Volume 3/3

Part III

Riccati equations and Optimal control

Chapter 8

Optimal Control ($\frac{\mathrm{d}}{\mathrm{du}}\mathcal{J} = 0$)

And if dearly that error hath cost me,
And more that I once could foresee,
I have found that, whatever it lost me,
It could not deprive me of thee.

— Lord Byron (1788–1824), "Stanzas to Augusta"

In this chapter we present an abstract theory on optimization and optimal control in state feedback form and the application of this theory to WPLSs with guidelines to more general or time-varying systems.

We shall study the critical points of a given cost function and the case where such control corresponds to a stabilizing state feedback pair. Such an "optimal" state feedback pair corresponds to a "stabilizing" solution of the Riccati equation, as shown in Chapter 9. The corresponding special control problems are solved in Chapters 10–12.

In Sections 8.1 and 8.2 we work in an abstract setting. The cost function $\mathcal{J}(x,u)$ is a quadratic function of vectors x and u, and we wish to find "controls" u that are critical points of $\mathcal{J}(x,\cdot)$ for a fixed "initial state" x; equivalently, for which the Fréchet derivative of $\mathcal{J}(x,\cdot)$ is zero. Such controls correspond to solutions of optimization problems (e.g., LQR, H^∞ or any other quadratic maximization, minimization or minimax problems). In the sequel, we prefer the word "critical" (or "J-critical") to "optimal", since in general critical points need not be optimums although the converse is always true.

We show that there is a unique J-critical control for each x iff there is a unique J-critical control for $x = 0$ and the abstract system is "stabilizable". Moreover, if this is the case, then the J-critical control can be written in "state feedback form".

In Section 8.2, we define and study "J-coercivity", which is a generalization of the standard nonsingularity assumptions of several control problems (including the "J-coercivity" assumptions defined in [S97b]–[S98d], the "Popov Toeplitz invertibility" assumption in the stable case and the "no transmission zeros" and "no invariant zeros" assumptions in the positive case). We show that any "stabilizable" J-coercive abstract system has a unique critical control for each initial state, so that the results of Section 8.1 can be applied. We also present some related results.

There are three reasons for the use of this abstract setting. First, this can be considered as a short-hand notation for WPLSs as to make the optimization theory simple, clear and neat. Second, this theory can be applied more generally, as indicated in Section 8.5 on time-varying systems and in Section 8.6 on systems whose input operator ("B") is allowed to be more unbounded than those of WPLSs. However, we only give guidelines for these settings since they go beyond the scope of this book.

The third and most important reason is that when working with the H^∞ problem in Section 11.7, we have to optimize the control for a fixed state *and* a fixed disturbance; the WPLS framework does not cover such optimization. Therefore, we solve the H^∞ full-information control problem in the abstract framework of Section 8.1, although the results will be applied to WPLSs only.

Naturally, in Chapters 9–12 we have to work hard on the "raw WPLS solutions" obtained as corollaries of the abstract theory, before we can turn them into direct generalizations of classical control theory results.

In Sections 8.3 and 8.4 we apply our abstract optimization theory to obtain a very general theory on control problems for WPLSs. Given a WPLS $\left[\begin{smallmatrix}\mathbb{A} & \mathbb{B}\\ \mathbb{C} & \mathbb{D}\end{smallmatrix}\right]$ and a cost operator $J = J^* \in \mathcal{B}(Y)$, we study the *cost function*

$$\mathcal{J}(x_0,u) := \int_0^\infty \langle y(t), Jy(t)\rangle_Y \, dt, \quad \text{where} \quad y := \mathbb{C}x_0 + \mathbb{D}u \quad (x_0 \in H,\ u : \mathbf{R}_+ \to U) \tag{8.1}$$

(for suitable $\mathbb{C}$ and $\mathbb{D}$ this covers all classical control problems mentioned above) and u is required to be exponentially stabilizing, strongly stabilizing, stabilizing or something similar, depending on how stable one wishes the closed-loop system to be.

Under J-coercivivity and stabilizability assumptions, there is a unique J-critical ("optimal") control for any initial state, and this optimal control can be given in a WPLS form (this generalizes the corresponding result in [FLT]). However, the corresponding feedback need not be well-posed without additional assumptions on the system, as illustrated in Examples 8.4.13 and 11.3.7. This leads to some additional difficulties in the Riccati equation theory (the situation is the same even in the case studied in [FLT]).

In Theorem 8.4.5, we extend the standard result that an optimization problem over exponentially stabilizing controllers can be solved by first finding a preliminary exponentially stabilizing controller and then optimizing over stable controls for the preliminary controlled system. We also give corresponding results over other forms of stabilization, but for them one needs an additional quasi-coprimeness ("q.r.c.-") assumption. Then we give further results on J-coercivity and recall its connection to spectral and coprime factorizations for maps in MTIC classes.

Sections 8.1 and 8.2 are written for a the abstract setting of Hypothesis 8.1.1 (see also Hypothesis 8.2.2). Hypothesis 8.3.1 is assumed throughout this Sections 8.3–8.5, and Hypothesis 8.6.1 is assumed through Sections 8.6.

8.1 Abstract J-critical control ($Jy_{ycrit} \perp \Delta y$)

The more control, the more that requires control.

In this section we define the set $\mathcal{U}$ of admissible controls and the cost function $\mathcal{J}$ and study their basic properties.

Standing Hypothesis 8.1.1 *Throughout this section and Section 8.2 we shall assume that U, X, Y^{s}, Z^{s} and are Banach spaces, that Y and Z are TVSs, and that the embeddings $Y^{\mathrm{s}} \subset Y$ and $Z^{\mathrm{s}} \subset Z$, are continuous. We also assume that $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \in \mathcal{B}(X \times U, Z \times Y)$ and $J = J^* \in \mathcal{B}(Y^{\mathrm{s}}, Y^{\mathrm{s}*})$.*

All results given in these two sections are valid whether we use linear or conjugate-linear duals, i.e., $X^* := X^{\mathrm{B}}$ or $X^* := X^{\mathrm{d}}$; see Remark A.3.22 for details. In particular, we may use Hilbert space adjoints instead of Banach space adjoints.

We use sesquilinear duals and adjoints (see Remark A.3.22) and the notation of Lemma A.3.5 (e.g., "$J = J^*$" means that $J^*|_{Y^{\mathrm{s}}} = J$), hence the results look as if the spaces in Standing Hypothesis 8.1.1 were Hilbert spaces.

Outside Sections 8.1 and 8.2, we shall apply these results only in the case where Y^{s} is a Hilbert space. Therefore, we recommend the reader to consider only this Hilbert space setting so that $Y^{\mathrm{s}*} = Y^{\mathrm{s}}$ and hence there is nothing special with inner products or self-adjointness.

Remark 8.1.2 *In this section one may allow $U, X, Y^{\mathrm{s}}, Z^{\mathrm{s}}$ to be arbitrary* F-spaces *(i.e., complete metrizable TVSs, see Theorem 1.24 of [Rud73]), because the Closed Graph Theorem (Theorem 2.15 of [Rud73]) is the only nongeneral TVS property that we use here.*

Given an *input* $u \in U$ and *initial state* $x \in X$, we call $z := Ax + Bu \in Z$ the *state*, $y := Cx + Du \in Y$ the *output* and $\langle y, Jy \rangle$ the *cost* of the "system" for u and x s.t. the state and output are *stable*, i.e., $z \in Z^{\mathrm{s}}$ and $y \in Y^{\mathrm{s}}$ As before, we set $\|y\|_{Y^{\mathrm{s}}} = +\infty$ for $y \notin Y^{\mathrm{s}}$, etc.

We could have dropped A, B, Z^{s} and Z from the theory without reducing generality (replace C by $\left[\begin{smallmatrix} A \\ C \end{smallmatrix}\right]$ and D by $\left[\begin{smallmatrix} B \\ D \end{smallmatrix}\right]$ etc.), but we have chosen this more explicit presentation to make later applications more obvious.

The simplest application of this theory to WPLSs is obtained by the substitutions $U \mapsto \mathrm{L}^2(\mathbf{R}_+; U)$, $Y^{\mathrm{s}} \mapsto \mathrm{L}^2(\mathbf{R}_+; Y)$, $Y \mapsto \mathrm{L}^2_\omega(\mathbf{R}_+; Y)$, $X \mapsto H$, $Z^{\mathrm{s}}, Z \mapsto \mathrm{L}^2_\omega(\mathbf{R}_+; H)$, $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathbb{A} & \mathbb{B}\tau \\ \mathbb{C} & \mathbb{D} \end{smallmatrix}\right]$ and $x \mapsto x_0 \in H$, where U, H, Y are Hilbert spaces, $\left[\begin{smallmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{smallmatrix}\right] \in \mathrm{WPLS}_\omega(U, H, Y)$ and $\omega \in \mathbf{R}$.

It follows that $\mathcal{U}(x) := \{u \in U \,|\, Cx + Du \in Y^{\mathrm{s}},\ Ax + Bu \in Z^{\mathrm{s}}\}$ becomes the space of stable (unless we take $U = \mathrm{L}^2_\omega$) controls that make the output $Cx + Du$ stable. Sometimes we also require the state $\mathbb{A}x_0 + \mathbb{B}\tau u$ to be stable, i.e., we set $Z^{\mathrm{s}} := \mathrm{L}^2(\mathbf{R}_+; H)$. See Remark 8.3.4 for further applications to WPLSs.

Now for the general definitions:

Definition 8.1.3 (J-critical control) *For each $x \in X$, we set*

$$\mathcal{U}(x) := \{u \in U \mid Cx + Du \in Y^s \ \& \ Ax + Bu \in Z^s\} \tag{8.2}$$
$$\mathcal{Y}(x) := \{Cx + Du \mid u \in \mathcal{U}(x)\}, \tag{8.3}$$
$$\mathcal{J}(x,u) := \langle Cx + Du, J(Cx + Du)\rangle \quad (u \in \mathcal{U}(x)). \tag{8.4}$$

We call $\mathcal{J}$ the cost function. *A control $u \in \mathcal{U}(x)$ (resp. output $Cx + Du \in \mathcal{Y}(x)$) is called J-*critical *for x (w.r.t. $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$) if $\langle Cx + Du, JD\eta\rangle = 0$ for all $\eta \in \mathcal{U}(0)$.*

Note that $\mathcal{U}(x) \subset U$ and $\mathcal{Y}(x) \subset Y^s \subset Y$ for all $x \in X$.

Given an initial state $x \in X$, we often wish to find a *stabilizing control* (i.e., $u \in \mathcal{U}$) s.t. the output $y := Cx + Du \in \mathcal{Y}$ is optimizes the cost $\mathcal{J}(x,u) = \langle y, Jy\rangle$ in some sense. Because optimality (in most reasonable settings) requires that the Fréchet derivative of the cost function is zero, or equivalently, that the control is J-critical (by Lemma 8.1.6), we shall concentrate on J-critical controls in this section. See Sections 8.2 and 11.7 for applications.

In the following few lemmas we list simple algebraic properties of the above concepts for future use:

Lemma 8.1.4 ($\mathcal{U}(x)$ and $\mathcal{Y}(x)$) *The sets $\mathcal{U}(0) \subset U$ and $\mathcal{Y}(0) \subset Y^s$ are linear subspaces. Let $x \in X$, $u \in \mathcal{U}(x)$, $y \in \mathcal{Y}(x)$. Then $\mathcal{U}(x) = u + \mathcal{U}(0)$ and $\mathcal{Y}(x) = y + \mathcal{Y}(0)$.*

Moreover, $\mathcal{U}(\alpha x_0 + \beta x_1) = \alpha\mathcal{U}(x_0) + \beta\mathcal{U}(x_1)$ and $\mathcal{Y}(\alpha x_0 + \beta x_1) = \alpha\mathcal{Y}(x_0) + \beta\mathcal{Y}(x_1)$ ($\alpha, \beta \in \mathbf{C} \setminus \{0\}$, $x_0, x_1 \in X$ s.t. $\mathcal{U}(x_0) \neq \emptyset$). □

(The very easy proof is left to the reader. The last two formulae need not hold for x_0, x_1 s.t. $\mathcal{U}(x_0) = \emptyset = \mathcal{U}(x_1)$.)

The case where $Z^s = Z$ and $Y^s = Y$ is called the stable case, because if(f) $\mathcal{U}(x) = U$ for all $x \in X$ (i.e., if(f) all controls are stabilizing), then we can replace Z^s by Z and Y^s by Y, w.l.o.g., by the following lemma:

Lemma 8.1.5 (Stable case) *If $Y^s = Y$ and $Z^s = Z$, then $\mathcal{U}(x) = U$ for all $x \in X$. We have $\mathcal{U}(x) = U$ for all $x \in X$ iff $C \in \mathcal{B}(X, Y^s)$, $D \in \mathcal{B}(U, Y^s)$, $A \in \mathcal{B}(X, Z^s)$, $B \in \mathcal{B}(U, Z^s)$.* □

(This follows from Standing Hypothesis 8.1.1 and Lemma A.3.6.)

With the standard substitutions mentioned above, the stable case is the case where each x and u produce a stable output, i.e., where $\mathbb{C}$ and $\mathbb{D}$ are stable (that is, $\Sigma \in$ SOS; the alternative substitutions with $Z^s \mapsto \mathrm{L}^2(\mathbf{R}_+; H)$ correspond to exponentially stable Σ).

The J-critical controls are exactly the zeros of the derivative of the cost:

Lemma 8.1.6 (J-critical $\Leftrightarrow \frac{d\mathcal{J}}{du} = 0$) *Let $x \in X$.*

A control $u_{\mathrm{crit}} \in \mathcal{U}(x)$ is J-critical for x iff $\frac{d\mathcal{J}(x,u)}{du}(u_{\mathrm{crit}}) = 0$. In particular, if $u_{\mathrm{crit}}(x)$ is a local extremal point of $\mathcal{J}(x,\cdot)$ (on $\mathcal{U}(x)$), then $u_{\mathrm{crit}}(x)$ is J-critical for x.

Here $\frac{d\mathcal{J}(x,u)}{du}$ is the *(real) Fréchet derivative* of $\mathcal{J}(x,\cdot)$ on its domain $\mathcal{U}(x)$. Thus, $\frac{d\mathcal{J}(x,u)}{du}(\widetilde{u}) : \mathcal{U}(0) \to \mathbf{R}$, where $\widetilde{u} \in \mathcal{U}(x)$, means the operator

$$\mathcal{U}(0) \ni \eta \mapsto \frac{d\mathcal{J}(x,\widetilde{u}+t\eta)}{dt}(0) = \lim_{t\to 0} \frac{\mathcal{J}(x,\widetilde{u}+t\eta) - \mathcal{J}(x,\widetilde{u})}{t} \in \mathbf{R}. \tag{8.5}$$

Proof: 1° *"Iff":* By Lemma 8.1.4, $\mathcal{U}(x) = u_{\text{crit}} + \mathcal{U}(0)$. By linearity and continuity, $\frac{d\mathcal{J}(x,u_{\text{crit}}+t\eta)}{dt}(0) = 2\,\text{Re}\langle Cx+Du, JD\eta\rangle$. This is zero for all $\eta \in \mathcal{U}(0)$ iff $\langle Cx+Du, JD\eta\rangle = 0$ for all $\eta \in \mathcal{U}(0)$, (apply the right-hand-side to η and $i\eta$), i.e., iff u_{crit} is J-critical for x.

It is obvious that $\mathcal{J}(x,\widetilde{u}+t\eta)$ has an extremum at $t=0$ if $\widetilde{u}$ is an extremum (or a saddle point) of $\mathcal{J}(x,\cdot)$. □

Given a critical control u_{crit}, the cost for $u_{\text{crit}}+\eta$ equals the critical cost plus the cost for η:

Lemma 8.1.7 (Critical cost $\mathcal{J}(x,u_{\text{crit}})$) *Let $x \in X$ and $u_{\text{crit}} \in \mathcal{U}(x)$. Set $y_{\text{crit}} := Cx + Du_{\text{crit}}$. Then the following are equivalent:*

(i) u_{crit} is J-critical for x;

(ii) $\mathcal{J}(x,u_{\text{crit}}+\eta) = \langle y_{\text{crit}}, Jy_{\text{crit}}\rangle + \langle D\eta, JD\eta\rangle \quad (\eta \in \mathcal{U}(0))$;

(iii) $\langle Cx+D(u_{\text{crit}}+\eta_1), J(Cx+D(u_{\text{crit}}+\eta_2))\rangle = \langle y_{\text{crit}}, Jy_{\text{crit}}\rangle + \langle D\eta_1, JD\eta_2\rangle$ $(\eta_1,\eta_2 \in \mathcal{U}(0))$.

Note that (ii) means that $\mathcal{J}(x,u_{\text{crit}}+\eta) = \mathcal{J}(x,u_{\text{crit}}) + \mathcal{J}(0,\eta)$.

Proof: By a direct computation, (i)⇒(iii)⇒(ii) ((i) implies that the cross terms of (iii) are zero).

Assume (ii). Then $2\,\text{Re}\langle y_{\text{crit}}, JD\eta\rangle = 0$ for all $\eta \in \mathcal{U}(0)$. An application to η and $i\eta$ shows that $\langle y_{\text{crit}}, JD\eta\rangle = 0$ for all $\eta \in \mathcal{U}(0)$, i.e., that (i) holds. □

For positive cost functions ($\mathcal{J}(0,\cdot) \geq 0$), "$J$-critical" is equivalent to "minimizing":

Corollary 8.1.8 (Minimizing ⇔ J-critical & ≥ 0) *A control $u \in \mathcal{U}(x)$ minimizes [strictly] $\mathcal{J}(x,\cdot)$ (on $\mathcal{U}(x)$) iff u is J-critical and $\langle D\eta, JD\eta\rangle \geq 0$ [> 0] for all nonzero* $\eta \in \mathcal{U}(0)$.

Note that $\mathcal{J}(0,\eta) = \langle D\eta, JD\eta\rangle$.

Proof: "If": This follows from Lemma 8.1.7(ii). "Only if": This follows from Lemma 8.1.6 and Lemma 8.1.7(ii). □

All critical controls produce the same cost and the same sensitivity of the cost to a disturbance:

Lemma 8.1.9 ($\mathcal{U}^{\text{crit}}$ and uniqueness) *Denote by $\mathcal{U}^{\text{crit}}(x)$ the set of J-critical controls for each $x \in X$. Then $\mathcal{U}^{\text{crit}}(0)$ is a linear subspace of $\mathcal{U}(0)$. If $u \in \mathcal{U}^{\text{crit}}(x)$, $x, x' \in X$, then $\mathcal{U}^{\text{crit}}(x+x') = u + \mathcal{U}^{\text{crit}}(x')$. Moreover, $\mathcal{J}(x,u) = \mathcal{J}(x,v)$, and $\mathcal{J}(x,u+\eta) = \mathcal{J}(x,v+\eta)$ for $u,v \in \mathcal{U}^{\text{crit}}(x)$, $\eta \in \mathcal{U}(0)$.*

In particular, there is at most one J-critical control for each $x \in X$ iff $\mathcal{U}^{\text{crit}}(0) = \{0\}$.

Proof: 1° $\mathcal{U}^{\text{crit}}(x+x') = u + \mathcal{U}^{\text{crit}}(x')$: Let $u \in \mathcal{U}^{\text{crit}}(x)$. For an arbitrary $u' \in U$, we have $u+u' \in \mathcal{U}^{\text{crit}}(x+x')$ iff

$$\langle Cx + Du + Cx' + Du', JD\eta\rangle = 0 \quad (\eta \in \mathcal{U}(0)), \tag{8.6}$$

equivalently, iff $u' \in \mathcal{U}^{\text{crit}}(x')$. Therefore, $\mathcal{U}^{\text{crit}}(x+x') = u + \mathcal{U}^{\text{crit}}(x')$.

2° Obviously, $\mathcal{U}^{\text{crit}}(0)$ is a subspace of $\mathcal{U}(0)$.

3° Let $u, u+\widetilde{u} \in \mathcal{U}^{\text{crit}}(x)$, (so that $\widetilde{u} \in \mathcal{U}^{\text{crit}}(0)$, by 1°). By Lemma 8.1.7(ii), we have $\mathcal{J}(x, u+\widetilde{u}) = \mathcal{J}(x,u) + \langle D\widetilde{u}, JD\widetilde{u}\rangle = \mathcal{J}(x,u)$. If $\eta \in \mathcal{U}(0)$, then

$$\mathcal{J}(x, u+\widetilde{u}+\eta) = \mathcal{J}(x, u+\widetilde{u}) + \langle D\eta, JD\eta\rangle = \mathcal{J}(x,u) + \langle D\eta, JD\eta\rangle = \mathcal{J}(x, u+\eta). \tag{8.7}$$

4° Since always $0 \in \mathcal{U}^{\text{crit}}(0)$, the last claim follows from the identity $\mathcal{U}^{\text{crit}}(x+x') = u + \mathcal{U}^{\text{crit}}(x')$ (with $x' = 0$). □

We shall later give several sufficient conditions for the existence of a unique J-critical control. Such a control and corresponding state and output are always produced by a kind of an abstract system (in "state feedback form"):

Theorem 8.1.10 ($\boldsymbol{\Sigma_{\text{crit}}}$) *Let there be a unique J-critical control $u_{\text{crit}}(x)$ for each $x \in X$. Define*

$$\Sigma_{\text{crit}} := \begin{bmatrix} A_{\text{crit}} \\ C_{\text{crit}} \\ K_{\text{crit}} \end{bmatrix} : x \mapsto \begin{bmatrix} Ax + Bu_{\text{crit}}(x) \\ Cx + Du_{\text{crit}}(x) \\ u_{\text{crit}} \end{bmatrix} =: \begin{bmatrix} z_{\text{crit}}(x) \\ y_{\text{crit}}(x) \\ u_{\text{crit}}(x) \end{bmatrix}. \tag{8.8}$$

Then $\Sigma_{\text{crit}} \in \mathcal{B}(X, Z^{\text{s}} \times Y^{\text{s}} \times U)$. Moreover, by setting $\mathcal{P} := C_{\text{crit}}^* J C_{\text{crit}} \in \mathcal{B}(X, X^*)$ we obtain

$$\mathcal{J}(x, u_{\text{crit}}(x) + \eta) = \langle x, \mathcal{P}x\rangle_{\langle X, X^*\rangle} + \mathcal{J}(0,\eta) = \langle y_{\text{crit}}(x), Jy_{\text{crit}}(x)\rangle_{Y^{\text{s}}} + \langle D\eta, JD\eta\rangle_{Y^{\text{s}}}. \tag{8.9}$$

for $x \in X$ and $\eta \in \mathcal{U}(0)$.

(Stable case) *If $A \in \mathcal{B}(X, Z^{\text{s}})$ and $C \in \mathcal{B}(X, Y^{\text{s}})$, or $B \in \mathcal{B}(U, Z^{\text{s}})$ and $D \in \mathcal{B}(U, Y^{\text{s}})$, then $C \in \mathcal{B}(X, Y^{\text{s}})$ and $\mathcal{P} = C^* J C_{\text{crit}} = C_{\text{crit}}^* J C$.*

Thus, we can consider $K_{\text{crit}} \in \mathcal{B}(X,U)$ as a "state feedback operator", and Σ_{crit} as the left column of the corresponding "closed-loop system".

In Theorem 8.3.9 we shall show that if $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ is a WPLS, then Σ_{crit} is a WPLS. For sufficiently smooth WPLSs (and for all wpls's), Σ_{crit} becomes the left column of the closed-loop system corresponding to a state feedback operator determined by the stabilizing solution of the corresponding Riccati equation (see, e.g., Section 9.2).

Proof: 1° "$C_{\text{crit}}, K_{\text{crit}} \in \mathcal{B}$": The map $K_{\text{crit}} : x \to u_{\text{crit}}(x)$ are obviously linear, hence so are $C_{\text{crit}} : x \to y_{\text{crit}}(x)$ and $A_{\text{crit}} : x \to z_{\text{crit}}(x)$, i.e., $\Sigma_{\text{crit}} \in \text{Hom}(X, Z^{\text{s}} \times Y^{\text{s}} \times U)$. To show that Σ_{crit} is bounded, we use the Closed Graph Theorem.

Assume that $x_n \to 0$ in X and $\Sigma_{\text{crit}} x_n \to \left[\begin{smallmatrix} z \\ y \\ u \end{smallmatrix}\right]$ in $Z^{\text{s}} \times Y^{\text{s}} \times U$ (by Lemma A.3.4(E1), we only need to show that $z, y, u = 0$).

Then $C_{\text{crit}}x_n \to y \in Y^{\text{s}}$, but $C_{\text{crit}}x_n = Cx_n + DK_{\text{crit}}x_n \to 0 + Du$ in Y, hence $y = Du$. Analogously, $z = Bu$, hence $u \in \mathcal{U}(0)$. But

$$\langle D\eta, JDu\rangle = \lim_{n\to+\infty}\langle D\eta, JC_{\text{crit}}x_n\rangle = \lim_{n\to+\infty} 0 = 0 \tag{8.10}$$

for all $\eta \in \mathcal{U}(0)$, hence $u = u_{\text{crit}}(0) = 0$; consequently, $y = Du = 0$ and $z = Bu = 0$. Thus, $\Sigma_{\text{crit}} \in \mathcal{B}$.

2° *(8.9):* This follows from the definitions of C_{crit} and $\mathcal{J}$, and from Lemma 8.1.7(ii).

3° *Case* $A \in \mathcal{B}(X, Z^{\text{s}})$, $C \in \mathcal{B}(X, Y^{\text{s}})$*:* Let $x \in X$. Obviously, $\mathcal{U}(x) = \mathcal{U}(0)$. Moreover, $DK_{\text{crit}}x = C_{\text{crit}}x - Cx \in Y^{\text{s}}$, and $\langle DK_{\text{crit}}x, JC_{\text{crit}}x\rangle = 0$, because $K_{\text{crit}}x \in \mathcal{U}(x) = \mathcal{U}(0)$. Therefore,

$$\langle Cx, JC_{\text{crit}}x\rangle = \langle C_{\text{crit}}x, JC_{\text{crit}}x\rangle - 0 = \langle x, \mathcal{P}x\rangle \tag{8.11}$$

Because $x \in X$ was arbitrary, we have $\mathcal{P} = C^* JC_{\text{crit}}$, hence $C^*_{\text{crit}}JC = \mathcal{P}^* = \mathcal{P}$.

4° *Case* $B \in \mathcal{B}(U, Z^{\text{s}})$, $D \in \mathcal{B}(U, Y^{\text{s}})$*:* Now $Ax = A_{\text{crit}}x - BK_{\text{crit}}x \in Z^{\text{s}}$, $Cx = C_{\text{crit}}x - DK_{\text{crit}}x \in Y^{\text{s}}$, for any $x \in X$. Therefore, $A \in \mathcal{B}(X, Z^{\text{s}})$, $C \in \mathcal{B}(X, Y^{\text{s}})$, by Lemma A.3.6. Thus, the rest follows from 3°. □

(See the notes on p. 362.)

8.2 Abstract J-coercivity ($\mathcal{J} \mapsto [u \ Du]$)

Coercion. The unpardonable crime.

— Dorothy Miller Richardson (1873–1957)

In this section we aim to give sufficient conditions to the existence of a unique "optimal" (J-critical) control, so that the abstract optimization theory of Section 8.1 can be applied. The most important of such conditions is that the system is "stabilizable" (i.e., $\mathcal{U}(x) \neq \emptyset$ for all $x \in X$) and J-coercive.

In the WPLS framework with $J = I$ (a minimization problem), J-coercivity means the condition $\|\mathbb{D}u\|_2 \geq \varepsilon \|u\|_2$ or something similar, depending on the desired set of admissible controls u. The concept J-coercivity is a generalization of the standard nonsingularity assumptions of several control problems.

Indeed, later in this section and in Section 11.7 (and in their WPLS applications in Chapters 10–12), we shall show that all optimization (sub)problems faced in rather general LQR and H^∞ settings are J-coercive. Moreover, in Propositions 10.3.1 and 10.3.2 we show that (the extensions to WPLSs of) all different classical nonsingularity assumptions on LQR settings are equivalent.

At the end of this section, we shall give applications to stable systems, including the direct formula for the optimal control given in Lemma 8.2.9. Recall the assumptions of Standing Hypothesis 8.1.1 and note also those of Hypothesis 8.2.2.

We start by defining J-coercivity. The map D is J-coercive iff $\mathcal{J}(0,u) \mapsto \begin{bmatrix} u & Du \end{bmatrix}$ is continuous:

Definition 8.2.1 (J-coercive) *We call D* J-coercive (over $\mathcal{U}$) *if there is $\varepsilon > 0$ s.t. for all nonzero $u \in \mathcal{U}(0)$ there is a nonzero $v \in \mathcal{U}(0)$ s.t.*

$$\langle Dv, JDu \rangle \geq \varepsilon \|u\|_D \|v\|_D. \tag{8.12}$$

where $\|u\|_D := \max\{\|u\|_U, \|Bu\|_{Z^s}, \|Du\|_{Y^s}\}$. If, in addition, $\langle Du, JDu \rangle \geq 0$ for each $u \in \mathcal{U}(0)$, then D is called positively J-coercive.

By Lemma 8.2.3(c2), $\mathbb{D}$ is positively J-coercive iff "we can take $v = u$", i.e., iff $\langle Du, JDu \rangle \geq \varepsilon \|u\|_D^2$ for all $u \in \mathcal{U}(0)$ and some $\varepsilon > 0$. Obviously, we can replace $\|\cdot\|_D$ by an equivalent norm in (8.12).

When applying the results of this section, we often take $Z^s = Z$ (or $A = 0 = B$), so that A and B become continuous and hence insignifant and can be dropped from all formulae. In particular, then $\|u\|_D$ becomes equivalent to $\|u\|_D' := \max\{\|u\|_U, \|Du\|_{Y^s}\}$.

We shall use certain techniques that require reflexivity:

Standing Hypothesis 8.2.2 *Throughout the rest of this section, except in Lemmas 8.2.4, 8.2.8 and 8.2.9, we assume that U, Z^s and Y^s are reflexive.*

This hypothesis allows us to fully apply Lemma A.3.5(c2):

Lemma 8.2.3 *Equip $\mathcal{U}(0)$ with norm $\|\cdot\|_D$. Then we have the following:*

(a) **(Stable case)** *Let $Y^{\mathrm{s}} = Y$ and $Z^{\mathrm{s}} = Z$. Then $\mathcal{U}(0)$ and U are (TVS) isomorphic, hence D is J-coercive iff $D^*JD \in \mathcal{GB}(U,U^*)$. Moreover, D is positively J-coercive iff $D^*JD \gg 0$.*

(b) $\mathcal{U}(0)$ is a reflexive Banach space (a Hilbert space if U, Z^{s} and Y^{s} are Hilbert spaces), and $\begin{bmatrix} B \\ D \end{bmatrix}|_{\mathcal{U}(0)} \in \mathcal{B}(\mathcal{U}(0), Z^{\mathrm{s}} \times Y^{\mathrm{s}})$.

(c1) D is J-coercive iff $D^*JD \in \mathcal{GB}(\mathcal{U}(0))$.

(c2) D is positively J-coercive iff $D^*JD \gg 0$ on $\mathcal{U}(0)$.

(d1) $\mathbb{D}$ is positively J-coercive iff there is $\varepsilon > 0$ s.t. for all $u \in \mathcal{U}(0)$ we have

$$\langle Du, JDu\rangle \geq \varepsilon(\|u\|_U^2 + \|Bu\|_{Z^{\mathrm{s}}}^2 + \|Du\|_{Y^{\mathrm{s}}}^2). \tag{8.13}$$

(d2) Let $J \gg 0$. Then D is positively J-coercive iff $\|Du\|_{Y^{\mathrm{s}}} \geq \varepsilon'(\|u\|_U + \|Bu\|_{Z^{\mathrm{s}}})$ for some $\varepsilon' > 0$ and all $u \in \mathcal{U}$.

(d3) If D is J-coercive, then D is injective on U.

(d4) Let D be J-coercive. Then $\|JDu\|_{Y^{\mathrm{s}*}} \geq \varepsilon'\|u\|_D$ for some $\varepsilon' > 0$ and all $u \in \mathcal{U}(0)$.

(e) Even if Standing Hypothesis 8.2.2 does not hold, claims (d1)–(d4) above hold and J-coercivity implies that there is at most one J-critical control for each $x_0 \in X$.

If $J \geq 0$, $Y^{\mathrm{s}} = Y$ and $Z^{\mathrm{s}} = Z$, then D is J-coercive iff $\langle Du, JDu\rangle \geq \varepsilon\|u\|_U^2$ $(u \in U)$ for some $\varepsilon > 0$, by (a). The operator D^*JD can be considered as the Popov Toeplitz operator.

Proof: (a) Obviously, $\|\cdot\|_D$ is now equivalent to $\|\cdot\|_U$ on $U = \mathcal{U}(0)$, hence the claims follow from (c1)&(c2).

(b) By Lemma A.3.15 (with $T := \begin{bmatrix} I^* & B^* & D^* \end{bmatrix}^*$), $\mathcal{U}(0)$ is a reflexive Banach space (a Hilbert space if U, Z^{s} and Y^{s} are Hilbert spaces) and $T \in \mathcal{B}(\mathcal{U}(0), U \times Z^{\mathrm{s}} \times Y^{\mathrm{s}})$.

(c1)&(c2) These follow from Lemma A.3.4(N4)(xi) and Lemma A.3.5(c2)&(d),

(d1) This a reformulation of (c2) (with an equivalent norm).

(d2) Now $D^*JD \gg 0 \Leftrightarrow D^*D \gg 0 \Leftrightarrow \|Du\|_{Y^{\mathrm{s}}} \geq \varepsilon'\|u\|_D$ $(u \in \mathcal{U}(0)) \Leftrightarrow \|Du\|_{Y^{\mathrm{s}}} \geq \varepsilon'(\|u\|_U + \|Bu\|_{Z^{\mathrm{s}}})$ $(u \in \mathcal{U}(0))$ (note that this does not hold for all $u \in U$ in general).

(d3) This follows from (c1).

(d4) Let $u \in \mathcal{U}(0)$. Let $\varepsilon > 0$ and v be as in Definition 8.2.1. Then

$$\varepsilon\|u\|_D\|Dv\| \leq \varepsilon\|u\|_D\|v\|_D \leq \langle Dv, JDu\rangle \leq \|JDu\|\|Dv\|. \tag{8.14}$$

Because $Dv \neq 0$, by (d3), we have $\|JDu\|_2 \geq \varepsilon\|u\|_D$. Since u was arbitrary, (d4) holds.

(e) For (d1)–(d4) this is obvious; the uniqueness claim follows from 2° of the proof of Theorem 8.2.5. □

A coordinate change in the input space affects J-coercivity and J-critical control in the expected way:

Lemma 8.2.4 (D vs. DE) *Let $E \in \mathcal{GB}(U)$. Let $\mathcal{U}_{DE}(x) := \{u \in U \mid \left[\begin{smallmatrix} A & BE \\ C & DE \end{smallmatrix}\right] \left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right] \in Y^{\mathrm{s}} \times Z^{\mathrm{s}}\}$. Then we have the following:*

(a) $\mathcal{U}_{DE}(x) = E^{-1}\mathcal{U}(x)$ for all $x \in X$.

(b) A control u is J-critical for x and $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ iff $E^{-1}u$ is J-critical for x and $\left[\begin{smallmatrix} A & BE \\ C & DE \end{smallmatrix}\right]$.

(c) D is J-coercive iff DE is J-coercive.

Naturally, here the J-coercivity of DE refers to the set $\{u \in U \mid DEu \in Y^{\mathrm{s}}\}$ in place of $\mathcal{U}(0)$ and to the norm $\|u\|_{DE} := \max\{\|u\|, \|BEu\|, \|DEu\|\}$ in place of $\|u\|_D$.

Proof: Claims (a) and (b) are trivial. Choose $\delta > 0$ is s.t. $\delta\|u\| \le \|Eu\| \le \delta^{-1}\|u\|$ $(u \in U)$. Then $\delta\|u\|_D \le \|E^{-1}u\|_{DE} \le \delta^{-1}\|u\|_D$ $(u \in U)$, hence these two norms are equivalent. From this and (a) we obtain (c) easily. □

Trivially, $\mathcal{U}(x) \neq \emptyset$ is a necessary control for the existence of a J-critical control for x. For J-coercive systems, this is also sufficient:

Theorem 8.2.5 (J-coercive $\Rightarrow \exists! J$-critical control) *Assume that D is J-coercive. If $x \in X$ is s.t. $\mathcal{U}(x) \neq \emptyset$, then there is a unique J-critical control for x.*

This follows from the fact that J-coercivity allows us to project an element of $J(Cx + D\mathcal{U}(x))$ continuously. However, J-coercivity is not the weakest possible assumption for the above theorem: whenever $J = I$, $C = 0$ and $D > 0$, then there is a unique $\mathcal{J}$-minimizing control for all $x \in X$ even if $D \not\gg 0$.

Nevertheless, even for $J = I$, conditions that are necessary and sufficient do not seem to be useful (e.g., "for some (hence all) $y \in \mathcal{Y}(x)$, the orthogonal projection of y to $\overline{\mathrm{Ran}(D)}$ is in $\mathrm{Ran}(D)$").

Proof: 1° *Existence:* We write $E := D_{|\mathcal{U}(0)}$ to clarify the proof. By Lemma 8.2.3(b)&(c1), we have $E \in \mathcal{B}(\mathcal{U}(0), Y^{\mathrm{s}})$ and $E^*JE \in \mathcal{GB}(\mathcal{U}(0))$.

Let $x \in X$. Choose $u' \in \mathcal{U}(x)$, and set $y' := Cx + Du' \in Y^{\mathrm{s}}$. Set $u'' := -(E^*JE)^{-1}E^*Jy' \in \mathcal{U}(0)$, $u := u' + u'' \in \mathcal{U}(x)$. Then

$$\langle Cx + Du, JD\eta\rangle_{Y^{\mathrm{s}}} = \langle y' + Eu'', JE\eta\rangle_{Y^{\mathrm{s}}} = \langle E^*Jy' + E^*JEu'', \eta\rangle_D = 0 \qquad (8.15)$$

for all $\eta \in \mathcal{U}(0)$, hence u is J-critical for x.

2° *Uniqueness:* If u is J-critical for $x = 0$, then $\langle Dv, JDu\rangle = 0$ for all $v \in \mathcal{U}(0)$, hence then $\|u\|_D = 0$, hence $u = 0$. Thus, the J-critical control is unique for any $x \in X$, by Lemma 8.1.9. □

Thus, we have solved a rather general minimization problem:

Corollary 8.2.6 (Minimization) *Assume that $\mathcal{U}(x) \neq \emptyset$ for all $x \in X$. If there is $\varepsilon > 0$ s.t.*

$$\mathcal{J}(0,u) \geq \varepsilon(\|u\|_U^2 + \|Bu\|_{Z^s}^2 + \|Du\|_{Y^s}^2) \qquad (u \in \mathcal{U}(0)), \tag{8.16}$$

then there is a unique minimizing control for each $x \in X$. The corresponding state, output and cost are given by (8.8) and (8.9).

Proof: By Lemma 8.2.3(c2), D is positively J-coercive, hence Theorems 8.2.5 and 8.1.10 apply. By Corollary 8.1.8, the unique J-critical control is strictly minimizing (on $\mathcal{U}(x)$). □

A special case of this is the standard LQR problem:

Corollary 8.2.7 (Standard LQR problem) *Let $J \gg 0$ and $\mathcal{U}(x) \neq \emptyset$ for all $x \in X$. Let there be $\varepsilon > 0$ s.t. $\|Du\|_{Y^s} \geq \varepsilon(\|u\|_U + \|Bu\|_{Z^s})$ for all $u \in \mathcal{U}(0)$.*

Then there is a unique minimizing control for each $x \in X$. The corresponding state, output and cost are given by (8.8) and (8.9). □

(This follows directly from Corollary 8.2.6.)

Analogously, if $Y^s = Y$ and $Z^s = Z$ (the stable case), $J \geq 0$, and $\|J^{1/2}Du\|_{Y^s} \geq \varepsilon\|u\|$ for all $u \in U$, then there is a unique minimizing control for each $x \in X$, and Theorems 8.2.5 and 8.1.10 and Lemma 8.2.8 apply.

The condition "$\mathcal{U}(x) \neq \emptyset$ for all $x \in X$" is called the *Finite Cost Condition*, because for $J = I$ this condition holds iff for each x there is a control $u \in \mathcal{U}(x)$ s.t. $\mathcal{J}(x,u) = \|Cx + Du\|_{Y^s}^2 < \infty$.

The proof and implications of Theorem 8.2.5 become easy in the stable case:

Lemma 8.2.8 (Stable J-critical control) *Assume that $\mathcal{U}(x) = U$ for all $x \in X$. Assume, in addition, that the "Popov Toeplitz operator" $T := D^*JD$ be invertible ($T \in \mathcal{GB}(U,U^*)$).*

Then there is a the unique J-critical control for each $x \in X$, and the system Σ_{crit} and J-critical cost operator $\mathcal{P}$ of Theorem 8.3.9 are given by

$$A_{\text{crit}} = A - BT^{-1}D^*JC, \tag{8.17}$$

$$C_{\text{crit}} = \big(I - DT^{-1}D^*J\big)C, \tag{8.18}$$

$$K_{\text{crit}} = -T^{-1}D^*JC, \tag{8.19}$$

$$\mathcal{P} = C^*\big(J - JDT^{-1}D^*J\big)C = C^*JC_{\text{crit}}. \tag{8.20}$$

The condition on T holds iff D is J-coercive, by Lemma 8.2.3(a).

Proof: Now $\langle Cx + Du, JD\eta\rangle = 0$ for all $\eta \in U = \mathcal{U}(0)$ iff $D^*J(Cx + Du) = 0$ ($\in U^*$), i.e., iff $u = -T^{-1}D^*JCx$. The formulae can be computed from (8.8). □

Standard stable LQR and H^∞ problems are of the following form and hence give us the following direct formulae for the solutions:

Lemma 8.2.9 ("Stable Σ with bounded C") *Assume that $\mathcal{U}(x) = U$ for all $x \in X$, $\widetilde{C} \in \mathcal{B}(Z^s, Y^s)$, "$y = \widetilde{C}z + \widetilde{D}u$", i.e., $C = \widetilde{C}A$, $D = \widetilde{C}B + \widetilde{D}$, and that $S := \widetilde{D}^* J \widetilde{D} \in \mathcal{GB}(U, U^*)$, $\widetilde{D}^* J \widetilde{C} = 0$, $T := D^* JD = S + B^* QB \in \mathcal{GB}(U, U^*)$, where $Q := \widetilde{C}^* J \widetilde{C} \in \mathcal{B}(Z^s)$.*

Then Lemma 8.2.8 applies, $K_{\text{crit}} = -S^{-1} B^* Q A_{\text{crit}}$ and $\mathcal{P} = A^* Q A_{\text{crit}}$.

Thus, $u_{\text{crit}}(x) = -S^{-1} B^* Q z_{\text{crit}}(x)$. See Proposition 8.3.10 (and equation (8.93)) for a WPLS application of the above two lemmas and Chapter 10 for further LQR results. Below the proposition we describe two methods for obtaining a direct formula for u_{crit} and $\mathcal{P}$ in the unstable case.

Proof: Note that $\widetilde{D} = D - \widetilde{C}B \in \mathcal{B}(U, Y)$ and $D^* JC = B^* QA$. Multiply (8.19) by T to the left to obtain

$$SK_{\text{crit}} + B^* QBK_{\text{crit}} = -B^* QA, \tag{8.21}$$

i.e., $-SK_{\text{crit}} = B^* Q(A + BK_{\text{crit}}) = B^* QA_{\text{crit}}$, as claimed.

Moreover, $C_{\text{crit}} = \widetilde{C}A_{\text{crit}} + \widetilde{D}K_{\text{crit}}$, hence $\mathcal{P} = C^* JC_{\text{crit}} = A^* QA_{\text{crit}}$. A common alternative formula is

$$\mathcal{P} = C_{\text{crit}}^* JC_{\text{crit}} = \mathbb{A}_{\text{crit}}^* Q \mathbb{A}_{\text{crit}} + K_{\text{crit}}^* SK_{\text{crit}}. \tag{8.22}$$

□

Notes for Sections 8.1 and 8.2

For stable WPLSs, the idea to use Fréchet differentiation for optimal control (cf. Lemma 8.1.6) and the stable Popov Toeplitz operator method of Lemma 8.2.8 were first used in [S97b], which also contains a variant of the stable case of Theorem 8.1.10 for WPLSs. These methods seem to have been used in control theory for several decades,

and the same holds for the alternative "completing the square" method (not presented here) for minimization problems; see [Zwart] for a WPLS application. We have not seen earlier unstable versions of Theorem 8.2.5 in any framework.

See the notes for Section 8.4 for "J-coercivity". As noted below Theorem 8.2.5, one could prove the existence of a unique J-critical control under weaker assumptions than J-coercivity. Such "singular control problems" are usually ruled out by the assumptions, because such settings are rarely encountered in practice and they cannot be solved as satisfactorily.

The abstract setting of these two sections would allow for further extension of control theory (e.g., feedback and coprimeness), and one could easily obtain results analogous to those in Chapter 6 (or to those in the other chapters). However, we do not have the need to address these concepts at this abstract level.

Further notes are given in Sections 8.3 and 8.4, where this abstract theory is applied to WPLSs

8.3 J-critical control for WPLSs

To drift with every passion till my soul
Is a stringed lute on which all winds can play,
Is it for this that I have given away
Mine ancient wisdom, and austere control?
— Oscar Wilde (1856–1900)

In this section we apply the abstract optimization theory of Sections 8.1 and 8.2 to WPLSs.

Standing Hypothesis 8.3.1 *Throughout this Sections 8.3–8.5, U, W, H and Y denote Hilbert spaces of arbitrary dimensions, $\Sigma = \left[\begin{smallmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{smallmatrix}\right] \in \mathrm{WPLS}(U,H,Y)$, and $J = J^* \in \mathcal{B}(Y)$.*

As explained on pp. 351-352, we consider the cost function (8.1).

In the minimization (LQR) problem, one often takes $J = I$ so that $\mathcal{J}(x_0,u) = \int_0^\infty \|y(t)\|_H^2\,dt$, and one wishes to find, for each initial state $x_0 \in H$, a "stabilizing" control $u : \mathbf{R}_+ \to U$ s.t. the cost $\mathcal{J}(x_0,u)$ is minimized over all $u \in \mathcal{U}_*^*(x_0)$, where $\mathcal{U}_*^*(x_0)$ denotes the set of "stabilizing" controls for the initial state x_0.

We may choose $\mathcal{U}_*^*(x_0)$ to be the set of those $u \in \mathrm{L}^2(\mathbf{R}_+;U)$ for which the output $y := \mathbb{C}x_0 + \mathbb{D}u$ is in L^2; we denote this set by $\mathcal{U}_{\mathrm{out}}(x_0)$. The subset of u's for which also the state $x := \mathbb{A}x_0 + \mathbb{B}\tau u$ belongs to L^2 is denoted by $\mathcal{U}_{\mathrm{exp}}(x_0)$. We also allow for other choices of $\mathcal{U}_*^*$, so that we are able to solve control problems with very different stability restrictions depending on the choice.

In other control problems, one may wish to maximize $\mathcal{J}(x_0,\cdot)$ or to find a minimax point of $\mathcal{J}(x_0,\cdot)$ (as for the H^∞ problem). Therefore, instead of minimums, we look for the critical points (the zeros of the Fréchet derivative) of $\mathcal{J}(x_0,\cdot)$ over $\mathcal{U}_*^*(x_0)$; we call these the "J-critical controls" (cf. Lemma 8.3.6). Naturally, all extremums and other saddle points of $\mathcal{J}(x_0,\cdot)$ are J-critical.

In this section we shall study such controls and show that if there is a unique J-critical control for each initial state $x_0 \in H$, then this control and the corresponding state and output can be represented as the state and output of a WPLS (Theorem 8.3.9). This WPLS is obtained by applying certain kind of state feedback to the original system, but this feedback need not be well-posed (unless the WPLS is sufficiently regular); we study necessary and sufficient conditions for its well-posedness (the maps from the state to the output and J-critical control (feedback) are always well-posed but the sensitivity of the feedback loop to external input/disturbance need not be). Such conditions are treated from another point of view in Section 9.14, and further sufficient regularity conditions are given in other sections of Chapter 9.

The "J-critical" system (the closed-loop system in the well-posed case) becomes output stable if we optimize over $\mathcal{U}_{\mathrm{out}}$ and exponentially stable if we optimize over $\mathcal{U}_{\mathrm{exp}}$; conversely, the control corresponding to an output-stabilizing (resp. exponentially stabilizing) state feedback is necessarily in $\mathcal{U}_{\mathrm{out}}$ (resp. $\mathcal{U}_{\mathrm{exp}}$). Thus, optimization over $\mathcal{U}_{\mathrm{out}}$ (resp. $\mathcal{U}_{\mathrm{exp}}$) corresponds to optimization over all output- (resp. exponentially) stabilizing state feedback pairs (under the regularity conditions mentioned above).

Such a J-critical state feedback pair corresponds to a unique "stabilizing" solution of the Riccati equation, see Chapter 9 for details. The corresponding special control problems are solved in Chapters 10–12.

We also present similar results for the case where the critical controls are not unique.

Now it is the time to define $\mathcal{U}_*^*$. Due to generality, the definition contains awfully many symbols, hence we recommend the reader to just note that $\mathcal{U}_{\text{exp}}$ and $\mathcal{U}_{\text{out}}$ are as explained above, observe the alternative (equivalent) definition of J-critical points (controls) of the cost function and ignore the rest of the definition until there is some need for the other $\mathcal{U}_*^*$'s. The general case will become more clear later with the applications:

Definition 8.3.2 ($\mathcal{U}_*^*$ and J-critical control) *Assume that Z^{s} is a Banach space, Z^{u} is a TVS s.t. $Z^{\text{s}} \underset{c}{\subset} Z^{\text{u}}$, $\mathbb{Q} \in \mathcal{B}(H, Z^{\text{u}})$, $\mathbb{R} \in \mathcal{B}(\mathrm{L}^2(\mathbf{R}_+;U), Z^{\text{u}})$ and $\vartheta \in \mathbf{R}$. Define*

$$\mathcal{U}^{\vartheta}_{[\mathbb{Q}\ \mathbb{R}]}(x_0) := \{u \in \mathrm{L}^2_{\vartheta}(\mathbf{R}_+;U) \,\big|\, \left[\begin{smallmatrix}\mathbb{C} & \mathbb{D}\\ \mathbb{Q} & \mathbb{R}\end{smallmatrix}\right]\left[\begin{smallmatrix}x_0\\ u\end{smallmatrix}\right] \in \mathrm{L}^2 \times Z^{\text{s}}\} \qquad (x_0 \in H); \tag{8.23}$$

$$\|u\|_{\mathcal{U}^{\vartheta}_{[\mathbb{Q}\ \mathbb{R}]}} := \max\{\|u\|_{\mathrm{L}^2_{\vartheta}}, \|\mathbb{R}u\|_{Z^{\text{s}}}, \|\mathbb{D}u\|_2\} \leq \infty \qquad (u \in \mathrm{L}^2_{\vartheta}(\mathbf{R}_+;U)); \tag{8.24}$$

$$x_{x_0,u} := \mathbb{A}x_0 + \mathbb{B}\tau u \qquad (u \in \mathrm{L}^2(\mathbf{R}_+;U)); \tag{8.25}$$

$$y_{x_0,u} := \mathbb{C}x_0 + \mathbb{D}u \qquad (u \in \mathrm{L}^2(\mathbf{R}_+;U)); \tag{8.26}$$

$$\mathcal{Y}_*^*(x_0) := \{y_{x_0,u} \,\big|\, u \in \mathcal{U}_*^*(x_0)\}; \tag{8.27}$$

$$\mathcal{J}(x_0,u) := \langle y_{x_0,u}, J y_{x_0,u}\rangle \qquad (u \in \mathcal{U}_*^*(x_0)). \tag{8.28}$$

We use $\mathcal{U}_*^* := \mathcal{U}^{\vartheta}_{[\mathbb{Q}\ \mathbb{R}]}$ when we do not wish to specify $\mathbb{Q}$, $\mathbb{R}$, Z^{u} and Z^{s} and ϑ. We are mainly interested in the following choices of $\mathcal{U}_*^*$:

$$\mathcal{U}_{\text{out}}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,\big|\, y_{x_0,u} \in \mathrm{L}^2\}, \tag{8.29}$$

$$\mathcal{U}_{\text{sta}}(x_0) := \{u \in \mathcal{U}_{\text{out}}(x_0) \,\big|\, \sup \|x_{x_0,u}\| < \infty\}, \tag{8.30}$$

$$\mathcal{U}_{\text{str}}(x_0) := \{u \in \mathcal{U}_{\text{out}}(x_0) \,\big|\, \|x_{x_0,u}(t)\|_H \to 0 \text{ as } t \to +\infty\}, \tag{8.31}$$

$$\mathcal{U}_{\text{exp}}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,\big|\, x_{x_0,u} \in \mathrm{L}^2\}; \tag{8.32}$$

i.e., then we assume that $\vartheta = 0$, $\left[\mathbb{Q} \quad \mathbb{R}\right] = \left[\begin{array}{c|c}\mathbb{A} & \mathbb{B}\tau\end{array}\right]$, $Z^{\text{u}} = \mathrm{L}^2_{\omega}(\mathbf{R}_+;H)$, and $Z^{\text{s}} = \mathrm{L}^2_{\omega}(\mathbf{R}_+;H)$, $Z^{\text{s}} = \mathrm{L}^{\infty}(\mathbf{R}_+;H)$, $Z^{\text{s}} = \mathcal{C}_0(\mathbf{R}_+;H)$ or $Z^{\text{s}} = \mathrm{L}^2(\mathbf{R}_+;H)$, respectively, where $\omega := \omega_A + 1$.

We call $\mathcal{J}$ the cost function. *A control $u \in \mathcal{U}_*^*(x_0)$ (resp. output $y_{x_0,u} \in \mathcal{Y}_*^*(x_0)$) is called J-*critical *(over $\mathcal{U}_*^*$) for x_0 (and Σ) if $\langle \mathbb{D}\pi_+\eta, J y_{x_0,u}\rangle_{\mathrm{L}^2} = 0$ for all $\eta \in \mathcal{U}_*^*(0)$.*

*We call an admissible state feedback pair $\left[\begin{array}{c|c}\mathbb{K} & \mathbb{F}\end{array}\right]$ (or a corresponding state feedback operator) J-*critical *(over $\mathcal{U}_*^*$ for Σ) if $\mathbb{K}_\flat x_0 := (I - \mathbb{F})^{-1}\mathbb{K}x_0$ is J-critical for all $x_0 \in H$.*

Thus, we usually write $\mathcal{U}_*^*$ instead of $\mathcal{U}^{\vartheta}_{[\mathbb{Q}\ \mathbb{R}]}$ (and we often omit "over $\mathcal{U}_*^*$"). Of course, we still assume that some $\begin{bmatrix}\mathbb{Q} & \mathbb{R}\end{bmatrix}$, Z^{u}, Z^{s} and ϑ are given. The optimal controls for most reasonable control problems are J-critical over $\mathcal{U}_*^*$ where $\mathcal{U}_*^*$ is the set over which one wishes to optimize, as explained above.

Note that $\|\cdot\|_{\mathcal{U}_*^*}$ is a norm on $\mathcal{U}_*^*(0)$; it will be used to define J-coercivity in the next section. Some simplifications of these norms are given in Lemma 8.4.2.

By Lemma 6.7.8, we have

$$\mathcal{U}_{\mathrm{exp}}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,\big|\, x_{x_0,u}, y_{x_0,u} \in \mathrm{L}^2\}. \tag{8.33}$$

Thus, $\mathcal{U}_{\mathrm{exp}}(x_0)$ is, indeed, the set of "exponentially stabilizing controls" ($x, y \in \mathrm{L}^2$) described before the Definition 8.3.2. Analogously, $\mathcal{U}_{\mathrm{out}}(x_0)$ is the set of "output stabilizing controls" ($y \in \mathrm{L}^2$); $\mathcal{U}_{\mathrm{sta}}(x_0)$ is the set of "stabilizing controls" ($y \in \mathrm{L}^2$, x bounded); and $\mathcal{U}_{\mathrm{str}}(x_0)$ is the set of "strongly stabilizing controls" ($y \in \mathrm{L}^2$, x strongly stable).

In finite-dimensional problems, one usually optimizes over $\mathcal{U}_{\mathrm{exp}}$. In applications, often physical quantities may determine a natural norm for the state, and this norm might be such that one does not want to require the "optimal" control to be exponentially stabilizing. Therefore, also set $\mathcal{U}_{\mathrm{out}}$ has often been used, particularly for infinite-dimensional problems (see, e.g., [Zwart], [WW], [LT00a]), and there are at least some kind of implicit applications of $\mathcal{U}_{\mathrm{str}}$ [Oostveen] and $\mathcal{U}_{\mathrm{sta}}$ [S97b]–[S98d].

Coercivity assumptions that guarantee the existence of a unique J-critical control over $\mathcal{U}_{\mathrm{out}}$ are very natural whereas their analogies for $\mathcal{U}_{\mathrm{exp}}$ are substantially stronger though still rather commonly used (compare Propositions 10.3.1 to 10.3.2). Nevertheless, in the literature one often uses just the former assumption and obtains an optimal control which is optimal over $\mathcal{U}_{\mathrm{exp}}$ too — how is this trick possible? The secret is to assume exponential detectability, since it implies that the four sets coincide:

Lemma 8.3.3 ($\boldsymbol{\mathcal{U}_{\mathrm{exp}} = \mathcal{U}_{\mathrm{out}}}$) *We have $\mathcal{U}_{\mathrm{exp}}(x_0) \subset \mathcal{U}_{\mathrm{out}}(x_0)$ and $\mathcal{U}_{\mathrm{str}}(x_0) \subset \mathcal{U}_{\mathrm{sta}}(x_0) \subset \mathcal{U}_{\mathrm{out}}(x_0)$ for all $x_0 \in H$.*

If Σ is estimatable or exponentially q.r.c.-stabilizable (e.g., exponentially stable), then $\mathcal{U}_{\mathrm{exp}} = \mathcal{U}_{\mathrm{str}} = \mathcal{U}_{\mathrm{sta}} = \mathcal{U}_{\mathrm{out}}$, and then Σ is [positively] J-coercive over $\mathcal{U}_{\mathrm{out}}$ iff Σ is [positively] J-coercive over $\mathcal{U}_{\mathrm{exp}}$.

If Σ is [strongly] q.r.c.-stabilizable (e.g., [strongly] stable), then $[\mathcal{U}_{\mathrm{str}} =]\, \mathcal{U}_{\mathrm{sta}} = \mathcal{U}_{\mathrm{out}}$. If u is J-critical for x_0 over $\mathcal{U}_{\mathrm{out}}$ and $\mathcal{U}_{\mathrm{str}}$, then u is J-critical for x_0 over $\mathcal{U}_{\mathrm{sta}}$.

(Since $\ell^2 \subset c_0$, we have $\mathcal{U}_{\mathrm{exp}} \subset \mathcal{U}_{\mathrm{str}}$ in discrete time (cf. Theorem 13.3.13).)

Thus, when Σ is estimatable, we may equivalently optimize over any of these sets for main domains for u, and we only have to look for J-coercivity over $\mathcal{U}_{\mathrm{out}}$ (and optimizability) to guarantee the existence of a unique J-critical control (see Definition 8.4.1 and Theorem 8.4.3).

Proof: (By $\mathcal{U}_{\mathrm{out}} = \mathcal{U}_{\mathrm{sta}}$ we mean equality as functions of $x_0 \in H$ (to L^2); thus, we could as well write $\mathcal{U}_{\mathrm{out}} \equiv \mathcal{U}_{\mathrm{sta}}$.)

The claims on estimatability follow from Theorem 6.7.7. The claims on q.r.c.-stabilizability are given in Theorem 8.4.5(g2).

The rest is rather obvious. (In fact, whenever two of these four spaces are equal for $x_0 = 0$, then they have equal norms, by Lemma A.3.6 (with, e.g., $\|u\|_{X_3} := \|u\|_2 + \|y\|_2 + \|x\|_{\mathrm{L}^2_1}$), and hence [positive] J-coercivity over them become equivalent, by (c1) and (c2) of Lemma 8.2.3.) □

Trivially, the condition that $\mathcal{U}^*_*(x_0) \neq \emptyset$ for all $x_0 \in H$ is necessary for the existence of a J-critical control for any initial state, hence for the solvability of any reasonable optimization problem (over $\mathcal{U}^*_*$). This condition is called the *Finite Cost Condition*, since, for $J \gg 0$, it corresponds to the existence of u s.t. "$\mathcal{J}(x_0,u) < \infty$". For $\mathcal{U}^*_* = \mathcal{U}_{\mathrm{exp}}$ this concept obviously coincides with optimizability.

After a "translation of notation", the results of previous sections can be read as results for WPLSs:

Remark 8.3.4 *We can apply the theory of Sections 8.1 and 8.2 by substitutions* $X \mapsto H$, $U \mapsto \mathrm{L}^2_\vartheta(\mathbf{R}_+;U)$, $Y^{\mathrm{s}} \mapsto \mathrm{L}^2(\mathbf{R}_+;Y)$, $Y \mapsto \mathrm{L}^2_\beta(\mathbf{R}_+;Y)$, $Z^{\mathrm{s}} \mapsto \mathrm{L}^2_\beta(\mathbf{R}_+;H) \times Z^{\mathrm{s}}$, $Z \mapsto \mathrm{L}^2_\beta(\mathbf{R}_+;H) \times Z^{\mathrm{u}}$ *and* $\left[\begin{array}{c|c} A & B \end{array}\right] \mapsto \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B}\tau \\ \mathbb{Q} & \mathbb{R} \end{array}\right]$, $\left[\begin{array}{c|c} C & D \end{array}\right] \mapsto \left[\begin{array}{c|c} \mathbb{C} & \mathbb{D} \end{array}\right]$, *where* $\beta > \max\{\vartheta, \omega_A\}$.

(Note that Σ *and* $\mathbb{B}\tau$ *are* β*-stable. Above, one can equivalently write explicitly* $B \mapsto \pi_+\mathbb{B}\tau\pi_+$, $D \mapsto \pi_+\mathbb{D}\pi_+$.*)*

Thus, the concepts $\mathcal{J}$*, "J-critical" and "J-coercive" of those two sections coincide with those of this section, and* $\mathcal{U}$ *becomes* $\mathcal{U}^*_*$.

Consequently, we obtain Lemmas 8.3.5, 8.3.6, 8.3.7 and 8.3.8, and Proposition 8.3.10 and Theorem 8.4.3 from corresponding results in Sections 8.1 and 8.2 (see those sections for further results). □

(By Remark 8.1.2, we might also allow completely unstable controls by substitution $U \mapsto \mathrm{L}^2_{\mathrm{loc}}(\mathbf{R}_+;U)$.)

Next we list the "translated" auxiliary lemmas mentioned in the remark:

Lemma 8.3.5 ($\mathcal{U}^*_*(x_0)$ and $\mathcal{Y}^*_*(x_0)$) *The sets* $\mathcal{U}^*_*(0) \subset \mathrm{L}^2_\vartheta(\mathbf{R}_+;U)$ *and* $\mathcal{Y}^*_*(0) \subset \mathrm{L}^2(\mathbf{R}_+;Y)$ *are linear subspaces. Let* $x_0 \in H$, $u \in \mathcal{U}^*_*(x_0)$, $y \in \mathcal{Y}^*_*(x_0)$. *Then* $\mathcal{U}^*_*(x_0) = u + \mathcal{U}^*_*(0)$ *and* $\mathcal{Y}^*_*(x_0) = y + \mathcal{Y}^*_*(0)$.

Moreover, $\mathcal{U}^*_*(\alpha x_0 + \beta x_1) = \alpha\mathcal{U}^*_*(x_0) + \beta\mathcal{U}^*_*(x_1)$ *and* $\mathcal{Y}^*_*(\alpha x_0 + \beta x_1) = \alpha\mathcal{Y}^*_*(x_0) + \beta\mathcal{Y}^*_*(x_1)$ *(*$\alpha,\beta \in \mathbf{C} \setminus \{0\}$, $x_0, x_1 \in H$ *s.t.* $\mathcal{U}^*_*(x_0) \neq \emptyset$*).* □

As claimed above, a control u_{crit} is J-critical iff the gradient of the cost $u \mapsto \langle y, Jy \rangle$ is zero at u_{crit}:

Lemma 8.3.6 (J-critical $\Leftrightarrow \frac{d\mathcal{J}}{du} = 0$) *A control* $u_{\mathrm{crit}}(x_0) \in \mathcal{U}^*_*(x_0)$ *is J-critical for* x_0 *iff* $\frac{d\mathcal{J}(x_0,u)}{du}(u_{\mathrm{crit}}(x_0)) = 0$.

In particular, if $u_{\mathrm{crit}}(x_0)$ *is a local extremal point or saddle point of* $\mathcal{J}(x_0,u)$ *(over* $\mathcal{U}^*_*(x_0)$*), then* $u_{\mathrm{crit}}(x_0)$ *is J-critical for* x_0. □

The expression $\frac{d\mathcal{J}(x_0,u)}{du}$ denotes the (real) *Fréchet derivative* of $\mathcal{J}(x_0,\cdot)$ on its domain $\mathcal{U}^*_*(x_0)$; see Lemma 8.1.6 for details.

Saddle points correspond to solutions of the H^∞ minimax problem.

Lemma 8.3.7 (Critical cost $\mathcal{J}(x_0, u_{\rm crit})$) *Let $x_0 \in H$ and $u_{\rm crit}(x_0) \in \mathcal{U}_*^*(x_0)$. Set $y_{\rm crit}(x_0) := y_{x_0, u_{\rm crit}(x_0)}$. Then the following are equivalent:*

(i) $u_{\rm crit}(x_0)$ is J-critical for x_0;

(ii) $\mathcal{J}(x_0, u_{\rm crit}(x_0) + \eta) = \langle y_{\rm crit}(x_0), J y_{\rm crit}(x_0)\rangle + \langle \mathbb{D}\eta, J\mathbb{D}\eta\rangle \quad (\eta \in \mathcal{U}_*^*(0))$;

(iii) $\langle y_{x_0, u_{\rm crit}(x_0)+\eta_1}, J y_{x_0, u_{\rm crit}(x_0)+\eta_2}\rangle = \langle y_{\rm crit}(x_0), J y_{\rm crit}(x_0)\rangle + \langle \mathbb{D}\eta_1, J\mathbb{D}\eta_2\rangle$ $(\eta_1, \eta_2 \in \mathcal{U}_*^*(0))$.

□

Note that (ii) means that $\mathcal{J}(x_0, u_{\rm crit}(x_0) + \eta) = \mathcal{J}(x_0, u_{\rm crit}(x_0)) + \mathcal{J}(0, \eta)$. Thus, given a critical control $u_{\rm crit}$, the cost for $u_{\rm crit} + \eta$ equals the critical cost plus the cost for η.

All critical controls produce the same cost and the same sensitivity of the cost to a disturbance:

Lemma 8.3.8 ($\mathcal{U}_*^{*,\rm crit}$ and uniqueness) *Let $\mathcal{U}_*^{*,\rm crit}(x_0)$ be the set of J-critical controls for $x_0 \in H$. Then $\mathcal{U}_*^{*,\rm crit}(0)$ is a linear subspace of $\mathcal{U}_*^*(0)$. If $u \in \mathcal{U}_*^{*,\rm crit}(x_0)$, then $\mathcal{U}_*^{*,\rm crit}(x_0) = u + \mathcal{U}_*^{*,\rm crit}(0)$. Moreover, $\mathcal{J}(x_0, u) = \mathcal{J}(x_0, v)$, and $\mathcal{J}(x_0, u + \eta) = \mathcal{J}(x_0, v + \eta)$ for $u, v \in \mathcal{U}_*^{*,\rm crit}(x_0)$, $\eta \in \mathcal{U}_*^*(0)$.*

In particular, there is at most one J-critical control for each $x \in X$ iff $\mathcal{U}^{\rm crit}(0) = \{0\}$. □

We shall later meet several sufficient conditions for the existence of a unique J-critical control. Such a control and corresponding state and output are always produced by a WPLS:

Theorem 8.3.9 ($\Sigma_{\rm crit}$) *Assume that there is a unique J-critical control $u_{\rm crit}(x_0)$ over $\mathcal{U}_*^*$ for each $x_0 \in H$, and define*

$$\Sigma_{\rm crit} := \left[\begin{array}{c|c} \mathbb{A}_{\rm crit} & \\ \hline \mathbb{C}_{\rm crit} & \\ \mathbb{K}_{\rm crit} & \end{array}\right] : x_0 \mapsto \left[\begin{array}{c|c} x_{\rm crit}(x_0) & \\ \hline y_{\rm crit}(x_0) & \\ u_{\rm crit}(x_0) & \end{array}\right] := \left[\begin{array}{c|c} \mathbb{A}x_0 + \mathbb{B}\tau u_{\rm crit}(x_0) & \\ \hline \mathbb{C}x_0 + \mathbb{D}u_{\rm crit}(x_0) & \\ u_{\rm crit}(x_0) & \end{array}\right]. \tag{8.34}$$

(Alternatively, we may assume that $\Sigma_{\rm crit}$ is any J-critical control in WPLS form (see Definition 8.3.15).)

Then the following hold except that in (a1)–(a5) and (b2) we assume, in addition, that $\vartheta = 0$ (e.g., that $\mathcal{U}_*^* \in \{\mathcal{U}_{\rm out}, \mathcal{U}_{\rm sta}, \mathcal{U}_{\rm str}, \mathcal{U}_{\rm exp}\}$).

(a1) The maps $\mathbb{C}_{\rm crit}$ and $\mathbb{K}_{\rm crit}$ are stable, and $\Sigma_{\rm crit} \in \mathrm{WPLS}(\{0\}, H, Y \times U)$.

(a2) If $\mathcal{U}_*^* = \mathcal{U}_{\rm exp}$, then $\Sigma_{\rm crit}$ is exponentially stable; if $\mathcal{U}_*^* = \mathcal{U}_{\rm str}$, then $\Sigma_{\rm crit}$ is strongly stable; if $\mathcal{U}_*^* = \mathcal{U}_{\rm sta}$, then $\Sigma_{\rm crit}$ is stable.

(a3) If $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \end{array}\right]$ is [strongly] stable, then $\Sigma_{\rm crit}$ is [strongly] stable.

(a4) If $\omega \geq 0$ is s.t. $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \end{array}\right] \in \mathrm{WPLS}_\omega$, then $\Sigma_{\rm crit} \in \mathrm{WPLS}_\omega$.

(a5) If Σ is estimatable, then $\Sigma_{\rm crit}$ is exponentially stable.

(a1') $\mathbb{C}_{\rm crit}$ is stable, $\mathbb{K}_{\rm crit}$ is ϑ-stable, $\mathbb{Q} + \mathbb{R}\mathbb{K}_{\rm crit} \in \mathcal{B}(H, Z^{\rm s})$, and $\Sigma_{\rm crit} \in \mathrm{WPLS}(\{0\}, H, Y \times U)$.

(a2') If $\mathcal{U}_*^{*,\mathrm{crit}}(x_0) \subset \mathcal{U}_{\mathrm{exp}}(x_0)$ for all $x_0 \in H$, then Σ_{crit} is exponentially stable.

(a4') If $\omega \geq \vartheta$ is s.t. $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \end{array}\right] \in \mathrm{WPLS}_\omega$, then $\Sigma_{\mathrm{crit}} \in \mathrm{WPLS}_\omega$.

(b1) We call $\mathcal{P} := \mathbb{C}_{\mathrm{crit}}^* J \mathbb{C}_{\mathrm{crit}} \in \mathcal{B}(H)$ the J-critical cost operator. *It satisfies*

$$\mathcal{J}(x_0, u_{\mathrm{crit}}(x_0) + \eta) = \langle x_0, \mathcal{P} x_0 \rangle_H + \mathcal{J}(0, \eta) \quad (x_0 \in H,\ \eta \in \mathcal{U}_*^*(0)). \tag{8.35}$$

(b2) **(Stable case)** *If 1. $\mathcal{U}_*^* = \mathcal{U}_{\mathrm{out}}$, and $\mathbb{C}$ or $\mathbb{D}$ is stable; 2. $\mathcal{U}_*^* = \mathcal{U}_{\mathrm{sta}}$, and $\mathbb{A}$ and $\mathbb{C}$ (or $\mathbb{B}$ and $\mathbb{D}$) are stable; 3. $\mathcal{U}_*^* = \mathcal{U}_{\mathrm{str}}$, and $\mathbb{A}$ and $\mathbb{C}$ (or $\mathbb{B}$ and $\mathbb{D}$) are strongly stable; or 4. $\mathcal{U}_*^* = \mathcal{U}_{\mathrm{exp}}$, and $\mathbb{A}$ is exponentially stable or $\mathbb{B}\tau$ stable; then $\mathbb{C}$ is stable and*

$$\mathcal{P} := \mathbb{C}_{\mathrm{crit}}^* J \mathbb{C}_{\mathrm{crit}} = \mathbb{C}^* J \mathbb{C}_{\mathrm{crit}} = \mathbb{C}_{\mathrm{crit}}^* J \mathbb{C}. \tag{8.36}$$

(b2') If $\mathbb{C}$ is stable and $\mathbb{Q}[H] \subset Z^{\mathrm{s}}$, or $\mathbb{D}$ is stable and $\mathbb{R} \in \mathcal{B}(\mathrm{L}^2_\vartheta(\mathbf{R}_+; U), Z^{\mathrm{s}})$, then $\mathbb{C}$ is stable and (8.36) holds.

The (optimal) control $\mathbb{K}_{\mathrm{crit}} x_0$ equals $(K_{\mathrm{crit}})_{\mathrm{w}} x$ a.e., where $x := \mathbb{A}_{\mathrm{crit}} x_0$ is the state of Σ_{crit} with $A_{\mathrm{crit}} = A + BK_{\mathrm{crit}}$, by Lemma 8.3.17(a). Thus, such a control corresponds to some kind of state feedback, but the feedback loop need not be well-posed (indeed, the "maps $\mathbb{K}$, $\mathbb{F}$, $\mathbb{B}_\flat$, $\mathbb{D}_\flat$, $\mathbb{F}_\flat$ of Definition 6.6.10" need not be well posed); see Remark 9.7.7 and Examples 8.4.13 and 11.3.7 (which also cover the stable setting of Proposition 8.3.10 below). This means that any external input (e.g., disturbance or modelling error) might "explode" the system.

The corresponding generalized Riccati equations are treated in Section 9.7; the one in [FLT] is a special case of these. The rest of Sections 9.1–9.12 treat the case where the "optimal state feedback" is well posed, by which we mean that Σ_{crit} is the left column of $\Sigma_\flat$ for some admissible state feedback pair $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$ for Σ.

In discrete-time, a unique minimizing control is always of this form, and it corresponds to the unique $\mathcal{U}_*^*$-stabilizing solution of the DARE, by Theorem 14.1.6. If Σ is sufficiently regular, then the same holds in continuous time too (see, e.g., Lemma 8.3.18 or Remark 9.9.14).

Proof of Theorem 8.3.9: (a1') If Σ_{crit} is a general J-critical control in WPLS form, then it is rather obvious that (a1') holds; therefore, we assume below that there is a unique J-critical control over $\mathcal{U}_*^*$.

Since the rest follows from Corollary 8.1.10, we only have to show that $\Sigma_{\mathrm{crit}} \in \mathrm{WPLS}$.

Let $x_0 \in H$, $t \geq 0$. We first show that $\pi_+ \tau^t \mathbb{K}_{\mathrm{crit}} x_0$ is J-critical for $\mathbb{A}_{\mathrm{crit}}^t x_0$, i.e., equal to $\mathbb{K}_{\mathrm{crit}} \mathbb{A}_{\mathrm{crit}}^t x_0$: For $\eta \in \mathcal{U}_*^*(0)$ we have $\tau^{-t}\eta \in \mathcal{U}_*^*(0)$, hence

$$\langle J \pi_+ \tau^t \mathbb{C}_{\mathrm{crit}} x_0, \mathbb{D}\eta \rangle_{\mathrm{L}^2} = \langle J \mathbb{C}_{\mathrm{crit}} x_0, \mathbb{D}\tau^{-t}\eta \rangle_{\mathrm{L}^2} = 0 \quad (\eta \in \mathcal{U}_*^*(0)). \tag{8.37}$$

But

$$\pi_+ \tau^t \mathbb{C}_{\mathrm{crit}} x_0 = \pi_+ \tau^t (\mathbb{C} x_0 + \mathbb{D}\mathbb{K}_{\mathrm{crit}} x_0) = \mathbb{C}\mathbb{A}^t x_0 + \pi_+ \mathbb{D}(\pi_+ + \pi_-)\tau^t \mathbb{K}_{\mathrm{crit}} x_0 \tag{8.38}$$

$$= \mathbb{C}\mathbb{A}^t x_0 + \mathbb{D}\pi_+ \tau^t \mathbb{K}_{\mathrm{crit}} x_0 + \mathbb{C}\mathbb{B}\tau^t \mathbb{K}_{\mathrm{crit}} x_0 = \mathbb{C}\mathbb{A}_{\mathrm{crit}}^t x_0 + \mathbb{D}\pi_+ \tau^t \mathbb{K}_{\mathrm{crit}} x_0. \tag{8.39}$$

This and (8.37) imply that $\pi_+\tau^t\mathbb{K}_{\rm crit}x_0$ is J-critical for $\mathbb{A}^t_{\rm crit}x_0$; thus

$$\pi_+\tau^t\mathbb{K}_{\rm crit}x_0 = u_{\rm crit}(\mathbb{A}^t_{\rm crit}x_0) = \mathbb{K}_{\rm crit}\mathbb{A}^t_{\rm crit}x_0,\ \ \pi_+\tau^t\mathbb{C}_{\rm crit}x_0 = y_{\rm crit}(\mathbb{A}^t_{\rm crit}x_0) = \mathbb{C}_{\rm crit}\mathbb{A}^t_{\rm crit}x_0. \tag{8.40}$$

By the dynamic programming principle, $\mathbb{A}$ is a semigroup; a detailed proof of this fact goes as follows, using (8.40):

$$\mathbb{A}^s_{\rm crit}\mathbb{A}^t_{\rm crit} = \mathbb{A}^s(\mathbb{A}^t + \mathbb{B}\tau^t\mathbb{K}_{\rm crit}) + \mathbb{B}\tau^s\mathbb{K}_{\rm crit}\mathbb{A}^t_{\rm crit} \tag{8.41}$$

$$= \mathbb{A}^s\mathbb{A}^t + \mathbb{B}\tau^s\pi_-\tau^t\mathbb{K}_{\rm crit} + \mathbb{B}\tau^s\pi_+\tau^t\mathbb{K}_{\rm crit} = \mathbb{A}^s\mathbb{A}^t + \mathbb{B}\tau^{s+t}\mathbb{K}_{\rm crit} = \mathbb{A}^{t+s}_{\rm crit}. \tag{8.42}$$

Obviously, $\mathbb{A}^0_{\rm crit} = \mathbb{A}^0 = I$. By Theorem 6.2.13(a1), $\mathbb{A}_{\rm crit}x_0 = x_{\rm crit}(x_0)$ is continuous for each $x_0 \in H$. Therefore, $\mathbb{A}_{\rm crit}$ is a C_0-semigroup. This and (8.40) imply that $\Sigma_{\rm crit}$ is a WPLS.

(a1)&(a4) These follow from (a1') and (a4'), respectively.

(a2) The exponentially stable case follows from (a2'); the rest is obvious.

(a2') Now $\mathbb{A}_{\rm crit}x_0 \in \mathrm{L}^2$ for all $x_0 \in H$, hence $\mathbb{A}_{\rm crit}$ is exponentially stable, by Lemma A.4.5.

(a3) See the proof of Lemma 6.6.8(a).

(a4') For any $u \in \mathrm{L}^2_\omega(\mathbf{R}_+;U)$, we have $\|\tau^t u\|_{\mathrm{L}^2_\omega} = e^{\omega t}\|u\|_{\mathrm{L}^2_\omega}$ hence $\|\mathbb{B}\tau^t u\|_H \le \|\mathbb{B}\|e^{\omega t}\|u\|_{\mathrm{L}^2_\omega}$. Thus, if $\mathbb{A}$, $\mathbb{B}$ and $\mathbb{K}_{\rm crit}$ are ω-stable, then so is $\mathbb{A}_{\rm crit}$.

(a5) *Exponential stability:* Let $x_0 \in H$. Because $u_{\rm crit}(x_0), y_{\rm crit}(x_0) \in \mathrm{L}^2$, we have $\mathbb{A}_{\rm crit}x_0 = \mathbb{A}x_0 + \mathbb{B}u_{\rm crit}(x_0) \in \mathrm{L}^2$, by Theorem 6.7.7. By Lemma A.4.5, $\mathbb{A}_{\rm crit}$ is exponentially stable, hence so is $\Sigma_{\rm crit}$.

(b1) This follows from Corollary 8.1.10 (or directly form the definitions of $\mathbb{C}_{\rm crit}$, $\mathcal{J}$ and $\mathcal{P}$).

(b2') This follows from Theorem 8.1.10.

(b2) This follows from (b2') except that for $\mathcal{U}_{\rm exp}$ we also used the following: if $\mathbb{B}\tau$ is stable and $\mathcal{U}^*_* = \mathcal{U}_{\rm exp}$, then $\mathbb{A} = \mathbb{A}_{\rm crit} - \mathbb{B}\tau\mathbb{K}_{\rm crit}$ is exponentially stable, by Lemma A.4.5; if $\mathbb{A}$ is exponentially stable, then so is Σ.

(N.B. if $\mathbb{A}$ or $\mathbb{B}$ is exponentially stable, then so is $\mathbb{B}\tau$, by Lemma 6.1.10.) □

The proof and implications of Theorem 8.4.3 become simple in the stable case:

Proposition 8.3.10 (Stable *J*-critical control) *Assume that $\mathcal{U}^*_*(x_0) = \mathrm{L}^2(\mathbf{R}_+;U)$ for all $x_0 \in H$. Assume, in addition, that the Popov Toeplitz operator $T := \pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible (i.e., $T \in \mathcal{GB}(\mathrm{L}^2(\mathbf{R}_+;U))$).*

Then there is a the unique J-critical control for each $x_0 \in H$, and the system $\Sigma_{\rm crit}$ and J-critical cost operator $\mathcal{P}$ of Theorem 8.3.9 are given by

$$\mathbb{A}_{\rm crit} = \mathbb{A} - \mathbb{B}\tau\pi_+T^{-1}\pi_+\mathbb{D}^*J\mathbb{C}, \tag{8.43}$$

$$\mathbb{C}_{\rm crit} = \left(I - \mathbb{D}\pi_+T^{-1}\pi_+\mathbb{D}^*J\right)\mathbb{C}, \tag{8.44}$$

$$\mathbb{K}_{\rm crit} = -T^{-1}\pi_+\mathbb{D}^*J\mathbb{C}, \tag{8.45}$$

$$\mathcal{P} = \mathbb{C}^*\left(J - J\mathbb{D}\pi_+T^{-1}\pi_+\mathbb{D}^*J\right)\mathbb{C} = \mathbb{C}^*J\mathbb{C}_{\rm crit}. \tag{8.46}$$

If, in addition, $C \in \mathcal{B}(H,Y)$, $D^*JC = 0$ and $R := D^*JD \in \mathcal{GB}(U)$, then

$\mathbb{K}_{\text{crit}} = -R^{-1}(\pi_+\mathbb{B}\tau\pi_+)^*Q\mathbb{A}_{\text{crit}}$, *i.e.*, $u_{\text{crit}}(x_0) = -R^{-1}(\pi_+\mathbb{B}\tau\pi_+)^*Qx_{\text{crit}}(x_0)$ *for all* $x_0 \in H$, *where* $Q := C^*JC$ *(so that* $\mathcal{J} = \langle x, Qx\rangle + \langle u, Ru\rangle$*).*

Note that $\mathcal{U}_{\text{out}} \equiv \mathrm{L}^2(\mathbf{R}_+;U)$ iff $\Sigma \in$ SOS, $\mathcal{U}_{\text{sta}} \equiv \mathrm{L}^2(\mathbf{R}_+;U)$ iff Σ is stable, $\mathcal{U}_{\text{str}} \equiv \mathrm{L}^2(\mathbf{R}_+;U)$ iff Σ is strongly stable, and $\mathcal{U}_{\text{exp}} \equiv \mathrm{L}^2(\mathbf{R}_+;U)$ iff Σ is exponentially stable.

The condition on T means that $\mathbb{D}$ is J-coercive, by Lemma 8.2.3(a). E.g., for $J \gg 0$, we have $T \in \mathcal{GB}$ iff $\|\mathbb{D}u\|_2 \geq \varepsilon\|u\|^2$ ($u \in \mathrm{L}^2$) for some $\varepsilon > 0$.

There are two well-known methods for obtaining a direct formula (as at the end of Proposition 8.3.10) also in the unstable case. One is to first solve the finite-time problem on $[0,T]$ and then take a limit of $x_{\text{crit}}(x_0)$, $y_{\text{crit}}(x_0)$, $u_{\text{crit}}(x_0)$ and $\mathcal{P}x_0$ as $T \to +\infty$; we take a quick glance at this in Section 8.5.

The other method is to derive the corresponding Riccati equation and use it to obtain more information on the solution. This works well when B is bounded (in particular, in the discrete-time case) or when Σ is otherwise regular, but the classical results cannot be completely generalized to the general case (only to the extent of Section 9.7), hence we shall present partial results for different generalities in Chapter 9.

Proof of Lemma 8.3.10: This follows from Lemmas 8.2.8 and 8.2.9. To apply the latter, we must set $\widetilde{C} := \begin{bmatrix} C \\ 0 \end{bmatrix}$, $\widetilde{D} := \begin{bmatrix} D \\ 0 \end{bmatrix}$, and use substitutions $U \mapsto \mathrm{L}^2(\mathbf{R}_+;U)$ and $Z^{\text{s}} \mapsto Z^{\text{s}}_2$ in Remark 8.3.4, where Z^{s}_2 is the closure of

$$Z^{\text{s}}_0 := \{\mathbb{A}x_0 + \mathbb{B}\tau u \,\big|\, x_0 \in H,\ u \in \mathrm{L}^2(\mathbf{R}_+;U)\} \quad \text{w.r.t.} \quad \|x\|_{Z^{\text{s}}_2} := \max\{\|x\|_{\mathrm{L}^2_\beta}, \|Cx\|_2\} \tag{8.47}$$

(indeed, $Cx = y - Du \in \mathrm{L}^2$, where $x := \mathbb{A}x_0 + \mathbb{B}\tau u$ and $y := \mathbb{C}x_0 + \mathbb{D}u$, for all $x_0 \in H$ and $u \in \mathrm{L}^2(\mathbf{R}_+;U)$, hence $\|\cdot\|_{Z^{\text{s}}_2}$ is a norm on the vector space Z^{s}_0). It follows that $\widetilde{C}$ becomes continuous, $\mathcal{U}^*_*$ is unchanged, and the assumptions of Lemma 8.2.9 are satisfied. □

If Σ is SOS-stable and $\mathcal{U}^*_* = \mathcal{U}_{\text{out}}$ (or $\mathcal{U}^*_* \in \{\mathcal{U}_{\text{exp}}, \mathcal{U}_{\text{str}}, \mathcal{U}_{\text{sta}}\}$ and Σ has the corresponding stability), then the existence of a spectral factorization leads to the existence of a stable, optimal state feedback pair (the converse holds under J-coercivity, by Corollary 9.9.11):

Corollary 8.3.11 (SpF ⇒ J-critical) *Assume that* $\Sigma = \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline \mathbb{C} & \mathbb{D} \end{array}\right] \in \text{SOS}(U,H,Y)$ *and that* $\mathcal{U}^*_*(x_0) = \mathrm{L}^2(\mathbf{R}_+;U)$ *for all* $x_0 \in H$. *Assume, in addition, that the Popov operator* $\mathbb{D}^*J\mathbb{D}$ *has a spectral factorization* $\mathbb{X}^*S\mathbb{X}$. *Then its Toeplitz operator has the inverse* $T^{-1} = \mathbb{X}^{-1}\pi_+S^{-1}\mathbb{X}^{-*}$, *hence then Proposition 8.3.10 applies.*

In fact, then also (Crit1)–(Crit4) of Theorem 9.9.10 hold; in particular, (9.140) defines a stable, J-critical state feedback pair and (8.43)–(8.46) can be written as and (9.141)–(9.145). □

(This is obvious.)

We now compute the I-critical (minimizing) cost operator and control for the delay line system of Example 6.2.14:

Example 8.3.12 ($\mathcal{J}$ and $\mathcal{P}$) Take again $\Sigma := \left[\begin{array}{c|c} \pi_+\tau & \pi_{[0,1)}\tau(-1) \\ \hline \pi_+ & \tau(-1) \end{array}\right]$ (Σ is strongly stable) with $U = \mathbf{C} = Y$, $H := \mathrm{L}^2(\mathbf{R}_+;Y)$, and $J = I$. By Proposition 8.3.10, we have (note that $T=\pi_+$)

$$\mathcal{P}=\pi_+(I-\tau(1)\pi_+\tau(1)^*)\pi_+=\pi_{[0,1)}, \qquad \mathbb{K}_{\mathrm{crit}}=-\pi_+\tau(1), \quad (8.48)$$

$$\mathbb{A}_{\mathrm{crit}}(t)=\pi_+\tau-\pi_{[0,1)}\tau\pi_{[1,\infty)}=(\pi_+-\pi_{[1-t,1)})\tau, \qquad \mathbb{C}_{\mathrm{crit}}=\pi_{[0,1)}. \quad (8.49)$$

Thus, for $x_0\in H$, the control $u=-\pi_+\tau(1)x_0$ is the unique J-critical (and minimizing) control (over $\mathcal{U}_{\mathrm{out}}$). Now $\mathcal{J}(x_0,u)=\|x_0+\tau(-1)u\|_2^2$, and the J-critical cost is

$$\mathcal{J}(x_0,u_{\min}(x_0))=\|\pi_{[0,1)}x_0\|_2^2=\langle x_0,\mathcal{P}x_0\rangle. \quad (8.50)$$

Naturally, this is the minimal cost, because $\mathcal{J}(x_0,u)=\|x_0+\tau(-1)u\|_2^2$ and hence u can cancel x_0 on $[1,+\infty)$ only. $\triangleleft$

See Example 9.8.15 for the corresponding Riccati equation.

Note that the cost function $\mathcal{J}$, the J-critical control and state u_{crit} and y_{crit}, and the J-critical cost operator $\mathcal{P}$ depend on $\mathbb{C}$, $\mathbb{D}$ and J only, whereas x_{crit} depends on $\mathbb{A}$ and $\mathbb{B}$ too.

We sometimes need the following useful formula (with terms corresponding to $\pi_{[0,t)}J$ and $\pi_{[t,\infty)}J$):

$$\langle\mathbb{D}v,J\mathbb{D}u\rangle_{\mathrm{L}^2}=\langle\mathbb{D}^t v,J\mathbb{D}^t u\rangle_{\mathrm{L}^2}+\langle\mathbb{C}\mathbb{B}^t v+\mathbb{D}\pi_+\tau^t v,J(\mathbb{C}\mathbb{B}^t u+\mathbb{D}\pi_+\tau^t u)\rangle_{\mathrm{L}^2} \quad (8.51)$$

for all $t>0$, $u,v\in\mathrm{L}^2_{\mathrm{loc}}(\mathbf{R}_+;U)$ s.t. $\mathbb{D}u,\mathbb{D}v\in\mathrm{L}^2$; (use the fact that $\langle\mathbb{D}v,\pi_{[t,\infty)}J\mathbb{D}u\rangle=\langle\mathbb{D}(\pi_++\pi_-)\tau^t v,\pi_+J\mathbb{D}(\pi_++\pi_-)\tau^t u\rangle$). In particular,

$$\mathcal{J}(0,u)=\langle\mathbb{D}^t u,J\mathbb{D}^t u\rangle+\mathcal{J}(\mathbb{B}^t u,\pi_+\tau^t u)\quad (t>0,\ u\in\mathrm{L}^2_{\mathrm{loc}}(\mathbf{R}_+;U),\ \mathbb{D}u\in\mathrm{L}^2). \quad (8.52)$$

We now give two necessary and sufficient conditions for a unique J-critical control to be of state feedback form, i.e., for Σ_{crit} to be the left column of some closed-loop system of Σ:

Theorem 8.3.13 ($\Sigma_{\mathrm{crit}}=\Sigma_\circlearrowleft$) *Let there be a unique J-critical control over $\mathcal{U}_*^*(x_0)$ for each $x_0\in H$. Assume that $\vartheta=0$. Set $\mathbb{T}:=\mathbb{K}_{\mathrm{crit}}\mathbb{B}$. Let $0\le\gamma>\omega_A$. Then the following hold:*

(a) $\mathbb{T}\in\mathcal{B}(\mathrm{L}^2_\gamma(\mathbf{R};U))$, and, for each $v\in\mathrm{L}^2_\gamma(\mathbf{R}_-;U)$, $\mathbb{T}v$ is uniquely defined by the conditions $\mathbb{T}v\in\mathcal{U}_*^*(\mathbb{B}v)$ and $\langle\mathbb{D}(v+\mathbb{T}v),J\mathbb{D}\eta\rangle=0$ for all $\eta\in\mathcal{U}_*^*(0)$.

If $\mathcal{U}_*^*=\mathcal{U}_{\mathrm{out}}$, then $\mathbb{T}v$ is uniquely defined by $\mathbb{T}v,\pi_+\mathbb{D}(v+\mathbb{T}v)\in\mathrm{L}^2$ and $\langle\mathbb{D}(v+\mathbb{T}v),J\mathbb{D}\eta\rangle=0$ for all $\eta\in\mathcal{U}_{\mathrm{out}}(0)$; in particular, then $\mathbb{T}$ depends only on $\mathbb{D}$ and J.

If $\mathcal{U}_*^*=\mathcal{U}_{\mathrm{exp}}$ (resp. $\mathcal{U}_{\mathrm{sta}},\mathcal{U}_{\mathrm{str}}$), then $\mathbb{T}v$ is uniquely defined by $\mathbb{T}v,\pi_+\mathbb{D}(v+\mathbb{T}v)\in\mathrm{L}^2$, $\pi_+\mathbb{B}\tau(v+\mathbb{T}v)\in\mathrm{L}^2$ (resp. $\in\mathrm{L}^\infty$, $\in\mathcal{C}_0$) and $\langle\mathbb{D}(v+\mathbb{T}v),J\mathbb{D}\eta\rangle=0$ for all $\eta\in\mathcal{U}_*^*(0)$; in particular, then $\mathbb{T}$ depends only on $\mathbb{B}$, $\mathbb{D}$ and J.

(b1) Conditions (i)–(iii) are equivalent:

(i) There is an admissible state feedback pair $\begin{bmatrix}\mathbb{K} \mid \mathbb{F}\end{bmatrix}$ for Σ s.t. the corresponding closed-loop system $\Sigma_\circlearrowleft$ satisfies $\mathbb{K}_\circlearrowleft=\mathbb{K}_{\mathrm{crit}}$.

(ii) There is $\mathbb{X}\in\mathcal{G}\mathrm{TIC}_\infty(U)$ s.t. $-\mathbb{X}\mathbb{T}=\pi_+\mathbb{X}\pi_-\in\mathcal{B}(\mathrm{L}^2_\alpha(\mathbf{R}_-;U),\mathrm{L}^2_\beta(\mathbf{R}_+;U))$ for some $\alpha,\beta\in\mathbf{R}$.

(iii) There is $\mathbb{M} \in \mathcal{G}\mathrm{TIC}_\infty(U)$ *s.t.* $\mathbb{T}\mathbb{M} = \pi_+\mathbb{M}\pi_- \in \mathcal{B}(\mathrm{L}^2_\alpha(\mathbf{R}_-;U), \mathrm{L}^2_\beta(\mathbf{R}_+;U))$ *for some* $\alpha, \beta \in \mathbf{R}$.

(Naturally, it suffices to have the equality on $\mathrm{L}^2_c(\mathbf{R}_-;U)$ *on (ii) or (iii) if* α *and* β *are big enough to make both sides continuous.)*

(b2) If $\mathbb{X}$ *solves (ii) and* $\mathcal{U}^*_* \in \{\mathcal{U}_{\mathrm{out}}, \mathcal{U}_{\mathrm{str}}, \mathcal{U}_{\mathrm{sta}}, \mathcal{U}_{\mathrm{exp}}\}$*, then all solutions of (ii) are given by* $E\mathbb{X}$ *(*$E \in \mathcal{G}\mathcal{B}(U)$*).*

(c1) Assume (i). Then $\mathbb{A}_\circlearrowleft = \mathbb{A}_{\mathrm{crit}}$ *and* $\mathbb{C}_\circlearrowleft = \mathbb{C}_{\mathrm{crit}}$*. Moreover,* $\mathbb{X} := I - \mathbb{F}$*,* $\mathbb{M} := \mathbb{X}^{-1} = \mathbb{F}_\circlearrowleft + I$ *and* $\mathbb{N} := \mathbb{D}_\circlearrowleft := \mathbb{D}\mathbb{M}$ *satisfy (ii) and (iii) (for any* $\alpha, \beta > \max(0, \omega_A)$*),* $\left[\begin{smallmatrix}\mathbb{N}\\ \mathbb{M}\end{smallmatrix}\right] \in \mathrm{TIC}_\omega$ *for all* $\omega > 0$ *and* $\left[\begin{smallmatrix}\mathbb{N}\\ \mathbb{M}\end{smallmatrix}\right]\mathrm{L}^2_c \subset \mathrm{L}^2$*. Naturally,* $\mathbb{K}$*,* $\mathbb{F}$ *and* $\mathbb{X}$ *are* ω*-stable for any* $\omega > \omega_A$*. If* $\mathcal{U}^*_* = \mathcal{U}_{\mathrm{exp}}$*, then* $\Sigma_\circlearrowleft$ *is exponentially stable.*

(c2) Conversely, if (ii) holds, then $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right] := \left[\begin{array}{c|c} \mathbb{X}\mathbb{K}_{\mathrm{crit}} & I - \mathbb{X} \end{array}\right]$ *satisfies (i). Moreover,* $\mathbb{X}$ *satisfies (ii) iff* $\mathbb{M} := \mathbb{X}^{-1}$ *satisfies (iii).*

Assume, in addition, that $\mathcal{U}^*_* \in \{\mathcal{U}_{\mathrm{out}}, \mathcal{U}_{\mathrm{str}}, \mathcal{U}_{\mathrm{sta}}, \mathcal{U}_{\mathrm{exp}}\}$*. Then also (d)–(f) hold:*

(d) Assume (i). We have $\mathbb{K}_\circlearrowleft x_0 + \mathbb{M}u_\circlearrowleft \in \mathcal{U}^*_*(x_0)$ *for all* $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R}_+;U)$*. If* $\mathcal{U}^*_* = \mathcal{U}_{\mathrm{exp}}$*, then*

$$\mathcal{U}_{\mathrm{exp}}(x_0) = \{\mathbb{K}_\circlearrowleft x_0 + \mathbb{M}u_\flat \,\big|\, u_\circlearrowleft \in \mathrm{L}^2(\mathbf{R}_+;U)\} \quad (x_0 \in H). \tag{8.53}$$

(e) Assume (i). Then there is a unique $S = S^* \in \mathcal{B}(U)$ *s.t.* $\langle \mathbb{N}u_\circlearrowleft, J\mathbb{N}u_\circlearrowleft\rangle = \langle u_\circlearrowleft, Su_\circlearrowleft\rangle$ *for all* $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R}_+;U)$*. Moreover,* S *is one-to-one,* $\langle \mathbb{D}v, J\mathbb{D}\mathbb{M}u_\circlearrowleft\rangle = \langle \mathbb{M}^{-1}v, Su_\circlearrowleft\rangle$ *for all* $v \in \mathcal{U}^*_*(0)$ *and* $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R};U)$*, and*

$$\mathcal{J}(x_0, \mathbb{K}_\circlearrowleft x_0 + \mathbb{M}u_\circlearrowleft) = \langle x_0, \mathcal{P}x_0\rangle + \langle u_\circlearrowleft, Su_\circlearrowleft\rangle \quad (x_0 \in H,\ u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R}_+;U)). \tag{8.54}$$

(f) If $(\mathcal{P}, S, \mathbb{K}, \mathbb{X})$ *satisfies (i) and (e), then all such quadruples are given by* $(\mathcal{P}, E^{-*}SE^{-1}, E\mathbb{K}, E\mathbb{X})$*,* $E \in \mathcal{G}\mathcal{B}(U)$*.*

We shall show in Theorem 9.9.1(a1) that (i)–(iii) hold iff the Riccati equation (eIARE) for Σ and J has a "$\mathcal{U}^*_*$-stabilizing" solution (and that solution is given by $(\mathcal{P}, S, \left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right])$; cf. (f)). As noted above, this is always the case when Σ is sufficiently regular. Conditions (i)–(iii) are treated from another point of view in Section 9.14.

Proof: (a) Both $\mathbb{K}_{\mathrm{crit}}$ and $\mathbb{B}$ are γ-stable, hence so is $\mathbb{T}$. Set $x_0 := \mathbb{B}v$, $u := \mathbb{T}v = \mathbb{K}_{\mathrm{crit}}x_0$, $y := \mathbb{C}x_0 + \mathbb{D}u$. Then $u_{\mathrm{crit}}(x_0)$ is the unique $u \in \mathcal{U}^*_*(x_0)$ satisfying $\langle y, J\mathbb{D}\eta\rangle = 0$ for all $\eta \in \mathcal{U}^*_*(0)$, and $y = \mathbb{C}x_0 + \mathbb{D}u = \mathbb{C}\mathbb{B}v + \mathbb{D}\mathbb{T}v = \pi_+\mathbb{D}(\pi_- v + \mathbb{T}v)$.

By definition, $u \in \mathcal{U}_{\mathrm{out}}(\mathbb{B}v)$ iff $u \in \mathrm{L}^2$ and $y \in \mathrm{L}^2$; for $\mathcal{U}_{\mathrm{exp}}$ we have the extra condition that $\pi_+(\mathbb{A}\mathbb{B}v + \mathbb{B}\tau\mathbb{T}v) = \pi_+\mathbb{B}\tau(v + \mathbb{T}v) \in \mathrm{L}^2(\mathbf{R}_+;U)$.

(b1)&(c2) If (i) holds, then $\mathbb{T}\mathbb{M} = \mathbb{K}_\circlearrowleft\mathbb{B}_\circlearrowleft = \pi_+\mathbb{M}\pi_-$ in $\mathcal{B}(\mathrm{L}^2_\alpha(\mathbf{R};U))$ for any $\alpha > \max(\omega_A, 0)$ (note that $\mathbb{B}$, $\mathbb{M}$ and $\mathbb{K}_\circlearrowleft$ are α-stable) hence then (iii) holds.

If $\mathbb{X} = \mathbb{M}^{-1} \in \mathcal{G}\mathrm{TIC}_\infty(U)$, then $\pi_+\mathbb{X}\pi_-\mathbb{M}\pi_- = \pi_+ I\pi_- - \pi_+\mathbb{X}\pi_+\mathbb{M}\pi_- = -\pi_+\mathbb{X}\pi_+\mathbb{M}\pi_-$. Therefore, (iii) implies (ii).

Assume (ii). Let $\omega > \max\{\omega_A, \alpha, \beta\}$. Then $\mathrm{L}^2_\omega(\mathbf{R}_-;U) \subset \mathrm{L}^2_\alpha(\mathbf{R}_-;U)$ and $\mathbb{X}, \mathbb{T}, \pi_+\mathbb{X}\pi_- \in \mathcal{B}(\mathrm{L}^2_\omega)$, hence then $-\mathbb{X}\mathbb{T} = \pi_+\mathbb{X}\pi_- \in \mathcal{B}(\mathrm{L}^2_\omega)$. Set $\mathbb{F} := I - \mathbb{X}$ and $\mathbb{K} := \mathbb{X}\mathbb{K}_{\mathrm{crit}}$. Then $\mathbb{K}\mathbb{B} = \pi_+\mathbb{F}\pi_-$, and

$$\mathbb{K}\mathbb{A}^t = \mathbb{X}\mathbb{K}_{\mathrm{crit}}\mathbb{A}^t = \mathbb{X}\mathbb{K}_{\mathrm{crit}}(\mathbb{A}^t_{\mathrm{crit}} - \mathbb{B}^t\mathbb{K}_{\mathrm{crit}}) = \mathbb{X}\pi_+\tau^t\mathbb{K}_{\mathrm{crit}} - \mathbb{X}\mathbb{K}_{\mathrm{crit}}\mathbb{B}\tau^t\mathbb{K}_{\mathrm{crit}} \tag{8.55}$$

$$= \pi_+\mathbb{X}\pi_+\tau^t\mathbb{K}_{\mathrm{crit}} - \pi_+\mathbb{X}\pi_-\tau^t\mathbb{K}_{\mathrm{crit}} = \pi_+\tau^t\mathbb{X}\mathbb{K}_{\mathrm{crit}} = \pi_+\tau^t\mathbb{K}. \tag{8.56}$$

Therefore, $\left[\begin{smallmatrix}\mathbb{A} & \mathbb{B}\\ \mathbb{K} & \mathbb{F}\end{smallmatrix}\right] \in \mathrm{WPLS}_\omega(U,H,U)$, hence $\Sigma_{\mathrm{ext}} \in \mathrm{WPLS}_{\omega'}(U,H,Y\times U)$ for any $\omega' > \omega_A$. Obviously, Σ_{crit} is the left column of the corresponding closed-loop system $\Sigma_\circlearrowleft$, since $\mathbb{K}_\circlearrowleft = \mathbb{M}\mathbb{K} = \mathbb{K}_{\mathrm{crit}}$.

(b2) This follows from (f).

(c1) Now $\mathbb{C}_\circlearrowleft = \mathbb{C} + \mathbb{D}\mathbb{K}_\circlearrowleft = \mathbb{C}_{\mathrm{crit}}$, $\mathbb{A}_\circlearrowleft = \mathbb{A} + \mathbb{B}^t\mathbb{K}_\circlearrowleft = \mathbb{A}_{\mathrm{crit}}$. The formulae for (ii) and (iii) were shown above. Because $\mathbb{C}_\circlearrowleft$ and $\mathbb{K}_\circlearrowleft$ are stable, the maps $\mathbb{N}$ and $\mathbb{M}$ are as above, by Lemma 6.1.11. The ω_A claim follows from Lemma 6.1.10, and the final claim from Theorem 8.3.9(a2).

(d) Now $u := \mathbb{K}_\circlearrowleft x_0 + \mathbb{M}u_\circlearrowleft \in \mathrm{L}^2$ and $y := \mathbb{C}x_0 + \mathbb{D}u = \mathbb{C}_\circlearrowleft x_0 + \mathbb{N}u_\circlearrowleft \in \mathrm{L}^2$, by (c1). If $\mathcal{U}^*_* = \mathcal{U}_{\mathrm{exp}}$, then also $x := \mathbb{A}x_0 + \mathbb{B}^t u = \mathbb{A}_\circlearrowleft x_0 + \mathbb{B}^t_\circlearrowleft u_\circlearrowleft \in \mathrm{L}^2$, by Lemma 6.1.10. Analogously, if $\mathcal{U}^*_* = \mathcal{U}_{\mathrm{sta}}$ (resp. $= \mathcal{U}_{\mathrm{str}}$) and $u_\circlearrowleft \in \mathrm{L}^2([0,T);U)$, then $\mathbb{B}_\circlearrowleft\tau^{T+t}u_\circlearrowleft = \mathbb{A}^t_\circlearrowleft\mathbb{B}_\circlearrowleft\tau^T u_\circlearrowleft$ is bounded (resp. goes to zero), as $t \to +\infty$, since $\mathbb{A}_\circlearrowleft$ is stable (resp. strongly stable). Thus, $u \in \mathcal{U}^*_*(x_0)$. Formula (8.53) will be proved in Theorem 8.4.5(e).

(e) Let $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R}_+;U)$, $v \in \mathrm{L}^2_c(\mathbf{R}_-;U)$. Then $u := \mathbb{M}u_\circlearrowleft \in \mathcal{U}^*_*(0)$, by (d), hence then $\langle \mathbb{N}\pi_- v, J\mathbb{N}u_\circlearrowleft\rangle = \langle \mathbb{C}_{\mathrm{crit}}\mathbb{B}v, J\mathbb{D}u\rangle = 0$. Consequently, we obtain S from Lemma 2.3.1. By (8.35), we have

$$\mathcal{J}(x_0, \mathbb{K}_\circlearrowleft x_0 + \mathbb{M}u_\circlearrowleft) = \langle x_0, \mathcal{P}x_0\rangle + \mathcal{J}(0, \mathbb{M}u_\circlearrowleft) = \langle x_0, \mathcal{P}x_0\rangle + \langle u_\circlearrowleft, Su_\circlearrowleft\rangle \tag{8.57}$$

for all $x_0 \in H$, $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R}_+;U)$.

Let $v \in \mathcal{U}^*_*(0)$ and $u_\circlearrowleft \in \mathrm{L}^2_c(\mathbf{R};U)$. Because $\langle \mathbb{D}v, J\mathbb{N}\pi_- u_\circlearrowleft\rangle = \langle \mathbb{D}v, J\mathbb{C}_{\mathrm{crit}}\mathbb{B}_\circlearrowleft u_\circlearrowleft\rangle = 0$, we may assume that $u_\circlearrowleft = \pi_+ u_\circlearrowleft$. Choose $T > 0$ s.t. $u_\circlearrowleft = \pi_{[0,T)}u_\circlearrowleft$. Set $v_\circlearrowleft := \mathbb{M}^{-1}v$, $v_1 := \mathbb{M}\pi_{[0,T)}v_\circlearrowleft \in \mathcal{U}^*_*(0)$, $v_2 := v - v_1 \in \mathcal{U}^*_*(0)$. Then $\pi_{[0,T)}v_2 = \pi_{[0,T)}\mathbb{M}\pi_{[T,\infty)}v_\circlearrowleft = 0$, hence

$$\langle \mathbb{D}v, J\mathbb{N}u_\circlearrowleft\rangle = \langle \pi_{[0,T)}v_\circlearrowleft, Su_\circlearrowleft\rangle + \langle \mathbb{D}v_2, J\mathbb{C}_{\mathrm{crit}}\mathbb{B}_\circlearrowleft u_\circlearrowleft\rangle = \langle \pi_{[0,T)}v_\circlearrowleft, Su_\circlearrowleft\rangle = \langle v_\circlearrowleft, Su_\circlearrowleft\rangle. \tag{8.58}$$

By taking $v_\circlearrowleft := \chi_{[0,1)}v_0$, $u_\circlearrowleft := \chi_{[0,1)}u_0$ for arbitrary $u_0, v_0 \in U$, we see that S is unique.

Let $u_0 \in U \setminus \{0\}$ be arbitrary. Set $u_\circlearrowleft := \chi_{[0,1)}u_0$, $u := \mathbb{M}u_\circlearrowleft \in \mathcal{U}^*_*(0) \setminus \mathcal{U}^{*,\mathrm{crit}}_*(0)$ (recall that $\mathcal{U}^{*,\mathrm{crit}}_*(0) = \{0\}$, by uniqueness), hence then $\langle \mathbb{D}v, J\mathbb{D}u\rangle \neq 0$ for some $v \in \mathcal{U}^*_*(0)$. Thus, $\langle \mathbb{M}^{-1}v, Su_\circlearrowleft\rangle \neq 0$. In particular, $Su_0 \neq 0$. Therefore, S is one-to-one.

(f) Let $(\mathcal{P}, \widetilde{S}, \widetilde{\mathbb{K}}, \widetilde{\mathbb{X}})$ also satisfy (i) and (e). Claim (e) and the proof of Lemma 9.10.1(c2) shows that (9.160) is satisfied by both $\mathbb{X}$ and $\widetilde{\mathbb{X}}$, hence

$\widetilde{\mathbb{X}} = E\mathbb{X}$ and $\widetilde{S} = E^{-*}SE^{-1}$ for some $E \in \mathcal{GB}(U)$, by Lemma 2.3.5 (recall that S is one-to-one). By (c2), $\widetilde{\mathbb{K}} = \widetilde{\mathbb{X}}\mathbb{K}_{\mathrm{crit}} = E\mathbb{X}\mathbb{K}_{\mathrm{crit}} = E\mathbb{K}$. □

The above assumption on uniqueness in the theorem is mostly superfluous for a control corresponding to a state feedback pair:

Lemma 8.3.14 *Let* $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$ *be a J-critical state feedback pair over* $\mathcal{U}_*^*$ *for* Σ *and J. Set* $\mathbb{T} := \mathbb{K}_\circlearrowleft \mathbb{B}$, $\Sigma_{\mathrm{crit}} := \Sigma_\circlearrowleft$.

Then (i)–(iii) and (c1)–(e) of Theorem 8.3.13 hold except that S need not one-to-one in (e). □

(The same proof applies.) In particular, then $\mathbb{A}_\circlearrowleft = \mathbb{A}_{\mathrm{crit}}$ and $\mathbb{C}_\circlearrowleft = \mathbb{C}_{\mathrm{crit}}$, so that $\Sigma_\circlearrowleft$ produces J-critical state, control and output (for zero input).

In fact, most of the rest of Theorem 8.3.13 holds in this more general case too, but we shall return to this in Chapter 9. See Theorem 9.9.1(a1)&(e2)&(f2)–(h) for details.

If $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$ is a J-critical state feedback pair over $\mathcal{U}_*^*$, then the left column of the corresponding closed-loop system is just like the one in (8.34) (except that it need not be unique). We shall call such a column a J-critical control in WPLS form in order to be able to treat both cases simultaneously:

Definition 8.3.15 (J-critical control in WPLS form) *We call the control* $x_0 \mapsto \mathbb{K}_{\mathrm{crit}} x_0$ *(and* Σ_{crit}*) a* control for Σ in WPLS form *if* $\mathbb{K}_{\mathrm{crit}} : H \to \mathrm{L}^2_{\mathrm{loc}}(\mathbf{R}_+;U)$ *is s.t.*

$$\Sigma_{\mathrm{crit}} := \left[\begin{array}{c|c} \mathbb{A}_{\mathrm{crit}} & \\ \hline \mathbb{C}_{\mathrm{crit}} & \\ \mathbb{K}_{\mathrm{crit}} & \end{array}\right] := \left[\begin{array}{c|c} \mathbb{A}+\mathbb{B}\mathbb{K}_{\mathrm{crit}} & \\ \hline \mathbb{C}+\mathbb{D}\mathbb{K}_{\mathrm{crit}} & \\ \mathbb{K}_{\mathrm{crit}} & \end{array}\right] \in \mathrm{WPLS}(\{0\}, H, Y \times U). \tag{8.59}$$

If $\mathbb{K}_{\mathrm{crit}} x_0$ *is J-critical for each* x_0*, then we say that* $\mathbb{K}_{\mathrm{crit}}$ *(or* Σ_{crit}*) is a* J-critical control in WPLS form *and that* u_{crit} *can be given in WPLS form, where* $\begin{bmatrix} x_{\mathrm{crit}}(x_0) \\ y_{\mathrm{crit}}(x_0) \\ u_{\mathrm{crit}}(x_0) \end{bmatrix} := \Sigma_{\mathrm{crit}} x_0$.

If $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$ *is an admissible state feedback pair for Σ with closed-loop system* $\Sigma_\flat$*, then we call* $\mathbb{K}_\circlearrowleft := (I-\mathbb{F})^{-1}\mathbb{K}$ *(an the left column of* $\Sigma_\flat$*) a* control in state feedback form.

As explained below Theorem 8.3.9, a control in WPLS form need not be of state feedback form unless, e.g., B is bounded, as shown in Lemma 8.3.18.

We start with some rather obvious facts:

Lemma 8.3.16

(a1) A unique J-critical control can always be given in WPLS form.

(a2) If there is a J-critical state feedback pair, then the corresponding J-critical control can be given in WPLS form.

(b) A control ("Σ_{crit}") in WPLS form has the properties described in Remark 9.7.7, Theorem 8.3.9, and Theorem 8.3.13(b1)&(c1)&(c2)&(d).

(c) A control in state feedback form is in WPLS form.

Thus, we need not assume a J-critical control to be unique in Theorem 8.3.9 as long as Σ_{crit} is a J-critical control for Σ in WPLS form. Therefore, we can and will use the latter (weaker) assumption in several results below to treat both cases simultaneously.

Proof: (a1) This is contained in Theorem 8.3.9.

(a2) In fact, given any admissible state feedback pair $\left[\ \mathbb{K}\ \middle|\ \mathbb{F}\ \right]$ for Σ, then the control $\mathbb{K}_{\circlearrowleft} := (I-\mathbb{F})^{-1}\mathbb{K}$ can be given in state feedback form (let Σ_{crit} be the left column of the corresponding closed-loop system).

(b) This is contained in Remark 9.7.7, Theorem 8.3.9, and the proof of Theorem 8.3.13(b1)&(c1)&(c2)&(d) (note that in (b1) we only use the fact that $\mathbb{K}_{\text{crit}}\mathbb{B}\mathbb{M} = \pi_+\mathbb{M}\pi_-$ (or $-\mathbb{X}\mathbb{K}_{\text{crit}}\mathbb{B} = \pi_+\mathbb{X}\pi_-$)).

(c) This is obvious. □

The generators of a control in WPLS form are analogous to those of a state feedback system:

Lemma 8.3.17 *Let Σ_{crit} be a control in WPLS form.*

(a) We have $A_{\text{crit}} = A + BK_{\text{crit}}$ and $C_{\text{crit}} = C_c + D_c K_{\text{crit}}$ on $\text{Dom}(A_{\text{crit}})$, hence $\text{Dom}(A_{\text{crit}}) \subset H_B$, where $\left[\ A_{\text{crit}}^{\text{T}}\ \middle|\ C_{\text{crit}}^{\text{T}}\quad K_{\text{crit}}^{\text{T}}\ \right]^{\text{T}}$ are the generators of Σ_{crit}, and (C_c, D_c) is any compatible pair for Σ.

(b) Let K_c be a compatible admissible state feedback operator for Σ, and let $\Sigma_\flat$ be the corresponding closed-loop system.

Then $K_c = K_{\text{crit}}$ on $\text{Dom}(A_{\text{crit}})$ iff $\mathbb{K}_\flat = \mathbb{K}_{\text{crit}}$. If $\mathbb{K}_\flat = \mathbb{K}_{\text{crit}}$, then $\mathbb{A}_\flat = \mathbb{A}_{\text{crit}}$, $\mathbb{C}_\flat = \mathbb{C}_{\text{crit}}$, and K_c is the unique compatible operator having $\mathbb{K}_\flat = \mathbb{K}_{\text{crit}}$.

Proof: (a) 1° As in the proof of Proposition 6.6 of [W94b], we take the Laplace transform of the equation $\mathbb{A}_{\text{crit}}^t x_0 - \mathbb{A}^t x_0 = \mathbb{B}\tau^t\mathbb{K}_{\text{crit}}x_0$ ($x_0 \in H$, $t \geq 0$) to obtain

$$(s-A_{\text{crit}})^{-1} - (s-A)^{-1} = (s-A)^{-1}BK_{\text{crit}}(s-A_{\text{crit}})^{-1} \in \mathcal{B}(H). \qquad (8.60)$$

Multiply this by $(s-A_{\text{crit}}) \in \mathcal{B}(A_{\text{crit}}, H)$ to the right and by $s-A \in \mathcal{B}(H, H_{-1})$ to the left to obtain

$$A_{\text{crit}} - A = (s-A) - (s-A_{\text{crit}}) = BK_{\text{crit}} \in \mathcal{B}(\text{Dom}(A_{\text{crit}}), H_{-1}). \qquad (8.61)$$

(N.B. by duality, $A^* = A_{\text{crit}}^* - K_{\text{crit}}^* B^*$ on $\text{Dom}(A^*)$.) If $x_0 \in \text{Dom}(A_{\text{crit}})$, then $Ax_0 + BK_{\text{crit}}x_0 \in H$, hence then $x_0 \in H_B$ (see Definition 6.1.17).

2° By Laplace transforming the equation $\mathbb{C}_{\text{crit}} = \mathbb{C} + \mathbb{D}\mathbb{K}_{\text{crit}}$ and using Theorem 6.2.11(c1), Lemma 6.3.10(a) and (8.60), we obtain that

$$\widehat{\mathbb{C}_{\text{crit}}}(s) = C_{\text{crit}}(s-A_{\text{crit}})^{-1} = \widehat{\mathbb{C}}(s) + \widehat{\mathbb{D}}(s)\widehat{\mathbb{K}_{\text{crit}}}(s) \qquad (8.62)$$

$$= C(s-A)^{-1} + D_c K_{\text{crit}}(s-A_{\text{crit}})^{-1} C_c (s-A)^{-1} BK_{\text{crit}}(s-A_{\text{crit}})^{-1} \qquad (8.63)$$

$$= (C_c + D_c K_{\text{crit}})(s-A_{\text{crit}})^{-1} \in \mathcal{B}(H, Y) \qquad (8.64)$$

for $s > \max\{\omega_A, \omega_{A_{\rm crit}}\}$. Because $s - A_{\rm crit}$ maps H onto $\mathrm{Dom}(A_{\rm crit})$, we must have $C_{\rm crit} = C_{\rm c} + D_{\rm c}K_{\rm crit}$.

(b) 1° If $\mathbb{K}_\flat = \mathbb{K}_{\rm crit}$, then $K_{\rm c} = K_{\rm crit}$ on $\mathrm{Dom}(A_{\rm crit})$, by Proposition 6.6.18(d2), because $\mathrm{Dom}(A_{\rm crit}) \subset H_B$, by (c).

2° Assume that $K_{\rm c} = K_{\rm crit}$ on $\mathrm{Dom}(A_{\rm crit})$. By Proposition 6.6.18(d2), $A_\flat = A + BK_{\rm c}$ and $K_\flat = K_{\rm c}$ on H_B and $\mathrm{Dom}(A_\flat) \subset H_B$. But $\mathrm{Dom}(A_{\rm crit}) \subset H_B$ and $A_{\rm crit} = A + BK_{\rm c}$, by (c), hence $A_\flat = A_{\rm crit}$ on $\mathrm{Dom}(A_{\rm crit})$. Choose $\omega > \max\{\omega_{A_\flat}, \omega_{A_{\rm crit}}\}$. Then

$$\mathrm{Dom}(A_{\rm crit}) = \{x_0 \in H_B \,\big|\, (\omega - A_{\rm crit})x_0 \in H\} \subset \{x_0 \in H \,\big|\, (\omega - A_\flat)x_0 \in H\} = \mathrm{Dom}(A_\flat), \tag{8.65}$$

hence $\mathbb{A}_{\rm crit} = \mathbb{A}_\flat$, by Lemma A.4.2(i). Consequently, $\mathbb{K}_\flat = K_{\rm c}\mathbb{A}_\flat = K_{\rm crit}\mathbb{A}_{\rm crit} = \mathbb{K}_{\rm crit}$ on $\mathrm{Dom}(A_{\rm crit})$, hence on H, by density. In particular, $K_{\rm c}$ is J-critical. It follows that $\mathbb{C}_\flat = \mathbb{C} + \mathbb{D}\mathbb{K}_\flat = \mathbb{C}_{\rm crit}$.

3° *Uniqueness:* If also $K_{\rm c}'$ is compatible and admissible and $\mathbb{K}_\flat' = \mathbb{K}_{\rm crit}$ (i.e., $K_{\rm c}' = K_{\rm crit}$ on $\mathrm{Dom}(A_{\rm crit})$), then $K_{\rm c}' = K_{\rm c}$ on H_B (i.e., $\left[\begin{array}{c|c} \mathbb{K}' & \mathbb{F}' \end{array}\right] = \left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$), by Proposition 6.6.18(g), i.e., $K_{\rm c}' = K_{\rm c}$. (Recall that we consider $K_{\rm c}$ and $K_{\rm c}'$ equal when $K_{\rm c}' = K_{\rm c}$ on H_B, since the values of $K_{\rm c}$ outside H_B do not affect $\Sigma_{\rm ext}$, nor $\Sigma_\flat$.) □

As stated above, when B is bounded, a control in WPLS form corresponds to an ULR state feedback operator:

Lemma 8.3.18 *Let $\Sigma_{\rm crit}$ be a control in WPLS form and let B be bounded.*

Then $\Sigma_{\rm crit}$ is of state feedback form.

Moreover, $\mathrm{Dom}(A_{\rm crit}) = H_B = \mathrm{Dom}(A)$, *$K_{\rm crit}$ is an ULR state feedback operator for Σ, and $\Sigma_{\rm crit}$ is the left column of the corresponding closed-loop system $\Sigma_\circlearrowleft$.*

Since, in discrete time, B is always bounded, any control in "wpls form" is necessarily induced by state feedback to the original system (the proof below applies mutatis mutandis; in fact, the discrete-time form of the proof is contained at the beginning of the proof of Theorem 14.1.6).

Proof: (Note that $\Sigma_{\rm crit}$ is not assumed to be J-critical.)

Let $A_{\rm crit}, C_{\rm crit}, K_{\rm crit}$ be the generators of $\Sigma_{\rm crit}$. By Lemma 6.3.16(b), $\left[\begin{array}{c|c} A_{\rm crit} & B \\ \hline K_{\rm crit} & 0 \end{array}\right]$ generate an ULR WPLS $\left[\begin{array}{c|c} \mathbb{A}_{\rm crit} & \mathbb{B}_{\rm crit} \\ \hline \mathbb{K}_{\rm crit} & \mathbb{F}_{\rm crit} \end{array}\right]$. Because $F_{\rm crit} = 0$ and $\mathbb{F}_{\rm crit} \in \mathrm{ULR}$, we have $I + \mathbb{F}_{\rm crit} \in \mathcal{G}\mathrm{TIC}_\infty$, by Proposition 6.3.1(c). Therefore, $\left[\begin{array}{c|c} -\mathbb{K}_{\rm crit} & -\mathbb{F}_{\rm crit} \end{array}\right]$ is an admissible state feedback pair for $\left[\begin{array}{c|c} \mathbb{A}_{\rm crit} & \mathbb{B}_{\rm crit} \end{array}\right]$,

By (6.145), the corresponding closed-loop system $\Sigma_\flat$ is generated by $\left[\begin{array}{c|c} A & B \\ \hline -(K_{\rm crit})_{\rm w} & 0 \end{array}\right]$, hence of form $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline -\mathbb{K} & -\mathbb{F} \end{array}\right]$ for some $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$, where $\mathbb{F} = I - (I + \mathbb{F}_{\rm crit})^{-1}$. Indeed, $H_{B_{\rm crit}} := \mathrm{Dom}(A_{\rm crit}) + (s - A_{\rm crit})^{-1}BU = \mathrm{Dom}(A_{\rm crit})$ with equivalent norms, by Lemma A.3.6. But

$$H_B = \mathrm{Dom}(A) = \{x_0 \in H \,\big|\, Ax_0 \in H\} = \mathrm{Dom}(A_{\rm crit}) = H_{B_{\rm crit}}, \tag{8.66}$$

(by the formula $A_{\rm crit} = A + BK_{\rm crit}$ from Lemma 8.3.17(a)) with equivalent norms, by Lemma A.3.6, hence $A_\flat := A_{\rm crit} - BK_{\rm crit} = A$ (with same domains).

Consequently, $K = (K_{\text{crit}})_{w|_{\text{Dom}(A)}} = K_{\text{crit}}$ is an admissible ULR state feedback operator for Σ, and Σ_{crit} is the left column of the corresponding closed-loop system (due to same generators). □

As noted in [FLT], optimizability is almost equivalent to exponential stabilizability:

Proposition 8.3.19 *Let Σ be optimizable. Then there is an exponentially stable WPLS of form (8.59).*

The difference is that the "exponentially stabilizing state feedback" of Theorem 8.3.9 need not be well-posed in general; see Theorem 9.2.12 for sufficient conditions.

Proof: Define $\Sigma_{\text{ext}} := \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline \mathbb{C}_{\text{ext}} & \mathbb{D}_{\text{ext}} \end{array}\right]$ by

$$\mathbb{C}_{\text{ext}} := \begin{bmatrix} \mathbb{C} \\ \mathbb{A} \\ 0 \end{bmatrix}, \quad \mathbb{D}_{\text{ext}} := \begin{bmatrix} \mathbb{D} \\ \mathbb{B} \\ I \end{bmatrix}, \quad J = I \tag{8.67}$$

to have $\mathcal{J}_{\text{ext}}(x_0,u) = \|u\|_2^2 + \|x\|_2^2 + \|y\|_2^2$, where $x := \mathbb{A}x_0 + \mathbb{B}\tau u$, $y := \mathbb{C}x_0 + \mathbb{D}u$.

Because Σ_{ext} is J-coercive over $\mathcal{U}_{\text{exp}}$, we obtain an exponentially stable system $(\Sigma_{\text{ext}})_{\text{crit}} = \left[\begin{array}{c} \mathbb{A}_{\text{crit}} \\ \hline \mathbb{C}_{\text{ext}} + \mathbb{D}_{\text{ext}}\mathbb{K}_{\text{crit}} \\ \mathbb{K}_{\text{crit}} \end{array}\right] \in \text{WPLS}(U,H,Y \times H \times U)$ from Theorems 8.4.3 and 8.3.9; and $\mathbb{C}_{\text{ext}} + \mathbb{D}_{\text{ext}}K_{\text{crit}} = \begin{bmatrix} \mathbb{C} + \mathbb{D}\mathbb{K}_{\text{crit}} \\ \mathbb{A}_{\text{crit}} \\ \mathbb{K}_{\text{crit}} \end{bmatrix}$, so that we only have to drop the second and third row from $(\Sigma_{\text{ext}})_{\text{crit}}$ to obtain an exponentially stable WPLS of form (8.59). □

Now we can give a postponed proof:

Lemma 8.3.20 *Theorem 6.7.7 holds.*

Proof: Let $\Sigma_\sharp := \left[\begin{array}{c|cc} \mathbb{A}_\sharp & \mathbb{B}_\sharp & \mathbb{H}_\sharp \end{array}\right]$ be the exponentially stable system of Proposition 8.3.19 for Σ^{d} (which is optimizable), and set $M := \|\mathbb{A}_\sharp\|_{\mathcal{B}(H,L^2)} + \|\mathbb{B}_\sharp\tau\|_{\text{TIC}} + \|\mathbb{H}_\sharp\tau\|_{\text{TIC}} < \infty$. One easily verifies that $x := \mathbb{A}x_0 + \mathbb{B}\tau u = \mathbb{A}_\sharp x_0 + \mathbb{B}_\sharp\tau u - \mathbb{H}_\sharp\tau y$. Since $\Sigma_\sharp$ is strongly stable, we have that $x \in \mathcal{C}_0(\mathbf{R}_+;H)$. □

Notes

Much of Lemmas 8.3.5–8.3.8 and Theorem 8.3.9 has been used for decades in some special cases; see [Zwart] for the standard unstable LQR (minimization) problem and [S98c] for the stable indefinite setting of Proposition 8.3.10 and Corollary 8.3.11; both articles only treat J-coercive WPLSs and only the latter treats the closed-loop system.

Section 3 of [S98c] essentially contains Proposition 8.3.10 except for its last paragraph, whose results resemble what is a key formula in several articles by Irena Lasiecka, Roberto Triggiani and others. They first use such a stable case result (a special case of Lemma 8.2.9, as in Section 8.5) to solve the finite-time LQR problem and then let the length of the time interval approach infinity to obtain the solution for the infinite-time problem as the limit of the finite-time solution (they use a very coercive cost function to guarantee the convergence). Their results in [FLT] include Proposition 8.3.19. See also [LT00a]–[LT00b] (also a third part of the trilogy is supposed to appear). This finite-time method has been used by several authors at least since seventies.

All the articles mentioned above optimize over $\mathcal{U}_{\text{out}}$ ($= \mathcal{U}_{\text{sta}}$ in [S97b] and [S98c], and $= \mathcal{U}_{\text{exp}}$ in much of [FLT] and [LT00a]–[LT00b], by Lemma 8.3.3), and so does, e.g., [CZ]. Nevertheless, $\mathcal{U}_{\text{exp}}$ is the most commonly used class for finite-dimensional systems (see, e.g., [IOW], [LR], [GL]) and possibly also for infinite-dimensional systems (see, e.g., [Keu], [Pandolfi] and [WR00]).

For finite-dimensional systems, strong stability is equivalent to exponential stability, but in the infinite-dimensional case the requirement of strong stability ($\mathcal{U}_{\text{str}}$) has often been used since the seventies; see, e.g., [Slemrod] and [Balakrishnan]. This case has been studied for WPLSs with bounded input and output operators (B and C) by Ruth Curtain and Job Oostveen in several articles (see [OC98]), and the monograph [Oostveen] contains a rather mature theory, further historical remarks on this case and examples where $\mathcal{U}_{\text{str}}$ is the most natural choice for $\mathcal{U}_*^*$. See p. 501 for a comparison of $\mathcal{U}_{\text{exp}}$, $\mathcal{U}_{\text{str}}$, $\mathcal{U}_{\text{sta}}$ and $\mathcal{U}_{\text{out}}$.

We give several sufficient conditions of a unique J-critical control to be of (regular) state feedback form in Remark 9.9.14 and in the notes to Section 9.9.

8.4 J-coercivity and factorizations

A person who is wise does nothing against their will, nothing with sighing or under coercion.

— Marcus Tullius Cicero (106 B.C. – 43 B.C.)

In this section, we apply J-coercivity to WPLSs and explore its connection to optimal (J-critical) control and spectral and inner coprime factorizations. This concept generalizes several general nonsingularity assumptions of control problems.

We shall list several equivalent conditions for positive J-coercivity in Section 10.3, such as the popular "no transmission zeros" ($\mathcal{U}_{\mathrm{out}}$) and "no invariant zeros" ($\mathcal{U}_{\mathrm{exp}}$) conditions. Most of these equivalent conditions have been used in classical minimization problems; we also show there that several other classical minimization assumptions are stronger than positive J-coercivity.

In the stable case, J-coercivity is equivalent to the condition that the *Popov Toeplitz operator* $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible, by Lemma 8.4.11(a1). The general definition below requires that "$\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$" is invertible on $\mathcal{U}_*^*(0)$, by Lemma 8.2.3(c1), hence also the general condition can be considered as a Popov Toeplitz invertibility condition.

General J-coercivity with the minimal stabilizability assumption $\mathcal{U}_*^*(x_0) \neq 0$ $(x_0 \in H)$ will be shown to be a sufficient condition for the existence of an optimal (i.e., J-critical) control and for the existence of a unique "stabilizing" solution of the Riccati equation (under sufficient regularity); in fact, these three are equivalent in some cases (see, e.g., Theorem 9.2.16).

At the end of this section, we shall show that J-coercivity over $\mathcal{U}_{\mathrm{out}}$ is implied by the existence of an inner coprime factorization (spectral factorization in the stable case) of the I/O map, with equivalence under sufficient regularity and stabilizability assumptions.

We also show that optimization over $\mathcal{U}_{\mathrm{exp}}$ can be reduced to the stable case, whereas for $\mathcal{U}_{\mathrm{out}}$, $\mathcal{U}_{\mathrm{sta}}$ and $\mathcal{U}_{\mathrm{str}}$, we need quasi-coprimeness for an analogous reduction.

In accordance to Definition 8.2.1, we generalize J-coercivity as follows:

Definition 8.4.1 (J-coercive) *We call* $\mathbb{D}$ J-coercive (over $\mathcal{U}_*^*$) *if there is* $\varepsilon > 0$ *s.t. for each nonzero* $u \in \mathcal{U}_*^*(0)$ *there is a nonzero* $v \in \mathcal{U}_*^*(0)$ *s.t.*

$$\langle \mathbb{D}v, J\mathbb{D}u \rangle \geq \varepsilon \|u\|_{\mathcal{U}_*^*} \|v\|_{\mathcal{U}_*^*}. \tag{8.68}$$

If, in addition, $\langle \mathbb{D}u, J\mathbb{D}u \rangle \geq 0$ *for each* $u \in \mathcal{U}_*^*(0)$*, then* $\mathbb{D}$ *is called* positively J-coercive (over $\mathcal{U}_*^*$).

(Note that for $\mathcal{U}_{\mathrm{out}}$, [positive] J-coercivity depends on $\mathbb{D}$ and J only, not on the rest of Σ.)

If $\widehat{\mathbb{D}}$ is a rational matrix-valued function or stable, then $\mathbb{D}$ is positively J-coercive over $\mathcal{U}_{\mathrm{out}}$ iff $\widehat{\mathbb{D}}^*J\widehat{\mathbb{D}} \geq \varepsilon I$ a.e. on $i\mathbf{R}$ for some $\varepsilon > 0$ ("$\widehat{\mathbb{D}}$ has a full column rank on $i\mathbf{R} \cup \{\infty\}$"), by Proposition 10.3.1(b)&(c), unless U is unseparable (see (c)).

By Lemma 8.2.3, we have the following when Z^s is a Hilbert space (here $\mathbb{D}$ stands for $\mathbb{D}|_{\mathcal{U}_*^*(0)}$):

(b) $\mathbb{D} \in \mathcal{B}(\mathcal{U}_*^*(0), \mathrm{L}^2(\mathbf{R}_+;Y))$ and $\mathcal{U}_*^*(0)$ is a Hilbert space (under an equivalent norm);

(c1) $\mathbb{D}$ is J-coercive iff $\mathbb{D}^* J \mathbb{D} \in \mathcal{GB}(\mathcal{U}_*^*(0))$

(c2) $\mathbb{D}$ is positively J-coercive iff $\mathbb{D}^* J \mathbb{D} \gg 0$ on $\mathcal{U}_*^*(0)$, i.e., iff $\langle \mathbb{D}u, J\mathbb{D}u \rangle \geq \varepsilon \|u\|_{\mathcal{U}_*^*}^2$ for all $u \in \mathcal{U}_*^*(0)$ and some $\varepsilon > 0$.

We can simplify the $\|\cdot\|_{\mathcal{U}_*^*}$ norms as follows:

Lemma 8.4.2 *The norm $\|\cdot\|_{\mathcal{U}_*^*}$ is a norm on $\mathcal{U}_*^*(0)$. The following norms are equivalent to $\|\cdot\|_{\mathcal{U}_{\text{out}}}$, $\|\cdot\|_{\mathcal{U}_{\text{sta}}}$, $\|\cdot\|_{\mathcal{U}_{\text{str}}}$ and $\|\cdot\|_{\mathcal{U}_{\text{exp}}}$, respectively:*

$$\|u\|'_{\mathcal{U}_{\text{out}}} := \max\{\|u\|_2, \|\mathbb{D}u\|_2\}, \tag{8.69}$$

$$\|u\|'_{\mathcal{U}_{\text{str}}} := \max\{\|u\|_2, \|\mathbb{D}u\|_2, \|\mathbb{B}\tau u\|_\infty\} =: \|u\|'_{\mathcal{U}_{\text{sta}}}, \tag{8.70}$$

$$\|u\|'_{\mathcal{U}_{\text{exp}}} := \max\{\|u\|_2, \|\mathbb{B}\tau u\|_2\}. \tag{8.71}$$

□

(For $\|u\|'_{\mathcal{U}_{\text{exp}}}$, this follows from Lemma 6.7.8; the other claims are obvious.)

If $\mathcal{U}_*^* = \mathcal{U}_{\text{out}}$ or $\mathcal{U}_*^* = \mathcal{U}_{\text{exp}}$, then Z^s is a Hilbert space, so that then (b)–(c2) above hold and J-coercivity implies the existence of a unique J-critical control:

Theorem 8.4.3 (J-coercive $\Rightarrow \exists! J$-critical control) *Assume that Z^s is a reflexive Banach space and $\mathbb{D}$ is J-coercive. If $x_0 \in H$ is s.t. $\mathcal{U}_*^*(x_0) \neq \emptyset$, then there is a unique J-critical control over $\mathcal{U}_*^*$ for x_0.* □

(This follows from Theorem 8.2.5. J-coercivity is not the weakest possible assumption, e.g., let $C = 0$, $D > 0$, $J = I$ (but not $D \gg 0$). However, with reasonable additional assumptions, we obtain the converse for $\mathcal{U}_*^* = \mathcal{U}_{\text{exp}}$, as in, e.g., Theorem 9.2.16.)

Since $\mathcal{U}_{\text{out}}$ and $\mathcal{U}_{\text{exp}}$ are the most common sets of admissible controls, and J-coercivity implies all standard classical coercivity assumptions for control problems (see Section 10.3), the above theorem suffices for most applications. However, we often obtain results for $\mathcal{U}_{\text{str}}$ and $\mathcal{U}_{\text{sta}}$ from those for $\mathcal{U}_{\text{out}}$ by suitable strong stabilizability assumptions that make these three equal.

The uniqueness part of the above theorem does not require reflexivity:

Lemma 8.4.4 *If $\mathbb{D}$ is J-coercive, then there is at most one J-critical control for each $x_0 \in H$.* □

(This follows from Lemma 8.2.3.)

If there is a J-critical state feedback pair over $\mathcal{U}_{\text{exp}}$, then Σ is exponentially stabilizable. On the other hand, if Σ is exponentially stabilizable, then optimization over $\mathcal{U}_{\text{exp}}$ can be reduced to optimization of the corresponding exponentially stable closed-loop system:

Theorem 8.4.5 (Reduce $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$ to $\mathcal{U}_{\rm exp}$) *Let $\left[\begin{array}{c|c} \widetilde{\mathbb{K}} & \widetilde{\mathbb{F}} \end{array}\right]$ be admissible for Σ, let $\Sigma_\flat$ be the corresponding closed loop system, and set $\widetilde{\mathbb{X}} := I - \widetilde{\mathbb{F}}$, $\widetilde{\mathbb{M}} := \widetilde{\mathbb{X}}^{-1}$. Set*

$$\mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,|\, \mathbb{A}_\flat x_0 + \mathbb{B}_\flat \tau u \in \mathrm{L}^2\}, \tag{8.72}$$

$$\mathcal{U}_{\rm out}^{\Sigma_\flat}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,|\, \mathbb{C}_\flat x_0 + \mathbb{D}_\flat \tau u \in \mathrm{L}^2\} \qquad (x_0 \in H). \tag{8.73}$$

Then the following hold:

(a) The system $\Sigma_\flat$ has a J-critical pair over $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$ iff Σ has a J-critical pair over $\mathcal{U}_{\rm exp}$.

Moreover, if $\left[\begin{array}{c|c} \mathbb{K}_\natural & \mathbb{F}_\natural \end{array}\right]$ is J-critical over $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$ for $\Sigma_\flat$, then $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right] := \left[\begin{array}{c|c} \mathbb{K}_\natural + \mathbb{X}_\natural \widetilde{\mathbb{K}} & I - \mathbb{X}_\natural \widetilde{\mathbb{X}} \end{array}\right]$ (here $\mathbb{X}_\natural := I - \mathbb{F}_\natural$) is J-critical over $\mathcal{U}_{\rm exp}$ for Σ.

Conversely, if $\left[\begin{array}{c|c} \mathbb{K} & \mathbb{F} \end{array}\right]$ is J-critical over $\mathcal{U}_{\rm exp}$ for Σ, then $\left[\begin{array}{c|c} \mathbb{K}_\natural & \mathbb{F}_\natural \end{array}\right] := \left[\begin{array}{c|c} \mathbb{K} - \mathbb{X}\mathbb{K}_\flat & I - \mathbb{X}\widetilde{\mathbb{M}} \end{array}\right]$ (here $\mathbb{X} := I - \mathbb{F}$) is J-critical over $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$ for $\Sigma_\flat$.

The corresponding closed-loop systems correspond to each other as in Lemma 6.7.12.

(b) $\mathbb{K}_{\rm crit}$ is a J-critical control over $\mathcal{U}_{\rm exp}$ in WPLS form for Σ iff $\mathbb{K}_{\rm crit}^\flat := \widetilde{\mathbb{X}}\mathbb{K}_{\rm crit} - \widetilde{\mathbb{K}}$ is a J-critical control over $\mathcal{U}_{\rm exp}$ in WPLS form for $\Sigma_\flat$.

(c1) Let $x_0 \in H$. If $u_\flat \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0)$, then $u := \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}} u_\flat \in \mathcal{U}_{\rm exp}(x_0)$ and $u_\flat = -\widetilde{\mathbb{K}} x_0 + \widetilde{\mathbb{X}} u$.

Conversely, if $u \in \mathcal{U}_{\rm exp}(x_0)$, then $u_\flat := -\widetilde{\mathbb{K}} x_0 + \widetilde{\mathbb{X}} u \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0)$ and $u = \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}} u_\flat$.

Thus, $\mathcal{U}_{\rm exp}(x_0) = \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}}[\mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0)]$ and $\mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0) = -\widetilde{\mathbb{K}} x_0 + \widetilde{\mathbb{X}}[\mathcal{U}_{\rm exp}(x_0)]$; in particular, $\mathcal{U}_{\rm exp}(0) = \widetilde{\mathbb{M}}\mathcal{U}_{\rm exp}^{\Sigma_\flat}(0)$. Moreover, $y := \mathbb{C} x_0 + \mathbb{D} u = \mathbb{C}_\flat x_0 + \mathbb{D}_\flat u_\flat$ and $x := \mathbb{A} x_0 + \mathbb{B}\tau u = \mathbb{A}_\flat x_0 + \mathbb{B}_\flat \tau u_\flat$ in either case.

(c2) If $u = \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}} u_\flat$, then u is J-critical over $\mathcal{U}_{\rm exp}(x_0)$ iff $u_\flat$ is J-critical over $\mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0)$.

(c3) If(f) there is a unique J-critical control over $\mathcal{U}_{\rm exp}(x_0)$ for Σ and for each $x_0 \in H$, then the same holds for $\left[\begin{array}{c|c} \mathbb{A}_\flat & \mathbb{B}_\flat \\ \hline \mathbb{C}_\flat & \mathbb{D}_\flat \end{array}\right]$ and $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$, and $\mathbb{A}_{\rm crit} = \mathbb{A}_{\rm crit}^\flat$, $\mathbb{C}_{\rm crit} = \mathbb{C}_{\rm crit}^\flat$ and $\mathcal{P}$ are common for Σ and $\Sigma_\flat$, but $\mathbb{K}_{\rm crit} = \mathbb{K}_\flat + \widetilde{\mathbb{M}}\mathbb{K}_{\rm crit}^\flat$, $\mathbb{K}_{\rm crit}^\flat = \widetilde{\mathbb{X}}\mathbb{K}_{\rm crit} - \widetilde{\mathbb{K}}$.

(c4) If(f) there is a J-critical control over $\mathcal{U}_{\rm exp}(x_0)$ for Σ and for all $x_0 \in H$, then the same holds for $\left[\begin{array}{c|c} \mathbb{A}_\flat & \mathbb{B}_\flat \\ \hline \mathbb{C}_\flat & \mathbb{D}_\flat \end{array}\right]$ and $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$, and $\mathcal{P}$ is common for Σ and $\Sigma_\flat$.

(c5) There is $\varepsilon > 0$ s.t. $\varepsilon\|u\|_{\mathcal{U}_{\rm exp}} \le \|u_\flat\|_{\mathcal{U}_{\rm exp}^{\Sigma_\flat}} \le \varepsilon^{-1}\|u\|_{\mathcal{U}_{\rm exp}}$ whenever $u \in \mathcal{U}_{\rm exp}(0)$ and $u_\flat \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}(0)$ are as in (c1).

(d) The map $\mathbb{D}$ is [positively] J-coercive over $\mathcal{U}_{\rm exp}$ iff $\mathbb{D}_\flat$ is [positively] J-coercive over $\mathcal{U}_{\rm exp}^{\Sigma_\flat} = \mathcal{U}_{\rm out}^{\Sigma_\flat}$.

(e) If $\left[\ \widetilde{\mathbb{K}}\ \middle|\ \widetilde{\mathbb{F}}\ \right]$ *is exponentially stabilizing, then*

$$\mathcal{U}_{\rm exp}(x_0) = \{\mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}} u_\flat \,\big|\, u_\flat \in \mathrm{L}^2(\mathbf{R}_+;U)\} \tag{8.74}$$

and $\mathcal{U}^{\Sigma_\flat}_{\rm exp}(x_0) = \mathcal{U}^{\Sigma_\flat}_{\rm out}(x_0) = \mathrm{L}^2(\mathbf{R}_+;U)$*, for all* $x_0 \in H$*.*

(f) Claims (a)–(e) also hold with replacements $\mathcal{U}_{\rm exp} \mapsto \mathcal{U}^{\vartheta}_{[\mathbb{Q}\ \mathbb{R}]}$*,* $\mathcal{U}^{\Sigma_\flat}_{\rm exp}, \mathcal{U}^{\Sigma_\flat}_{\rm out} \mapsto \mathcal{U}^{\gamma,\Sigma_\flat}_{[\mathbb{Q}_\flat\ \mathbb{R}_\flat]}$ *and* $\mathrm{L}^2(\mathbf{R}_+;U) \mapsto \mathcal{U}^{\gamma,\Sigma_\flat}_{[\mathbb{Q}_\flat\ \mathbb{R}_\flat]}(x_0)$*, where* $\begin{bmatrix}\mathbb{Q}_\flat & \mathbb{R}_\flat\end{bmatrix} := \begin{bmatrix}\mathbb{Q} & \mathbb{R}\\ 0 & I\end{bmatrix}\begin{bmatrix}I & 0\\ \mathbb{K}_\flat & \widetilde{\mathbb{M}}\end{bmatrix}$*,* $Z^{\rm u}_\flat := Z^{\rm u} \times \mathrm{L}^2_\gamma$*,* $Z^{\rm s}_\flat := Z^{\rm s} \times \mathrm{L}^2_\vartheta$ *and* $\gamma > \max\{\omega_A, \omega_{A_\flat}, \vartheta\}$*.*

(g1) If $\left[\ \widetilde{\mathbb{K}}\ \middle|\ \widetilde{\mathbb{F}}\ \right]$ *is q.r.c.-SOS-stabilizing, then (a)–(e) also hold with replacements* $\mathcal{U}_{\rm exp} \mapsto \mathcal{U}_{\rm out}$ *and* $\mathcal{U}^{\Sigma_\flat}_{\rm exp} \mapsto \mathcal{U}^{\Sigma_\flat}_{\rm out} = \mathrm{L}^2(\mathbf{R}_+;U)$*.*

Moreover, then $\left[\ \mathbb{K}\ \middle|\ \mathbb{F}\ \right]$ *(in (a)) is q.r.c.-SOS-stabilizing iff* $\left[\ \mathbb{K}_\natural\ \middle|\ \mathbb{F}_\natural\ \right]$ *is q.r.c.-SOS-stabilizing (equivalently, stable and [r.c.-]SOS-stabilizing).*

(g2) If $\left[\ \widetilde{\mathbb{K}}\ \middle|\ \widetilde{\mathbb{F}}\ \right]$ *is [[exponentially] strongly] q.r.c.-stabilizing, then* $\mathcal{U}_{\rm out} = \mathcal{U}_{\rm sta}[= \mathcal{U}_{\rm str}[= \mathcal{U}_{\rm exp}]]$*.*

Thus, if we are optimizing over $\mathcal{U}_{\rm exp}$, we only need to stabilize Σ exponentially and then find an J-critical control for $\Sigma_\flat$ w.r.t. $\mathcal{U}_{\rm out} = \mathcal{U}_{\rm exp}$ corresponding to $\Sigma_\flat$. If the original system is J-coercive, then we end up with the situation of Proposition 8.3.10, by (d).

The key to the Theorem is (c1), the fact that $u, x \in \mathrm{L}^2 \Leftrightarrow u_\flat, x \in \mathrm{L}^2$ (this follows from Lemma 6.1.10). As shown by Example 9.13.2 (for $\mathcal{U}^*_* = \mathcal{U}_{\rm out}$), an analogous reduction cannot be made for general $\mathcal{U}^*_*$. Indeed, we do not have a similar equivalence "$u, y \in \mathrm{L}^2 \Leftrightarrow u_\flat, y \in \mathrm{L}^2$" for $\mathcal{U}_{\rm out}$ unless $\left[\ \widetilde{\mathbb{K}}\ \middle|\ \widetilde{\mathbb{F}}\ \right]$ is q.r.c.-SOS-stabilizing (cf. Theorem 9.9.10). Fortunately, part (f) is helpful in certain technical proofs.

Proof of Theorem 8.4.5: (c1) 1° Let $u_\flat \in \mathcal{U}^{\Sigma_\flat}_{\rm exp}(x_0)$, $x_0 \in H$. Then $u, x \in \mathrm{L}^2$, by Lemma 6.1.10, where $u := \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}} u_\flat = \mathbb{K}_\flat x_0 + \mathbb{F}_\flat u_\flat + u_\flat$ and $x := \mathbb{A} x_0 + \mathbb{B}\tau u = \mathbb{A}_\flat x_0 + \mathbb{B}_\flat \tau u_\flat$. Thus, $u \in \mathcal{U}_{\rm exp}(x_0)$.

2° By exchanging the roles of Σ and $\left[\begin{array}{c|c}\mathbb{A}_\flat & \mathbb{B}_\flat\\ \hline \mathbb{C}_\flat & \mathbb{D}_\flat\end{array}\right]$ (note that the pair $-\left[\ \mathbb{K}_\flat\ \middle|\ \mathbb{F}_\flat\ \right]$ is admissible for $\left[\begin{array}{c|c}\mathbb{A}_\flat & \mathbb{B}_\flat\\ \hline \mathbb{C}_\flat & \mathbb{D}_\flat\end{array}\right]$, and Σ with the added row $-\left[\ \mathbb{K}\ \middle|\ \mathbb{F}\ \right]$ is the corresponding closed-loop system, by Lemma 6.6.14), we note that if $u \in \mathcal{U}_{\rm exp}(x_0)$, then $u_\flat \in \mathcal{U}^{\Sigma_\flat}_{\rm exp}(x_0)$, where $u_\flat := -\mathbb{K} x_0 - \mathbb{F} u + I = -\mathbb{K} x_0 + \mathbb{X} u$.

3° We noted above that "$x = x$"; the same holds for y: $\mathbb{C} x_0 + \mathbb{D} u = \mathbb{C} x_0 + \mathbb{D}\mathbb{K}_\flat x_0 + \mathbb{D}\widetilde{\mathbb{M}} u_\flat = \mathbb{C}_\flat x_0 + \mathbb{D}_\flat u_\flat$.

(c2) Now $u_\flat$ is J-critical over $\mathcal{U}^{\Sigma_\flat}_{\rm exp}(x_0)$ iff $\langle y, J\mathbb{D}_\flat \eta_\circlearrowleft\rangle = \langle y, J\mathbb{D}\eta\rangle = 0$ for all $\eta_\circlearrowleft \in \mathcal{U}_{\rm exp}(0)$, i.e., for all $\eta := \widetilde{\mathbb{M}}\eta_\circlearrowleft \in \mathcal{U}_{\rm exp}(x_0)$, i.e., iff u is J-critical over $\mathcal{U}_{\rm exp}(x_0)$.

(c3) By (c1)–(c2), there is a unique J-critical control over $\mathcal{U}^{\Sigma_\flat}_{\rm exp}(x_0)$ for each x_0, and (c3) holds. (Exchange the roles of Σ and $\Sigma_\flat$ for the converse.)

(c4) The first claim follows from (c2). By (c1), $\langle y, Jy\rangle$ $(=: \langle x_0, \mathcal{P}x_0\rangle)$ is the same for $\mathcal{U}_{\rm exp}$ and $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$, hence so is $\mathcal{P}$.

(c5) Set $x := \mathbb{B}\tau u = \mathbb{B}_\flat \tau u_\flat$ (by (c1), we have $u = \widetilde{\mathbb{M}} u_\flat$). Then $\|u_\flat\|_{\mathcal{U}_{\rm exp}^{\Sigma_\flat}} := \max(\|u_\flat\|_2, \|x\|_2) \le 2M \max(\|u\|_2, \|x\|_2) =: \|u\|_{\mathcal{U}_{\rm exp}}$ (here we have used (8.71); for an equivalent norm we need to divide ε by an equivalence constant) for some $M := M_{\Sigma'} < \infty$, by Lemma 6.7.8, where $\Sigma' := \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline -\widetilde{\mathbb{K}} & \widetilde{\mathbb{X}} \end{array}\right]$; analogously, $\max(\|u\|_2, \|x\|_2) \le 2M' \max(\|u_\flat\|_2, \|x\|_2)$ for some fixed $M' < \infty$; take $\varepsilon := \min\{(2M)^{-1}, (2M')^{-1}\}$.

(a) This follows from Lemma 6.7.12 (either directly (since $\mathbb{C}_\circlearrowleft$ and $\mathbb{D}_\circlearrowleft$ are common for both closed-loop systems) or from the fact that $\mathbb{K}_\circlearrowleft = \widetilde{\mathbb{M}}\mathbb{K}_\circlearrowleft^\flat + \mathbb{K}_\flat$, $\mathbb{K}_\circlearrowleft^\flat = -\widetilde{\mathbb{K}} + \widetilde{\mathbb{X}}\mathbb{K}_\circlearrowleft$, as in (c3)).

(b) Set $\mathbb{K}_{\rm crit}^\flat := \widetilde{\mathbb{M}}^{-1}(\mathbb{K}_{\rm crit} - \mathbb{K}_\flat)$, $\mathbb{A}_{\rm crit} := \mathbb{A} + \mathbb{B}\tau\mathbb{K}_{\rm crit}$, and $\mathbb{C}_{\rm crit} := \mathbb{C} + \mathbb{D}\mathbb{K}_{\rm crit}$. Then $\mathbb{A}_\flat + \mathbb{B}_\flat\tau\mathbb{K}_{\rm crit}^\flat = \mathbb{A}_{\rm crit}$ and $\mathbb{C}_\flat + \mathbb{D}_\flat\mathbb{K}_{\rm crit}^\flat = \mathbb{C}_{\rm crit}$. By a straightforward computation using the above formulae (and the identity $\widetilde{\mathbb{X}}\pi_+\widetilde{\mathbb{M}}\pi_- = -\pi_+\widetilde{\mathbb{X}}\pi_-\widetilde{\mathbb{M}}$), one verifies that $\mathbb{K}_{\rm crit}^\flat \mathbb{A}_{\rm crit}^t = \pi_+\tau^t\mathbb{K}_{\rm crit}^\flat$ for any $t \ge 0$, so that $\left[\begin{array}{c|cc} \mathbb{A}_{\rm crit}^{\rm T} & \mathbb{C}_{\rm crit}^{\rm T} & {\mathbb{K}_{\rm crit}^\flat}^{\rm T} \end{array}\right]^{\rm T} \in$ WPLS; thus, $\mathbb{K}_{\rm crit}^\flat$ is a control in WPLS form.

Now $\langle J\mathbb{C}_{\rm crit}x_0, \mathbb{D}\eta\rangle = 0$ for all $\eta \in \mathcal{U}_{\rm exp}(0)$ iff $\langle J\mathbb{C}_{\rm crit}x_0, \mathbb{D}_\flat\eta_{\flat'}\rangle = 0$ for all $\eta_{\flat'} \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}(0) = \widetilde{\mathbb{M}}^{-1}\,\mathcal{U}_{\rm exp}(0)$, hence (b) holds (since we can interchange Σ and $\Sigma_\flat$ for the converse).

(d) (Note that the J-coercivity of $\mathbb{D}_\flat$ over $\mathcal{U}_{\rm exp}^{\Sigma_\flat}$ is equivalent to the J-coercivity of $\mathbb{D}_\flat$ over $\mathcal{U}_{\rm out}^{\Sigma_\flat}$, by Lemma 8.3.3.)

This follows from (c5) and (c1): Let $\mathbb{D}$ be J-coercive over $\mathcal{U}_{\rm exp}$, and let $\varepsilon > 0$ as in Definition 8.4.1. Given a nonzero $u_\flat \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}(0)$, set $u := \widetilde{\mathbb{M}}u_\flat \in \mathcal{U}_{\rm exp}(0) \setminus \{0\}$ (by (c1)), $x := \mathbb{B}\tau u = \mathbb{B}_\flat\tau u_\flat$. Choose a nonzero $v \in \mathcal{U}_{\rm exp}(0)$ as in Definition 8.4.1, and set $v_\flat := \widetilde{\mathbb{X}}v \in \mathcal{U}_{\rm exp}^{\Sigma_\flat}$, $\widetilde{x} := \mathbb{B}\tau v$. Then $\langle\mathbb{D}_\flat v_\flat, J\mathbb{D}_\flat u_\flat\rangle = \langle\mathbb{D}v, J\mathbb{D}u\rangle \ge \varepsilon\|u\|\|v\| \ge \varepsilon'\|u_\flat\|\|v_\flat\|$. Since $u_\flat$ was arbitrary, $\mathbb{D}_\flat$ is J-coercive. Exchange Σ and $\Sigma_\flat$ for the converse. [By (c1), $\langle\mathbb{D}\cdot, J\mathbb{D}\cdot\rangle \ge 0 \Leftrightarrow \langle\mathbb{D}_\flat\cdot, J\mathbb{D}_\flat\cdot\rangle \ge 0$.]

(e) By Lemma 6.1.10, we have $\mathcal{U}_{\rm exp}^{\Sigma_\flat}(x_0) = \mathcal{U}_{\rm out}^{\Sigma_\flat}(x_0) = \mathrm{L}^2(\mathbf{R}_+; U)$ for all $x_0 \in H$. By (c1), we obtain (8.74).

(f) 1° *The definition of* $\begin{bmatrix}\mathbb{Q}_\flat & \mathbb{R}_\flat\end{bmatrix}$ *implies that (c1) holds:* Indeed, $\widetilde{\mathbb{K}}$, $\widetilde{\mathbb{X}}$, $\mathbb{K}_\flat$ and $\widetilde{\mathbb{M}}$ are γ-stable, the upper row of $\begin{bmatrix}\mathbb{Q}_\flat & \mathbb{R}_\flat\end{bmatrix}$ corresponds to condition $\begin{bmatrix}\mathbb{Q} & \mathbb{R}\end{bmatrix}\begin{bmatrix}x_0 \\ u\end{bmatrix} \in Z^{\rm s}$ and the lower row to condition $u \in \mathrm{L}^2_\vartheta$, so that $\mathcal{U}^{\gamma,\Sigma_\flat}_{[\mathbb{Q}_\flat\ \mathbb{R}_\flat]}$ is independent on γ (since $u := \mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}}u_\flat \in \mathrm{L}^2_\vartheta \Rightarrow u_\flat = \widetilde{\mathbb{K}}x_0 + \widetilde{\mathbb{X}}u \in \mathrm{L}^2_\gamma$).

2° *The rest:* The proofs of (a)–(e) above apply with slight changes ((c5) becomes easier).

(g1) By Lemma 6.5.6(f)&(a1), we have $\mathcal{U}_{\rm out}(0) = \mathbb{M}\mathrm{L}^2(\mathbf{R}_+; U)$. Since $\mathbb{K}_\flat x_0 \in \mathcal{U}_{\rm out}(x_0)$ for all $x_0 \in H$, we obtain "$\mathcal{U}_{\rm out} = \{\mathbb{K}_\flat x_0 + \widetilde{\mathbb{M}}u_\flat \,|\, u_\flat \in \mathrm{L}^2(\mathbf{R}_+; U)\}$" (cf. (8.74) from Lemma 8.3.5. The proofs of (a)–(d) above apply with slight changes (use, e.g., Lemma 8.4.11(b1) for J-coercivity). The last claim follows from Lemma 6.7.11(a2) and Lemma 6.6.17(b).

(g2) Now the proof and conlusion of (g1) applies also to $\mathcal{U}_{\rm sta}$ [and $\mathcal{U}_{\rm str}$ [and

$\mathcal{U}_{\text{exp}}$]] in place of $\mathcal{U}_{\text{out}}$. [[Note that it suffices that $\left[\begin{array}{c|c} \widetilde{\mathbb{K}} & \widetilde{\mathbb{F}} \end{array}\right]$ is exponentially stabilizing and q.r.c.-stabilizing.]] □

We will often need the assumption that J-coercivity implies the existence of a spectral factorization. It is well-known that this is true for any rational transfer function (hence for any stable I/O map of a system with $\dim H < \infty$); in fact, this is true for any element of $\text{MTIC}_{T\mathbf{Z}}$, as we shall show in Theorem 8.4.9. Since there is a wide variety of classes satisfying this assumption, we shall write below three hypotheses with differing strengths, and then use these as the assumptions of our results in optimal control theory, to avoid dependence on the current state of spectral factorization theory (or on the part included in this book).

We start by the weakest formulation:

Definition 8.4.6 (J-coercive ⇒ SpF) *Let $\mathbb{D} \in \text{TIC}(U,Y)$ and $J = J^* \in \mathcal{B}(Y)$. We write $(\mathbb{D},J) \in \text{SpF}$ iff either $\mathbb{D}$ is not J-coercive or $\mathbb{D}^*J\mathbb{D}$ has a spectral factorization.*

Thus, $(\mathbb{D},J) \in \text{SpF}$ means that if $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible, then $\mathbb{D}^*J\mathbb{D} = \mathbb{X}^*S\mathbb{X}$ for some $\mathbb{X} \in \mathcal{G}\text{TIC}(U)$ and $S \in \mathcal{G}\mathcal{B}(U)$. This (and the stronger requirement below) is satisfied by any of the classes in Theorem 8.4.9 (alternatively, by any $\mathbb{D} \in \text{TIC}$ if $J \gg 0$). Recall from Lemma 6.4.7(b) that the converse holds for any $\mathbb{D} \in \text{TIC}(U,Y)$ and $J = J^* \in \mathcal{B}(Y)$.

However, often we also need to know that $\mathbb{D}$ belongs not only to the class TIC but also to some subclass "$\mathcal{A}$" whose every element is ULR and has the above property (for each J) and which is closed w.r.t. spectral factorization. We formulate this as follows (see Definition 6.2.4 for "$\underset{a}{\subset}$"):

Hypothesis 8.4.7 (ULR classes $\mathcal{A}(U)$ that admit spectral factorization)

(1.) We have $\mathcal{B} \subset \mathcal{A} \underset{a}{\subset} \text{TIC} \cap \text{ULR}$;

(2.) if Y is an arbitrary Hilbert space, $\mathbb{D} \in \mathcal{A}(U,Y)$, $J = J^* \in \mathcal{B}(Y)$, and the Popov Toeplitz operator $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible on $\pi_+\mathrm{L}^2$, then $\mathbb{D}^*J\mathbb{D} = \mathbb{X}^*S\mathbb{X}$ for some $S = S^* \in \mathcal{G}\mathcal{B}(U)$ and $\mathbb{X} \in \mathcal{G}\mathcal{A}(U)$.

If Hypothesis 8.4.7 holds for $\mathcal{A}(U)$, then, trivially, $(\mathbb{D},J) \in \text{SpF}$ for any Y, $\mathbb{D} \in \mathcal{A}(U,Y)$ and $J = J^* \in \mathcal{B}(Y)$ See Hypothesis 10.6.6 and Lemma 10.6.7 for the positive case.

A sufficient condition for (2.) is that if $\mathbb{E} = \mathbb{E}^* \in \mathcal{A}(U)$ and $\pi_+\mathbb{E}\pi_+$ is invertible on $\pi_+\mathrm{L}^2(U)$, then $\mathbb{E}$ has a spectral factorization over $\mathcal{A}(U)$. However, the weaker formulation above has the advantage to cover also exponentially stable classes (cf. Theorem 8.4.9) and still be strong enough for applications.

Much of our theory is valid even without the assumption (1.), but because all classes listed in Theorem 8.4.9 satisfy (1.), we have assumed it to simplify the presentation.

Sometimes we also wish to have $D^*JD = X^*SX$ (cf. Example 6.3.7):

Hypothesis 8.4.8 (Classes $\mathcal{A}(U)$ that admit spectral factorization with $D^*JD = X^*SX$) *We require that $\mathcal{A}(U)$ satisfies Hypothesis 8.4.7 with*

(3.) $D^*JD = X^*SX$.

By Lemma 6.4.5(a), condition (3.) holds for some $\mathbb{X}$ and S satisfying (2.) iff (3.) holds for all such $\mathbb{X}$ and S (for fixed $\mathbb{D}$ and J). A sufficient condition is that $\mathcal{A} \subset \text{SHPR}$, by Lemma 6.3.6(b).

Now we cite the main results of Chapter 5:

Theorem 8.4.9 (Classes satisfying Hypothesis 8.4.7) *Let U be a Hilbert space, let $\mathbf{S} = \mathbf{S} - \mathbf{S} \subset \mathbf{R}$, and let (α), (β) or (γ) hold, where*

- *(α) $\mathcal{A}$ is one of the classes* MTIC^{L^1}, $\text{MTIC}^{\text{L}^1,\mathcal{BC}}$, $\text{MTIC}_{T\mathbf{Z}}$, $\text{MTIC}^{\mathcal{BC}}_{T\mathbf{Z}}$, $\text{MTIC}_{\text{d},T\mathbf{Z}}$, *and* $\text{MTIC}^{\mathcal{BC}}_{\text{d},T\mathbf{Z}}$*;*
- *(β)* $\dim U < \infty$ *and $\mathcal{A}$ is one of the classes* MTIC, MTIC_{d}, $\text{MTIC}_{\mathbf{S}}$, *and* $\text{MTIC}_{\text{d},\mathbf{S}}$.
- *(γ) $\mathcal{A}(U,Y) = \mathcal{B}(U,Y) + \{\mathbb{D} \,|\, \widehat{\mathbb{D}} \in \text{H}^2_{\text{strong}}(\mathbf{C}^+_\omega; \mathcal{B}(U,Y))$ for some $\omega < 0\}$ (this is the set of maps having an exponentially stable realization with a bounded input operator).*

Then Hypothesis 8.4.7 holds for $\mathcal{A}(U)$ and for the class $\mathcal{A}_{\text{exp}}(U)$ of exponentially stable $\mathcal{A}(U)$ maps, *where*

$$\mathcal{A}_{\text{exp}}(U,Y) := \cap_{r<0}\{e^{r\cdot}\mathbb{D}e^{-r\cdot} \,\big|\, \mathbb{D} \in \mathcal{A}(U,Y)\} \tag{8.75}$$

Moreover,

- *(a) If $\mathcal{A} = \text{MTIC}^{\text{L}^1}$, $\mathcal{A} = \text{MTIC}^{\text{L}^1,\mathcal{BC}}$ or (γ) holds, then also Hypothesis 8.4.8 holds for $\mathcal{A}(U)$ as well as for $\mathcal{A}_{\text{exp}}(U)$ (for any Hilbert space U).*
- *(b) We have $X \in \mathcal{GB}(U)$, and we can choose $\mathbb{X}$ and S s.t. $X = I$.*
- *(c) If (α) or (β) holds, then $\mathcal{A} = \mathcal{A}^{\text{d}}$. We have $\mathcal{A} = \mathcal{A}^{\text{d}}$ also for the class*

  (γ′) $\mathcal{A}(U,Y) = \mathcal{B}(U,Y) + \{\mathbb{D} \,|\, \widehat{\mathbb{D}}, \widehat{\mathbb{D}}(\bar{\cdot})^* \in \text{H}^2_{\text{strong}}(\mathbf{C}^+_\omega; \mathcal{B}(*,*))$ for some $\omega < 0\}$ (this is the set of I/O maps having an exponentially stable PS-realization).

Thus, Hypothesis 8.4.7 holds for $\mathcal{A}(U) = \text{MTIC}_{T\mathbf{Z}}(U)$ and $\mathcal{A}(U) = \text{MTIC}(\mathbf{C}^n)$ (and for their subclasses of exponentially stable maps) for any $n \in \mathbf{N}$ and any Hilbert space U (we hope that the future study will show the reference to U superfluous for MTI (and its subclasses), i.e., that Hypothesis 8.4.7 holds for $\text{MTI}(U)$ for any Hilbert space U).

See Lemma 14.3.5 for four more classes (the Cayley images of discrete ℓ^1 classes). Also Theorem 9.2.14 contains analogous results, with requirements posed on the whole system (e.g., we may allow for any $\mathbb{D}, \mathbb{B}\tau \in \text{SMTIC}^{\text{L}^1}_{-\varepsilon}$ if C is bounded and $D^*JC = 0$, to obtain "(1.)–(3.)").

Proof of Theorem 8.4.9: *Case (γ):* Now Hypothesis 8.4.8 holds, by Theorem 9.2.14(c2) (by its proof we have $\mathbb{X} \in \mathcal{A}(U)$).

(It is not a problem that we refer here to later results; part (γ) of this lemma is not used in this monograph before Chapter 12 except in phrases like "if $\mathcal{A}$ is any of the classes of Theorem 8.4.9"; in particular, part (γ) is not used in derivation of any earlier result of this monograph, we just want to record it next to (α) and (β).)

The assumption means that $\widehat{\mathbb{D}} : \mathbf{C}_\omega^+ \to \mathcal{B}(U,Y)$ is s.t. $\widehat{\mathbb{D}}u_0 \in \mathrm{H}^2_\omega(\mathbf{C}^+;Y)$ for all $u_0 \in U$ and $y_0 \in Y$ (see Lemma F.3.2(a)); the number $\omega < 0$ may depend on $\mathbb{D}$. (Thus, $\mathcal{A}_{\text{exp}} = \mathcal{A}$.)

The correspondence to bounded B was shown in Theorem 6.9.1.

Case (γ′): (The remarks of case (γ) apply. The correspondence to PS-systems was shown in Theorem 6.9.6.)

We obtain the result from case (γ) (since obviously $\mathcal{A} = \mathcal{A}^{\mathrm{d}}$) except for the fact that $\widehat{\mathbb{X}}(\bar{\cdot})^* - X^* \subset \mathrm{H}^2_{\text{strong}}(\mathbf{C}_\varepsilon^+;\mathcal{B}(U))$ for some $\varepsilon < 0$, which was recorded in the proof of Theorem 9.2.14(c2).

Cases (α) and (β):

For the rest of the proof, we assume that (α) or (β) holds.

Hypotheses 8.4.7(1.)&(2.) (and Hypothesis 8.4.8(4.) if $\mathcal{A} = \mathrm{MTIC}^{\mathrm{L}^1}$ or $\mathcal{A} = \mathrm{MTIC}^{\mathrm{L}^1,\mathcal{BC}}$) is satisfied by Theorem 2.6.4.

Moreover, if $\mathbb{D}$ and J are as in Hypothesis 8.4.7(3.) and we set $\mathbb{E} := \mathbb{D}^* J \mathbb{D} \in \mathcal{A}'(U)$, where $\mathcal{A}' := \mathcal{A} + \mathcal{A}^*$ is the corresponding noncausal (MTI) class, then $\mathbb{E}$ has a spectral factorization in $\mathcal{A}(U)$, by Theorem 5.2.7 (and Lemma 5.2.1(d)), hence (3.) holds for $\mathcal{A}(U)$.

If, in addition, $\mathbb{D} \in \mathcal{A}_{\text{exp}}(U)$, i.e., $\mathbb{D} \in \mathcal{A}_\omega(U)$ for some $\omega < 0$, then $\mathbb{D} \in \mathcal{A}_\omega(U) \cap \mathcal{A}_{-\omega}(U)$ and $\mathbb{D}^* \in \mathcal{A}'_{-\omega} \cap \mathcal{A}'_\omega(U)$, by Theorem 2.6.4(g1)&(g2), hence then $\mathbb{E} := \mathbb{D}^* J \mathbb{D} \in \mathcal{A}'_{-\omega} \cap \mathcal{A}'_\omega(U)$; thus, then the spectral factorization of $\mathbb{E}$ is in fact a spectral factorization in $\mathcal{A}_{\text{exp}}(U)$, by Theorem 5.2.2. Therefore, (3.) holds for $\mathcal{A}_{\text{exp}}(U)$ too.

(a) This was noted above.

(b) This follows from Proposition 6.3.1(c). □

The reason for mentioning also subclasses of classes mentioned above is that in many theorems using Hypothesis 8.4.7 some kind of controllers are constructed within the same class, hence stricter conditions guarantee smoother controllers.

By [Treil94], the class $\mathrm{CTIC}(\mathbf{C})$ does not satisfy Hypothesis 8.4.7: There is $\mathbb{D} = \begin{bmatrix} \mathbb{D}_1 \\ \mathbb{D}_2 \end{bmatrix} \in \mathrm{CTIC}(\mathbf{C},\mathbf{C}^2)$ s.t. the spectral factor $\mathbb{X} \in \mathcal{G}\mathrm{TIC}(\mathbf{C})$ satisfying $\mathbb{X}^*\mathbb{X} = \mathbb{D}^*\mathbb{D}$ (i.e., $|\widehat{\mathbb{X}}|^2 = |\widehat{\mathbb{D}}_1|^2 + |\widehat{\mathbb{D}}_2|^2$ on $i\mathbf{R}$) does not belong to CTIC. We can even take $\widehat{\mathbb{D}}_1$ and $\widehat{\mathbb{D}}_2$ to have no zeros on $\overline{\mathbf{C}^+} \cup \{\infty\}$.

Lemma 8.4.10 *Let $\mathcal{A}(U)$ satisfy Hypothesis 8.4.7. Then $\mathcal{A}(U)$ is inverse closed in* $\mathrm{TIC}(U)$*, i.e.,* $\mathbb{X} \in \mathcal{A}(U) \cap \mathcal{G}\mathrm{TIC}(U) \Rightarrow \mathbb{X} \in \mathcal{G}\mathcal{A}(U)$.

(Analogously, $\mathcal{A}(U,Y)$ is inverse closed.)

Proof: Because the Toeplitz operator $\pi_+\mathbb{X}^*\mathbb{X}\pi_+$ has the inverse $\mathbb{X}^{-1}\pi_+\mathbb{X}^{-*}$ on $\pi_+\mathrm{L}^2$, we have $\mathbb{X}^*\mathbb{X} = \mathbb{Z}^*S\mathbb{Z}$ for some $S = S^* \in \mathcal{G}\mathcal{B}(U)$, $\mathbb{Z} \in \mathcal{G}\mathcal{A}(U)$, by (2.). By Lemma 6.4.5(a), $\mathbb{X} = E\mathbb{Z} \in \mathcal{G}\mathcal{A}(U)$ for some $E \in \mathcal{G}\mathcal{B}(U)$. □

In the stable or r.c.-stabilizable case (see Lemma 8.3.3), J-coercivity can be easily verified:

Lemma 8.4.11 (J-coercive) *Let $J = J^* \in \mathcal{B}(Y)$ and $\mathbb{D} \in \mathrm{TIC}_\infty(U,Y)$. The following holds for $\mathcal{U}_*^* := \mathcal{U}_{\mathrm{out}}$:*

- *(a1) Let $\mathbb{D} \in \mathrm{TIC}$. Then $\mathbb{D}$ is J-coercive iff the Popov Toeplitz operator $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible on $\mathrm{L}^2(\mathbf{R}_+;U)$.*
- *(a2) Let $\mathbb{D} \in \mathrm{TIC}$. Then $\mathbb{D}$ is positively J-coercive iff $\mathbb{D}^*J\mathbb{D} \gg 0$.*
- *(b1) Let $\mathbb{D} = \mathbb{N}\mathbb{M}^{-1}$ be a q.r.c.f. Then $\mathcal{U}_{\mathrm{out}}(0) = \mathbb{M}\mathrm{L}^2(\mathbf{R}_+;U)$. Moreover, $\mathbb{D}$ is [positively] J-coercive iff $\mathbb{N}$ is [positively] J-coercive.*
- *(b2) Let $\mathbb{D} = \mathbb{N}\mathbb{M}^{-1}$ be a q.r.c.f. and $S = \mathbb{N}^*J\mathbb{N}$. Then $\mathbb{D}$ is [positively] J-coercive over $\mathcal{U}_{\mathrm{out}}$ iff $S \in \mathcal{GB}$ [$S \gg 0$].*
- *(c) The space $\mathcal{U}_{\mathrm{out}}(0)$ is a Hilbert space, and $\mathbb{D} \in \mathcal{B}(\mathcal{U}_{\mathrm{out}}(0), \mathrm{L}^2(\mathbf{R}_+;Y))$.*

  Moreover, $\mathbb{D}$ is [positively] J-coercive iff $\mathbb{D}^*J\mathbb{D}$ is invertible [$\gg 0$] on $\mathcal{U}_{\mathrm{out}}(0)$.
- *(d1) $\mathbb{D}$ is positively J-coercive iff there is $\varepsilon > 0$ s.t. for all $u \in \mathrm{L}^2(\mathbf{R}_+;U)$ we have*
$$\langle \mathbb{D}u, J\mathbb{D}u\rangle \geq \varepsilon(\|u\|_2^2 + \|\mathbb{D}u\|_2^2) \tag{8.76}$$
- *(d2) Let $J \gg 0$. Then $\mathbb{D}$ is positively J-coercive iff $\|\mathbb{D}u\|_2 \geq \varepsilon\|u\|_2$ for some $\varepsilon > 0$ and all $u \in \mathrm{L}^2(\mathbf{R}_+;U)$.*
- *(d3) If $\mathbb{D}$ is J-coercive, then $\mathbb{D}$ is injective on $\mathrm{L}^2(\mathbf{R}_+;U)$.*
- *(d4) Let $\mathbb{D}$ be J-coercive. Then $\|J\mathbb{D}u\|_2 \geq \varepsilon\|u\|_{\mathbb{D}}$ for some $\varepsilon > 0$ and all $u \in \mathcal{U}_{\mathrm{out}}(0)$.*

Recall that $\mathcal{U}_{\mathrm{out}}(0) = \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,\big|\, \mathbb{D}u \in \mathrm{L}^2\}$, and that the [positive] J-coercivity of $\mathbb{D}$ (over $\mathcal{U}_{\mathrm{out}}$) depends only on $\mathbb{D}$ and J. Thus, we can define the [positive] J-coercivity of $\mathbb{N}$ analogously (we used this implicitly in (b1)). Consequently, $\mathbb{N}$ is [positively] J-coercive iff $\pi_+\mathbb{N}^*J\mathbb{N}\pi_+ \in \mathcal{GB}$ [$\gg 0$], by (a1) [(a2)].

Obviously, when $\mathbb{D} \in \mathrm{TIC}(U,Y)$, the space $\mathcal{U}_{\mathrm{out}}(0)$ equals $\mathrm{L}^2(\mathbf{R}_+;U)$ with an equivalent norm ($\|\cdot\|_{\mathcal{U}_{\mathrm{out}}}$). Contrary to (b1), we have no control on zeros of $\widehat{\mathbb{N}}$ if $\mathbb{N}$ and $\mathbb{M}$ are not required to be q.r.c. (e.g., take $\widehat{\mathbb{N}}(s) = s/(s+1) = \widehat{\mathbb{M}}(s)$, $\mathbb{D} = I$).

The condition $\mathbb{D}^*J\mathbb{D} \gg 0$ in (a2) holds iff $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+ \gg 0$ on $\mathrm{L}^2(\mathbf{R}_+;U)$, or equivalently, iff $\widehat{\mathbb{D}}^*J\widehat{\mathbb{D}} \geq \varepsilon I$ a.e. on $i\mathbf{R}$ (in $\mathrm{L}^\infty_{\mathrm{strong}}$). See Propositions 10.3.1 and 10.3.2 for further equivalent conditions for positive J-coercivity and Lemma 2.2.2 for the invertibility of $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$.

Proof of Lemma 8.4.11: Part (b1) follows from Lemma 6.5.6(f)&(a1)&(a2). Part (b2) follows from (a1)–(b1), because $\pi_+S\pi_+ \in \mathcal{G}\mathrm{L}^2(\mathbf{R}_+;U)$ iff $S \in \mathcal{GB}$. The rest follows from Lemma 8.2.3. □

For classes satisfying Hypothesis 8.4.7, the existence of a spectral factorization and the invertibility of the Popov Toeplitz operator (that is, the J-coercivity of $\mathbb{D}$) are equivalent:

Theorem 8.4.12 (MTI spectral factorization) *Let $\mathcal{A} \subset \mathrm{TIC}$ and $J = J^* \in \mathcal{B}(Y)$. For $\mathbb{D} \in \mathrm{TIC}(U,Y)$ we have (iv)⇒(iii)⇔(ii)⇒(i), where* *(i)* $\mathbb{D}$ *is J-coercive over* $\mathcal{U}_{\text{out}}$ *(i.e.., $\pi_+\mathbb{D}^*J\mathbb{D}\pi_+$ is invertible on* $\mathrm{L}^2(\mathbf{R}_+;U)$*);*

(ii) $\mathbb{D}^*J\mathbb{D} = \mathbb{X}^*S\mathbb{X}$ *for some* $\mathbb{X} \in \mathcal{G}\mathrm{TIC}(U)$, $S \in \mathcal{GB}(U)$*;*

(iii) $\mathbb{D}^*J\mathbb{D} = \mathbb{Y}^*\mathbb{X}$ *for some* $\mathbb{X}, \mathbb{Y} \in \mathcal{G}\mathrm{TIC}(U)$*;*

(iv) $\mathbb{D}^*J\mathbb{D} = \mathbb{X}^*S\mathbb{X}$ *for some* $\mathbb{X} \in \mathcal{GA}(U)$, $S \in \mathcal{GB}(U)$.

If $\mathbb{D} \in \mathcal{A}(U,Y)$ *and* $\mathcal{A}(U)$ *satisfies Hypothesis 8.4.7, then (i)–(iv) are equivalent and any spectral factorization of* $\mathbb{D}^*J\mathbb{D}$ *is over* $\mathcal{A}(U)$.

If $\mathbb{D}^*J\mathbb{D} \geq 0$*, then (i)⇔(ii)⇔(iii).*

In fact, in discrete-time, (i)–(iv) are equivalent for $\mathbb{D} \in \mathcal{A} := \mathrm{tic}_{\mathrm{exp}}$ (the set of exponentially stable discrete-time maps), by Theorem 14.3.2. Unfortunately, the Cayley images of $\widehat{\mathrm{tic}_{\mathrm{exp}}}$ cover only those continuous-time maps which are H^∞ outside some disc in the left half-plane, and that requirement is rather strong. By Example 8.4.13, (i) does not imply (ii) for general $\mathbb{D} \in \mathrm{TIC}$.

Proof: "(iii)⇔(ii)" holds by Lemma 6.4.7(a3). "(iv)⇒(ii)" is trivial and "(ii)⇒(i)" follows from $\pi_+\mathbb{X}^{-1}S^{-1}\mathbb{X}^{-*}\pi_+ = (\pi_+\mathbb{X}^*S\pi_+\mathbb{X}\pi_+)^{-1}$.

The missing implication (i)$\Longrightarrow$(iv) is contained in Hypothesis 8.4.7, and the last sentences follow from Lemma 6.4.5(a) Lemma 6.4.7(a). □

The solvability of several control problems implies the invertibility of the corresponding Popov Toeplitz operator (condition (i) above) and is implied by the existence of a spectral factorization of the Popov operator (condition (ii) above). Thus, the above equivalence makes all these equivalent. Even better, the regularity of $\mathcal{A}$ (see Hypothesis 8.4.7(1.)) makes a complete Riccati equation theory possible.

This is why we obtain complete solutions for the classes of Theorem 8.4.9, but only sufficient conditions (in terms of spectral factorizations and Riccati equations) in the general case (e.g., compare Theorem 11.3.3 to Proposition 11.3.4(f)). See Remark 9.9.14 for other classes of systems for which similar (even better) optimality, factorization and Riccati equation results can be established.

Ilya Spitkovsky has constructed an example showing that the invertibility of the Toeplitz operator (i.e., J-coercivity over $\mathcal{U}_{\text{out}}$) does not imply the existence of a (bounded) spectral factorization in the indefinite case, as mentioned in [S98c, Remark 4.8]. We give here an extended version of that example:

Example 8.4.13 ((minimax) *J*-coercive $\not\Rightarrow$ SpF) (We give this example for discrete time, use Cayley transform (see Lemma 13.2.1–Theorem 13.2.3) for the continuous-time counterpart.)

(a) Let $J := J_\gamma^{2,1} := \mathrm{diag}(1,1,-\gamma^2)$, where $\gamma := \sqrt{2}$. By Lemma 6.4.7(a), there is $h \in \mathcal{G}\mathrm{H}^\infty(\mathbf{D})$ s.t. $|h|$ equals $1/2$ on the left hemicircle and $\sqrt{3}/2$ on the right hemicircle. Set

$$\widehat{\mathbb{D}} := \begin{bmatrix} ih & h \\ h(-\cdot) & h(-\cdot) \\ 0 & 1 \end{bmatrix} \in \mathrm{H}^\infty(\mathbf{D};\mathbf{C}^{3\times 2}). \quad \text{Then} \tag{8.77}$$

$$\widehat{\mathbb{E}} := \widehat{\mathbb{D}}^*J\widehat{\mathbb{D}} = \begin{bmatrix} 1 & f \\ \overline{f} & -1 \end{bmatrix} \in \mathrm{L}^\infty(\partial\mathbf{D};\mathbf{C}^{2\times 2}), \tag{8.78}$$

where $f \in \mathrm{L}^\infty(\partial\mathbf{D};\mathbf{C})$ assumes exactly two values, $f_r := \frac{1}{4} - i\frac{3}{4}$ on the right hemicircle and $f_l := \frac{3}{4} - i\frac{1}{4}$ on the left hemicircle.

Therefore, $\widehat{\mathbb{E}}_{11} = 1 \gg 0$ and $\widehat{\mathbb{E}}_{22} - \widehat{\mathbb{E}}_{21}\widehat{\mathbb{E}}_{11}^{-1}\widehat{\mathbb{E}}_{12} = -1 - |f|^2 \ll 0$, so that $\mathbb{D}$ is minimax J-coercive, hence J-coercive over $\mathcal{U}_{\text{out}}$, by Lemma 11.4.2. By Theorem 9.15.3 and Lemma 9.15.2, there is a unique $\widehat{\mathbb{X}} \in \mathcal{G}\mathrm{H}^2(\partial\mathbf{D};\mathbf{C}^{2\times 2})$ modulo a constant s.t. $\widehat{\mathbb{X}}^* \left[\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right] \widehat{\mathbb{X}} = \widehat{\mathbb{E}}$ a.e. on $\partial\mathbf{D}$.

Let E_r and E_l be the two values of $\widehat{\mathbb{E}}$. Since the eigenvalues of $E_r^{-1}E_l$ are not positive (not even real), it follows that $\widehat{\mathbb{X}} \notin \mathcal{G}\mathrm{H}^\infty(\mathbf{D};\mathbf{C}^{2\times 2})$; in fact, one can show that both $\widehat{\mathbb{X}}$ and $\widehat{\mathbb{X}}^{-1}$ are unbounded. By uniqueness, there can be no ($\mathcal{G}\mathrm{H}^\infty$) spectral factorization of $\mathbb{D}^* J\mathbb{D}$.

(b) Furthermore, there is a minimax $\widetilde{J}$-coercive $\widehat{\mathbb{D}_0} \in \mathrm{H}^\infty(\mathbf{D};\mathbf{C}^{6\times 4})$, where $\widetilde{J} := J_\gamma^{4,2} := \begin{bmatrix} I_{4\times 4} & 0 \\ 0 & -2I_{2\times 2} \end{bmatrix}$, s.t. $\widehat{\mathbb{D}_0}$ is of form $\left[\begin{smallmatrix} * & * \\ 0 & I_{2\times 2} \end{smallmatrix}\right]$, and $\widehat{\mathbb{D}_0}^* \widetilde{J}\widehat{\mathbb{D}_0} = \widehat{\mathbb{X}_0}^* S\widehat{\mathbb{X}_0}$, where $S := J_1^{2,2} = \begin{bmatrix} I_{2\times 2} & 0 \\ 0 & -I_{2\times 2} \end{bmatrix}$, $\widehat{\mathbb{X}_0} \in \mathcal{G}\mathrm{H}^2(\mathbf{D};\mathbf{C}^4)$ and both $\widehat{\mathbb{X}_0}$ and $\widehat{\mathbb{X}_0}^{-1}$ are unbounded near -1.

Consequently, the corresponding continuous-time $\widetilde{J}$-critical "state feedback controller" over $\mathcal{U}_{\text{out}}$ (the H^∞ full-information minimax controller over $\mathcal{U}_{\text{out}}$) is non-well-posed (alternatively, unstable, by (c)) in both its open-loop and closed-loop forms, as shown in Example 11.3.7.

(c) Set $K := \overline{\mathbf{D}} \setminus \{z \,\big|\, |z \pm i| < \varepsilon\}$ for some $\varepsilon > 0$. Then there is a neighborhood $\Omega \supset \mathbf{D}$ of K s.t. $\widehat{\mathbb{D}}, \widehat{\mathbb{X}}, \widehat{\mathbb{X}}^{-1} \in \mathrm{H}^\infty(\Omega;\mathcal{B}(\mathbf{C}^2,*))$ in (a). In particular, the corresponding continuous-time maps (Cayley inverses) are uniformly half-plane regular.

Consequently, if we drop the rotation from the proof of (b), the maps $\mathbb{D}_0$, $\mathbb{X}_0$, $\mathbb{X}_0^{-1}$ become (well-posed and) uniformly half-plane-regular (but $\mathbb{X}_0^{\pm 1}$ become unstable, althought $\mathbb{X}_0^{\pm 1}\mathrm{L}^2_c \subset \mathrm{L}^2$). $\triangleleft$

The above example shows that the J-critical control is not always of state feedback form (in continuous time); cf. Remark 9.7.7(a3). (Due to minimax J-coercivity, the example shows that the problem cannot be avoided even in connection with H^∞ problems.)

In discrete-time that cannot happen, but the implication (i)$\Rightarrow$(iii) in Theorem 8.4.12 is nevertheless false in general in discrete-time too, as shown by the above example (see also Section 9.15 and Example 11.3.7) unless $\mathbb{D}$ is exponentially stable (in discrete time).

Proof of Example 8.4.13: (a) (Ilya Spitkovsky has skecthed the proof; this is a modified version of that sketch.)

1° *Constructing $\mathbb{D}$ s.t. (8.78) holds:* By Theorem 3.1.3(a1)&(e1), any $\widehat{\mathbb{F}} \in \mathrm{L}^\infty(i\mathbf{R})$ satisfying $\widehat{\mathbb{F}} \geq \varepsilon$ a.e. on $i\mathbf{R}$ corresponds to some $\mathbb{F} \in \mathrm{TI}(\mathbf{C})$ with $\mathbb{F} \geq \varepsilon I$, so that $\mathbb{F} = |\mathbb{Z}|^2$ for some $\mathbb{Z} \in \mathcal{G}\mathrm{TIC}(\mathbf{C})$ (i.e., some $\widehat{\mathbb{Z}} \in \mathrm{H}^\infty(\mathbf{C}^+)$), by Lemma 6.4.7(a). Apply Cayley transform to this result to show the existence of $h \in \mathrm{H}^\infty(\mathbf{D})$ (see the example above). (This also shows that h is invertible in H^∞, but we do not need the invertibility of h.) One easily verifies that the eigenvalues of

$$E_r^{-1}E_l = \frac{-1}{1+|f_r|^2}\begin{bmatrix} -1 - \underline{f_r}\overline{f_l} & f_r - \underline{f_l} \\ \overline{f_l} - f_r & -1 - f_r f_l \end{bmatrix} \tag{8.79}$$

are given by $\lambda_\pm := \mathrm{Re}\,t \pm \sqrt{(\mathrm{Re}\,t)^2 - 4(|t|^2+|s|^2)}$, where $t := -1 - f_r\overline{f_l}$, $s := f_r - f_l$, hence these values are not real.

2° *The other claims:* The other claims are explained in the example except the fact that that $\widehat{\mathbb{X}}, \widehat{\mathbb{X}}^{-1} \notin \mathrm{H}^\infty(\mathbf{D};\mathbf{C}^{2\times 2})$.

If we had $\widehat{\mathbb{X}} \in \mathcal{G}\mathrm{H}^\infty$, then the factorization would exist in all L^p spaces; however, it does not exist for $p = 2\pi/\arg\lambda_\pm$ (since $\lambda_\pm$ are not real; see [LS] for details).

Choose $z_0 \in \partial\mathbf{D}$ is s.t. $\widehat{\mathbb{X}}$ or $\widehat{\mathbb{X}}^{-1}$ is unbounded in each neighborhood of z_0. (N.B. $z_0 = \pm i$, because $\widehat{\mathbb{X}}$ and $\widehat{\mathbb{X}}^{-1}$ have holomorphic extensions around each point of $\overline{\mathbf{D}} \setminus \{\pm i\}$, by Lemma 9.15.5. For the same reason it seems that the Cayley inverse of the function $\widehat{\mathbb{D}_0}$ constructed below will not be (weakly) regular.)

Set $\widehat{\mathbb{X}}^{\mathrm{d}}(z) := \widehat{\mathbb{X}}(\overline{z})^*$ $(z \in \mathbf{D})$ (cf. Lemma 13.1.8). Obviously, $\widehat{\mathbb{X}}^{\mathrm{d}} \in \mathcal{G}\mathrm{H}^2(\mathbf{D};\mathbf{C}^{2\times 2})$ and

$$(\widehat{\mathbb{X}}^{-\mathrm{d}}(z))^* J_1^{-1} \widehat{\mathbb{X}}^{-\mathrm{d}}(z) = (\widehat{\mathbb{X}}(\overline{z})^* J_1 \widehat{\mathbb{X}}(\overline{z}))^{-1} = \widehat{\mathbb{E}}(\overline{z})^{-1} = \widehat{\mathbb{E}}(z)^{-1} \tag{8.80}$$

(since $\widehat{\mathbb{E}}(\overline{z}) = \widehat{\mathbb{E}}(z)$, because $z \mapsto \overline{z}$ maps the right and left hemicircles onto themselves).

Set $a := \sqrt{8/13}$, so that $a^*a = (1+|f|^2)^{-1} = 8/13$. One easily verifies that $\widehat{\mathbb{E}}^{-1} = a^*a\widehat{\mathbb{E}}$, hence $(a\widehat{\mathbb{X}}^*)J_1(a\widehat{\mathbb{X}}) = \widehat{\mathbb{E}}^{-1}$. We conclude from Lemma 6.4.5(a) that $a\widehat{\mathbb{X}} = E\widehat{\mathbb{X}}^{-\mathrm{d}} = E\widehat{\mathbb{X}}(\overline{\cdot})^{-*}$ for some $E \in \mathcal{G}\mathbf{C}^{2\times 2}$. Therefore, both $\widehat{\mathbb{X}}$ and $\widehat{\mathbb{X}}^{-1}$ are unbounded on $\mathbf{D}$ (one near z_0 and the other near $\overline{z_0}$ (at least)).

(b) 1° *Constructing $\widehat{\mathbb{D}_a}$ and $\widehat{\mathbb{X}_a}$:* Set $\widehat{\mathbb{F}}(z) := \widehat{\mathbb{D}}(-z)$, $\widehat{\mathbb{Z}}(z) := \widehat{\mathbb{X}}(-z)$ $(z \in \mathbf{D})$, so that $\widehat{\mathbb{F}}^* J \widehat{\mathbb{F}} = \widehat{\mathbb{Z}}^* J_1 \widehat{\mathbb{Z}}$. By (a)2°, $\widehat{\mathbb{X}_a} = \begin{bmatrix} \widehat{\mathbb{X}} & 0 \\ 0 & \widehat{\mathbb{Z}} \end{bmatrix} \in \mathcal{G}\mathrm{H}^2(\mathbf{D};\mathbf{C}^{4\times 4})$ and its inverse are both unbounded at $\pm z_0 = \pm i$. Obviously, $\widehat{\mathbb{X}_a}^* \begin{bmatrix} J_1 & 0 \\ 0 & J_1 \end{bmatrix} \widehat{\mathbb{X}_a} = \widehat{\mathbb{D}_a}^* J' \widehat{\mathbb{D}_a}$, where $\widehat{\mathbb{D}_a} := \begin{bmatrix} \mathbb{D} & 0 \\ 0 & \mathbb{F} \end{bmatrix}$, $J' := \begin{bmatrix} J & 0 \\ 0 & J \end{bmatrix}$, $J_1 := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

2° *Rotation to -1:* Replace $\widehat{\mathbb{D}_a}$ by $\widehat{\mathbb{D}_a}(i\cdot)$ (and hence $\widehat{\mathbb{X}_a}$ by $\widehat{\mathbb{X}_a}(i\cdot)$). Then both $\widehat{\mathbb{X}_a}$ and $\widehat{\mathbb{X}_a}^{-1}$ become unbounded near 1 and near -1. (Recall that (our) Cayley transform maps -1 to ∞; this is why this is important.)

3° *Constructing $\widehat{\mathbb{D}_0}, \widehat{\mathbb{X}_0}$:* Set

$$\widehat{\mathbb{D}_0} := T'\widehat{\mathbb{D}_a}T, \quad \widehat{\mathbb{X}_0} := T\widehat{\mathbb{X}_a}T, \text{ where} \tag{8.81}$$

$$T := \begin{bmatrix} I&0&0&0\\0&0&I&0\\0&I&0&0\\0&0&0&I \end{bmatrix} = T^* = T^{-1} \in \mathbf{C}^{4\times 4}, \quad T' := \begin{bmatrix} I&0&0&0&0&0\\0&I&0&0&0&0\\0&0&0&0&I&0\\0&0&0&I&0&0\\0&0&I&0&0&0\\0&0&0&0&0&I \end{bmatrix} \in \mathbf{C}^{6\times 6} \tag{8.82}$$

(note that $T\cdot$ (resp. $\cdot T$) permutes second and third rows (resp. columns) and $T'\cdot$ (resp. $\cdot T'$) permutes third and fifth rows (resp. columns)), so that

$$S := J_1^{2,2} := \mathrm{diag}(1,1,-1,-1) = T \begin{bmatrix} J_1 & 0 \\ 0 & J_1 \end{bmatrix} T, \quad \widetilde{J} = T'J'T', \quad \text{and} \tag{8.83}$$

$$\widehat{\mathbb{D}_0}^* \widetilde{J} \widehat{\mathbb{D}_0} = T\widehat{\mathbb{D}_a}^* J' \widehat{\mathbb{D}_a} T = T\widehat{\mathbb{X}_a}^* \begin{bmatrix} J_1 & 0 \\ 0 & J_1 \end{bmatrix} \widehat{\mathbb{X}_a} T = \widehat{\mathbb{X}_0}^* S \widehat{\mathbb{X}_0}. \tag{8.84}$$

Thus, $\mathbb{X}_0^* S \mathbb{X}_0$ is a spectral factorization of $\mathbb{D}_0^* \widetilde{J} \mathbb{D}_0$. Moreover, $\widehat{\mathbb{D}_0}$ is of form

$\left[\begin{smallmatrix} * & * \\ 0 & I_{2\times 2}\end{smallmatrix}\right]$ and

$$\widehat{\mathbb{D}_0}^* \widetilde{J} \widehat{\mathbb{D}_0} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} f & 0 \\ 0 & f \end{bmatrix} \\ \begin{bmatrix} \bar{f} & 0 \\ 0 & \bar{f} \end{bmatrix} & \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix}, \tag{8.85}$$

hence $\mathbb{D}_0$ is obviously minimax $\widetilde{J}$-coercive w.r.t. $\mathbf{C}^2 \times \mathbf{C}^2$ (see Definition 11.4.1). See Example 11.3.7 for the final claims. For clarity, we write out $\widehat{\mathbb{D}_0}$ (here $g := h(i\cdot)$, $G := h(-i\cdot)$):

$$\widehat{\mathbb{D}_0} = \begin{bmatrix} ig & 0 & g & 0 \\ G & 0 & G & 0 \\ 0 & ig & 0 & g \\ 0 & G & 0 & G \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8.86}$$

(c) 1° *Constructing* Ω*:* Since $|h|^2$ has a holomorphic extension (namely $1/2$ and $\sqrt{3}/2$ on two disjoint open sets) to a neighborhood of $K \cap \partial\mathbf{D}$, so has its spectral factor, that is, h (and its inverse h^{-1}), by Lemma 9.15.5.

Consequently, $\widehat{\mathbb{D}}$ and hence also $\widehat{\mathbb{X}}$ and $\widehat{\mathbb{X}}^{-1}$ have holomorphic extensions to a neighborhood of $K \cap \partial\mathbf{D}$. Let Ω be the union of that neighborhood and $\mathbf{D}$.

2° *The rest:* Take Cayley transforms to obtain that for any $\varepsilon > 0$, there is a compact set $K_2 \subset \mathbf{C}^- \cup \{z \in \mathbf{C} \,\big|\, |z \pm i| < \varepsilon\}$ s.t. $\widehat{\mathbb{D}}, \widehat{\mathbb{X}}, \widehat{\mathbb{X}}^{-1} \in \mathrm{H}^\infty(K_2^c; \mathcal{B}(\mathbf{C}^2, *))$. In particular, these functions are continuous at infinity, hence ULR, even uniformly half-plane-regular (to be exact, $\mathbb{X}$ and/or $\mathbb{X}^{-1}$ is not stable (at $\pm i$) but it is otherwise uniformly half-plane-regular). Consequently, $\widehat{\mathbb{D}_0}, \widehat{\mathbb{X}_0}^{\pm 1} \in \mathrm{H}^\infty(K_2^c; \mathcal{B}(\mathbf{C}^4, *))$. □

Next we state the unstable version of Theorem 8.4.12:

Corollary 8.4.14 (MTI (J,S)-inner r.c.f.) *Let* $\mathbb{D} \in \mathrm{TIC}_\infty(U,Y)$, $\mathcal{A} \subset \mathrm{TIC}$, $J = J^* \in \mathcal{B}(Y)$ *and* $\mathcal{U}_*^* = \mathcal{U}_{\mathrm{out}}$.

(a) If $\mathbb{D}$ *has a* $(J,*)$*-inner q.r.c.f.* $\mathbb{N}\mathbb{M}^{-1}$*, then* $\mathbb{D}$ *and* $\mathbb{N}$ *are J-coercive.*

Let, in addition, $\mathbb{D}$ *have a [q.]r.c.f.* $\mathbb{D} = \mathbb{N}'\mathbb{M}'^{-1}$*. Then we have the following:*

(b1) We have (iii)⇔(ii)⇒(i), where

(i) $\mathbb{N}'$ *is J-coercive (i.e.,* $\mathbb{D}$ *is J-coercive);*
(ii) $\mathbb{N}'^* J \mathbb{N}'$ *has a spectral factorization;*
(ii') $\mathbb{N}'^* J \mathbb{N}'$ *has a spectral factorization over* $\mathcal{A}$*;*
(iii) $\mathbb{D}$ *has a* $(J,*)$*-inner [q.]r.c.f.*
(iii') $\mathbb{D}$ *has a* $(J,*)$*-inner [q.]r.c.f.* $\mathbb{N}\mathbb{M}^{-1}$ *with* $\mathbb{N}, \mathbb{M} \in \mathcal{A}$*.*

(b2) If $\mathbb{N}', \mathbb{M}' \in \mathcal{A}$*, then (ii')⇒(iii')⇒(iii)⇔(ii)⇒(i).*

(b3) Assume that $\mathbb{N}', \mathbb{M}' \in \mathcal{A}$, and that $\mathcal{A}(U)$ satisfies Hypothesis 8.4.7.

Then (i)–(iii') are equivalent. Moreover, if some factorization $\mathbb{D} = \mathbb{N}\mathbb{M}^{-1}$ satisfies (ii) or (iii), then $\mathbb{N}, \mathbb{M} \in \mathcal{A}$, $\mathbb{D}, \mathbb{M}^{-1} \in$ ULR, and we can choose them so that $\widehat{\mathbb{M}}(+\infty) = I$.

(b4) Assume that Hypothesis 8.4.7 holds for $\mathcal{A}(U)$, and that (iii') holds. Set $\mathbb{X} := \mathbb{M}^{-1}$. Then $N^*JN = S$ and $D^*JD = X^*SX$.

(b5) If $J \geq 0$, then (i)⇔(ii)⇔(iii).

Theorem 9.9.10 gives further equivalent conditions in terms of Riccati equations. Example 8.4.13 shows that implication (i)⇒(iii) does not hold in general.

The above corollary will be applied in Chapters 9 and 11 to reduce inner coprime factorizations of unstable maps to spectral factorizations of their stabilized counterparts.

Note that, by (b3), any $(J,*)$-inner r.c.f. $\mathbb{D} = \mathbb{N}\mathbb{M}^{-1}$ with $\mathbb{N}, \mathbb{M} \in \mathcal{A}$ is actually a r.c.f. *over* $\mathcal{A}$ (i.e., also $\widetilde{\mathbb{X}}, \widetilde{\mathbb{Y}} \in \mathcal{A}$ for some stable $\widetilde{\mathbb{X}}, \widetilde{\mathbb{Y}}$ satisfying $\widetilde{\mathbb{X}}\mathbb{M} - \widetilde{\mathbb{Y}}\mathbb{N} = I$).

If $\mathbb{N}$ and $\mathbb{M}$ are exponentially stable discrete-time maps (cf. the remark after Theorem 8.4.12), then (i)–(iii) are equivalent; see Corollary 14.3.3.

Proof of Corollary 8.4.14: (a) This follows from "(iii)⇒(i)" of (b1) (take $\mathbb{N}' = \mathbb{N}$, $\mathbb{M}' = \mathbb{M}$) and Lemma 8.4.11(b1).

(b1) We have (ii')⇒(ii)⇒(i) (even (ii')⇔(ii)⇔(i) provided Hypothesis 8.4.7 is satisfied), by Theorem 8.4.12.

"(ii)⇔(iii)": If (ii) holds, i.e., $\mathbb{U}^*S\mathbb{U} = \mathbb{N}'^*J\mathbb{N}'$ for some $\mathbb{U} \in \mathcal{G}\mathrm{TIC}(U)$, $S \in \mathcal{G}\mathcal{B}(U)$, then $\mathbb{N}'\mathbb{U}^{-1}$ is (J,S)-inner; thus, $\mathbb{D} = (\mathbb{N}'\mathbb{U}^{-1})(\mathbb{M}'\mathbb{U}^{-1})^{-1}$ is as in (iii).

Conversely, if $\mathbb{D} = \mathbb{N}\mathbb{M}^{-1}$ is a $(J,*)$-inner [q.]r.c.f., then, by Lemma 6.4.5(c), $\mathbb{N}' = \mathbb{N}\mathbb{U}$ for some $\mathbb{U} \in \mathcal{G}$TIC, hence then $\mathbb{N}'^*J\mathbb{N}' = \mathbb{U}^*\mathbb{N}^*J\mathbb{N}\mathbb{U} = \mathbb{U}^*S\mathbb{U}$ is a spectral factorization of $\mathbb{N}'$.

(b2) Implication "(ii')⇒(iii')" follows from the proof of "(ii)⇔(iii)". Obviously, (iii')⇒(iii); the rest follows from (b1).

(b3) By Theorem 8.4.12, we have (i)⇔(ii'), hence all five claims are equivalent, by (b2). Theorem 8.4.12 also implies that necessarily $\mathbb{U} \in \mathcal{G}\mathcal{A}$ in (ii), which provides the $\mathbb{N}, \mathbb{M} \in \mathcal{A}$ claim, by the proof of "(ii)⇔(iii)" above.

Because $\mathcal{A} \subset$ ULR, we have $\mathbb{M}^{-1}, \mathbb{N}\mathbb{M}^{-1} \in$ ULR, and $M := \widehat{\mathbb{M}}(+\infty) \in \mathcal{G}\mathcal{B}$, by Lemma 6.2.5 and Proposition 6.3.1(c). Therefore, we can take $M = I$, by Lemma 6.4.5(a).

(b4) This follows from Lemma 6.3.6(b).

(b5) This follows from (b1) and the last claim in Theorem 8.4.12. □

Notes for Section 8.4

The Popov Toeplitz invertibility condition of Proposition 8.3.10 and Lemma 8.4.11(a1) is very common in the literature of stable control problems [WW] [S98c]. Our definition of J-coercivity generalizes this concept to general control problems for WPLSs. In a sense, it is the weakest assumption that guarantees the existence of a unique solution (with equivalence for smooth systems with $\mathcal{U}_*^* = \mathcal{U}_{\exp}$ and $\dim U < \infty$ or $D^*JD \in \mathcal{G}\mathcal{B}(U)$); see Theorems 9.2.16 and 10.2.11 and Corollaries 10.2.3 and 9.2.19. However there are singular counterexamples at least when U is infinite-dimensional (e.g., take $C = 0$, $D > 0$, $D \not\gg 0$, $J = I$). Singular control problems are rare in practice and in the literature, and they cannot be solved as satisfactorily; see, e.g., [Stoorvogel] for the finite-dimensional case.

Propositions 10.3.1 and 10.3.2 provide several conditions that are equivalent to J-coercivity. Many popular assumptions in the literature are special cases of these assumptions [ZDG] [Keu] and most others are stronger [FLT]. We have seen nothing similar in the indefinite unstable case.

The name "J-coercivity" is from [S97b]–[S98c], where it means the invertibility of the Popov Toeplitz operator for a stable system, and for a jointly stabilizable and detectable system it means the J-coercivity of the corresponding stabilized subsystem. From Lemma 8.4.11(a1)–(b1) and Theorem 6.6.28 we observe that these definitions are special cases of that of ours.

In the stable case, the method of Theorem 8.4.3 (as the first part of Proposition 8.3.10) is very old. The same holds for the reduction of unstable $\mathcal{U}_{\text{exp}}$ problems to the stable case, as in Theorem 8.4.5; its coprimeness method for $\mathcal{U}_{\text{sta}}$ has been used in [S98b] and in [S98c, Section 7] for jointly stabilizable and detectable systems. See Chapter 5 for notes on the spectral factorization results on which Theorems 8.4.9 and 8.4.12 are based.

Lemma 8.4.11(a1)–(b1) are essentially from [S98c]. Part of Example 8.4.13 was mentioned in [S98c, Remark 4.8], and its original form is due to Ilya Spitkovsky in a communication with Joseph Ball and Olof Staffans.

8.5 Problems on a finite time interval

Lord of the far horizons,
Give us the eyes to see
Over the verge of the sundown
The beauty that is to be.

— Bliss Carman (1861–1929)

In this section, we swiftly review how the abstract optimization of Sections 8.1–8.2 can be applied to finite-horizon problems. The derivation of further details and differential Riccati equation theory is analogous to that in the infinite-horizon case, but it requires a lengthy treatment, hence we omit it.

Throughout this section, the letters U, H, and Y denote (complex) Hilbert spaces of arbitrary dimensions, $T > 0$, $J = J^* \in \mathcal{B}(Y)$ and $\left[\begin{smallmatrix}\mathbb{A} & \mathbb{B}\\ \mathbb{C} & \mathbb{D}\end{smallmatrix}\right] \in \mathrm{WPLS}(U,H,Y)$. In finite time interval (finite-horizon) problems, the cost to be optimized is given by

$$\mathcal{J}(x_0,u) = \int_0^T \langle y(t), Jy(t)\rangle_Y \, dt, \tag{8.87}$$

(as opposed to $T = \infty$ in other sections, i.e., in infinite-horizon problems), where $x_0 \in H$ is the initial state, $u \in \mathrm{L}^2([0,T);U)$ is the control and $y := \mathbb{C}x_0 + \mathbb{D}u$ is the corresponding output. Often one wishes to add to $\mathcal{J}$ an end cost $\langle x(T), Gx(T)\rangle_H$, where $x(T) := \mathbb{A}(T)x_0 + \mathbb{B}\tau(T)u$ is the (terminal) state at time T.

The abstract optimization theory can be applied also to finite-horizon problems by using the following substitutions:

Remark 8.5.1 (Sections 8.1–8.2 apply also to finite-horizon problems) *Standing Hypothesis 8.1.1 is satisfied with substitutions*

$$U \mapsto \mathrm{L}^2([0,T);U), \qquad X \mapsto H \tag{8.88}$$

$$Y, Y^{\mathrm{s}} \mapsto \mathrm{L}^2([0,T);Y), \qquad Z, Z^{\mathrm{s}} \mapsto \mathrm{L}^2([0,T);H), \tag{8.89}$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \mapsto \left[\begin{array}{c|c} \pi_{[0,T)}\mathbb{A} & \pi_{[0,T)}\mathbb{B}\tau\pi_{[0,T)} \\ \hline \pi_{[0,T)}\mathbb{C} & \pi_{[0,T)}\mathbb{D}\pi_{[0,T)} \end{array}\right] \qquad J \mapsto J. \tag{8.90}$$

It follows that $\mathcal{J}(x_0,u) = \int_0^T \|y(t)\|_Y^2 \, dt$, where $y := \mathbb{C}x_0 + \mathbb{D}\tau u$, for all $u \in \mathcal{U}(x_0) = \mathrm{L}^2([0,T);U)$ and $x_0 \in H$.

To add an end cost $\langle x(T), Gx(T)\rangle_H$, $G \in \mathcal{B}(H)$, one could substitute $C \mapsto \begin{bmatrix}\pi_{[0,T)}\mathbb{C} \\ \mathbb{A}^T\end{bmatrix}$, $D \mapsto \begin{bmatrix}\pi_{[0,T)}\mathbb{D}\pi_{[0,T)} \\ \mathbb{B}^T\end{bmatrix}$, $J \mapsto \left[\begin{smallmatrix}J & 0\\ 0 & G\end{smallmatrix}\right]$, $Y, Y^{\mathrm{s}} \mapsto \mathrm{L}^2([0,T);Y) \times H$.

In particular, “the stable case” applies (in both cases, although we refer below to the case $G = 0$). (The substitution $\mathcal{U}(x_0) = \mathrm{L}^2([0,T);U)$ above corresponds to “$\mathcal{U}_{\mathrm{exp}} = \mathcal{U}_{\mathrm{str}} = \mathcal{U}_{\mathrm{sta}} = \mathcal{U}_{\mathrm{out}}$”.)

Assume that “D^*JD”$= \pi_{[0,T)}\mathbb{D}^* J \pi_{[0,T)}\mathbb{D}\pi_{[0,T)}$ is invertible (this is the case if, e.g., $J \geq 0$ and there is a quadratic cost on u). Then the system is J-coercive, so that Lemma 8.2.8 provides us direct formulae for $\mathbb{A}_{\mathrm{crit},T}$, $\mathbb{C}_{\mathrm{crit},T}$, $\mathbb{K}_{\mathrm{crit},T}$, and “$\mathcal{P} = C^*JC_{\mathrm{crit}}$”.

In particular, $\mathcal{P}$ becomes (we shall set $\mathcal{P}_T := \mathcal{P}$ to distinguish between cost operators for problems on different intervals, i.e., for different values of T)

$$\langle x_1, \mathcal{P}_T x_0 \rangle = \int_0^T \langle J(\mathbb{C}x_1)(t), ((\mathbb{C}+\mathbb{D}\mathbb{K}_{\mathrm{crit},T})x_0)(t)\rangle_H \, dt. \tag{8.91}$$

Assume, in addition, that $C \in \mathcal{B}(H,U)$, $D^*JC = 0$. Then, as in the proof of Lemma 8.2.9, we have

$$\langle x_1, \mathcal{P}_T x_0 \rangle = \int_0^T \langle C\mathbb{A}^t x_1, JC\mathbb{A}^t_{\mathrm{crit},T} x_0 \rangle_H \, dt, \tag{8.92}$$

and $\mathbb{A}_{\mathrm{crit}} = \mathbb{A} + \mathbb{B}\tau\mathbb{K}_{\mathrm{crit}} \in \mathcal{B}(H, \mathcal{C}([0,T];H))$.

Again by Lemma 8.2.9, with the additional assumption that $D^*JD \in \mathcal{GB}(U)$, we also have

$$\mathbb{K}_{\mathrm{crit},T} x_0 = -(D^*JD)^{-1}(\pi_{[0,T)}\mathbb{B}\tau\pi_{[0,T)})^* C^*JC\mathbb{A}_{\mathrm{crit},T} x_0. \tag{8.93}$$

It seems that one can rewrite Sections 9.7 and 9.2 for finite time; in particular, that a unique J-critical control is of the state feedback form whenever $\mathbb{D} \in \mathrm{MTIC}^{\mathrm{L}^1}_{\infty}$ and C is bounded, or $\mathbb{D} \in \mathcal{B} + \mathrm{L}^2_{\infty}*$ (we do not have to pose any requirements on $\mathbb{A}B$ (or $\mathbb{B}\tau$), since on a finite time horizon we always have "the stable case", as in Hypothesis 9.2.2(6.)–(7.)), and that the J-critical state feedback operator corresponds then to a unique solution of the differential Riccati equation.

At least when $\mathcal{J}(x_0,u) = \|y\|^2$, $y = \left[\begin{smallmatrix} y_1 \\ u \end{smallmatrix}\right]$, the solution converges strongly to the infinite-horizon solution, as $T \to +\infty$ (possibly we have to restrict T to a sequence converging to infinity), and $\mathcal{P}_T \to \mathcal{P}$ strongly. This follows by standard arguments (see Lemma 4.2 of [FLT] in the case of a bounded C; the same arguments combined with Lemma A.3.1(i1)&(i4) also apply to a general WPLS).

The situation is analogous also in the case of an extended linear system (see Section 8.6).

However, we leave the details to the reader, because it seems that this approach does not provide us with any additional information $\mathcal{P}$ (e.g., we do not know whether $\mathcal{P}$ converges in the way that $\mathcal{P}[H] \subset \mathrm{Dom}(B_{\mathrm{w}}^*)$ even for the special $\mathbb{D}$'s mentioned above). This is an interesting open problem.

Notes

It is well-known that, under certain regularity (see, e.g., [GL] for the finite-dimensional and [Jacob98], [Jacob01], [Jacob99], [FLT] or [LT00b] for the infinite-dimensional case), 1. one can show that $\mathcal{P}_T \in \mathcal{B}(H, \mathrm{Dom}(B_{\mathrm{w}}^*))$ and $(\mathbb{K}_{\mathrm{crit},T} x_0)(s) = -B_{\mathrm{w}}^* \mathcal{P}_{T-s} \mathbb{A}^s_{\mathrm{crit},T} x_0$, i.e., that $u_{\mathrm{crit}} = -B_{\mathrm{w}}^* \mathcal{P}_{T-s} x_{\mathrm{crit}}$; moreover, $\mathcal{P}_T$ satisfies the corresponding differential Riccati equation (as a function of $T > 0$); 2. by letting $T \to +\infty$, we obtain that $\mathcal{P}_T$, u_{crit}, x_{crit} and y_{crit} converge (for a suitable cost function; for our techniques one sufficient condition seems to be that $J \geq 0$).

Analogously, one could apply the abstract optimization theory of Sections 8.1–8.2 to time-variant systems. Birgit Jacob [Jacob98] has formulated "tv-systems", which are a natural (time-variant) generalization of WPLSs, and solved the standard optimal control problem on a finite time interval (the same is done for a generalization of PS-systems in [Jacob01] and for a generalization of WPLSs with bounded input and output operators in [Jacob99]; thus, [Jacob98] is the most general one). Her solution coincides with the implications of Lemma 8.2.8 combined with an equivalent of Remark 8.5.1 for tv-systems, but assuming the equivalent of weak regularity (with $D^*JC = 0$), she obtains the formula $(\mathbb{K}_{\mathrm{crit},T}x_0)(t) = -B_{\mathrm{w}}^*\mathcal{P}_{T-t}\mathbb{A}_{\mathrm{crit},T}^t x_0$ (where B_{w}^* is allowed to be time-varying). In [Jacob99], Jacob has solved the finite-horizon minimization problem for a nonstandard (general) cost function for time-variant systems but under the assumption that the input and output operators are bounded. For time-invariant WPLSs with bounded output operators, the nonstandard minimization problem is solved in [BP] in terms of an integral Riccati inequality, by Francesca Bucci and Luciano Pandolfi.

Historical remarks on the LQR problem are given in the notes to Chapter 6 of [CZ]; also finite-horizon problems are treated therein.

Over a finite time horizon we have "$\mathcal{U}_{\mathrm{exp}} = \mathcal{U}_{\mathrm{out}}$", because $\mathrm{L}^2_{\mathrm{loc}}([0,T];H) = \mathrm{L}^2([0,T];H)$, hence the minimization results in something resembling more $\mathcal{U}_{\mathrm{out}}$ than $\mathcal{U}_{\mathrm{exp}}$ but having the nicest properties of both settings.

In particular, if we let $T \to +\infty$, we end up with a minimizing (hence J-critical) control over $\mathcal{U}_{\mathrm{out}}$ (not over $\mathcal{U}_{\mathrm{exp}}$, and the closed-loop system need not be exponentially stable), and the state feedback formula $u = -B_{\mathrm{w}}^*\mathcal{P}_T$ converges to $u = -B_{\mathrm{w}}^*\mathcal{P}$ (or to some other operator, denoted by "$-B^*\mathcal{P}$" in [FLT], if we reduce regularity assumptions), where $\mathcal{P}$ refers to the J-critical cost operator over $\mathcal{U}_{\mathrm{out}}$. Naturally, this can be fixed by assuming the system to be estimatable (see Lemma 8.3.3).

It seems that the methods of this monograph could be generalized to tv-systems. This might be an interesting subject for future studies.

8.6 Extended linear systems (ELS)

Human beliefs, like all other natural growths, elude the barrier of systems.

— George Eliot (1819–1880)

In this section, we give guidelines on how to extend our optimization and Riccati equation results for more general systems than WPLSs. Most readers probably wish to skip this section (the results are not used elsewhere in this monograph).

The optimization results given above, such as Theorem 8.3.9 (often combined with Theorem 8.4.3) require that $\mathcal{U}_*^*(x_0) \neq \emptyset$ for all $x_0 \in H$. What if this holds only for all $x_0 \in H_{1/2}$, for some Hilbert space $H_{1/2}$ s.t. $H_1 \underset{c}{\subset} H_{1/2} \underset{c}{\subset} H$? If "$B \in \mathcal{B}(U, H_{1/2-1})$", then we could replace H by $H_{1/2}$ and go on — except that $\mathbb{B}$ need not be well-posed anymore.

Despite the generality of WPLSs, there are some interesting PDE-based systems for which there are no known finite cost condition results for any choice of state space that would make $\mathbb{B}$ well-posed, as noted in [LT93]. Therefore, in this section, we shall give guidelines on how to relax the well-posedness requirement of $\mathbb{B}$.

For systems mentioned above, to guarantee the finite cost condition the state space must be chosen to be very small (this causes the high unboundedness of B), and hence the output operator becomes bounded; therefore we assume this, i.e., that $C \in \mathcal{B}(H, Y)$, although basic results of this section hold for general well-posed C too (see Remark 8.6.8).

Thus, we are again solving the problem "$x' = Ax + Bu$, $y = Cx + Du$, $x(0) = x_0$", but this time we do not require $\mathbb{B}$ to be well-posed, i.e., to have values in H (equivalently, to satisfy $\mathbb{B}\pi_{[-t,0)} \in \mathcal{B}(\mathrm{L}^2([-t,0);U),H)$). The solution $\Sigma := \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline \mathbb{C} & \mathbb{D} \end{array}\right] : \left[\begin{smallmatrix} x_0 \\ u \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ is called an *Extended Linear System (ELS)* under the assumptions stated below:

Standing Hypothesis 8.6.1 *Throughout this section, we shall assume that U, H and Y are Hilbert spaces, that $\mathbb{A}$ is a C_0-semigroup on H, $B \in \mathcal{B}(U, H_{-1})$, $C \in \mathcal{B}(H, Y)$, $D \in \mathcal{B}(U, Y)$, $J = J^* \in \mathcal{B}(Y)$ and* $0 \le \beta > \omega_A$.

By Lemma 6.3.16, $\left(\begin{smallmatrix} A \\ C \end{smallmatrix}\right)$ generate $\left[\begin{smallmatrix} \mathbb{A} \\ \mathbb{C} \end{smallmatrix}\right] \in \mathrm{WPLS}_\beta(\{0\}, H, Y)$ (this is trivial, with $\mathbb{C} = C\mathbb{A} : H \to \mathcal{C}(\mathbf{R}_+; U)$) and $\left(\begin{array}{c|c} A & B \\ \hline I & 0 \end{array}\right)$ generate $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline \mathbb{A} & \mathbb{B}\tau \end{array}\right] \in \mathrm{WPLS}_\beta(U, H_{-1}, H_{-1})$. Thus, we can and will have this section based heavily on WPLS results and on the abstract optimization theory Sections 8.1 and 8.2.

Moreover, $\mathbb{B}\tau$ maps $\mathrm{L}^2_\beta + \pi_+ \mathrm{L}^2_{\mathrm{loc}}$ to $\mathcal{C}(\mathbf{R}; H_{-1})$ (because $\pi_- \tau u \in \mathcal{C}(\mathbf{R}; \mathrm{L}^2_\beta)$ and $\mathbb{B} \in \mathcal{B}(\mathrm{L}^2_\beta; H_{-1})$) and $\mathrm{W}^{1,2}_\beta(\mathbf{R}; U) \to \mathcal{C}(\mathbf{R}; H)$ (by Theorem 6.2.13(b1), because $(H_{-1})_B = H$).

We now define some spaces of allowable inputs u with graph norms:

Definition 8.6.2 *For all $T \in [0,\infty)$ we define*

$$\widetilde{U}_{\text{loc}} := \{u \in \mathrm{L}^2_{\text{loc}}(\mathbf{R};U) \,|\, \mathbb{B}\tau u \in \mathrm{L}^2_{\text{loc}}(\mathbf{R};H)\}, \tag{8.94}$$

$$\mathcal{U}_{\exp}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,|\, \mathbb{A}x_0 + \mathbb{B}\tau u \in \mathrm{L}^2(\mathbf{R}_+;H)\} \quad (x_0 \in H), \tag{8.95}$$

$$\mathcal{U}^{\mathcal{C}}_{\exp}(x_0) := \{u \in \mathrm{L}^2(\mathbf{R}_+;U) \,|\, \mathbb{A}x_0 + \mathbb{B}\tau u \in \mathrm{L}^2(\mathbf{R}_+;H) \cap \mathcal{C}(\mathbf{R}_+;H)\} \quad (x_0 \in H), \tag{8.96}$$

$$\widetilde{U}_\beta := \{u \in \mathrm{L}^2_\beta(\mathbf{R};U) \,|\, \mathbb{B}\tau u \in \mathrm{L}^2_\beta(\mathbf{R};H)\}, \tag{8.97}$$

$$\widetilde{U}_{[0,T)} := \{u \in \mathrm{L}^2([0,T);U) \,|\, \mathbb{B}\tau u \in \mathrm{L}^2([0,T);H),\ \mathbb{B}\tau^T u \in H\}, \tag{8.98}$$

$$\|u\|^2_{\widetilde{U}_\beta} := \|u\|_2^2 + \|\mathbb{B}\tau u\|^2_{\mathrm{L}^2_\beta}, \tag{8.99}$$

$$\|u\|^2_{\widetilde{U}_{[0,T)}} := \|u\|_2^2 + \|\mathbb{B}\tau u\|^2_{\mathrm{L}^2([0,T);H)} + \|\mathbb{B}\tau^T u\|_H^2, \tag{8.100}$$

$$\mathbb{D}u := (C\mathbb{B}\tau + D)u \in \mathrm{L}^2_{\text{loc}}(\mathbf{R};Y) \quad (u \in \widetilde{U}_{\text{loc}}). \tag{8.101}$$

As for WPLSs, we extend $\mathbb{B}$ to $\mathcal{B}(\mathrm{L}^2_\omega(\mathbf{R};U), H_{-1})$ for those ω for which $\mathbb{B}$ is continuous; note that $\mathrm{L}^2_{\omega'}(\mathbf{R}_-;U) \underset{c}{\subset} \mathrm{L}^2_\omega$ for all $\omega' > \omega$. (Thus, $u \in \widetilde{U}_{\text{loc}}$ implies that $\pi_- u \in \mathrm{L}^2_\omega$ for some such ω.)

Obviously, $\mathbb{D}$ is time-invariant and causal $\mathrm{Dom}(\mathbb{D}) \to \mathrm{L}^2_{\text{loc}}$, and $\mathrm{W}^{1,2}_\beta(\mathbf{R};U) \to \mathcal{C}(\mathbf{R};H)$, although $\mathrm{L}^2_\beta \setminus \mathrm{Dom}(\mathbb{D})$ may be nonempty (and $\mathbb{D} \not\in \mathrm{TIC}_\infty$ is possible; take, e.g., $U = H$, $C = I$, $B = A$ s.t. $\widehat{\mathbb{D}}(s) = \widehat{\mathbb{B}\tau}(s) = (s-A)^{-1}A = s(s-A)^{-1} - I$ is unbounded on each right half-plane (i.e., A must be non-analytic; see Section 9.4).

Now we list the basic properties of ELSs:

Lemma 8.6.3 *Let $x_0 \in H$ and $T \in \mathbf{R}_+$. Then*

(a1) $\mathbb{B}\tau \in \mathcal{B}(\widetilde{U}_\beta, \mathrm{L}^2_\beta(\mathbf{R};H))$, $\mathbb{D} \in \mathcal{B}(\widetilde{U}_\beta, \mathrm{L}^2_\beta(\mathbf{R};Y))$, *and* $\mathbb{C}x_0 + \mathbb{D}u \in \mathrm{L}^2(\mathbf{R}_+;Y)$ *for all* $u \in \mathcal{U}_{\exp}(x_0)$.

(b1) $\left[\begin{smallmatrix}\mathbb{A}\\ \mathbb{C}\end{smallmatrix}\right] \in \mathrm{WPLS}_\beta(\{0\}, H, Y)$ *and* $\left[\begin{smallmatrix}\mathbb{A} & \mathbb{B}\\ \mathbb{A} & \mathbb{B}\tau\end{smallmatrix}\right] \in \mathrm{WPLS}_\beta(U, H_{-1}, H_{-1})$.

(b2) $\pi_+\mathbb{D}\pi_- = \mathbb{C}\mathbb{B}$ *on* $\widetilde{U}_\beta$.

(c1) $\mathbb{B}\tau \in \mathcal{B}(\mathrm{W}^{n+1,2}_\omega(\mathbf{R};U), \mathrm{W}^{n,2}_\omega(\mathbf{R};H))$, *hence* $\mathbb{B}\tau : \mathrm{W}^{n+1,2}_\omega(\mathbf{R};U) \to \mathcal{C}^n(\mathbf{R};H)$, *and* $\|(\mathbb{B}\tau u)^{(n)}(t)\|_H \le M e^{\omega t}\|u\|_{\mathrm{W}^{n+1,2}}$.

(c2) $\mathbb{C}x_0 \in \mathcal{C}(\mathbf{R}_+;Y)$ *($\in \mathcal{C}^k(\mathbf{R}_+;Y)$ if $x_0 \in H_k := \mathrm{Dom}(A^k)$).*

(d) $\widehat{\mathbb{B}\tau u}(s) = (s-A)^{-1}B\hat{u}(s)$, $\widehat{\mathbb{C}x_0}(s) = C(s-A)^{-1}x_0$, *and* $\widehat{\mathbb{D}u}(s) = (D + C(s-A)^{-1}B)\hat{u}(s)$ *(if $\mathbb{B}\tau u \in \mathrm{L}^2_{\text{loc}}$) for all $s \in \mathbf{C}^+_{\omega_A}$, and $u \in \mathrm{L}^2_\omega(\mathbf{R}_+;U)$, $\omega < \mathrm{Re}\, s$.*

(e) When $\omega > \mathrm{Re}\, s > \omega_A$ and $u_0 \in U$, we have

$$\mathbb{B}\tau\pi_+ e^{s\cdot} u_0 = (e^{s\cdot} - \mathbb{A})(s-A)^{-1}Bu_0 \in \mathcal{C}(\mathbf{R}_+;H) \cap \mathrm{L}^2_\omega, \tag{8.102}$$

$$\mathbb{D}\pi_+ e^{s\cdot} u_0 = e^{s\cdot}\widehat{\mathbb{D}}(s)u_0 - \mathbb{C}(s-A)^{-1}Bu_0 \in \mathcal{C}(\mathbf{R}_+;Y) \cap \mathrm{L}^2_\omega. \tag{8.103}$$

In particular, $\pi_+ e^{s\cdot} u_0 \in \widetilde{U}_{[0,T)}$.

(f) $\pi_{[0,T)}\mathcal{U}^{\mathcal{C}}_{\exp}(x_0) \subset \widetilde{U}_{[0,T)} \underset{c}{\subset} \widetilde{U}_\beta \subset \widetilde{U}_{\text{loc}}$ *and* $\widetilde{U}_{[0,T)}$ *and* $\widetilde{U}_\beta$ *are Hilbert spaces.*

Proof: (a1) Obviously, $\mathbb{B}\tau \in \mathcal{B}$; hence $\mathbb{D} := C\mathbb{B}\tau + D \in \mathcal{B}$. The last claim is also obvious.

(b1) This was noted above.

(b2) Now $(\pi_+ \mathbb{D}\pi_- u)(t) = \pi_+ C\mathbb{B}\tau^t \pi_- u = \pi_+ C\mathbb{A}^t \mathbb{B}u = (\mathbb{C}\mathbb{B}u)(t)$, by 2. and 3. of Definition 6.1.1.

(c1) This follows from Theorem 6.2.13(c2) and Lemma 6.3.19 (because now $(H_{-1})_B = H$, since $B \in \mathcal{B}(U, H_{-1})$, i.e., B is bounded to H_{-1}).

(c2) This follows from Theorem 6.2.13(c1).

(d) This follows from (b1) and the fact that $\mathbb{D} := C\mathbb{B}\tau + D$.

(e) This follows from (b1), Lemma 6.2.10, and the fact that $\mathbb{D} := C\mathbb{B}\tau + D$.

(f) Let $x_0 \in H$ and $T \geq 0$. Because $\mathbb{A}x_0 \in \mathcal{C} \cap \mathrm{L}^2_\beta$, we have $\pi_{[0,T)} \mathcal{U}^{\mathcal{C}}_{\exp}(x_0) \subset \widetilde{U}_{[0,T)}$. Obviously, $\widetilde{U}_\beta \subset \widetilde{U}_{\text{loc}}$.

Assume that $u \in \widetilde{U}_{[0,T)}$. Set $x(t) := \mathbb{B}^t u = \mathbb{B}^t u$ $(t \geq 0)$. Then

$$x(T+t) = \mathbb{B}\tau^{T+t} u = \mathbb{B}\tau^t \pi_- \tau^T u = \mathbb{A}^t \mathbb{B}\tau^T u = \mathbb{A}x(T) \quad (t \geq 0). \tag{8.104}$$

hence $\|x(T+\cdot)\|_{\mathrm{L}^2_\beta} \leq M\|x(T)\|_H$, where $M := \|\mathbb{A}\|_{\mathcal{B}(H, \mathrm{L}^2_\beta(\mathbf{R};H))}$. Therefore, $\widetilde{U}_{[0,T)} \underset{c}{\subseteq} \widetilde{U}_\beta$. □

The fact that $\left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \end{array}\right]$ and $\left[\begin{smallmatrix} \mathbb{A} \\ \mathbb{C} \end{smallmatrix}\right]$ are WPLSs offers us a wide range of additional useful facts. Note that if $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ generate a WPLS (i.e., $\mathbb{B} \in \mathcal{B}(\mathrm{L}^2([-\varepsilon,0);U),H)$, $\mathbb{D} \in \mathcal{B}(\mathrm{L}^2([0,\varepsilon);U), \mathrm{L}^2([0,\varepsilon);Y))$ for some $\varepsilon > 0$), then all above definitions coincide with corresponding WPLS definitions.

Remark 8.6.4 *Substituting* $X \mapsto H$, $U \mapsto \widetilde{U}_\beta$, $Y^{\mathrm{s}}, Z^{\mathrm{s}} \mapsto \mathrm{L}^2$, $Y, Z \mapsto \mathrm{L}^2_\beta$ *and* $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] \mapsto \left[\begin{smallmatrix} \mathbb{A} & \mathbb{B}\tau \\ \mathbb{C} & \mathbb{D} \end{smallmatrix}\right]$, *(cf. Remark 8.3.4), we obtain that* $\mathcal{U} = \mathcal{U}_{\exp}$ *and that the results of Sections 8.1 and 8.2 become applicable.* □

For example, the finite cost condition together with J-coercivity imply the existence of a unique J-critical control, and $\mathcal{J}(0,\cdot) \geq 0$ implies that a J-critical control is minimizing. Also most of Section 8.3 and much of further WPLS theory remain valid.

To give some examples, we mention explicitly below two more results from Sections 8.1 and 8.2:

Theorem 8.6.5 (LQR) *Assume that* $\mathcal{U}_{\exp}(x_0) \neq 0$ *for all* $x_0 \in H$. *Let* $C = \left[\begin{smallmatrix} I \\ 0 \end{smallmatrix}\right] \in \mathcal{B}(H, H \times U)$, $D = \left[\begin{smallmatrix} 0 \\ I \end{smallmatrix}\right]$, $J = \left[\begin{smallmatrix} Q & 0 \\ 0 & S \end{smallmatrix}\right] \in \mathcal{B}(H \times U)$, $0 \ll Q \in \mathcal{B}(H)$, $0 \ll S \in \mathcal{B}(U)$.

Then, for each $x_0 \in H$, *there is a unique* J*-critical control, and this control is strictly minimizing over* $\mathcal{U}_{\exp}(x_0)$. □

(This follows from Remark 8.6.4 and Corollary 8.2.7.) See Theorem 8.6.6 and Remark 8.6.7 for corresponding closed-loop systems and Riccati equations.

The assumptions correspond to $D^*JC = 0$, $C^*JC = Q \gg 0$, $D^*JD = S \gg 0$, and the corresponding cost is given by $\mathcal{J}(x_0, u) = \langle x, Qx \rangle_{\mathrm{L}^2} + \langle u, Su \rangle_{\mathrm{L}^2}$. Thus, standard LQR problems are covered; for the indefinite case and H^∞ problems we need slight changes.

The ELS variant of Theorem 8.3.9 takes the following form:

Theorem 8.6.6 (Σ_{crit}) *Assume that there is a unique J-critical control $u_{\text{crit}}(x_0)$ over $\mathcal{U}_*^*$ for each $x_0 \in H$. Define*

$$\Sigma_{\text{crit}} := \left[\begin{array}{c|c} \mathbb{A}_{\text{crit}} & \\ \hline \mathbb{C}_{\text{crit}} & \\ \mathbb{K}_{\text{crit}} & \end{array}\right] : x_0 \mapsto \left[\begin{array}{c|c} x_{\text{crit}}(x_0) & \\ \hline y_{\text{crit}}(x_0) & \\ u_{\text{crit}}(x_0) & \end{array}\right] := \left[\begin{array}{c|c} \mathbb{A}x_0 + \mathbb{B}\tau u_{\text{crit}}(x_0) & \\ \hline \mathbb{C}x_0 + \mathbb{D}u_{\text{crit}}(x_0) & \\ u_{\text{crit}}(x_0) & \end{array}\right]. \tag{8.105}$$

Then the following hold:

(a) We have $\Sigma_{\text{crit}} \in \mathrm{WPLS}_0^{-1}(\{0\}, H_{-1}, Y \times U)$ *and* $\Sigma_{\text{crit}} \in \mathcal{B}(H, \mathrm{L}^2(\mathbf{R}_+; H) \times \mathrm{L}^2(\mathbf{R}_+; Y) \times \mathrm{L}^2(\mathbf{R}_+; U))$.

(b) By setting $\mathcal{P} := \mathbb{C}_{\text{crit}}^* J \mathbb{C}_{\text{crit}} \in \mathcal{B}(H)$ *we obtain*

$$\mathcal{J}(x_0, u_{\text{crit}}(x_0) + \eta) = \langle x_0, \mathcal{P}x_0 \rangle_H + \mathcal{J}(0, \eta) \quad (x_0 \in H,\ \eta \in \mathcal{U}_{\exp}(0)). \tag{8.106}$$

(c) $\mathbb{C}_{\text{crit}} = C\mathbb{A}_{\text{crit}} + D\mathbb{K}_{\text{crit}}$.

(d) If $x_{\text{crit}}(x_0) \in \mathcal{C}(\mathbf{R}_+; H)$ *(equivalently,* $u_{\text{crit}}(x_0) \in \mathcal{U}_{\exp}^{\mathcal{C}}(x_0)$*) for all* $x_0 \in H$*, then* $\Sigma_{\text{crit}} \in \mathrm{WPLS}_{-\varepsilon}(\{0\}, H, Y \times U)$ *for some* $\varepsilon > 0$*,* $A_{\text{crit}} = A + BK_{\text{crit}}$ *and* $C_{\text{crit}} = C + DK_{\text{crit}}$ *on* $\mathrm{Dom}(A_{\text{crit}})$*, hence then* $\mathrm{Dom}(A_{\text{crit}}) \subset H_B$*, where* $\left[\begin{smallmatrix} A_{\text{crit}} \\ \hline C_{\text{crit}} \\ K_{\text{crit}} \end{smallmatrix}\right]$ *are the generators of* Σ_{crit}.

Here $\mathrm{WPLS}_\omega^{-1}$ is defined as WPLS_ω, except that the semigroup $\mathbb{A}$ needs to be strongly continuous (i.e., "C_0") in H_{-1} norm only, i.e., $\|\mathbb{A}x_0 - x_0\|_{H_{-1}} \to 0$ for all $x_0 \in H$.

However, from the fact that $\mathbb{A}_{\text{crit}} x_0 \in \mathrm{L}^2(\mathbf{R}_+; H) \cap \mathcal{C}(\mathbf{R}_+; H_{-1})$ one can usually derive the fact that actually $\mathbb{A}_{\text{crit}} x_0 \in \mathcal{C}(R_+; H)$, i.e., that (d) applies; see Examples 1.4.1–1.4.4 of [LT93] for details.

Proof of Theorem 8.6.6: By Corollary 8.1.10 (and Remark 8.6.4), $\Sigma_{\text{crit}} \in \mathcal{B}(H, \mathrm{L}^2)$.

Also the rest of the proof goes as the proof of Theorem 8.3.9 (use Lemma 8.6.3) except (c), which obvious, and (d), which is given below.

(d) Assume that $x_{\text{crit}}(x_0) \in \mathcal{C}(\mathbf{R}_+; H)$ for all $x_0 \in H$, equivalently, that $\mathbb{A}_{\text{crit}}$ is strongly continuous. Then $\mathbb{A}_{\text{crit}}$ is a C_0-semigroup. Now the exponential stability of $\mathbb{A}_{\text{crit}}$ and the rest of (c) follows as in the proof of Theorem 8.3.9. □

Having given the above basic theory, we make two remarks on natural extensions:

Remark 8.6.7 *It seems that most of the Riccati equation theory "on* $\mathrm{Dom}(A_{\text{crit}})$*", as given in Section 9.7, can be extended to this more general setting. (The operators* $\mathbb{B}^t$ *and* $\mathbb{D}^t$ *are defined on* $\widetilde{U}_{[0,T)}$ *and the equations only need to hold on* $\widetilde{U}_{[0,T)}$*.) Moreover, one can go on to derive "extended* B_w^**-CARE theory", the ELS extension of the theory of Section 9.2 under similar assumptions.*

Finally, by using Remark 8.6.4, one can write ELS counterparts for several of the optimization results of Chapters 10 and 11 (at least). In particular, instead of the special (standard LQR) cost function of Theorem 8.6.5, we can use any J-coercive cost function, over, e.g., $\mathcal{U}_{\text{out}}$ *too.* □

(The original proofs will do, mutatis mutandis.)

Remark 8.6.8 (More general systems than ELSs) *Instead of Hypothesis 8.6.1, the theory of this section can be adapted for general well-posed* $\mathbb{C}$ *(i.e., s.t.* $\left[\begin{smallmatrix}\mathbb{A}\\ \mathbb{C}\end{smallmatrix}\right] \in \mathrm{WPLS}$*) and for general unbounded* B *(e.g.,* $B \in \mathcal{B}(U, H_{-827})$*).* $\square$

(We omit the details; the methods are roughly the same.)

Notes

Some special cases of the theory are given in [LT93], by Irena Lasiecka and Roberto Triggiani, who treats the standard LQR problem and present examples of physical ELSs that are not WPLSs. Luciano Pandolfi [Pandolfi] has treated minimization of a more general cost function, although under a special coercivity condition and the assumptions that $\mathbb{A}$ is analytic with the unboundedness of B being less than 1 ("$\beta > -1$") and (A, B) having a bounded exponentially stabilizing state feedback operator.