

$$L_2 = \frac{1}{2} \sum_{n=0}^{\infty} (A_{as}^{\times})^{-(n+1)} B_1 B_1^T (A_{as}^{\times T})^{-(n+1)}$$

as required. ∎

Recall that in order for the formulas (6.1)–(6.5) to generate a solution of (FC) (including the stability constraint), we require also that the function e given by (6.6) has nonnegative values. It is interesting to analyze this explicitly for the linear case

$(M(\sigma) = \sigma,\ W(\sigma) = \sigma)$. Then $e(x)$ can be written out explicitly as

$$\begin{aligned} e(x) &= \frac{1}{4}\sum_{k=0}^{\infty} \langle L^{-1}(A_{as}^{\times})^{-(k+1)}B_{2as}D_{12}^{-1}D_{12}^{T-1}B_{2as}^{T}(A_{as}^{\times T})^{-(k+1)}L^{-1}P_{as}^{\times}x,\ P_{as}^{\times}x\rangle \\ &\quad - \frac{1}{4}\sum_{k=0}^{\infty} \langle L^{-1}(A^{\times})^{-k}B_1B_1^T(A^{\times T})^{-k}L^{-1}P_{as}^{\times}x,\ P_{as}^{\times}x\rangle \\ &= \frac{1}{4}\left\langle L^{-1}\left\{\sum_{k=0}^{\infty}(A_{as}^{\times})^{-(k+1)}B_{2as}D_{12}^{-1}(D_{12}^T)^{-1}B_{2as}^T(A_{as}^{\times T})^{-(k+1)}L^{-1}\right.\right. \\ &\qquad \left.\left. - \sum_{k=0}^{\infty}(A_{as}^{\times})^{-k}B_1B_1^T(A_{as}^{\times T})^{-k}\right\}L^{-1}P_{as}^{\times}x,\ P_{as}^{\times}x\right\rangle \\ &= -\tfrac{1}{2}\langle LP_{as}^{\times}x,\ P_{as}^{\times}x\rangle. \end{aligned}$$

Thus the condition that $e(x) \geq 0$ for all x is equivalent to $-L$ being positive definite. From the formula for L in Theorem 6.1 we see that, in general, L satisfies the nonlinear Stein equation

$$\begin{aligned} [L - (A_{as}^{\times})^{-1}\cdot L\cdot(A_{as}^{\times T})^{-1}](p) &= (A_{as}^{\times})^{-1}B_{2as}D_{12}^{-1}W^{-1}(-\tfrac{1}{2}(D_{12}^T)^{-1}B_2^T(A^{\times T})^{-1}p) \\ &\quad + \tfrac{1}{2}(A_{as}^{\times})^{-1}B_1B_1^T(A_{as}^{\times T})^{-1}p. \end{aligned} \tag{6.29}$$

For the linear case where W = identity map, we see that $(-2L)$ satisfies

$$\begin{aligned} &(-2L) - (A_{as}^{\times})^{-1}(-2L)(A_{as}^{\times T})^{-1} \\ &\quad = (A_{as}^{\times})^{-1}B_1B_1^T(A_{as}^{\times T})^{-1} - (A_{as}^{\times})^{-1}B_{2as}D_{12}^{-1}(D_{12}^T)^{-1}B_2^T(A^{\times T})^{-1}. \end{aligned} \tag{6.30}$$

This Stein equation is the discrete-time analogue of one of the Riccati equations occurring in [DGKF] (actually $(-2L)^{-1}$ corresponds to the solution of one of the Riccati equations in [DGKF]); a second Stein equation is not relevant for the case here since we are assuming that $\sigma(A) \subset \Delta$. In any case, our nonlinear theory specialized to the linear case recovers the result that solutions of the H^∞ control problem for the linear plant given by (ML) (with M = identity) exist if and only if the solution $(-2L)$ of the Stein equation (6.30) is positive definite.

Appendix. Proof of Lemma 2.2

In this appendix we present the proof of Lemma 2.2.

Let the function f and the points $(r^*, s^*) = ((r_+^*, r_-^*), (s_+^*, s_-^*))$ and $\varphi(r) = (\varphi_+(r), \varphi_-(r))$ be as in the statement of the lemma.

We prove that (r_+^*, r_-^*) is a max–min point for Ψ by verifying the conditions of Lemma 2.1. We first verify that $r^* = (r_+^*, r_-^*)$ is a critical point for Ψ. By assumption (r^*, s^*) is a max–min point for f and for each fixed r, $\varphi(r)$ is a max–min point for ψ_r. By necessity in Lemma 2.1, we see that (r^*, s^*) is a critical point for f, so

$$Df(r^*, s^*)[(h, k)] = 0 \qquad \text{for all} \quad (h, k) \in T_{r^*}\mathscr{R} \times T_{s^*}\mathscr{S}, \tag{A.1}$$

and, in particular, $\varphi(r)$ is a critical point for ψ_r, so

$$Df(r, \varphi(r))[(0, k)] = 0 \qquad \text{for all} \quad k \in T_{\varphi(r)}\mathscr{S}. \tag{A.2}$$

In particular, when $r = r^*$ from A.1 we conclude that

$$Df(r_*, s_*)[(0, k)] = 0 \qquad \text{for all} \quad k \in T_{s_*}\mathscr{S}.$$

By the assumed uniqueness of critical points, we conclude that $\varphi(r^*) = s^*$. From the chain rule we have

$$D\Psi(r)[h] = Df(r, \varphi(r))[(h, D\varphi(r)[h])]. \tag{A.3}$$

From (A.2) we see that this vanishes for all h if and only if $Df(r, \varphi(r)) = 0$, i.e., if and only if $r = r^*$. Thus, Ψ has a unique critical point at $r = r^*$.

To check that r^* is a max–min point for Ψ we must check the Hessian for Ψ at r^*. Differentiating (A.3) gives

$$\begin{aligned} D^2\Psi(r)[k, h] &= D_r\{Df(r, \varphi(r))[(h, D\varphi(r)[h])]\}[k] \\ &= D^2 f(r, \varphi(r))[(k, D\varphi(r)[k]), (h, D\varphi(r)[h])] \\ &\quad + Df(r, \varphi(r))[h, D^2\varphi(r)[k, h]]. \end{aligned}$$

At $r = r^*$, this collapses to

$$D^2\Psi(r^*)[k, h] = D^2 f(r^*, \varphi(r^*))[(k, D\varphi(r^*)[k]), (h, D\varphi(r^*)[h])].$$

Write the full Hessian of f at $(r^*, \varphi(r^*))$ in the form

$$D^2 f(r^*, \varphi(r^*))[(h, k), (h, k)] = [h^T, k^T]\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix},$$

where

$$\begin{aligned} D^2 f(r^*, \varphi(r^*))[(h, 0), (h, 0)] &= h^T A h, \\ D^2 f(r^*, \varphi(r^*))[(0, k), (h, 0)] &= h^T B k, \\ D^2 f(r^*, \varphi(r^*))[(0, k), (0, k)] &= k^T C k. \end{aligned}$$

The assumptions on the Hessians imply that $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ and C are invertible.

On the other hand, if we differentiate the identity

$$Df(r, \varphi(r))[(0, h)] = 0,$$

with respect to r we get

$$D^2 f(r, \varphi(r))[(k, D\varphi(r)[k]), (0, h)] = 0,$$

i.e.,

$$0 = D^2 f(r, \varphi(r))[(k, 0), (0, h)] + D^2 f(r, \varphi(r))[(0, D\varphi(r)[k]), (0, h)].$$

Evaluation at $r = r^*$ gives

$$h^T B^T k + h^T C D\varphi(r^*)[k] = 0$$

from which we conclude

$$CD\varphi(r^*) = -B^T.$$

As we already observed that C is invertible, this gives

$$D\varphi(r^*) = -C^{-1}B^T.$$

Thus $D^2\Psi(r^*)[k, h]$ takes the form

$$\begin{aligned} D^2\Psi(r^*)[k, h] &= h^T[I, D\varphi(r^*)^T]\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}\begin{bmatrix} I \\ D\varphi(r^*) \end{bmatrix}k \\ &= h^T\{A - 2BC^{-1}B^T + (BC^{-1})C(C^{-1}B^T)\}k \qquad (\text{A.4}) \\ &= h^T\{A - BC^{-1}B^T\}k. \end{aligned}$$

By a Schur complement argument (see, e.g., page 656 of [K]) the invertibility of $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ and of C implies that of $A - BC^{-1}B^T$. Thus $D^2\Psi(r^*)$ is invertible.

We now write $h \in T_{r^*}\mathscr{R}$ in the finer form of $h = (h_+, h_-) \in T_{r_+^*}\mathscr{R}_+ \times T_{r_-^*}\mathscr{R}_-$ and similarly $k = (k_+, k_-) \in T_{s_+^*}\mathscr{S}_+ \times T_{s_-^*}\mathscr{S}_-$. This induces natural finer decompositions of the matrices A, B, C as

$$A = \begin{bmatrix} A_{++} & A_{+-} \\ A_{+-}^T & A_{--} \end{bmatrix},$$

$$B = \begin{bmatrix} B_{++} & B_{+-} \\ B_{-+} & B_{--} \end{bmatrix},$$

$$C = \begin{bmatrix} C_{++} & C_{+-} \\ C_{+-}^T & C_{--} \end{bmatrix}.$$

The assumption that (r^*, s^*) is a max–min point for f implies by Lemma 2.1 that

$$\begin{bmatrix} A_{--} & B_{--} \\ B_{--}^T & C_{--} \end{bmatrix} > 0, \qquad (\text{A.5})$$

and that

$$\begin{bmatrix} A_{++} & B_{++} \\ B_{++}^T & C_{++} \end{bmatrix} - \begin{bmatrix} A_{+-} & B_{+-} \\ B_{+-}^T & C_{-+}^T \end{bmatrix}\begin{bmatrix} A_{--} & B_{--} \\ B_{--}^T & C_{--} \end{bmatrix}^{-1}\begin{bmatrix} A_{-+} & B_{-+} \\ B_{-+}^T & C_{-+} \end{bmatrix} < 0. \qquad (\text{A.6})$$

The assumption that s^* is a max–min point for Ψ_{r^*} means

$$C_{--} > 0 \qquad (\text{A.7})$$

and

$$C_{++} - C_{+-}C_{--}^{-1}C_{+-}^T < 0. \qquad (\text{A.8})$$

From (A.5) and (A.7) we get

$$A_{--} - B_{--}C_{--}^{-1}B_{--}^T > 0, \qquad (\text{A.9})$$

by a Schur complement argument. To show that r^* is a max–min point for Ψ, by Lemma 2.1 and formula (A.4) for $D^2\Psi(r^*)$, we must check that

$$[A - BC^{-1}B^T]_{--} > 0.$$

Since C_{--} is invertible, the blocks of C^{-1} can be computed explicitly; we find

$$C^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}C_{+-}C_{--}^{-1} \\ -C_{--}^{-1}C_{+-}^T\Delta^{-1} & C_{--}^{-1} + C_{--}^{-1}C_{+-}^T\Delta^{-1}C_{+-}C_{--}^{-1} \end{bmatrix},$$

where Δ is the Schur complement

$$\Delta \triangleq C_{++} - C_{+-}C_{--}^{-1}C_{+-}^T.$$

We can then compute

$$\begin{aligned}[A - BC^{-1}B^T]_{--} &= A_{--} - [B_{-+}B_{--}]C^{-1}\begin{bmatrix} B_{-+}^T \\ B_{--}^T \end{bmatrix} \\ &= [A_{--} - B_{--}C_{--}^{-1}B_{--}^T] \\ &\quad - (B_{-+} - B_{--}C_{--}^{-1}C_{+-}^T)\Delta^{-1}(B_{-+} - B_{--}C_{--}^{-1}C_{+-}^T)^T > 0\end{aligned}$$

by (A.9) and (A.8). The second Hessian check in Lemma 2.1, namely, that

$$D_{++} - D_{+-}D_{--}^{-1}D_{-+} < 0,$$

where $D = A - BC^{-1}B^T$, follows since $A - BC^{-1}B^T$, as the Schur complement of C in $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$, has the appropriate number of positive and negative eigenvalues. ∎

In the proof of Proposition 3.2 we apply Lemma 2.2 to a case where the manifold $\mathscr{S}$ is infinite dimensional. The proof above goes through in this level of generality except for the last step, where the negative definiteness of

$$D_{++} - D_{+-}D_{--}^{-1}D_{-+} \text{ (where } D = A - BC^{-1}B^T)$$

was argued based on counting the number of positive and negative eigenvalues of $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$. Nevertheless we can bootstrap to the infinite-dimensional case by restricting $\mathscr{S}$ to finite-dimensional submanifolds and applying the above results. The details of this we leave to the reader.

References

[BH1] J. A. Ball and J. W. Helton, Shift invariant manifolds and nonlinear analytic function theory, *Integral Equations Operator Theory*, **11** (1988), 615–725.

[BH2] J. A. Ball and J. W. Helton, Factorization of nonlinear systems: Toward a theory for nonlinear H^∞ control, *Proceedings of the 27th IEEE Conference on Decision and Control*, Austin, TX, 1988, pp. 2376–2381.

[BH3] J. A. Ball and J. W. Helton, Inner–outer factorization of nonlinear operators, preprint.

[BH4] J. A. Ball and J. W. Helton, H_∞ control for nonlinear plants: Connections with differential games, *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, FL, 1989, pp. 956–962.

[BHW] J. A. Ball, J. W. Helton, and M. Walker, A variational approach to nonlinear H^∞ control, preprint.

[BR] J. A. Ball and A. C. M. Ran, Global inverse spectral problems for rational matrix functions, *Linear Algebra Appl.*, **86** (1987), 237–282.

[B] T. Basar, A dynamic games approach to controller design: Disturbance rejection in discrete time, *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, FL, 1989, pp. 407–414.

[BO] T. Basar and G. J. Olsder, *Dynamic Non-Cooperative Game Theory*, Academic Press, New York, 1977.

[DGKF] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, State-space solutions to standard H_2 and H_∞ control problems, *IEEE Trans. Automat. Control*, **34** (1989), 831–847.

[F] B. A. Francis, *A Course in H_∞ Control Theory*, Lecture Notes in Control and Information Science, Vol. 88, Springer-Verlag, Berlin, 1987.

[HM] D. J. Hill and P. J. Moylan, Connections between finite gain and asymptotic stability, *IEEE Trans. Automat. Control*, **25** (1980), 931–936.

[IA] A. Isidori and A. Astolfi, Nonlinear H_∞-Control via Measurement Feedback, *J. Math. Systems Estimation Control*, **2** (1992), 31–44.

[K] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[PS] G. P. Papavassilopoulos and M. G. Safanov, Robust control via game-theoretic methods, *Proceedings of the 28th IEEE Conference on Decision and Control*, Tampa, FL, 1989, pp. 382–387.

[S1] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Springer-Verlag, New York, 1990.

[S2] E. D. Sontag, Smooth stabilization implies coprime factorization, *IEEE Trans. Automat. Control*, **34** (1989), 435–443.

[W] J. C. Willems, Dissipative dynamical systems, part I: General theory, *Arch. Rational Mech. Anal.*, **45** (1972), 321–351.