Chapter 7

Q Parameterization

This section presents *Q-Parameterization* (or, more or less equivalently, *Youla parameterization*): a fundamental result of linear control theory. In its original form it provides an *affine* parameterization of the set of all closed loop transfer matrices which can be obtained by using a stabilizing finite order LTI controller in a standard finite order LTI feedback optimization setup.

In essence, derivation of the classical Q-parameterization result relies on the fact that the plant is linear, its coefficients are perfectly known, and that the structure and the order of the controller are not restricted in any way. Multiple generalizations of Q-parameterization are available. In particular, the assumption of time invariance of the plant is not really essential: it can be dropped as long as requirement for the controller to be time invariant is dropped as well. We will discuss a (relatively) recent development which allows one to derive a Q-parameterization in cases when a set of "structure" contraints is imposed on admissible controllers, as long as the structure is "compatible" with the plant according to a criterion of *quadratic invariance*.

The term "Q-parameterization" is due to the notation commonly used for the free parameter ("Q"). Since the "Q-terminology" would essentially remove letter Q from circulation in all situations when Youla parameterization is mentioned, this text also uses the name of the person associated with the discovery of the parameterization (D.C. Youla).

7.1 The Classical Q-Parameterization

This section derives Q-parameterization in its classical form, i.e. for finite order LTI plants being stabilized by finite order LTI controllers.

7.1.1 Motivation

Not every stable transfer matrix $G = G(s)$ (from w to e) can be obtained by closing a stabilizing LTI feedback loop $K = K(s)$ in the canonical feedback design setup shown on Figure 7.1. In this section, we investigate the limitations of LTI feedback ability to

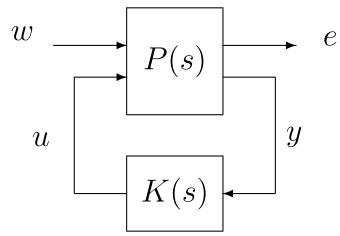

Figure 7.1: LTI feedback

change the closed loop transfer matrix, by parameterizing, in convenient terms, the set of all closed loop systems. Of course, the expression

$$G = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}, \quad \text{where} \quad P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \tag{7.1}$$

provides *some* parameterization of G (in terms of parameter K), but it is seriously deficient:

(a) the transfer matrix K in (7.1) is not arbitrary: it has to be a *stabilizing* controller, which by itself is a rather complicated highly nonlinear constraint;

(b) the dependence of G on K is not linear.

Essentially, Youla parameterization "untangles" the feedback loop, describing all achievable closed loop dependencies of e on w in the format shown on Figure 7.2, where S_1, S_2 are two fixed stable finite order LTI systems (which can be derived once a state space model of P is given, and have same order as P), and Q (the *Youla parameter*) is an *arbitrary* stable finite order LTI system (same number of inputs and outputs as K). Moreover, there is a one-to-one correspondence between *transfer matrices* of stable Youla parameters Q in Figure 7.2 and stabilizing controllers K in Figure 7.1.

In terms of transfer matrices, Youla parameterization means that

$$G(s) = G_0(s) + G_1(s)Q(s)G_2(s), \tag{7.2}$$

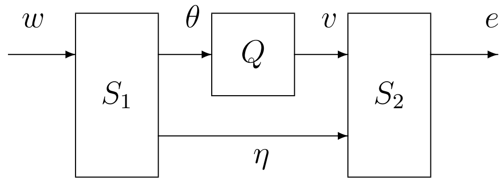

Figure 7.2: Affine parameterization of the closed loop

where G_0, G_1, G_2 are fixed stable transfer matrices, and Q is an arbitrary stable transfer matrix. There is also a convenient time domain interpretation, to be presented later in this section.

For design purposes, the result is extremely important: instead of thinking in terms of the controller transfer matrix $K = K(s)$, one can be much better off by designing $Q(s)$.

7.1.2 Main Statement

We consider LTI systems P from Figure 7.1 (where u and y are vectors of size m and k respectively) given in the state space format

$$P: \dot{x}(t) = Ax(t) + B_1 w(t) + B_2 u(t), \tag{7.3}$$
$$e(t) = C_1 x(t) + D_{11} w(t) + D_{12} u(t), \tag{7.4}$$
$$y(t) = C_2 x(t) + D_{21} w(t), \tag{7.5}$$

as well as the corresponding *feedback controllers*

$$K: \dot{x}_f(t) = A_f x_f(t) + B_f y(t), \tag{7.6}$$
$$u(t) = C_f x_f(t) + D_f y(t). \tag{7.7}$$

The resulting *closed loop system* has state space model

$$G: \dot{x}_{cl}(t) = A_{cl} x_{cl}(t) + B_{cl} w(t), \tag{7.8}$$
$$e(t) = C_{cl} x_{cl}(t) + D_{cl} w(t), \tag{7.9}$$

where

$$x_{cl}(t) = \begin{bmatrix} x(t) \\ x_f(t) \end{bmatrix}, \; A_{cl} = \begin{bmatrix} A + B_2 D_f C_2 & B_2 C_f \\ B_f C_2 & A_f \end{bmatrix}, \; B_{cl} = \begin{bmatrix} B_1 + B_2 D_f D_{21} \\ B_f D_{21} \end{bmatrix},$$

$$C_{cl} = \begin{bmatrix} C_1 + D_{12} D_f C_2 & D_{12} C_f \end{bmatrix}, \quad D_{cl} = D_{11} + D_{12} D_f D_{21}.$$

Recall that controller (7.6),(7.7) is said to stabilize system (7.3)-(7.5) if and only if A_{cl} is a Hurwitz matrix, i.e. if all of its eigenvalues have negative real part.

In the following statetement of Youla parameterization two state space models are called *equivalent* if they correspond to the same transfer matrix.

Theorem 7.1. *Let F, L be constant matrices such that $A+B_2F$ and $A+LC_2$ are Hurwitz matrices. Define system S_1 with input w and outputs θ, $\eta = [\Delta; w]$ according to*

$$S_1: \ \dot{\Delta}(t) = (A+LC_2)\Delta(t) + (B_1 + LD_{21})w(t), \tag{7.10}$$
$$\theta(t) = C_2\Delta(t) + D_{21}w(t). \tag{7.11}$$

Define system S_2 with inputs v and $\eta = [\Delta; w]$ according to

$$S_2: \ \dot{x}(t) = (A+B_2F)x(t) - B_2F\Delta(t) + B_1w(t) + B_2v(t), \tag{7.12}$$
$$e(t) = (C_1 + B_2F)x(t) - D_{12}F\Delta(t) + D_{11}w(t) + D_{12}v(t). \tag{7.13}$$

(a) If controller (7.6),(7.7) stabilizes system (7.3)-(7.5) then the closed loop model (7.8),(7.9) is equivalent to the model defined by the block diagram on Figure 7.2 where Q is the stable system with input $\theta = \theta(t)$, output $v = v(t)$, and state $x_q = x_q(t)$ defined by

$$Q: \ \dot{x}_q(t) = A_q x_q(t) + B_q\theta(t), \tag{7.14}$$
$$v(t) = C_q x_q(t) + D_q\theta(t) \tag{7.15}$$

with

$$A_q = A_{cl}, \ B_q = \begin{bmatrix} B_2D_f - L \\ B_f \end{bmatrix}, \ C_q = \begin{bmatrix} D_fC_2 - F & C_f \end{bmatrix}, \ D_q = D_f. \tag{7.16}$$

(b) If A_q is a Hurwitz matrix, and B_q, C_q, D_q are matrices of compatible dimensions then the controller K (with input $y = y(t)$, output $u = u(t)$, and state $x_f = [x_q; x_e]$) defined by (7.14),(7.15), and

$$E: \ \dot{x}_e(t) = Ax_e(t) + Bu(t) + L(c_2x_e(t) - y(t)), \tag{7.17}$$
$$\theta(t) = C_2x_e(t) - y(t), \tag{7.18}$$
$$u(t) = v(t) + Fx_e(t), \tag{7.19}$$

stabilizes system (7.3)-(7.5) and produces a closed loop which is equivalent to the model defined by the block diagram on Figure 7.2.

Though possibly intimidating, all state space models introduced while formulating Theorem 7.1 allow straightforward "systems" interpretations:

(a) equation (7.17) defines a standard state estimator (see Figure 7.3): using $y = y(t)$ and $u = u(t)$ as inputs, it produces an estimate $x_e = x_e(t)$ of $x = x(t)$ such that the dynamics of the estimation error $\Delta(t) = x(t) - x_e(t)$, described by equation (7.10), are stable;

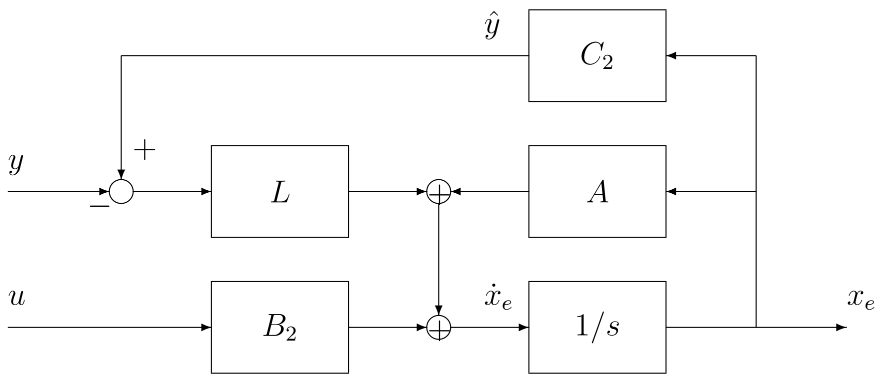

Figure 7.3: state estimator

(b) the output $\theta = \theta(t)$ of S_1 is the "sensor output estimation error", i.e. the difference $\theta(t) = C_2 x_e(t) - y(t)$ between the "sensor output estimate" $C_2 x_e(t)$ and the actual measurement $y(t)$;

(c) equations (7.14),(7.15) with the coefficients defined by (7.16) describe the closed loop dynamics relating sensor noise θ to the difference $v = u - F x_e$ between actual control u and the "nominal" stabilizing control value $F x_e$, so that the overall controller model is as shown on Figure 7.4;

(d) the equations for S_2 describe the relation between inputs $u(t) - F x(t) = v(t) - F\Delta(t)$, $w = w(t)$ and the output $e = e(t)$.

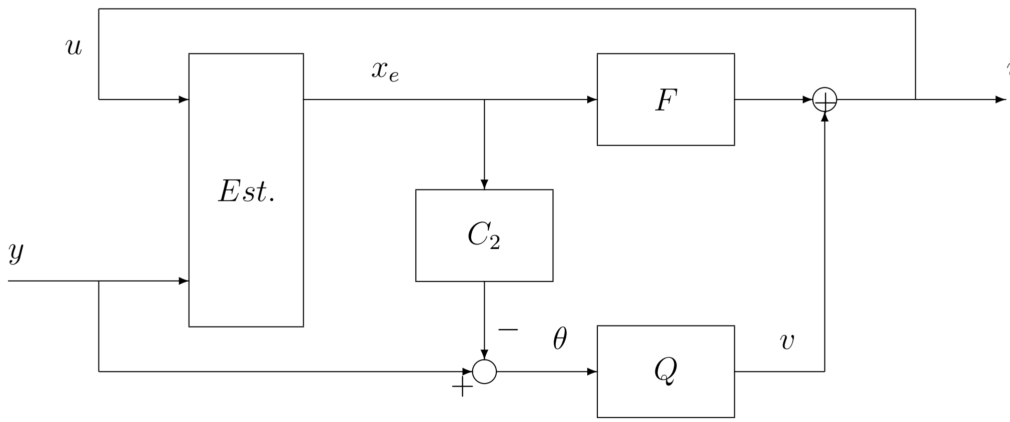

Figure 7.4: controller structure

(e) Transfer matrices G_i in (7.2) correspond to state space models

$$\begin{aligned} G_2: \quad \dot{\Delta}(t) &= (A + LC_2)\Delta(t) + (B_1 + LD_{21})w(t), \\ \theta(t) &= C_2\Delta(t) + D_{21}w(t), \end{aligned}$$

$$\begin{aligned} G_1: \quad \dot{\tilde{x}}(t) &= (A + B_2F)\tilde{x}(t) + B_2v(t), \\ \tilde{e}(t) &= (C_1 + D_{12}F)\tilde{x}(t) + D_{12}v(t), \end{aligned}$$

$$\begin{aligned} G_0: \quad \dot{x}(t) &= (A + B_2F)x(t) - B_2F\Delta(t) + B_1w(t), \\ \dot{\Delta}(t) &= (A + LC_2)\Delta(t) + (B_1 + LD_{21})w(t), \\ \tilde{e}(t) &= (C_1 + D_{12}F)x(t) - D_{12}F\Delta(t) + D_{11}w(t). \end{aligned}$$

7.1.3 Open Loop Zeros and Duality

The restrictions on the closed loop transfer function imposed by the Q parameterization can be understood in terms of the so-called *open loop zeros*.

To define zeros of MIMO systems, consider state space models

$$\dot{x} = ax + bu, \quad y = cx + du,$$

where the dimensions of x_1, u_1 and y_1 are n, m and k respectively. Consider the complex matrix

$$M(s) = \begin{bmatrix} a - sI & b \\ c & d \end{bmatrix}$$

where $s \in \mathbb{C}$ is a scalar complex variable. Assuming that $\ker M(s_0) = \{0\}$ for *some* $s_0 \in \mathbb{C}$, the system is said to have a *right zero* at a point $s \in \mathbb{C}$ if $\ker M(s) \neq \{0\}$, i.e. if $M(s)$ is not left invertible, or, equivalently, if there exists a non-zero pair (X, U) of complex vectors such that

$$sX = aX + bU, \quad 0 = cX + dU. \tag{7.20}$$

Similarly, assuming that the range of $M(s_0)$ is $\mathbb{C}^{n+k}$ for *some* $s_0 \in \mathbb{C}$, the system is said to have a *left zero* at s if the range of $M(s)$ is not the whole vector space $\mathbb{C}^{n+k}$, i.e. if $M(s)$ is not right invertible, or, equivalently, if there exists a non-zero pair (p, q) of complex vectors such that

$$sp = pa + qc, \quad 0 = pb + qd. \tag{7.21}$$

As it can be seen from the formulae for G_1 and G_2, the restrictions on the closed loop transfer matrix are caused by the unstable *right zeros* of the system

$$\dot{x} = Ax + B_1 w, \quad y = C_2 x + D_{21} w, \tag{7.22}$$

and by the unstable *left zeros* of the system

$$\dot{x} = Ax + B_2 u, \quad z = C_1 x + D_{12} u. \tag{7.23}$$

Indeed, the right zeros of (7.22) are the same as the right zeros of the state space model for G_2. Since $A + LC_2$ is a Hurwitz matrix, a complex number s_0 with non-negative real part is a right zero of G_2 if and only $G_2(s_0)w_0 = 0$ for some non-zero vector w_0 (possibly with complex entries). Similarly, s_0 such that $\text{Re}(s_0) \geq 0$ is a left zero of G_1 if and only $q_0 G_1(s_0) = 0$ for some non-zero row vector q_0. Therefore the closed loop transfer matrix $G = G_0 + G_1 Q G_2$ satisfies the *interpolation constraints* $G(s_0)w_0 = G_0(s_0)w_0$ for unstable right zeros of (7.22), and $q_0 G(s_0) = q_0 G_0(s_0)$ for unstable left zeros of (7.23). It can be shown that a complete list of such interpolation constraints yields an exact description of the set of all achievable closed loop transfer matrices $G = G(s)$ in the stabilized system. Such description can be viewed as *dual* to (7.16).

It can be seen immediately that the right zeros of (7.22) cause problems by obstructing the *observation process*, while the left zeros of (7.23) describe problems which a control action will experience even in the case of a complete knowledge of w and x.

7.1.4 Affine Parameterization of Impulse Responses

The following statement, where

$$\lambda(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0, \end{cases}$$

denotes the unit step function, is similar to Theorem 7.1, but expresses the affine parameterization in a time-domain format.

Theorem 7.2. *Let F, L be constant matrices such that $A+B_2F$ and $A+LC_2$ are Hurwitz matrices. A matrix-valued function $E = E(t)$ is a closed-loop impulse response from w to e generated in the closed loop system (7.3)-(7.7) by some stabilizing feedback controller (7.6),(7.7) if and only if there exists a function $Z = Z(t)$, elements of which are finite linear combinations of generalized one-sided exponents $t^k e^{st}\lambda(t)$ with* $\mathrm{Re}(s) < 0$ *(and, in case of a controller which is not strictly proper, Dirac deltas $\delta = \delta(t)$), such that*

$$E(t) = C_1 X(t) + D_{12} U(t) + D_{11}\delta(t), \tag{7.24}$$

where X, U are defined by

$$\begin{aligned} \dot{X}(t) &= AX(t) + B_2 U(t) + B_1\delta(t), &(7.25)\\ U(t) - FX(t) &= H(t)B_1 + V(t)D_{21}, &(7.26)\\ V(t) - H(t)L &= Z(t), &(7.27)\\ \dot{H}(t) &= H(t)A + V(t)C_2 - F\delta(t), &(7.28) \end{aligned}$$

and $X(\cdot), H(\cdot)$ are zero for $t < 0$.

By construction, the matrix-valued function $Z = Z(t)$ in Theorem 7.2 is the impulse response of system Q from Theorem 7.1. Once Z is selected, equtions (7.27),(7.28) determine $H = H(t)$ and $V = V(t)$ (the Dirac Delta functions dictate the initial conditions for the differential equation), after which equations (7.25),(7.26) determine $X = X(t)$ and $U = U(t)$. Naturally, X, U are the closed loop responses from w to x and u respectively (defined by (7.3)-(7.7)), and H, V are the closed loop impulse responses from respectively f and g to $u - Fx$ subject to the controller equations (7.6),(7.7) combined with

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_2 u(t) + f(t), &(7.29)\\ y(t) &= C_2 x(t) + g(t). &(7.30) \end{aligned}$$

Just as their frequency domain counterpart (7.16), the *linear* equations (7.24)-(7.28) define an affine set of stable impulse responses. It is important to understand that, despite their complicated appearance, equations (7.24)-(7.28) define a simple stable state space model

$$\tilde{e} = c\tilde{x} + d\tilde{u}, \quad \dot{\tilde{x}} = a\tilde{x} + b\tilde{u}, \quad \tilde{x}(0) = \tilde{x}_0, \tag{7.31}$$

where $\tilde{e} = \tilde{e}(t)$ is the column vector made out of the entries of matrix $E = E(t)$, $\tilde{x} = \tilde{x}(t)$ is the column vector made out of the entries of matrices $X = X(t)$ and $H = H(t)$ (while the components of $\tilde{x}_0$ are zeros and the entries of B_1 and F), and $\tilde{u} = \tilde{u}(t)$ is the column vector made out of the entries of matrix $Z = Z(t)$.

The proof of Theorem 7.2 is similar to the proof of Theorem 7.1, and uses the following elementary statements about the relation between state space models and impulse responses of CT LTI systems:

(S1) if the output y of system L is obtained by applying LTI transformation with impulse response $q = q(t)$ to the output $g = g(t)$ of state space model

$$\dot{x}_L(t) = ax_L(t) + bf(t), \quad g(t) = cx_L(t) + df(t)$$

with input $f = f(t)$ then its impulse response $l = l(t)$ is given by

$$l(t) = h(t)b + q(t)d, \text{ where } \dot{h}(t) = h(t)a + q(t)c, \ h(t) = 0 \text{ for } t < 0;$$

(S2) if the output y of system L is obtained by

$$\dot{x}_L(t) = ax_L(t) + bg(t), \quad y(t) = cx_L(t) + dg(t)$$

where $g = g(t)$ is the output of LTI transformation with impulse response $v = v(t)$ and input $f = f(t)$ then its impulse response $l = l(t)$ of L is given by

$$l(t) = cx(t) + dv(t), \text{ where } \dot{x}(t) = ax(t) + bv(t), \ x(t) = 0 \text{ for } t < 0.$$

A dual description of the set of all functions $\tilde{e}$ defined by (7.31) with a square integrable (over $\{t\} = (0, \infty)$) input $\tilde{u}$ can be given by characterizing all square integrable row vector functions q and constants v_0 such that the identity

$$\int_0^\infty q(t)\tilde{e}(t)dt = v_0 \tag{7.32}$$

holds for all $\tilde{e}$.

Lemma 7.1. *Let matrices a, b, c, d and vector $\tilde{x}_0$ be fixed. Assume a is a Hurwitz matrix. Identity (7.32) is satisfied for all $\tilde{e} = \tilde{e}(t)$ defined by (7.31) with a square integrable $\tilde{u} = \tilde{u}(t)$ if and only if*

$$p(t)b + q(t)d = 0, \quad p(0)\tilde{x}_0 = v_0, \tag{7.33}$$

where

$$p(t) = \int_t^\infty q(\tau)ce^{(\tau-t)a}d\tau$$

is the only solution of the differential equation

$$\dot{p}(t) = -p(t)a - q(t)c$$

which converges to zero as $t \to \infty$.

To prove Lemma 7.1, simply notice that

$$\begin{aligned}
\int_0^\infty q\tilde{e}dt &= \int_0^\infty q(c\tilde{x}+d\tilde{u})dt \\
&= \int_0^\infty \{(-\dot{p}-pa)\tilde{x}+qd\tilde{u}\}dt \\
&= p(0)\tilde{x}_0 + \int_0^\infty (p\dot{\tilde{x}} - pa\tilde{x} + qd\tilde{u})dt \\
&= p(0)\tilde{x}_0 + \int_0^\infty (pb+qd)\tilde{u}dt.
\end{aligned}$$

7.2 Q Parametrization for Structured Controllers

A relatively recent development in the theory of linear feedback control is an extension of the classical Q parametrization rezult to the case when the feedback is allowed to have *structure*. As a rule, a *structure* is a set of linear constraints to be imposed on the linear functions which defined the feedback, i.e. a linear subspace $\mathcal{K}$ in the space of all possible linear feedback relations. For an "arbitrarily" selected structure subspace $\mathcal{K}$, the set of all closed loop systems is unlikely to be affine, in which case a Q parameterization is obviously impossible. On the other hand, some interesting structures do allow an affine parameterization of all closed loop systems. The objective of this section is to explore conditions, to be imposed on $\mathcal{K}$ and on the plant P, to guarantee validity of a Q parametrization.

To streamline the presentation, only the discrete time case will be considered. Most of the ideas to be discussed generalize easily to the continuous time case.

7.2.1 A "Zero Initial Conditions" Setup

For positive integers m, k let $\mathcal{L}_c^{m,k}$ denote the set of all linear functions $S:\ \ell^m \to \ell^k$ which are *causal*, in the sense that $v(t) = 0$ for all $t < T$ whenever $v = Sw$ and $w(t) = 0$ for $t < T$. Such functions can be viewed as models of causal linear discrete time systems with zero initial conditions and with output $v = Sw \in \ell^k$ uniquely defined by the input $w \in \ell^m$. Similarly, let $\mathcal{L}_{sc}^{m,k}$ denote the set of all linear functions $S:\ \ell^m \to \ell^k$ which are *strictly causal*, in the sense that $v(t) = 0$ for all $t \le T$ whenever $v = Sw$ and $w(t) = 0$ for $t < T$.

Consider DT feedback design setup described by signal relations

$$y = Pu + f, \quad u = Ky + g, \tag{7.34}$$

where

(a) $P \in \mathcal{L}_c^{m,k}$ is a given causal linear model, the *plant*;

(b) $K \in \mathcal{L}_c^{k,m}$ is a linear model to be designed, the *controller*;

(c) f, g are input signals, and e, y, u are outputs signals, i.e. equations (7.34) are to be solved for $y \in \ell^k$ and $u \in \ell^m$ for every possible combination of $f \in \ell^d$, $g \in \ell^k$.

Feedback system (7.34) (defined by $P \in \mathcal{L}_c^{m,k}$ and $K \in \mathcal{L}_c^{k,m}$) is called *well posed* if equations (7.34) have a unique solution $(y, u) \in \ell^k \times \ell^m$ for all $f \in \ell^k$ and $g \in \ell^m$. In this case the map $(f, g) \to (y, u)$ is a linear system $G \in \mathcal{L}_c^{k+m,k+m}$, to be referred to as the *closd loop system* defined by (7.34).

Let us refer to a linear subspace $\mathcal{K} \subset \mathcal{L}_c^{k,m}$ as *feedback structure*. We are interested in studying, for a given pair $(P, \mathcal{K})$, the set $\mathcal{G} = \mathcal{G}(P, \mathcal{K})$ of all closed loop systems G resulting from using controllers K for which feedback system (7.34) is well posed. Specifically, it is important to describe the circumstances under which $\mathcal{G}$ is an affine set of the special type:

$$\mathcal{G} = \{G_0 + G_1 Q G_2 : \quad Q \in \mathcal{K}\} \tag{7.35}$$

for some $G_0 \in \mathcal{L}_c^{k+m,k+m}$, $G_1 \in \mathcal{L}_c^{m,k+m}$, $G_2 \in \mathcal{L}_c^{k+m,k}$.

Note that parametrization (7.35) does not discriminate between stabilizing and non-stabilizing controllers. This is not expected to cause any practical difficulties, as stability of feedback interconnections associated with (7.34) is usually a *linear* constraint to be imposed on G (in other words, scaling and addition of stable linear system models leads to a stable system model).

Example: Conversion to Zero Initial Conditions Setup

Consider the task of designing a linear feedback law

$$u(t) = k_0(t)y(t) + k_1(t)y(t-1), \quad y(-1) = y_0$$

for the system

$$x(t+1) = x(t) + w(t) + u(t), \quad x(0) = x_0, \quad e(t) = x(t), \quad y(t) = x(t) + w(t),$$

with the objective of achieving stability of the closed loop state space model

$$x_{cl}(t+1) = \begin{bmatrix} 1 + k_0(t) & k_1(t) \\ 1 & 0 \end{bmatrix} x_{cl}(t) + \begin{bmatrix} 1 + k_0(t) \\ 1 \end{bmatrix} w(t), \quad x_{cl}(t) \stackrel{\text{def}}{=} \begin{bmatrix} x(t) \\ y(t-1) \end{bmatrix},$$

which is understood as finiteness of the L2 gain from $[w; g]$ to x_{cl} in

$$x_{cl}(t+1) = \begin{bmatrix} 1+k_0(t) & k_1(t) \\ 1 & 0 \end{bmatrix} x_{cl}(t) + \begin{bmatrix} 1+k_0(t) & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} w(t) \\ g(t) \end{bmatrix},$$

and minimizing closed loop L2 gain from w to e.

To convert this setup to a "zero initial conditions" format, begin with expressing the sensor output $y(t)$ as a linear function of $u(\cdot)$ and other parameters. We have

$$y(t) = \sum_{\tau=0}^{t-1} u(\tau) + x(0) + \sum_{\tau=0}^{t} w(\tau) \quad \forall\ t \in \mathbb{Z}_+,$$

which suggects that $P:\ \ell \to \ell$ should be the linear function, defined in such a way that $y = Pu$ means

$$y(t) = \sum_{\tau=0}^{t-1} u(\tau) \quad \forall\ t \in \mathbb{Z}_+,$$

and the external signal $f = f(t)$ is expected to be of the form

$$f(t) = x_0 + \sum_{\tau=0}^{t} w(\tau) \quad \forall\ t \in \mathbb{Z}_+,$$

where $x_0 \in \mathbb{R}$.

Similarly $K:\ \ell \to \ell$ is a linear function defined in such a way that $u = Ky$ means

$$u(t) = \begin{cases} k_0(t)y(t) + k_1(t)y(t-1), & t > 0, \\ k_0 y(t), & t = 0, \end{cases}$$

while $g \in \ell$ is a signal such that $g(t) = 0$ for $t > 0$, with $g(0)$ being arbitrary unless $k_1(0) = 0$, in which case $g(0) = 0$.

The set $\mathcal{K}$ of such linear functions $K:\ \ell \to \ell$ is closed under the operations of addition and scaling by a real number. In addition, due to strict causality of P, and causality of K, equations (7.34) are guaranteed to have unique solutions $u, y \in \ell$ for all $f, g \in \ell$, i.e. the feedback loop is well-posed for all $K \in \mathcal{K}$.

Note that the optimization objective is formally defined as a property of the function mapping $(w, x_0, y_0) \in \ell \times \mathbb{R} \times \mathbb{R}$ to $x = y - w$, but, subject to closed loop stability, it equals L2 gain of $G_{fy}L$, where G_{fy} is the "upper left block" of G, i.e. the closed loop map from f to y, and $L:\ \ell \to \ell$ is defined in such a way that $v = Lw$ means

$$v(t) = \sum_{\tau=0}^{t} w(\tau).$$

In other words, the optimization objective is a convex function of G.

Similarly, the stabilization objective is formally defined as a property of the function mapping $(w, g, x_0, y_0) \in \ell \times \mathbb{R} \times \mathbb{R}$ to x_{cl}, but it is actually equivalent to L2 gain of G being finite, because

(a) $x_{cl}(t)$ is a linear combination of $y(t)$, $y(t-1)$, and $w(t)$;

(b) $y(t)$, $u(t)$ are linear combinations of $x_{cl}(t)$, $x_{cl}(t+1)$, and $w(t)$;

(c) the dependence of y, u on the initial conditions x_0, y_0 can be expressed in terms of dependence of y, u on $w(0)$ and $g(0)$.

Since addition and scaling of finite L2 gain systems results in a finite L2 gain system, the stabilization objective will only impose *linear* constraints on G, or, equivalently, on the parameter Q in (7.35).

7.2.2 Quadratic Invariance

Let us call a linear subset $\mathcal{K} \subset \mathcal{L}_c^{k,m}$ *quadratically invariant* with respect to a given $P \in \mathcal{L}_c^{m,k}$ if $K_1 P K_2 \in \mathcal{K}$ whenever $K_1, K_2 \in \mathcal{K}$.

Example 7.1. *$\mathcal{K} = L_c^{k,m}$ is quadratically invariant with respecct to every strictly causal P (i.e. every $P \in \mathcal{L}_{sc}$).*

Example 7.2. *A system $K \in \mathcal{L}_c^{k,m}$ is called* memoryless *if there exists a sequence $\{k(t)\}_{t=0}^{\infty}$ of m-by-k real matrices $k(t)$ such that $u = Ky$ means $u(t) = k(t)y(t)$ for all $t \in \mathbb{Z}_+$. The subset $\mathcal{K}_{ml}$ of all* memoryless *systems $K \in \mathcal{L}_c^{k,m}$ is quadratically invariant with respect to a given $P \in \mathcal{L}_c^{m,k}$ if and only if P itself is memoryless.*

Let us call a subset $\mathcal{K} \subset \mathcal{L}_c^{k,m}$ *local* if membership in $\mathcal{K}$ is defined "locally", time-wise, in the sense that $K \in \mathcal{K}$ if and only if $\pi_T K \in \pi_T \mathcal{K}_T$ for all $T \in \mathbb{Z}_+$, where π_T is the "first $T+1$ samples" projection, i.e.

$$v = \pi_T w \quad \text{means} \quad v(t) = \begin{cases} w(t), & t \le T, \\ 0, & t > T. \end{cases}$$

Example 7.3. *For example, the subset $\mathcal{K}_{ml}$ of all* memoryless *systems $K \in \mathcal{L}_c^{k,m}$ is local.*

Example 7.4. *The subset $\mathcal{K}_{mlf} \subset \mathcal{K}_{ml}$ of all memoryless functions for which the coefficients sequence $\{k(t)\}_{t=0}^{\infty}$ has a finite number of non-zero elements, is not local, because $\pi_T \mathcal{K}_{ml} = \pi_T \mathcal{K}_{mlf}$, i.e. there is no way of distinguishing between elements of $\mathcal{K}_{ml}$ and $\mathcal{K}_{mlf}$ by looking at their finite subsequences.*

Example 7.5. *The set $\mathcal{K}_{mlti} \subset \mathcal{K}_{ml}$ of all memoryless functions for which the coefficients sequence $\{k(t)\}_{t=0}^{\infty}$ is constant, i.e. satisfies $k(t) = k(t+1)$ for all $t \in \mathbb{Z}_+$, is local.*

The following theorem combines quadratic invariance, locality, and a strict causality assumption to produce a sufficient condition for the existence of Q parametrization (7.35).

Theorem 7.3. *Assume that causal system $P \in \mathcal{L}_c^{m,k}$ and a linear local subset $\mathcal{K} \subset \mathcal{L}_c^{k,m}$ are such that*

(a) $\mathcal{K}$ is quadratically invariant, i.e. $K_1PK_2 \in \mathcal{K}$ whenever $K_1, K_2 \in \mathcal{K}$;

(b) $PK \in \mathcal{L}_{sc}^{k,k}$ is strictly causal for all $K \in \mathcal{K}$.

Then the feedback interconnection (7.34) is well-posed for all $K \in \mathcal{K}$, and the set $\mathcal{G}$ of all closed loop systems has Q parametrization (7.35) with $Q = K(I - PK)^{-1}$ and

$$G_0 = \begin{bmatrix} I & P \\ 0 & I \end{bmatrix}, \quad G_1 = \begin{bmatrix} P \\ I \end{bmatrix}, \quad G_2 = \begin{bmatrix} I & P \end{bmatrix}.$$

Proof. Since PK is strictly causal for all $K \in \mathcal{K}$, equation $y - PKy = w$ has unique solution $y = (I - PK)^{-1}w \in \ell^k$ for every $w \in \ell^k$. Hence equations (7.34) have unique solution

$$y = (I - PK)^{-1}(f + Pg) \in \ell^k, \quad u = K(I - PK)^{-1}(f + Pg) + g \in \ell^m$$

for all $f \in \ell^k$ and $g \in \ell^m$, $K \in \mathcal{K}$. Therefore representation $G = G_0 + G_1QG_2$ with $Q = K(I - PK)^{-1}$ takes place for all $K \in \mathcal{K}$.

Note that, for $y \in \ell^k$ and $v, w \in \ell^m$, equalities $v = Ky$ and $y - PKy = Pw$ imply

$$(I - PK)(w + v) = w + v - KPw - KPv = w + Ky - K(y - PKy) - KPKy = w.$$

Similarly, $u - KPu = w$ implies $u - w = Ky$ and $y - PKy = Pw$ for $y = Pu$. Hence $I - KP$ is invertible, and

$$K = (I + QP)^{-1}Q \quad \text{whenever} \quad Q = K(I - PK)^{-1}.$$

Now, to complete the proof it is sufficient to show that

(i) $K(I - PK)^{-1} \in \mathcal{K}$ whenever $K \in \mathcal{K}$;

(ii) $(I + QP)^{-1}Q \in \mathcal{K}$ whenever $Q \in \mathcal{K}$.

Let us prove (i) (the proof of (ii) follows the same pattern).

The idea of the proof is to rely on the observation that, due to quadratic invariance, $K(PK)^t \in \mathcal{K}$ for all $t \in \mathbb{Z}_+$ and $K \in \mathcal{K}$. Indeed, for $t = 0$ we have $K(PK)^t = K \in \mathcal{K}$, and, if $K(PK)^t \in \mathcal{K}$ for some $t \in \mathbb{Z}_+$ then

$$K(PK)^{t+1} = (K(PK)^t)PK = K_1PK \in \mathcal{K} \quad \text{because} \quad K_1 = K(PK)^t \in \mathcal{K}.$$

Hence, by induction, $K(PK)^t \in \mathcal{K}$ for all $t \in \mathbb{Z}_+$.

Now, an informal application of the "power expansion"

$$(I - PK)^{-1} = \sum_{t=0}^{\infty}(PK)^t,$$

would yield

$$Q \overset{\text{def}}{=} K(I - PK)^{-1} = \sum_{t=0}^{\infty} K(PK)^t \in \mathcal{K},$$

as an "infinite sum" of elements from $\mathcal{K}$. Formally speaking, making an argument like this rigorously requires introduction of some suitable notion of *convergence* on $\mathcal{L}_c^{k,k}$ and $\mathcal{L}_c^{m,k}$ (to define the infinite sums), as well as an assertion of *closedness* of $\mathcal{K}$ in this topology. However, under the assumptions made, there is a way around such complications.

Indeed, since PK is strictly causal, we have $\pi_T(PK)^T = 0$ for all $T \in \mathbb{Z}_+$. Hence

$$X_T \overset{\text{def}}{=} \pi_T Y_T, \quad \text{where} \quad Y_T \overset{\text{def}}{=} K \sum_{t=0}^{T-1}(PK)^t,$$

satisfies

$$X_T = \pi_T K \sum_{t=0}^{T}(PK)^t = \pi_T K + \pi_T K \sum_{t=1}^{T}(PK)^t = \pi_T K + X_T PK,$$

which implies

$$X_T = \pi_T K(I - PK)^{-1} = \pi_T Q.$$

Sine $Y_T \in \mathcal{K}$ by construction, we have $X_T \in \mathcal{K}_T$ for all $T \in \mathbb{Z}_+$. Hence $\pi_T Q \in \mathcal{K}_T$ for all $T \in \mathbb{Z}_+$, which implies $Q \in \mathcal{K}$ by the *locality* assumption. ■

7.2.3 Network Information Delay Structure

Quadratic invariance is a restictive conditions. For many controller structures (such as the "memoryless controllers" structure $\mathcal{K}_{ml}$) the set of plants of P for which quadratic invariane holds is so narrow that Theorem 7.3 becomes quite useless, as it covers no interesting tasks.

In this subsection we discuss a class of structures $\mathcal{K}$ for which quadratic invariance is not overly restrictive.

For a given k-by-m matrix T with entries from $\mathbb{Z}_+ \cup \{\infty\}$ let $\mathcal{N}[T]$ be the set of all linear systems $K \in \mathcal{L}_c^{k,m}$ such that the i-th component $u_i(t)$ if $u(t)$ is zero whenever $u = Ky$ and the the l-th component $y_l(\tau)$ of $y(\tau)$ equals zero for all l, τ such that $t \geq \tau + T_{il}$. The resulting structure $\mathcal{K} = \mathcal{N}[T]$ can be interpreted as a *network propagation delay* structure, where T_{il} represents the minimal time it takes for the information from *sensor node* $l \in \{1, \dots, k\}$ to reach *actuator node* $i \in \{1, \dots, m\}$.

A simple sufficient condition for quadratic invariance can be formulated in terms of network propagation delays.

Theorem 7.4. *Let T be a non-negative integer m-by-k matrix. Let A be the k-by-m matrix with entries from $\mathbb{Z}_+ \cup \{\infty\}$ defined by*

$$A_{li} = \max_{a,b} \max\{0, T_{ab} - T_{al} - T_{ib}\}.$$

Then $\mathcal{N}[T]$ is quadratically invariant with respect to $P \in \mathcal{L}_c^{m,k}$ if and only if $P \in \mathcal{N}[A]$.

In contrast with matrix T, matrix A in Theorem 7.4 specifies lower bounds for *action* propagation delay in a network, i.e. how fast action of a given actuator node (a component of u) reaches a given sensor note (a component of y). Essentially, the requirement is for the information to travel at least as fast as action.

Example: a Simple Network Structure

Consider plant $P \in \mathcal{L}_c^{2,2}$ defined by equations

$$\begin{aligned} y_1(t+1) &= 2y_1(t) + w(t) + u_1(t), \\ y_2(t+1) &= y_1(t) + u_2(t), \end{aligned}$$

where $u = [u_1; u_2]$ is actuator input, and $y = [y_1; y_2]$ is sensor output. The objective is to design feedback control of the form

$$\begin{aligned} u_1(t) &= \sum_{\tau=0}^{t} \alpha_{11}(\tau) y_1(t-\tau), \\ u_2(t) &= \sum_{\tau=0}^{t} \alpha_{22}(\tau) y_2(t-\tau) + \sum_{\tau=d}^{t} \alpha_{12}(\tau) y_1(t-\tau), \end{aligned}$$

where $d \geq 0$ is a given integer, to minimize closed loop L2 gain from w to y.

Here the index $i \in \{1, 2\}$ in u_i and y_i refers to one of the two nodes in a network shown on Figure 7.5. While sensor measurement information $y_i(t)$ is made available immediately at node i (which is reflected in the possibility of making $u_i(t)$ directly dependent on $y_i(t)$ by selecting $\alpha_{ii}(0) \neq 0$), it incurs a fixed delay of d time steps getting from node 1 to node 2, and it does not travel from node 2 to node 1 at all.

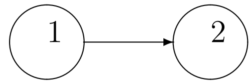

Figure 7.5: A two-node network

The resulting matrix T of information propagation delays is

$$T = \begin{bmatrix} 0 & \infty \\ d & 0 \end{bmatrix}.$$

The corresponding matrix A from Theorem 7.4 is given by $A = T$. Since the true action propagation matrix A_P for the plant is

$$A_P = \begin{bmatrix} 1 & \infty \\ 2 & 1 \end{bmatrix},$$

condition of $P \in \mathcal{N}[A]$ is satisfied if and only if $d \leq 2$. Since the plant is strictly causal, a combination of Theorem 7.3 and Theorem 7.4 guarantees existence of a Q parametrization of all closed loop systems when $d \leq 2$.

For $d \leq 1$ the minimal possible closed loop L2 gain from w to y is 1, achieved, for example, with feedback

$$u_1(t) = -2y_1(t), \quad u_2(t) = -y_1(t).$$

For $d \geq 2$ the minimal possible closed loop L2 gain from w to y is $\sqrt{2}$, achieved, for example, with feedback

$$u_1(t) = -2y_1(t), \quad u_2(t) = 0.$$