Chapter 6
Set-theoretic analysis of dynamic systems

In this section, several applications of set-theoretic methods to the performance analysis of dynamic systems will be presented. Although, in principle, the proposed techniques are valid for general systems, their application is computationally viable in the case of (uncertain) linear systems and thus we restrict the attention to this case.

6.1 Set propagation

6.1.1 Reachable and controllable sets

Consider a dynamic system of the form

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

or of the form

$$x(t+1) = f(x(t), u(t))$$

where $u(t) \in \mathcal{U}$. The following classical definitions of reachability and controllability sets are reported.

Definition 6.1 (Reachability set). Given the set $\mathcal{P}$, the reachability set $\mathcal{R}_T(\mathcal{P})$ from $\mathcal{P}$ in time $T < +\infty$ is the set of all vectors x for which there exists $x(0) \in \mathcal{P}$ and $u(\cdot) \in \mathcal{U}$ such that $x(T) = x$.

Definition 6.2 (Controllability set). Given the set $\mathcal{S}$, the controllability set $\mathcal{C}_T(\mathcal{S})$ to $\mathcal{S}$ in time $T < +\infty$ is the set of all vectors x for which there exists $u(\cdot) \in \mathcal{U}$ such that if $x(0) = x$ then $x(T) \in \mathcal{S}$.

F. Blanchini, S. Miani, *Set-Theoretic Methods in Control*, Systems & Control: Foundations & Applications, DOI 10.1007/978-3-319-17933-9_6

More in general, $\mathcal{S}$ is said to be reachable from $\mathcal{P}$ in time T if for all $x \in \mathcal{S}$ there exists $x(0) \in \mathcal{P}$ and $u(\cdot)$ such that $x(T) = x$. Similarly, $\mathcal{P}$ is said to be controllable to $\mathcal{S}$ in time T if, for all $x(0) \in \mathcal{P}$, there exists $u(\cdot)$ such that $x(T) \in \mathcal{S}$. Unless for very specific cases, the fact that $\mathcal{P}$ is reachable from $\mathcal{S}$ does not imply that $\mathcal{S}$ is controllable to $\mathcal{P}$ and vice versa.

However, if backward systems are considered, namely systems that evolve backward in time of the form

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

or of the form

$$x(t+1) = f^{-1}(x(t), u(t))$$

where f^{-1} is the inverse of f with respect to x (if it exists at all), precisely the map such that $f^{-1}(f(x,u),u) = x$ for all x and $u \in \mathcal{U}$, then the set $\mathcal{P}$ is reachable from (controllable to) $\mathcal{S}$ if and only if $\mathcal{S}$ is controllable to (reachable from) $\mathcal{P}$ for the backward system.

Controllable sets have the following composition property[1]. If $\mathcal{S}_0$ is controllable in time T_1 to $\mathcal{S}_1$ and $\mathcal{S}_1$ is controllable in time T_2 to $\mathcal{S}_2$, then $\mathcal{S}_0$ is controllable in time $T_1 + T_2$ to $\mathcal{S}_2$. The analogous composition property holds for reachability.

Reachable sets are useful to describe the effects of a bounded disturbance on a dynamical system or to describe the range of effectiveness of a bounded control. Unfortunately, the computation of reachable sets is, in general, very hard even in the discrete-time case, although effort is currently put in this direction [RKML06]. For simple systems, typically planar ones, they can be computed (approximately) by simulation and the approximated reachable and controllable sets can be visualized by appealing computer graphics. Unfortunately, as the dimension grows, our mind gets somehow lost, besides the inherent intractability of reachable set computation.

From the theoretical point of view, some results that characterize the closedness or compactness of controllability/reachability sets which are available in the mathematical literature. For instance, in the discrete-time case, if the map f is assumed to be continuous and $\mathcal{U}$ compact, then the expression of the one-step reachability set of a compact set $\mathcal{P}$, precisely

$$f(\mathcal{P}, \mathcal{U})$$

is compact. Therefore the reachable set in k steps, that can be recursively computed by setting $\mathcal{R}_0 := \mathcal{P}$ and

$$\mathcal{R}_{k+1} = f(\mathcal{R}_k, \mathcal{U})$$

[1] A semi-group property.

is compact. Some compactness results for the controllability sets can be easily inferred under some assumptions, such as the system reversibility (i.e., that $f^{-1}(x, u)$ is defined for all u and continuous). This kind of closedness–compactness results are valid also for continuous-time systems under suitable regularity assumptions.

The next theorem shows that, at least for the case of linear systems (which are those we will be mostly interested in), some reasonable procedures can be devised.

Theorem 6.3. *Consider the system*

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (or\ \ x(t+1) = Ax(t) + Bu(t))$$

where $u \in \mathcal{U}$, with $\mathcal{U}$ a convex and compact set and, in the discrete-time case, A is assumed to be invertible. Let $\mathcal{P}$ be a convex and compact set. Then, for all $T < +\infty$,

- *the controllability set in time T to $\mathcal{P}$ is convex and compact.*
- *the reachability set in time T from $\mathcal{P}$ is convex and compact.*

In the discrete-time case, if matrix A is singular, the reachability set is still convex and compact whereas the controllability set is convex, closed but not necessarily bounded.

Proof. The proof of compactness will be reported in the continuous-time case only, whereas the proof of convexity and the discrete-time case are left as an exercise.

The reachability set is given by the set of all vectors

$$x = e^{AT}\bar{x} + \int_0^T e^{A(T-\sigma)}Bu(\sigma)d\sigma, \tag{6.1}$$

(with $u(\cdot)$ a measurable function) namely the sum of the image of $\mathcal{P}$ with respect to e^{AT} (which is compact) and the set of all vectors reachable from 0 in time T:

$$\mathcal{R}_T(\mathcal{P}) = e^{AT}\mathcal{P} + \mathcal{R}_T(\{0\})$$

The set of states reachable from 0 in a finite time is compact as shown in [PN71], and, since the sum of two compact sets is compact, $\mathcal{R}_T(\mathcal{P})$ is compact. The analogous proof for controllable sets is derived by noticing that the controllable set is the set of all $\bar{x}$ for which (6.1) holds with x in $\mathcal{P}$. By multiplying both sizes by e^{-AT} one immediately derives

$$\mathcal{C}_T(\mathcal{P}) = e^{-AT}\mathcal{P} + \mathcal{R}_T^-(\{0\})$$

where we denoted by $\mathcal{R}_T^-(\{0\})$ the set of 0-reachable states of the time-reversed sub-system $(-A, -B)$, hence the claim.

Convexity of reachability and controllability sets in the case of linear systems is a strong property which allows to obtain practical results, as it will be seen later.

An important problem that can be solved, in principle, in a set-theoretic framework is the analysis of uncertainty effects via set propagation. The literature on this kind of investigation is spread in different areas. A classical approach to the problem is that based on the concept of differential inclusion. As we have seen, a system of the form $\dot{x}(t) = f(x(t), w(t)),\ w(t) \in \mathcal{W}$, is a special case of differential inclusion. Given an initial condition $x(0)$, if one is able to determine the reachable set $\mathcal{R}_t(\{x(0)\})$, then one can actually have an idea of the uncertainty effect. The literature presents some effort in this sense, however, most of the work is effective only for special classes of systems, typically of low dimensions. A survey of numerical methods for differential inclusions can be found in [DL92].

6.1.2 Computation of set propagation under polytopic uncertainty

Let us now consider the discrete-time system

$$x(t+1) = A(w(t))x(t) + E(w(t))d(t) \tag{6.2}$$

with

$$A(w) = \sum_{i=1}^{s} A_i w_i(t), \quad E(w) = \sum_{i=1}^{s} E_i w_i(t)$$

$$w \in \mathcal{W} = \{w : w_i \geq 0,\ \sum_{i=1}^{s} w_i = 1\}$$

and $d \in \mathcal{D}$, also a polytope. Here, no control action is considered and w and d are both external uncontrollable signals.

Consider the problem of computing the propagation of the uncertainty for this system, starting from a set $\mathcal{X}_0$ of initial conditions which is a polytope. This set can be propagated forward in time, keeping into account the effect of the uncertainty and disturbance, by considering the set:

$$\mathcal{X}_1 = \mathcal{R}_1(\mathcal{X}_0) = \{A(w)x + E(w)d :\ w \in \mathcal{W},\ d \in \mathcal{D}\} \tag{6.3}$$

Even from the first step it is not difficult to see that the one step reachable set $\mathcal{X}_1$ is not convex (then it cannot be a polytope). The lack of convexity is shown in the next example.

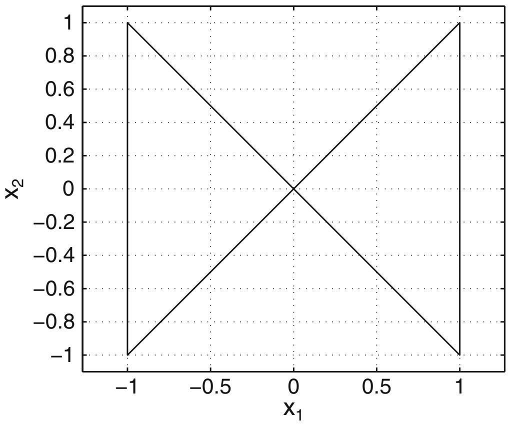

Fig. 6.1 Butterfly shaped non-convex one step reachable set for example 6.4

Example 6.4. Consider the autonomous system whose matrices are

$$A = \begin{bmatrix} 1 & w \\ 0 & 1 \end{bmatrix}, \quad E = 0,$$

and $|w| \leq 1$. Let $\mathcal{X}_0 = \{(x_1,\ x_2):\quad x_1 = 0,\ x_2 \in\ [-1,\ \ 1]\}$. The reachable set is the set of all the possible images of the considered segment with respect to matrix $A(w)$, which turns out to be the union of two triangles with center in the origin (the one having vertices $[0\ 0]^T$, $[1\ 1]^T$ $[1\ -1]^T$ and its opposite), depicted in Figure 6.1 and clearly non-convex.

Though the reachable set is non-convex, Barmish and Sankaran [BS79] showed that the convex hull of the reachable sets can be propagated recursively, as per the next result.

Proposition 6.5. *Let $\mathcal{X}_0$ be a polytope and let $\mathcal{X}_k$ be the k-step reachability set of (6.2) from $\mathcal{X}_0$. Let $\hat{\mathcal{X}}_k = conv\{\mathcal{X}_k\}$ be its convex hull. Then the sequence of convex hulls can be generated recursively as*

$$\hat{\mathcal{X}}_{k+1} = conv\left\{\mathcal{R}_1\left(\hat{\mathcal{X}}_k\right)\right\},$$

roughly, as the convex hulls of image sets of convex hulls.

The remarkable property evidenced by the previous theorem is that one can compute the convex hulls of the reachability sets by just propagating the vertices. Precisely, assume a vertex representation of the polytope $\hat{\mathcal{X}} = \mathcal{V}(x_1, x_2, \ldots, x_s)$ is known. Let A_i and E_i, $i = 1, 2, \ldots, r$, be the matrices generating $A(w)$ and $E(w)$ and let $\mathcal{D} = \mathcal{V}(D)$, where $D = [d_1, d_2, \ldots, d_h]$. Then the convex hull of the one-step reachability set (which might be non-convex) is given by the convex hull of all the points of the form $A_i x_k + E_i d_j$, say its expression is

$$conv\left\{\mathcal{R}_1\left(\hat{\mathcal{X}}\right)\right\} =$$
$$= conv\left\{A_i x_k + E_i d_j,\ i = 1,2,\ldots,r,\ k = 1,2,\ldots,s,\ \ j = 1,2,\ldots,h\right\} \tag{6.4}$$

Therefore, a routine which propagates the vertices of the sets $\hat{\mathcal{X}}_k$ can be easily constructed. Its application is a different story. Indeed, the complexity of the problem is enormous, since the number of candidate vertices grows exponentially. One can apply the mentioned methods to remove internal points, but still the number of true vertices might explode in few steps even for small dimensional problems. The reader is referred to [RKKM05b, RKK$^+$05] for more recent results on this construction.

Example 6.6. As previously mentioned, the one-step forward reachability set of a convex set is in general non-convex [BS79]. Here, by means of a simple two-dimensional system, another graphical representation of such lack of convexity is reported. Consider the two-dimensional autonomous uncertain system

$$x(k+1) = A(w(k))x(k)$$

with $|w(k)| \leq 1$ and

$$A(w) = \begin{bmatrix} 1/3 & -2/3\,w \\ -2/3 + 2/3\,w & 1/2 \end{bmatrix}$$

and the set

$$\mathcal{X} = \{x : \|x\|_\infty \leq 1\}$$

The one step forward reachability set (computed on a grid of points) is depicted in figure 6.2

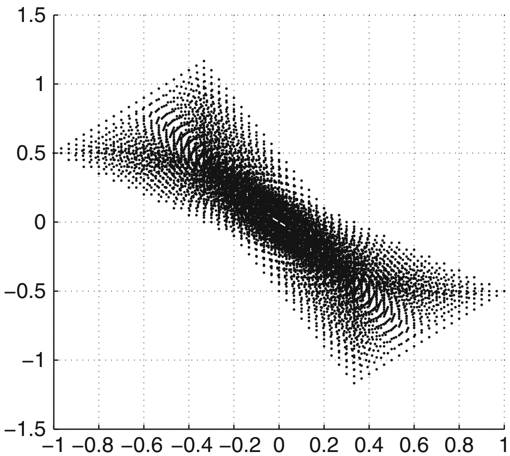

Fig. 6.2 Image (computed on a grid) of the one step reachable set for example 6.6

It is immediately seen that this set is non-convex (to double check such sentence one can try to determine whether there exist $x \in \mathcal{X}$ and a value $-1 \leq w \leq 1$ such that $A(w)x = [1/2,\ 0]^T$).

Conversely the preimage set is convex and precisely

$$\mathcal{S} = \{x : \|A(1)x\|_\infty \leq 1,\ \|A(-1)x\|_\infty \leq 1\}$$

If A and E are certain matrices (i.e., the family of matrices are singletons), then the following result, reported without proof, holds:

Proposition 6.7. *Let $\mathcal{X}_0$ and $\mathcal{D}$ be polytopes. Consider the system*

$$x(t+1) = Ax(t) + Ed(t), \quad \textit{with } d \in \mathcal{D}$$

Then $\mathcal{X}_k$, the k-step reachability set from $\mathcal{X}_0$, is a polytope.

Again, if $\mathcal{X}$ and $\mathcal{D}$ are known, then (6.4), as a special case, provides the expression of the one-step reachability set. It is worth mentioning that the propagation of the uncertainty effect cannot be achieved by considering ellipsoidal sets. Indeed, even in the case in which no parametric uncertainty is present, the one step reachable set from an ellipsoid is convex, but it is not an ellipsoid.

We have seen that the attempt of propagating the disturbance effect forward in time can be frustrating even in the case of linear systems, if parameter uncertainties are to be considered. Thus, working with reachable sets forward in time, unless for the special case of linear systems with no parameter uncertainties, is very hard. The reader is referred to [RKKM05a, RK07, LO05] for recent results on the topic. Luckily enough, there is another bullet to shoot, the controllability one. It will soon be shown that, by working backward in time, it is possible to keep convexity, a property which allows to derive efficient numerical algorithms. We will consider this aspect later when the concept of worst case-controllability will be considered.

6.1.3 Propagation of uncertainties via ellipsoids

A known method to investigate the effect of uncertainty is the adoption of ellipsoids. However, as already mentioned, the methods based on ellipsoids are conservative, since they are usually unfit to describe the true reachability set. However, they typically require less computational effort. We remind the an ellipsoid with center c, radius 1 and characterizing matrix $G^{-1} \succ 0$ is denoted by

$$\mathcal{E}(c, G^{-1}, 1) = \{x : (x-c)^T G^{-1} (x-c) \leq 1\}$$

Let us consider the case of the following linear system

$$\dot{x}(t) = Ax(t) + Bu(t) + Ed(t)$$

where $u(t)$ is a known input and $d(t)$ is an uncertain input, bounded as

$$d \in \mathcal{E}(0, G^{-1}, 1)$$

(i.e., $d^T G^{-1} d \leq 1$). Let us assume that the state is initially confined in the following ellipsoid with center c_0

$$x(0) \in \mathcal{E}(c_0, Q^{-1}, 1)$$

Then the state of the system is confined at each time in the ellipsoid[2] (see [Sch73], section 4.3.3)

$$x(t) \in \mathcal{E}(c(t), Q^{-1}(t), 1) \tag{6.5}$$

where the center $c(t)$ and the matrix $Q^{-1}(t)$ describing the ellipsoid satisfy the following equations

$$\dot{c}(t) = Ac(t) + Bu(t) \tag{6.6}$$

$$\dot{Q}(t) = AQ(t) + Q(t)A^T + \beta(t)Q(t) + \beta(t)^{-1}EGE^T \tag{6.7}$$

where $\beta(t)$ is an arbitrary positive function. A discussion on how to choose the free function β to achieve some optimality conditions for the ellipsoid $\mathcal{E}(c(t), Q^{-1}(t), 1)$ is proposed in [Che81, CO04, Sch73]. The reader is referred to the recent survey books [Che94, KV97] for a more complete overview. It is worth noticing that, in the case of a stable A, assuming $u = 0$ and a constant function β, the asymptotic value of Q is achieved by setting $\dot{Q} = 0$, thus achieving, as a particular case, equation (4.23). Note also that, by setting $Q(0) = 0$ (in this case the expression $\mathcal{E}(0, Q^{-1}(t), 1)$ has no significance for $t = 0$), the initial state is set to 0. Then the corresponding set $\mathcal{E}(0, Q^{-1}(t), 1)$ (defined for $t > 0$) includes the set of states reachable in time t from the origin. It will be seen how to compute, at least approximately, the reachability set from 0.

There is a corresponding equation for discrete-time ellipsoidal confinement. The reader is referred to specialized literature (see, for instance, [Sch73], Section 4.3.2)

[2] In general the inclusion is quite conservative.

6.2 0-Reachable sets with bounded inputs

In this section, a specific problem, and precisely that of estimating the 0-reachable set of a linear time invariant system, will be considered. We consider the reachable sets with pointwise bounded inputs for both discrete and continuous-time systems. We will consider also the problem of determining reachable sets with energy-bounded inputs (a problem elegantly solvable via ellipsoidal sets) although such a class of signals have not been considered in this book so far.

6.2.1 *Reachable sets with pointwise-bounded noise*

Consider initially the discrete-time system

$$x(t+1) = Ax(t) + Ed(t).$$

Assume that $d \in \mathcal{D}$ is a convex and closed set including the origin. Denote by $\mathcal{R}_T$ the set of all reachable states in T steps. It is not difficult to see that, since $0 \in \mathcal{D}$,

$$\mathcal{R}_T \subseteq \mathcal{R}_{T+1}$$

namely the sets $\mathcal{R}_T$ are nested. The T step reachability set is given by

$$\mathcal{R}_T = \sum_{k=0}^{T-1} A^k E \mathcal{D}.$$

and it can be recursively computed as follows:

$$\mathcal{R}_{T+1} = A\mathcal{R}_T + E\mathcal{D}$$

This involves known operations amongst sets, such as computing the sum and the image of a set (see Section 3.1.1, page 96). These operations can be done, in principle, in the environment of convex sets. However, for computational purposes, sticking to polyhedral sets is of great help. Let us assume that $\mathcal{D}$ is a polytope. Then, assuming the following vertex representation,

$$\mathcal{R}_T = \mathcal{V}[x_1^{(T)}, x_2^{(T)}, \ldots, x_{r_T}^{(T)}], \;\; \mathcal{D} = \mathcal{V}[d_1, d_2, \ldots, d_s]$$

the set $\mathcal{R}_{T+1}$ has the points $Ax_j^{(T)} + Ed_k$ as candidate vertices, precisely

$$\mathcal{R}_T = conv\left\{Ax_j^{(T)} + Ed_k, \; j = 1, \ldots, r_T, \; k = 1, \ldots, s\right\}$$

Again the number of candidate vertices grows exponentially. Therefore the algorithm may result difficult to apply to systems of high dimension. It is worth noticing that the generation of 0-reachability sets for polytopic systems $x(t+1) = A(w(t))x(t) + E(w(t))d(t)$, in view of the previous consideration, leads to sets that are non-convex. However, if one is satisfied with the convex hulls $conv\{\mathcal{R}_T\}$, these computations can be done in an exact way according to Proposition 6.5 and the operations in (6.4).

A different approach that may be used for evaluating reachable sets is based on the hyperplane representation of the set. Since $\mathcal{R}_T$ is convex and, in general, closed, it can be described by its support functional as

$$\mathcal{R}_T = \{x : z^T x \leq \phi_T(z), \ \ \forall z\}$$

The support functional $\phi_T(z)$ can be computed as follows. Denote by $\mathcal{D}_T = \mathcal{D} \times \mathcal{D} \times \cdots \times \mathcal{D}$, ($T$ times), the convex and compact set of finite sequences of T vectors in $\mathcal{D}$. The T-step reachability set is given by

$$\mathcal{R}_T = \left\{ x = [\, E \; AE \; A^2E \; \ldots \; A^{T-1}E \,] d_T, \quad d_T \in \mathcal{D}_T \right\}.$$

Therefore

$$\begin{aligned}
\phi_T(z) &= \sup_{d_T \in \mathcal{D}^T} \left\{ z^T [\, E \; AE \; A^2E \; \ldots \; A^{T-1}E \,] d_T \right\} = \\
&= \sum_{i=0}^{T-1} \sup_{d \in \mathcal{D}} z^T A^i E d \\
&= \sum_{i=0}^{T-1} \phi_{\mathcal{D}}(z^T A^i E),
\end{aligned}$$

where $\phi_{\mathcal{D}}(\cdot)$ is the support functional of $\mathcal{D}$. Therefore the evaluation of $\phi_T(z)$ at a point z requires the solution of the programming problem $\sup_{d \in \mathcal{D}} \ z^T A^i E d$. If $\mathcal{D}$ is a C-set, then "sup" is actually a "max." Remarkable cases are those in which $\mathcal{D}$ is the unit box of the p norm, with $1 \leq p \leq \infty$

$$\mathcal{D} = \{d : \|d\|_p \leq 1\}$$

For the above,

$$\max_{d \in \mathcal{D}} z^T A^i E d = \|z^T A^i E\|_q,$$

where q is such that $1/p + 1/q = 1$. In particular, if $\mathcal{D}$ is assumed to be a hyperbox, the components of d are all bounded as $|d_i| \leq \bar{d}_i$. Without restrictions, we can assume $|d_i| \leq 1$ a condition always achievable by scaling the columns of matrix E.

Then

$$\phi_T(z) = \sum_{i=0}^{T-1} \|z^T A^i E\|_1$$

Example 6.8. Let us consider a very simple example in which the computation can be carried out by hand (perhaps an isolated case in this book). Consider the matrices

$$A = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and let the disturbance set be $\mathcal{D} = [-1, 1]$. The reachable sets for the first three steps, also depicted in Fig. 6.3, are

$$\mathcal{R}_1 = \bar{\mathcal{V}}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right),$$

$$\mathcal{R}_2 = \bar{\mathcal{V}}\left(\begin{bmatrix} 3/2 & -1/2 \\ -1/2 & -1/2 \end{bmatrix}\right),$$

$$\mathcal{R}_3 = \bar{\mathcal{V}}\left(\begin{bmatrix} 3/2 & -1/2 & -3/2 \\ -1 & -1 & 0 \end{bmatrix}\right).$$

For the above system, consider the problem of determining the largest absolute value of the output $y(t) = x_2(t)$. This problem may be recast as follows: consider constraints of the form $z^T x \leq \mu$ and $-z^T x \leq \mu$, where $z = \begin{bmatrix} 0 & 1 \end{bmatrix}^T$ and determine the smallest value of μ such that the reachable set is inside a proper μ-thick strip (see Fig. 6.3).

$$\bar{\mathcal{P}}[z, \mu] = \left\{x : |z^T x| \leq \mu\right\}$$

It is immediately seen that such a value is the support functional of $\mathcal{R}_t$ computed in z. In this example the smallest value in three steps is $\mu_{min} = 1$.

This value can be computed by considering the expression

$$\mu_{min} = \phi_3(z) = \|z^T E\|_1 + \|z^T A E\|_1 + \|z^T A^2 E\|_1 = 0 + \frac{1}{2} + \frac{1}{2}$$

It is clear that, in principle, one could compute in an approximate way the infinite-time reachability set

$$\mathcal{R}_\infty = \bigcup_{k=0}^{\infty} \mathcal{R}_k$$

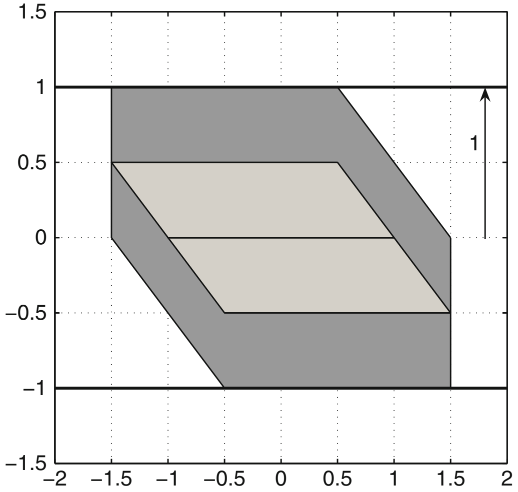

Fig. 6.3 The reachability sets for example 6.8

by computing $\mathcal{R}_k$ with k large. Clearly, this set would be an internal approximation. The following convergence property holds.

Proposition 6.9. *Assume that $\mathcal{D}$ is a C-set, that (A, E) is a reachable pair and that A is asymptotically stable. Then the set $\mathcal{R}_\infty$ is bounded and convex and its support functional is given by*

$$\phi_\infty(z) = \sum_{i=0}^{\infty} \phi_{\mathcal{D}}(z^T A^i)$$

Furthermore, $\mathcal{R}_k \to \mathcal{R}_\infty$ in the sense that for all $\epsilon > 0$ there exists $\bar{k}$ such that for $k \geq \bar{k}$

$$\mathcal{R}_k \subseteq \mathcal{R}_\infty \subseteq (1+\epsilon)\mathcal{R}_k$$

Proof. There are several proofs of the previous result in the literature, for instance [GC86b]. We just sketch the proof. The support functional is

$$\phi_\infty(z) = \sup_{d(\cdot)\in\mathcal{D}} z^T \sum_{h=0}^{\infty} A^h E d(h) =$$

$$= \sum_{h=0}^{\infty} \sup_{d(\cdot)\in\mathcal{D}} z^T A^h E d(h) = \sum_{h=0}^{\infty} \phi_{\mathcal{D}}(z^T A^h E)$$

This value is finite because $\|z^T A^i E\| \leq \|z^T\| \|A\|^i \|E\|$ converges to 0 exponentially, and then $\mathcal{R}_\infty$ is bounded. As far as the inclusions are concerned, the first one is obvious. The second can be proved as follows. Fix $\bar{k} > 0$. Any reachable state in

$k \geq \bar{k}$ steps can be written as

$$x = \sum_{h=0}^{\bar{k}-1} A^h E d(h) + A^{\bar{k}} \sum_{h=0}^{k-\bar{k}} A^h E d(h+\bar{k}) =$$
$$= \sum_{h=0}^{\bar{k}-1} A^h E d(h) + A^{\bar{k}} s_k \in \mathcal{R}_{\bar{k}} + A^{\bar{k}} \mathcal{R}_{k-\bar{k}}$$

where $s_k \in \mathcal{R}_{k-\bar{k}}$ (it is understood that if $k \leq \bar{k}$ then $s_k = 0$), so $\|s_k\| \leq \mu$ for some positive μ. Then

$$\nu_{\bar{k}} \doteq \|A^{\bar{k}}\|\mu \to 0 \quad \text{as } \bar{k} \to \infty$$

and, denoting by $\mathcal{B} = \mathcal{N}[\|\cdot\|, 1]$ the unit ball of $\|\cdot\|$ we have that the k-step reachable state is

$$\mathcal{R}_k \subseteq \mathcal{R}_{\bar{k}} + \nu_{\bar{k}} \mathcal{B} \subseteq (1+\epsilon)\mathcal{R}_{\bar{k}}$$

This, in turn, implies that any reachable state is in $(1+\epsilon)\mathcal{R}_{\bar{k}}$.

By means of the just reported property one can compute an internal approximation of the set $\mathcal{R}_\infty$ by computing $\mathcal{R}_k$. By the way, it turns out that each of the sets $\mathcal{R}_k$, under the assumption of the theorem, is a C-set as long as $\mathcal{D}$ is a C-set. The same property does not hold for the set $\mathcal{R}_\infty$, which is convex and bounded, but in general is not closed. This assertion is easily proved by the next scalar counterexample.

Example 6.10. The infinite time reachability set for the system

$$x(t+1) = \frac{1}{2} x(t) + d(t), \quad |d(t)| \leq 1$$

is clearly the open interval $\mathcal{R}_\infty = (-2, 2)$. We stress that we defined the reachability set as the set of all states that can be reached in *finite time* $0 < T < \infty$, although for T arbitrary. This is why the extrema are not included.

The situation is different if one considers the set $\bar{\mathcal{R}}_\infty = \sum_{k=0}^{\infty} A^k E \mathcal{D}$ which is closed, indeed the closure of $\mathcal{R}_\infty$ [RK07].

To achieve an external approximation one can use several tricks. The first one is that of "enlarging" $\mathcal{D}$. Indeed, if the reachability set $\mathcal{R}_t^\epsilon$ with disturbances $d \in (1+\epsilon)\mathcal{D}$ is considered, by linearity the condition

$$\mathcal{R}_t^\epsilon = (1+\epsilon)\mathcal{R}_t$$

is obtained, thus achieving an external approximation. A different trick is that of computing the reachable set for the modified system

$$x(t+1) = \frac{A}{\lambda}x(t) + \frac{E}{\lambda}d(t) \tag{6.8}$$

Denoting by $\bar{\lambda}_{max} = \max\{|\lambda_i|, \quad \lambda_i \in eig(A)\}$, for $\bar{\lambda}_{max} < \lambda \leq 1$, the system remains stable, so that the reachability sets of the modified system $\mathcal{R}_k^\lambda$ are bounded. For any $0 < \lambda < 1$, $A^k E\mathcal{D} \subset (A/\lambda)^k(E/\lambda)\mathcal{D}$, and then

$$\mathcal{R}_T \subset \mathcal{R}_T^\lambda = \sum_{k=0}^{T-1} \left(\frac{A}{\lambda}\right)^k \frac{E}{\lambda}\mathcal{D}$$

For λ approaching 1, $\mathcal{R}_k^\lambda$ approaches $\mathcal{R}_k$ from outside. An interesting property is the following.

Proposition 6.11. *Assume that $\mathcal{D}$ is a C-set, that (A, E) is a reachable pair and that A is asymptotically stable. Let $\lambda < 1$ be such that A/λ is stable. Then there exists $\bar{k}$ such that, for $k \geq \bar{k}$, the set $\mathcal{R}_k^\lambda$, computed for the modified system (6.8), is robustly positively invariant for the original system and*

$$\mathcal{R}_\infty \subset \mathcal{R}_k^\lambda$$

The proof of the above proposition can be deduced by the fact that:

- $\mathcal{R}_\infty^\lambda$ is positively invariant for the modified system (this fact will be reconsidered later) and then, in view of Lemma 4.31, it is contractive for the original system;
- $\mathcal{R}_k^\lambda \to \mathcal{R}_\infty^\lambda$, which has been shown in Proposition 6.9.

We refer the reader to [RKKM05a, RKK+05] for further details on this kind of approximations.

Let us now consider the problem of evaluating the reachability set for continuous-time system

$$\dot{x}(t) = Ax(t) + Ed(t).$$

It is at this point rather clear that the problem cannot be solved by considering a sequence $\mathcal{R}_t$, because such a set is not polyhedral even if $\mathcal{D}$ is such. Therefore the hyperplane method, previously considered for discrete-time systems and based on the support functional, seems the most appropriate. Let $\mathcal{R}_t$ be the set of all the states reachable in time t from the origin, with $\mathcal{D}$ a C-set. Let us consider the support functional $\phi_t(z)$ of $\mathcal{R}_t$. Then, in view of the following chain of equalities

$$\phi_t(z) = \sup_{d \in \mathcal{D}} z^T \int_0^t e^{A\sigma} Ed(\sigma)d\sigma = \int_0^t \sup_{d(\sigma) \in \mathcal{D}} z^T e^{A\sigma} Ed(\sigma)d\sigma =$$
$$\int_0^t \phi_{\mathcal{D}}(z^T e^{A\sigma} E)d\sigma$$

the reachability set in time t turns out to be the convex set characterized in terms of support functional as

$$\mathcal{R}_t = \left\{ x : z^T x \le \phi_t(z), \ \forall z \right\}$$

The reader can enjoy her/himself in investigating special cases of $\mathcal{D}$ given by her/his preferred norms. Let us consider the single input case and the set $\mathcal{D} = [-1, 1]$. In this case the support functional of $\mathcal{D}$ is $\phi_{\mathcal{D}}(\delta) = |\delta|$ and then

$$\phi_t(z) = \int_0^t |z^T e^{A\sigma} E| \, d\sigma \tag{6.9}$$

By (possibly numerical) integration, as done and reported graphically in the next example, it is possible to determine $\phi_t(z)$ and $\phi_\infty(z)$, the support functional of $\mathcal{R}_\infty$, at least approximately.

Example 6.12. By using Eq. (6.9), the reachable sets $\mathcal{R}_t$ for the continuous-time dynamic system

$$\dot{x}(t) = \begin{bmatrix} -0.3 & 1 \\ -1 & -0.3 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} d(t)$$

when the disturbance is bounded as $|d(t)| \le 1$, were computed for $t = 1, \ 2, \ 4, \ 100$. Such sets are depicted in Figure 6.4.

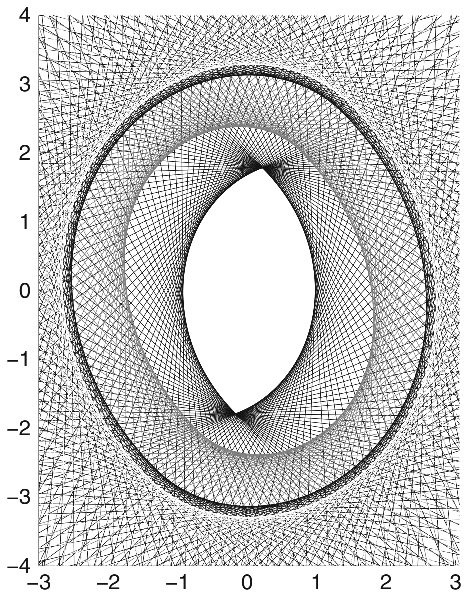

Fig. 6.4 Reachable sets with pointwise-bounded noise for Example 6.12, computed for different values of t

Henceforth we establish some properties that concern the reachability sets in both continuous and discrete-time case. We start with the following basic fact.

Proposition 6.13. *Assume that $\mathcal{D}$ is a C-set, (A,E) is a reachable pair and that A is asymptotically stable. Then $\mathcal{R}_\infty$ is the smallest robustly-positively-invariant set for the system, precisely any robustly-positively-invariant set includes $\mathcal{R}_\infty$.*

Proof. The discrete-time case only is considered (the continuous-time case proof is in practice the same). The fact that $\mathcal{R}_\infty$ is robustly-positively-invariant is obvious.

Proving minimality is equivalent to showing that any invariant set $\mathcal{P}$ contains $\mathcal{R}_\infty$, say $x \in \mathcal{R}_\infty$ implies $x \in \mathcal{P}$. Assume then that $\mathcal{P}$ is invariant, let k be arbitrary and let $\bar{x} \in \mathcal{R}_k$ be an arbitrary vector. Let $\mathcal{B}$ be any neighborhood of the origin such that any state of $\mathcal{B}$ can be driven to 0 in finite time with a signal $d(t) \in \mathcal{D}$. Consider any $x(0) \in \mathcal{P}$ and let $d(t) = 0$ until the state $x(t_1) \in \mathcal{B}$. For $t \geq t_1$ take the sequence $d(t) \in \mathcal{D}$ which drives the state to zero at t_2, $x(t_2) = 0$. Then consider the sequence of further k inputs $d(t) \in \mathcal{D}$ which drive the state $x(t_3) = \bar{x}$. Since $\mathcal{P}$ is robustly invariant, $\bar{x} \in \mathcal{P}$. Since $\bar{x}$ is arbitrary, $\mathcal{P}$ contains any point of $\mathcal{R}_k$ and since k is also arbitrary, $\mathcal{P}$ contains any point of $\mathcal{R}_\infty$.

The set $\mathcal{R}_\infty$ is the limit set of the stable system (A,E). In other words, for any $x(0)$, $x(t) \rightarrow \mathcal{R}_\infty$[3]. This fact is important because it allows to characterize the asymptotic behavior of a system. As an example of application, let us consider the problem of characterizing the worst case state estimation error.

Example 6.14 (Observer asymptotic error). Let us consider an observer-based control for the system

$$(x(t+1)) \;\; \dot{x}(t) = Ax(t) + Bu(t) + Ed(t)$$
$$y(t) = Cx(t) + v(t)$$

in which d and v are external inputs subject to $d(t) \in \mathcal{D}$ and $v \in \mathcal{V}$. In most cases these inputs represent noise and cannot be measured. If a standard linear observer is designed,

$$(z(t+1)) \;\; \dot{z}(t) = (A+LC)z(t) + Bu(t) - Ly(t)$$
$$e(t) = z(t) - x(t)$$

where $e(t)$ is the error, the error equation results in

$$(e(t+1)) \;\; \dot{e}(t) = (A+LC)e(t) - Lv(t) - Ed(t) \tag{6.10}$$

It is apparent that, under persistent noises d and v, the observer error does not vanish asymptotically. The asymptotic effect of the noise can be clearly evidenced

[3] In the sense that $\delta(x(t), \mathcal{R}_\infty)$, the distance from $x(t)$ to $\mathcal{R}_\infty$ converges to 0.

by computing the reachability set of the error system (6.10). If an invariant set $\mathcal{E}$ for this system is computed, then it is possible to assure that, whenever $e(0) \in \mathcal{E}$,

$$x(t) \in z(t) - \mathcal{E}$$

$t > 0$. We will come back on this problem in Chapter 11.

So far the problem of determining the reachability set has been considered under the following assumptions: reachability of (A, E), asymptotic stability of A, and $\mathcal{D}$ a C-set. The assumption that $\mathcal{D}$ is a C-set is reasonable. If $\mathcal{D}$ has an empty interior, but 0 is inside the relative interior[4] then it is possible to reconsider the problem by redefining E. to this subspace, namely by involving a new matrix ED, where D is any basis of such a subspace.

Conversely there are cases in which the constraint set does not include 0 as an interior point. In this case the problem has to be managed in a different way. For instance, one can decompose $d \in \mathcal{D}$ by choosing a constant $d_0 \in int\mathcal{D}$. Then $d = d_0 + d_1$, where $d_1 \in \mathcal{D}_1 = \mathcal{D} - d_0$. Now, the translated disturbance d_1 is in a C-set $\mathcal{D}_1$. The effect of d_0 and d_1 can be investigated separately. An interesting case is that of systems with positive controls. We do not analyze this case but we refer the reader to specialized literature, such as [FB97].

Then let us still assume that $\mathcal{D}$ is a C-set, but let us remove the stability or the reachability assumption.

Proposition 6.15. *For the* 0*-reachability sets* $\mathcal{R}_k$ *the following properties hold:*

- $\mathcal{R}_k \subseteq$ *reach*(A, E)*, the reachable space of* (A, E)*.*
- $\mathcal{R}_\infty$ *is bounded if and only if the reachable sub-system of* (A, E) *is asymptotically stable.*
- *Assume that* (A, E) *is reachable, and denote by* X_{sta} *and* X_{uns} *the eigen-spaces of* A *associated with the stable and the unstable modes. Then the reachable set is given by*

$$\mathcal{R}_\infty = \mathcal{R}_\infty^{sta} + X_{uns}$$

where $\mathcal{R}_\infty^{sta}$ *denotes the set of reachable states in the subspace* X_{sta}*.*

Proof. The first statement is obvious. The second statement is obvious in its sufficient part because, by the previous statement we can consider the reachable sub-system and conclude that the reachability set is bounded. As far as necessity is concerned, assume that the reachable sub-system is unstable. Then, by means of a bounded input, it is possible to reach from $x(0) = 0$ an eigenvector $\bar{v}$ associated with an unstable eigenvalue λ (in general an unstable subspace) in time $[0, \bar{t}]$ and, assuming $d(t) = 0$ for $t > \bar{t}$ so that $x(t) = e^{\lambda(t-\bar{t})}\bar{v}$, it is immediate to see that

[4] 0 is in the interior relatively to the smallest subspace including $\mathcal{D}$.

$x(t)$ cannot be bounded. The third statement requires more work and its proof is postponed to the problem of controllability of systems with bounded-control.

6.2.2 *Infinite-time reachability and l_1-norm*

We now investigate an important connection between the infinite-time reachability set $\mathcal{R}_\infty$ and the l_1-norm, often referred to as ∞ to ∞ induced norm or peak-to-peak norm of a system. Consider the SISO stable system (A, E, H)

$$\begin{aligned} x(t+1) &= Ax(t) + Ed(t) \\ y(t) &= Hx(t) \end{aligned}$$

The ∞ to ∞ induced norm is defined as

$$\|H(zI-A)^{-1}E\|_{\infty,\infty} \doteq \sup_{t\geq 0, x(0)=0, |d(t)|\leq 1} |y(t)|$$

The reason why this norm is referred to as l_1-norm is that it turns out to be the l_1-norm [DP87] of the sequence of Markov parameters

$$\|H(zI-A)^{-1}E\|_{\infty,\infty} = \|H(zI-A)^{-1}E\|_{l_1} \doteq \sum_{k=0}^{\infty} |HA^kE|$$

In the general case of a MIMO (possibly not strictly proper) system the l_1-norm can be defined by replacing $|\cdot|$ by $\|\cdot\|$, precisely

$$\|H(zI-A)^{-1}E + D\|_{\infty,\infty} \doteq \sup_{t\geq 0, x(0)=0, \|d(t)\|\leq 1} \|y(t)\|$$

Such a norm can be evaluated as the sum of a series [DP87]

$$\begin{array}{c} \|H(zI-A)^{-1}E + D\|_{\infty,\infty} = \\ \|H(zI-A)^{-1}E + D\|_{l_1} \doteq \max_i\{\|D_i\|_1 + \sum_{k=0}^{\infty} \|[HA^kE]_i\|_1\} \end{array} \tag{6.11}$$

where D_i and $[HA^kE]_i$ denote the ith row of the matrices D and HA^kE, respectively. A set-theoretic equivalent condition is given in the next proposition.

Proposition 6.16. *Consider the asymptotically stable system* (A, E, H) *(i.e., assume* $D = 0$*). Then the smallest value* μ_{inf} *of* μ *such that* $\mathcal{R}_\infty$ *is included in the strip*

$$\bar{\mathcal{P}}[H, \mu\bar{1}] = \{x : \|Hx\| \leq \mu\}$$

is equal to the l_1-norm of the system:

$$\mu_{inf} = \inf\{\mu : \mathcal{R}_\infty \subset \bar{\mathcal{P}}[H, \mu\bar{1}]\} = \|H(zI - A)^{-1}E\|_{\infty,\infty} = \|H(zI - A)^{-1}E\|_{l_1}$$

In the single output case this is the support functional of $\mathcal{R}_\infty$ evaluated in H, i.e. $\phi_\infty(H) = \phi_\infty(-H)$ (by symmetry).

When D is non-zero, the following holds:

Proposition 6.17. *Consider the asymptotically stable system (A, E, H, D), with p outputs. Then the l_1-norm of $\|(A, E, H, D)\|_{l_1}$ is the smallest value of μ for which the 0-reachability set $\mathcal{R}_\infty$ is included in the set*

$$\bar{\mathcal{P}}[H, \tilde{\mu}\bar{1}]\,\{x : \|H_i x\| \leq \mu - \|D\|_1, \;\; i = 1, 2, \ldots, p\}$$

where $\tilde{\mu} = \mu - \|D\|_1$

Proof. It is known [DP87] hat the l_1-norm condition $\|H(zI - A)^{-1}E + D\|_{l_1} = \mu$ is equivalent to the fact that for $x(0) = 0$, the condition $\|y(t)\|_\infty \leq \mu$ holds for all $\|d(t)\|_\infty \leq 1$, namely,

$$-\mu \leq y_i(t) \leq \mu,$$

which is, in turn, equivalent to

$$-\mu \leq H_i x(t) + D_i d(t) \leq \mu$$

for all $\|d(t)\|_\infty \leq 1$. Since the current value of $d(t)$ does not depend on $x(t)$ and can be any arbitrary vector with ∞-norm not greater than 1, it is possible to write

$$-\min_{\|d\|_\infty \leq 1} D_i d - \mu \leq H_i x(t) \leq \mu - \max_{\|d\|_\infty \leq 1} D_i d$$

Then the proof is completed since

$$-\min_{\|d\|_\infty \leq 1} D_i d = \max_{\|d\|_\infty \leq 1} D_i d = \|D_i\|_1$$

The previous proposition represents an interesting interpretation of the $\|\cdot\|_{l_1}$ norm of a system in terms of reachability. In practice, the $\|\cdot\|_{l_1}$ norm less than μ is equivalent to the inclusion of $\mathcal{R}$ in $\bar{\mathcal{P}}[H, \mu\bar{1}]$. It will be soon shown that this interpretation is very useful to compute the peak-to-peak induced norm in those cases (i.e., polytopic systems) in which the computation via Markov parameters is not possible.

6.2.3 Reachable sets with energy-bounded noise

In this section, a characterization of the disturbances which is unusual in the book is analyzed. Precisely, the focus of the present section are linear dynamic systems of the form

$$\dot{x}(t) = Ax(t) + Ed(t).$$

with disturbances bounded as follows:

$$\int_0^\infty d^T(t)Rd(t)\,dt \leq 1, \quad \text{with} \quad R \succ 0$$

To avoid unnecessary complications, it is assumed, without lack of generality, that $R = I$, since if this is not the case one can replace the matrix E by $ER^{-1/2}$ and consider the input $\hat{d} = R^{1/2}d$. Let us then assume

$$\int_0^\infty d^T(t)d(t)\,dt \leq 1. \tag{6.12}$$

Denote by $\mathcal{B}(t)$ the set of all the functions having energy bounded by 1 on the interval $[0,t]$, precisely such that

$$\mathcal{B}(t) = \left\{ d(t) : \int_0^t d^T(t)d(t)\,dt \leq 1 \right\}$$

Note that the set of reachable states with inputs $d \in \mathcal{B}(t)$ is non-decreasing with t, precisely, $\mathcal{B}(t')$ includes $\mathcal{B}(t)$ for $t' > t$. Let us consider the set of all 0-reachable states with inputs bounded as above. It turns out that this set is an ellipsoid according to the following theorem. We remind that an ellipsoid $\mathcal{D}(Q) = \mathcal{D}(Q,1)$ can be described as in (3.14)

$$\mathcal{D}(Q) = \left\{ x : z^T x \leq \sqrt{z^T Q z}, \;\; \text{for all } z \right\}$$

where $\sqrt{z^T Q z}$ is its support functional [5]

Theorem 6.18. *Let A be a stable matrix and let (A,E) be a reachable pair. The closure of the set of all the states reachable from $x(0) = 0$ with inputs bounded as in (6.12) is given by the ellipsoid $\mathcal{D}(Q)$, where Q is the reachability Gramian, i.e. the unique solution of*

$$QA^T + AQ = -EE^T$$

---

[5] Note that $\mathcal{D}$ has not the same meaning of the previous subsection, but represents now the ellipsoid.

Proof. Consider any state $x(t)$ reachable at time t with energy bounded as

$$\int_0^t d^T(\sigma)d(\sigma)\, d\sigma \leq 1.$$

Take any vector z and consider the following optimization problem

$$\mu_t = \sup_{d \in \mathcal{B}(t)} z^T x(t) = \sup_{d \in \mathcal{B}(t)} \int_0^t z^T e^{A\sigma} E d(t-\sigma)\, d\sigma = \sup_{d \in \mathcal{B}(t)} \left(z^T e^{A(\cdot)} E, d(\cdot) \right)$$

where $(\cdot,\cdot)$ is a scalar product in the Hilbert space of the square measurable functions defined on the time interval $[0,t]$ with values in $\mathbb{R}^p$ [Lue69]. Such an optimization problem has solution

$$\mu_t = \|z^T e^{A(\cdot)} E\|_2 = \sqrt{\int_0^t z^T e^{A\sigma} E E^T e^{A^T\sigma} z d\sigma} = \sqrt{z^T Q(t) z}$$

where

$$Q(t) \doteq \int_0^t e^{A\sigma} E E^T e^{A^T\sigma} d\sigma$$

Therefore the set of all reachable states in time t is the ellipsoid $\mathcal{D}(Q(t))$. By obvious mathematical speculations, such an ellipsoid is non-decreasing with t, precisely, $z^T Q(t) z \leq z^T Q(t') z$ for $t \leq t'$. Now consider the identity

$$\int_0^t \frac{d}{d\sigma}\left[e^{A\sigma} E E^T e^{A^T\sigma}\right] d\sigma = \int_0^t \left[A e^{A\sigma} E E^T e^{A^T\sigma} + e^{A\sigma} E E^T e^{A^T\sigma} A^T\right] d\sigma$$
$$= A\left[\int_0^t e^{A\sigma} E E^T e^{A^T\sigma} d\sigma\right] + \left[\int_0^t e^{A\sigma} E E^T e^{A^T\sigma} d\sigma\right] A^T = AQ(t) + Q(t)A^T$$

On the other hand, we can write the same quantity as

$$\int_0^t \frac{d}{d\sigma}\left[e^{A\sigma} E E^T e^{A^T\sigma}\right] d\sigma = e^{At} E E^T e^{A^T t} - E E^T =$$
$$= \frac{d}{dt}\int_0^t e^{A\sigma} E E^T e^{A^T\sigma}\, d\sigma - E E^T = \dot{Q}(t) - E E^T$$

and notice that $Q(t)$ is solution of the following equation

$$\dot{Q}(t) = AQ(t) + Q(t)A^T + E E^T \tag{6.13}$$

Now, since A is stable, a finite limit value $Q = \lim_{t\to\infty} Q(t)$ exists and is achievable by setting $\dot{Q} = 0$. Then the theorem is proved, if we remind that $\mathcal{D}(Q(t))$ is non-decreasing and then included in the limit value $\mathcal{D}(Q)$. Moreover, for any $\bar{x}$ in the boundary of $\mathcal{D}(Q)$, we can find points in $\mathcal{D}(Q(t))$ arbitrarily close to $\bar{x}$, so $\mathcal{D}(Q)$ is the closure of all $\mathcal{D}(Q(t))$.

The discrete-time version of the theorem is the following:

Theorem 6.19. *Consider the system*

$$x(t+1) = Ax(t) + Ed(t)$$

with A stable and (A,E) reachable. The closure of the set of all the states reachable from $x(0) = 0$ with inputs bounded as

$$\sum_{t=0}^{\infty} d(t)^T d(t) \leq 1$$

is given by the ellipsoid $\mathcal{D}(Q)$, where Q is the discrete-time reachability Gramian which is the unique solution of

$$AQA^T - Q = -EE^T$$

Proof. (Sketch). The proof of the theorem is basically the same as the previous one. Let

$$Q(t) = \sum_{k=0}^{t-1} A^k EE^T (A^T)^k$$

so that

$$z^T [Ed(t-1)\; AEd(t-2)\; A^2Ed(t-3) \ldots A^{t-1}Ed(0)] = \sqrt{z^T Q(t) z},$$

say the ellipsoid $\mathcal{D}(Q(t))$ is the t-step reachability set with bounded energy. The matrix $Q(t)$ clearly satisfies the equation

$$Q(t+1) = AQ(t)A^T + EE^T.$$

and its limit value is the solution of the Lyapunov equation in the theorem statement.

Remark 6.20. The same results might have been obtained, both in the discrete and the continuous-time case, by resorting to the adjoint operator theory. We have skipped that powerful and elegant approach, since the main focus here has been put on set-theoretic aspects.

6.2.4 Historical notes and comments

The history of set-propagation is wide, especially as far as the computation of reachable sets is concerned. The first contribution are due to the seminal works of Bertesekas and Rhodes [BR71a] and Glower and Schweppe [GS71, Sch73], followed by several further contributions of which only a portion is mentioned in this book. We have mentioned the work by Chernousko and Kurzhanski[Che81, KV97, Che94], which provided techniques for ellipsoidal approximation. Considering the problem of the computation of the reachability sets, the available literature is enormous and providing a survey is a major challenge. Among the first contributions, it has to be mentioned [PN71], where several types of input bounds have been considered and [HR71], where a numerical algorithm for the determination of reachable sets via amplitude bounded inputs is provided. In [GC86b] and [GC87] it has been exploited the fact that the 0-reachable sets are the 0-controllable set for the inverse system and an algorithm based on polyhedral sets has been proposed. The hyperplane method idea is due to [SS90b] and [GK91b]. See also [Gay86, Las87, SS90a, Las93] for further results on the topic.

For further references, we refer to the survey [Gay91] or to [RKKM05a] for some more recent contributions concerning the computation and approximation of the minimal invariant set [RKKM05a].

6.3 Stability and convergence analysis of polytopic systems

Stability analysis is a fundamental problem in system theory. For linear systems this trivial task requires the computation of the eigenvalues of a matrix. This method cannot be applied when dealing with an uncertain system. Let us consider again a system of the form

$$\begin{array}{c} x(t+1) = A(w(t))x(t), \quad \text{(respectively, } \dot{x}(t) = A(w(t))x(t)) \\ A(w) = \sum_{i=1}^{s} A_i w_i, \quad \sum_{i=1}^{s} w_i = 1, \quad w_i \geq 0 \end{array} \tag{6.14}$$

with the basic questions:

- is the system stable?
- assumed that it is, how fast does it converge?

These questions will be faced next by means of quadratic and non-quadratic Lyapunov functions.

6.3.1 Quadratic stability

One approach to the problem is inspired by the well-known fact that any stable linear time-invariant system admits a quadratic Lyapunov function, leading to the following criterion.

Theorem 6.21. *The discrete-time (resp. continuous-time) system (6.14) is stable if all the systems share a common quadratic Lyapunov function, equivalently, if there exists a positive definite matrix P such that*

$$A_i^T P A_i - P \prec 0, \quad \textit{(respectively } A_i^T P + P A_i \prec 0\textit{)}$$

for $i = 1, 2, \ldots, s$.

Corollary 6.22. *The condition of the theorem is equivalent to the existence of $\epsilon > 0$, $\beta > 0$ or $0 \leq \lambda < 1$ such that, for all i*

$$A_i^T P A_i - \lambda^2 P \preceq 0, \quad \textit{(respectively } A_i^T P + P A_i + 2\beta P \preceq 0\textit{)}, \quad P \succ \epsilon I$$

The easy proof of this theorem (the corollary follows obviously) is not reported here. We will come back on it later, when we will show that the provided condition is sufficient, but not necessary at all. To provide necessary and sufficient conditions one might think about resorting to another family of Lyapunov functions. The class of polyhedral Lyapunov functions is an appropriate one as we will show soon.

6.3.2 Joint spectral radius

To provide non-conservative and constructive solutions to the stability analysis of a Linear Parameter-Varying (LPV) system one can consider the procedure for the construction of the largest invariant and the basic finite determination of Theorem 5.17. To investigate on this matter, a connection with the joint spectral radius is established.

Given a square matrix A its spectral radius is defined as the largest modulus of its eigenvalues $\Sigma(A) = \max\{|\lambda| : \lambda \in eig(A)\}$. For a set of matrices the joint spectral radius of the set is defined as the supremum of the spectral radius of all possible products of generating matrices.

Definition 6.23 (Joint spectral radius). Given a finite set of square matrices $[A_1, A_2, \ldots, A_s]$, the quantity

$$\Sigma(A_1, A_2, \ldots, A_s) \doteq \limsup_{k \geq 0} \max_{C_k \in \mathcal{I}_k} \Sigma(\Pi_{C_k})^{\frac{1}{k}} \tag{6.15}$$

is said the joint spectral radius of the family [RS60].

We remind that C_k is a string of k elements of $\{1, 2, \ldots, s\}$ and Π_{C_k} is the product of the matrices A_i indexed by the corresponding elements. The above quantity can be equivalently defined as

$$\Sigma(A_1, A_2, \ldots, A_s) = \limsup_{k\to\infty} \max_{C_k \in \mathcal{I}_k} \|\Pi_{C_k}\|^{1/k}.$$

(the quantity does not depend on the adopted norm) and it is related to the notion of Lyapunov exponent [Bar88a, Bar88b, Bar88c]. The following property is well known:

Proposition 6.24. *The robust exponential stability of the discrete-time system* $x(t+1) = A(w(t))x(t)$ *as in (6.14) is equivalent to* $\Sigma(A_1, A_2, \ldots, A_s) < 1$.

Proof. It is obvious that $x(t+1) = A(w(t))x(t)$ stable implies that the switching

$$x(t+1) = A(k)x(t), \quad A(k) \in \mathcal{A} = \{A_1, A_2, \ldots, A_s\}$$

is stable hence $\Sigma(A_1, A_2, \ldots, A_s) < 1$. The converse statement can be proved by using Proposition 6.5. Consider any initial polytopic set $\mathcal{X}_0$. The T-steps reachable set of the discrete inclusion is included in the convex hull of the points

$$A_{i_{T-1}} A_{i_{T-2}} \ldots A_{i_0} v_j, \quad A_{i_t} \in \mathcal{A},\ v_j \in vert\{\mathcal{X}_0\}$$

Thus, if $\Sigma(A_1, A_2, \ldots, A_s) < 1$, these points converge to 0 as $T \to \infty$.

The following theorem holds.

Theorem 6.25. *Assume that the matrices in the set have no common proper non-trivial invariant subspaces*[6]*. Then the following implications hold.*

i) If the spectral radius $\Sigma(A_1, A_2, \ldots, A_s) < 1$, *then for any initial polyhedral C-set* $\mathcal{X}$, *the largest invariant set* $\mathcal{S}$ *included in* $\mathcal{X}$ *is represented by a finite number of inequalities.*

ii) Conversely, if $\Sigma(A_1, A_2, \ldots, A_s) > 1$, *then there exists* $\bar{k}$ *such that*

$$\mathcal{S}^{(\bar{k})} \subset int\{\mathcal{X}\}$$

Proof. See [BM96b].

It has to be stressed that claim i) of the theorem holds even in the case in which the A_i share a common invariant subspace, which is the case of a single matrix A [GT91]. Statement ii) requires the assumption (see Exercise 11).

As previously mentioned, this implies that the procedure for computing $\mathcal{S}$ can be used to check the stability of a system. The following theorem formalizes this fact.

[6] Say there is no proper subspace $\mathcal{G}$, $\{0\} \neq \mathcal{G} \subset \mathbb{R}^n$ such that $A_i\mathcal{G} \subset \mathcal{G}$, for all i.

Let us consider the sequence (5.25) of sets $\mathcal{S}^{(k)}$ computed for the modified system

$$z(t+1) = \frac{A(w(t))}{\lambda} z(t) \tag{6.16}$$

(note that $x(t) = z(t)\lambda^t$ if $x(0) = z(0)$), which turns out to be

$$\mathcal{S}_\lambda^{(k)} = \{x : F\frac{\Pi_{C_h}}{\lambda^h} x \leq 1, \;\; C_h \in \mathcal{I}_h \;\; h = 0, 1, 2, \ldots, k\}$$

The next theorem formalizes some of the properties concerning the spectral radius.

Theorem 6.26. *Define the following numbers:*

$$\lambda_1 = \inf\{\lambda > 0 : \;\; \textit{the modified system (6.16) is stable}\}$$
$$\lambda_2 = \inf\{\lambda > 0 : \|x(t)\| \leq C\|x(0)\|\lambda^t, \;\; \textit{for some} \;\; C > 0\}$$
$$\lambda_3 = \inf\{\lambda > 0 : \;\; \mathcal{S}_\lambda^\infty \;\; \textit{is a C-set}\}$$
$$\lambda_4 = \inf\{\lambda > 0 : \;\; \mathcal{S}_\lambda^\infty = \mathcal{S}_\lambda^{(k)} \;\; \textit{for a finite } k\}$$

Then

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \Sigma(A_1, A_2, \ldots, A_s)$$

Proof. The fact that $\lambda_1 = \lambda_2 = \Sigma(A_1, A_2, \ldots, A_k)$ is a well-known result, see, for instance, [Bar88a, Bar88b, Bar88c]. The remaining equalities are immediate consequence of the previous theorem.

It follows immediately from Theorems 6.25 and 6.26 that, in the case of a single linear time-invariant system $x(t) = Ax(t)$, the procedure stops in a finite number of steps if $\Sigma(A) < \lambda$ for any C-set $\mathcal{X}$ or determines a set $\mathcal{S}^{(k)}$ which is in the interior of $\mathcal{X}$ if $\Sigma(A) > \lambda$ (this is in perfect agreement with the earlier result in [GT91]).

A remarkable consequence which can be drawn from Theorem 6.26 is that, in principle, the joint spectral radius can be approximately computed by bisection, by increasing (resp. decreasing) λ if the numeric procedures, applied to the modified system, stops unsuccessfully (resp. successfully). As previously pointed out, the procedure produces a number of constraints which increases enormously when $\lambda \simeq \Sigma$. This is in agreement with the work presented in [TB97, BT00] which analyzes the complexity of computing or approximating, the joint spectral radius of matrices and which can provide an explanation of this phenomenon (although there are particular interesting cases in which the complexity can be reduced, see [BNT05]) The reader is referred to [BN05] for more details and references on this topic. We will show later also that the considered type of procedures can be used to compute, beside the spectral radius, other performance indices for uncertain systems.

6.3.3 Polyhedral stability

To face the problem of robust stability analysis we can exploit the fact that polyhedral stability is equivalent to stability for an LPV system, as stated next.

Theorem 6.27. *The following statements are equivalent.*

1. *The discrete-time (continuous-time) system (6.14) is asymptotically stable.*
2. *The discrete-time (continuous-time) system (6.14) is exponentially stable, namely there exists (C,λ), $0 \leq \lambda < 1$ (resp. (C,β), $\beta > 0$) such that*

$$\|x(t)\| \leq C\|x(0)\|\ \lambda^t \tag{6.17}$$

(respectively

$$\|x(t)\| \leq C\|x(0)\|\ e^{-\beta t}) \tag{6.18}$$

3. *The system admits a polyhedral norm $\|Fx\|_\infty$ as a Lyapunov function. Precisely, there exists $0 \leq \lambda < 1$ (resp. $\beta > 0$) such that*

$$\|FA(w)x\|_\infty \leq \lambda\|Fx\|_\infty, \ \ (resp.\ \ D^+\|FA(w)x\|_\infty \leq -\beta\|Fx\|_\infty,)\ \forall w$$

4. *All the vertex systems $x(t+1) = A_i x(t)$ share a common polyhedral Lyapunov function $\|Fx\|_\infty$.*
5. *For any signal $v(t)$, with $\|v(t)\| \leq \nu$, there exist β, C_1 and C_2 such that the solution of the system $x(t+1) = A(w(t))x(t) + v(t)$ (resp. $\dot{x}(t) = A(w(t))x(t) + v(t)$) is bounded as*

$$\|x(t)\|_* \leq C_1\|x(0)\|\ \lambda^t + C_2, \ \ (resp.\ \|x(t)\|_* \leq C_1\|x(0)\|\ e^{-\beta t} + C_2)$$

Proof. The proof of the equivalence of the first three statements is reported in [MP86a, MP86b, MP86c] and [Bar88a, Bar88b, Bar88c]. See also the work in [BT80]. The equivalence 3–4 is easy, while the equivalence of statement 5 to the other ones is a tedious exercise (suggested but not required to the reader).

The theorem, as stated, is non-constructive. To check stability of an assigned discrete-time polytopic system (along with the determination of a proper polyhedral Lyapunov function, whenever stable) it is possible to proceed iteratively as previously mentioned. Indeed it is possible to use the procedure described in Section 5.4, starting from an *arbitrary* polyhedral set $\mathcal{X}^{(0)}$. Precisely, given the initial set $\mathcal{X}^{(0)} = \{x : \ \|F^{(0)}x\|_\infty \leq 1\}$ it is possible to recursively compute the sets

$$\begin{aligned}
\mathcal{X}^{(k+1)} &= \{x :\in \mathcal{X}^{(k)} : \ A_i\, x \in \mathcal{X}^{(k)}, \ \ i = 1,2,\ldots,s\} \\
&= \{x : \|F^{(k)}x\|_\infty \leq 1, \ \ \|F^{(k)}A_i\, x\|_\infty \leq 1, \ \ i = 1,2,\ldots,s\} \\
&\doteq \{x : \ \ \|F^{(k+1)}x\|_\infty \leq 1\}
\end{aligned}$$

Theorem 6.25 assures that if the system is stable then no matter how the polyhedral C-set $\mathcal{X}$ is chosen the largest invariant set included in it is also a polyhedral C-set and can be determined by a recursive procedure in a finite number of steps.

This theorem can be used "the other way around." Precisely, one can try to compute $\Sigma(A_1, A_2, \ldots, A_s)$ by computing the largest invariant set for the system

$$x(t+1) = \frac{A(w(t))}{\lambda} x(t)$$

and to reduce/increase λ if the procedure stops successfully/unsuccessfully. In detail, given a tentative λ one runs the procedure and

- decreases λ if for some $\bar{k}$

$$\mathcal{S}^{(\bar{k})} = int\mathcal{S}^{(\bar{k}-1)} (= \mathcal{S})$$

- increases λ if for some $\bar{k}$

$$\mathcal{S}^{(\bar{k})} \subset int\{\mathcal{X}^{(0)}\}$$

According to Theorem 6.25, under the assumption that the matrices do not admit a common proper invariant subspace (unless for the critical value $\lambda = \Sigma$), both conditions are detected in a finite number of steps. We will come back on this later, when we will deal with the more general problem of computing the best transient estimate.

For the continuous-time case one can, once again, resort to the EAS

$$x(t+1) = [I + \tau A(w)]x(t)$$

supported by the next proposition.

Proposition 6.28. *The following two statements are equivalent.*

- *The continuous-time system is stable and admits the Lyapunov function* $\|Fx\|_\infty$.
- *There exists* $\tau > 0$ *such that the EAS is stable and admits the Lyapunov function* $\|Fx\|_\infty$.

Proof. See [BM96a].

Therefore, the stability of a continuous-time polytopic system can be established by applying the previously described bisection algorithm to the EAS. In this case, there are two parameters on which it is necessary to iterate: λ and τ. One possibility to avoid this double iteration is that of iterating over the parameter τ only by assuming $\lambda(\tau) = 1 - \rho\tau^2$, as already mentioned in Section 5.2.

A possibility of reducing the complexity of the computation of the Lyapunov function is based on the following Proposition, which basically states that the stability of a differential inclusion is unchanged if we multiply it by a positive function.

Proposition 6.29. *Consider the differential inclusion*

$$\dot{x}(t) = \rho(t)A(w(t))x(t), \quad 0 < \rho^- \leq \rho(t) \leq \rho^+ \tag{6.19}$$

Then its stability does not depend on the bounds $0 < \rho^- \leq \rho^+$*. In particular it is stable iff* $\dot{x}(t) = A(w(t))x(t)$ *is stable.*

Proof. We prove sufficiency, since necessity is obvious. If $\dot{x}(t) = A(w(t))x(t)$ is stable, then it admits a polyhedral Lyapunov function $\Psi(x)$ such that $D^+\Psi(x, A(w)x) \leq -\beta\Psi(x)$. If we consider this function for (6.19) we get, denoting by $h' = h\rho$,

$$D^+\Psi(x, \rho A(w)x) = \lim\sup_{h\to 0^+} \frac{\Psi(x + h\rho A(w)x) - \Psi(x)}{h}$$
$$= \lim\sup_{h'\to 0^+} \frac{\Psi(x + h'A(w)x) - \Psi(x)}{h'}\rho = \rho D^+\Psi(x, A(w)x) \leq -\beta\rho\Psi(x)$$

Note that multiplication by $\rho > 0$ is equivalent to a time scaling: it changes the speed of convergence, but cannot compromise stability.

As a simple corollary, in the case of polytopic systems we can replace the generating matrices by scaled matrices

$$A(w) = \sum_{i=1}^{s} \rho_i A_i\, w_i$$

with positive scalars $\rho_i > 0$ without affecting the stability/instability properties. As an immediate consequence, when we consider the EAS for the computation of a polyhedral function, we can adopt different τ_i for different matrices. Precisely stability of the continuous-time system can be proven by computing a polyhedral function for the "EAS".

$$x(t+1) = [I + \sum_{i=1}^{s} \tau_i A_i w_i]x(t)$$

This property can be applied as follows. Given a single stable A, the eigenvalues of the EAS are $1 + \tau\lambda_i$, where λ_i are the eigenvalues of A. If τ is small enough, then $I + \tau A$ is stable, but if τ is too small, then the discrete-time eigenvalues are squeezed to 1, so that the discrete-time contractivity is very low. In general, different matrices A might suggest different values of τ. We can take advantage of this fact in computing a Lyapunov function, reducing both the computation time and the function complexity.

Example 6.30. Consider the polytopic system generated by the two matrices

$$A_1 = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 \\ -\frac{7}{4} & -1 \end{bmatrix}.$$

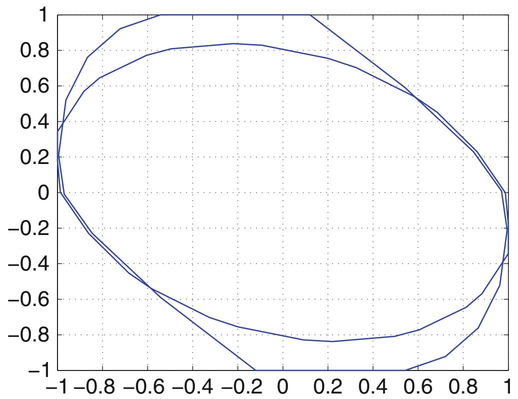

Fig. 6.5 The computed regions with a single $\tau = 0.2$ (external) and with two different values $\tau_1 = 0.5$ and $\tau_2 = 0.2$ (internal)

The polyhedral Lyapunov function computed with $\tau = 0.2$ considering the unit square (unit ball of $\|\cdot\|_\infty$) as initial set, produced a unit ball with 28 delimiting planes. The eigenvalues of A_1 are $-0.5, -0.5$ while those of A_2 are $-0.5 \pm 1.225j$. Those of the EAS are $0.9, 0.9$, and $0.9 \pm 0.24495j$. If we notice that it is reasonable to take a smaller τ for A_1 than for A_2, then we can take $\tau_1 = 0.5$, So the EAS has eigenvalue $0.75, 0.75$, and $\tau_2 = 0.2$. The resulting function is represented by a unit ball of 14 delimiting planes (Fig. 6.5). Clearly by no means the stability of the two discrete-time matrices assures convergence and continuous-time stability. In general, we will have to reduce all the τ_i when the procedure stops unsuccessfully.

6.3.4 The robust stability radius

Let us now consider the problem of computing the "robustness measure." Consider the system $\dot{x}(t) = A(w(t))x(t)$ (or $x(t+1) = A(w(t))x(t)$), with

$$A(w) = [A_0 + \Delta(w(t))], \quad \Delta(w) \in \rho\mathcal{W}$$

where $\mathcal{W}$ is compact and $\Delta(w)$ is continuous. The robustness measure we are thinking about is reported in the next definition.

Definition 6.31. Assuming A_0 a stable matrix

$$\rho_{ST} = \sup\{\rho : \text{ the system is robustly stable}\}$$

In the discrete-time case it is possible to apply the bisection procedure, precisely by starting with a tentative ρ, and to increase/reduce it if the computed set includes/does-not-include the origin in its interior. Thus, by applying the proposed procedure and by iterating by bisection on ρ, it is possible to derive an upper and lower bound on ρ_{ST}.

This algorithm may be directly applied to polytopic discrete-time systems in which

$$A(w) = A_0 + \rho[\sum_{i=1}^{s} \Delta_i w_i], \quad \sum_{i=1}^{s} w_i = 1, \;\; w_i \geq 0$$

with Δ_i assigned. As mentioned above, to consider continuous-time systems, one can use the EAS and iterate over τ.

6.3.5 *Best transient estimate*

Detecting stability only can be a non-sufficient task. One could be interested in evaluating the transient quality. To this aim, one can evaluate the evolution with respect to a given norm $\|\cdot\|_*$ by computing a transient estimate.

Definition 6.32. A transient estimate is a pair (C, λ) (respectively (C, β)) for which (6.17) (respectively (6.18)) holds for the solution.

Note that, in principle, λ may be any non-negative number and β any number. In other words, it is possible to estimate the transient of an unstable system (thus determining the "speed of divergence").

Let us consider the problem of computing a transient estimate with respect to the ∞-norm $\|x\|_\infty$ (the procedure immediately generalizes to any polyhedral norm of the form $\|Fx\|_\infty$). This can be done, in the discrete-time case, by performing the following steps.

Procedure. Computation of a transient estimate, given a contraction factor λ.

1. Fix a positive $\lambda < 1$.
2. Compute the largest invariant set $\mathcal{P}_\lambda$ inside the unit ball of the ∞-norm $\mathcal{X} = \mathcal{N}[\|\cdot\|_\infty, 1]$, for the modified system $x(t+1) = (A(w)/\lambda)x(t)$. Note that $\mathcal{P}_\lambda$ is the largest λ-contractive set for the considered system.
3. If $\mathcal{P}_\lambda$ has empty interior, then the transient estimate does not exist for the given λ (then one can increase λ and go back to Step 2).
4. Determine $C_\lambda > 0$ as the inverse of the largest factor μ such that $\mu\mathcal{X}$ is included inside $\mathcal{P}_\lambda$

$$C_\lambda^{-1} = \max_{\mu > 0} \;\; \text{s.t.} \;\; \mu\mathcal{X} \subseteq \mathcal{P}_\lambda$$

It can be shown that C_λ is the smallest constant such that (C_λ, λ) is a transient estimate. It is then clear that, by iterating over λ, it is possible to determine the "best transient estimate" (see [BM96a] for details). It turns out that if the system converges (or diverges) with speed $\lambda_0 < \lambda$, then the set $\mathcal{P}_\lambda$ is a polyhedral set and, as we have already seen, the procedure for its generation converges in a finite number of steps.

The same procedure can be used for continuous-time systems as follows. We can fix $\beta > 0$ and consider a small τ such that $\lambda(\tau) = 1 - \tau\beta < 1$. Then apply the procedure with such a λ to the EAS. It turns out that if the system converges with speed of convergence $\lambda_0 > \beta$ then, for sufficiently small τ, it is possible to compute a λ-contractive polyhedral set for the EAS and then a β-contractive polyhedral set with $\beta = (1 - \lambda)/\tau$.

Note that in principle, the transient estimate could be computed by means of any Lyapunov function, possibly quadratic, as shown later on. However the results are conservative.

Example 6.33. We report as an example the continuous-time system considered in [Zel94], with $m = 2$

$$A_1 = \begin{bmatrix} 0 & 1 \\ -2 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 \\ -2-\Delta & -1 \end{bmatrix}.$$

$\Delta = \Delta(t) \geq 0$. Quadratic stability is assured for this system if and only if $0 \leq \Delta < \Delta_Q \approx 3.82$ [Zel94] (this bound can be also obtained via standard continuous-time $\mathcal{H}_\infty$ analysis, as it will shown later). Zelentsowsky [Zel94] found the stability limit $\Delta_Z = 5.73$, say a 50% improvement. By using homogeneous polynomial Lyapunov functions and LMI techniques, in [CGTV03] it was shown that stability is assured for $\Delta_S = 6.7962$. Though not explicitly dealing with transient estimates, it is worth recalling that those techniques can be applied to the problem as well. Using the (EAS) with $\tau = 2.5 \times 10^{-4}$, $\lambda = 1 - 1 \times 10^{-9}$ and the polyhedral Lyapunov function construction, we were able to determine a polyhedral function for $\Delta_P = 6.97$. The computed transient estimate corresponding to Δ_Q, Δ_Z, and Δ_P are $(C_Q, \beta_Q) = (2.5439, 0.14)$, $(C_Z, \beta_Z) = (2.7068, 0.02)$ $(C_P, \beta_P) = (2.7805, 4.0 \times 10^{-6})$. The unit ball $\{x : \|Fx\|_\infty \leq 1\}$ of the Lyapunov function corresponding to Δ_P is reported in Fig. 6.6.

As it has been underlined several times, polyhedral Lyapunov functions are non-conservative. However, they generally require algorithms for the generation of their unit ball that are extremely heavy from the computational standpoint. The number of planes necessary for the description of such sets can drive out-of-range the most powerful machines, even for trivial instances. Clearly a transient estimate can be computed by means of quadratic function. If a positive definite matrix P such that

$$A_i^T P + PA_i + 2\beta I \prec 0$$

is found, then the corresponding family of ellipsoids $\mathcal{E}(P, \nu)$ is β-contractive. This in turn implies that one can take a β-contractive ellipsoid $\mathcal{E}(P, \nu)$ included in the box $\mathcal{X}$ and including $\mu\mathcal{X}$ for a proper $\mu \leq 1$. Then (C, β) with $C = 1/\mu$ is a transient estimate. Clearly such a transient estimate is, in general, conservative, not only because β is smaller, but also because C is quite greater than the best

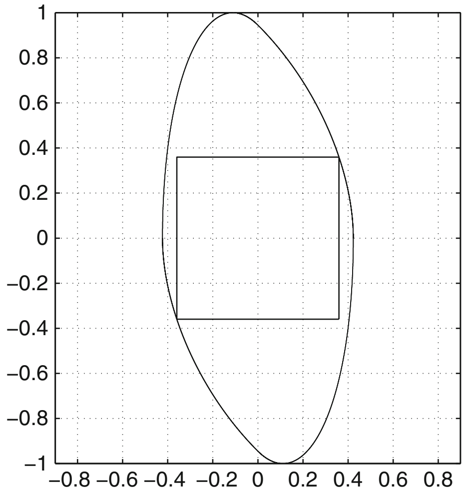

Fig. 6.6 The computed region for $\beta = 4.0 \times 10^{-6}$ and the inscribed region $1/C_P\mathcal{X}$

transient estimate (see Exercise 12). On the other hand, the computation is much easier. Indeed in the inclusion and containment constraints

$$\mu\mathcal{X} \subset \mathcal{E}(P,\nu) \subset \mathcal{X} \tag{6.20}$$

only the variables μ, ν, and P come into play. Besides, there are several problems, such as finding the smallest invariant set including a polytope or the largest invariant set in a polytope (in the sense of volume), that have the further important property of being convex in their variable and therefore very efficient algorithms are available. The reader is referred to [BEGFB04] for further details. As a final remark, it should be mentioned that the proposed analysis does not take into account variation of speed limits in the parameter. Taking into account these limits makes the problem harder (see, for instance, [Ran95, ACMS97]).

6.3.6 Comments about complexity and conservativity

Polyhedral functions are non-conservative, but computationally demanding[7]. Thus considering polyhedral functions instead of quadratic ones can be dramatic since the former might be extremely complex. A legitimate question is whether this is always the case. We show by means of a simple example that there are systems which are not quadratically stabilizable, but they admit a polyhedral function whose representation is not more complex than the representation of a quadratic function.

[7]Perhaps the reader will find this a tedious repetition in the book, still this conservativeness issue was not well known in the control literature for a long period [Ola92, Bla95].

Example 6.34 (A low complexity polyhedral function). Consider the four matrices

$$A_1 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix},$$

and the system

$$\dot{x}(t) = \sum_{i=1}^{4} w_i[-\varepsilon I + A_i]x(t), \quad \sum_{i=1}^{4} w_i = 1, \;\; w_i \geq 0 \tag{6.21}$$

with $\varepsilon > 0$ sufficiently small. The system admits $V(x) = \|x\|_1$ as a common Lyapunov function. Indeed any of the generating matrices $-\varepsilon I + A_i$ has $\|\cdot\|_1$ as an LF because it is strictly diagonally dominant, with negative diagonal entries.

We show that there are no common quadratic positive definite Lyapunov functions. To prove this, we first note that the set of matrices $\{A_k, \;\; k = 1,2,3,4\}$ is invariant with respect to the following transformations

$$T_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad T_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

namely changes of signs or reflections along the bisectors, since, for every choice of $T_i, i = 1,2,3$, we have $\{T_i^{-1}A_kT_i, \;\; k = 1,2,3,4\} = \{A_k, \;\; k = 1,2,3,4\}$. This amounts to saying that for every $i = 1,2,3$ and every $k = 1,2,3,4$ there exists $j = 1,2,3,4$ such that $T_i^{-1}A_kT_i = A_j$. This same property applies to the matrices $A_k - \varepsilon I, k = 1,2,3,4$. Consequently, if the positive definite matrix

$$P_1 = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

defines a common quadratic Lyapunov function for the matrices $A_k - \varepsilon I, k = 1,2,3,4$, so does

$$P_2 = \begin{bmatrix} a & -b \\ -b & c \end{bmatrix} = T_1^{-1}P_1T_1.$$

Since the set of common Lyapunov matrices for $\{A_k - \varepsilon I, \;\; k = 1,2,3,4\}$ is a convex cone, then

$$P_3 = \frac{P_1 + P_2}{2} = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}$$

defines a common quadratic Lyapunov function for the matrices $A_k - \varepsilon I, k = 1,2,3,4$. But since the set $\{A_k, \;\; k = 1,2,3,4\}$ is also invariant over bisector reflections, the positive definite matrix

$$P_4 = T_3^{-1} P_3 T_3 = \begin{bmatrix} c & 0 \\ 0 & a \end{bmatrix}$$

and hence the scalar matrix

$$\frac{P_4 + P_3}{2} = \begin{bmatrix} a+c & 0 \\ 0 & c+a \end{bmatrix} \frac{1}{2}$$

obtained as the average of P_3 and P_4, both define common quadratic Lyapunov functions. This implies that $P = I_2$ defines a common Lyapunov matrix, in other words that $\hat{A}_k^T + \hat{A}_k < 0$ for every $k = 1,2,3,4$. To verify that such condition is not true it is sufficient to compute

$$\det\left(-\hat{A}_1^T - \hat{A}_1\right) = \det\left(\begin{bmatrix} 2\,(1+\varepsilon) & -1 \\ -1 & 2\varepsilon \end{bmatrix}\right) = 4\varepsilon^2 + 4\varepsilon - 1$$

which is clearly negative for $0 < \varepsilon < \frac{-1+\sqrt{2}}{2}$.

6.3.7 Robust stability/contractivity analysis via system augmentation

A possibility to investigate stability contractivity in a less conservative way than using quadratic Lyapunov functions is based on system augmentation. Let us consider the discrete-time first. Assume we wish to establish the stability of the system

$$x(t+1) = [\sum_{i}^{s} A_{i=1} w_i(t)] x(t), \qquad \sum_{i=1} w_i = 1, \;\; w_i \geq 0$$

or equivalently, we wish to check if $\Sigma(A_1, A_2, \ldots, A_s) < 1$.

We can consider the T step system defined as follows:

$$x(k+T) = [A_{i_{T-1}} A_{i_{T-2}} \ldots A_{i_0}] x(t) = \Phi_t x(t) \tag{6.22}$$

where

$$\Phi_t \in \bar{\mathcal{A}}_T \doteq [A_{i_{T-1}} A_{i_{T-2}} \ldots A_{i_0}]$$

are the matrices formed by all possible T-products of the given A_i. The following proposition holds.

Proposition 6.35. *The difference inclusion is stable (or $\Sigma(A_1, A_2, \dots, A_k) < 1$) if and only if for T large enough (6.22) is quadratically stable. Moreover, for T large enough, any quadratic positive definite function $x^T Px$, $P \succ 0$ is a suitable quadratic Lyapunov function for system (6.22).*

Proof. The proof is similar to that of Proposition 6.24 and left to the reader as an exercise.

It should be noticed that the previous proposition is essentially a re-statement of old theory, for instance [GY93].

To apply the criterion we have two possibilities. One possibility is to fix an horizon T and check if all the matrices $\Phi \in \bar{\mathcal{A}}_T$ share a common Lyapunov function. The second one is to fix an horizon (usually much larger) to check if $\Phi \in \bar{\mathcal{A}}_T$ have all norms $\|\Phi\| < 1$. The shortcoming of the approach is that the number of matrices $\Phi \in \bar{\mathcal{A}}_T$ grows exponentially.

In the continuous-time case we can use a different system augmentation. This technique was used for the first time in [Zel94] and deeply investigated later [CGTV03, Che10, CGTV09, CCG$^+$12]. The idea, explained in brief, sounds as follows. Instead of $x(t)$ we introduce the variable $x^{(m)}(t)$ formed by all the monomials of order m. For instance, for $x(t) \in \mathbb{R}^2$

$$x^{(3)}(t) = [x_1^3(t), x_1^2(t)x_2(t), x_1(t)x_2^2(t), x_2^3(t)]^T$$

Consider the linear system

$$\dot{x}(t) = Ax(t)$$

Then the system in the new variable $x^{(m)}$ is described by the following "expanded" dynamic system

$$\dot{x}^{(m)}(t) = A^{(m)}x^{(m)}(t)$$

where the matrices $A^{(m)}$ can be computed as shown in [CGTV09]. Now it is obvious that the stability of the original system and of the expanded one are equivalent.

If we consider a quadratic candidate Lyapunov function for the new system

$$\Psi(x^{(m)}) = (x^{(m)})^T Px^{(m)}$$

this function turns out to be a polynomial Lyapunov function for the original system [CGTV09]. We know that the class of positive polynomials are universal, hence non-conservative for the robust stability problem [MP86a, MP86b, MP86a, BM99c]. It has recently been proved that the stability of the original system is equivalent to the quadratic stability of the extended system for m large enough [Che11b, CCG$^+$12, Che13].

6.4 Performance analysis of dynamical systems

In this section, it is shown how several problems related to the performance analysis of dynamic systems may be solved via a set-theoretic approach. We start by considering the fact that, as we have seen, for Linear Time-Invariant (LTI) systems, basic properties such as the effect of bounded inputs on the system can be in practice solved without set-computing. For instance, the evaluation of the l_1 norm of a system (i.e., the worst-case output peak for all possible peak-bounded inputs) requires the computation of the sum of a series. We have given a set-theoretic interpretation which has its own interest but it does not provide practical or theoretical advantages. Here it is shown how the set-theoretic formulation can be used to solve some analysis problems for uncertain systems for which the formulas known for LTI systems are no longer useful.

6.4.1 *Peak-to-peak norm evaluation*

Let us consider the problem of evaluating the largest output value achievable by the constrained inputs with 0 initial conditions for the discrete-time polytopic system

$$\begin{aligned} x(t+1) &= A(w(t))x(t) + Ed(t) \\ y(t) &= Hx(t) \end{aligned}$$

where, again, $A(w) = \sum_{i=1}^{s} A_i w_i$, with $w \in \mathcal{W}$, namely, $\sum_{i=1}^{s} w_i = 1$, $w_i \geq 0$ and d belongs to the compact set $\mathcal{D}$.

The paradigm consists in the following question: assume $x(0) = 0$ and let $d(t) \in \mathcal{D}$. Is the constraint

$$\|y(t)\|_* \leq \mu,$$

(with $\|\cdot\|_*$ a given norm) satisfied for all $t \geq 0$?

In the case in which also $\mathcal{D}$ is the unit ball of the same norm $\|\cdot\|_*$, we are evaluating the system induced norm. Formally the question is

- Q0:

$$\|(A(w), E, H)\|_{*,*} = \sup_{\substack{w(t) \in \mathcal{W} \\ x(0) = 0 \\ \|d(t)\|_* \leq 1,\ t \geq 0}} \ \sup_{t>0} \|y(t)\|_* \leq \mu\,?$$

The actual system norm can be estimated by iterating over μ. One way to proceed is that of computing the convex hulls of the 0-reachable sets. From the results previously presented it is indeed apparent that, denoting by $\mathcal{R}_t$ the 0 reachable set in t steps, the sequence of the convex hulls $conv\{\mathcal{R}_t\}$ can be computed as shown in Proposition 6.5. In view of the above consideration, an "yes" answer is equivalent to checking that

$$conv\{\mathcal{R}_t\} \in \mathcal{Y}(\mu) \doteq \mathcal{N}[\|Hx\|_*, \mu],$$

(roughly the set of all x such that $\|Hx\|_* \leq \mu$), for all t. This way of proceeding has the drawback that if the previous condition is satisfied till a certain $\bar{t}$, there is no guarantee that the same condition will be satisfied in the future. As it often happens, inverting the reasoning can be helpful. This is equivalent to reverting time, in this case. The problem can be solved in two steps as follows.

- Compute the largest robustly invariant set $\mathcal{P}_\mu$ for system $x(t+1) = A(w(t))x(t) + Ed(t)$ inside $\mathcal{Y}(\mu)$.
- If $0 \in \mathcal{Y}(\mu)$ then the answer to Q0 is "yes", otherwise it is "no".

In principle we should assume that the system has passed the stability test. Under some assumptions, such as the existence of an observable pair$(A(w), H)$ the stability test is actually included in the procedure according to following theorem.

Theorem 6.36. *Assume that there exists $w' \in \mathcal{W}$ such that $(A(w'), H)$ is an observable pair and that there exists $w'' \in \mathcal{W}$ such that $(A(w''), E)$ is reachable. The following statements are equivalent.*

- *All the reachable sets (equivalently, their convex hulls) are inside $\mathcal{Y}(\mu)$, say $\mathcal{R}_t \subset \mathcal{Y}(\mu)$, for all $t > 0$.*
- *The largest robustly invariant set $\mathcal{P}_\mu$ included in $\mathcal{Y}(\mu)$ is a C-set.*
- *The system is stable and question Q0 has answer "yes" (in the case of the induced norm $\|(A(w), E, H)\|_{*,*} \leq \mu$).*

Proof. The set $\mathcal{P}_\mu$ is the region of initial states starting from which the condition $x(t) \in \mathcal{Y}(\mu)$ is guaranteed for all $t \geq 0$. Therefore, the first two statements are obviously equivalent to the third statement, with the exception of the "stability claim." To include stability, we need to consider the observability and reachability assumption. Indeed, if we assume that $(A(w'), H)$ is observable, then the closed and convex set $\mathcal{P}_\mu$ is necessarily bounded [GT91]. Furthermore, if $(A(w), E)$ is reachable, then the reachable set $\mathcal{R}_T$ includes the origin as an interior point for all $T > 0$ (for T large enough in the discrete-time case) and then $\mathcal{P}_\mu$ is a $C - set$. Then we are in the position of proving stability.

Take any initial condition x_0 on the boundary of the C-set $\mathcal{P}_\mu$. The corresponding solution is given by $x(t) = x_f(t) + x_d(t)$ where $x_f(t)$ is the free response (i.e., such that $x_f(t+1) = A(w(t))x_f(t)$ and $x_f(0) = x_0$) and $x_d(t)$ is the response driven by d (precisely $x_d(0) = 0$ and $x_d(t+1) = A(w(t))x_d(t) + Ed(t)$. Since $x_d(T) \in \mathcal{R}_T$, then

$$x(T) \in \{x_f(T)\} + \mathcal{R}_T \subseteq \mathcal{P}_\mu$$

Denote by $\mathcal{S}_T = conv\{\mathcal{R}_T\}$ the convex hull of $\mathcal{R}_T$. Being $\mathcal{P}_\mu$ convex the last inclusion can be replaced by

$$x(T) \in \{x_f(T)\} + \mathcal{S}_T \subseteq \mathcal{P}_\mu$$

therefore $\{x_f(t)\}$ is in the erosion, $[\tilde{\mathcal{P}}_\mu]_{\mathcal{S}_T}$, of (see Definition 3.8) $\mathcal{P}_\mu$ with respect to $\mathcal{S}_T$

$$x_f(T) \in [\mathcal{P}_\mu]_{\mathcal{S}_T}$$

Since $\mathcal{S}_T$ is a C-set there exists $\lambda < 1$ such that $x_f(T) \in \lambda\mathcal{P}_\mu$.

We have proved that, for all $x_0 \in \partial\mathcal{P}_\mu$, we have that in T steps $x_f(T) \in \lambda\mathcal{P}_\mu$. Consider the T-step forward system

$$x_f(t+T) = [A(w(t+T-1))A(w(t+T-2))\dots A(w(t))]\, x_f(t)$$

which is linear, hence homogeneous. By applying Theorem 4.18 and Lemma 4.31 we can see $\mathcal{P}_\mu$ (which is invariant for $x_f(t+1) = A(w(t))x_f(t)$) is λ-contractive for such a system which implies stability.

We sketch now the algorithm proposed in [FG95] and [BMS97] that can be used for the $\|\cdot\|_\infty$ norm. Precisely we assume that $\|d(t)\|_\infty \leq 1$ and we seek for the largest possible $\|y(t)\|_\infty$.

1. Fix an initial guess $\mu > 0$ and a tolerance $\epsilon > 0$.
2. Set $F^{(0)} = H$, $g^{(0)} = \mu\bar{1}$, $k = 0$, $\mu^+ = +\infty$ and $\mu^- = 0$.
3. If $\mu^+ - \mu^- \leq \epsilon$ STOP. Else
4. Given the set $\mathcal{S}_k = \{x : |F_i^{(k)}x| \leq g_i^{(k)},\ \ i = 1, 2, \dots, r^{(k)}\}$, where $F_i^{(k)}$ is the ith row of matrix $F^{(k)}$ and $g_i^{(k)}$ is the ith component of vector $g^{(k)}$, compute the pre-image set $\mathcal{P}_{k+1}$ as
$$\mathcal{P}_{k+1} = \{x : |F_i^{(k)}A_j x| \leq \mu^{(k)} - \|F_i^{(k)}E\|_1,\ j = 1, 2, \dots, s,\ \ i = 1, 2, \dots, r^{(k)}\}$$
5. Compute the intersection
$$\mathcal{S}_{k+1} \doteq \mathcal{P}_{k+1} \bigcap \mathcal{S}_k$$
to form the matrix $F_i^{(k+1)}$ and the vector $g^{(k+1)}$.
6. If $0 \not\in \mathcal{S}_{k+1}$, then set $\mu^- = \mu$, increase μ and GOTO step 3.
7. If $\mathcal{S}_k = \mathcal{S}_{k+1}$, then set $\mu^+ = \mu$, reduce μ and GOTO step 3,

The previous results can be applied to continuous-time systems by means of the EAS. It can be shown that the ∞-to-∞ induced norm of the EAS system is always an upper bound for the corresponding induced norm of the continuous-time system

$$\|(A, E, H)\|_{\infty,\infty} \leq \|((I + \tau A), \tau E, H)\|_{\infty,\infty}$$

This fact can be inferred from the property that $\|((I + \tau A), \tau E, H)\|_{\infty,\infty} \leq \mu$ implies the existence of an invariant set for the EAS included in $\mathcal{Y}(\mu)$. In view of Lemma 4.26 and Proposition 5.10, such an invariant set is positively invariant for the continuous-time system.

The computation of the norm of system

$$(A(w_A), E(w_E), H(w_H)) = \left(\sum A_i w_{A,i}, \sum E_i w_{E,i}, \sum H_i w_{H,i}\right)$$

with polytopic structure can be handled as follows. The input $E(w_E)d$ is replaced by $v \in \mathcal{V}$ the convex hull of all possible points of the form $E_k d_h$, with E_k and d_h on their vertices. It is quite easy to see that the convex hulls of the reachability sets of $x(t+1) = A(w_A)x(t) + v(t)$ are the same as those of the original system. As far as the output uncertainty is concerned, the condition to be faced is

$$\|y\|_\infty = \|H(w_H)x\|_\infty \leq \mu \Leftrightarrow \|y^{(k)}\|_\infty = \|H_j x\|_\infty \leq \mu, \ \ \forall j$$

Therefore the problem requires repeating the iteration for all matrices H_j and retaining the minimum value. Note that, in this extension, it has been assumed that the uncertainties affecting $(A(w_A)$,$E(w_E)$,$H(w_H))$ are independent.

We remind the reader that the induced norm for the time-varying uncertain system we are considering here, say $\|(A(w), E, H)\|_{\infty,\infty}$, is quite different from the time-invariant norm, namely $\|(A(\bar{w}), E, H)\|_{\infty,\infty}$, the norm computed for the time-invariant system achieved by fixing $w = \bar{w}$. Clearly the time-invariant norm is not greater than the time-varying worst case norm:

$$\|(A(\bar{w}), E, H)\|_{\infty,\infty} \leq \|(A(w), E, H)\|_{\infty,\infty}$$

Example 6.37. Let us consider the following system

$$A(w) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -(3+w) & -2 & -3 \end{bmatrix}, \quad E = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},$$

$w \in [0, 1]$, and let us consider the corresponding EAS $(I+\tau A, \tau E, H)$. The algorithm provided the following limits for the norm

$$13.5 = \mu^- \leq \|((I + \tau A(w)), \tau E, H)\|_{\infty,\infty} \leq \mu^+ = 13.6$$

Note that the upper bound is actually an upper bound for the continuous-time system, while the lower bound is not. The algorithm required 3638 iterations to detect that the origin was not included in the largest invariant set for $\mu = \mu^- = 13.5$ and required 144 iterations to find an invariant set $\mu = \mu^+ = 13.6$. Such an invariant set is generated by 1292 constraints (by symmetry these correspond to

646 rows for the describing matrix F, not reported for paper-saving reasons). In the next table we report the number of iterations to detect the "yes/no" answer and the number of residual constraints (the actual constraints forming $\mathcal{P}_\mu$ in the "YES" answer case) as a function of the "guess" μ. We point out that the norms

μ	8	10	12	13.5	13.6	14	16	18
$\| \| \leq \mu$?	NO	NO	NO	NO	YES	YES	YES	YES
iterations	646	969	1554	3638	144	53	44	43
constraints	226	276	270	288	1292	586	530	524

of the extreme systems are quite smaller, $\|((I + \tau(A(0)), \tau E, H)\|_{\infty,\infty} \approx 1.98$ $\|((I + \tau(A(1)), \tau E, H)\|_{\infty,\infty} \approx 5.95$, and this means that high values of the norm are not due to a special "critical" value of w, but mainly to its variation inside $\mathcal{W}$.

Clearly the performance of the system might be estimated via ellipsoids. Let us consider the following problem. Consider the system

$$\dot{x}(t) = A(w)x(t) + Ed(t), \quad y(t) = Hx(t)$$

with

$$\|d(t)\| \leq \frac{1}{\mu}.$$

Now we assume that the norm is the Euclidean one (in the single input case it does not matter). Then we can consider the condition (4.23) (see [Sch73, USGW82]) to state that the ellipsoid $\mathcal{E}(P, 1)$ is positively invariant if, denoting by $Q = P^{-1}$, we have, for all i:

$$QA_i^T + A_iQ + \alpha Q + \frac{1}{\alpha}EE^T\frac{1}{\mu^2} \preceq 0, \quad \text{for some} \quad \alpha > 0 \tag{6.23}$$

The condition (4.23) has been stated for a single system $\dot{x} = Ax + Ed$, but the generalization above is obvious. The problem is that of including the ellipsoid $\mathcal{E}(Q^{-1}, 1)$ inside the strip $\mathcal{Y}(1)$. Note that for convenience we are iterating over μ by scaling the control disturbance rather than changing the size of $\mathcal{Y}$ which is obviously equivalent.

The condition $\mathcal{E}(Q^{-1}, 1) \subset \mathcal{Y}(1)$ can be easily expressed. Let us consider the single-input case for brevity. Then $\mathcal{Y}(1) = \{x : |Hx| \leq 1\}$, so that $\mathcal{E}(Q^{-1}, 1) \subset \mathcal{Y}(1)$ iff

$$HQH^T \leq 1. \tag{6.24}$$

Then, if we find a matrix $Q \succ 0$ such that conditions (6.23) and (6.24) are satisfied, then we are sure that the induced norm of the system is less than μ. If such an ellipsoid does not exist, however, we cannot conclude that the induced norm of the system is greater than μ.

Example 6.38. To show that the previous condition can be conservative, consider the example in [USGW82], namely the system $\dot{x}(t) = Ax(t) + Bu(t) + Ed(t)$ with matrices

$$A = \begin{bmatrix} -0.0075 & -0.0075 & 0 \\ 0.1086 & -0.149 & 0 \\ 0 & 0.1415 & -0.1887 \end{bmatrix}, \quad E = \begin{bmatrix} 0 \\ -0.0538 \\ 0.1187 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0037 \\ 0 \\ 0 \end{bmatrix}$$

to which the linear feedback control

$$u = Kx = -37.85x_1 - 4.639x_2 + 0.475x_3$$

is applied. Four outputs were considered: the state components and the control input. On all these variables, constraints are imposed as follows:

$$|x_1| \leq 0.1, \quad |x_2| \leq 0.01, \quad |x_3| \leq 0.1, \quad |u| \leq 0.25$$

which can be written as $\|Hx\|_\infty \leq 1$, where

$$H = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 10 \\ -151.40 & -18.55 & 1.90 \end{bmatrix}.$$

The disturbance input is bounded as $|d| \leq \alpha$. The ellipsoidal method provides the bound

$$\alpha_{ell} = 1.27$$

which implies the bound for the induced norm equal to $\|(A, E, H)\|_{\infty,\infty} \leq (1.27)^{-1} = 0.787$ By considering the EAS with $\tau = 1$, we achieved the bound

$$\alpha_{EAS} = 1.45$$

which implies $\|(A, E, H)\|_{\infty,\infty} \leq (1.45)^{-1} = 0.685$. Clearly, by reducing τ, tighter bounds can be achieved. We will reconsider this example later as a synthesis benchmark. We remind that the condition is also necessary [BC98] for positive invariance (if there exists a reachable pair $(A(w), E)$), therefore conservativeness is not due to condition (6.23), but to the adoption of ellipsoids.

6.4.2 Step response evaluation

We consider now the problem of computing the peak of the step response of the system

$$x(t+1) = A(w(t))x(t) + Ed(t), \qquad y(t) = Hx(t) + Gd(t)$$

where it is assumed that $d(t) \equiv 1$ and $x(0) = 0$. Basically the values one would like to evaluate for this system are the worst case peak and the asymptotic error and precisely, for given positive μ and ν, the questions now are:

Q1 : is the largest peak bound less than μ, $\sup_{t\geq 0} \|y(t)\| \leq \mu$?
Q2 is the largest asymptotic value less than ν, $\overline{\lim}\sup_{t\to\infty} y(t) \leq \nu$?

We can answer this questions as follows. We assume that $A(w)$ is stable. Let us consider the following sets

$$\mathcal{Y}(\xi) = \{x : \|Hx + G\| \leq \xi\}$$

(remind that $d \equiv 1$). Then we can claim the following.

Proposition 6.39.

- *The answer to question Q1 is yes if and only if the largest invariant set included in $\mathcal{Y}(\mu)$ includes the origin.*
- *The answer to question Q2 is yes if and only if the largest invariant set included in $\mathcal{Y}(\nu)$ is non-empty.*

The proof of this proposition can be found in [BMS97] where a more general case with both disturbances and constant inputs is considered.

Again, in terms of ellipsoids, a bound can be given as suggested in [BEGFB04] (see notes and references of Chapter 6). Indeed, the unit step is a particular case of norm-bounded input. However, as pointed out in [BEGFB04], the method is conservative (see Exercise 8).

Example 6.40. Consider the system

$$\dot{y}(t) = -[1 + w(t)/2]y(t) + u(t)$$

with the integral control

$$\dot{u}(t) = -\kappa(y(t) - r(t)).$$

Assume $r(t) \equiv \bar{r} = 1$, $\kappa = 5$ and $u(0) = y(0) = 0$. We use the EAS with $\tau = 0.1$, so achieving the system

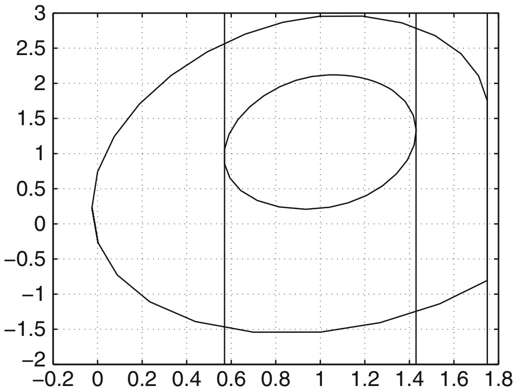

Fig. 6.7 The maximal invariant set (external) and the limit set (internal)

$$A = \begin{bmatrix} [0.85, 0.9] & 1 \\ -0.5 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix},$$

We first compute the output peak with respect to output y, asking

Q1: is condition $\sup_{t\geq 0} |y(t)| \leq \mu$ true for all $w(t)$?

It turns out that $\mu^+ = 1.75$, $\mu^- = 1.74$ are the upper and lower limits for the "yes" answer. In Figure 6.7 the largest invariant set included in $\mathcal{Y}_{max}(1.75)$ is depicted. Its margin is the rightmost vertical line, which includes the origin and certifies that $\mu^+ = 1.75$ is actually an upper limit. Then we compute the asymptotic behavior of the error $y - r$, the question now is

Q2: is the condition $\limsup_{t\to\infty} |y(t) - \bar{r}| \leq \nu$, true for all $w(t)$?

It turns out that the limit for the "yes" answer is between $\nu^+ = 0.430$ and $\nu^- = 0.429$. The smaller set in Fig. 6.7 is the largest invariant set included in the set $\mathcal{Y}_{lim}(0.430)$, the strip included between the two leftmost lines which certifies that the asymptotic behavior of y is between $\bar{r} + \nu^+ = 1.430$ and $\bar{r} - \nu^+ = 0.570$ (we remind that $\bar{r} = 1$)

The conclusions that can be drawn are the following. The step response of the system with the considered control does not exceed 1.75 as a peak, no matter how $0 \leq w(t) \leq 1$ changes. The asymptotic error is clearly not constant unless w has a limit value $0 \leq \bar{w} \leq 1$ (in which case the integrator assures $e(t) \to 0$). For persistent fluctuating values of w, in agreement with the considerations in Subsection 2.1.3. the error fluctuates and the worst case (for the EAS) is 0.430, which assures that the worst case for the continuous-time system does not exceed 0.430. It is intuitively clear that by taking τ smaller and smaller one converges to the actual value for the continuous-time system. Such intuition is supported by the results in [BS94], where such an assertion is proved. In Figure 6.8 a simulated step response is proposed. The value of $w(t)$ is alternatively taken equal to 0 and 1 starting with $w(0) = 0$ and by switching at $t = 3,\ 13,\ 15,\ 17,\ 19,\ 21$. It appears that the estimated values are sensibly larger than the actual ones. These are essentially due to two reasons. First, the realization of $w(t)$ considered in the simulation is not necessarily the "worst

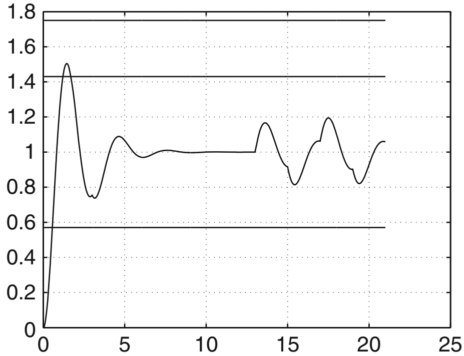

Fig. 6.8 The proposed simulation

case." Second, the provided bounds are non-conservative as $\tau \to 0$. Thus, we could reduce the value of τ (the considered one is 0.1), at the price of a noticeable increase in the number of planes delimiting the set.

6.4.3 Impulse and frequency response evaluation

It is possible to analyze impulse responses in the set-theoretic framework. Consider the SISO system

$$x(t+1) = A(w(t))x(t) + Ed(t), \quad y(t) = Hx(t)$$

with $w(t) \in \mathcal{W}$, $x(0) = 0$, $d(t) = \delta_0(t) = \{1, 0, 0, \ldots\}$, and assume that $(A(\tilde{w}), H)$ is observable for some $\tilde{w} \in \mathcal{W}$. The question is to find

$$\sup_{t \geq 0} |y(t)|_\infty$$

The problem can be reformulated as by fixing a $\mu > 0$ and checking if $\sup_{t \geq 0} |y(t)|_\infty \leq \mu$. By iterating over μ we can solve the problem up to a numerical approximation. We have the following.

Proposition 6.41. *Assume that the system is asymptotically stable. Then the impulse response y is such that $\sup_{t \geq 0} |y(t)|_\infty \leq \mu$ if and only if the (finitely determined) largest invariant set in the strip $\{x : |Hx| \leq \mu\}$ for the system includes the vector E.*

Note that, in principle the step response analysis proposed in the previous subsection could be carried out by augmenting the (stable) system

$$x(t+1) = A(w(t))x(t) + Ed(t), \tag{6.25}$$

by adding a fictitious equation

$$d(t+1) = d(t), \quad d(0) = 1$$

and testing the impulse response for the resulting system with output $y(t) = Hx(t) + 0d(t)$. The only problem is that, in this way the augmented system is not stable anymore and then there is no way to assure that the algorithm which computes the largest invariant set converges in finite time. To fix the problem, we can decide to accept the approximation achieved by replacing the equation $d(t+1) = d(t)$ by a slow decay

$$d(t+1) = \lambda d(t)$$

with $0 < \lambda < 1$ and $\lambda \approx 1$. With this kind of tricks we can manage other kind of problems. For instance, we can augment system (6.25) by adding the second order system

$$z(t+1) = R(\theta)z(t) + Pr(t), \quad d(t) = z_1(t)$$

where $R(\theta)$ is the θ-rotation matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

and $P = [\,1 \;\; 0\,]^T$. By means of the impulse response of this plant we can test the frequency response amplitude of the original plant (at frequency θ). Again the convergence of the algorithm is an open issue, since the augmented system is not asymptotically stable.

6.4.4 Norm evaluation via LMIs

We briefly now discuss some problems that can be solved by means of methods which are in some sense related to the set-theoretic approach we are dealing with. An important performance index for a system is the so-called $\mathcal{L}_2$-to-$\mathcal{L}_2$ induced gain which can be defined as follows:

$$\|(A,E,H,G)\|_{2,2} = \sup_{w \neq 0} \frac{\|y\|_2}{\|d\|_2}$$

where the $\mathcal{L}_2$ norm is defined as

$$\|u\|_2 = \sqrt{\int_0^\infty u(\sigma)^T u(\sigma)\, d\sigma}$$

It is well known that, if A is stable, such a norm has the fundamental frequency response characterization

$$\|(A,E,H,G)\|_{2,2} = \sup_{\omega\geq 0} \sqrt{\max \mathrm{eig}\,[W(j\omega)^T\ W(-j\omega)]}$$

where $W(s) = H(SI-A)^{-1}E+G$ is the transfer function matrix and $\mathrm{eig}(W^TW)$ is the set of the eigenvalues of W^TW (which are real non-negative) This norm is also referred to as $\mathcal{H}_\infty$-norm. It is known that the property $\|(A,E,H,G)\|_{2,2} < 1$ has an LMI characterization [SPS98]. Let us assume now that (A,E,H,G) has a polytopic structure

$$[A,E,H,G] = \sum_{i=1}^{s}\ w_i\ [A_i,E_i,H_i,G_i]$$

$\sum_{i=1}^{s}\ w_i = 1$, $w_i \geq 0$, Then the induced norm condition $\|(A,E,H,G)\|_{2,2} < 1$ is assured if there exists a positive definite P such that

$$\begin{bmatrix} A_i^TP+PA_i & PE_i & H_i^T \\ E_i^TP & -I & G_i^T \\ H_i & G_i & -I \end{bmatrix} \prec 0$$

Again this condition is a complete characterization (i.e., it provides a necessary and sufficient condition) for a single system (A,E,H,G), but for polytopic systems it is only sufficient when the condition is far to be necessary. Indeed as it will be seen later, there exist stable systems, therefore with finite induced gains, which are not quadratically stable (a condition which is implied by the previous LMI condition). Clearly, the discrete-time version of the problem has also an LMI characterization and the reader is referred to specialized literature.

Similar considerations can be done for the computation of the impulse response energy. Precisely, one might be interested in the evaluation of $\|(A,E,H)\|_2$, the $\mathcal{L}_2$ system norm, defined as the $\mathcal{L}_2$ norm of the system impulse response (for simplicity, the SISO case only is considered). Such norm is then equal to

$$\|(A,E,H)\|_2 = \|y_{imp}\|_2 = \sqrt{\int_0^\infty\ (He^{At}E)^T He^{At}E\,dt} = \sqrt{E^TPE}$$

where $P \succ 0$ is the unique (assuming (A,H) observable) solution to the equation

$$A^TP + PA + H^TH = 0$$

Let us consider a polytopic system. Assume that there exists $P \succ 0$ such that

$$A_i^TP + PA_i + H^TH \prec 0$$

and define the function $\Psi(x) = x^T Px$. Then, for any initial condition $x(0)$, the free response is such that

$$\dot{\Psi}(x) = x^T[A(w)^T P + PA(w)]x \leq -x^T H^T Hx = -y^T y$$

say $y^T y \leq -\dot{\Psi}(x)$. By integrating we have

$$\int_0^t y(t)^T y(t)\, dt \leq \Psi(x(0)) - \Psi(x(t))$$

Consider the system impulse response y, namely the free evolution with initial condition $x(0) = E$. Then, in view of the assumed asymptotic stability, $\Psi(x(t)) \to 0$ as $t \to \infty$, then

$$\|y_{imp}\|_2^2 \leq E^T PE$$

Again this is not a tight bound since the condition implies quadratic stability, which is stronger than stability.

6.4.5 *Norm evaluation via non-quadratic functions*

It is clear that if we consider bounds based on quadratic functions, then the system has to be quadratically stable. So the criterion is conservative for polytopic systems.

In general, given a stable system of the form

$$\dot{x}(t) = A(w(t))x(t) + Ed(t), \tag{6.26}$$

$$y(t) = Hx(t) \tag{6.27}$$

and a positive definite positively homogeneous function of the second order $\Psi(x)$, from a condition of the form

$$D^+\Psi(x) \leq -y^2(t) + \gamma d^2(t)$$

by integration we get, assuming $d(t) \to 0$ and $x(t) \to 0$

$$\int_0^\infty y^2(t)dt \leq \gamma \int_0^\infty d^2(t)dt + \Psi(x_0)$$

where x_0 is the initial state. The function $\Psi(x)$ is not necessarily quadratic and we can derive a polytopic bound on the output energy of the impulse response as follows. For brevity we consider the SISO case and $d \equiv 0$.

Consider the set

$$\mathcal{Y} = \{x : |Hx| \leq 1\}$$

and compute a β-contractive (possibly the largest) set $\mathcal{S}$ inside $\mathcal{Y}$. Consider the corresponding Minkowski functional $\Psi(x)$, for which, if $d = 0$, we get $D^+\Psi(x) \leq -\beta\Psi(x)$. Let $\psi(x) = \Psi^2(x)$. We get

$$D^+\psi(x) \leq -\frac{1}{\mu}\psi(x)$$

where $\mu \doteq (2\beta)^{-1}$. On the other hand, by construction, $\mathcal{S} \subset \mathcal{Y}$, so $\psi(x) \geq y^2$ (because the 1-level surface of ψ, $\mathcal{S} = \mathcal{N}[\psi, 1]$ is included in the 1-level surface of y^2, namely $\mathcal{Y}$. Hence

$$D^+\psi(x) \leq -\psi(x)/\mu \leq -y^2/\mu$$

By integrating we get for $d = 0$

$$\int_0^\infty y^2(t)dt \leq \mu\Psi(x_0)$$

which provides a bound for the output energy with initial condition x_0, so $\Psi(E)$ is a bound for the energy of the impulse response (say, when $d(t) = \delta(t)$).

The computation of the set $\mathcal{S}$ can be performed as previously described.

6.5 Periodic system analysis

We briefly consider the analysis problem of periodic systems. It is a known problem in the mathematical literature and we sketch some basic results. Consider the system

$$\dot{x}(t) = f(x(t), w(t))$$

and assume that f is Lipschitz and that $w(t)$ is a periodic signal of period T. A basic question considered in the literature is the existence of periodic trajectories. Clearly, the periodicity of $w(t)$ does not imply the existence of a periodic trajectory. However, there are some sufficient condition. Assume that there exist a C-set $\mathcal{X}$, t_0 and a period $T > 0$ such that, for all $x(t_0) \in \mathcal{X}$, $x(t_0 + T) \in \mathcal{X}$. Then, there exists a periodic trajectory. This fact can be shown by considering the Brouwer fixed-point theorem. Consider the map $F : \mathcal{X} \to \mathcal{X}$ which associates to $x \in \mathcal{X}$ the solution of the equation with initial condition $x(t_0)$ at time $t_0 + T$ $x(t_0 + T) = F(x(t_0))$. The map

F is continuous in view of the continuous dependence on the initial condition. Therefore there exists $\bar{x} \in \mathcal{X}$ such that $F(\bar{x}) = \bar{x}$, which implies that the solutions which starts from $\bar{x}$ at t_0 is T-periodic.

However, this basic result does not characterize the behavior of the periodic solution, for instance as far as its stability is concerned. Here we propose some results for systems of the form

$$\dot{x}(t) = A(t, w(t))x(t), \qquad \text{or, as usual } \; x(t+1) = A(t, w(t))x(t) \tag{6.28}$$

with $A(t, w)$ periodic in t. For this class of systems, stability implies exponential stability, as proved below.

Theorem 6.42. *Assume that in Eq. (6.28) $A(t, w)$ is continuous and periodic of period T, for any fixed $w \in \mathcal{W}$, with $\mathcal{W}$ compact. Assume that (6.28) is globally uniformly asymptotically stable (GUAS), according to Definition 2.16. Then it is exponentially stable.*

Proof. If the system is GUAS, for all $\mu > 0$ and $\epsilon > 0$, there exists an integer $\kappa = \kappa(\mu, \epsilon) > 0$ such that if $\|x(0)\| \leq \mu$ then $\|x(t)\| \leq \epsilon$, for all $t \geq \kappa T$ and it is bounded as $\|x(t)\| \leq \nu$ $0 \leq t \leq \kappa T$, for some $\nu > 0$. Take $\mu = 1$ and $\epsilon = \mu/2 = 1/2$. Then $\|x(\kappa T)\| \leq 1/2$. Consider the modified system

$$\dot{z}(t) = [\beta I + A(t, w(t))]z(t),$$

and recall that if $x(0) = z(0)$ then, $z(t) = e^{\beta t}x(t)$ is the solution of the modified system, since

$$\frac{d}{dt}(xe^{\beta t}) = \beta e^{\beta t}x + e^{\beta t}\dot{x} = [\beta I + A(t, w)](xe^{\beta t}).$$

Take $\beta > 0$ small enough to assure that $\|z(\kappa T)\| \leq 1$. Then, since $\|z(0)\| \leq 1$ implies $\|z(\kappa T)\| \leq 1$ and since $z(t)$ is bounded for $0 \leq t \leq \kappa T$, by the assumed periodicity we have that $\|z(r\kappa T)\| \leq 1$ for all integer r and that z is bounded, say $\|z(t)\| \leq \rho$ for some $\rho > 0$. Therefore

$$\|x(t)\| = \|e^{-\beta t}z(t)\| = e^{-\beta t}\|z(t)\| \leq e^{-\beta t}\rho$$

for all $\|x(0)\| \leq 1$, and thus also for $\|x(0)\| = 1$. In view of the linearity, in general we have

$$\|x(t)\| \leq e^{-\beta t}\rho\|x(0)\|$$

thus exponential stability.

The previous result, as a particular case, proves that an LPV system is stable if and only if it is exponentially stable, precisely the equivalence of the first two items of Theorem 6.27.

In the case of discrete-time periodic systems, stability can be checked by algorithms, which are based on the approach previously described. Indeed, one can start the backward construction (see Section 5.1.2) of the sets

$$\mathcal{X}_{-k-1} = \left\{ x : \frac{A(t,w)}{\lambda} x \in \mathcal{X}_{-k} \right\} \bigcap \mathcal{X}_0$$

starting from any arbitrary C-set $\mathcal{X}_0$. It can be shown that the sequence of sets, which is nested in T steps

$$\mathcal{X}_{-k-T} \subseteq \mathcal{X}_{-k}$$

either collapses to the origin or converges to a periodic sequence. The occurrence of the latter proves stability of the system under consideration. Precisely, assume that

$$\mathcal{X}_{-k-T} = \mathcal{X}_{-k}$$

(a condition which is typically met within a certain tolerance). The above states the fact that $x(t) \in \mathcal{X}_{-k}$ implies $x(t+T) \in \lambda^T \mathcal{X}_{-k}$, where λ is the contractivity factor.

The provided set-theoretic approach to performance evaluation can be easily extended to non-autonomous periodic systems. Consider, for instance, the system

$$x(t+1) = A(t, w(t))x(t) + Ed(t), \quad y(t) = Hx(t)$$

with $A(t,w)$ periodic in t with period T and d belonging to the C-set $\mathcal{D}$. Assume that one wishes to check if the worst case magnitude is $\|y(t)\|_\infty \leq \mu$. Then, setting $\mathcal{X}_0 = \mathcal{Y}(\mu) = \{x : \|Hx\| \leq \mu\}$, it is possible to start a similar backward construction:

$$\mathcal{X}_{-k-1} = \{x : A(t,w)x + Ed \in \mathcal{X}_{-k}\} \bigcap \mathcal{X}_0.$$

Again the sequence of sets is nested in T steps, say $\mathcal{X}_{-k-T} \subseteq \mathcal{X}_{-k}$. The sequence either stops due to an empty element, $\mathcal{X}_{-k} = \emptyset$, or converges to a periodic sequence [BU93].

6.6 Exercises

1. Show an example of a set $\mathcal{S}$ controllable to $\mathcal{P}$ such that $\mathcal{P}$ is not reachable form $\mathcal{S}$ and vice versa.
2. Show that if f is continuous, and if $\mathcal{P}$ and $\mathcal{U}$ are compact, then
$$f(\mathcal{P},\mathcal{U})$$
is compact (too easy?).
3. Assume that $\mathcal{P}$ is controlled-invariant. Show that, for $T_1 \leq T_2$, $\mathcal{C}_{T_1}(\mathcal{P}) \subseteq \mathcal{C}_{T_2}(\mathcal{P})$ where $\mathcal{C}_T(\mathcal{P})$ is the controllability set in time T. Show that the implication $\mathcal{R}_{T_1}(\mathcal{P}) \subseteq \mathcal{R}_{T_2}(\mathcal{P})$ is not true in general.
4. Explain why $\mathcal{C}_T(\mathcal{P})$ is not compact, even for a compact $\mathcal{P}$, in the case of discrete-time linear systems (Hint: take A singular ...).
5. Prove Proposition 6.7.
6. Show, by means of an example, that the one step reachable set from an ellipsoid $\mathcal{E}$ for the system $x(t+1) = Ax(t) + Ed(t)$, $d \in \mathcal{D}$ is convex, but in general it is not an ellipsoid, no matter if $\mathcal{D}$ is an ellipsoid or a polytope.
7. The set $\mathcal{R}_\infty$ with bounded input $d \in \mathcal{D}$ is robustly positively invariant. Is the set $\mathcal{R}_\infty(\bar{x})$ of all states reachable from $\bar{x} \neq 0$, for some arbitrary $T > 0$, positively invariant? Is the set$\mathcal{C}_\infty(\bar{x})$ of all states controllable to $\bar{x} \neq 0$ for arbitrary $T > 0$, positively invariant?
8. Given a stable system, the ratio between a) the maximum (worst case) output peak persistent disturbance inputs $|d(t)| \leq 1$, and b) unit step output, may be arbitrarily large. Can you show a sequence of LTI systems for which this ratio grows to infinity?
9. The l_1-norm of a MIMO system (A, E, H) is defined as follows. Denote by $Q^{(1)}, Q^{(2)}, \ldots, Q^{(k)}, \ldots$ the sequence of Markov parameters ($p \times m$ matrices). Then the l_1 norm is defined as
$$\|H(zI - A)^{-1}E\|_{l_1} = \sup_i \sum_{k=1}^{\infty} \sum_{j=1}^{m} |Q_{ij}^{(k)}|$$
This norm is known to be equal to
$$\|H(zI - A)^{-1}E\|_{\infty,\infty} \doteq \sup_{t \geq 0, x(0)=0, \|d(k)\|_\infty \leq 1} \|y(t)\|_\infty$$
Provide the "reachability set" characterization of this norm which is the MIMO version of Proposition 6.16.
10. Formulate a "convex" optimization problem to find P, μ, and ν which satisfy (6.20).

11. The statement ii) of Theorem 6.25 does not hold, in general, if the matrices share a proper invariant subspace. Show this by considering the single matrix $A = diag\{2, 1/2\}$ and $\mathcal{X}$ the unit square.
12. Consider the system $x(t+1) = Ax(t)$ with

$$A = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

and the norm $\|\cdot\|_\infty$ and $1/\sqrt{2} \leq \lambda < 1$. Find the best transient estimate (the largest λ-contractive set is delimited by 8 planes). What about the transient estimate evaluated with the Lyapunov norm $\|\cdot\|_2$?