Complementarity methods in the analysis of piecewise linear dynamical systems

Camlibel, M.K.

CentER

Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems

Kanat Çamlıbel

Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems

Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems

Proefschrift

ter verkrijging van de graad van doctor aan de
Katholieke Universiteit Brabant,
op gezag van de
Rector Magnificus, prof. dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen ten overstaan
van een door het college voor promoties
aangewezen commissie in de aula van de Universiteit op
maandag 28 mei 2001 om 16.15 uur
door

Mehmet Kanat Çamlıbel

geboren op 17 mei 1970
te Istanbul, Turkije.

Promotor: Prof. dr. J.M. Schumacher

to my parents
İlknur and Doğan

Acknowledgement

As a master's student, I became aware of the Dutch Systems and Control community while I was reading the book *Three Decades of Mathematical System Theory* edited by Henk Nijmeijer and Hans Schumacher at the occasion of the 50th birthday of Jan Willems. It was then that I decided to study in the Netherlands. After meeting Hans Schumacher at a workshop in Istanbul and obtaining a NATO fellowship, I started working with him at CWI in Amsterdam. Four years passed and I am now writing the last lines of my Ph. D. dissertation which would not exist without contributions, help, advice and support of a number of people that I would like to thank in what follows.

First and foremost, I wish to express my deep gratitude to my *promotor* Hans Schumacher, for his constant support from the first days, for the inspiring/stimulating discussions and ideas, and for his efforts to improve my writing skills. I have learnt a lot from his professionalism and his way of looking at mathematical problems. I feel both privileged and fortunate to have had such an advisor. Thank you Hans!

I am also very much indebted to Maurice Heemels. Our pleasant and fruitful cooperation enormously contributed to this thesis. I do remember that he was always ready when I needed help. Once, he even had to improvise a talk at the Benelux meeting since I could not show up. *Hartelijk bedankt* Maurice!

Further, I would like to thank Arjan van der Schaft who made an important impact on my research first at regular CWI meetings and later on every occasion we could discuss.

I profited a lot from numerous discussions that I had with my former teacher Külmiz Çevik during his one year stay in the Netherlands. Besides our academic collaboration, there were chess games and lots of fun. *Teşekkürler* Külmiz!

I am also grateful to the members of my promotion committee Bart De Moor, Jacob Engwerda, Henk Nijmeijer, Arjan van der Schaft, Stef Tijs, and Jan Willems for reading the draft version and making valuable comments.

It is also a great pleasure for me to thank those people with whom I shared so many things and from whom I received a lot of help: Stefi Cavallar, Tamás Fleiner, Ebru Angün, Bram van den Broek, Kaifeng Chen, Gül Gürkan, Norbert Hari, Attila Korpos, and Amol Sasane.

Last but not least, I should mention Ulviye Başer, Külmiz Çevik, İbrahim Eksin, Vasfi Eldem, Cem Göknar, Leyla Gören, Müjde Güzelkaya, Kadri Özçaldıran, Kemal Sarıoğlu, and Hasan Selbuz who encouraged and helped me to come to the Netherlands.

Finally, I acknowledge NWO (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) and TÜBİTAK (The Scientific and Technical Research Council of Turkey) for their financial support of my research.

Kanat Çamlıbel

Contents

Chapter 1

Introduction and Preliminaries

1.1 Introduction

Piecewise linear modeling has been a widely used technique in many engineering areas for a long time. By means of piecewise linear models, nonlinear phenomena can be approximated as accurately as desired. In general, the cost is sacrificing the *smoothness* and/or having large models. However, the properties offered by linearity, even in a piecewise manner, still make it one of the most natural options. Other ways in which piecewise linear systems may emerge include for instance gain scheduling type of controllers [39, 58, 59], variable structure systems [66] and bang-bang control [9, 43]. Of course, piecewise linear systems form a very general class. Inevitably, one sometimes has to sacrifice generality and consider specific subclasses in order to establish reasonably significant results. By following this idea, our treatment will focus on a subclass which allows us to employ complementarity methods of mathematical programming. With a slight abuse of terminology, we sometimes use the term *complementarity systems* (see [33, 38, 55, 56]) for this subclass of piecewise linear systems that can be dealt with by means of complementarity methods. It is possible to find lots of application areas in various fields such as electrical engineering, mechanical systems, operations research, economics etc. We refer to [32, 33, 57] for more detailed discussion of (potential) application areas. Since our treatment is based on complementarity theory, we can roughly say that our work lies in the junction of the system theory and the mathematical programming. To put/fit this thesis into a place within the existing literature, we discuss related areas and approaches in what follows.

Motivated, to a great extent, by the applications in mechanical systems (see for instance [42, 52] for classical treatments of unilateral constraints, and see also [8] for a survey on nonsmooth mechanics), and in circuit theory and control systems theory (see e.g. [9,43,53,66]), discontinuous dynamical systems have been studied extensively since the fifties and sixties. As a fundamental work in this area, we can mention Filippov [28] where differential equations with discontinuous right hand sides have been under consideration with an emphasis on the existence and uniqueness of solutions in the sense of Carathéodory. In the first part of this thesis, we will address similar questions for complementarity systems. Our development differs from Filippov's since complementarity systems do not fit into the framework of [28] in general. The work on differential inclusions (see e.g. [1]) is another branch of research on discontinuous dynamical systems. The combinations of differential equations and inequalities, and hence piecewise linear systems, can be easily cast as differential inclusions which usually have nonunique solutions by their nature. On the other hand, the uniqueness of solutions is of great importance from our model validation perspective.

Another way of looking at piecewise linear systems is to consider them as a subfamily of the huge family of hybrid systems. Indeed, piecewise linear systems can be regarded as hybrid systems (what cannot be?) just by translating the piecewise linearity to the language of hybrid systems. Embedding the piecewise linear nature into a hybrid automaton model would be one of such translations. Suppose that the piecewise linear system is given in the following explicit form

$$\dot{x} = A^i x + b^i \text{ if } x \in \mathcal{X}^i$$

for $i = 1, 2, \ldots, m$. For the corresponding hybrid automaton model, one can choose m *modes* in the natural way. The state space partition determined by the sets $\mathcal{X}^i$ directly indicates the *invariants* and *guards*.[1] It is hard to come up with tractable analysis methods for general hybrid systems. Naturally, some researches have focused on special subclasses of hybrid dynamical systems. In particular, the work that has been done on mixed logical dynamical systems ([3, 4]), first order linear hybrid systems with saturation ([24]) and piecewise affine systems ([60, 61]) is closely related to complementarity systems. Indeed, in a recent report [31] it has been shown for discrete systems that these subclasses and complementarity systems are equivalent under certain assumptions.

Among the fields that stimulated the work on piecewise linear systems, circuit theory has a special place because of the fact that the piecewise linear modeling idea comes up rather naturally in this context. One branch of research (see e.g. [5,19–22,29,40,41,45,67]) is mainly focused on canonical representations of piecewise linear characteristics/functions. In the cited references only analysis of static piecewise linear systems (resistive piecewise linear circuits in network theoretical terminology) has been considered. The main goal of those works was to represent resistive piecewise linear circuits in a canonical form and to propose methods to find the solutions (driving points) of the circuit. The employment of the complementarity setting separates [5, 40, 41, 45, 67] from the others. The first part of our thesis can be viewed as the continuation of this strand of work towards dynamical

[1] We follow the terminology used in [57].

systems.

Another direction of research in the community of circuit theory which our work can be connected with is the simulation of switching circuits (see e.g. [2, 5, 27, 44, 45, 48, 54, 68]). Roughly speaking, there are three main approaches, namely *event-tracking methods*, *time-stepping methods* and *smoothing methods.*[2] While the papers [2,48] are examples of work on event-tracking methods, [5, 44, 45, 54] give examples of studies on time-stepping methods. At this point, we should mention the work on time-stepping methods that has been done for unilaterally constrained mechanical systems with friction phenomena [47,49,51,62–64]. It seems that the question of convergence for these methods is usually not considered in the literature of circuit theory. With the inspiration of the cited work on mechanical systems, we have attempted to emphasize the need of justification of the time-stepping methods for switching circuits in the last two chapters of the second part of the thesis. The first chapter of the second part deals with smoothing methods. As related work in the context of mechanical systems, one can refer to [8, Chapter 2] and references therein.

After their introduction by Dupuis and Nagurney [25] (see also [50] for further development), projected dynamical systems have been used for studying the behavior of oligopolistic markets, urban transportation networks, traffic networks, international trade, agricultural and energy markets. Variational inequalities have been employed to characterize the stationary points of the projected dynamical systems. The well-known close relationship (see e.g. [30]) between complementarity problems and variational inequalities suggests that complementarity systems and projected dynamical systems are related to each other. Indeed, this relation has been addressed in [33, Chapter 6].

In the operations research community, several variations/extensions/generalizations of complementarity problems have been under consideration. Among all those variations/extensions/generalizations, the topological complementarity problem (TCP) (see [6,7]) is of considerable importance for us. In the second part of the thesis, we employ TCP as a general framework to investigate the convergence of approximations. Well-posedness of complementarity systems can be formulated in a pure TCP framework as well. Indeed, finding a solution of a complementarity system is nothing but finding a solution of a certain TCP. However, the available conditions which guarantee solvability of TCPs are very restrictive and are not satisfied in general by the systems we are looking at in this thesis. In this respect, our well-posedness results provide solvability conditions for a special class of TCPs.

In an infinite-dimensional systems setting, the book [26] addresses well-posedness issues as well as convergence of smoothing and time-stepping methods for *partial differential inequalities* that arise from mechanics and physics. Since we work in a finite dimensional

[2]For the explanations of these terms see Chapters 5 and 6.

framework here, the treatment in the cited reference is clearly more general. However, its development has been based on some *coerciveness* condition and hence it has implications for only a rather restrictive subclass of linear passive complementarity systems.

1.1.1 Outline of the thesis

The thesis is divided into two parts each containing three chapters. While Part I deals with the well-posedness of complementarity systems, Part II investigates convergence of approximations of complementarity systems.

In Chapter 2 we consider the well-posedness (in the sense of existence and uniqueness of solutions) of linear complementarity systems with external inputs where the underlying linear system is of index 1 as defined in Definition 2.3.1.

Linear passive complementarity systems (LPCS) are the objects of Chapter 3. The properties that are offered by passivity make it possible to derive stronger well-posedness results in the sense that the solutions are unique in larger spaces. The chapter contains comparisons of several solution concepts for LPCS. All the results that are obtained for LPCS will be extended to the class of systems that are passifiable by pole shifting (see Definition 3.4.2). After investigating Zeno behavior of this newly introduced class of systems, we will pass to the discussion on nonregular initial states. Finally, the chapter will be closed with results on well-posedness for distributional versions of two previously defined solution concepts.

Chapter 4 is devoted to a class of piecewise systems that can be formulated in a complementarity setting. Its main goal is to establish well-posedness results for this class of systems. It will be shown that linear complementarity systems and linear relay systems can be treated within the framework used in this chapter.

We consider some continuity properties of linear complementarity systems in Chapter 5. The idea is to replace the non-Lipschitzian complementarity characteristic by a Lipschitzian characteristic and investigate the convergence of the sequence of trajectories produced by approximating systems that have Lipschitzian characteristic as the Lipschitzian characteristic *tends to* the non-Lipschitzian complementarity characteristic. We will present sufficient conditions for the convergence of approximating trajectories to the trajectories of the actual system. The chapter will be closed by a discussion on more general approximations.

In Chapter 6 we will show that a time-stepping method, namely the backward Euler method, is consistent (in the sense that the approximations generated by the method converge to the actual solution of the original system in a suitable sense) for LPCS. As a side result, it will be proven that the solutions depend on the initial data continuously for that class of systems.

By employing the general framework presented in the previous chapter, we will investigate the consistency of the backward Euler method for relay systems in Chapter 7. This chapter will be followed by the conclusions in Chapter 8.

1.1.2 Origins of the chapters

Chapter 2 is mainly based on [15], which has been presented at the 14th International Symposium of Mathematical Theory of Networks and Systems in Perpignan (France), with slight changes. The only addition is Theorem 2.3.4 which provides a necessary condition for well-posedness of the systems under consideration.

The material of Chapter 3 is a cocktail of [10, 13, 34, 36]. Indeed, the results on the existence and uniqueness of solutions to LPCS were presented, for the first time, at the 38th IEEE Conference on Decision and Control in Phoenix (USA) (see [13] where one can also find the characterization of regular initial states). The notion of passifiability by pole shifting (PPS) has been introduced in [12] which has been presented at the 39th IEEE Conference on Decision and Control in Sydney (Australia). The necessary and sufficient conditions for PPS property are again due to [12]. The results on Zeno behavior can be found in [18]. Section 3.6 is based on [10] which is an improved version of the paper [34] that has been presented at the 4th International Conference on Automation of Mixed Processes: Hybrid Dynamic Systems in Dortmund (Germany).

Chapter 4 is basically based on [11] which is an outgrowth of the paper [46]. An early attempt, with weaker results, in this direction was presented at the European Control Conference'99 in Karlsruhe (Germany)(see [17]).

Chapter 5 is an extended version of the paper [12].

The report [16], after a minor revision, has been included as Chapter 6. It has already been submitted to IEEE Transactions on Circuits and Systems. For a less technical (without proofs) exposition, we refer to [14] which has been presented at the 4th International Conference on Automation of Mixed Processes: Hybrid Dynamic Systems in Dortmund (Germany).

The paper [35], which was presented at the 39th IEEE Conference on Decision and Control in Sydney (Australia), has been appended as Chapter 7 after including the proofs.

1.2 Preliminaries

1.2.1 Notation

Every text that contains a bit of mathematics, like this thesis, is written in two languages. One *natural* language, for instance English herein, is accompanied by the language of mathematical notations. In *Mathesis Biceps vetus et nova* (1670), Johann Caramuel[3] writes $102 = 857$ where the sign '=' is employed as the separatrix in decimal fractions. Although such severe complications are very unlikely to arise, we devote this subsection to *the second language*: mathematical notations.

Sets

The symbols $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$, $\mathbb{R}(s)$ and $\mathbb{C}$ denote the sets of real numbers, nonnegative real numbers, positive real numbers, real coefficient rational functions and complex numbers, respectively. For a given integer n, we write $\overline{n}$ for the set $\{1, 2, \ldots, n\}$. Let $\mathcal{X}$ be a set. The notations $\mathcal{X}^{n \times m}$ where n and m are integers denote the sets of n-tuples and $n \times m$ matrices of the elements of $\mathcal{X}$. The set of subsets of $\mathcal{X}$ will be denoted by $2^{\mathcal{X}}$. We write $|\mathcal{X}|$ for the number of elements of $\mathcal{X}$.

Matrices

Let $A \in \mathcal{X}^{n \times m}$ be a matrix of the elements of the set $\mathcal{X}$. We write A_{ij} for the (i, j)th element of A. The transpose of A is denoted by $A^\top$. For $J \subseteq \overline{n}$, and $K \subseteq \overline{m}$, A_{JK} denotes the submatrix $\{A_{ij}\}_{j \in J, k \in K}$. If $J = \overline{n}$ ($K = \overline{m}$), we also write $A_{\bullet K}$ ($A_{J \bullet}$). In order to avoid bulky notation, we use $A^\top_{JK}$ and A^{-1}_{JK} instead of $(A_{JK})^\top$ and $(A_{JK})^{-1}$, respectively. Given two matrices $A \in \mathcal{X}^{n_a \times m}$ and $B \in \mathcal{X}^{n_b \times m}$, the matrix obtained by stacking A over B is denoted by $\mathrm{col}(A, B)$. The diagonal matrix with the diagonal element $a_1, a_2, \ldots, a_n$ is denoted by $\mathrm{diag}(a_1, a_2, \ldots, a_n)$.

A rational matrix $A(s) \in \mathbb{R}^{n \times m}(s)$ is said to be *proper* if $\lim_{s \to \infty} A(s)$ is finite. If $\lim_{s \to \infty} A(s) = 0$ it is said to be *strictly proper*. A square rational matrix $A(s) \in \mathbb{R}^{n \times m}(s)$ is called *biproper* if it is proper, invertible as a rational matrix and its inverse is also proper.

Mappings

Given a mapping $f : \mathcal{U} \to \mathcal{V}$, we denote the image of f by $\mathrm{im}\, f := \{v \in \mathcal{V} \mid v = f(u) \text{ for some } u \in \mathcal{U}\}$ and the kernel of A by $\ker f := \{u \in \mathcal{U} \mid f(u) = 0\}$. $f|_{\mathcal{W}}$ will denote the restriction of f to $\mathcal{W} \subseteq \mathcal{U}$.

Function spaces

The notation $\mathcal{F}(\mathcal{U}, \mathcal{V})$ stands for the functions defined from $\mathcal{U}$ to $\mathcal{V}$. When $\mathcal{U} \subset \mathbb{R}$, we define the *reverse* operator $\mathrm{rev}_{[t', t'']} : \mathcal{F}([t', t''], \mathcal{V}) \to \mathcal{F}([t', t''], \mathcal{V})$ by

$$(\mathrm{rev}_{[t', t'']}\, v)(t) = v(t' + t'' - t).$$

[3]Quoted from F. Cajori, A History of Mathematical Notations, Dover Pub. Inc., New York, 1993.

The most often utilized function space will be the space of Bohl functions. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is called *Bohl function* if it has a rational Laplace transform. Every Bohl function is of the form $He^{F\cdot}G$ for some matrices $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{n\times 1}$ and $H \in \mathbb{R}^{1\times n}$. The set of all Bohl functions will be denoted by $\mathcal{B}$. As one can expect from their definition, Bohl functions are related to linear constant coefficient homogeneous differential equations and hence linear (time-invariant) dynamical systems. In our treatment of piecewise linear dynamical systems, piecewise Bohl functions play a similar role to the one is played in the study of linear systems by Bohl functions. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is said to be a *piecewise Bohl function* if for each $t \in \mathbb{R}_+$ there exist $\epsilon > 0$ and a Bohl function g such that $f|_{[t,t+\epsilon)} = g|_{[0,\epsilon)}$. The set of all such functions is denoted by $\mathcal{PB}$. Note that $\mathcal{PB}$ is not closed under time reversal. Since Bohl functions are real-analytic, the corresponding Bohl function to a piecewise Bohl function for a given (time instant) t is uniquely determined and the quantity $\max\{\epsilon > 0 \mid f|_{[t,t+\epsilon)} = g|_{[0,\epsilon)}\}$ is well-defined. For convenience, we define $\alpha : \mathcal{PB}^n \times \mathbb{R}_+ \to \mathcal{B}^n$ as

$$\alpha(f,t) = g$$

and $\beta : \mathcal{PB}^n \times \mathbb{R}_+ \to \mathbb{R}_{++} \cup \{\infty\}$ as

$$\beta(f,t) = \max\{\epsilon > 0 \mid f|_{[t,t+\epsilon)} = g|_{[0,\epsilon)}\}$$

where the Bohl function g is such that $f|_{[t,t+\rho)} = g|_{[0,\rho)}$ for some $\rho > 0$. The set of bounded piecewise Bohl functions, denoted by $\mathcal{PBB}$, consists of piecewise Bohl functions that are bounded on $[0,T]$ for each $T > 0$.

Another class of functions that appears later is the space of one-variable real-valued (locally) square integrable functions. In the standard way, we say a Lebesgue measurable function $f : \Omega \to \mathbb{R}^n$ is *square integrable* if

$$\int_\Omega f^\top(\tau) f(\tau)\, d\tau < \infty$$

holds where $\Omega \subset \mathbb{R}$. This class will be denoted by $\mathcal{L}_2(\Omega,\mathbb{R}^n)$. It is well-known that $\mathcal{L}_2(\Omega,\mathbb{R}^n)$ is a Hilbert space with the inner product

$$\langle f,g\rangle = \int_\Omega f^\top(\tau) g(\tau)\, d\tau$$

where $f,\ g \in \mathcal{L}_2(\Omega,\mathbb{R}^n)$. The norm that induced by this inner product can be given by

$$\|f\| = \langle f,f\rangle^{\frac{1}{2}} = (\int_\Omega f^\top(\tau) f(\tau)\, d\tau)^{\frac{1}{2}}.$$

A sequence $\{f_n\} \subset \mathcal{L}_2(\Omega, \mathbb{R}^n)$ is said to converge (strongly) to $f \in \mathcal{L}_2(\Omega, \mathbb{R}^n)$ if

$$\lim_{n\to\infty} \|f_n - f\| = 0,$$

and it is said to converge weakly to $f \in \mathcal{L}_2(\Omega, \mathbb{R}^n)$ if

$$\lim_{n\to\infty} \langle f_n, g\rangle = \langle f, g\rangle$$

for all $g \in \mathcal{L}_2(\Omega, \mathbb{R}^n)$.

Two particular subspaces of distributions will be of interest. We denote the set of distributions that are supported on a point $\{t\}$ by $\mathcal{D}'_t$. It is well-known from the distributional theory (see e.g. [65, Theorem 24.6]) that $v \in \mathcal{D}'_t$ is of the form

$$v = \sum_{i=0}^{N} v^i \delta^{(i)}$$

where N is a natural number, v^i is a real number for all $i \in \overline{N}$ and $\delta^{(i)}$ denotes the ith derivative of the Dirac distribution δ with the convention $\delta^{(0)} = \delta$. Later on, we restrict our attention to rather special classes of distributions, more specifically direct sums of $\mathcal{D}'_0$ and some function spaces. With an abuse of terminology, we say a distribution v is a *Bohl distribution* if it is of the form $v = v_{imp} + v_{reg}$ where the *impulsive part* $v_{imp} \in \mathcal{D}'_0$ and the *regular part* $v_{reg} \in \mathcal{B}$. The set of all such distributions is denoted by $\mathcal{B}_\delta$. Note that $\mathcal{B}_\delta = \mathcal{D}'_0 \oplus \mathcal{B}$. The leading coefficient of the impulsive part of a Bohl distribution v is defined by

$$\operatorname{lead}(v_{imp}) = \begin{cases} 0 & \text{if } v_{imp} = 0, \\ v^N & \text{if } v_{imp} = \sum_{i=0}^{N} v^i \delta^{(i)} \text{ with } v^N \neq 0. \end{cases}$$

We say that a Bohl distribution v is initially nonnegative if

$$(\operatorname{lead}(v_{imp}) > 0) \text{ or } (\operatorname{lead}(v_{imp}) = 0 \text{ and } v_{reg}(t) \geq 0 \text{ for all } t \in [0, \epsilon) \text{ for some } \epsilon > 0).$$

It is known ([37, Lemma 5.3]) that v is initially nonnegative if and only if $\hat{v}(\sigma) \geq 0$ for all sufficiently large σ where $\hat{v}(s)$ is its Laplace transform.

In parallel to the definition of piecewise Bohl distributions, we define the space $\mathcal{L}_2^\delta([0, T], \mathbb{R})$ consisting of distributions $v = v_{imp} + v_{reg}$ where the *impulsive part* $v_{imp} \in \mathcal{D}'_0$ and $v_{reg} \in \mathcal{L}_2([0, T], \mathbb{R})$.

Miscellaneous

The notations $\{x_n\}$ and $[y_i]_{i=1}^k$ denotes the sequence $x_1, x_2, \dots$ and the ordered set of the elements $y_1, y_2, \dots, y_k$, respectively.

All inequalities involving vectors must be understood componentwise. For two vectors x, $y \in \mathbb{R}^n$, $\max(x, y)$ and $\min(x, y)$ denote the componentwise maximum and minimum, respectively. The nonnegative and nonpositive parts of a vector x are denoted by x^+ and x^-, i.e., $x^+ = \max(x, 0)$ and $x^- = -\min(x, 0)$. Note that $x^+ \geq 0$, $x^- \geq 0$ and $x^+ \perp x^-$. We say that a proposition $\mathcal{P}(\alpha)$ holds for all sufficiently small (large) $\alpha \in \mathbb{R}_+$ if there exists $\alpha_0 \in \mathbb{R}_+ > 0$ such that $\mathcal{P}(\alpha)$ holds for all $0 < \alpha \leq \alpha_0$ ($\alpha_0 \leq \alpha$).

1.2.2 Linear complementarity problem

We briefly recall the linear complementarity problem (LCP) of mathematical programming. For an extensive survey on the problem, the reader is referred to [23].

Problem 1.2.1 (LCP(q, M)) Given $q \in \mathbb{R}^m$ and $M \in \mathbb{R}^{m \times m}$, find $z \in \mathbb{R}^m$ such that

$$z \geq 0 \tag{1.1a}$$

$$q + Mz \geq 0 \tag{1.1b}$$

$$z^\top (q + Mz) = 0. \tag{1.1c}$$

We say that z is *feasible* if it satisfies (1.1a)-(1.1b). Similarly, we say z *solves* LCP(q, M) if it satisfies (1.1). The set of all solutions of LCP(q, M) will be denoted by SOL(q, M). In general, SOL(q, M) may be the empty set. The notation $K(M)$ denotes the set $\{q \mid \text{SOL}(q, M) \neq \emptyset\}$. It is easy to see that $\mathbb{R}^m_+ \subseteq K(M)$ for all $M \in \mathbb{R}^{m \times m}$. The following fact on the closedness of $K(M)$ will be used several times in the sequel.

Fact 1.2.2 The set $K(M)$ (possibly empty) is closed for any matrix M.

The LCP leads to the study of a substantial number of matrix classes that relate to several aspects of the problem such as feasibility, solvability, unique solvability. The following ones will be of particular interest for our purposes.

Definition 1.2.3 A matrix $M \in \mathbb{R}^{m \times m}$ is called

- *nondegenerate* if all its principal matrices are nonzero.
- a $\mathcal{P}$-*matrix* if all its principal minors are positive.
- a $\mathcal{P}_0$-*matrix* if all its principal minors are nonnegative.

- *positive (nonnegative) definite* if $x^\top Mx > 0\ (\geq 0)$ for all $0 \neq x \in \mathbb{R}^m$.
- *copositive* if $x^\top Mx \geq 0$ for all $x \geq 0$.
- *copositive-plus* if it is copositive and the following implication holds:

$$x^\top Mx = 0 \text{ and } x \geq 0 \Rightarrow (M + M^\top)x = 0.$$

For a given nonempty set $\mathcal{S}$, we say that the set $\{v \mid v^\top w \geq 0 \text{ for all } w \in \mathcal{S}\}$ is the *dual cone* of $\mathcal{S}$. It is denoted by $\mathcal{S}^*$. The next lemma states some of the standard results on the matrix classes defined above.

Lemma 1.2.4 *Let $M \in \mathbb{R}^{m\times m}$ be given. The following statements hold.*

1. [23, Theorem 3.3.7] *LCP(q, M) has a unique solution for all $q \in \mathbb{R}^m$ if and only if M is a $\mathcal{P}$-matrix.*
2. [23, Corollary 3.8.10] *If M is copositive-plus then $K(M) = (SOL(0, M))^*$.*

Note that the last implication holds in particular when M is nonnegative definite.

1.2.3 Solution concepts

It is already well-known that the selection of *universum*, the space where all possible solutions live, is of great importance for the existence and uniqueness issues. We aim to illustrate this fact by means of an example in this subsection. Consider the following example due to Filippov [28, p. 116]

$$\begin{aligned}\dot{x}_1 &= \operatorname{sgn} x_1 - 2\operatorname{sgn} x_2\\ \dot{x}_2 &= 2\operatorname{sgn} x_1 + \operatorname{sgn} x_2\end{aligned}$$

where $\operatorname{sgn} x$ is the set-valued function given by

$$\operatorname{sgn} x = \begin{cases}-1 & \text{if } x < 0\\ [-1, 1] & \text{if } x = 0\\ 1 & \text{if } x > 0\end{cases}.$$

Its time-reversed version can be given by

$$\begin{aligned}\dot{y}_1 &= -\operatorname{sgn} y_1 + 2\operatorname{sgn} y_2\\ \dot{y}_2 &= -2\operatorname{sgn} y_1 - \operatorname{sgn} y_2\end{aligned}$$

Solutions of the time-reversed version are spiraling towards the origin, which is an equi-

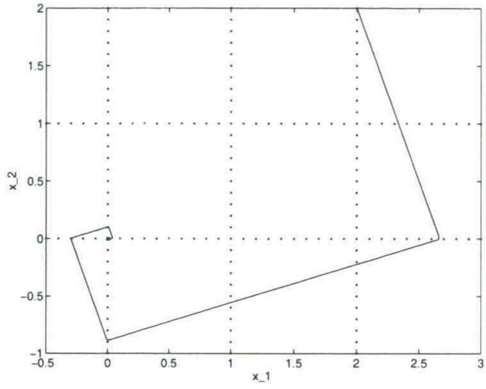


Figure 1.1: Trajectory with initial state $(2,2)^\top$.

librium. Since $\frac{d}{dt}(|y_1(t)| + |y_2(t)|) = -2$ when $y(t) \neq 0$ along trajectories x of the system, solutions reach the origin in finite time (see Figure 1.1 for a trajectory). Therefore, time-reversals of all these trajectories qualify as a solution (starting from the origin) to the original system in the sense of Definition 3.3.8 below for which the universum is $\mathcal{L}_2$-functions that are defined on a bounded interval. However, if one requires solutions to be right continuous (as in Definition 3.3.1 below) then there is a unique solution, namely the *zero* solution. As this example shows, a system might be well-posed for one solution concept but not for another one.

References

[1] J. P. Aubin and A. Cellina. *Differential Inclusions.* Springer, Berlin, 1984.

[2] D. Bedrosian and J. Vlach. Time-domain analysis of networks with internally controlled switches. *IEEE Trans. Circuits and Systems-I*, 39(3):199–212, 1992.

[3] A. Bemporad, G. Ferrari-Trecate, and M. Morari. Observability and controllability of piecewise affine and hybrid systems. Technical Report AUT99-16, ETH-Zürich, Switzerland, 1999.

[4] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427, 1999.

[5] W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis.* Kluwer, Deventer, the Netherlands, 1981.

[6] J. M. Borwein. Alternative theorems for general complementarity problems. In *Infinite Programming*, Lecture Notes in Econ. & Math. Systems 259, pages 194–203. Springer, Berlin, 1985.

[7] J. M. Borwein and M. A. H. Dempster. The linear order complementarity problem. *Mathematics of Operations Research*, 14(3):534–558, 1989.

[8] B. Brogliato. *Nonsmooth Impact Mechanics*. Springer-Verlag, London, 1996.

[9] D.W. Bushaw. *Differential Equations with Discontinuous Forcing Term*. PhD thesis, Dept. of Math., Princeton Univ., 1952.

[10] M. K. Çamlıbel, W. P. M. H. Heemels, and J.M. Schumacher. On linear passive complementarity systems. 2001, submitted for publication.

[11] M. K. Çamlıbel and J.M. Schumacher. Existence and uniqueness of solutions for a class of piecewise linear dynamical systems. 2001, submitted for publication.

[12] M.K. Çamlıbel, M.K.K. Cevik, W.P.M.H. Heemels, and J.M. Schumacher. From Lipschitzian to non-Lipschitzian characteristics: continuity of behaviors. In *Proc. of the 39th IEEE Conference on Decision and Control*, Sydney (Australia), 2000.

[13] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. The nature of solutions to linear passive complementarity systems. In *Proc. of the 38th IEEE Conference on Decision and Control*, pages 3043–3048, Phoenix (USA), 1999.

[14] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Simulation of linear networks with ideal diodes: consistency of a time-stepping method. In *Proc. of the 4th International Conference on Automation of Mixed Processes: Hybrid Dynamic Systems*, pages 265–270, Dortmund (Germany), 2000.

[15] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Well-posedness of a class of linear network with ideal diodes. In *Proc. of the 14th International Symposium of Mathematical Theory of Networks and Systems*, Perpignan (France), 2000.

[16] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Dynamical analysis of linear passive networks with diodes. Part II: Consistency of a time-stepping method. Technical Report 00 I/03, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted for publication.

[17] M.K. Çamlıbel and J.M. Schumacher. Well-posedness of a class of piecewise linear systems. In *Proc. of the European Control Conference*, Karlsruhe (Germany), 1999.

[18] M.K. Çamlıbel and J.M. Schumacher. Do the complementarity systems exhibit Zeno behavior? 2001, submitted for presentation at CDC'01.

[19] L. O. Chua and A. C. Deng. Canonical piecewise-linear analysis. *IEEE Transactions on Circuits and Systems-I*, 30:201–229, 1983.

[20] L. O. Chua and A. C. Deng. Canonical piecewise-linear analysis: Part II — tracing driving-point and transfer characteristics. *IEEE Transactions on Circuits and Systems-I*, 32:417–443, 1985.

[21] L. O. Chua and A. C. Deng. Canonical piecewise-linear modeling. *IEEE Transactions on Circuits and Systems-I*, 33:511–525, 1986.

[22] L. O. Chua and A. C. Deng. Canonical piecewise-linear representation. *IEEE Transactions on Circuits and Systems-I*, 35:101–111, 1988.

[23] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem.* Academic Press, Inc., Boston, 1992.

[24] B. De Schutter. Optimal control of a class of linear hybrid systems with saturation. Technical report, Accepted for publication in *SIAM Journal on Control and Optimization*, 1999.

[25] P. Dupuis and A. Nagurney. Dynamical systems and variational inequalities. *Annals of Operations Research*, 44:9–42, 1993.

[26] G. Duvaut and J. L. Lions. *Inequalities in Mechanics and Physics.* Springer, Berlin, 1976.

[27] J.T.J. van Eijndhoven. Solving the linear complementarity problem in circuit simulation. *SIAM Journal on Control and Optimization*, 24(5):1050–1062, 1986.

[28] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides.* Mathematics and Its Applications. Prentice-Hall, Dordrecht, The Netherlands, 1988.

[29] C. Güzeliş and İ. C. Göknar. A canonical representation for piecewise-affine maps and its applications to circuit analysis. *IEEE Transactions on Circuits and Systems-I*, 38:1342–1354, 1991.

[30] P. T. Harker and J.-S. Pang. Finite-dimensional variational inequalities and nonlinear complementarity problems: a survey of theory, algorithm and applications. *Math. Progr. Ser. B*, 40:161–220, 1990.

[31] W. P. M. H. Heemels and B. De Schutter. On the equivalence of classes of hybrid systems: Mixed logical dynamical and complementarity systems. Technical Report 00 I/04, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted for publication.

[32] W. P. M. H. Heemels, J. M. Schumacher, and S. Weiland. Applications of complementarity systems. In *European Control Conference*, Kalsruhe, Germany, 1999.

[33] W.P.M.H. Heemels. *Linear complementarity systems: a study in hybrid dynamics.* PhD thesis, Dept. of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands, 1999.

[34] W.P.M.H. Heemels, M.K. Çamlıbel, and J.M. Schumacher. On dynamics, complementarity and passivity: electrical networks with ideal diodes. In *Proc. of the 4th International Conference on Automation of Mixed Processes: Hybrid Dynamic Systems*, pages 197–202, Dortmund (Germany), 2000.

[35] W.P.M.H. Heemels, M.K. Çamlıbel, and J.M. Schumacher. A time-stepping method for relay systems. In *Proc. of the 39th IEEE Conference on Decision and Control*, Sydney (Australia), 2000.

[36] W.P.M.H. Heemels, M.K. Çamlıbel, and J.M. Schumacher. Dynamical analysis of linear passive networks with diodes. Part I: Well-posedness. Technical Report 00 I/02, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted for publication.

[37] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. The rational complementarity problem. *Linear Algebra and Its Applications*, 294:93–135, 1999.

[38] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics*, 60(4):1234–1269, 2000.

[39] K. J. Hunt and T. A. Johansen. Design and analysis of gain-scheduled control using local controller networks. *International Journal of Control*, 66:619–651, 1997.

[40] T. A. M. Kevenaar and D. M. W. Leenaerts. A comparison of piecewise-linear model descriptions. *IEEE Transactions on Circuits and Systems-I*, 39:996–1004, 1992.

[41] T. A. M. Kevenaar, D. M. W. Leenaerts, and W. M. G. van Bokhoven. Extensions to Chua's explicit piecewise-linear function descriptions. *IEEE Transactions on Circuits and Systems-I*, 41:308–314, 1994.

[42] C.W. Kilmister and J.E. Reeve. *Rational Mechanics.* Longmans, London, 1966.

[43] J. P. LaSalle. Time optimal control systems. *Proc. Natl. Acad. Sci. U.S.*, 45:573-577, 1959.

[44] D.M.W. Leenaerts. On linear dynamic complementarity systems. *IEEE Transactions on Circuits and Systems-I*, 46(8):1022–1026, 1999.

[45] D.M.W. Leenaerts and W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[46] Y.J. Lootsma, A.J. van der Schaft, and M.K. Çamlıbel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[47] P. Lötstedt. Numerical simulation of time-dependent contact and friction problems in rigid body mechanics. *SIAM Journal on Scientific and Statistical Computing*, 5:370–393, 1984.

[48] A. Massarini, U. Reggiani, and K. Kazimierczuk. Analysis of networks with ideal switches by state equations. *IEEE Trans. Circuits and Systems-I*, 44(8):692–697, 1997.

[49] J.J. Moreau. Numerical aspects of the sweeping process. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):329–349, 1999.

[50] A. Nagurney and D. Zhang. *Projected Dynamical Systems and Variational Inequalities with Applications.* Kluwer, Boston, 1996.

[51] L. Paoli and M. Schatzman. Schéma numérique pour un modèle de vibrations avec contraintes unilatérales et perte d'énergie aux impacts, en dimension finie. *C.R. Acad. Sci. Paris Sér. I Math.*, 317:211–215, 1993.

[52] J. Pérès. *Mécanique Générale.* Masson & Cie., Paris, 1953.

[53] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes.* Interscience, New York, 1962.

[54] I.W. Sandberg. Theorems on the computation of the transient response of nonlinear networks containing transistors and diodes. *Bell System Technical Journal*, 49:1739-1776, 1970.

[55] A.J. van der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9:266–301, 1996.

[56] A.J. van der Schaft and J.M. Schumacher. Complementarity modelling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[57] A.J. van der Schaft and J.M. Schumacher. *An Introduction to Hybrid Dynamical Systems.* Springer-Verlag, London, 2000.

[58] J. S. Shamma. *Analysis and Design of Gain-scheduled Control Systems.* PhD thesis, Dept. of Mech. Eng., Massuchusetts Institute of Technology, 1988.

[59] J. S. Shamma and M. Athans. Analysis of gain scheduled control for nonlinear plants. *IEEE Trans. on Automatic Control*, 35:898–907, 1990.

[60] E. Sontag. Interconnected automata and linear systems: a theoretical framework in discrete-time. *Hybrid Systems III*, volume 1066 of *Lecture Notes in Computer Science*, Springer, pages 436–448, 1999.

[61] E. Sontag. Nonlinear regulation: the piecewise linear approach. *IEEE Trans. on Automatic Control*, 26(2):346–357, 1999.

[62] D.E. Stewart. Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problem. *Archive for Rational Mechanics and Analysis*, 145(3):215–260, 1998.

[63] D.E. Stewart. Time-stepping methods and the mathematics of rigid body dynamics. Chapter 1 of *Impact and Friction*, A. Guran, J.A.C. Martins and A. Klarbring (eds.), Birkhäuser, 1999.

[64] D.E. Stewart and J.C. Trinkle. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and Coulomb friction. *Int. Journal for Numerical Methods in Engineering*, 39:2673–2691, 1996.

[65] F. Trèves. *Topological Vector Spaces, Distributions and Kernels.* Academic Press, New York, 1967.

[66] V. I. Utkin. *Sliding Modes in Control Optimization.* Springer, Berlin, 1992.

[67] L. Vandenberghe, B. L. De Moor, and J. Vandewalle. The generalized linear complementarity problem applied to the complete analysis of resistive piecewise-linear circuits. *IEEE Trans. Circuits Syst.*, CAS-36:1382–1391, 1989.

[68] J. Vlach, J.M. Wojciechowski, and A. Opal. Analysis of nonlinear networks with inconsistent initial conditions. *IEEE Transactions on Circuits and Systems-I*, 42(4):195–200, 1995.

Part I

Well-posedness

Chapter 2

Well-posedness of Linear Complementarity Systems with Inputs: Low Index Case

2.1 Introduction

The appropriateness of a proposed mathematical model for a given physical system can be tested in various ways. A very basic test is the following: if the physical system that is being modeled is deterministic in the sense that it shows identical behavior under identical circumstances, then the mathematical model should have the same property. Model validity would be put into serious doubt if it would turn out that the equations of the mathematical model allow multiple solutions for some initial data. With any model formulation for a deterministic physical system it is therefore important to establish well-posedness of the model, i. e., existence and uniqueness of solutions for feasible initial conditions.

This chapter considers the well-posedness of a class of linear complementarity systems, i.e., linear systems coupled to complementarity conditions. The most typical examples of these systems are linear electrical networks with ideal diodes. In the engineering literature, mathematical models that make use of the ideal diode characteristic are routinely used for such networks. Remarkably enough, it seems that the well-posedness of such models has not been rigorously established before. Although general results from the theory of ordinary differential equations may be used to establish well-posedness of network models containing elements with Lipschitzian characteristics (see for instance [12]) or in special cases even for non-Lipschitzian characteristics (see for instance [2, 7]), such results do not cover the ideal diode characteristic since it cannot be reformulated as a current or voltage-controlled resistor. Neither does it seem possible to derive general well-posedness results for network models with ideal diodes from the theory of differential equations with discontinuous right hand sides [3], which in network terminology is concerned with models involving ideal relay elements. The theory that we develop below will be based on the theory of *complementarity systems* that has been worked out in a series of recent papers [4–6, 9, 10]; see also [11].

It is easy to come up with examples of mathematical models involving ideal diode characteristics (which are equivalent to complementarity conditions) that are not well-posed; see for instance [9]. Therefore, some restrictions need to be imposed. We will study this class of models in the more general setting of complementarity conditions coupled to linear dynamical systems with a special zero structure at infinity. Some might say that it is "intuitively clear" that such network models are well-posed; nevertheless, ideal diodes are only approximations to real diodes and so the fact that actual networks with diodes behave deterministically does not make it evident that the corresponding mathematical models with idealized elements have unique solutions. Rather, as argued above, one should consider well-posedness as a test of model validity.

The chapter is organized as follows. In Section 2 we first of all develop a precise notion of solution for linear complementarity systems. Then in Section 3 we briefly discuss the linear complementarity problem (LCP) of mathematical programming that plays an important role in our development. The main results follow in Section 4. The chapter will be closed by conclusions in Section 5 and proofs in Section 6.

2.2 Linear Complementarity Systems

As interconnection of a continuous, time-invariant, linear system and complementarity conditions, a *linear complementarity system* can be given by

$$\dot{x}(t) = Ax(t) + Bu(t) + Ew(t) \tag{2.1a}$$

$$y(t) = Cx(t) + Du(t) \tag{2.1b}$$

$$0 \leq u(t) \perp y(t) \geq 0. \tag{2.1c}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$, $w(t) \in \mathbb{R}^p$, and A, B, C, D and E are matrices with appropriate sizes. We denote the above system by LCS(A, B, C, D, E). For the previous study on this class of systems, the reader is referred to [4–6, 9, 10]. From a *hybrid system* point of view, one can distinguish 2^m *modes* depending on complementarity conditions (2.1c). Every index set $K \subseteq \overline{m}$ determines one of these modes by imposing the constraints $y_K = 0$ and $u_{\overline{m} \setminus K} = 0$. Associated to each mode K, there are a linear dynamics given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Ew(t) \\ y(t) &= Cx(t) + Du(t) \\ y_K(t) = 0, &\qquad u_{\overline{m} \setminus K}(t) = 0 \end{aligned}$$

and a set called *invariants* given by

$$y_{\overline{m}\setminus K}(t) \geq 0, \qquad u_K(t) \geq 0. \tag{2.2}$$

Starting at a given mode, the system trajectories must obey the dynamics corresponding to this mode as long as they belong to the invariant set, i.e., satisfy the inequalities (2.2). Time instants at which the state variables tend to leave the invariant set are called *event times*. Whenever an event occurs, another mode will become active depending on the state variables x and inputs w at the event time. Before giving a precise definition of the solution concept, we illustrate the above features of the systems under consideration in the following example.

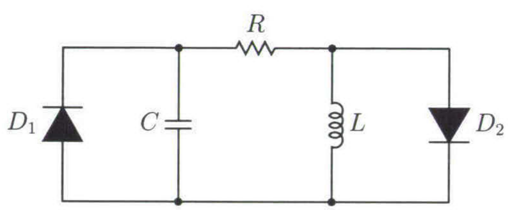


Figure 2.1: RLC circuit with ideal diodes

Example 2.2.1 Consider the linear RLC circuit (with $R = 1$ Ohm, $L = 1$ Henry and $C = 1$ Farad) coupled to two ideal diodes as shown in Figure 2.1. By choosing the voltage across the capacitor and the current through the inductor as the state variables and by taking into account the ideal diode characteristic depicted in Figure 2.2, the governing equations of the network can be given by

$$C\tfrac{d}{dt}v_C = i_L - i_{D_1} + i_{D_2} \tag{2.3a}$$

$$L\tfrac{d}{dt}i_L = -v_C - Ri_L - Ri_{D_2} \tag{2.3b}$$

$$v_{D_1} = v_C \tag{2.3c}$$

$$v_{D_2} = -v_C - Ri_L - Ri_{D_2} \tag{2.3d}$$

$$0 \leq i_{D_1} \perp -v_{D_1} \geq 0 \tag{2.3e}$$

$$0 \leq i_{D_2} \perp -v_{D_2} \geq 0. \tag{2.3f}$$

Depending on whether the diodes are blocking or conducting, the system has 4 *modes*.

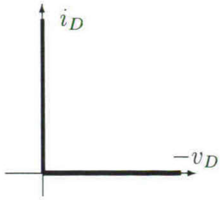


Figure 2.2: Ideal diode characteristic

- *Mode BB:* In this mode, both diodes are blocking, i. e., $i_{D_1} = i_{D_2} = 0$. Hence, the conditions (2.3e)-(2.3f) yield

$$\begin{aligned} 0 = i_{D_1} \quad & -v_{D_1} \geq 0 \\ 0 = i_{D_2} \quad & -v_{D_2} \geq 0. \end{aligned}$$

The *activities*, or *circuit topology* (see Figure 2.3 (a)) as it is called in network theory terminology, can be given by

$$\begin{aligned} C\tfrac{d}{dt}v_C &= i_L \\ L\tfrac{d}{dt}i_L &= -v_C - Ri_L. \end{aligned}$$

The corresponding *invariants* (the conditions that ensure the diodes to keep blocking state) are

$$\begin{aligned} -v_{D_1} &= -v_C \geq 0 \\ -v_{D_2} &= v_C + Ri_L \geq 0. \end{aligned}$$

- *Mode BC:* The first diode is blocking while the second one is conducting, i. e., $i_{D_1} = v_{D_2} = 0$ in this mode. Hence, the conditions (2.3e)-(2.3f) yield

$$\begin{aligned} 0 = i_{D_1} \quad & -v_{D_1} \geq 0 \\ 0 \leq i_{D_2} \quad & v_{D_2} = 0. \end{aligned}$$

The activities can be given by

$$\begin{aligned} C\tfrac{d}{dt}v_C &= i_L + i_{D_2} \\ L\tfrac{d}{dt}i_L &= -v_C - Ri_L - Ri_{D_2} \\ v_{D_2} &= v_C + Ri_L + Ri_{D_2} = 0. \end{aligned}$$

The corresponding circuit topology is shown in Figure 2.3 (b). The invariants, as being the conditions that ensure the first diode to keep blocking state and the second conducting state, are

$$\begin{aligned} -v_{D_1} &= -v_C \geq 0 \\ i_{D_2} &= -\tfrac{1}{R}v_C - i_L \geq 0. \end{aligned}$$

- *Mode CB:* The first diode is conducting and the second one is blocking, i. e., $v_{D_1} = i_{D_2} = 0$ in this mode. Hence, the conditions (2.3e)-(2.3f) yield

$$\begin{aligned} 0 \leq i_{D_1} &\quad v_{D_1} = 0 \\ 0 = i_{D_2} &\quad -v_{D_2} \geq 0. \end{aligned}$$

The activities can be given by

$$\begin{aligned} C\tfrac{d}{dt}v_C &= i_L - i_{D_1} \\ L\tfrac{d}{dt}i_L &= -v_C - Ri_L \\ v_{D_1} &= v_C = 0. \end{aligned}$$

The corresponding circuit topology is shown in Figure 2.3 (c). The invariants are

$$\begin{aligned} i_{D_1} &= i_L \geq 0 \\ -v_{D_2} &= Ri_L \geq 0. \end{aligned}$$

- *Mode CC:* In this mode both diodes are conducting, i. e., $v_{D_1} = v_{D_2} = 0$. Hence, the conditions (2.3e)-(2.3f) yield

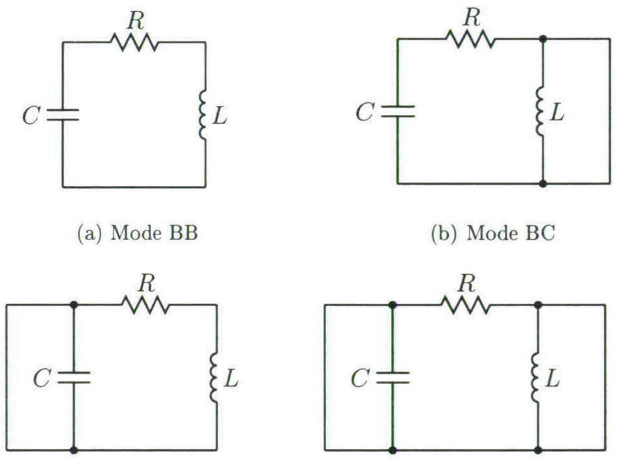
$$\begin{aligned} 0 \leq i_{D_1} &\quad v_{D_1} = 0 \\ 0 \leq i_{D_2} &\quad v_{D_2} = 0. \end{aligned}$$

The corresponding circuit topology is depicted in Figure 2.3 (d) and the activities of the mode can be given by

$$\begin{aligned} C\tfrac{d}{dt}v_C &= i_L - i_{D_1} + i_{D_2} \\ L\tfrac{d}{dt}i_L &= -v_C - Ri_L - Ri_{D_2} \\ v_{D_1} &= v_C = 0 \\ v_{D_2} &= -v_C - Ri_L - Ri_{D_2} = 0. \end{aligned}$$

The invariants can be obtained as

$$
\begin{aligned}
i_{D_1} &= 0 \\
i_{D_2} &= -i_L \geq 0.
\end{aligned}
$$



(a) Mode BB (b) Mode BC

(c) Mode CB (d) Mode CC

Figure 2.3: Circuit topologies for the modes

We investigate the behaviour of the network for the initial condition $(v_C(0), i_L(0)) = (-e, 1)$. Note that the first diode must be blocking initially since $v_{D_1}(0) = v_C(0) \neq 0$ and the second one must be conducting initially since $v_C(0) + i_L(0) < 0$. Then, the mode BC is active at the beginning. It can be checked that the dynamics of this mode yields

$$
\begin{aligned}
v_C(t) &= -e^{1-t} \\
i_L(t) &= 1.
\end{aligned}
$$

The first inequality of those describing the invariants of this mode holds for all t while the second one holds only if $t \in [0, 1]$. Therefore, $t_1 = 1$ is the first event time. At the event time, the state of the system is given by $v_C(1) = -1$ and $i_L(1) = 1$. In the next mode, the first diode still must be blocking initially since $v_{D_1}(1) = v_C(1) \neq 0$ but the second one cannot be conducting anymore. Hence, the next mode in which the system will evolve should be the mode BB. It can be computed that the dynamics of this mode yields

$$v_C(t) = -e^{-\frac{1}{2}(t-1)}[\cos(\tfrac{\sqrt{3}}{2}(t-1)) - \tfrac{\sqrt{3}}{3}\sin(\tfrac{\sqrt{3}}{2}(t-1))]$$
$$i_L(t) = e^{-\frac{1}{2}(t-1)}[\cos(\tfrac{\sqrt{3}}{2}(t-1)) + \tfrac{\sqrt{3}}{3}\sin(\tfrac{\sqrt{3}}{2}(t-1))]$$

for $t \geq 1$. It can be verified that $v_C(t) + i_L(t) \geq 0$ and $v_C(t) \leq 0$ for all $1 \leq t \leq 1 + \frac{2\sqrt{3}}{9}\pi$, and also that $v_C(1 + \frac{2\sqrt{3}}{9}\pi) = 0$ and $\frac{dv_C}{dt}(1 + \frac{2\sqrt{3}}{9}\pi) > 0$. Consequently, the first diode cannot be blocking anymore and this means that the second event takes place at event time $t_2 = 1 + \frac{2\sqrt{3}}{9}\pi$. At the event time, the state of the system can be given by $v_C(t_2) = 0$ and $i_L(t_2) = e^{\frac{-\sqrt{3}}{9}\pi}$. The next mode should be the mode CB and its dynamics result in

$$v_C(t) = 0$$
$$i_L(t) = e^{t-(1-\frac{2\sqrt{3}}{9}\pi)}$$

for $t \geq t_1$. It can be easily verified that invariants of this mode are satisfied for all $t \geq t_1$, i.e., there will be no mode change anymore. The trajectories are depicted in Figure 2.4.

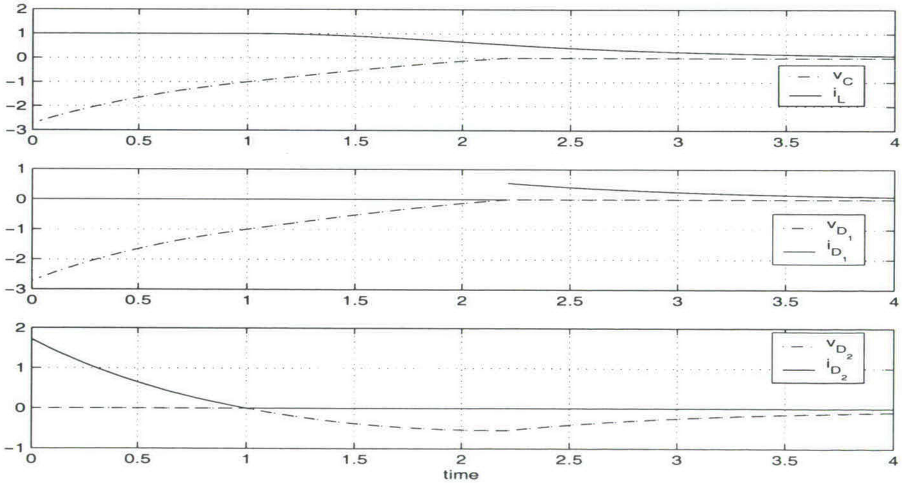


Figure 2.4: Trajectories for the initial state $(-e, 1)$.

Later on, we will employ 'hybrid system' thinking to construct solutions to LCSs. However, the concept of solution will be clarified first. In what follows, we propose a solution notion by keeping in mind the hybrid features of the system. Indeed, the 'universum' we consider, namely the space of piecewise Bohl functions, is asymmetric in time in the sense that the time reverse of a piecewise Bohl function is not piecewise Bohl in general.

Definition 2.2.2 A triple $(u, x, y) \in \mathcal{PB}^{m+n+m}$ is a *solution* on $[0, T]$ of LCS(A, B, C, D, E) for the input $w \in \mathcal{PBB}^p$ and the initial state x_0 if the following conditions hold

$$x(t) = x_0 + \int_0^t [Ax(s) + Bu(s) + Ew(s)]\ ds$$
$$y(t) = Cx(t) + Du(t)$$
$$0 \leq u(t) \perp y(t) \geq 0.$$

for all $t \in [0, T]$.

Notice that x-trajectory is continuous by definition. In the sequel, we will derive sufficient conditions under which linear complementarity systems have unique solutions. Before doing this, we will review some facts from complementarity theory in order to be self-contained.

2.3 Main Results

In this section, we present sufficient conditions for well-posedness, in the sense of existence and uniqueness of solutions, of linear complementarity systems. One of our main assumptions will be on the *index* of the underlying system. The following definitions will make clear what is meant by the index of a linear system.

Definition 2.3.1 A rational matrix $H(s) \in \mathbb{R}^{l \times l}(s)$ is said to be *of index k* if it is invertible as a rational matrix and $s^{-k}H^{-1}(s)$ is proper.

Definition 2.3.2 A rational matrix $H(s) \in \mathbb{R}^{l \times l}(s)$ is said to be *totally of index k* if all its principal submatrices are of index k.

Now, we can state the main result concerning the well-posedness of the linear complementarity systems.

Theorem 2.3.3 *Consider a matrix quintuple (A, B, C, D, E). Suppose that $G(s) = D + C(sI - A)^{-1}B$ is totally of index 1 and $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then, the following two statements are equivalent.*

1. *For each $w \in \mathcal{PBB}^p$, there exists a unique solution on $[0, \infty)$ of LCS(A, B, C, D, E) for the input w and the initial state x_0.*

2. *$Cx_0 \in K(D)$.*

Note that $G(s)$ is totally of index 1 if and only if $D + CBs^{-1}$ is. Since $\det(\cdot)$ is a continuous function, if $D + CBs^{-1}$ is of index 1 then we have $\operatorname{sign}(\det(G_{JJ}(\sigma))) = \operatorname{sign}(\det(D_{JJ} + C_{J\bullet}B_{\bullet J}\sigma^{-1}))$ for all sufficiently large σ. This means that the $\mathcal{P}$-matrix assumption on the transfer matrix holds if $D + CB\sigma^{-1}$ is a $\mathcal{P}$-matrix for all sufficiently large σ. In general, there are no explicit characterizations of the set $K(D)$. However, if D is copositive-plus the set $K(D)$ can be characterized explicitly as stated in Lemma 1.2.4 item 2. Note that all nonnegative definite matrices are copositive-plus.

The above theorem provides sufficient conditions for well-posedness. In the next theorem we will present a necessary condition.

Theorem 2.3.4 *Consider a matrix quintuple (A, B, C, D, E). Suppose that D is nondegenerate and C is of full row rank. If D is not a $\mathcal{P}_0$-matrix then for some $x_0 \in \mathbb{R}^n$ and $T > 0$, there exist at least two different solutions on $[0, T]$ of LCS(A, B, C, D, E) for the zero input and the initial state x_0.*

2.4 Conclusions

We showed that a class of linear complementarity systems including electrical networks with diodes as typical examples passes the validity test of well-posedness. Using complementarity theory, we were able to prove the existence and uniqueness of solution trajectories under a condition on the zero structure of the underlying state space description. As an additional result we gave an explicit characterization of the regular states, i.e., the initial states for which the linear complementarity systems admit solutions in the sense of Definition 2.2.2.

2.5 Proofs

This section is devoted to the proofs of the presented results.

2.5.1 Lipschitzian properties of LCP

We begin with stating some results on Lipschitzian properties of LCP. For our purposes, it is important to relate the index of the system and Lipschitzian properties of a series of LCPs involving the transfer function of the system. First, we present a rather general result on locally Lipschitz continuous functions.

Lemma 2.5.1 *Let the sets $\mathcal{Q}^i \subset \mathbb{R}^n$ for $i = 1, 2, \ldots, p$ be such that $\mathcal{Q}^i$ is closed and convex, and*

$$\bigcup_{i=1}^{p} \mathcal{Q}^i = \mathbb{R}^n.$$

Assume that $f : \mathbb{R}^n \to \mathbb{R}^m$ is a continuous function which is Lipschitz on each set $\mathcal{Q}^i$ with a Lipschitz constant α^i. Then, f is Lipschitz continuous with the Lipschitz constant $\max_i \alpha^i$.

Proof: Let x_a and $x_b \in \mathbb{R}^n$. Consider the line segment $[x_a, x_b]$ in $\mathbb{R}^n$. Since the number of $\mathcal{Q}^i$s is finite and they are all closed convex sets, one can find finite number of points in $\mathbb{R}^n$, say $x_a =: x^1, x^2, \ldots, x^l := x_b$, such that for each $i \in \overline{l-1}$ the line segment $[x^i, x^{i+1}] \subset \mathcal{Q}^{j_i}$ for some j_i. Note that due to the continuity of f we have

$$\begin{aligned}
\|f(x^1) - f(x^l)\| &\leq \|f(x^1) - f(x^2)\| + \|f(x^2) - f(x^3)\| + \cdots + \|f(x^{l-1}) - f(x^l)\| \\
&\leq \alpha^{j_1}\|x^1 - x^2\| + \alpha^{j_2}\|x^2 - x^3\| + \cdots + \alpha^{j_{l-1}}\|x^{l-1} - x^l\| \\
&\leq (\max_i \alpha)(\|x^1 - x^2\| + \|x^2 - x^3\| + \cdots + \|x^{l-1} - x^l\|).
\end{aligned}$$

Since all x^is are on the line segment $[x^1, x^l]$, it is obvious that $\sum_{i=1}^{l-1} \|x^i - x^{i+1}\| = \|x^1 - x^l\|$. Consequently, we get $\|f(x^1) - f(x^l)\| \leq (\max_i \alpha^i)\|x^1 - x^l\|$. ■

In the sequel, for a given nondegenerate matrix $M \in \mathbb{R}^{n\times n}$, $d(M)$ is defined as follows:

$$d(M) = (\max_{J \subseteq \overline{n}} \|M_{JJ}^{-1}\|).$$

It is known (see [1, Theorem 7.3.10]) that if LCP(q, M) is uniquely solvable for each q then the mapping $q \mapsto z$ where z is the unique solution of LCP(q, M) is Lipschitz continuous. However, to compute the Lipschitz constant given in [1] is not so easy. By making use of above lemma, we will show that the quantity $d(M)$ can be taken as Lipschitz constant for the LCP(q, M) whenever M is a $\mathcal{P}$-matrix.

Lemma 2.5.2 *Assume that $M \in \mathbb{R}^{n\times n}$ is a $\mathcal{P}$-matrix. Let z^i be the unique solution of LCP(q^i, M) for $i = 1, 2$. Then, we have*

$$\|z^1 - z^2\| \leq d(M)\|q^1 - q^2\|.$$

Proof: Since M is a $\mathcal{P}$-matrix, Lemma 1.2.4 item 1 implies that LCP(q, M) is uniquely solvable for all q. Consider the function $q \mapsto z$ where z is the unique solution of LCP(q, M). For a given index set $J \in \overline{n}$, define the set $\mathcal{Q}^J$ as

$$\mathcal{Q}^J = \{q \in \mathbb{R}^n \mid -M_{JJ}^{-1} q_J \geq 0 \text{ and } q_{\overline{n}\setminus J} - M_{(\overline{n}\setminus J)J} M_{JJ}^{-1} q_J \geq 0\}.$$

i. Clearly, $\mathcal{Q}^J$ is closed and convex for each J.

ii. Note that LCP(q, M) is solvable for all $q \in \mathbb{R}^n$, and if z is the unique solution of LCP(q, M) and $J = \{j \in \overline{n} \mid z_j > 0\}$ then $q \in \mathcal{Q}^J$. Thus, we have $\underset{J\subset\overline{n}}{\cup} \mathcal{Q}^J = \mathbb{R}^n$.

iii. Note that if $q \in \mathcal{Q}^J$ then z with $z_J = -M_{JJ}^{-1}q_J$ and $z_{\overline{n}\setminus J} = 0$ is the (unique) solution of LCP(q, M). Then, the function $q \mapsto z$ can be given by

$$z = A^J q \text{ if } q \in \mathcal{Q}^J$$

where $A_{JJ}^J = -M_{JJ}^{-1}$ and $A_{KL}^J = 0$ for $J \cap K \cap L = \emptyset$. Moreover, it is continuous due to the uniqueness of the solution of the corresponding LCP and Lipschitz continuous on $\mathcal{Q}^J$ with the constant $\|A^J\|$.

iv. Notice that $\|A^J\| = \|M_{JJ}^{-1}\|$.

The facts i-iv enables us to get the required result by applying Lemma 2.5.1. ■

2.5.2 Rational matrices with index 1

We will characterize the index of a rational matrix in terms of its power series expansion around infinity in the following Lemma.

Lemma 2.5.3 *Let $H(s) \in \mathbb{R}^{l\times l}(s)$ be given and let its power series expansion around infinity be given by*

$$H(s) = H^0 + H^1 s^{-1} + \cdots .$$

Then, the following statement are equivalent.

1. *$H(s)$ is of index 1.*

2. *$H^0 + H^1 s^{-1}$ is of index 1.*

3. *$\operatorname{im} H^0 \oplus H^1(\ker H^0) = \mathbb{R}^l$.*

4. *There exist matrices $P \in \mathbb{R}^{p\times l}$ and $Q \in \mathbb{R}^{(l-p)\times l}$ such that*

$$\begin{bmatrix} P \\ Q \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} PH^0 \\ QH^1 \end{bmatrix}$$

are both nonsingular and $QH^0 = 0$.

Proof: We achieve the proof of this lemma in the following order.

$$1 \Leftrightarrow 2$$
$$2 \Rightarrow 3$$
$$3 \Rightarrow 4$$
$$4 \Rightarrow 2$$

We denote the rational matrices $H^0 + H^1 s^{-1}$ and $H^2 s^{-2} + H^3 s^{-3} + \cdots$ by $H_{low}(s)$ and $H_{high}(s)$, respectively.

1 ⇒ 2: Since $H(s)$ is of index 1, it is invertible as a rational matrix. Thus, we have

$$[H_{low}(s) + H_{high}(s)]H^{-1}(s) = I. \tag{2.4}$$

Since $H(s)$ is of index 1, $s^{-2}H^{-1}(s)$ is strictly proper and so is $H_{high}(s)H^{-1}(s)$. It follows from (2.4) that

$$\lim_{s\to\infty} H_{low}(s)H^{-1}(s) = I. \tag{2.5}$$

Therefore, $H_{low}(s)H^{-1}(s)$ is biproper. This means that $H_{low}(s)$ is also invertible as a rational matrix. On the other hand, (2.5) can be rewritten as

$$\lim_{s\to\infty} sH(s)\ s^{-1}H_{low}^{-1}(s) = I. \tag{2.6}$$

Since $H(s)$ is of index 1, $\lim_{s\to\infty} s^{-1}H^{-1}(s)$ is well-defined. Left multiplying (2.6) by $\lim_{s\to\infty} s^{-1}H^{-1}(s)$ results in

$$\lim_{s\to\infty} s^{-1}H_{low}^{-1}(s) = \lim_{s\to\infty} s^{-1}H^{-1}(s).$$

Clearly, this implies that $s^{-1}H_{low}^{-1}(s)$ is proper and hence $H_{low}(s)$ is of index 1.

2 ⇒ 1: Note that

$$H(s) = H_{low}(s)[I + H_{low}^{-1}(s)H_{high}(s)]. \tag{2.7}$$

Since $H_{low}(s)$ is of index 1, $s^{-2}H_{low}^{-1}(s)$ is strictly proper and so is $H_{low}^{-1}(s)H_{high}(s)$. Therefore, the second factor on the right hand side of (2.7) is biproper. Consequently, $H(s)$ is of index 1.

2 $\Rightarrow$ *3:* Since $H^0 + H^1 s^{-1}$ is of index 1, the power series expansion of its inverse can be given by

$$(H^0 + H^1 s^{-1})^{-1} = N^{-1} s + N^0 + N^1 s^{-1} + \cdots . \tag{2.8}$$

Note that (2.8) gives

$$H^0 N^{-1} = 0 \tag{2.9}$$

$$N^{-1} H^0 = 0 \tag{2.10}$$

$$N^{-1} H^1 + N^0 H^0 = I. \tag{2.11}$$

i. Suppose that $u \in \operatorname{im} H^0 \cap H^1(\ker H^0)$. Then, we have

$$u = H^0 v \tag{2.12}$$

$$u = H^1 w \tag{2.13}$$

$$0 = H^0 w \tag{2.14}$$

for some v and w. It follows that

$$(H^0 + H^1 s^{-1}) w \overset{(2.14)}{=} s^{-1} H^1 w \overset{(2.13)}{=} s^{-1} u \overset{(2.12)}{=} s^{-1} H^0 v.$$

Then, (2.8) yields

$$w = (N^{-1} s + N^0 + N^1 s^{-1} + \cdots) H^0 v s^{-1} \overset{(2.10)}{=} N^0 H^0 v s^{-1} + \cdots .$$

Since the right hand side of the above equation is strictly proper, w is zero and so is u due to (2.13). Hence, $\operatorname{im} H^0 \cap H^1(\ker H^0) = \{0\}$.

ii. We have

a. (2.9) $\Rightarrow \operatorname{im} N^{-1} \subseteq \ker H^0$,

b. (2.11) $\Rightarrow (v \in \ker H^0 \Rightarrow v \in \operatorname{im} N^{-1}) \Rightarrow \ker H^0 \subseteq \operatorname{im} N^{-1}$.

Obviously, (a) and (b) imply that $\ker H^0 = \operatorname{im} N^{-1}$. Thus, one gets $H^1(\ker H^0) = \operatorname{im} H^1 N^{-1}$. Suppose that $u \in (\operatorname{im} H^0 + H^1(\ker H^0))^\perp$, i. e.,

$$u^\top H^0 = 0 \tag{2.15}$$

$$u^\top H^1 N^{-1} = 0. \tag{2.16}$$

Then, we get

$$\begin{aligned} u^\top &\overset{(2.8)}{=} u^\top(H^0 + H^1 s^{-1})(N^{-1}s + N^0 + N^1 s^{-1} + \cdots) \\ &\overset{(2.15)}{=} u^\top H^1 s^{-1}(N^{-1}s + N^0 + N^1 s^{-1} + \cdots) \\ &\overset{(2.16)}{=} u^\top H^1 N^0 s^{-1} + \cdots . \end{aligned}$$

The fact that the right hand side of the above equation is strictly proper implies that u is zero. Hence, $\operatorname{im} H^0 + H^1(\ker H^0) = \mathbb{R}^l$.

It follows from (i) and (ii) that $\operatorname{im} H^0 \oplus H^1(\ker H^0) = \mathbb{R}^l$.

3 ⇒ 4: Let $Q \in \mathbb{R}^{q\times l}$ be such that $\ker Q = \operatorname{im} H^0$. Take any $P \in \mathbb{R}^{(l-q)\times l}$ such that $\operatorname{col}(P, Q)$ is nonsingular. Suppose that

$$\begin{bmatrix} PH^0 \\ QH^1 \end{bmatrix} x = 0 \tag{2.17}$$

for some $x \in \mathbb{R}^l$. Since $\ker Q = \operatorname{im} H^0$, we have $\operatorname{col}(P, Q)H^0 x = 0$ from (2.17). This implies that, $H^0 x = 0$, i.e., $x \in \ker H^0$. Hence, $H^1 x \in H^1(\ker H^0)$. On the other hand, (2.17) also yields $H^1 x \in \ker Q = \operatorname{im} H^0$. Therefore, $H^1 x \in \operatorname{im} H^0 \cap H^1(\ker H^0)$. It follows from the hypothesis that $H^1 x = 0$. Note that

$$\dim(\operatorname{im} H^0) + \dim(H^1(\ker H^0)) = l = \dim(\operatorname{im} H^0) + \dim(\ker H^0).$$

Thus, we have $\dim(H^1(\ker H^0)) = \dim(\ker H^0)$. In other words, $\ker(H^1|_{\ker H^0}) = \{0\}$. It follows from $H^0 x = H^1 x = 0$ that $x = 0$ and hence $\operatorname{col}(PH^0, QH^1)$ is nonsingular.

4 ⇒ 2: Note that

$$\begin{aligned} H^0 + s^{-1}H^1 &= \begin{pmatrix} P \\ Q \end{pmatrix}^{-1} \begin{pmatrix} P \\ Q \end{pmatrix} (H^0 + s^{-1}H^1) = \begin{pmatrix} P \\ Q \end{pmatrix}^{-1} [\begin{pmatrix} PH^0 \\ 0 \end{pmatrix} + s^{-1} \begin{pmatrix} PH^1 \\ QH^1 \end{pmatrix}] \\ &= \begin{pmatrix} P \\ Q \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ 0 & s^{-1}I \end{pmatrix} [\begin{pmatrix} PH^0 \\ QH^1 \end{pmatrix} + s^{-1} \begin{pmatrix} PH^1 \\ 0 \end{pmatrix}]. \end{aligned} \tag{2.18}$$

It follows from the hypothesis that the matrix $\operatorname{col}(PH^0, QH^1)$ is nonsingular that $H^0 + s^{-1}H^1$ is of index 1. ∎

In particular, the constant $d(\cdot)$ of index 1 rational matrices will be of interest.

Lemma 2.5.4 *Let $H(s) \in \mathbb{R}^{l\times l}(s)$ be totally of index 1. Then, there exists an $\alpha > 0$ such that $d(H(\sigma)) \leq \alpha\sigma$ for all sufficiently large σ.*

Proof: Note that $H_{JJ}(s)$ is of index 1 for each index set $J \subseteq \overline{m}$ by the definition of total index. Hence, $s^{-1}H_{JJ}^{-1}(s)$ is proper for each $J \subseteq \overline{m}$. Therefore, for each $J \subseteq \overline{m}$ there exists $\alpha_J > 0$ such that

$$\|H_{JJ}^{-1}(\sigma)\| \leq \alpha_J\sigma$$

for all sufficiently large σ. As a consequence, we have $d(H(\sigma)) \leq \alpha\sigma$ for all sufficiently large σ where $\alpha = \max_{J\subseteq\overline{m}} \alpha_J$. ■

2.5.3 Towards to the proof of Theorem 2.3.3

In this subsection, we make necessary preparations for the Proof of Theorem 2.3.3. We begin with defining the concept of *initial solution*. We say that a continuous function $v : \mathbb{R}_+ \to \mathbb{R}$ is *initially nonnegative* if there exists $\epsilon > 0$ such that $v(t) \geq 0$ for all $t \in [0, \epsilon)$.

Definition 2.5.5 A triple $(u, x, y) \in \mathcal{B}^{m+n+m}$ is an *initial solution* of LCS(A, B, C, D, E) for the input $w \in \mathcal{B}$ and the initial state x_0 if there exists an index set $K \subseteq \overline{m}$ such that

$$\begin{aligned} \dot{x} &= Ax + Bu + Ew, \ x(0) = x_0 \\ y &= Cx + Du \\ y_K &= 0 \\ u_{\overline{m}\setminus K} &= 0 \end{aligned}$$

holds, and u and y are initially nonnegative.

Next, we recall the so-called the *Rational Complementarity Problem* (see [4] for a detailed discussion in the case with no external inputs).

Problem 2.5.6 (RCP$(x_0, \hat{w}(s), A, B, C, D, E)$) Given $x_0 \in \mathbb{R}^n$, $\hat{w}(s) \in \mathbb{R}^p(s)$, and (A, B, C, D, E) with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times m}$ and $E \in \mathbb{R}^{n\times p}$ find $\hat{u}(s) \in \mathbb{R}^m(s)$ such that

1. $\hat{u}(s) \perp \hat{y}(s)$ for all $s \in \mathbb{C}$.

2. $\hat{u}(\sigma) \geq 0$ and $\hat{y}(\sigma) \geq 0$ for all sufficiently large $\sigma \in \mathbb{R}$.

where

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + C(sI - A)^{-1}E\hat{w}(s) + [D + C(sI - A)^{-1}B]\hat{u}(s).$$

For brevity of notation, we denote RCP$(x_0, \hat{w}(s), A, B, C, D, E)$ by RCP$(x_0, \hat{w}(s))$ if (A, B, C, D, E) is clear from the context. There is one-to-one correspondence between the strictly proper solutions of RCP and initial solutions of LCS as described in the following lemma.

Lemma 2.5.7 *Consider a given matrix quintuple (A, B, C, D, E). The following statements hold.*

1. *Let $\hat{u}(s)$ be a strictly proper solution of RCP$(x_0, \hat{w}(s))$ for some x_0 and strictly proper $\hat{w}(s)$. Define $\hat{x}(s)$ and $\hat{y}(s)$ as follows*

$$\begin{aligned}\hat{x}(s) &= (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s) + (sI - A)^{-1}E\hat{w}(s),\\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s).\end{aligned}$$

Then, the inverse Laplace transform (u, x, y) of $(\hat{u}(s), \hat{x}(s), \hat{y}(s))$ is an initial solution of LCS(A, B, C, D, E) for the input w and the initial state x_0 where w is the inverse Laplace transform of $\hat{w}(s)$.

2. *Let (u, x, y) be an initial solution of LCS(A, B, C, D, E) for the input w and the initial state x_0 and let $\hat{u}(s)$ be the Laplace transform of u. Then, $\hat{u}(s)$ solves RCP$(x_0, \hat{w}(s))$ where $\hat{w}(s)$ is the Laplace transform of w.*

Proof: Evident from the proof of [5, Theorem 5.3]. ∎

The following lemma will play a key role in the proof of Theorem 2.3.3.

Lemma 2.5.8 *Consider a matrix quintuple (A, B, C, D, E). Suppose that $G(s) := D + C(sI - A)^{-1}B$ is totally of index 1 and $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then the following statements hold.*

1. *RCP$(x_0, \hat{w}(s))$ has a unique solution for all $x_0 \in \mathbb{R}^n$ and for all $\hat{w}(s) \in \mathbb{R}^p(s)$.*

2. *For a given strictly proper $\hat{w}(s)$, the unique solution of RCP$(x_0, \hat{w}(s))$ is strictly proper if and only if $Cx_0 \in K(D)$.*

Proof:

1: Since $D + C(\sigma I - A)^{-1}B$ is a $\mathcal{P}$-matrix for all sufficiently large σ, the statement follows from [4, Theorem 4.1] and Lemma 1.2.4 item 1.

2: Let $\hat{u}(s)$ be the unique solution of RCP$(x_0, \hat{w}(s))$. For the 'only if' part, suppose that $\hat{u}(s)$ is strictly proper. Let the power series expansion around infinity of $\hat{u}(s)$ and

$\hat{w}(s)$ be of the form

$$\hat{u}(s) = u_1 s^{-1} + u_2 s^{-2} + \dots \tag{2.19a}$$
$$\hat{w}(s) = w_1 s^{-1} + w_2 s^{-2} + \dots . \tag{2.19b}$$

Define

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + C(sI - A)^{-1}E\hat{w}(s) + [D + C(sI - A)^{-1}B]\hat{u}(s).$$

By substituting (2.19) into the above equation, we get

$$\hat{y}(s) = (Cx_0 + Du_1)s^{-1} + (CAx_0 + CEw_1 + CBu_1)s^{-2} + \dots .$$

It follows from the formulation of $\mathrm{RCP}(x_0, \hat{w}(s))$ that $u_1 \geq 0$, $Cx_0 + Du_1 \geq 0$ and $u_1^\top(Cx_0 + Du_1) = 0$. Consequently, $\mathrm{LCP}(Cx_0, D)$ is solvable. In other words, $Cx_0 \in K(D)$. To show the 'if' part, suppose that $Cx_0 \in K(D)$. Let $\bar{u}$ be a solution of $\mathrm{LCP}(Cx_0, D)$. It is clear that $\sigma^{-1}\bar{u}$ solves $\mathrm{LCP}(\sigma^{-1}Cx_0, D)$ for all $\sigma > 0$. Then, it also solves $\mathrm{LCP}(\sigma^{-1}Cx_0 - \sigma^{-1}C(\sigma I - A)^{-1}B\bar{u}, G(\sigma))$. Lemma 2.5.2 together with Lemma 2.5.4 gives

$$\begin{aligned} \|\hat{u}(\sigma) - \sigma^{-1}\bar{u}\| \leq \alpha\sigma \| & C[(\sigma I - A)^{-1} - \sigma^{-1}I]x_0 \\ & + C(\sigma I - A)^{-1}E\hat{w}(\sigma) + \sigma^{-1}C(\sigma I - A)^{-1}B\bar{u}\| \end{aligned} \tag{2.20}$$

for all sufficiently large σ. Note that for some $\beta > 0$ the final factor at the last term of the right hand side is less than $\beta\sigma^{-2}$ for all sufficiently large σ. Therefore, it follows from (2.20) that $\|\hat{u}(\sigma) - \sigma^{-1}\bar{u}\| \leq \alpha\beta\sigma^{-1}$ for all sufficiently large σ. This implies that $\hat{u}(s)$ is strictly proper. ∎

As a final ingredient of the proof of Theorem 2.3.3, we need the following lemma on the elimination of algebraic constraints.

Lemma 2.5.9 *Consider a given matrix quintuple (A, B, C, D, E). Suppose that $G(s) = D + C(sI - A)^{-1}B$ is totally of index 1. For all $K \subseteq \overline{m}$ there exist matrices F^K, G^K, H^K and J^K such that if $(w, u, x, y) \in \mathcal{F}(\mathbb{R}, \mathbb{R}^{p+m+n+m})$ satisfies*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Ew(t) \\ y(t) &= Cx(t) + Du(t) \\ y_K(t) &= 0 \\ u_{\overline{m}\setminus K}(t) &= 0 \end{aligned}$$

for all $t \in (t', t'')$ *then they also satisfy*

$$\begin{aligned}\dot{x}(t) &= F^K x(t) + G^K w(t)\\ u(t) &= H^K x(t) + J^K w(t)\\ y(t) &= Cx(t) + Du(t).\end{aligned}$$

Proof: Clearly, the triple (u, x, y) satisfies

$$\begin{aligned}\dot{x} &= Ax + B_{\bullet K} u_K + Ew\\ 0 &= C_{K\bullet} x + D_{KK} u_K.\end{aligned} \tag{2.21}$$

Since $G(s)$ is totally of index 1, Lemma 2.5.3 implies that there exist matrices P^K and Q^K such that $\mathrm{col}(P^K, Q^K)$ and $\mathrm{col}(P^K D_{KK}, Q^K C_{K\bullet} B_{\bullet K})$ are both nonsingular and $Q^K D_{KK} = 0$. By premultiplying (2.21) by the first matrix above, we get

$$P^K C_{K\bullet} x + P^K D_{KK} u_K = 0 \tag{2.22}$$

$$Q^K C_{K\bullet} x = 0. \tag{2.23}$$

Differentiating (2.23) with respect to time, one gets

$$Q^K C_{K\bullet} Ax + Q^K C_{K\bullet} B_{\bullet K} u_K + Q^K C_{K\bullet} Ew = 0. \tag{2.24}$$

By combining (2.22) and (2.24), one can obtain

$$\begin{bmatrix} P^K D_{KK} \\ Q^K C_{K\bullet} B_{\bullet K} \end{bmatrix} u_K = -\begin{bmatrix} P^K C_{K\bullet} \\ Q^K C_{K\bullet} A \end{bmatrix} x - \begin{bmatrix} 0 \\ Q^K C_{K\bullet} E \end{bmatrix} w. \tag{2.25}$$

Since the factor of u_K is nonsingular, the matrices H^K and J^K can be found by solving u_K from (2.25). F^K and G^K can be given as $F^K = A + BH^K$ and $G^K = E + BJ^K$. ■

2.5.4 Proofs of Theorem 2.3.3 and Theorem 2.3.4

After all these preparations, we can finally prove Theorem 2.3.3.

Proof of Theorem 2.3.3:

2 $\Rightarrow$ *1:* Let the input $w^{\bar{x}} \in \mathcal{PBB}^p$ and the initial state $\bar{x}$ with $C\bar{x} \in K(D)$ be given. Define $v^{\bar{x}} = \alpha(w^{\bar{x}}, 0)$. Note that we have $w^{\bar{x}}|_{[0,\epsilon)} = v^{\bar{x}}|_{[0,\epsilon)}$ whenever $\epsilon \leq \beta(w^{\bar{x}}, 0)$. It follows from Lemma 2.5.8 items 1 and 2 that $\mathrm{RCP}(\bar{x}, \hat{v}^{\bar{x}}(s))$ has a unique strictly proper solution where $\hat{v}^{\bar{x}}(s)$ is the Laplace transform of $v^{\bar{x}}$. Hence, Lemma 2.5.7 item 1 implies that there exists an initial solution $(u^{\bar{x}}, x^{\bar{x}}, y^{\bar{x}})$ of $\mathrm{LCS}(A, B, C, D, E)$ for the input $v^{\bar{x}}$ and the initial state $\bar{x}$. We define $\iota : \mathbb{R}^n \times \mathcal{PBB}^p \to 2^{\overline{m}}$ as $\iota(\bar{x}, w^{\bar{x}}) = K$ where the index set $K = \{k \in \overline{m} \mid u_k^{\bar{x}} \not\equiv 0\}$, $\tau : \mathbb{R}^n \times \mathcal{PBB}^p \to \mathbb{R}_{++}$ as

$$\tau(\bar{x}, w^{\bar{x}}) = \sup\{T \mid T < \beta(w^{\bar{x}}, 0) \text{ and } \mathrm{col}(u_j^{\bar{x}}(t), y_j^{\bar{x}}(t)) \geq 0 \text{ for all } t \in [0, T]\},$$

and $\kappa : \mathbb{R}^n \times \mathcal{PBB}^p \to \mathbb{R}^n$ as

$$\kappa(\bar{x}, w^{\bar{x}}) = x^{\bar{x}}(\tau(\bar{x})).$$

Note that $t \mapsto (u^{\bar{x}}, x^{\bar{x}}, y^{\bar{x}})(t + \rho)$ forms an initial solution of LCS(A, B, C, D, E) for the input $t \mapsto w^{\bar{x}}(t + \rho)$ and the initial state $x^{\bar{x}}(\rho)$ whenever $\rho \in [0, \tau(\bar{x}, w^{\bar{x}}))$. Hence, we have $Cx^{\bar{x}}(\rho) \in K(D)$ for all $\rho \in [0, \tau(\bar{x}, w^{\bar{x}}))$ due to Lemma 2.5.7 item 2 and Lemma 2.5.8 item 2. It follows from Fact 1.2.2 (i.e., the closedness of the set $K(D)$) and continuity of $x^{\bar{x}}$ that $\kappa(\bar{x}, w^{\bar{x}}) \in K(D)$.

<u>existence:</u> For a given input $w \in \mathcal{PBB}^p$, define $x_{i+1} = \kappa(x_i, w^i)$ for $i = 0, 1, \ldots$ where $w^i = w^{i-1}|_{[\tau(x_{i-1}, w^{i-1}), \infty)}$ if $i \neq 0$ and $w^0 = w$. From the previous discussion, we know that $Cx_0 \in K(D)$ implies that $Cx_i \in K(D)$ for all $i = 0, 1, \ldots$. Hence, LCS(A, B, C, D, E) admits a unique solution for the input $\alpha(w^i, 0)$ and the initial state x_i for all $i = 0, 1, \ldots$ due to Lemma 2.5.8 item 2 and Lemma 2.5.7 item 1. Let $(u^{x_i}, x^{x_i}, y^{x_i})$ denote the initial solution of LCS(A, B, C, D, E) for the input $\alpha(w^i, 0)$ and the initial state x_i. Define $\tau_k = \sum_{i=1}^{k} \tau(x_{k-1}, w^{k-1})$ for $k > 0$ and $\tau_0 = 0$ and also define

$$(u, x, y)|_{[\tau_k, \tau_{k+1}]} = (u^{x_k}, x^{x_k}, y^{x_k})|_{[0, \tau(x^k)]}.$$

It can be verified that (u, x, y) is a solution on $[0, T)$ for some $T > 0$ of LCS(A, B, C, D, E) for the input w and the initial state x_0. Suppose that $T < \infty$ is such that there is no solution on $[0, T')$ whenever $T' > T$. Note that (w, u, x, y) satisfies

$$\dot{x}(t) = F^{\iota(x_k, w^k)} x(t) + G^{\iota(x_k, w^k)} w(t)$$

for $t \in (\tau_k, \tau_{k+1})$ due to Lemma 2.5.9. Since x and $t \mapsto e^{F^L t} G^L$ for $L \subseteq \overline{m}$ is continuous $[0, T)$ and $w \in \mathcal{PBB}^p$, they are all bounded on $[0, T)$, i.e., there exists a $\mu > 0$ such that $\|x(t)\| \leq \mu$ and $\|e^{F^L t} G^L w(t)\| \leq \mu$ for all $t \in [0, \tau)$ and for all $L \subseteq \overline{m}$. Then, we have

$$\begin{aligned} \|\tilde{x}(t) - \tilde{x}(\rho)\| &\leq \|e^{F^{\iota(x_k, w^k)}(t-\rho)} \tilde{x}(\rho) - \tilde{x}(\rho)\| + \|\int_\rho^t e^{F^{\iota(x_k, w^k)}(t-s)} G^{\iota(x_k, w^k)} w(s) ds\| \\ &\leq (1 + \nu_{\iota(x_k, w^k)}) \mu |t - \rho| \end{aligned} \tag{2.26}$$

for all $\rho, t \in (\tau_k, \tau_{k+1})$ since the function $t \mapsto \frac{e^{F^K t} - I}{t}$ is bounded, say by ν_K. Hence, for $\rho, t \in [0, T)$, we get from (2.26) $\|x(t) - x(\rho)\| \leq \mu[\max_{K \subseteq \overline{m}} (1 + \nu_K)]|t - \rho|$. It follows that x is Lipschitz continuous on $[0,T)$ and thus uniformly continuous. A standard result in mathematical analysis [8, Exercise 4.13] implies that $x^* := \lim_{t\uparrow\tau} x(t)$ exists. Since $Cx(t) \in K(D)$ for all $t \in [0,T)$ and x is continuous, $Cx^* \in K(D)$ which means one can extend the solution (u,x,y) beyond $[0,T)$ by using the initial solution of LCS(A,B,C,D,E) for the input $w|_{[T,\infty)}$ and the initial state x^*. This contradicts the definition of T. Thus, we can conclude that there exists a solution on $[0,\infty)$ of LCS(A,B,C,D,E) for the input w and the initial state x_0.

uniqueness: Let $(u^i,x^i,y^i) \in \mathcal{PB}^{m+n+m}$ for $i=1,2$ denote two different solutions of LCS(A,B,C,D,E) for the input w and the initial state x_0. Clearly, $(u^1,x^1,y^1)-(u^2,x^2,y^2)$ is a piecewise Bohl function as well. If it is not identically zero then there should exists $\tau \geq 0$ and $\epsilon > 0$ such that $((u^1,x^1,y^1)-(u^2,x^2,y^2))|_{[0,\tau]} \equiv 0$ and $((u^1,x^1,y^1)-(u^2,x^2,y^2))(t) \neq 0$ for all $t \in (\tau,\tau+\epsilon)$ due to the definition of piecewise Bohl functions. For (u^i,x^i,y^i) and $\tau \geq 0$, one can find $\epsilon_i > 0$ and Bohl functions $(\bar{u}^i,\bar{x}^i,\bar{y}^i)$ such that $(u^i,x^i,y^i)|_{[\tau,\tau+\epsilon_i)} = (\bar{u}^i,\bar{x}^i,\bar{y}^i)|_{[0,\epsilon_i)}$ with $i=1,2$ again by the definition of piecewise Bohl functions. It is easy to see that $(\bar{u}^i,\bar{x}^i,\bar{y}^i)$ forms two different initial solutions of LCS(A,B,C,D,E) for the input $\beta(w|_{[\tau,\infty)})$ and the same initial state, $x^1(\tau)=x^2(\tau)$. Then, Lemma 2.5.7 item 2 implies that Laplace transforms of $\bar{u}^i$ are solutions of RCP$(x^1(\tau),\beta(w|_{[\tau,\infty)}))$. However, it is known from Lemma 2.5.8 item 1 that RCP$(x^1(\tau),\beta(w|_{[\tau,\infty)}))$ has a unique solution since $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Therefore, $\bar{u}^1=\bar{u}^2$. It follows that $\bar{x}^1=\bar{x}^2$ and $\bar{y}^1=\bar{y}^2$. Thus, we have $((u^1,x^1,y^1)-(u^2,x^2,y^2))(t)=0$ for all $t \in [\tau,\tau+\min(\epsilon_1,\epsilon_2))$. This contradicts the definition of τ.

1 ⇒ 2: Let the unique solution of LCS(A,B,C,D,E) for the input w and the initial state x_0 be (u,x,y). Since $w \in \mathcal{PBB}^p$ and $(u,x,y) \in \mathcal{PB}^{m+n+m}$, there exist ϵ_w, ϵ_u, ϵ_x, ϵ_y and $(w',u',x',y') \in \mathcal{B}^{p+m+n+m}$ such that $w|_{[0,\epsilon_w)} = w'|_{[0,\epsilon_w)}$, $u|_{[0,\epsilon_u)} = u'|_{[0,\epsilon_u)}$, $x|_{[0,\epsilon_x)} = x'|_{[0,\epsilon_x)}$ and $y|_{[0,\epsilon_y)} = y'|_{[0,\epsilon_y)}$. Define $\epsilon = \min(\epsilon_w,\epsilon_u,\epsilon_x,\epsilon_y)$. Then, (w',u',x',y') satisfies

$$x'(t) = x_0 + \int_0^t [Ax'(s) + Bu'(s) + Ew'(s)]\, ds \tag{2.27a}$$

$$y'(t) = Cx'(t) + Du'(t) \tag{2.27b}$$

$$0 \leq u'(t) \perp y'(t) \geq 0 \tag{2.27c}$$

for all $t \in [0,\epsilon)$. Note that (2.27) implies that there exists an index set K such that

$$\dot{x}' = Ax' + Bu' + Ew', \; x'(0) = x_0 \tag{2.28a}$$

$$y' = Cx' + Du' \tag{2.28b}$$

$$y'_K \equiv 0 \quad u'_{\overline{m}\setminus K} \equiv 0 \tag{2.28c}$$

since Bohl functions are real-analytic. Note also that (2.27c) implies that both u' and y' are initially nonnegative. This fact together with (2.28) reveals that (u', x', y') is an initial solution of LCS(A, B, C, D, E) for the input w' and the initial state x_0. It follows from Lemma 2.5.7 item 2 and Lemma 2.5.8 item 2 that $Cx_0 \in K(D)$. ∎

Proof of Theorem 2.3.4: Since D is not a $\mathcal{P}_0$-matrix, there exists a principal submatrix of D, say D_{NN}, which has a negative eigenvalue due to [1, Theorem 3.4.2(c)]. Let $\lambda < 0$ denote such an eigenvalue of D_{NN} and let v be such that v_N is an eigenvector corresponding this eigenvalue. Define the index sets $J = \{j \in N \mid v_j > 0\}$, $K = \{k \in N \mid v_k < 0\}$ and $L = \overline{m} \setminus (J \cup K)$. Then, we have

$$\begin{pmatrix} D_{JJ} & D_{JK} \\ D_{KJ} & D_{KK} \end{pmatrix} \begin{pmatrix} v_J \\ v_K \end{pmatrix} = \lambda \begin{pmatrix} v_J \\ v_K \end{pmatrix}. \tag{2.29}$$

Since C is of full row rank, the equation

$$\begin{pmatrix} C_{J\bullet} \\ C_{K\bullet} \\ C_{L\bullet} \end{pmatrix} x_0 = \begin{pmatrix} -D_{JJ}v_J \\ D_{KK}v_K \\ \max(-D_{LJ}v_J, D_{LK}v_K) + e \end{pmatrix} \tag{2.30}$$

is solvable for x_0 where max is meant for componentwise maximum and e denotes the vector consisting of ones. Take (u^1, x^1, y^1) and (u^2, x^2, y^2) as

$$x^1 = e^{(A - B_{\bullet J} D_{JJ}^{-1} C_{J\bullet})t} x_0 \tag{2.31a}$$

$$u_J^1 = -D_{JJ}^{-1} C_{J\bullet} x^1 \qquad u_K^1 \equiv 0 \qquad u_L^1 \equiv 0 \tag{2.31b}$$

$$y_J^1 \equiv 0 \qquad y_K^1 = C_{K\bullet} x^1 + D_{KJ} u_J^1 \qquad y_L^1 = C_{L\bullet} x^1 + D_{LJ} u_J^1 \tag{2.31c}$$

$$x^2 = e^{(A - B_{\bullet K} D_{KK}^{-1} C_{K\bullet})t} x_0 \tag{2.32a}$$

$$u_K^2 = -D_{KK}^{-1} C_{K\bullet} x^2 \qquad u_J^2 \equiv 0 \qquad u_L^2 \equiv 0 \tag{2.32b}$$

$$y_K^2 \equiv 0 \qquad y_J^2 = C_{J\bullet} x^2 + D_{JK} u_K^2 \qquad y_L^2 = C_{L\bullet} x^2 + D_{LK} u_K^2. \tag{2.32c}$$

Note that (u^i, x^i, y^i) satisfies

$$\dot{x}^i = Ax^i + Bu^i \tag{2.33a}$$

$$y^i = Cx^i + Du^i \tag{2.33b}$$

$$0 \leq u^i \perp y^i \geq 0 \tag{2.33c}$$

for $i = 1, 2$. Furthermore, we have

$$u_J^1(0) \overset{(2.31b)}{=} -D_{JJ}^{-1} C_{J\bullet} x^1(0) \overset{(2.31a)}{=} -D_{JJ}^{-1} C_{J\bullet} x_0 \overset{(2.30)}{=} v_J > 0 \tag{2.34a}$$

$$y_K^1(0) \overset{(2.31c)}{=} C_{K\bullet} x^1(0) + D_{KJ} u_J^1(0) \overset{(2.31)}{=} C_{K\bullet} x_0 - D_{KJ} D_{JJ}^{-1} C_{J\bullet} x_0 \tag{2.34b}$$

$$\overset{(2.30)}{=} D_{KK} v_K + D_{KJ} v_J \overset{(2.29)}{=} \lambda v_K > 0$$

$$y_L^1(0) \overset{(2.31c)}{=} C_{L\bullet} x^1(0) + D_{LJ} u_J^1(0) \overset{(2.31)}{=} C_{L\bullet} x_0 - D_{LJ} D_{JJ}^{-1} C_{J\bullet} x_0 \tag{2.34c}$$

$$\overset{(2.30)}{=} C_{L\bullet} x_0 + D_{LJ} v_J \overset{(2.30)}{\geq} e \tag{2.34d}$$

and

$$u_K^2(0) \overset{(2.32b)}{=} -D_{KK}^{-1} C_{K\bullet} x^2(0) \overset{(2.32a)}{=} -D_{KK}^{-1} C_{K\bullet} x_0 \overset{(2.30)}{=} -v_K > 0 \tag{2.35a}$$

$$y_J^2(0) \overset{(2.32c)}{=} C_{J\bullet} x^1(0) + D_{JK} u_K^2(0) \overset{(2.32)}{=} C_{J\bullet} x_0 - D_{JK} D_{KK}^{-1} C_{K\bullet} x_0 \tag{2.35b}$$

$$\overset{(2.30)}{=} -D_{JJ} v_J - D_{JK} v_K \overset{(2.29)}{=} -\lambda v_J > 0$$

$$y_L^2(0) \overset{(2.32c)}{=} C_{L\bullet} x^2(0) + D_{LK} u_K^2(0) \overset{(2.32)}{=} C_{L\bullet} x_0 - D_{LK} D_{KK}^{-1} C_{K\bullet} x_0 \tag{2.35c}$$

$$\overset{(2.30)}{=} C_{L\bullet} x_0 - D_{LK} v_K \overset{(2.30)}{\geq} e. \tag{2.35d}$$

It follows from (2.33), (2.34) and (2.35) that (u^i, x^i, y^i) for $i = 1, 2$ is an initial solution of LCS(A, B, C, D, E) with the zero input and the initial state x_0. Note that $0 < u_J^1(0) \neq u_J^2(0) = 0$. Define $T_i = \sup\{\tau \mid u^i(t) \geq 0 \text{ and } y^i(t) \geq 0 \text{ for all } t \in [0, \tau]\}$ for $i = 1, 2$. Take $T = \min(T_1, T_2)$. ■

References

[1] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem.* Academic Press, Inc., Boston, 1992.

[2] C.A. Desoer and J. Katzenelson. Nonlinear RLC networks. *The Bell System Technical Journal*, 44:161–198, 1965.

[3] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides.* Mathematics and Its Applications. Prentice-Hall, Dordrecht, The Netherlands, 1988.

[4] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. The rational complementarity problem. *Linear Algebra and its Applications*, 294:93–135, 1999.

[5] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics*, 60(4):1234–1269, 2000.

[6] Y.J. Lootsma, A.J. van der Schaft, and M.K. Çamlıbel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[7] T. Ohtsuki and H. Watanabe. State-variable analysis of RLC networks containing nonlinear coupling elements. *IEEE Trans. on Circuit Theory*, 18(1):26–38, 1969.

[8] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1976.

[9] A.J. van der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9:266–301, 1996.

[10] A.J. van der Schaft and J.M. Schumacher. Complementarity modelling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[11] A.J. van der Schaft and J.M. Schumacher. *An Introduction to Hybrid Dynamical Systems*. Springer-Verlag, London, 2000.

[12] M. Vidyasagar. *Nonlinear System Analysis*. Prentice-Hall, London, 1992.

Chapter 3

Linear Passive Complementarity Systems

3.1 Introduction

In this chapter, we will continue to discuss the well-posedness of linear complementarity systems with an emphasis on the case when the underlying linear system is passive. In this case, we call the overall system a *linear passive complementarity system* (LPCS). The most typical example of LPCSs are the electrical networks consisting of linear resistors, inductors, capacitors, gyrators, transformers (RLCGT) and ideal diodes.

In circuit theory, most of the effort that has been invested in considering existence and uniqueness of solutions to electrical networks is focused on *static* (DC) models of networks [1, 2, 5–7, 13, 14, 16–19]. The studies of dynamic equivalent are rare. The only papers known to the author dealing with existence and uniqueness of solutions of (dynamic) RLC-networks with non-Lipschitzian elements are [4, 15]. Since an ideal diode cannot be formulated as a current- or voltage-controlled resistor, the obtained results in [4, 15] do not cover the networks containing diodes.

Not surprisingly, the selection of *universum*, the space where all possible solutions live, plays a key role in the study of existence and uniqueness of solutions. The universum that was proposed in the previous chapter has been motivated from a hybrid system point of view. Generally speaking, the solution concepts that have been developed in this context impose a *direction* on time. For instance, the whole idea of hybrid automaton modeling is based on what we call *forward* thinking. More precisely, the system is presumed to evolve onward by starting from a point in time. Although it might be reasonable from a computer science viewpoint, there are no obvious reasons to treat time asymmetrically in modeling of physical systems. Our goals in this chapter are

- to propose a new solution concept in which the time will be treated symmetrically,
- to compare it with the previous one,
- to establish the existence and uniqueness of solutions and to characterize the set of regular initial states of LPCSs,
- to investigate the so-called Zeno behavior of LPCS,
- to extend the new solution concept in order to treat nonregular initial states as well.

The outline of the chapter is as follows. We begin with recalling passivity and the Kalman-Yakubovich-Popov lemma in Section 3.2. In Section 3.3, three different solution concepts will be proposed and their relations will be discussed. This will be followed by the introduction of a new class of systems, namely the systems that are *passifiable by pole shifting* in Section 3.4. After establishing necessary and sufficient conditions for passifiability by pole shifting, all the existing results on linear passive complementarity systems will be generalized to this new class of systems. Section 3.5 is devoted to a brief discussion on the so-called Zeno behavior of linear complementarity systems. We start with considering nonregular initial states of a linear passive complementarity system in Section 3.6. After proposing a *jump rule* in terms of the stored energy of the passive system, we reach a new (distributional) solution concept for linear passive complementary systems which treats nonregular initial states as well. Then, a number of equivalent characterizations of the jump rule will be in order. One of those characterizations will open the possibility to motivate a jump rule for general linear (possibly nonpassive) complementarity systems and this section will be completed by extending the distributional framework to these systems. As usual the chapter will be closed by conclusions in Section 3.7 and proofs in Section 3.8.

3.2 Passive Systems

Ever since it was introduced in system theory by V. M. Popov, the notion of passivity has played an important role in various contexts such as stability issues, adaptive control, identification etc. Particularly, the interest in stability issues led to the theory of dissipative systems [22] due to J. C. Willems. Before going further, we will quickly recall the notion of passivity as it is defined in [22].

Consider a continuous-time, linear and time-invariant system given by

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3.1a}$$

$$y(t) = Cx(t) + Du(t) \tag{3.1b}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$ and A, B, C, and D are matrices with appropriate sizes. We denote (3.1) by $\Sigma(A, B, C, D)$.

A triple $(u, x, y) \in \mathcal{L}_2((t_0, t_1), \mathbb{R}^{m+n+m})$ is said to be an *$\mathcal{L}_2$-solution* on (t_0, t_1) of $\Sigma(A, B, C, D)$ with the initial state x_0 if it satisfies (3.1a) in the sense of Carathéodory, i.e., for almost all $t \in [t_0, t_1]$,

$$x(t) = x_0 + \int_{t_0}^{t} [Ax(s) + Bu(s)]ds \tag{3.2}$$

and (3.1b) holds.

Definition 3.2.1 [22] The system $\Sigma(A, B, C, D)$ given by (3.1) is said to be *passive (dissipative with respect to the supply rate $u^\top y$)* if there exists a function $V : \mathbb{R}^n \to \mathbb{R}_+$ (a *storage function*), such that

$$V(x(t_0)) + \int_{t_0}^{t_1} u^\top(t)y(t)dt \geq V(x(t_1)) \tag{3.3}$$

holds for all t_0 and t_1 with $t_1 \geq t_0$, and for all $\mathcal{L}_2$-solutions $(u, x, y) \in \mathcal{L}_2((t_0, t_1), \mathbb{R}^{m+n+m})$ of $\Sigma(A, B, C, D)$.

Next, we quote a very well-known characterization of passivity.

Theorem 3.2.2 [22] *Assume that (A, B, C) is minimal. Let $G(s) = C(sI - A)^{-1}B + D$ be the transfer matrix of $\Sigma(A, B, C, D)$. Then the following statements are equivalent:*

1. *$\Sigma(A, B, C, D)$ is passive.*

2. *The matrix inequalities*

$$K = K^\top > 0 \text{ and } \begin{bmatrix} A^\top K + KA & KB - C^\top \\ B^\top K - C & -(D + D^\top) \end{bmatrix} \leq 0$$

   have a solution.

3. *$G(s)$ is positive real, i.e., $G(\lambda) + G^\top(\bar{\lambda}) \geq 0$ for all $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$.*

Moreover, $V(x) = \frac{1}{2}x^\top Kx$ defines a quadratic storage function if and only if K satisfies the above system of linear matrix inequalities.

The equivalence of the statements 2 and 3 is sometimes called the positive real lemma or the Kalman-Yakubovich-Popov lemma.

3.3 Linear Passive Complementarity Systems

As the interconnection of a continuous, time-invariant, linear system and complementarity conditions, a *linear complementarity system* can be given by

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3.4a}$$

$$y(t) = Cx(t) + Du(t) \tag{3.4b}$$

$$0 \leq u(t) \perp y(t) \geq 0 \tag{3.4c}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$, and A, B, C and D are matrices with appropriate sizes. We denote the above system by LCS(A, B, C, D).

In what follows we will define several solution concepts for LCSs and investigate their relations. The first one is the zero-input version of Definition 2.2.2.

Definition 3.3.1 The triple $(u, x, y) \in \mathcal{PB}^{m+n+m}$ is a *$\mathcal{PB}$-solution* on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 if the following conditions hold for all $t \in [0, T]$:

$$x(t) = x_0 + \int_0^t [Ax(s) + Bu(s)]\ ds$$

$$y(t) = Cx(t) + Du(t)$$

$$0 \leq u(t) \perp y(t) \geq 0.$$

We often make the following assumption on the system matrices.

Assumption 3.3.2 (A, B, C) is minimal and $\operatorname{col}(B, D + D^\top)$ is of full column rank.

We define the set $\mathcal{Q}_D = \text{SOL}(0, D) = \{v \mid v \geq 0, Dv \geq 0 \text{ and } v^\top Dv = 0\}$ for a given matrix D. It is known from the complementarity theory (see Lemma 1.2.4 item 2) that the dual cone of this set $\mathcal{Q}_D^*$ coincides with $K(D)$ if D is nonnegative definite. Then the following lemma follows as a direct implication of Theorem 2.3.3.

Theorem 3.3.3 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passive. Then, there exists a unique $\mathcal{PB}$-solution on $[0, \infty)$ of LCS(A, B, C, D) with the initial state x_0 if and only if $Cx_0 \in \mathcal{Q}_D^*$.*

As explained in the previous chapter, one way of looking at linear complementarity systems is to regard them as hybrid systems. A popular model for hybrid systems is the hybrid automaton model which combines finite automata with continuous dynamics. Basically, a hybrid automaton consists of a number of modes, dynamics associated to these modes and mode transition rules. Starting from a mode, the trajectories of the system evolve according to the dynamics of that mode until the mode transition rules trigger a mode change (called *event*). After the mode change, the dynamics of the new mode shapes the behavior of the system until the next event takes place. Our approach will put emphasis on the solution concept rather than the hybrid automaton model. Since our interest is focused on a rather special class of hybrid systems, our *hybrid* solution concept will be a trimmed version of a solution concept one needs for more general classes of hybrid systems. Nevertheless, our solution concept is more general than some existing ones in the sense that it allows existence of both *left* and *right accumulations* of event times. We begin with the definition of event times set.

Definition 3.3.4 A set $\mathcal{E} \subset \mathbb{R}_+$ is called an *admissible event times set* if it is closed and countable, and $0 \in \mathcal{E}$. To each admissible event times set $\mathcal{E}$, we associate a collection of intervals between events $\tau_\mathcal{E} = \{(t_1, t_2) \subset \mathbb{R}_+ \mid t_1, t_2 \in \mathcal{E} \cup \{\infty\}$ and $(t_1, t_2) \cap \mathcal{E} = \emptyset\}$.

Next, we define a *hybrid* solution concept which is general enough for linear complementarity systems with index 1.

Definition 3.3.5 A quintuple $(\mathcal{E}, \mathcal{S}, u, x, y)$ where $\mathcal{E}$ is an admissible event times set, $\mathcal{S} : \tau_\mathcal{E} \to 2^{\overline{m}}$, and $(u, x, y) \in \mathcal{PC}(\mathbb{R}_+, \mathbb{R}^{m+n+m})$ is said to be a *hybrid solution* of LCS(A, B, C, D) with the initial state x_0 if the following conditions hold.

1. x is continuous, piecewise differentiable and $x(0) = x_0$.

2. For each $\tau \in \tau_\mathcal{E}$ and for all $t \in \tau$, it holds that

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t) \\
y_{\mathcal{S}(\tau)}(t) = 0 &\quad u_{\overline{m} \setminus \mathcal{S}(\tau)}(t) = 0 \\
u_{\mathcal{S}(\tau)}(t) \geq 0 &\quad y_{\overline{m} \setminus \mathcal{S}(\tau)}(t) \geq 0.
\end{aligned}$$

Moreover, we say that a hybrid solution $(\mathcal{E}, \cdot, u, x, y)$ is *redundant* if there exists $t \in \mathcal{E}$ and t', t'' with $t' < t < t''$ such that (u, x, y) is analytic on (t', t''). It is said to be *nonredundant* otherwise.

While the first condition corresponds the continuous dynamics, the second one indicates that the *reset maps* for the hybrid automaton model corresponding to linear complementarity system are identity.

An interesting phenomenon that occurs in hybrid automaton modeling methodology is the accumulation of event times. This type of behavior is called *Zeno behavior* referring to

Zeno's paradox of Achilles and the turtle. We need to set up a language for accumulation points of event times.

Definition 3.3.6 An element t of an admissible set $\mathcal{E}$ is said to be a *left (right) accumulation point* if for all $t' > t$ ($t' < t$) $(t, t') \cap \mathcal{E}$ ($(t', t) \cap \mathcal{E}$) is not empty. An admissible event times set $\mathcal{E}$ is said to be *left (right) Zeno free* if it does not contain any left (right) accumulation points. A hybrid solution is said to be *left (right) Zeno* if the corresponding event times set contains at least one left (right) accumulation point and *non-Zeno* if the corresponding event times set contains no left or right accumulation points.

In the following proposition, the relation between the two solution concepts defined so far is established.

Proposition 3.3.7 *Consider a matrix quadruple (A, B, C, D). If (u, x, y) is a $\mathcal{PB}$-solution on $[0, \infty)$ of LCS(A, B, C, D) with some initial state then there exist a left Zeno free admissible event times set $\mathcal{E}$ and a mode indicator $\mathcal{S}$ such that $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a nonredundant hybrid solution of LCS(A, B, C, D) with the same initial state.*

Undoubtedly, the function space $\mathcal{PB}$ is not the most *natural* one to work with. Next, we introduce a solution concept in $\mathcal{L}_2$ which is clearly more natural. Later on, we will state stronger (in the sense that LCS admits unique solutions from a larger space) existence and uniqueness results by exploiting the structure provided by passivity.

Definition 3.3.8 The triple $(u, x, y) \in \mathcal{L}_2([0, T], \mathbb{R}^{m+n+m})$ is an *$\mathcal{L}_2$-solution* on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 if the following conditions hold

$$
\begin{aligned}
x(t) &= x_0 + \int_0^t [Ax(s) + Bu(s)] \, ds \\
y(t) &= Cx(t) + Du(t) \\
0 \leq u(t) &\perp y(t) \geq 0
\end{aligned}
$$

for almost all $t \in [0, T]$.

Similar to Proposition 3.3.7, we can state the following proposition.

Proposition 3.3.9 *Consider a matrix quadruple (A, B, C, D). Assume that $D + C(sI - A)^{-1}B$ is totally of index 1. If $(\cdot, \cdot, u, x, y)$ is a hybrid solution of LCS(A, B, C, D) with some initial state then for any $T > 0$ (u, x, y) is an $\mathcal{L}_2$-solution on $[0, T]$ of LCS(A, B, C, D) with the same initial state.*

The above proposition together with Proposition 3.3.7 and Theorem 3.3.3 implies existence of $\mathcal{L}_2$-solutions for LCS(A, B, C, D). By exploiting the passivity, we can establish uniqueness of $\mathcal{L}_2$-solutions as presented in the following theorem.

Theorem 3.3.10 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passive. Let $T > 0$ be given. Then, there exists a unique $\mathcal{L}_2$-solution on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 if and only if $Cx_0 \in \mathcal{Q}_D^*$.*

3.4 Passifiability by Pole Shifting

Consider a given system $\Sigma(A, B, C, D)$ and its pole-shifted version $\Sigma(A + \rho I, B, C, D)$. Note that if (u, x, y) is a solution of the former one then $e^{\rho\cdot}(u, x, y)$ is a solution of the latter one. By using this correspondence, we reach the following rather obvious fact.

Fact 3.4.1 If the triple (u, x, y) is a $\mathcal{PB}$-solution ($\mathcal{L}_2$-solution) on some interval of LCS(A, B, C, D) with some initial state then $e^{\rho\cdot}(u, x, y)$ is a $\mathcal{PB}$-solution ($\mathcal{L}_2$-solution) on the same interval of LCS$(A + \rho I, B, C, D)$ with the same initial state.

This fact opens the possibility of applying Theorem 3.3.10 to a class of nonpassive systems. Indeed, one can find ρ such that $\Sigma(A + \rho I, B, C, D)$ is passive although $\Sigma(A, B, C, D)$ is not. In what follows, we will investigate under what conditions $\Sigma(A, B, C, D)$ can be made passive by pole shifting.

Definition 3.4.2 A system $\Sigma(A, B, C, D)$ is said to be *passifiable by pole shifting* if there exists $\rho \in \mathbb{R}$ such that $\Sigma(A + \rho I, B, C, D)$ is passive.

Next, we give necessary and sufficient conditions for passifiability by pole shifting in the following theorem.

Theorem 3.4.3 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Let E be such that $\ker E = \{0\}$ and $\operatorname{im} E = \ker(D + D^\top)$. Then (A, B, C, D) is passifiable by pole shifting if and only if D is nonnegative definite and $E^\top CBE$ is symmetric positive definite.*

We are in a position to apply Theorem 3.3.3 and Theorem 3.3.10 to the class of systems that are passifiable by pole shifting as stated in the following corollary.

Corollary 3.4.4 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is minimal and $\operatorname{col}(B, D + D^\top)$ is of full column rank. Let E be such that $\ker E = \{0\}$ and $\operatorname{im} E = \ker(D + D^\top)$. Suppose that D is nonnegative and $E^\top CBE$ is symmetric positive definite. Then, the following statements are equivalent.*

1. *There exists a unique $\mathcal{PB}$-solution on $[0, \infty)$ of LCS(A, B, C, D) with the initial state x_0.*

2. *There exists a unique $\mathcal{L}_2$-solution on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 for any $T > 0$.*

3. *$Cx_0 \in \mathcal{Q}_D^*$.*

Moreover, the unique $\mathcal{PB}$-solution and $\mathcal{L}_2$-solution for a fixed initial state x_0 with $Cx_0 \in \mathcal{Q}_D^*$ are the same.

Especially, the case when D is positive definite is worth stating separately.

Corollary 3.4.5 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is minimal, and D is positive definite. Then, the following statements hold for any $T > 0$.*

1. *There exists a unique $\mathcal{PB}$-solution on $[0, T]$ of LCS(A, B, C, D) for all initial states.*

2. *There exists a unique $\mathcal{L}_2$-solution on $[0, T]$ of LCS(A, B, C, D) for all initial states.*

Moreover, the unique $\mathcal{PB}$-solution and $\mathcal{L}_2$-solution for a fixed initial state are the same.

3.5 Zeno Behavior

Ever since the linear complementarity systems were introduced (see [20,21]), Zeno behavior has been an interesting open problem. In this section, existence of accumulation points of the event times set will be investigated. In [10] the problem is addressed in a very general context of hybrid systems.

Our first result rules out left accumulation points for the systems that are passifiable by pole shifting.

Lemma 3.5.1 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Assume that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Then, there is no left Zeno solution of LCS(A, B, C, D).*

As an immediate consequence, we can state the following proposition.

Proposition 3.5.2 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Assume that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. If $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a hybrid solution of LCS(A, B, C, D) with some initial state then (u, x, y) is a $\mathcal{PB}$-solution on $[0, \infty)$ with the same initial state.*

As another implication of the previous Lemma, the class of passifiable systems for which the passifiability is invariant under *time-reversion* enjoy the non-Zenoness property. To show this, we need to point out the following fact.

Fact 3.5.3 If $(u, x, y) \in \mathcal{L}_2([0, T], \mathbb{R}^{m+n+m})$ is an $\mathcal{L}_2$-solution on $[0, T]$ of $\Sigma(A, B, C, D)$ then $\mathrm{rev}_{[0,T]}(u, x, y)$ is an $\mathcal{L}_2$-solution on $[0, T]$ of $\Sigma(-A, -B, C, D)$.

Lemma 3.5.4 *Consider a matrix quadruple (A, B, C, D). Suppose that (A, B, C) is minimal and D is positive definite. Then, there is no Zeno solution of LCS(A, B, C, D).*

Note that *Zeno states* (i.e., the states at the accumulation points) are well-defined due to the fact that x is continuous for a hybrid solution $(\cdot, \cdot, \cdot, x, \cdot)$. Intuitively, the most natural candidates for Zeno states are equilibrium states, in particular the zero state, of the system. The following theorem indicates that the zero state cannot be a Zeno state for a class of passifiable complementarity systems.

Theorem 3.5.5 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Assume that $\Sigma(A, B, C, D)$ is passifiable by pole shifting and there exists an index set $J \subseteq \overline{m}$ such that D_{JJ} is positive definite, $D_{J,\overline{m}\setminus J} = 0$, $D_{\overline{m}\setminus J,J} = 0$ and $D_{\overline{m}\setminus J,\overline{m}\setminus J}$ is skew-symmetric. Let $(\mathcal{E}, \cdot, \cdot, x, \cdot)$ be a nonredundant hybrid solution of LCS(A, B, C, D) with some initial state. If t^* is a right accumulation point of $\mathcal{E}$ then $x(t^*) \neq 0$.*

3.6 Nonregular Initial States

Our aim is to propose a method of *re-initialization* for nonregular initial states in terms of the stored energy of the underlying linear passive system. To do so, consider a passive system $\Sigma(A, B, C, D)$. Let the set of all positive definite matrices that generate a quadratic storage function, i.e., $\{K \mid x \mapsto x^\top Kx \text{ is a storage function for } \Sigma(A, B, C, D)\}$ be denoted by $\mathcal{K}$. It is known from [22] that $\mathcal{K}$ is convex and has a minimal and a maximal element (called *required supply* and *available storage*, respectively) with respect to the order induced by positive definiteness. We propose the jump rule $x_0 \mapsto x^+$ where x^+ is the solution of the following minimization problem

$$\text{minimize } \|x - x_0\|_K \text{ subject to } Cx \in \mathcal{Q}_D^*.$$

where $K \in \mathcal{K}$. In the next theorem, it will be shown that this proposal is justified in the sense that the above minimization problem admits unique solutions regardless of the choice of the storage function.

Theorem 3.6.1 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passive. The following statements hold.*

1. *The minimization problem*

$$\text{minimize } \|x - x_0\|_K \text{ subject to } Cx \in \mathcal{Q}_D^* \tag{3.5}$$

admits the same unique solution for all $K \in \mathcal{K}$.

2. *Let this unique solution be denoted by x^+. Then, there exists a unique $u^0 \in \mathcal{Q}_D$ such $x^+ - x_0 = Bu^0$.*

We call x^+ and u^0 enjoying the properties in the second item of the previous theorem as the *re-initialized state* and the *jump multiplier*, respectively.

The second statement of the above theorem tells us that the jump occurs along $\text{im}B$. In other words, the jump $x_0 \mapsto x^+$ can be represented by the effect of $u = u^0\delta$. This fact will be exploited to establish a solution concept that covers also nonregular initial states in a natural way. Before passing to the introduction a new solution concept, we need some preparations. The first thing we shall recall is the initial solution concept.

Definition 3.6.2 The triple $(u, x, y) \in \mathcal{B}_\delta^{m+n+m}$ is an *initial solution* of LCS(A, B, C, D) with the initial state x_0 if there exists an index set $K \subseteq \overline{m}$ such that

$$\begin{aligned} \dot{x} &= Ax + Bu + x_0\delta \\ y &= Cx + Du \\ y_K &= 0 \quad u_{\overline{m}\setminus K} = 0 \end{aligned}$$

holds in the distributional sense, and u and y are initially nonnegative.

Under some index and sign conditions, it can be shown that there exist unique initial solutions as stated in the following lemma.

Lemma 3.6.3 *Consider a matrix quadruple (A, B, C, D). Let $G(s)$ be the transfer matrix, i.e., $G(s) = D + C(sI - A)^{-1}B$. Suppose that $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then the following statements hold.*

1. *There exists a unique initial solution of LCS(A, B, C, D) with any initial state.*

2. *If $G(s)$ is totally of index 1 then the impulsive part of this unique initial solution is of the form $(\bar{u}, 0, D\bar{u})$ for some $\bar{u} \in \mathcal{Q}_D$.*

As a direct application of this lemma, we have the following corollary for the class of passifiable systems.

Corollary 3.6.4 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Then, there exists a unique initial solution of LCS(A, B, C, D) with any initial state. Moreover, the impulsive part of this unique initial solution is of the form $(\bar{u}, 0, D\bar{u})$ for some $\bar{u} \in \mathcal{Q}_D$.*

Now, we can state the following theorem which will lead to a solution concept that covers nonregular initial states as well.

Theorem 3.6.5 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passive. Let N be such that $\mathrm{pos}(N) = \mathcal{Q}_D$. Also let the re-initialized state x^+ be the unique solution of minimization problem (3.5) for the initial state x_0 and let the jump multiplier u^0 be the unique vector such that $x^+ = x_0 + Bu^0$. Then the following statements hold.*

1. *The jump multiplier $u^0 = N\bar{v}$ where $\bar{v}$ is the unique solution of the linear complementarity problem*

$$
\begin{aligned}
v &\geq 0 \\
N^\top C x_0 + N^\top CBNv &\geq 0 \\
v^\top (N^\top C x_0 + N^\top CBNv) &= 0.
\end{aligned}
$$

2. *The jump multiplier u^0 is the unique solution of the generalized linear complementarity problem*

$$
\begin{aligned}
z &\in \mathcal{Q}_D \\
Cx_0 + CBz &\in \mathcal{Q}_D^* \\
z^\top (Cx_0 + CBz) &= 0.
\end{aligned}
$$

3. *The jump multiplier is the common unique solution of the minimization problems*

$$
\text{minimize } (x_0 + Bz)^\top K (x_0 + Bz) \text{ subject to } z \in \mathcal{Q}_D
$$

   with $K \in \mathcal{K}$.

4. *The impulsive part of the initial solution of LCS(A, B, C, D) with the initial state x_0 is $(u^0\delta, 0, Du^0\delta)$.*

Although the jump rule proposed in Theorem 3.6.1 is applicable only for passive (or passifiable) systems, we can think of providing a jump rule for general low index systems via item 4. Indeed, such a proposition would be independent of passivity related concepts such as storage function. With this motivation, we introduce a new solution concept in what follows. We denote the set of all distributions $v = v_{imp} + v_{reg}$ where $v_{imp} \in \mathcal{D}'_0$ and v_{reg} is a piecewise Bohl function by $\mathcal{PB}_\delta$. They will be called *piecewise Bohl distributions*.

Definition 3.6.6 The triple $(u, x, y) \in \mathcal{PB}_\delta^{m+n+m}$ is a *distributional $\mathcal{PB}$-solution* on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 if the impulsive part of (u, x, y) coincides with the impulsive part of an initial solution of LCS(A, B, C, D) with the initial state x_0 and the regular part is a $\mathcal{PB}$-solution on $[0, T]$ of LCS(A, B, C, D) with the re-initialized state.

By merging Theorem 2.3.3 (for zero input) and Theorem 3.6.5, we reach the following well-posedness result for low index LCSs.

Theorem 3.6.7 *Consider a matrix quadruple (A, B, C, D). Suppose that $G(s) = D + C(sI - A)^{-1}B$ is totally of index 1 and $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then, there exists a unique distributional $\mathcal{PB}$-solution on $[0, \infty)$ of LCS(A, B, C, D) with all initial states.*

Earlier in this chapter, stronger (in the sense that the function space in which solutions live is larger) well-posedness results were presented for passifiable linear complementarity systems in Corollary 3.4.4. In a similar fashion to Theorem 3.6.7, a generalization of those results is possible.

Definition 3.6.8 The triple $(u, x, y) \in \mathcal{L}_2^\delta([0, T], \mathbb{R}^{m+n+m})$ is an *$\mathcal{L}_2^\delta$-solution* on $[0, T]$ of LCS(A, B, C, D) with the initial state x_0 if the impulsive part of (u, x, y) coincides with the impulsive part of an initial solution of LCS(A, B, C, D) with the initial state x_0 and the regular part is a $\mathcal{L}_2$-solution on $[0, T]$ of LCS(A, B, C, D) with the re-initialized state.

As a natural consequence, we have the following theorem.

Theorem 3.6.9 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Then, for any $T > 0$ there exists a unique $\mathcal{L}_2^\delta$-solution on $[0, T]$ of LCS(A, B, C, D) with all initial states.*

3.7 Conclusions

We have continued to deal with the well-posedness of linear complementarity systems. After proposing a more general solution concept than the one studied in the previous chapter, we have proven that the linear passive complementarity systems have unique solutions. Moreover, the characterization of regular initial states has been established. As a generalization of passivity, the notion of passifiability by pole shifting (PPS) was introduced. The necessary and sufficient conditions for a system to be PPS have been presented. We have extended all available well-posedness and regularity results to this new class. An interesting phenomenon that can occur in hybrid systems is the so-called Zeno behavior. The introduction of PPS systems makes it possible to show absence of Zeno behavior for LCS under some conditions. Finally, we have proposed a jump rule for nonregular initial states and extended the solution concept in such a way that nonregular initial states can be treated as well.

The $\mathcal{L}_2$-uniqueness of solutions is of considerable importance not only from a well-posedness viewpoint but also because its immediate consequences for the work in the second part of the thesis. So one of the most obvious open questions is whether the $\mathcal{L}_2$-uniqueness of solutions can be extended to the systems which are well-posed in the sense of $\mathcal{PB}$-solutions. Further study of Zeno behavior may have interesting impacts for the convergence issues. Indeed, absence of Zeno behavior would imply convergence in stronger senses for many cases.

3.8 Proofs

3.8.1 Some facts from matrix theory

In the sequel, we need the following technical lemmas.

Lemma 3.8.1 *Let $X \in \mathbb{R}^{p\times q}$ be given and $\mathcal{X}$ be defined by*

$$\mathcal{X} = \begin{pmatrix} I + XX^\top & X \\ X^\top & I \end{pmatrix}.$$

Then, the estimates $(2 + \|X\|^{\frac{1}{2}})I \geq \mathcal{X} \geq (2 + \|X\|^{\frac{1}{2}})^{-1}I$ hold.

Proof: Note that $\mathcal{X}$ is positive definite since

$$\mathcal{X} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix}\begin{pmatrix} I & 0 \\ X^\top & I \end{pmatrix}.$$

Let $\operatorname{rank}(X) = r$. Then, we get $\dim(\ker(X)) = q - r$ and $\dim(\ker(X^\top)) = p - r$. Note that $X^\top u = 0$ implies that

$$\begin{pmatrix} I + XX^\top & X \\ X^\top & I \end{pmatrix}\begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}$$

and $Xv = 0$ implies that

$$\begin{pmatrix} I + XX^\top & X \\ X^\top & I \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix}.$$

Therefore, $\mathcal{X}$ has at least $p + q - 2r$ eigenvalues at $\lambda = 1$. On the other hand, $XX^\top$ has r nonzero eigenvalues and if $\mu \neq 0$ is its eigenvalue with an eigenvector u then

$$\begin{pmatrix} I + XX^\top & X \\ X^\top & I \end{pmatrix} \begin{pmatrix} (\lambda_-^\mu - 1)u \\ X^\top u \end{pmatrix} = \lambda_-^\mu \begin{pmatrix} (\lambda_-^\mu - 1)u \\ X^\top u \end{pmatrix} \tag{3.6}$$

and

$$\begin{pmatrix} I + XX^\top & X \\ X^\top & I \end{pmatrix} \begin{pmatrix} (\lambda_+^\mu - 1)u \\ X^\top u \end{pmatrix} = \lambda_+^\mu \begin{pmatrix} (\lambda_+^\mu - 1)u \\ X^\top u \end{pmatrix} \tag{3.7}$$

where $\lambda_-^\mu = \frac{2+\mu}{2} - \frac{\sqrt{(2+\mu)^2-4}}{2}$ and $\lambda_+^\mu = \frac{2+\mu}{2} + \frac{\sqrt{(2+\mu)^2-4}}{2}$. Note that $\lambda_-^\mu \leq 1 \leq \lambda_+^\mu$ and $\lambda_-^\mu \lambda_+^\mu = 1$. Since the function $\mu \mapsto \lambda_+^\mu$ is an increasing function on $[0, \infty)$, it follows from (3.6) and (3.7) that $\lambda_{\max}(\mathcal{X}) = \lambda_+^{\mu_{\max}}$ where $\mu_{\max} = \lambda_{\max}(XX^\top)$ and $\lambda_{\min}(\mathcal{X}) = (\lambda_{\max}(\mathcal{X}))^{-1}$. Therefore, we get

$$\lambda_+^{\mu_{\max}} I \geq \mathcal{X} \geq (\lambda_+^{\mu_{\max}})^{-1} I.$$

Note that $2 + \mu \geq \lambda_+^\mu$ whenever $\mu \geq 0$ and $\mu_{\max} = \|X\|^{\frac{1}{2}}$. Consequently, we get

$$(2 + \|X\|^{\frac{1}{2}}) I \geq \mathcal{X} \geq (2 + \|X\|^{\frac{1}{2}})^{-1} I.$$

■

Lemma 3.8.2 *Consider matrices A, B, and C such that $A = A^\top \in \mathbb{R}^{p\times p}$, $B \in \mathbb{R}^{p\times q}$, and $C = C^\top \in \mathbb{R}^{q\times q}$ and C is invertible. The following estimates hold:*

$$(2 + \|BC^{-1}\|^{\frac{1}{2}}) \max(\lambda_{\max}(D), \lambda_{\max}(C))\, I \geq \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \geq \frac{\min(\lambda_{\min}(D), \lambda_{\min}(C))}{2 + \|BC^{-1}\|^{\frac{1}{2}}} I$$

where the Schur complement $D = A - BC^{-1}B^\top$.

Proof: Note that

$$\begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} = \begin{pmatrix} I & BC^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A - BC^{-1}B^\top & 0 \\ 0 & C \end{pmatrix} \begin{pmatrix} I & 0 \\ C^{-1}B^\top & I \end{pmatrix}$$

and

$$\max(\lambda_{\max}(D), \lambda_{\max}(C))\, I \geq \begin{pmatrix} A - BC^{-1}B^\top & 0 \\ 0 & C \end{pmatrix} \geq \min(\lambda_{\min}(D), \lambda_{\min}(C))\, I.$$

The rest follows from Lemma 3.8.1 by taking $X = BC^{-1}$. ■

Lemma 3.8.3 *Let $M \in \mathbb{R}^{p\times p}$ and $N \in \mathbb{R}^{p\times p}$ be given. Suppose that M is nonnegative definite and the following implication holds*

$$x \neq 0,\, x^\top M x = 0 \Rightarrow x^\top N x > 0. \tag{3.8}$$

Then, there exists $\mu > 0$ such that $M + \epsilon N \geq \epsilon\, \mu\, I$ for all sufficiently small ϵ.

Proof: Let $\bar{M}$ and $\bar{N}$ denote the symmetric parts of M and N, i.e., $\bar{M} = \frac{1}{2}(M + M^\top)$ and $\bar{N} = \frac{1}{2}(N + N^\top)$. Let $Q \in \mathbb{R}^{m\times q}$ be such that $\operatorname{im} Q = \ker \bar{M}$ and $\ker Q = \{0\}$. Take any $P \in \mathbb{R}^{m\times(m-q)}$ such that $\begin{pmatrix} P & Q \end{pmatrix}$ is nonsingular. Note that

$$\begin{pmatrix} P^\top \\ Q^\top \end{pmatrix} (\bar{M} + \epsilon \bar{N}) \begin{pmatrix} P & Q \end{pmatrix} = \begin{pmatrix} P^\top \bar{M} P + \epsilon P^\top \bar{N} P & \epsilon P^\top \bar{N} Q \\ \epsilon Q^\top \bar{N} P & \epsilon Q^\top \bar{N} Q \end{pmatrix}$$

It follows from Lemma 3.8.2 that

$$\begin{pmatrix} P^\top \\ Q^\top \end{pmatrix} (\bar{M} + \epsilon \bar{N}) \begin{pmatrix} P & Q \end{pmatrix} \geq \frac{\min(\lambda_{\min}(X + \epsilon Y_{11} - \epsilon Y_{12} Y_{22}^{-1} Y_{12}^\top), \lambda_{\min}(\epsilon Y_{22}))}{2 + \|Y_{12} Y_{22}^{-1}\|^{\frac{1}{2}}} I \tag{3.9}$$

where $X := P^\top \bar{M} P$, $Y_{11} := P^\top \bar{N} P$, $Y_{12} := P^\top \bar{N} Q$ and $Y_{22} := Q^\top \bar{N} Q$. Note that X is positive definite due to the definition of P and Q and Y_{22} is positive definite due to the implication (3.8). Therefore, it follows from (3.9) that there exists $\mu > 0$ such that $\bar{M} + \epsilon \bar{N} \geq \epsilon\, \mu\, I$ for all sufficiently small ϵ. ■

3.8.2 Some implications of passivity

For ease of reference we list the following well-known facts.

Lemma 3.8.4 *Let $M = M^\top \in \mathbb{R}^{m\times m}$ be nonnegative definite. The following statements hold.*

1. *$N^\top M N = 0 \Rightarrow MN = 0$.*

2. *For any index set $J \subseteq \overline{m}$, $v^\top M_{JJ} v = 0 \Rightarrow M_{\bullet J} v = 0$.*

Proof:

1: Evident.

2: Let the index set $J \subseteq \overline{m}$ and the vector v be such that $v^\top M_{JJ} v = 0$. Clearly, we have

$$\begin{pmatrix} v \\ 0 \end{pmatrix}^\top \begin{pmatrix} M_{JJ} & M_{J,\overline{m}\setminus J} \\ M_{\overline{m}\setminus J,J} & M_{\overline{m}\setminus J,\overline{m}\setminus J} \end{pmatrix} \begin{pmatrix} v \\ 0 \end{pmatrix} = 0.$$

Hence, item 1 implies that

$$\begin{pmatrix} M_{JJ} & M_{J,\overline{m}\setminus J} \\ M_{\overline{m}\setminus J,J} & M_{\overline{m}\setminus J,\overline{m}\setminus J} \end{pmatrix} \begin{pmatrix} v \\ 0 \end{pmatrix} = 0.$$

Equivalently, $M_{\bullet J} v = 0$. ■

Passivity of a system has some useful implications for the *subsystems*. In what follows, we collect all such implications that will be employed later on.

Lemma 3.8.5 *Consider a matrix quadruple (A, B, C, D) satisfying Assumption 3.3.2. Let the matrices P^J and Q^J be such that $\ker P^J = \ker Q^J = \{0\}$, $\operatorname{im} Q^J = \ker(D_{JJ} + D_{JJ}^\top)$ and $\operatorname{im} P^J \oplus \operatorname{im} Q^J = \mathbb{R}^{|J|}$ for each index set $J \subseteq \overline{m}$. If the system $\Sigma(A, B, C, D)$ is passive then the following statements hold for each $J \subseteq \overline{m}$.*

1. *D_{JJ} is nonnegative definite.*
2. *$(P^J)^\top D_{JJ} P^J$ is positive definite.*
3. *$(Q^J)^\top C_{J\bullet} B_{\bullet J} Q^J$ is symmetric positive definite.*
4. *$D_{JJ} + C_{J\bullet} B_{\bullet J} \sigma^{-1}$ is positive definite for all sufficiently large σ.*
5. *$D_{JJ} + C_{J\bullet} B_{\bullet J} s^{-1}$ is of index 1.*

Proof: Since the system $\Sigma(A, B, C, D)$ is passive and (A, B, C) is minimal, Theorem 3.2.2 implies that the system of linear matrix inequalities

$$K = K^\top > 0 \text{ and } \begin{bmatrix} A^\top K + KA & KB - C^\top \\ B^\top K - C & -(D + D^\top) \end{bmatrix} \leq 0 \tag{3.10}$$

is feasible. It follows that for each index set $J \subseteq \overline{m}$ we have

$$\begin{bmatrix} A^\top K + KA & KB_{\bullet J} - C_{J\bullet}^\top \\ B_{\bullet J}^\top K - C_{J\bullet} & -(D_{JJ} + D_{JJ}^\top) \end{bmatrix} \leq 0. \tag{3.11}$$

1: Evident from (3.11).

2: It follows from 1 that $(P^J)^\top D_{JJ} P^J$ is nonnegative definite. Let v be such that $v^\top (P^J)^\top D_{JJ} P^J v = 0$. Hence, we have $v^\top (P^J)^\top (D_{JJ} + D_{JJ}^\top) P^J v = 0$. It follows from Lemma 3.8.4 item 1 that $P^J v \in \ker(D_{JJ} + D_{JJ}^\top) = \operatorname{im} Q^J$. Thus, $P^J v \in \operatorname{im} P^J \cap \operatorname{im} Q^J$. By the hypothesis, $v = 0$. Consequently, $(P^J)^\top D_{JJ} P^J$ is positive definite.

3: For any real number $\alpha \in \mathbb{R}$ and matrix $M^J \in \mathbb{R}^{|J| \times |J|}$, we have

$$\begin{aligned} 0 &\geq \begin{pmatrix} \alpha M^J \\ Q^J \end{pmatrix}^\top \begin{pmatrix} A^\top K + KA & KB_{\bullet J} - C_{J\bullet}^\top \\ B_{\bullet J}^\top K - C_{J\bullet} & -(D_{JJ} + D_{JJ}^\top) \end{pmatrix} \begin{pmatrix} \alpha M^J \\ Q^J \end{pmatrix} \\ &\geq \alpha^2 (M^J)^\top (A^\top K + KA) M^J + \alpha (M^J)^\top (KB_{\bullet J} - C_{J\bullet}^\top) Q^J + \alpha (Q^J)^\top (B_{\bullet J}^\top K - C_{J\bullet}) M^J. \end{aligned}$$

The absence of constant term in the above nonpositive quadratic form implies that the factor of α in the above equation must be zero for all M^J. In particular, the choice $M^J = (KB_{\bullet J} - C_{J\bullet}^\top) Q^J$ results in $(Q^J)^\top (B_{\bullet J}^\top K - C_{J\bullet}) = 0$. Hence, we have

$$(Q^J)^\top C_{J\bullet} B_{\bullet J} Q^J = (Q^J)^\top B_{\bullet J}^\top K B_{\bullet J} Q^J. \tag{3.12}$$

Since K is positive definite, the right hand side of the above equation is (at least) nonnegative definite. Let v be such that $v^\top (Q^J)^\top (B_{\bullet J})^\top K B_{\bullet J} Q^J v = 0$. Clearly, $B_{\bullet J} Q^J v = 0$. Note that $v^\top (Q^J)^\top (D_{JJ} + D_{JJ}^\top) Q^J v = 0$ implies from Lemma 3.8.4 item 2 that $(D_{\bullet J} + (D^\top)_{\bullet J}) Q^J v = 0$. Thus, we have

$$\begin{pmatrix} B_{\bullet J} & B_{\bullet J^c} \\ D_{\bullet J} + (D^\top)_{\bullet J} & D_{\bullet J^c} + (D^\top)_{\bullet J^c} \end{pmatrix} \begin{pmatrix} Q^J v \\ 0 \end{pmatrix} = 0.$$

It follows from the hypothesis that $\operatorname{col}(B, D + D^\top)$ is of full column rank that $Q^J v = 0$. Since $\ker Q^J = \{0\}$, v must be zero. Consequently, $(Q^J)^\top B_{\bullet J}^\top K B_{\bullet J} Q^J$ is positive definite and so is $(Q^J)^\top C_{J\bullet} B_{\bullet J} Q^J$ due to (3.12).

4: The previous item implies that the implication ($u \neq 0$ and $u^\top D_{JJ} u = 0$) $\Rightarrow$ $u^\top C_{J\bullet} B_{\bullet J} u > 0$ holds. It follows from Lemma 3.8.3 that there exists $\mu > 0$ such that

$$D_{JJ} + C_{J\bullet} B_{\bullet J} \sigma^{-1} \geq \mu \sigma^{-1} I \tag{3.13}$$

for all sufficiently large σ. Therefore, $D_{JJ} + C_{J\bullet} B_{\bullet J} \sigma^{-1}$ is positive definite for all sufficiently large σ.

5: It follows from (3.13) that $D_{JJ} + C_{J\bullet}B_{\bullet J}s^{-1}$ is invertible as a rational matrix and $s^{-1}(D_{JJ} + C_{J\bullet}B_{\bullet J}s^{-1})^{-1}$ is proper. Hence, $D_{JJ} + C_{J\bullet}B_{\bullet J}s^{-1}$ is of index 1. ■

3.8.3 Proofs for Section 3.3

Proof of Theorem 3.3.3: From Lemma 3.8.5 items 5 and 4, we know $G(s) = D + C(sI - A)^{-1}B$ is totally of index 1 and $G(\sigma)$ is positive definite and hence a $\mathcal{P}$-matrix for all sufficiently large σ. Besides, $K(D) = \mathcal{Q}_D^*$ due to Lemma 3.8.5 item 1 since D is nonnegative definite and hence is a copositive-plus matrix. Evidently, the rest follows by applying Theorem 2.3.3 for the zero input. ■

To prove Proposition 3.3.7, we need the following technical lemma.

Lemma 3.8.6 *Let $\mathcal{I}$ be a set of nonoverlapping intervals of real numbers with positive length. Then, $\mathcal{I}$ is a countable set.*

Proof: Each $I \in \mathcal{I}$ must contain a rational number. Therefore, the cardinality of the set $\mathcal{I}$ cannot exceed the cardinality of the rational numbers. Clearly, this means that $\mathcal{I}$ is a countable set. ■

Proof of Proposition 3.3.7: Assume that (u, x, y) is a $\mathcal{PB}$-solution on $[0, \infty)$ of LCS(A, B, C, D) with some initial state x_0. Let the mapping β be as defined in 1.2.1. Define $\mathcal{E}^* = \{t + \beta((u, x, y), t) \mid t \in \mathbb{R}_+\}$. Take $\mathcal{E} = \overline{\mathcal{E}^* \cup \{0\}}$. We first prove some properties of the set $\mathcal{E}$.

i. $\mathcal{E}$ is closed and $0 \in \mathcal{E}$.

ii. Consider the set $\tau_{\mathcal{E}} \cup \mathcal{E}$. It consists of nonoverlapping intervals of $\mathbb{R}$. It follows from Lemma 3.8.6 that it is a countable set and hence so is $\mathcal{E}$.

iii. Suppose that t^{left} is a left accumulation point of $\mathcal{E}$. By definition, $t^{\text{left}} + \beta((u, x, y), t^{\text{left}}) = t + \beta((u, x, y), t)$ whenever $t^{\text{left}} \leq t < t^{\text{left}} + \beta((u, x, y), t^{\text{left}})$. In other words, $(t^{\text{left}}, t^{\text{left}} + \beta((u, x, y), t^{\text{left}}) \cap \mathcal{E} = \emptyset$. However, this contradicts the fact that t^{left} is a left accumulation point of $\mathcal{E}$.

Clearly, i.-iii. yield that $\mathcal{E}$ is a left Zeno free admissible event times set. Let $\tau_{\mathcal{E}}$ be the associated collection of intervals. Let $(t', t'') \in \tau_{\mathcal{E}}$. Define $\operatorname{col}(\bar{u}, \bar{x}, \bar{y}) = \alpha((u, x, y), t')$.

Since (u, x, y) is a $\mathcal{PB}$-solution, we get

$$\begin{aligned}\bar{x}(t) &= x(t') + \int_{t'}^{t} [A\bar{x}(s) + B\bar{u}(s)]\ ds \\ \bar{y}(t) &= C\bar{x}(t) + D\bar{u}(t) \\ 0 &\leq \bar{u}(t) \perp \bar{y}(t) \geq 0.\end{aligned}$$

for all $t \in (t', t'')$. Then, $(\bar{u}, \bar{x}, \bar{y})$ satisfies

$$\begin{aligned}\dot{\bar{x}} &= A\bar{x} + B\bar{u},\ \bar{x}(0) = x(t') \\ \bar{y} &= C\bar{x} + D\bar{u} \\ \bar{y}_K &= 0 \\ \bar{u}_{\overline{m}\setminus K} &= 0\end{aligned}$$

for some $K \subseteq \overline{m}$ since they are Bohl functions. Define $\mathcal{S}((t', t'')) = K$. It can be easily verified that $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a hybrid solution of LCS(A, B, C, D) with the initial state x_0. It remains to show that it is nonredundant. Suppose, on the contrary, that it is redundant, i.e., there exist $t \in \mathcal{E}$ and t', t'' with $t' < t < t''$ such that (u, x, y) is analytic on (t', t''). Define $t_1 = t' + \beta((u, x, y), t')$. Clearly, we have

$$t' < t_1 \leq t < t''. \tag{3.14}$$

It is known that $(u, x, y)|_{[t', t_1)} = w|_{[0, \beta((u,x,y),t'))}$ where $w = \alpha((u, x, y), t')$. Since w is a Bohl function and hence analytic, we have even $(u, x, y)|_{[t', t'')} = w|_{[0, t''-t')}$. Therefore, $t' - t_1 = \beta((u, x, y), t') \geq t'' - t'$ which contradicts (3.14). ∎

Proof of Proposition 3.3.9: Assume that $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a hybrid solution of LCS(A, B, C, D) with some initial state x_0. Let $T > 0$. It follows from the continuity of x that it is bounded on the interval $[0, T]$. As a hybrid solution $(\mathcal{E}, \mathcal{S}, u, x, y)$ is such that for each $\tau \in \tau_{\mathcal{E}}$ the conditions

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \\ y_{\mathcal{S}(\tau)}(t) = 0 &\quad u_{\overline{m}\setminus\mathcal{S}(\tau)}(t) = 0 \\ u_{\mathcal{S}(\tau)}(t) \geq 0 &\quad y_{\overline{m}\setminus\mathcal{S}(\tau)}(t) \geq 0\end{aligned}$$

hold for all $t \in \tau \cap [0, T]$. Lemma 2.5.9 implies that (by taking $w = 0$) there exists $F^{\mathcal{S}(\tau)}$ such that $u(t) = F^{\mathcal{S}(\tau)}x(t)$ and $y(t) = (C + DF^{\mathcal{S}(\tau)})x(t)$ for all $t \in \tau$. Hence, both u and y

are bounded on $\tau' := (\underset{\tau \in \tau_{\mathcal{E}}}{\cup} \tau) \cap [0,T]$ since $\mathcal{S}(\tau)$ assumes values in the finite set $2^{\overline{m}}$. This means that they are essentially bounded on the interval $[0,T]$ since $\tau' = ([0,T] \cap \mathcal{E})^c$ and $\mathcal{E}$ is a countable set hence has measure zero. Boundedness of (u,x,y) on $[0,T]$ immediately implies $(u,x,y) \in \mathcal{L}_2([0,T], \mathbb{R}^{m+n+m})$. It is not hard to see that (u,x,y) is an $\mathcal{L}_2$-solution on $[0,T]$ of LCS(A,B,C,D) with the initial state x_0 for any $T > 0$. ■

Proof of Theorem 3.3.10:

if: Let $Cx_0 \in \mathcal{Q}_D^*$. Theorem 3.3.3, and Propositions 3.3.7 and 3.3.9 imply together that on any interval $[0,T]$ there exists an $\mathcal{L}_2$-solution of LCS(A,B,C,D) with the initial state x_0. Suppose that (u^i, x^i, y^i) for $i = 1,2$ are two different $\mathcal{L}_2$-solutions of LCS(A,B,C,D) with the same initial state. Then, $(u^1 - u^2, x^1 - x^2, y^1 - y^2)$ is an $\mathcal{L}_2$-solution of $\Sigma(A,B,C,D)$ with the initial state zero. Since $\Sigma(A,B,C,D)$ is passive and (A,B,C,D) satisfies Assumption 3.3.2, Theorem 3.2.2 implies that there exists $K = K^\top > 0$ such that

$$\int_0^t [u^1(s) - u^2(s)]^\top [y^1(s) - y^2(s)]\, ds \geq [x^1(t) - x^2(t)]^\top K [x^1(t) - x^2(t)] \tag{3.15}$$

for all $t \in [0,T]$. Since (u^i, x^i, y^i) are $\mathcal{L}_2$-solutions of LCS(A,B,C,D), we have $u^i(t) \geq 0$ and $y^i(t) \geq 0$ for all $t \in [0,T]$, and $\int_0^t (u^i(s))^\top y^i(s)\, ds = 0$. Therefore, the left hand side of (3.15) is nonpositive. It follows from the positive definiteness of K that

$$x^1(t) - x^2(t) = 0 \tag{3.16}$$

for all $t \in [0,T]$. Then, we get

$$B(u^1(t) - u^2(t)) = 0 \tag{3.17}$$

$$y^1(t) - y^2(t) = D(u^1(t) - u^2(t)) \tag{3.18}$$

for all $t \in [0,T]$ by the definition of $\mathcal{L}_2$-solution of a LCS. Left multiplying (3.18) by $(u^1(t) - u^2(t))^\top$ results in $(u^1(t) - u^2(t))^\top D(u^1(t) - u^2(t)) \leq 0$ for almost all $t \in [0,T]$. However, D is nonnegative definite due to Lemma 3.8.5 item 1. Therefore the above inequality holds as an equality. Hence, we get for almost all $t \in [0,T]$, $(u^1(t) - u^2(t))^\top (D + D^\top)(u^1(t) - u^2(t)) = 0$ which immediately gives

$$(D + D^\top)(u^1(t) - u^2(t)) = 0 \tag{3.19}$$

due to Lemma 3.8.4 item 1. The equations (3.17) and (3.19), together with Assumption 3.3.2, imply that $u^1(t) - u^2(t) = 0$ for almost all $t \in [0,T]$. It follows from (3.16) that $y^1(t) - y^2(t) = 0$ for almost all $t \in [0,T]$.

only if: Let (u, x, y) be an $\mathcal{L}_2$-solution on some interval $[0, T]$ of LCS(A, B, C, D) with some initial state x_0. Suppose that $Cx_0 \notin \mathcal{Q}_D^*$. Since $D \geq 0$ due to Lemma 3.8.5 item 1, it follows from Lemma 1.2.4 item 2 that $K(D) = \mathcal{Q}_D^*$. We know from Fact 1.2.2 that $K(D)$ is closed. Consequently, the fact that $Cx_0 \notin \mathcal{Q}_D^*$ implies together with the continuity of x that for some $\epsilon > 0$, $Cx(t) \notin \mathcal{Q}_D^*$ whenever $t \in [0, \epsilon]$. In other words, LCP$(Cx(t), D)$ is not solvable on $t \in [0, \epsilon]$. A contradiction follows from Definition 3.3.8. ■

3.8.4 Proofs for Section 3.4

We begin with a technical lemma which be employed later on.

Lemma 3.8.7 *Let A, $B \in \mathbb{R}^{m \times n}$ and let A be of full row rank. Then, there exists a symmetric positive definite matrix X such that $AX = B$ if and only if $BA^\top$ is symmetric positive definite.*

Proof:

only if: Postmultiplying $AX = B$ by $A^\top$, we get $AXA^\top = BA^\top$. Since $X = X^\top > 0$, $BA^\top = AB^\top > 0$.

if: Note that A can be written as $A = \begin{bmatrix} I & 0 \end{bmatrix} V$ for some nonsingular $V \in \mathbb{R}^{n \times n}$. Postmultiplying both sides of $AX = B$ by $V^\top$ and defining $Y := VXV^\top$, we get $\begin{bmatrix} I & 0 \end{bmatrix} Y = BV^\top$. Clearly, finding a solution to the latter equation with $Y = Y^\top > 0$ is equivalent to finding a solution to $AX = B$ with $X = X^\top > 0$. Let Y and $BV^\top$ be partitioned as follows:

$$Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \qquad BV^\top = \begin{bmatrix} B_1 & B_2 \end{bmatrix}.$$

To satisfy $\begin{bmatrix} I & 0 \end{bmatrix} Y = BV^\top$, we can take $Y_{12} = B_2$ and $Y_{11} = B_1 = BV^\top \begin{bmatrix} I & 0 \end{bmatrix}^\top = BA^\top$. Hence, by the hypothesis $Y_{11} = Y_{11}^\top > 0$. It remains to determine Y_{21} and Y_{22} in such a way that $Y = Y^\top > 0$. Choose $Y_{21} = Y_{12}^\top$ and $Y_{22} = I + Y_{12}^\top Y_{11}^{-1} Y_{12}$. Then, it follows from

$$Y = \begin{bmatrix} I & 0 \\ Y_{12}^\top Y_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} - Y_{12}^\top Y_{11}^{-1} Y_{12} \end{bmatrix} \begin{bmatrix} I & Y_{11}^{-1} Y_{12} \\ 0 & I \end{bmatrix}$$

that $Y = Y^\top > 0$. ■

Proof of Theorem 3.4.3:

if: Assumption 3.3.2 guarantees that BE is of full column rank. Then, the equation $E^\top C = E^\top B^\top K$ has a symmetric positive definite solution K according to Lemma 3.8.7. Define $\mu = \lambda_{max}(K)$. Let F be such that $\ker F = \{0\}$ and $\operatorname{im} E \oplus \operatorname{im} F = \mathbb{R}^m$. It follows from Lemma 3.8.5 item 2 that $F^\top DF$ is positive definite. Define $\alpha = \frac{1}{2\mu}\lambda_{max}(A^\top K + KA)$, $\beta = \frac{1}{2\mu}\|KBF - C^\top F\|$ and $\gamma = -\frac{1}{2\mu}\lambda_{min}(F^\top(D+D^\top)F)$. Note that $\gamma < 0$. Take $\rho \leq \frac{\beta^2}{\gamma} - \alpha$ and note that $\left[\begin{smallmatrix}\alpha+\rho & \beta \\ \beta & \gamma\end{smallmatrix}\right]$ is nonpositive definite. It can be verified that $(A+\rho I, B, C, D)$ is passive with the storage function $V(x) = x^\top Kx$. Indeed,

$$\begin{aligned}
&\begin{bmatrix} x \\ u \end{bmatrix}^\top \begin{pmatrix} (A+\rho I)^\top K + K(A+\rho I) & KB - C^\top \\ B^\top K - C & -(D+D^\top) \end{pmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
&= x^\top(A^\top K + KA)x + 2\rho x^\top Kx + 2x^\top(KB - C^\top)u - u^\top(D+D^\top)u \\
&= x^\top(A^\top K + KA)x + 2\rho x^\top Kx + 2x^\top(KB - C^\top)Fu_f - u_f^\top F^\top(D+D^\top)Fu_f
\end{aligned}$$

where $u = Eu_e + Fu_f$. From the Rayleigh-Ritz (see e.g. [11, Theorem 5.2.2.2]) and Cauchy-Schwarz inequalities, we get

$$\begin{aligned}
&\begin{bmatrix} x \\ u \end{bmatrix}^\top \begin{pmatrix} (A+\rho I)^\top K + K(A+\rho I) & KB - C^\top \\ B^\top K - C & -(D+D^\top) \end{pmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
&\quad \leq \lambda_{max}(A^\top K + KA)\|x\|^2 + 2\rho\lambda_{max}(K)\|x\|^2 + 2\|KBF - C^\top F\|\|u_f\|\|x\| \\
&\qquad - \lambda_{min}(F^\top(D+D^\top)F)\|u_f\|^2 \\
&\quad \leq 2\mu \begin{bmatrix} \|x\| \\ \|u_f\| \end{bmatrix}^\top \begin{bmatrix} \alpha+\rho & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} \|x\| \\ \|u_f\| \end{bmatrix} \leq 0.
\end{aligned}$$

Since K is positive definite and minimality of (A, B, C) implies that $(A+\rho I, B, C)$ is also minimal, we can conclude that $(A+\rho I, B, C, D)$ is passive due to Lemma 3.2.2.

only if: If (A, B, C, D) is passifiable by pole shifting then there exists a ρ such that $\Sigma(A+\rho I, B, C, D)$ is passive. Then, it follows from Lemma 3.8.5 items 1 and 3 that D is nonnegative definite and $E^\top CBE$ is symmetric positive definite. ■

Proof of Corollary 3.4.4: It follows from Theorems 3.3.3, 3.3.10 and Fact 3.4.1. ■

Proof of Corollary 3.4.5: Note that the matrix quadruple satisfies Assumption 3.3.2 and $\Sigma(A, B, C, D)$ is passifiable by pole shifting since D is positive definite. Note also that $\mathcal{Q}_D = \{0\}$ which implies $\mathcal{Q}_D^* = \mathbb{R}^m$. Then, we get the desired result by applying Corollary 3.4.4 . ■

3.8.5 Proofs for Section 3.5

Proof of Lemma 3.5.1: Suppose that $(\mathcal{E}^{\text{left}}, \mathcal{S}^{\text{left}}, u^{\text{left}}, x^{\text{left}}, y^{\text{left}})$ is a left Zeno nonredundant hybrid solution of LCS(A, B, C, D) with some initial state x_0. From Proposition 3.3.9, it is known that $(u^{\text{left}}, x^{\text{left}}, y^{\text{left}})|_{[0,T]}$ is an $\mathcal{L}_2$-solution on $[0, T]$ with the same initial state for any $T > 0$. Then, Corollary 3.4.4 implies that there exists a $\mathcal{PB}$-solution, say (u, x, y), on $[0, \infty)$ with the same initial state. It follows from Proposition 3.3.7 that there exists a left Zeno free event times set $\mathcal{E}$ and $\mathcal{S}$ such that $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a nonredundant hybrid solution with the same initial state. Due to Proposition 3.3.9, $(u, x, y)|_{[0,T]}$ is an $\mathcal{L}_2$-solution on $[0, T]$ with the same initial state for any $T > 0$. Therefore, we get $(u^{\text{left}}, x^{\text{left}}, y^{\text{left}})|_{[0,T]} = (u, x, y)|_{[0,T]}$ for all $T > 0$ from Theorem 3.3.10. Let T^{left} be a left accumulation point of $\mathcal{E}^{\text{left}}$. Since (u, x, y) are piecewise Bohl functions, they are analytic on $(T^{\text{left}}, T^{\text{left}} + \beta(u, y))$. But, this contradicts the nonredundancy. ■

Proof of Lemma 3.5.4: Since $D > 0$, $\text{col}(B, D + D^\top)$ is of full column rank. Hence, Assumption 3.3.2 is satisfied by the hypotheses. Furthermore, $D > 0$ implies that both $\Sigma(A, B, C, D)$ and $\Sigma(-A, -B, C, D)$ are passifiable by pole shifting due to Theorem 3.4.3. Let $(\mathcal{E}^1, \mathcal{S}^1, u^1, x^1, y^1)$ be a hybrid solution of LCS(A, B, C, D) with some initial state x_0. Proposition 3.3.9 and Corollary 3.4.4 imply that $Cx_0 \in K(D)$. Therefore, LCS$(-A, -B, C, D)$ admits a $\mathcal{PB}$-solution, say (u^2, x^2, y^2), on $[0, \infty)$ with the initial state x_0 due to Corollary 3.4.4. According to Proposition 3.3.7, there exist $\mathcal{E}^2$ and $\mathcal{S}^2$ such that $(\mathcal{E}^2, \mathcal{S}^2, u^2, x^2, y^2)$ is a hybrid solution of LCS$(-A, -B, C, D)$ with the initial state x_0. On the other hand, we know from Lemma 3.5.1 that $\mathcal{E}^1$ is left Zeno free. Suppose that it is not right Zeno free, i.e., there exists a right accumulation point of $\mathcal{E}^1$. Let $t^* \in \mathcal{E}^1$ be such a point. Define $\mathcal{E} = (t^* - (\mathcal{E}^1 \cap [0, t^*])) \cup (t^* + \mathcal{E}^2)$ where $t + \mathcal{E}$ denotes the set $\{t + t' \mid t' \in \mathcal{E}\}$. Clearly, $\mathcal{E}$ is an admissible event times set and $\tau_{\mathcal{E}} = \tau_{t^* - (\mathcal{E}^1 \cap [0,t^*])} \cup \tau_{t^* + \mathcal{E}^2}$. Define $\mathcal{S}$ and (u, x, y) by

$$S(\tau) = \begin{cases} \mathcal{S}^1(t^* + \tau) & \text{if } \tau \in \tau_{t^* - (\mathcal{E}^1 \cap [0,t^*])} \\ \mathcal{S}^2(t^* - \tau) & \text{if } \tau \in \tau_{t^* + \mathcal{E}^2} \end{cases}$$

$$(u, x, y)|_{[0,t^*]} = \text{rev}_{[0,t^*]}(u^1, x^1, y^1) \tag{3.20}$$

$$(u, x, y)|_{(t^*,\infty)} = (u^2, x^2, y^2). \tag{3.21}$$

It can be verified that $(\mathcal{E}, \mathcal{S}, u, x, y)$ is a hybrid solution of LCS$(-A, -B, C, D)$ with the initial state $x^1(t^*)$. Lemma 3.5.1 reveals that $\mathcal{E}$ is left Zeno free. By construction t^* is a left accumulation point of $\mathcal{E}$. Therefore, we reach a contradiction. ■

Proof of Theorem 3.5.5: Let $(\mathcal{E}, \mathcal{S}, u, x, y)$ be a nonredundant hybrid solution of LCS$(A,$

$B, C, D)$ with some initial state such that t^* is a right accumulation point of $\mathcal{E}$ and $x(t^*) = 0$. According to Proposition 3.3.9, (u, x, y) is an $\mathcal{L}_2$-solution on $[0, t^*]$ of LCS(A, B, C, D) with the same initial state and hence

$$u_i(t)y_i(t) = 0 \tag{3.22}$$

for almost all t and for all $i \in \overline{m}$. Define $(\bar{u}, \bar{x}, \bar{y}) = \mathrm{rev}_{[0,t^*]}(u, x, y)$. From Fact 3.5.3, we know that $(\bar{u}, \bar{x}, \bar{y})$ is an $\mathcal{L}_2$-solution on $[0, t^*]$ of $\Sigma(-A, -B, C, D)$ with the initial state $\bar{x}(0) = x(t^*) = 0$. Define $z_J = \bar{y}_J$ and $z_{J^c} = -\bar{y}_{J^c}$ where $J^c = \overline{m} \setminus J$. Clearly, the triple

$$(u^*, x^*, y^*) := \left(\begin{pmatrix} \bar{u}_J \\ \bar{u}_{J^c} \end{pmatrix}, \bar{x}, \begin{pmatrix} z_J \\ z_{J^c} \end{pmatrix} \right)$$

is an $\mathcal{L}_2$-solution on $[0, t^*]$ of the system Σ_0^* where

$$\Sigma_\rho^* = \Sigma(-A + \rho I, \begin{pmatrix} -B_{\bullet J} & -B_{\bullet J^c} \end{pmatrix}, \begin{pmatrix} C_{\bullet J} \\ -C_{\bullet J^c} \end{pmatrix}, \begin{pmatrix} D_{JJ} & 0 \\ 0 & -D_{J^c J^c} \end{pmatrix}).$$

On the other hand, the passifiability of $\Sigma(A, B, C, D)$ immediately implies that the system

$$\Sigma' = \Sigma(A, \begin{pmatrix} B_{\bullet J} & B_{\bullet J^c} \end{pmatrix}, \begin{pmatrix} C_{\bullet J} \\ C_{\bullet J^c} \end{pmatrix}, \begin{pmatrix} D_{JJ} & 0 \\ 0 & D_{J^c J^c} \end{pmatrix})$$

is passifiable. This means that the system Σ_0^* is also passifiable since $D_{J^c J^c}$ is skew-symmetric. Let ρ be such that Σ_ρ^* is passive. It follows from Fact 3.4.1 that $e^{\rho \cdot}(u^*, x^*, y^*)$ is an $\mathcal{L}_2$-solution of Σ_ρ^* with the zero initial state. Then, the dissipation inequality results in for all $t \in [0, t^*]$

$$\int_0^t e^{2\rho s}(u^*(s))^\top y^*(s)\, ds \geq e^{2\rho t}(x^*(t))^\top K x^*(t) \tag{3.23}$$

where $x \mapsto x^\top K x$ is a storage function for the system Σ_ρ^*. It follows from (3.22) and the definition of y^* that

$$\int_0^t e^{2\rho s}(u^*(s))^\top y^*(s)\, ds = \sum_{i=1}^m \int_0^t e^{2\rho s} u_i^*(s) y_i^*(s)\, ds = 0$$

Therefore, $x^*(t) = 0$ for all $t \in [0, t^*]$ due to (3.23) and the positive definiteness of K. However, this means that $x(0) = x^*(t^*) = 0$ and hence $(u, x, y) \equiv 0$ since (u, x, y) is the unique $\mathcal{L}_2$-solution of LCS(A, B, C, D) with the zero initial state. Consequently, $\mathcal{E} = \{0\}$ since the hybrid solution is nonredundant. Hence, we reach a contradiction. ∎

3.8.6 On quadratic programming

In this subsection, we quote some rather standard facts for the sake of completeness.

Lemma 3.8.8 *Consider the quadratic program*

$$\text{minimize } \tfrac{1}{2}x^\top Qx + b^\top x \text{ subject to } x \geq 0.$$

If Q is nonnegative definite then the Karush-Kuhn-Tucker conditions

$$x \geq 0,\ b + Qx \geq 0, \text{ and } x^\top(b + Qx) = 0$$

are necessary and sufficient for the vector x to be a globally optimal solution of the above quadratic program.

For a detailed discussion on this lemma, the reader is referred to [3, Section 1.2].

Theorem 3.8.9 *Consider the following two quadratic programs*

$$\begin{aligned} QP_1: &\quad \text{minimize } \tfrac{1}{2}x^\top Qx + b^\top x \text{ subject to } Ax \geq c \\ QP_2: &\quad \text{minimize } \tfrac{1}{2}x^\top Qx - c^\top u \text{ subject to } A^\top u - Qx = b \text{ and } u \geq 0. \end{aligned}$$

The following statements hold.

1. (Dorn's duality theorem) [12, Theorem 8.2.4] *Let Q be nonnegative definite. If $\bar{x}$ solves QP_1 then $(\bar{x}, \bar{u})$ solves QP_2 for some $\bar{u}$, and the two extrema are equal.*

2. (Strict converse duality theorem) [12, Theorem 8.2.5] *Let Q be positive definite. If $(\bar{x}, \bar{u})$ solve QP_2 then $\bar{x}$ solves QP_1, and the two extrema are equal.*

3.8.7 Proofs for Section 3.6

In the sequel, $\mathrm{MIN}(f(x), \mathcal{X})$ will denote the minimization problem

$$\text{minimize } f(x) \text{ subject to } x \in \mathcal{X}.$$

The following theorem will be employed later.

Theorem 3.8.10 *Consider a matrix quadruple (B, C, D, K) such that $\mathrm{col}(B, D + D^\top)$ is of full column rank, K is positive definite and the following implication holds*

$$v \in \ker(D + D^\top) \Rightarrow KBv = C^\top v. \tag{3.24}$$

Let N be such that $\mathrm{pos}(N) = \mathcal{Q}_D$. The following statements hold for all $x_0 \in \mathbb{R}^n$, $w \in \mathbb{R}^m$.

1. *If the vector $\bar{x}$ solves MIN$(\|x - x_0\|_K^2, \{x \mid Cx + w \in \mathcal{Q}_D^*\})$ then there exists $\bar{v}$ such that the pair $(\bar{x}, \bar{v})$ solves MIN$(\|x\|_K^2 + 2w^\top Nv, \{(x, v) \mid v \geq 0$ and $x = x_0 + BNv\})$.*

2. *If the pair $(\bar{x}, \bar{v})$ solves MIN$(\|x\|_K^2 + 2w^\top Nv, \{(x, v) \mid v \geq 0$ and $x = x_0 + BNv\})$ then the vector $\bar{x}$ solves MIN$(\|x - x_0\|_K^2, \{x \mid Cx + w \in \mathcal{Q}_D^*\})$.*

3. *The vector $\bar{x}$ solves MIN$(\|x - x_0\|_K^2, \{x \mid Cx + w \in \mathcal{Q}_D^*\})$ if and only if the vector $\bar{v}$ solves MIN$(\frac{1}{2}v^\top N^\top CBNv + (Cx_0 + w)^\top Nv, \{v \mid v \geq 0\})$ and $\bar{x} = x_0 + BN\bar{v}$.*

4. *The vector $\bar{v}$ solves MIN$(\frac{1}{2}v^\top N^\top CBNv + (Cx_0 + w)^\top Nv, \{v \mid v \geq 0\})$ if and only if it solves the linear complementarity problem*

$$\left.\begin{array}{c} v \geq 0 \\ N^\top(Cx_0 + w) + N^\top CBNv \geq 0 \\ v^\top(N^\top(Cx_0 + w) + N^\top CBNv) = 0. \end{array}\right\} \qquad (*)$$

5. *The vector $\bar{v}$ solves the linear complementarity problem $(*)$ if and only if $z = N\bar{v}$ solves the generalized linear complementarity problem*

$$\left.\begin{array}{c} z \in \mathcal{Q}_D \\ Cx_0 + w + CBz \in \mathcal{Q}_D^* \\ z^\top(Cx_0 + w + CBz) = 0. \end{array}\right\} \qquad (**)$$

6. *The generalized linear complementarity problem $(**)$ has at most one solution.*

Proof: Note that $\mathrm{pos}(N) = \mathcal{Q}_D$ implies that $\mathcal{Q}_D^* = \{v \mid N^\top v \geq 0\}$.

1: It follows from Theorem 3.8.9 item 1.

2: It follows from Theorem 3.8.9 item 2.

3: For the 'only if' part suppose that the vector $\bar{x}$ solves MIN$(\|x - x_0\|_K^2, \{x \mid Cx + w \in \mathcal{Q}_D^*\}$. It follows from the first item that there exists $\bar{v}$ such that the pair $(\bar{x}, \bar{v})$ solves MIN$(\|x\|_K^2 + 2w^\top Nv, \{(x, v) \mid v \geq 0$ and $x = x_0 + BNv\})$.

i. Clearly, $\bar{x} = x_0 + BN\bar{v}$.

ii. Since $Nv \in \mathcal{Q}_D \subseteq \ker(D + D^\top)$ for all $v \geq 0$, we have $KBNv = C^\top Nv$ for all $v \geq 0$. Then,

$$\|x\|_K^2 = \|x_0 + BNv\|_K^2 = v^\top N^\top CBNv + 2(Cx_0)^\top Nv + \|x_0\|_K^2 \qquad (3.25)$$

whenever $v \geq 0$. So the vector $\bar{v}$ solves $\mathrm{MIN}(\frac{1}{2}v^\top N^\top CBNv + (Cx_0 + w)^\top Nv)$.

Thus, i. and ii. imply the 'only if' part. For the 'if' part, suppose that the vector $\bar{v}$ solves $\mathrm{MIN}(\frac{1}{2}v^\top N^\top CBNv + (Cx_0 + w)^\top Nv)$ and $\bar{x} = x_0 + BN\bar{v}$. By using (3.25), we get that the vector $\bar{x}$ solves $\mathrm{MIN}(\|x - x_0\|_K^2, \{x \mid Cx + w \in \mathcal{Q}_D^*\})$.

4: It follows from a direct application of Lemma 3.8.8.

5: It is evident from $\mathcal{Q}_D = \{Nv \mid v \geq 0\}$ and $\mathcal{Q}_D^* = \{v \mid N^\top v \geq 0\}$.

6: Suppose that z^i is a solution of the generalized linear complementarity problem

$$\begin{aligned} z &\in \mathcal{Q}_D \\ Cx_0 + w + CBz &\in \mathcal{Q}_D^* \\ z^\top(Cx_0 + w + CBz) &= 0 \end{aligned}$$

for $i = 1, 2$. Note that

$$\begin{aligned} (z^1 - z^2)^\top CB(z^1 - z^2) &= (z^1 - z^2)^\top[(Cx_0 + w + CBz^1) - (Cx_0 + w + CBz^2)] \\ &= -(z^1)^\top(Cx_0 + w + CBz^2) - (z^2)^\top(Cx_0 + w + CBz^1)] \\ &\leq 0. \end{aligned} \tag{3.26}$$

Since $\mathcal{Q}_D \subseteq \ker(D + D^\top)$, we have $z^1 - z^2 \in \ker(D + D^\top)$. Hence, $(z^1 - z^2)^\top CB(z^1 - z^2) = (z^1 - z^2)^\top B^\top KB(z^1 - z^2) \geq 0$ by the hypothesis. Together with the above inequality, this gives $(z^1 - z^2)^\top CB(z^1 - z^2) = (z^1 - z^2)^\top B^\top KB(z^1 - z^2) = 0$. Since $\mathrm{col}(B, D + D^\top)$ is of full column rank and K is positive definite, we get $z^1 = z^2$. ■

Proof of Theorem 3.6.1: Take any $K \in \mathcal{K}$. It follows from [3, Theorem 2.8.1] that the minimization problem in (3.5), i.e., $\mathrm{MIN}(\|x - x_0\|_K^2, \{x \mid Cx \in \mathcal{Q}_D^*\})$ is solvable since $\{x \mid Cx \in \mathcal{Q}_D^*\}$ is a polyhedron and $\|x - x_0\|_K^2$ is bounded below. Let x_K^+ denote one of its solutions. It follows from Theorem 3.8.10 items 3-6 that $x_K^+ = x_0 + Bu^0$ where u^0 is the unique solution of the generalized linear complementarity problem

$$\begin{aligned} z &\in \mathcal{Q}_D \\ Cx_0 + CBz &\in \mathcal{Q}_D^* \\ z^\top(Cx_0 + CBz) &= 0. \end{aligned}$$

Note that u^0 and hence x_K^+ is independent of K. ■

To prove the remaining claims, we need to do some preparations. The rational complementarity problem and its relation to initial solutions will be in order.

Problem 3.8.11 (RCP(x_0, A, B, C, D)) Given $x_0 \in \mathbb{R}^n$ and (A, B, C, D) with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$ and $D \in \mathbb{R}^{m\times m}$ find $\hat{u}(s) \in \mathbb{R}^m(s)$ such that

1. $\hat{u}(s) \perp \hat{y}(s)$ for all $s \in \mathbb{C}$
2. $\hat{u}(\sigma) \geq 0$ and $\hat{y}(\sigma) \geq 0$ for all sufficiently large $\sigma \in \mathbb{R}$

where

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + [D + C(sI - A)^{-1}B]\hat{u}(s).$$

For brevity of notation, we denote RCP(x_0, A, B, C, D) by RCP(x_0) if (A, B, C, D) is clear from the context. There is one-to-one correspondence between the solutions of RCP and initial solutions of LCS as described in the following lemma.

Lemma 3.8.12 *Consider a matrix quadruple (A, B, C, D). The following statements hold.*

1. *Let $\hat{u}(s)$ be a solution of RCP($x_0, \hat{w}(s)$) for some x_0 and strictly proper $\hat{w}(s)$. Define $\hat{x}(s)$ and $\hat{y}(s)$ as follows*

$$\begin{aligned} \hat{x}(s) &= (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s), \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s). \end{aligned}$$

   Then, the inverse Laplace transform (u, x, y) of $(\hat{u}(s), \hat{x}(s), \hat{y}(s))$ is an initial solution of LCS(A, B, C, D) with the initial state x_0.

2. *Let (u, x, y) be an initial solution of LCS(A, B, C, D) with the initial state x_0 and let $\hat{u}(s)$ be the Laplace transform of u. Then, $\hat{u}(s)$ solves RCP(x_0).*

Proof: Evident from [9, Theorem 5.3]. ■

The following properties of RCP will be used later.

Lemma 3.8.13 *Consider a matrix quadruple (A, B, C, D). Let $G(s)$ be the transfer matrix of the system $\Sigma(A, B, C, D)$, i.e., $G(s) = D + C(sI - A)^{-1}B$. Suppose that $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then the following statements hold.*

1. *There exists a unique solution of RCP(x_0) for all initial states x_0.*

2. *If $G(s)$ is totally of index 1 then the solution of RCP(x_0) is proper for all x_0.*

Proof: The first claim follows from [8, Theorem 4.1 and Corollary 4.10]. For the second one, let the solution of RCP(x_0) be $\hat{u}(s)$. Then, there exists an index set $K \subseteq \overline{m}$ such that

$$\hat{u}_K(s) = -G_{KK}^{-1}(s)C_{K\bullet}(sI - A)^{-1}x_0 \tag{3.27}$$

$$\hat{u}_{\overline{m}\setminus K}(s) \equiv 0 \tag{3.28}$$

by the formulation of RCP. According to Lemma 2.5.4, there exists $\alpha > 0$ such that

$$||G_{KK}^{-1}(\sigma)|| \leq \alpha\sigma$$

for all sufficiently large σ since $G(s)$ is totally of index 1. Then, the equation (3.27) implies that $\hat{u}_K(\sigma)$ is bounded for all sufficiently large σ due to the fact that $C_{K\bullet}(sI - A)^{-1}$ is strictly proper. Consequently, $\hat{u}(s)$ is proper. ■

Proof of Lemma 3.6.3: The first claim follows from Lemma 3.8.12 and Lemma 3.8.13 item 1. It remains to prove the second one. Let (u, x, y) be the unique solution of LCS(A, B, C, D) with some initial state x_0 and let $(\hat{u}(s), \hat{x}(s), \hat{y}(s))$ be the Laplace transform of (u, x, y). It follows from Lemma 3.8.12 item 2 that $\hat{u}(s)$ solves RCP(x_0). According to Lemma 3.8.13 item 2, $\hat{u}(s)$ must be proper. Let the power series expansion of $\hat{u}(s)$ be given by

$$\hat{u}(s) = \bar{u} + u^1 s^{-1} + u^2 s^{-2} + \cdots. \tag{3.29}$$

Since $\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)$, $\hat{x}(s)$ is strictly proper, i.e., $\lim_{s\to\infty} \hat{x}(s) = 0$. Furthermore, the equation $\hat{y}(s) = C\hat{x}(s) + G(s)\hat{u}(s)$ implies that $\hat{y}(s)$ is proper and $\lim_{s\to\infty} \hat{y}(s) = D\bar{u}$. ■

Proof of Corollary 3.6.4: Since the system $\Sigma(A, B, C, D)$ is passifiable by pole shifting there exists ρ such that $\Sigma(A + \rho I, B, C, D)$ is passive. It follows from Lemma 3.8.5 items 5 and 4 that $G(s)$ is totally of index 1 and $G(\sigma)$ is positive definite for all sufficiently large σ. Then, the claims follow by applying Lemma 3.6.3. ■

Proof of Theorem 3.6.5:

1: It follows from Theorem 3.6.1 and application of Theorem 3.8.10 items 3 and 4 for $w = 0$.

2: It follows from Theorem 3.6.1 and application of Theorem 3.8.10 items 3 and 6 for $w = 0$.

3: It follows from Theorem 3.6.1 and application of Theorem 3.8.10 item 1 for $w = 0$.

4: It is known from Corollary 3.6.4 that there exists a unique initial solution of LCS(A, B, C, D) with the initial state x_0 and the impulsive part of this solution is of the form $(\bar{u}, 0, D\bar{u})$. Hence, it remains to show that $\bar{u} = u^0$. Since (u, y) is initially nonnegative due to Definition 3.6.2, we have

$$\bar{u} \geq 0 \quad D\bar{u} \geq 0 \tag{3.30a}$$

Note that $\hat{u}$ is the unique solution of RCP(x_0) due to Lemma 3.8.12 item 2. The power series expansion of $\hat{u}(s)$ around infinity is of the form

$$\hat{u}(s) = \bar{u} + u^1 s^{-1} + u^2 s^{-2} + \cdots .$$

Then, the power series expansion of $\hat{y}(s)$, the Laplace transform of y, around infinity is of the form

$$\begin{aligned}\hat{y}(s) &= C(sI - A)^{-1}x_0 + [D + C(sI - A)^{-1}B]\hat{u}(s)\\ &= D\bar{u} + (Cx_0 + CB\bar{u} + Du^1)s^{-1} + (CAx_0 + CAB\bar{u} + CBu^1 + Du^2) + \cdots .\end{aligned}$$

Hence, we get from the formulation of RCP that $0 = \hat{u}(s)^\top \hat{y}(s) = \bar{u}^\top D\bar{u} + \bar{u}^\top[Cx_0 + CB\bar{u} + (D + D^\top)u^1]s^{-1} + \cdots$ for all $s \in \mathbb{C}$. Thus, we have

$$\bar{u}^\top D\bar{u} = 0 \tag{3.31a}$$

$$\bar{u}^\top[Cx_0 + CB\bar{u} + (D + D^\top)u^1] = 0. \tag{3.31b}$$

Note that (3.31a), together with (3.30), implies that $\bar{u} \in \mathcal{Q}_D$. Besides, (3.31a) implies that $(D + D^\top)\bar{u} = 0$. This results in

$$\bar{u}^\top(Cx_0 + CB\bar{u}) = 0 \tag{3.32}$$

due to (3.31b). Since $\hat{u}(s)$ is a solution of RCP(x_0), we get

$$\bar{u} + u^1\sigma^{-1} \geq 0 \tag{3.33a}$$

$$D\bar{u} + (Cx_0 + CB\bar{u} + Du^1)\sigma^{-1} \tag{3.33b}$$

for all sufficiently large σ, and

$$(\bar{u} + u^1\sigma^{-1})^\top[D\bar{u} + (Cx_0 + CB\bar{u} + Du^1)\sigma^{-1}] = 0 \tag{3.34}$$

for all $\sigma \in \mathbb{R}$. It follows from (3.33) and (3.34) that $\bar{u} + u^1\sigma^{-1}$ is a solution of LCP$((Cx_0 + CB\bar{u})\sigma^{-1}, D)$ for all sufficiently large σ. Consequently, $Cx_0 + CB\bar{u} \in K(D)$ and even $Cx_0 + CB\bar{u} \in \mathcal{Q}_D^*$ due to Lemma 1.2.4 item 2. Finally, we can conclude that $\bar{u}$ is a solution of the generalized linear complementarity problem

$$\begin{aligned} z &\in \mathcal{Q}_D \\ Cx_0 + CBz &\in \mathcal{Q}_D^* \\ z^\top(Cx_0 + CBz) &= 0. \end{aligned}$$

Note that it is shown in 2 that u^0 is a solution of this problem. However, it is already known from Theorem 3.8.10 item 6 that the above problem has at most one solution. Hence, $\bar{u} = u^0$. ■

Proof of Theorem 3.6.7: Evident from Definition 3.6.6, Theorem 2.3.3 (for the zero input) and Theorem 3.6.5. ■

Proof of Theorem 3.6.9: Evident from Definition 3.6.8, Corollary 3.4.4 and Theorem 3.6.5. ■

References

[1] M.J. Chien. Piecewise-linear theory and computation of solutions of homeomorphic resistive networks. *IEEE Trans. Circuits and Systems*, 24(3):118–127, 1977.

[2] M. Ciampa, P. Terreni, and M. Poletti. Conditions for the existence and uniqueness of dc solutions of networks containing nonlinear opamps with ideal models. In *Proc. of IEEE International Symposium on Circuits and Systems*, volume 4, pages 2498–2501, 1993.

[3] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem.* Academic Press, Inc., Boston, 1992.

[4] C.A. Desoer and J. Katzenelson. Nonlinear RLC networks. *The Bell System Technical Journal*, 44:161-198, 1965.

[5] T. Fujisawa and E. S. Kuh. Some results on existence and uniqueness of solutions of nonlinear networks. *IEEE Trans. Circuit Theory*, 18(5):501–506, 1971.

[6] T. Fujisawa and E. S. Kuh. Piecewise-linear theory of nonlinear networks. *SIAM Journal on Applied Mathematics*, 22:307–328, 1972.

[7] M. Green and A. N. Wilson Jr. On the uniqueness of a circuit's DC operating point when its transistors have variable current gains. *IEEE Trans. Circuits and Systems*, 36(12):1521–1528, 1989.

[8] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. The rational complementarity problem. *Linear Algebra and Its Applications*, 294:93–135, 1999.

[9] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics*, 60(4):1234–1269, 2000.

[10] K.H. Johansson, J. Lygeros, S. Sastry, and M. Egerstedt. Simulation of Zeno hybrid automata. In *Proc. of the 38th IEEE Conference on Decision and Control*, pages 3538–3543, Phoenix (USA), 1999.

[11] H. Lütkepohl. *Handbook of Matrices*. Wiley, New York, 1996.

[12] O.L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, New York, 1969.

[13] R. O. Nielsen and A. N. Wilson Jr. Topological criteria for establishing the uniqueness of solutions. *IEEE Trans. Circuits and Systems*, 24(7):349–362, 1977.

[14] T. Ohtsuki, T. Fujisawa, and S. Kumagai. Existence theorems and a solution algorithm for piecewise-linear networks. *SIAM Journal of Mathematical Analysis*, 8(1):69–99, 1977.

[15] T. Ohtsuki and H. Watanabe. State-variable analysis of RLC networks containing nonlinear coupling elements. *IEEE Trans. Circuit Theory*, 18(1):26–38, 1969.

[16] V. C. Prasad. Necessary and sufficient condition for uniqueness of solutions of certain piecewise linear resistive networks containing transistors and diodes. *Electronic Letters*, 28(13), 1992.

[17] V. C. Prasad and V. P. Prakash. Existence, uniqueness and determination of solutions of certain piecewise linear resistive networks. In *Proc. of IEEE International Symposium on Circuits and Systems*, volume 2, pages 1470–1473, 1990.

[18] I. W. Sandberg and A. N. Wilson Jr. Some theorems on the properties of DC equations of nonlinear networks. *Bell System Technical Journal*, 48:1–35, 1969.

[19] I. W. Sandberg and A. N. Wilson Jr. Existence and uniqueness of solutions for the equations of nonlinear networks. *SIAM Journal on Applied Mathematics*, 22:173–186, 1972.

[20] A.J. van der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9:266–301, 1996.

[21] A.J. van der Schaft and J.M. Schumacher. Complementarity modelling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[22] J. C. Willems. Dissipative dynamical systems. *Arch. Rational Mech. Anal.*, 45:321–393, 1972.

Chapter 4

Systems with Piecewise Linear Elements

4.1 Introduction

The consideration of dynamical systems with external variables can be motivated in several ways. In control theory, external variables occur as actuators and sensors. In a hierarchical modeling context, external variables arise as the variables through which subsystems may be connected to each other. In studies of dissipative systems, inputs and outputs are used in pairs to quantify energy exchange. Some interesting classes of systems may be obtained by connecting inputs and outputs to each other in a particular way; for instance, letting certain inputs and outputs be linked by a feedback with given maximal L_2-gain has been popular in recent years as a way of describing model uncertainty. The many uses that can be made of inputs and outputs (or more generally, of external variables) shows the strength of systems theory as it has been developed in the past decades.

In this chapter we will be concerned with yet another way of using external variables. Similar to the way that inputs and outputs are used in the model uncertainty description that we just mentioned, the class of systems that we discuss below is obtained by connecting inputs and outputs in a specific way. Similar to the use of inputs and outputs in the context of dissipativity, the links that we specify are defined for pairs of (scalar) inputs and outputs. Below we consider the class of dynamical systems that is obtained by combining a dynamic linear input/output system with a static piecewise linear input/output relation. The properties of the systems that are obtained in this way are coded entirely into the usual (A, B, C, D) parameters of the linear system and the parameters of the piecewise linear relation. As a result, we study a class of nonlinear and nonsmooth dynamical systems using notions from linear systems theory.

Although of course many properties of the systems considered here are of potential interest, we concentrate on the most basic properties, namely existence and uniqueness of solutions. As already mentioned, we consider piecewise linear relations between pairs of variables. These relations do not necessarily specify one variable as a function of the other (see e.g. the characteristics shown in Figure 4.1(b) or Figure 4.3(b) below). For this reason, standard theorems on the well-posedness of feedback connections do not apply. Nevertheless, it will be shown below that existence and uniqueness of solutions do hold if certain conditions are satisfied. Of course one may consider piecewise linear systems with additional inputs and outputs that are not connected by piecewise linear relations; here however we study the basic situation in which there are no additional external variables so that solutions, if uniqueness holds, are parametrized by initial conditions.

The systems that we consider can also be studied in the framework of differential inclusions, as developed for instance in [1]. Indeed this is a very general framework that allows the study of many kinds of systems. We believe that the approach taken in this chapter is more tailored to the specific structure of piecewise linear systems; as will be shown below, some ideas from linear systems theory actually play an important role in the formulation of necessary and sufficient conditions for existence and uniqueness of solutions. The work by Filippov (see for instance [7]) is closer to the approach we take here than the general differential inclusion framework. On the one hand discontinuous dynamical systems in Filippov's sense may be described in terms of relays, which are a special case of the piecewise linear characteristics considered here; on the other hand Filippov allows nonlinear dynamics, whereas we restrict ourselves to couplings between piecewise linear characteristics and linear dynamics. As a result we obtain conditions for well-posedness that are different in nature from those considered by Filippov.

Piecewise linear systems are important for several reasons:

- They form a limited class which nevertheless can approximate nonlinear phenomena as accurately as desired.
- As quite natural extensions of linear systems, they allow already well-established linear analysis/synthesis methods to be applied locally.
- They arise naturally in many applications ranging from circuit theory to economics and from mechanics to control systems.

To give a quick impression of applications areas, we mention linear electrical circuits with piecewise resistive elements [2, 15, 24], systems with relays [22] and/or saturation characteristics, mechanical systems with Coulomb friction [17], variable structure systems [23], and bang-bang control [3, 14].

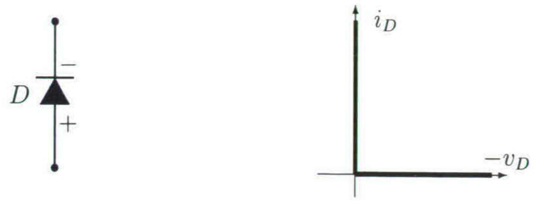
In many of the application areas mentioned, one encounters piecewise linear relations between two variables that cannot be rewritten as functions from one variable to the other. For instance, the relay characteristic is of such a type. Although it would be possible to apply a change of coordinates so as to obtain a functional relationship (for instance, rotation by 45 degrees in the relay example), such a transformation will affect the feedthrough term in the linear system component of the overall system description. As is well-known, even Lipschitzian feedback may not be well-posed for linear systems with feedthrough terms. In the development below, we allow piecewise linear relations of non-functional type as well as nonzero feedthrough terms in the part of the system that is specified by parameters (A, B, C, D).

As is well-known (see for instance [6]), piecewise linear relations may be described in terms of the linear complementarity problem (LCP) of mathematical programming. The LCP is briefly described in Section 4 below, together with one of its generalizations, the *horizontal* complementarity problem (HLCP). The complementarity formulation has been used for *static* piecewise linear systems in [15, 24]; this chapter may be viewed as an extension of the cited work in the sense that we consider *dynamic* systems. The chapter can also be viewed as a generalization of earlier work which was concerned with well-posedness of linear systems coupled to the ideal diode (pure complementarity) characteristic [11, 20] or to the relay characteristic [16], although the approach taken here is somewhat different from the one in [16].

The organization of the chapter is as follows. We begin with a very quick look at motivational examples in Section 4.2. Section 4.3 is devoted to the introduction of the piecewise linear characteristics that will be under investigation in the sequel. This will be followed by recalling the related complementarity problems in Section 4.4. In Section 4.5, we propose a definition of solution for the linear systems with piecewise linear characteristics and derive sufficient conditions under which the solutions do exist and are unique. Section 4.6 provides the connections with the previous chapters by showing that some of the well-posedness results obtained earlier can be deduced from our new general framework. It also illustrates the implications of our results on the linear systems with relays and saturation characteristics. Finally, the conclusions in Section 4.7 will be followed by the proofs in Section 4.8.

4.2 Motivational Examples

In circuit theory, piecewise linear modeling is a widely used technique. For instance, ideal modeling of a diode yields a voltage-current characteristic depicted in Figure 4.1. Similar-looking characteristics can be obtained from parallel/series connections of linear resistors, ideal diodes and batteries. Such a circuit and its voltage-current characteristic are shown in Figure 4.2. We can think of many other piecewise resistive elements such as saturation characteristics (see Figure 4.3) or dynamical elements such as capacitors/inductors with piecewise linear charge-voltage/flux-current characteristics. Of course, piecewise lin-



(a) Ideal diode (b) Ideal diode characteristic

Figure 4.1: Ideal diode and its voltage-current characteristic

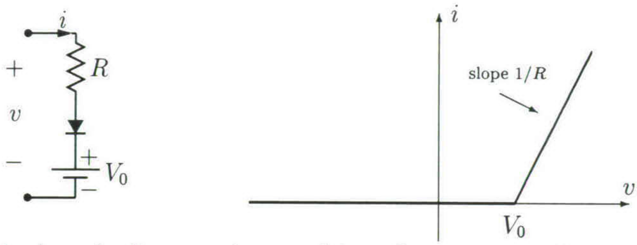
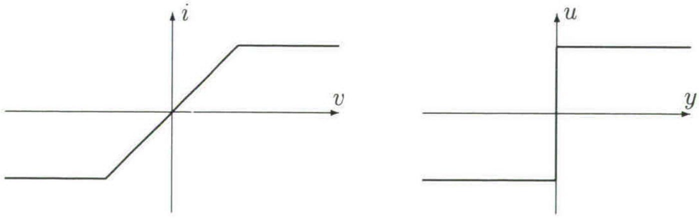


Figure 4.2: A piecewise linear resistor and its voltage-current characteristic

ear elements also occur in various other engineering areas. For instance, the ideal relay characteristic (see Figure 4.3) serves as an idealized model of Coulomb friction in mechanical systems and it arises as well in switching control schemes. Many other examples



(a) Saturation characteristic (b) Ideal relay characteristic

Figure 4.3: Saturation and ideal relay characteristics

and potential application areas of piecewise linear phenomena can be found. With these wide-range application areas in our mind, we will address the well-posedness (in the sense of existence and uniqueness of solutions) issues of models consisting of a linear (dynamical) system coupled with elements that are of a piecewise linear nature. As discussed in Chapter 2, we consider well-posedness from a model validation point of view.

4.3 Piecewise Linear Characteristics

The main ingredients of this section are *piecewise linear characteristics*. We consider only those characteristics which are piecewise affine curves in $\mathbb{R}^2$ as it is defined in the following.

Definition 4.3.1 A set $\mathcal{G}$ is called a *k-piecewise linear characteristic* if there exist *(directions)* d^-, $d^+ \in \mathbb{R}^2$ with $\|d^-\| = \|d^+\| = 1$ and *(vertices)* $[v^i]_{i=1}^{k-1} \in (\mathbb{R}^2)^k$ such that the two half lines

$$\mathcal{G}_1 = \{\lambda d^- + v^1 \mid 0 \leq \lambda\} \tag{4.1a}$$

$$\mathcal{G}_k = \{v^{k-1} + \lambda d^+ \mid 0 \leq \lambda\} \tag{4.1b}$$

and $k-2$ line segments

$$\mathcal{G}_i = \{\lambda v^{i-1} + (1-\lambda)v^i \mid 0 \leq \lambda \leq 1\} \text{ for } i = 2, 3, \ldots, k-1 \tag{4.1c}$$

satisfy the following conditions

1. $\mathcal{G}_i \cap \mathcal{G}_{i+1} = \{v_i\}$ for $i = 1, 2, \ldots, k-1$,
2. $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$ if $\mid i - j \mid > 1$,
3. $\bigcup_{i=1}^{k} \mathcal{G}_i = \mathcal{G}$.

If the above conditions hold we write $\mathcal{G} = \text{plc}(d^-, [v^i]_{i=1}^k, d^+)$. We say that $(d^-, [v^i]_{i=1}^k, d^+)$ is a *minimal description* of $\mathcal{G}$ if $\mathcal{G} = \text{plc}(d^-, [v^i]_{i=1}^k, d^+)$ and $\mathcal{G}$ is not a $(k-1)$-piecewise linear characteristic. We say that the vertex $v \in [v^i]_{i=1}^k$ of $\text{plc}(d^-, [v^i]_{i=1}^k, d^+)$ is *redundant* if $\text{plc}(d^-, [v^i]_{i=1}^k \setminus v, d^+) = \text{plc}(d^-, [v^i]_{i=1}^k, d^+)$.

Remark 4.3.2 Notice that $\text{plc}(d^-, [v^1, v^2, \ldots, v^k], d^+) = \text{plc}(d^+, [v^k, v^{k-1}, \ldots, v^1], d^-)$, i.e., the d's and v's are not unique. Notice also that every k-piecewise linear characteristic can be regarded as a $k+p$-piecewise linear characteristic by adding p artificial (redundant) vertices. It can be verified that there are exactly two minimal descriptions for every $\mathcal{G}$.

An example of a k-piecewise linear characteristic is depicted in Figure 4.4. It is known that such piecewise linear curves can be represented by using *complementarity* variables. The next definition is a first step towards introducing *complementarity representations* for k-piecewise linear characteristics.

Definition 4.3.3 An ordered set $[z^i]_{i=1}^k \in (\mathbb{R}^m)^k$ is called *k-horizontal complementary* if

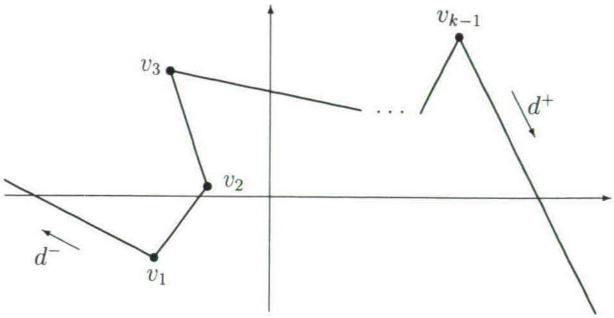


Figure 4.4: An example of k-piecewise linear characteristic

the following conditions hold:

$$0 \le z^1 \tag{4.2}$$

$$0 \le z^i \le e \text{ for } i = 2, 3, \dots, k-1 \tag{4.3}$$

$$0 \le z^k \tag{4.4}$$

$$(z^1)^\top z^2 = 0 \tag{4.5}$$

$$(e - z^i)^\top z^{i+1} = 0 \text{ for } i = 2, 3, \dots, k \tag{4.6}$$

where e denotes the vector of ones. The set of all such k-horizontal complementary ordered sets is denoted by $\mathcal{HC}_k^m$.

We will often use the following particular description of the set $\mathcal{HC}_k$.

Proposition 4.3.4 *Let the sets* $[\zeta^i]_{i=1}^k$ *be defined as*

$$\zeta^1 = \{[z^i]_{i=1}^k \in \mathbb{R}^k \mid 0 \le z^1,\ z^2 = z^3 = \cdots = z^k = 0\}, \tag{4.7a}$$

$$\zeta^k = \{[z^i]_{i=1}^k \in \mathbb{R}^k \mid 0 \le z^k,\ z^1 = 0,\ z^2 = z^3 = \cdots = z^{k-1} = 1\}, \tag{4.7b}$$

and for $j = 2, 3, \dots, k-1$,

$$\zeta^j = \{[z^i]_{i=1}^k \in \mathbb{R}^k \mid 0 \le z^j \le 1 \text{ and } z^i = \begin{cases} 0, & i = 1 \\ 1, & i = 2, 3, \dots, j-1 \\ 0, & i = j+1, j+2, \dots, k \end{cases} \}. \tag{4.7c}$$

Then the following statements hold.

1. *For* $i = 1, 2, \dots, k-1$, $\zeta^i \cap \zeta^{i+1}$ *is a singleton.*

2. $\zeta^i \cap \zeta^j = \emptyset$ *if* $|i-j| > 1$.

3. $\bigcup_{i=1}^{k} \zeta^i = \mathcal{HC}_k$.

The proof of the above proposition directly follows from the definitions of the sets ζ^j.

There is a correspondence between k-piecewise linear characteristics and affine functions defined on the set $\mathcal{HC}_k$. To see this, consider a k-piecewise linear characteristic $\mathcal{G} = \mathrm{plc}(d^-, [v^1, v^2, \ldots, v^k], d^+)$ and the affine function $f : \mathcal{HC}_k \to \mathbb{R}^2$ given by

$$f([z^i]_{i=1}^k) := v^1 + d^- z^1 + \sum_{j=2}^{k-1} (v^j - v^{j-1}) z^j + d^+ z^k. \tag{4.8}$$

Let the sets $[\zeta^i]_{i=1}^k$ be as in Proposition 4.3.4. Note that $f(\zeta^i) = \mathcal{G}_i$. Moreover, it can be verified that f is a bijection. We will represent piecewise linear characteristics by exploiting this correspondence. With this aim, consider m k-piecewise linear characteristics $[\mathcal{G}^i]_{i=1}^m$. We associate to each characteristic $\mathcal{G}^i = \mathrm{plc}(d^{i,-}, [v^{i,1}, v^{i,2}, \ldots, v^{i,k}], d^{i,+})$ two vectors

$$r^i = \mathrm{col}(-d_1^{i,-}, v_1^{i,2} - v_1^{i,1}, \ldots, v_1^{i,k-1} - v_1^{i,k-2}, d_1^{i,+}) \tag{4.9a}$$

$$s^i = \mathrm{col}(-d_2^{i,-}, v_2^{i,2} - v_2^{i,1}, \ldots, v_2^{i,k-1} - v_2^{i,k-2}, d_2^{i,+}). \tag{4.9b}$$

and a function $f^i : \mathcal{HC}_k \to \mathbb{R}^2$ defined by

$$f^i([z^i]_{i=1}^k) := v^{i,1} - \begin{pmatrix} r_1^i \\ s_1^i \end{pmatrix} z^1 + \begin{pmatrix} r_2^i \\ s_2^i \end{pmatrix} z^2 + \begin{pmatrix} r_3^i \\ s_3^i \end{pmatrix} z^3 + \cdots + \begin{pmatrix} r_k^i \\ s_k^i \end{pmatrix} z^k. \tag{4.10}$$

Define $q^u, q^y \in \mathbb{R}^m$ as $q^u = \mathrm{col}(v_1^{1,1}, v_1^{2,1}, \ldots, v_1^{m,1})$ and $q^y = \mathrm{col}(v_2^{1,1}, v_2^{2,1}, \ldots, v_2^{m,1})$. Also define $[R^j]_{j=1}^k, [S^j]_{j=1}^k \in (\mathbb{R}^{m\times m})^k$ as $R^j = \mathrm{diag}(r_j^1, r_j^2, \ldots, r_j^m)$ and $S^j = \mathrm{diag}(s_j^1, s_j^2, \ldots, s_j^m)$.

Fact 4.3.5 Consider m k-piecewise linear characteristics $[\mathcal{G}^i]_{i=1}^m$. Let $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ be as defined above. Then, the following statements are equivalent.

1. For each $i \in \overline{m}$,

$$\begin{pmatrix} u_i \\ y_i \end{pmatrix} \in \mathcal{G}^i. \tag{4.11}$$

2. For some $[z^i]_{i=1}^k \in \mathcal{HC}_k^m$,

$$u = q^u - R^1 z^1 + R^2 z^2 + R^3 z^3 + \cdots + R^k z^k \tag{4.12a}$$

$$y = q^y - S^1 z^1 + S^2 z^2 + S^3 z^3 + \cdots + S^k z^k. \tag{4.12b}$$

Moreover, the mapping $\mathrm{col}(u, y) \mapsto [z^i]_{i=1}^k$ is a bijection.

Indeed, the assertion follows immediately from the fact that each f^i is a bijection. In the light of Fact 4.3.5, we say that $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ is a *horizontal complementarity representation* of $[\mathcal{G}^i]_{i=1}^m$. It is clear from the discussion following Definition 4.3.1 that these representations are not unique.

4.4 Complementarity Problems

There are a number of interesting generalizations of the linear complementarity problem (LCP) of mathematical programming. Particularly, the (Extended) Horizontal LCP will play a key role in representing piecewise linear characteristics.

Problem 4.4.1 (HLCP($q, [M^i]_{i=1}^k$)) Given $q \in \mathbb{R}^m$ and $[M^i]_{i=1}^k \in \mathbb{R}^{m \times m}$, find $[z^i]_{i=1}^k \in \mathcal{HC}_k^m$ such that $M^1 z^1 = q + \sum_{i=2}^k M^i z^i$.

The HLCP was introduced in [12] with $k = 3$ and $M^1 = I$, and further developed in [13] with an eye towards piecewise linear functions. We briefly recall some facts from [21] and state a result on solvability of the problem which is parallel to Lemma 1.2.4 item 1. To do this we need to make some definitions.

Definition 4.4.2 A matrix $R \in \mathbb{F}^{m \times m}$ is called *a column representative* of $[M^i]_{i=1}^k \in (\mathbb{F}^{m \times m})^k$ if

$$R_{\bullet i} \in \{M^1_{\bullet i}, M^2_{\bullet i}, \ldots, M^k_{\bullet i}\} \text{ for all } i \in \bar{m}.$$

For a given $l \in \overline{k}^m$, the matrix $([M^i]_{i=1}^k)^l$ is defined by

$$([M^i]_{i=1}^k)^l_{\bullet j} = M^{l_j}_{\bullet j} \text{ for } j = 1, 2, \ldots, m.$$

Definition 4.4.3 We say that an ordered set of matrices $[M^i]_{i=1}^k$

- is *nondegenerate* if all column representative matrices are nondegenerate.
- has the *column $\mathcal{W}$-property* if the determinants of the column representative matrices are either all positive or all negative.

For the sake of completeness, we quote the following theorem from [21].

Theorem 4.4.4 [21] *HLCP($q, [M^i]_{i=1}^k$) has a unique solution for all $q \in \mathbb{R}^n$ if and only if $[M^i]_{i=1}^k$ has the column $\mathcal{W}$-property.*

Remark 4.4.5 The LCP can be regarded as a special case of the HLCP. Indeed, LCP(q, M) is nothing but HLCP$(q, [I, M])$. In this case, Lemma 1.2.4 item 1 and Theorem 4.4.4 coincide since M is a $\mathcal{P}$-matrix if and only if the determinants of all column representative matrices of $[I, M]$ are positive, i.e., $[I, M]$ has the column $\mathcal{W}$-property. On the other hand, HLCP$(q, [M^i]_{i=1}^k)$ can be written as an LCP whenever M^1 is invertible. For this purpose, we define

$$r(q) := \begin{pmatrix} q \\ e \\ e \\ \vdots \\ e \end{pmatrix} \quad \text{and } N([M^i]_{i=1}^k) := \begin{pmatrix} \tilde{M}^1 & \tilde{M}^2 & \cdots & \tilde{M}^{k-1} \\ -I & 0 & \cdots & 0 \\ 0 & -I & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -I \end{pmatrix}$$

where $\tilde{M}^i = (M^1)^{-1} M^{i+1}$ for $i = \overline{k-1}$. There is a one-to-one correspondence between the solutions of HLCP$(q, [M^i]_{i=1}^k)$ and LCP$(r((M^1)^{-1}q), N([M^i]_{i=1}^k))$. In fact, if $[z^i]_{i=1}^k$ solves the former then $\operatorname{col}(z^2, z^3, \ldots, z^k)$ solves the latter and vice versa.

4.5 Piecewise Linear Systems

Consider continuous-time, linear and time-invariant systems given by

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{4.13a}$$

$$y(t) = Cx(t) + Du(t) \tag{4.13b}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$ and A, B, C, and D are matrices with appropriate sizes. We denote (4.13) by $\Sigma(A, B, C, D)$. Let $[\mathcal{G}^i]_{i=1}^m$ be a given family of k-piecewise linear characteristics. Let the variables u and y be coupled via these characteristics as depicted in Figure 4.5, i.e.,

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \mathcal{G}^i \tag{4.14}$$

for all t. We denote the resulting piecewise linear system ($\Sigma(A, B, C, D)$ together with (4.14)) by PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. The following definition will make clear what is understood by a solution of PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$.

Definition 4.5.1 A triple $(u, x, y) \in \mathcal{PB}^{m+n+m}$ is said to be a solution on $[0, T)$ of PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0 if the following conditions hold for all

$t \in [0, T)$

$$x(t) = x_0 + \int_0^t [Ax(s) + Bu(s)]\, ds, \tag{4.15}$$

$$y(t) = Cx(t) + Du(t), \tag{4.16}$$

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \mathcal{G}^i \text{ for } i = 1, 2, \ldots, k. \tag{4.17}$$

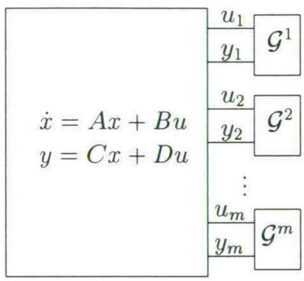


Figure 4.5: Overall system

In the sequel, we will be dealing with systems having *low index* in the sense as it will be defined in the following definitions.

Definition 4.5.2 A rational matrix $M(s) \in \mathbb{R}^{m \times m}(s)$ is said to be of *index* k if it is invertible as a rational matrix and $s^{-k}M^{-1}(s)$ is proper rational.

The notion of index will be generalized to families of matrices via column representatives in what follows.

Definition 4.5.3 A family of rational matrices $[M^i(s)]_{i=1}^k$ is said to be of *index* k if all its column representative matrices are of index k.

We can now present the main result of this chapter.

Theorem 4.5.4 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Suppose that* $[G(\sigma)R^j - S^j]_{j=1}^k$ *has the column* $\mathcal{W}$*-property for all sufficiently large* σ. *Then, the following statements hold.*

1. *Assume that* $[G(s)R^j - S^j]_{j=1}^k$ *is of index* 1. *There exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *with the initial state* x_0 *if and only if HLCP(*$Cx_0 + Dq^u - q^y, [DR^j - S^j]_{j=1}^k$*) is solvable.*

2. *If* $[DR^1 - S^1, DR^k - S^k]$ *is nondegenerate then there exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *for all initial states.*

Notice that the horizontal complementarity representations of a family of piecewise linear characteristics are not unique in general. However, the (sufficient) condition presented above for well-posedness depend on those representations. Naturally, one might ask whether it is possible that the condition holds for one representation but not for another one. As stated in the following theorem, the answer of this question is negative. In other words, the above theorem is independent of the choice of the representations.

Theorem 4.5.5 *Consider a matrix pair* $(M, N) \in \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times m}$ *and* k*-piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Let* $(\cdot, \cdot, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *and* $(\cdot, \cdot, [\bar{R}^j]_{j=1}^k, [\bar{S}^j]_{j=1}^k)$ *be horizontal complementarity representations of* $[\mathcal{G}^j]_{j=1}^m$. *If* $[MR^j + NS^j]_{j=1}^k$ *has the column* $\mathcal{W}$*-property then so does* $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$.

4.6 Examples

In this section we apply Theorem 4.5.4 to subclasses of piecewise linear system.

4.6.1 Linear complementarity systems

The well-posedness results for linear complementarity systems that have been presented earlier can be obtained as a special case of Theorem 4.5.4.

Corollary 4.6.1 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *where the piecewise linear characteristic* $\mathcal{G}^i$ *is as depicted in Figure 4.6 for each* $i \in \overline{m}$. *Suppose that* $G(s)$ *is totally of index* 1 *and* $G(\sigma)$ *is a* $\mathcal{P}$*-matrix for all sufficiently large* σ. *Then, there exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *with the initial state* x_0 *if and only if* $Cx_0 \in K(D)$.

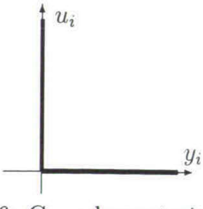


Figure 4.6: Complementarity characteristic

4.6.2 Linear relay systems

The existence and uniqueness of solutions of linear relay systems (linear systems coupled with relay characteristics) are addressed in [16] (see also [10]). The following corollary states a parallel result to those stated in [10, 16].

Corollary 4.6.2 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *where the piecewise linear characteristic* $\mathcal{G}^i$ *is as depicted in Figure 4.7 with* $e_2^i > e_1^i$ *for each* $i \in \overline{m}$*. Suppose that* $G(\sigma)$ *is a* $\mathcal{P}$*-matrix for all sufficiently large* σ*. Then, there exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *for all initial states* x_0*.*

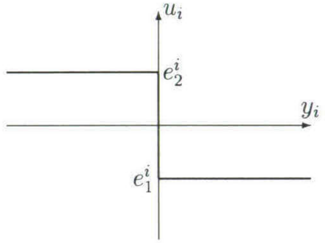


Figure 4.7: Relay characteristic

We can take one step further and consider relays with dead zone. The next corollary shows that the condition presented for relay systems is also sufficient for the well-posedness of linear systems coupled with relays having dead zone.

Corollary 4.6.3 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *where the piecewise linear characteristic* $\mathcal{G}^i$ *is as depicted in Figure 4.8 with* $0 \leq e_2^i > e_1^i \leq 0$ *and* $f_1^i > f_2^i$ *for each* $i \in \overline{m}$*. Suppose that* $G(\sigma)$ *is a* $\mathcal{P}$*-matrix for all sufficiently large* σ*. Then, there exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *for all initial states* x_0*.*

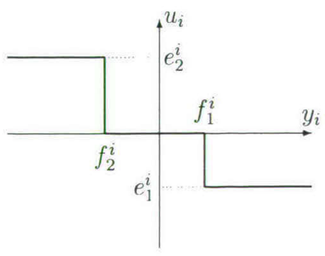


Figure 4.8: Relay with deadzone characteristic

4.6.3 Linear systems with saturation

Consider the single-input single-output system

$$\dot{x} = Ax + Bu \tag{4.18}$$

$$y = Cx \tag{4.19}$$

where u and y restricted by the saturation characteristic $\mathcal{G}$ depicted in Figure 4.9. The variable u is a piecewise linear function of the variable y, i.e., the function $y \mapsto u$ is Lipschitz continuous. Therefore, one can guarantee existence and uniqueness of the solutions of $\text{PLS}(A, B, C, 0, \mathcal{G})$ by employing standard results from the theory of ordinary differential equations. However, the presence of a feedthrough term D makes it impossible to employ such Lipschitz continuity arguments. In fact, one can find ill-posed examples for this case like the following.

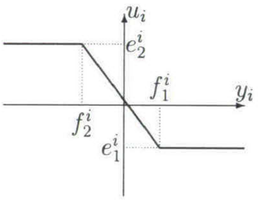


Figure 4.9: Saturation characteristic

Example 4.6.4 Consider the single-input single-output system

$$\dot{x} = u \tag{4.20}$$

$$y = x - 2u \tag{4.21}$$

where u and y restricted by a saturation characteristic $\mathcal{G}$ with $e_1 = -f_1 = -e_2 = f_2 = \frac{1}{2}$ shown in Figure 4.9. Let the periodic function $\tilde{u} : \mathbb{R}_+ \to \mathbb{R}$ be defined by

$$\tilde{u}(t) = \begin{cases} 1/2 & \text{if } 0 \leq t < 1 \\ -1/2 & \text{if } 1 \leq t < 3 \\ 1/2 & \text{if } 3 \leq t < 4 \end{cases}$$

and $\tilde{u}(t-4) = \tilde{u}(t)$ whenever $t \geq 4$. By using this function define $\tilde{x} : \mathbb{R}_+ \to \mathbb{R}$ as

$$\tilde{x}(t) = \int_0^t \tilde{u}(s)\, ds,$$

and $\tilde{y} : \mathbb{R}_+ \to \mathbb{R}$ as

$$\tilde{y} = \tilde{x} - 2\tilde{u}.$$

It can be verified that $(-\tilde{u}, -\tilde{x}, -\tilde{y})$, $(0,0,0)$ and $(\tilde{u}, \tilde{x}, \tilde{y})$ are all solutions of PLS$(0, 1, 1, -2, \mathcal{G})$ with the zero initial state.

As illustrated in the example, the Lipschitz continuity argument does not work in general. The following corollary gives a sufficient condition for the well-posedness of linear systems with saturation characteristics.

Corollary 4.6.5 *Consider a piecewise linear system* PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *where the piecewise linear characteristic $\mathcal{G}^i$ is as depicted in Figure 4.9 with $e_2^i > e_1^i$ for each $i \in \overline{m}$. Let $R = \operatorname{diag}(e_2^i - e_1^i)$ and $S = \operatorname{diag}(f_2^i - f_1^i)$. Suppose that $G(\sigma)R - S$ is a $\mathcal{P}$-matrix for all sufficiently large σ. Then, there exists a unique solution on $[0, \infty)$ of* PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *for all initial states x_0.*

4.7 Conclusions

In this paper we have considered linear systems with piecewise linear characteristics that can be represented by horizontal complementarity variables. We have proposed a solution concept for this class of systems, and we have presented sufficient conditions under which solutions exist and are unique. In particular we have given, under some conditions, a characterization of the set of initial states for which a solution exists.

We have worked with the class of piecewise Bohl functions, which in a sense is tailored for piecewise linear systems without external inputs. The class of piecewise Bohl functions may however be too small for some applications. A recent paper [18] reports existence and uniqueness results in a larger function space for linear systems with a single relay. To provide the same type of results for arbitrary piecewise linear characteristics is an open problem.

The systems considered in this paper are "closed" dynamical systems (i.e. systems without external variables), even though they are constructed with the aid of an "open" linear system and in fact we made extensive use of input/output system theory. Of course it would be of interest to consider piecewise linear systems with additional external variables, such as would be obtained by taking a linear system and connecting some but not all of its inputs and outputs by means of a piecewise linear relation. As an example, the i/o relation $y(t) = \max_{\tau \leq t} u(\tau)$ can be realized (assuming proper initialization) by a system that is obtained in this way. More generally, it might be asked which input/output relationships can be realized by means of piecewise linear systems with external variables.

Given that one has established existence and uniqueness of solutions, a natural next question is how to compute these solutions. Numerical procedures may be constructed on the basis of locating the points in time where transfer to another branch of a characteristic takes place, and re-starting the integration with the new data at each such time point ("event tracking schemes"). When there are many switches between branches this method may become awkward. There are indications that schemes may be devised that will asymptotically (as the time step goes to zero) converge to the true solution, even when no attempt is made to locate the switch times from one branch to another. Such a consistency result has been recently proven under a passivity assumption for systems with ideal diode characteristics [4], and a similar result has been obtained for relay systems in [9]. Extensions to arbitrary piecewise linear systems are currently under investigation.

4.8 Proofs

4.8.1 Some Lipschitzian results on HLCP

This subsection is devoted to Lipschitzian properties of HLCP. It is known that the solutions of LCP have the upper Lipschitzian property as shown in [5, Theorem 7.2.1]. Moreover, the solution is even a Lipschitz continuous function of the problem data under certain assumptions (see [5, Theorem 7.3.10]). In what follows, we will extend the Lipschitz continuity property to HLCP. We denote $\|\mathrm{col}(z^1, z^2, \dots, z^k)\|$ by $\|[z]_{j=1}^k\|$ for simplicity.

Theorem 4.8.1 *Assume that $[M^i]_{i=1}^k \subset \mathbb{R}^{m\times m}$ has the column $\mathcal{W}$-property. The function $q \mapsto [z^i]_{i=1}^k$ where $[z^i]_{i=1}^k$ is the unique solution of HLCP($q, [M^i]_{i=1}^k$) is Lipschitz continuous with the Lipschitz constant $d([M^i]_{i=1}^k)$ given by $d([M^i]_{i=1}^k) := \max\limits_{l\in \overline{k}^m} \|\{([M^i]_{i=1}^k)^l\}^{-1}\|$.*

define the sets $\mathcal{Z}^l \subset \mathcal{HC}_k^m$ and $\mathcal{Q}^l \subset \mathbb{R}^m$ as

$$\mathcal{Z}^l = \{[z^i]_{i=1}^k \in \mathcal{HC}_k^m \mid [z_j^i]_{i=1}^k \in \zeta^{l_j} \text{ for } j = 1, 2, \dots, m\} \tag{4.22}$$

$$\mathcal{Q}^l = \{q \in \mathbb{R}^m \mid q = M^1 z^1 - M^2 z^2 - M^3 z^3 - \dots - M^k z^k \text{ for some } [z^i]_{i=1}^k \in \mathcal{Z}^l\} \tag{4.23}$$

where $[\zeta^i]_{i=1}^k$ is as in Proposition 4.3.4. Suppose that $[z^i]_{i=1}^k$ is the unique solution of HLCP($q, [M^i]_{i=1}^k$) for some $q \in \mathcal{Q}^l$ with $l \in \overline{k}^m$. Then, we have

$$q = M^1 z^1 - M^2 z^2 - M^3 z^3 - \dots - M^k z^k. \tag{4.24}$$

$j = 1, 2, \ldots, k$. It follows that

$$z^i_{K_j} = \begin{cases} 0 & \text{if } i < j \text{ and } i = 1, \\ e & \text{if } i < j \text{ and } i \geq 2, \\ 0 & \text{if } i > j. \end{cases} \tag{4.25}$$

By substituting the above equations into (4.24), we get

$$q = M^1_{\bullet K_1} z^1_{K_1} - M^2_{\bullet K_2} z^2_{K_2} - M^3_{\bullet K_3} z^3_{K_3} - \cdots - M^k_{\bullet K_k} z^k_{K_k} - \sum_{i=2}^{j-1} \sum_{j=3}^{k} M^i_{\bullet K_j} e_{K_j}. \tag{4.26}$$

Consequently,

$$\begin{pmatrix} M^1_{\bullet K_1} & -M^2_{\bullet K_2} & -M^3_{\bullet K_3} & \cdots & -M^k_{\bullet K_k} \end{pmatrix} \begin{pmatrix} z^1_{K_1} \\ z^2_{K_2} \\ z^3_{K_3} \\ \vdots \\ z^k_{K_k} \end{pmatrix} = q + \sum_{i=2}^{j-1} \sum_{j=3}^{k} M^i_{\bullet K_j} e_{K_j}. \tag{4.27}$$

Note that $K_i \cap K_j = \emptyset$ if $i \neq j$ and $\bigcup_{j=1}^{k} K_i = \overline{m}$. It follows from the fact that $[M^i]_{i=1}^k$ has the column $\mathcal{W}$-property that the matrix $\begin{pmatrix} M^1_{\bullet K_1} & -M^2_{\bullet K_2} & -M^3_{\bullet K_3} & \cdots & -M^k_{\bullet K_k} \end{pmatrix}$ is invertible. Hence, (4.27) can be written as

$$\begin{pmatrix} z^1_{K_1} \\ z^2_{K_2} \\ z^3_{K_3} \\ \vdots \\ z^k_{K_k} \end{pmatrix} = \begin{pmatrix} M^1_{\bullet K_1} & -M^2_{\bullet K_2} & -M^3_{\bullet K_3} & \cdots & -M^k_{\bullet K_k} \end{pmatrix}^{-1} (q + \sum_{i=2}^{j-1} \sum_{j=3}^{k} M^i_{\bullet K_j} e_{K_j}). \tag{4.28}$$

The equations (4.25) and (4.28) imply that the function $q \mapsto [z^j]_{j=1}^k$ is affine on the set $\mathcal{Q}^l$. The column $\mathcal{W}$-property of $[M^i]_{i=1}^k$ implies from Theorem 4.4.4 that for each $q \in \mathbb{R}^m$ there exists a solution of HLCP$(q, [M^i]_{i=1}^k)$, i.e., $\bigcup_{l \in \overline{k}^m} \mathcal{Q}^l = \mathbb{R}^m$ and this solution is unique, i.e., $(\mathcal{Q}^{l^1} \cap \mathcal{Q}^{l^2})^\circ = \emptyset$ if $l^1 \neq l^2$. Furthermore, uniqueness of solutions implies that the function $q \mapsto [z^j]_{j=1}^k$ is continuous. Then, the claim of the theorem follows from Lemma 2.5.1 since

$$\| \begin{pmatrix} M^1_{\bullet K_1} & -M^2_{\bullet K_2} & -M^3_{\bullet K_3} & \cdots & -M^k_{\bullet K_k} \end{pmatrix} \| = \| \begin{pmatrix} M^1_{\bullet K_1} & M^2_{\bullet K_2} & M^3_{\bullet K_3} & \cdots & M^k_{\bullet K_k} \end{pmatrix} \|.$$

■

In particular, we will need to establish lower bounds for certain transfer matrices. The lemma below gives such a bound for low index transfer matrices.

Lemma 4.8.2 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^k$ *and* $G(s) = D + C(sI - A)^{-1}B$. *Suppose that* $[G(s)R^j - S^j]_{j=1}^k$ *is of index* 1. *Then, there exists a real number* α *such that for all sufficiently large* σ

$$d([G(\sigma)R^j - S^j]_{j=1}^k) \leq \alpha\sigma. \tag{4.29}$$

Proof: By hypothesis, we know that $s^{-1}\{([G(s)R^j - S^j]_{j=1}^k)^l\}^{-1}$ is proper for all $l \in \overline{m}^k$. Hence, $\|\{([G(\sigma)R^j - S^j]_{j=1}^k)^l\}^{-1}\| \leq \alpha_l\sigma$ for all sufficiently large σ. Clearly, (4.29) holds for $\alpha = \max_{l \in \overline{k}^m} \alpha_l$. ■

4.8.2 On the invertibility of rational matrices

In this subsection, we state the following lemma on the invertibility of rational matrices which will be employed later on.

Lemma 4.8.3 *Consider a matrix quadruple* (A, B, C, D)*such that* $G(s) = D + C(sI - A)^{-1}B$ *is invertible as a rational matrix. Suppose that the function pair* $(u, x) \in \mathcal{F}([0, T], \mathbb{R}^{m+n})$ *where* x *is differentiable satisfies*

$$\dot{x} = Ax + Bu + e \tag{4.30a}$$

$$0 = Cx + Du + f \tag{4.30b}$$

for some $e \in \mathbb{R}^n$ *and* $f \in \mathbb{R}^m$. *Then,* x *is uniquely determined, and there exist a matrix* $K \in \mathbb{R}^{m \times n}$ *and a vector* $l \in \mathbb{R}^m$ *both depending only in* (A, B, C, D, e, f) *such that* $u = Kx + l$.

Proof: It follows from [8, Theorem 3.24]. ■

4.8.3 Initial solutions and their characterizations

To prove Theorem 4.5.4, we shall first define the notion of *initial solution* for piecewise linear systems.

Definition 4.8.4 A triple $(u, x, y) \in \mathcal{B}^{m+n+m}$ is said to be an *initial solution* of $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0 if the following conditions hold.

1. The triple (u, x, y) satisfies

$$\dot{x} = Ax + Bu\ ,\ x(0) = x_0,$$
$$y = Cx + Du$$

2. For each $i = 1, 2, \dots, m$,

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \mathcal{G}^i \text{ for all sufficiently small } t.$$

We shall derive sufficient conditions for existence and uniqueness of initial solutions. To do this, we first need some preparations. There are several definitions in order.

Definition 4.8.5 The family of continuous functions $[f^i]_{i=1}^n \subset \mathcal{C}^m$ is said to be *initially k-complementary* if the following conditions hold.

1. For all sufficiently small t,

$$0 \leq f^1(t)$$
$$0 \leq f^i(t) \leq e \text{ for } i = 2, 3, \dots, n-1$$
$$0 \leq f^n(t).$$

2. For all $t \in \mathbb{R}_+$,

$$(f^1(t))^\top f^2(t) = 0$$
$$(e - f^i(t))^\top f^{i+1}(t) = 0 \text{ for } i = 2, 3, \dots, n.$$

Lemma 4.8.6 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Assume that all column representatives of* $[G(s)R^j - S^j]_{j=1}^k$ *are invertible as a rational matrix and the triple* (u, x, y) *is an initial solution of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *with some initial state. Then the following statements hold.*

1. *Let* $[\mathcal{G}_j^i]_{j=1}^k$ *be as in Definition 4.3.1 for each* $i = 1, 2, \dots, m$. *Then, there exists* $l \in \overline{k}^m$ *such that*

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \operatorname{affn} \mathcal{G}_{l_i}^i, \text{ for each } i = 1, 2, \dots, m \text{ and } t \in \mathbb{R}_+.$$

2. *Let* $l \in \overline{k}^m$ *be as in the previous item.*

(a) There exist vectors $\bar{u}^l$, $\bar{y}^l \in \mathbb{R}^m$ *and* $z \in \mathcal{B}^m$ *such that*

$$u = \bar{u}^l + \mathcal{R}^l z$$
$$y = \bar{y}^l + \mathcal{S}^l z$$

where $\mathcal{R} = [-R^1, R^2, R^3, \ldots, R^k]$ *and* $\mathcal{S} = [-S^1, S^2, S^3, \ldots, S^k]$.

(b) There exist initially k*-complementary Bohl functions* $[z^j]_{j=1}^k \subset \mathcal{B}^m$ *such that*

$$u = q^u - R^1 z^1 + R^2 z^2 + R^3 z^3 + \cdots + R^k z^k$$
$$y = y^u - S^1 z^1 + S^2 z^2 + S^3 z^3 + \cdots + S^k z^k.$$

(c) There exist matrices $F^l \in \mathbb{R}^{n\times n}$ *and* $G^l \in \mathbb{R}^{m\times n}$, *and vectors* $v^l \in \mathbb{R}^n$ *and* $w^l \in \mathbb{R}^m$ *depending only on* l *such that*

$$\dot{x} = F^l x + v^l$$
$$u = G^l x + w^l.$$

(d) For a given $T > 0$, *there exists* α^l *depending only on* l *and* T *such that*

$$\|x(t) - x(s)\| \leq \alpha^l \|t - s\|$$

for all $t, s \in [0, T]$.

To prove Lemma 4.8.6, we need the following technical proposition.

Proposition 4.8.7 *Let* $\mathcal{G} \subset \mathbb{R}^2$ *be an affine set. There exist real numbers* α, β, *and* γ *such that*

$$\begin{pmatrix} v \\ w \end{pmatrix} \in \mathcal{G} \Leftrightarrow \alpha v + \beta w + \gamma = 0.$$

Proof: Evident. ■

Proof of Lemma 4.8.6:

1: Since they are Bohl functions, both u_i and y_i are continuous. It follows from Definition 4.8.4 item 2 together with continuity that for each $i \in \overline{m}$ there exists $l_i \in \overline{k}$ such that

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \mathcal{G}^i_{l_i} \text{ for all } t \in [0, \epsilon)$$

for some $\epsilon > 0$. Since $\mathcal{G}^i_{l_i} \subseteq \text{affn}\, \mathcal{G}^i_{l_i}$, we have

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \text{affn}\mathcal{G}^i_{l_i} \text{ for all } t \in [0, \epsilon).$$

Then, it follows from Proposition 4.8.7 that for each $i = 1, 2, \ldots, m$ there exist real numbers α^i, β^i and γ^i such that $\alpha^i u_i(t) + \beta^i y_i(t) + \gamma^i = 0$ for $t \in [0, \epsilon)$. The real-analyticity of Bohl functions implies that $\alpha^i u_i(t) + \beta^i y_i(t) + \gamma^i = 0$ for $t \in \mathbb{R}_+$. Hence,

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} \in \text{affn}\, \mathcal{G}^i_{l_i}, \text{ for each } i = 1, 2, \ldots, m \text{ and } t \in \mathbb{R}_+.$$

2a: Define the sets $[\xi^i]_{i=1}^k$

$$\xi^1 = \{[z^i]_{i=1}^k \subset \mathbb{R} \mid z^2 = z^3 = \cdots = z^k = 0\}, \tag{4.31a}$$

$$\xi^k = \{[z^i]_{i=1}^k \subset \mathbb{R} \mid z^1 = 0,\, z^2 = z^3 = \cdots = z^{k-1} = 1\}, \tag{4.31b}$$

$$\xi^j = \{[z^i]_{i=1}^k \subset \mathbb{R} \mid z^i = \begin{cases} 0, & i = 1 \\ 1, & i = 2, 3, \ldots, j-1 \\ 0, & i = j+1, j+2, \ldots, k \end{cases} \}. \tag{4.31c}$$

Note that they are similar to ζ^J's as defined in (4.7) but without inequalities. Define also the sets $\mathcal{Y}^l = \{[z^j]_{j=1}^k \subset \mathbb{R}^m \mid [z^i_j]_{i=1}^k \in \xi^{l_j} \text{ for } j = 1, 2, \ldots, m\}$. Let $\mathcal{Z}^l$ be defined as in (4.22). It follows from definition of horizontal complementarity representations that

$$\mathcal{G}^i_{l_i} = \{ \begin{pmatrix} q^u_i \\ q^y_i \end{pmatrix} - \begin{pmatrix} R^1_{ii} \\ S^1_{ii} \end{pmatrix} z^1_i + \sum_{j=2}^k \begin{pmatrix} R^j_{ii} \\ S^j_{ii} \end{pmatrix} z^j_i \mid [z^j]_{j=1}^k \in \mathcal{Z}^l \}. \tag{4.32}$$

Moreover, it can be verified that

$$\text{affn}\, \mathcal{G}^i_{l_i} = \{ \begin{pmatrix} q^u_i \\ q^y_i \end{pmatrix} - \begin{pmatrix} R^1_{ii} \\ S^1_{ii} \end{pmatrix} z^1_i + \sum_{j=2}^k \begin{pmatrix} R^j_{ii} \\ S^j_{ii} \end{pmatrix} z^j_i \mid [z^j]_{j=1}^k \in \mathcal{Y}^l \}. \tag{4.33}$$

Then, Lemma 4.8.6 item 1 implies that there exist functions $z^j : \mathbb{R} \to \mathbb{R}^m$ such that

$$\begin{pmatrix} u_i(t) \\ y_i(t) \end{pmatrix} = \begin{pmatrix} q^u_i \\ q^y_i \end{pmatrix} - \begin{pmatrix} R^1_{ii} \\ S^1_{ii} \end{pmatrix} z^1_i(t) + \sum_{j=2}^k \begin{pmatrix} R^j_{ii} \\ S^j_{ii} \end{pmatrix} z^j_i(t) \tag{4.34a}$$

$$[z^j(t)]_{j=1}^k \in \mathcal{Y}^l \tag{4.34b}$$

for all $t \in \mathbb{R}_+$. Note that the functions z_i^j with $j \neq l_i$ are constant functions due to the definition of the set $\mathcal{Y}^l$. Define the function $z : \mathbb{R} \to \mathbb{R}^m$, and vectors $\bar{u}^l$ and $\bar{y}^l$ as

$$z = \begin{pmatrix} z_1^{l_1} \\ \vdots \\ z_m^{l_m} \end{pmatrix}, \; \bar{u}^l = q^u + \begin{pmatrix} \sum_{j=2}^{l_1-1} R_{11}^j \\ \vdots \\ \sum_{j=2}^{l_m-1} R_{mm}^j \end{pmatrix}, \text{ and } \bar{y}^l = q^y + \begin{pmatrix} \sum_{j=2}^{l_1-1} S_{11}^j \\ \vdots \\ \sum_{j=2}^{l_m-1} S_{mm}^j \end{pmatrix}.$$

One can check that (4.34) yields that

$$u = \bar{u}^l + \mathcal{R}^l z \tag{4.35a}$$

$$y = \bar{y}^l + \mathcal{S}^l z. \tag{4.35b}$$

It remains to prove that z is a Bohl function. It follows from (4.35) that the pair (z, x) satisfies

$$\dot{x} = Ax + B\mathcal{R}^l z + B\bar{u}^l \tag{4.36a}$$

$$0 = Cx + (D\mathcal{R}^l - \mathcal{S}^l)z + D\bar{u}^l - \bar{y}^l \tag{4.36b}$$

Since $G(s)\mathcal{R}^l - \mathcal{S}^l$ is a column representative of $[G(s)R^j - S^j]_{j=1}^k$, it is invertible as a rational matrix due to the hypothesis. Consequently, Lemma 4.8.3 implies that $z = E^l x + o^l$ for some E^l and o^l. This implies together with (4.36a) that x is Bohl and hence so is z.

2b: It has already been shown in the proof of previous item that the function z is Bohl. For each $j \in \overline{m}$ define

$$z_j^1 = \begin{cases} z_j & \text{if } l_j = 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.37}$$

$$z_j^i = \begin{cases} 0 & \text{if } l_j < i \\ z_j & \text{if } l_j = i \\ 1 & \text{otherwise} \end{cases} \quad \text{for } i = 2, 3, \ldots, k-1 \tag{4.38}$$

$$z_j^k = \begin{cases} z_j & \text{if } l_j = k \\ 0 & \text{otherwise} \end{cases} \tag{4.39}$$

where z is as in the previous item. Clearly, (4.34) holds. Since (u, x, y) is an initial solution, we know $\operatorname{col}(u_i(t), y_i(t)) \in \mathcal{G}_{l_i}^i$ for each $i \in \overline{m}$ and for all sufficiently small t. It follows from Fact 4.3.5 that $[z_i^j(t)]_{j=1}^k \in \zeta^{l_i}$ for all sufficiently small t where ζs defined as in (4.7). Consequently, $[z^i]_{i=1}^k$ is initially k-complementary.

2c: The matrices F^l, G^l, v^l, and w^l can be given as $F^l = A + B\mathcal{R}^l E^l$, $G^l = \mathcal{R}^l E^l$, $v^l = B\bar{u}^l + B\mathcal{R}^l o^l$, and $w^l = \bar{u}^l + \mathcal{R}^l o^l$ by substituting z into (4.35a) and (4.36a).

2d: From the previous item, it is known that x satisfies $\dot{x} = F^l x + v^l$ for some $F^l \in \mathbb{R}^{n\times n}$ and $v^l \in \mathbb{R}^n$. Since x is continuous, it is bounded on every finite interval $[0,T]$. It follows that $\dot{x}$ is also bounded on the interval $[0,T]$. Therefore, it is Lipschitz continuous on $[0,T]$ with a Lipschitz constant depending on only l and T. ■

By following the footsteps of the characterization of the initial solutions of linear complementarity systems in [10, 11], we define the *horizontal* version of the *rational complementarity problem.*

Problem 4.8.8 (HRCP($q(s), [M^i(s)]_{i=1}^k$)) Given $q(s) \in \mathbb{R}^n(s)$ and $[M^i(s)]_{i=1}^k \subset \mathbb{R}^{n\times n}(s)$, find $[z^i(s)]_{i=1}^k \subset \mathbb{R}^n(s)$ such that the following conditions hold.

1. $M^1(s)z^1(s) = q(s) + \sum_{i=2}^k M^i(s)z^i(s)$.

2. For all $s \in \mathbb{C}$,

$$
\begin{aligned}
z^1(s) &\perp z^2(s) \\
(s^{-1}e - z^i(s)) &\perp z^{i+1}(s) \text{ for } i = 2,3,\dots,k.
\end{aligned}
$$

3. For all sufficiently large σ,

$$
\begin{aligned}
0 &\leq z^1(\sigma) \\
0 \leq z^i(\sigma) &\leq e\sigma^{-1} \text{ for } i = 2,3,\dots,k-1 \\
0 &\leq z^k(\sigma).
\end{aligned}
$$

Notice that the conditions 3 imply that $z^i(s)$ is strictly proper for $i = 2,3,\dots,k-1$.

The initial solutions of piecewise linear system can be characterized by the strictly proper solutions of corresponding HRCPs as stated in the following lemma.

Lemma 4.8.9 *Consider a piecewise linear system* PLS($A,B,C,D,[\mathcal{G}^i]_{i=1}^m$). *Let* ($q^u, q^y$, $[R^j]_{j=1}^k, [S^j]_{j=1}^k$) *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *The following statements are equivalent:*

1. PLS($A,B,C,D,[\mathcal{G}^i]_{i=1}^m$) *has an initial solution with the initial state* x_0.

2. HRCP$(C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$ has a strictly proper solution.

To prove Lemma 4.8.9, we need the following technical lemma.

Lemma 4.8.10 *The family of Bohl functions $[f^i]_{i=1}^n \subset \mathcal{B}^m$ is initially k-complementary if and only if their Laplace transforms $[\hat{f}^i(s)]_{i=1}^n \subset \mathbb{R}^m(s)$ satisfy the following conditions.*

1. *For all $s \in \mathbb{C}$,*

$$\begin{aligned} \hat{f}^1(s) &\perp \hat{f}^2(s) \\ (s^{-1}e - \hat{f}^i(s)) &\perp \hat{f}^{i+1}(s) \text{ for } i = 2, 3, \dots, k. \end{aligned}$$

2. *For all sufficiently large σ,*

$$\begin{aligned} 0 &\le \hat{f}^1(\sigma) \\ 0 \le \hat{f}^i(\sigma) &\le e\sigma^{-1} \text{ for } i = 2, 3, \dots, k-1 \\ 0 &\le \hat{f}^k(\sigma). \end{aligned}$$

Proof: It follows directly from the initial value theorem of Laplace transformation. ■

Proof of Lemma 4.8.9:

$1 \Rightarrow 2$: Let (u, x, y) be an initial solution of PLS. It follows from Lemma 4.8.6 item 2b that there exist initially k-complementary Bohl functions $[z^j]_{j=1}^k$ such that

$$u = q^u - R^1 z^1 + R^2 z^2 + R^3 z^3 + \dots + R^k z^k \tag{4.40a}$$

$$y = q^y - S^1 z^1 + S^2 z^2 + S^3 z^3 + \dots + S^k z^k. \tag{4.40b}$$

Lemma 4.8.10 implies that the Laplace transforms of $[z^j]_{j=1}^k$, $[\hat{z}^j(s)]_{j=1}^k$ satisfy items 2 and 3 of Problem 4.8.8. On the other hand, the Laplace transform of (u, y), $(\hat{u}(s), \hat{y}(s))$, satisfies $\hat{y}(s) = C(sI-A)^{-1}x_0 + G(s)\hat{u}(s)$. This equation together with the Laplace domain versions of (4.40) results in

$$[G(s)R^1 - S^1]\hat{z}^1(s) = C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y + \sum_{j=2}^{k}[G(s)R^j - S^j]\hat{z}^j(s).$$

Hence, $[\hat{z}^j(s)]_{j=1}^k$ is a solution of HRCP$(C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$. It is clear that $[\hat{z}^j(s)]_{j=1}^k$ is strictly proper since these functions are Laplace transforms of Bohl functions.

$2 \Rightarrow 1$: Let $[\hat{z}^j(s)]_{j=1}^k$ be a strictly proper solution of HRCP$(C(sI-A)^{-1}x_0+s^{-1}G(s)q^u-s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$. Let $[z^j]_{j=1}^k$ denote the inverse Laplace transform of $[\hat{z}^j(s)]_{j=1}^k$. Define $u = q^u - R^1z^1 + R^2z^2 + R^3z^3 + \cdots + R^kz^k$ and $y = q^y - S^1z^1 + S^2z^2 + S^3z^3 + \cdots + S^kz^k$. Since $[\hat{z}^j(s)]_{j=1}^k$ satisfies

$$[G(s)R^1 - S^1]\hat{z}^1(s) = C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y + \sum_{j=2}^k [G(s)R^j - S^j]\hat{z}^j(s),$$

the Laplace transform of (u, y), $(\hat{u}, \hat{y})$ satisfies $\hat{y}(s) = C(sI-A)^{-1}x_0 + G(s)\hat{u}(s)$. Define $\hat{x}(s) = (sI-A)^{-1}x_0 + (sI-A)^{-1}B\hat{u}(s)$. It can be easily checked that (u, x, y) is an initial solution of PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0 where x denotes the inverse Laplace transform of $\hat{x}(s)$. ■

Based on the results of [10], we can make a connection between HRCPs and parametrized families of HLCPs.

Theorem 4.8.11 *Consider a piecewise linear system* PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Then the statements 1 and 3 are equivalent, and so are the statements 2 and 4.*

1. *HRCP$(C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$ is solvable.*

2. *HRCP$(C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$ is uniquely solvable.*

3. *HLCP$(\sigma C(\sigma I-A)^{-1}x_0 + G(\sigma)q^u - q^y, [G(\sigma)R^j - S^j]_{j=1}^k)$ is solvable for all sufficiently large σ.*

4. *HLCP$(\sigma C(\sigma I-A)^{-1}x_0 + G(\sigma)q^u - q^y, [G(\sigma)R^j - S^j]_{j=1}^k)$ is uniquely solvable for all sufficiently large σ.*

Proof:

$1 \Leftrightarrow 3$: It follows from Remark 4.4.5 and [10, Theorem 4.1].

$2 \Leftrightarrow 4$: It follows from Remark 4.4.5 and [10, Corollary 4.10]. ■

It is already known from Lemma 4.8.9 that strictly proper solutions of HRCP play a key role in the analysis of initial solutions. The following theorem establishes an equivalence between strictly proper solvability of a HRCP and solvability of a HLCP under the assumption that HRCP is uniquely solvable.

Theorem 4.8.12 *Consider a piecewise linear system* PLS$(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Suppose that* $[G(\sigma)R^j - S^j]_{j=1}^k$ *has the column* $\mathcal{W}$*-property for all sufficiently large* σ. *Then the following statements hold.*

1. *Assume that* $[G(s)R^j - S^j]_{j=1}^k$ *is of index* 1. *The following two statements are equivalent.*
   (a) *HRCP(*$C(sI-A)^{-1}x_0+s^{-1}G(s)q^u-s^{-1}q^y, [G(s)R^j-S^j]_{j=1}^k$*) has a strictly proper solution.*
   (b) *HLCP(*$Cx_0 + Dq^u - q^y, [DR^j - S^j]_{j=1}^k$*) has a solution.*
2. *If* $[DR^1 - S^1, DR^k - S^k]$ *is nondegenerate then HRCP(*$C(sI-A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k$*) has a strictly proper solution for all initial states* x_0.

Proof:

1a $\Rightarrow$ *1b:* Let $[z^j(s)]_{j=1}^k$ be a strictly proper solution of HRCP$(C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$, i.e., $[z^j(s)]_{j=1}^k$ satisfies the items 2 and 3 of Problem 4.8.8, and

$$(G(s)R^1 - S^1)z^1(s) = C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y + \sum_{j=2}^{k}(G(s)R^j - S^j)z^j(s) \tag{4.41}$$

for all $s \in \mathbb{C}$. Define $[\bar{z}^j]_{j=1}^k = \lim_{s\to\infty}[sz^j(s)]_{j=1}^k$. It follows from the items 2 and 3 of Definition 4.8.8 that $[\bar{z}^j]_{j=1}^k$ is k-horizontal complementary. By multiplying (4.41) by s and letting s tend to ∞, we get

$$(DR^1 - S^1)\bar{z}^1 = Cx_0 + Dq^u - q^y + \sum_{j=2}^{k}(DR^j - S^j)\bar{z}^j.$$

Consequently, $[\bar{z}^j]_{j=1}^k$ is a solution of HLCP$(Cx_0 + Dq^u - q^y, [DR^j - S^j]_{j=1}^k)$.

1b $\Rightarrow$ *1a:* Observe that we have the following two facts.

i. Since $[G(\sigma)R^j-S^j]_{j=1}^k$ has the column $\mathcal{W}$-property for all sufficiently large σ, HLCP$(C(\sigma I - A)^{-1}x_0 + \sigma^{-1}G(\sigma)q^u - \sigma^{-1}q^y, [G(\sigma)R^j - S^j]_{j=1}^k)$ is uniquely solvable for all sufficiently large σ. Hence, it follows from Lemma 4.8.11 that HRCP$(C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$ has a unique solution, say $[z^j(s)]_{j=1}^k$. Clearly, $[\sigma z^j(\sigma)]_{j=1}^k$ is a solution of HLCP$(\sigma C(\sigma I - A)^{-1}x_0 + G(\sigma)q^u - q^y, [G(\sigma)R^j - S^j]_{j=1}^k)$ for all sufficiently large σ.

ii. Let $[\bar{z}^j]_{j=1}^k$ be a solution of HLCP$(Cx_0 + Dq^u - q^y, [DR^j - S^j]_{j=1}^k)$. Clearly, it is also a solution of HLCP$(Cx_0 + Dq^u - q^y + \sum_{j=1}^k H^j(\sigma)\bar{z}^j, [G(\sigma)R^j - S^j]_{j=1}^k)$ where $H^1(s) = (G(s) - D)R^1$ and $H^j(s) = (D - G(s))R^j$ for $j = 2, 3, \ldots, k$.

By using Lemma 4.8.2, Theorem 4.8.1 and the triangle inequality, we get

$$\|[\sigma z^j(\sigma) - \bar{z}^j]_{j=1}^k\| \leq \alpha\sigma(\|\sigma C(\sigma I - A)^{-1}x_0 - \sigma^{-1}Cx_0\| \\ + \|G(\sigma)q^u - Dq^u\| + \sum_{j=2}^{k} \|H^j(\sigma)\|\|\bar{z}^j\|)$$

for all sufficiently large σ. Note that the right hand side of this inequality converges to a constant term as σ tends to infinity. This implies that $[z^j(s)]_{j=1}^k$ is strictly proper.

2: Suppose that $[DR^1 - S^1, DR^k - S^k]$ is nondegenerate but the solution of HRCP$(C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y, [G(s)R^j - S^j]_{j=1}^k)$, $[z^j(s)]_{j=1}^k$ is not strictly proper for some x_0. This means that $[z^1(s), z^k(s)]$ is not strictly proper since $[z^j(s)]_{j=2}^{k-1}$ is strictly proper by definition of Problem 4.8.8. Let l be an integer such that $\lim_{s\to\infty} s^{-l}[z^1(s), z^k(s)] = [\bar{z}^1, \bar{z}^k] \neq 0$. Clearly, $l \geq 0$. Note that $[z^1(s), z^k(s)]$ is a solution of HRCP$(C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y + \sum_{j=2}^{k-1}(G(s)R^j - S^j)z^j(s), [G(s)R^1 - S^1, G(s)R^k - S^k])$. Hence, $\sigma^{-l}[z^1(\sigma), z^k(\sigma)]$ is a solution of HLCP$(\sigma^{-l}C(\sigma I - A)^{-1}x_0 + \sigma^{-l-1}G(\sigma)q^u - \sigma^{-l-1}q^y + \sigma^{-l}\sum_{j=2}^{k-1}(G(\sigma)R^j - S^j)z^j(\sigma), [G(\sigma)R^1 - S^1, G(\sigma)R^k - S^k])$. Since $[z^j(s)]_{j=2}^{k-1}$ is strictly proper, it follows that $[\bar{z}^1, \bar{z}^k]$ is a solution of HLCP$(0, [DR^1 - S^1, DR^k - S^k])$. Then, we have

$$(DR^1 - S^1)\bar{z}^1 = (DR^k - S^k)\bar{z}^k. \tag{4.42}$$

Note that $(\bar{z}^1)^\top \bar{z}^k = 0$. Define the index sets J, K as $J = \{j \mid \bar{z}^1_j \neq 0\}$ and $K = \{j \mid j \notin J\}$. The equation (4.42) can be written as

$$\begin{pmatrix}(DR^1 - S^1)_{\bullet J} & (DR^k - S^k)_{\bullet K}\end{pmatrix}\begin{pmatrix}\bar{z}^1_J \\ -\bar{z}^k_K\end{pmatrix} = 0.$$

Note that the matrix on the left hand side is a column representative of $[DR^1 - S^1, DR^k - S^k]$ and hence nonsingular by the hypothesis. Then, $\bar{z}^1 = \bar{z}^2 = 0$ which contradicts the definition of the integer l. ∎

4.8.4 Proof of Theorem 4.5.4

We begin with the following lemma which is the last piece of preparations to prove the main result.

Lemma 4.8.13 *Consider a piecewise linear system* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$. *Let* $(q^u, q^y, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ *be a horizontal complementarity representation of the piecewise linear characteristics* $[\mathcal{G}^j]_{j=1}^m$. *Suppose that* $[G(\sigma)R^j - S^j]_{j=1}^k$ *has the column* $\mathcal{W}$*-property for all sufficiently large* σ. *Suppose also that the set*

$$\mathcal{R} = \{x_0 \in \mathbb{R}^n \mid HRCP(q_{x_0}(s), [G(s)R^i - S^i]_{i=1}^k) \text{ has a strictly proper solution}\}$$

is closed where $q_{x_0}(s) = C(sI - A)^{-1}x_0 + s^{-1}G(s)q^u - s^{-1}q^y$. *Then, there exists a unique solution on* $[0, \infty)$ *of* $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ *with the initial state* x_0 *if and only if* $x_0 \in \mathcal{R}$.

Proof:

if: Let the initial state $\bar{x}$ be given such that $\bar{x} \in \mathcal{R}$. Hence, it follows from Theorem 4.8.12 and Lemma 4.8.9 that $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ has an initial solution with the initial state $\bar{x}$. Let $(u^{\bar{x}}, x^{\bar{x}}, y^{\bar{x}})$ denote this initial solution. We define $\iota : \mathbb{R}^n \to \overline{k}^m$ as

$$\iota(\bar{x}) = l$$

where l is as in Lemma 4.8.6 item 1 for the initial solution $(u^{\bar{x}}, x^{\bar{x}}, y^{\bar{x}})$, $\tau : \mathbb{R}^n \to \mathbb{R}$ as

$$\tau(\bar{x}) = \sup\{T \mid \begin{pmatrix} u_j^{\bar{x}}(t) \\ y_j^{\bar{x}}(t) \end{pmatrix} \in \mathcal{G}_{\iota(\bar{x})_j} \text{ for all } j \in \overline{m} \text{ and } t \in [0, T]\},$$

and $\kappa : \mathbb{R}^n \to \mathbb{R}^n$ as

$$\kappa(\bar{x}) = x^{\bar{x}}(\tau(\bar{x})).$$

Note that $t \mapsto (u^{\bar{x}}, x^{\bar{x}}, y^{\bar{x}})(t + \rho)$ forms an initial solution of $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ with the initial state $x^{\bar{x}}(\rho)$ whenever $\rho \in [0, \tau(\bar{x}))$. Hence, we have $x^{\bar{x}}(\rho) \in \mathcal{R}$ for all $\rho \in [0, \tau(\bar{x}))$. It follows from the closedness of the set $\mathcal{R}$ and continuity of $x^{\bar{x}}$ that $\kappa(\bar{x}) \in \mathcal{R}$.

existence: Define $x_{i+1} = \kappa(x_i)$ for $i = 0, 1, \ldots$. From the previous discussion, we know that $x_i \in \mathcal{R}$ and hence $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ admits initial solutions for all initial states x_i due to Lemma 4.8.9. Let $(u^{x_i}, x^{x_i}, y^{x_i})$ denote an initial solution with the initial state x_i. Define $\tau_k = \sum_{i=1}^k \tau(x_{k-1})$ for $k > 0$ and $\tau_0 = 0$. Also define

$$(u, x, y)|_{[\tau_k, \tau_{k+1}]} = (u^{x_k}, x^{x_k}, y^{x_k})|_{[0, \tau(x^k)]}.$$

It can be verified that (u, x, y) is a solution on $[0, T)$ for some $T > 0$ of $\mathrm{PLS}(A, B, C, D, [\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0. Suppose that T is such that there is no solution on $[0, T')$ whenever $T' > T$. However, Lemma 4.8.6 item 2c implies that x is Lipschitz continuous with the Lipschitz constant $\max_{l \in \overline{k}^m} \alpha^l$ where α^l is as in the same item. Hence, x is uniformly continuous on $[0, T)$ and $x^* := \lim_{t \to T^-} x(t)$ exists due to [19, exercise 4.13].

Since $x(t) \in \mathcal{R}$ for all $t \in [0,T)$ and x is continuous, $x^* \in \mathcal{R}$ which means one can extend the solution (u,x,y) beyond $[0,T)$ by using the initial solution of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the initial state x^*. This contradicts the definition of T. Thus, we can conclude that there exists a solution on $[0,\infty)$ of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0.

uniqueness: Let $(u^i,x^i,y^i) \in \mathcal{PB}^{m+n+m}$ for $i=1,2$ denote two solutions of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0. Clearly, $(u^1,x^1,y^1)-(u^2,x^2,y^2)$ is a piecewise Bohl function as well. If it is not identically zero then there should exist $t \geq 0$ and $\epsilon > 0$ such that $((u^1,x^1,y^1)-(u^2,x^2,y^2))|_{[0,t]} \equiv 0$ and $((u^1,x^1,y^1)-(u^2,x^2,y^2))(s) \neq 0$ for all $s \in (t,t+\epsilon)$ due to the definition of piecewise Bohl functions. For (u^i,x^i,y^i) and $t \geq 0$, one can find $\epsilon_i > 0$ and Bohl functions $(\bar{u}^i,\bar{x}^i,\bar{y}^i)$ such that $(u^i,x^i,y^i)|_{[t,t+\epsilon_i)} = (\bar{u}^i,\bar{x}^i,\bar{y}^i)|_{[0,\epsilon_i)}$ with $i=1,2$ again by the definition of piecewise Bohl functions. It is easy to see that $(\bar{u}^i,\bar{x}^i,\bar{y}^i)$ form two different initial solutions of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the same initial state, $x^1(t)=x^2(t)$. Then, Lemma 4.8.9 and Theorem 4.8.12 imply that HLCP$(C(\sigma I-A)^{-1}x^1(t)+s^{-1}G(\sigma)q^u-s^{-1}q^y,[G(\sigma)R^j-S^j]_{j=1}^k)$ has at least two different solutions for all sufficiently large σ which is ruled out by Theorem 4.4.4 since $[G(\sigma)R^j-S^j]_{j=1}^k$ has the column $\mathcal{W}$-property for all sufficiently large σ.

only if: Let $(u,x,y) \in \mathcal{PB}^{m+n+m}$ be the unique solution of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0. By the definition of piecewise Bohl functions, we know that there exists $\epsilon > 0$ and $(\bar{u},\bar{x},\bar{y}) \in \mathcal{B}^{m+n+m}$ such that $(u,x,y)|_{[0,\epsilon)} = (\bar{u},\bar{x},\bar{y})|_{[0,\epsilon)}$. Obviously, $(\bar{u},\bar{x},\bar{y})$ is an initial solution of PLS$(A,B,C,D,[\mathcal{G}^i]_{i=1}^m)$ with the initial state x_0. Hence, $x_0 \in \mathcal{R}$ due to Lemma 4.8.9. ■

Proof of Theorem 4.5.4:

1: Let $\mathcal{R}$ be defined as in Lemma 4.8.13. It follows from Theorem 4.8.12 item 1 that $\mathcal{R} = \{x_0 \mid HLCP(Cx_0+Dq^u-q^y,[DR^j-S^j]_{j=1}^k)$ has a strictly proper solution$\}$. Since the set $\{q \in \mathbb{R}^n \mid HLCP(q,[M^i]_{i=1}^k)$ is solvable$\}$ is closed, $\mathcal{R}$ is closed. Then, Lemma 4.8.13 proves the statement.

2: Let $\mathcal{R}$ be defined as in Lemma 4.8.13. It follows from Theorem 4.8.12 item 2 that $\mathcal{R}=\mathbb{R}^n$. Then, Lemma 4.8.13 proves the statement. ■

4.8.5 Proof of Theorem 4.5.5

To devise a proof of Theorem 4.5.5, we need some preparations. Three rather technical lemmas on piecewise linear characteristics are in order. The first one presents equivalent conditions for redundancy of a vertex.

Lemma 4.8.14 *Let $\mathcal{G} = \mathrm{plc}(d^-, [v^i]_{i=1}^{k-1}, d^+)$ be a k-piecewise linear characteristic and $\mathcal{G}_i$ be as in Definition 4.3.1 for $i \in \overline{k}$. Also let the vectors r and s be defined for the piecewise characteristic $\mathcal{G}$ in accordance with (4.9). The following statements are equivalent.*

1. *The vertex v^i is redundant.*
2. *The set $\mathcal{G}_i \cup \mathcal{G}_{i+1}$ is convex.*
3. *There exists $\alpha > 0$ such that $r_j = \alpha r_{j+1}$ and $s_j = \alpha s_{j+1}$.*

Proof:

1 ⇔ 2: Evident.

2 ⇒ 3: We prove the statement only for $1 \neq i \neq k-1$. The other two cases can be proven in a similar fashion. Note that $v^i = \mathcal{G}_i \cap \mathcal{G}_{i+1}$ and $\mathcal{G}_i = \{\lambda v^{i-1} + (1-\lambda)v^i \mid 0 \leq \lambda \leq 1\}$. Since $\mathcal{G}_i \cup \mathcal{G}_{i+1}$ is convex, we can conclude that $\mathcal{G}_i \cup \mathcal{G}_{i+1} = \{\lambda v^{i-1} + (1-\lambda)v^{i+1} \mid 0 \leq \lambda \leq 1\}$. By writing v^i as a convex combination of v^{i-1} and v^{i+1}, we get

$$v^i = \lambda v^{i-1} + (1-\lambda)v^{i+1}.$$

Hence,

$$\underbrace{v^i - v^{i-1}}_{\begin{pmatrix} r_i \\ s_i \end{pmatrix}} = \underbrace{\tfrac{1-\lambda}{\lambda}}_{>0} \underbrace{(v^{i+1} - v^i)}_{\begin{pmatrix} r_{i+1} \\ s_{i+1} \end{pmatrix}}.$$

3 ⇒ 2: It is enough to show that v^i can be written as the convex combination of v^{i-1} and v^{i+1}. Since there exists $\alpha > 0$ such that $r_j = \alpha r_{j+1}$ and $s_j = \alpha s_{j+1}$, we get

$$v^i - v^{i-1} = \alpha(v^{i+1} - v^i).$$

It follows that

$$v^i = \tfrac{1}{1+\alpha} v^{i+1} + \tfrac{\alpha}{1+\alpha} v^{i-1}.$$

Note that $0 \leq \frac{1}{1+\alpha} \leq 1$. ∎

By utilizing these properties of redundant vertices, it can be shown that arbitrary descriptions of a piecewise linear characteristic must have some common properties in terms of minimal descriptions as stated in the following lemma.

Lemma 4.8.15 *Let $\mathcal{G} = \mathrm{plc}(d^-, [v^i]_{i=1}^{k-1}, d^+)$ be a k-piecewise linear characteristic and $(d^-, [v^{l_i}]_{i=1}^{k'-1}, d^+)$ be one of its minimal descriptions. Also let the vector pairs (r, s) and*

$(r^{\min}, s^{\min})$ *be defined for* $\operatorname{plc}(d^-, [v^i]_{i=1}^{k-1}, d^+)$ *and* $\operatorname{plc}(d^-, [v^{l_i}]_{i=1}^{k'-1}, d^+)$ *in accordance with* (4.9), *respectively. Then, the following statements hold.*

1. *For each* $j \in \overline{k}$ *there exist* $\alpha > 0$ *and* $p \in \overline{k'}$ *such that* $r_j = \alpha r_p^{\min}$ *and* $s_j = \alpha s_p^{\min}$.

2. *For each* $p \in \overline{k'}$ *there exist* $\alpha > 0$ *and* $j \in \overline{k}$ *such that* $r_p^{\min} = \alpha r_j$ *and* $s_p^{\min} = \alpha s_j$.

Proof: All the statements that will be made for the vectors r and $r^{\min}$ are equally valid for the vectors s and $s^{\min}$ in the rest of this proof.

1: We distinguish four cases.

- *Case 1:* $j \in \{1, k\}$. Obviously,

$$p = \begin{cases} 1 & \text{if } j = 1, \\ k' & \text{if } j = k. \end{cases}$$

  and $\alpha = 1$ do the job.

- *Case 2:* $j \in \{2, 3, \ldots, l_1\}$. Note that $v^{j'}$ is redundant for all $j' \in \{1, 2, \ldots, l_1 - 1\}$. It follows from Lemma 4.8.14 that there exists $\alpha_{j'}$ such that $r_{j'+1} = \alpha_{j'} r_{j'}$. Therefore, $r_j = (\prod_{j'=1}^{j-1} \alpha_{j'}) r_1$. Consequently, $p = 1$ and $\alpha = \prod_{j'=1}^{j-1} \alpha_{j'}$ do the job.

- *Case 3:* $j \in \{l_{p-1} + 1, l_{p-1} + 2, \ldots, l_p\}$. Note that $v^{j'}$ is redundant for all $j' \in \{l_{p-1} + 1, l_{p-1} + 2, \ldots, l_p - 1\}$. It follows from Lemma 4.8.14 that there exists $\alpha_{j'}$ such that $r_{j'+1} = \alpha_{j'} r_{j'}$. Thus, we get

$$r_{j''} = \begin{cases} (\prod_{j'=j''}^{j-1} \alpha_{j'})^{-1} r_j & \text{if } j'' \in \{l_{p-1}, l_{p-1} + 1, \ldots, j - 1\}, \\ r_j & \text{if } j'' = j, \\ (\prod_{j'=j}^{j''-1} \alpha_{j'}) r_j & \text{if } j'' \in \{j + 1, j + 2, \ldots, l_p\}. \end{cases} \tag{4.43}$$

  On the other hand, we have

$$\begin{aligned} r_p^{\min} &= v_1^{l_p} - v_1^{l_{p-1}} \\ &= \underbrace{v_1^{l_p} - v_1^{l_p - 1}}_{r_{l_p}} + \underbrace{v_1^{l_p - 1} - v_1^{l_p - 2}}_{r_{l_p - 1}} + - \cdots + \underbrace{v_1^{l_{p-1}+1} - v_1^{l_{p-1}}}_{r_{l_{p-1}+1}}. \end{aligned}$$

  By using (4.43), we get $r_p^{\min} = \beta r_j$ for some $\beta > 0$. Therefore, p and $\alpha = 1/\beta$ do the job.

- *Case 4:* $j \in \{l_{k'-1} + 1, l_{k'-1} + 2, \ldots, k - 1\}$. Note that $v^{j'}$ is redundant for all $j' \in \{l_{k'-1} + 1, l_{k'-1} + 2, \ldots, k - 1\}$. It follows from Lemma 4.8.14 that there exists

$\alpha_{j'}$ such that $r_{j'} = \alpha_{j'} r_{j'+1}$. Therefore, $r_j = (\prod_{j'=j}^{k-1} \alpha_{j'}) r_k$. Consequently, $p = k$ and $\alpha = \prod_{j'=j}^{k-1} \alpha_{j'}$ do the job.

2: The proof of the previous item shows that one can find positive α's for the following choices of j's.

- *Case 1:* $p \in \{1, k'\}$. Take $j = 1$ if $p = 1$ and $j = k$ if $p = k'$.
- *Case 2:* $p \in \{2, 3, \ldots, k'-1\}$. Take any $j \in \{l_{p-1}+1, l_{p-1}+2, \ldots, l_p\}$. ■

As indicated in Remark 4.3.2, there are exactly two minimal descriptions. The following lemma depicts how those descriptions are related to each other.

Lemma 4.8.16 *Let $(d^-, v^1, v^2, \ldots, v^{k-1}, d^+)$ be a minimal description of a k-piecewise linear characteristic $\mathcal{G}$. Also let the vector pairs $(r^{\min}, s^{\min})$ and $(r^{\min'}, s^{\min'})$ be defined for* $\operatorname{plc}(d^-, v^1, v^2, \ldots, v^{k-1}, d^+)$ *and* $\operatorname{plc}(d^+, v^{k-1}, v^{k-2}, \ldots, v^1, d^-)$ *in accordance with* (4.9), *respectively. Then, $r_j^{\min} = r_{k+1-j}^{\min'}$ and $s_j^{\min} = s_{k+1-j}^{\min'}$ for each $j \in \overline{k}$.*

Proof: Evident. ■

Finally, we can proof Theorem 4.5.5 by employing the above lemmas.

Proof of Theorem 4.5.5: Let $\operatorname{plc}(d^{i,-}, [v^{i,j}]_{j=1}^{k-1}, d^{i,+})$ and $\operatorname{plc}(\bar{d}^{i,-}, [\bar{v}^{i,j}]_{j=1}^{k-1}, \bar{d}^{i,+})$ for $i \in \overline{m}$ be the descriptions of the piecewise linear characteristics $[\mathcal{G}^i]_{i=1}^m$ corresponding to the horizontal complementarity representations $(\cdot, \cdot, [R^j]_{j=1}^k, [S^j]_{j=1}^k)$ and $(\cdot, \cdot, [\bar{R}^j]_{j=1}^k, [\bar{S}^j]_{j=1}^k)$, respectively.

i. Assume that for each $i \in \overline{m}$ the minimal descriptions of $\operatorname{plc}(d^{i,-}, [v^{i,j}]_{j=1}^{k-1}, d^{i,+})$ and $\operatorname{plc}(\bar{d}^{i,-}, [\bar{v}^{i,j}]_{j=1}^{k-1}, \bar{d}^{i,+})$ are the same and $(d^{i,-}, [v^{i,l_j}]_{j=1}^{k_i'-1}, d^{i,+})$. Note that every column representative matrix of $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$ is of the form

$$\begin{aligned}([M\bar{R}^j + N\bar{S}^j]_{j=1}^k)^l &= \left((M\bar{R}^{l_1} + N\bar{S}^{l_1})_{\bullet 1} \cdots (M\bar{R}^{l_m} + N\bar{S}^{l_m})_{\bullet m}\right) \\ &= M\operatorname{diag}(\bar{r}_1^{l_1}, \bar{r}_2^{l_2}, \ldots, \bar{r}_m^{l_m}) + N\operatorname{diag}(\bar{s}_1^{l_1}, \bar{s}_2^{l_2}, \ldots, \bar{s}_m^{l_m}) \qquad (4.44)\end{aligned}$$

for some $l \in \overline{k}^m$. However, (4.44) and Lemma 4.8.15 imply that for each column representative matrix of $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$ one can find a column representative matrix of $[MR^j + NS^j]_{j=1}^k$ such that the determinants of these two representative matrices have the same sign. Since $[MR^j + NS^j]_{j=1}^k$ enjoys the column $\mathcal{W}$-property, so does $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$.

ii. As already noted in Remark 4.3.2, there are exactly two minimal descriptions of a k-piecewise linear characteristic. Let q of the minimal descriptions of $\mathrm{plc}(d^{i,-}, [v^{i,j}]_{j=1}^{k-1}, d^{i,+})$ and $\mathrm{plc}(\bar{d}^{i,-}, [\bar{v}^{i,j}]_{j=1}^{k-1}, \bar{d}^{i,+})$ be different. The equation (4.44), Lemma 4.8.16 and Lemma 4.8.15 imply that for each column representative matrix of $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$ one can find a column representative matrix of $[MR^j + NS^j]_{j=1}^k$ such that the sign of the determinant of the former is equal to $(-1)^q$ times the sign of the determinant of the latter. Since $[MR^j + NS^j]_{j=1}^k$ enjoys the column $\mathcal{W}$-property, so does $[M\bar{R}^j + N\bar{S}^j]_{j=1}^k$. ■

4.8.6 Proofs for Section 4.6

Proof of Corollary 4.6.1: Note that $\mathcal{G}^i = \mathrm{plc}(d^{i,-}, v^{i,1}, d^{i,+})$ where

$$d^{i,-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; v^{i,1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ and } d^{i,+} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore,

$$r^i = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ and } s^i = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

A horizontal complementarity representation $(q^u, q^y, [R^1, R^2], [S^1, S^2])$ of $[\mathcal{G}^i]_{i=1}^m$ can be given by

$$q^u = q^y = 0, \; R^1 = 0, \; R^2 = I, \; S^1 = -I, \text{ and } S^2 = 0$$

Hence $[G(s)R^1 - S^1, G(s)R^2 - S^2] = [I, G(s)]$. Note that there is a natural correspondence between the column representation matrices of $[I, G(s)]$ and the submatrices of $G(s)$. This fact results in

- The ordered matrix set $[I, G(s)]$ is of index 1 if and only if $G(s)$ is totally of index 1.
- The ordered matrix set $[I, G(\sigma)]$ has the column $\mathcal{W}$-property for all sufficiently large σ if and only if $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ.

Note also that $[DR^1 - S^1, DR^2 - S^2] = [I, D]$. Hence, due to Remark 4.4.5, $\mathrm{HLCP}(Cx_0, [I, D])$ is solvable if and only if $\mathrm{LCP}(Cx_0, D)$ is solvable. The assertion follows immediately from the facts listed above together with Theorem 4.5.4 item 1. ■

Proof of Corollary 4.6.2: Note that $\mathcal{G}^i = \mathrm{plc}(d^{i,-}, v^{i,1}, v^{i,2}, d^{i,+})$ where

$$d^{i,-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; v^{i,1} = \begin{pmatrix} e_1^i \\ 0 \end{pmatrix}, \; v^{i,2} = \begin{pmatrix} e_2^i \\ 0 \end{pmatrix}, \text{ and } d^{i,+} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Therefore,

$$r^i = \mathrm{col}(0, e_2^i - e_1^i, 0) \text{ and } s^i = \mathrm{col}(-1, 0, -1).$$

A horizontal complementarity representation $(q^u, q^y, [R^1, R^2, R^3], [S^1, S^2, S^3])$ of $[\mathcal{G}^i]_{i=1}^m$ can be given by

$$q^u = \mathrm{col}(e_1^1, e_1^2, \dots, e_1^m),\ q^y = 0,$$
$$R^1 = R^3 = 0,\ R^2 = \mathrm{diag}(e_2^1 - e_1^1, e_2^2 - e_1^2, \dots, e_2^m - e_1^m),$$
$$S^1 = S^3 = -I,\ S^2 = 0.$$

Hence $[G(s)R^j - S^j]_{j=1}^3 = [I, G(s)R^2, I]$. Since R^2 is a diagonal matrix with positive elements on the diagonal and $[DR^1 - S^1, DR^3 - S^3] = [I, I]$, the following facts can be inferred.

- The ordered matrix set $[I, G(s)R^2, I]$ has the column $\mathcal{W}$-property for all sufficiently large σ if and only if $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ.
- The ordered matrix set $[DR^1 - S^1, DR^3 - S^3]$ is nondegenerate.

The assertion follows immediately from the facts listed above together with Theorem 4.5.4 item 2. ∎

Proof of Corollary 4.6.3: Note that $\mathcal{G}^i = \mathrm{plc}(d^{i,-}, v^{i,1}, v^{i,2}, v^{i,3}, v^{i,4}, d^{i,+})$ where

$$d^{i,-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ v^{i,1} = \begin{pmatrix} e_1^i \\ f_1^i \end{pmatrix},\ v^{i,2} = \begin{pmatrix} 0 \\ f_1^i \end{pmatrix},\ v^{i,3} = \begin{pmatrix} 0 \\ f_2^i \end{pmatrix},\ v^{i,4} = \begin{pmatrix} e_2^i \\ f_2^i \end{pmatrix}, \text{ and } d^{i,+} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Therefore,

$$r^i = \mathrm{col}(0, -e_1^i, 0, e_2^i, 0) \text{ and } s^i = \mathrm{col}(-1, 0, f_2^i - f_1^i, 0, -1).$$

A horizontal complementarity representation $(q^u, q^y, [R^1, R^2, \dots, R^5], [S^1, S^2, \dots, S^5])$ of $[\mathcal{G}^i]_{i=1}^m$ can be given by

$$q^u = \mathrm{col}(e_1^1, e_1^2, \dots, e_1^m),\ q^y = \mathrm{col}(f_1^1, f_1^2, \dots, f_1^m),$$
$$R^1 = R^3 = R^5 = 0,\ R^2 = -\mathrm{diag}(e_1^1, e_1^2, \dots, e_1^m),\ R^4 = \mathrm{diag}(e_2^1, e_2^2, \dots, e_2^m),$$
$$S^1 = S^5 = -I,\ S^2 = S^4 = 0,\ S^3 = \mathrm{diag}(f_2^1 - f_1^1, f_2^2 - f_1^2, \dots, f_2^m - f_1^m).$$

Hence $[G(s)R^j - S^j]_{j=1}^5 = [I, G(s)R^2, S^3, G(s)R^4, I]$. Since R^2, S^3 and R^4 are all diagonal matrices with positive elements on the diagonal and $[DR^1 - S^1, DR^5 - S^5] = [I, I]$, the following facts can be inferred.

- The ordered matrix set $[I, G(s)R^2, S^3, G(s)R^4, I]$ has the column $\mathcal{W}$-property for all sufficiently large σ if and only if $G(\sigma)$ is a $\mathcal{P}$-matrix for all sufficiently large σ.

- The ordered matrix set $[DR^1 - S^1, DR^5 - S^5]$ is nondegenerate.

The assertion follows immediately from the facts listed above together with Theorem 4.5.4 item 2. ■

Proof of Corollary 4.6.5: Note that $\mathcal{G}^i = \mathrm{plc}(d^{i,-}, v^{i,1}, v^{i,2}, d^{i,+})$ where

$$d^{i,-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ v^{i,1} = \begin{pmatrix} e_1^i \\ f_1^i \end{pmatrix}, \ v^{i,2} = \begin{pmatrix} e_2^i \\ f_2^i \end{pmatrix}, \text{ and } d^{i,+} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Therefore,

$$r^i = \mathrm{col}(0, e_2^i - e_1^i, 0) \text{ and } s^i = \mathrm{col}(-1, f_2^i - f_1^i, -1).$$

A horizontal complementarity representation $(q^u, q^y, [R^1, R^2, R^3], [S^1, S^2, S^3])$ of $[\mathcal{G}^i]_{i=1}^m$ can be given by

$$\begin{aligned} q^u &= \mathrm{col}(e_1^1, e_1^2, \dots, e_1^m), \ q^y = \mathrm{col}(f_1^1, f_1^2, \dots, f_1^m), \\ R^1 &= R^3 = 0, \ R^2 = \mathrm{diag}(e_2^1 - e_1^1, e_2^2 - e_1^2, \dots, e_2^m - e_1^m), \\ S^1 &= S^3 = -I, \ S^2 = \mathrm{diag}(f_2^1 - f_1^1, f_2^2 - f_1^2, \dots, f_2^m - f_1^m). \end{aligned}$$

Hence $[G(s)R^j - S^j]_{j=1}^3 = [I, G(s)R - S, I]$. Note that $[DR^1 - S^1, DR^3 - S^3] = [I, I]$. Then, the following facts can be inferred.

- The ordered matrix set $[I, G(s)R - S, I]$ has the column $\mathcal{W}$-property for all sufficiently large σ if and only if $G(\sigma)R - S$ is a $\mathcal{P}$-matrix for all sufficiently large σ.

- The ordered matrix set $[DR^1 - S^1, DR^3 - S^3]$ is nondegenerate.

The assertion follows immediately from the facts listed above together with Theorem 4.5.4 item 2. ■

References

[1] J. P. Aubin and A. Cellina. *Differential Inclusions.* Springer, Berlin, 1984.

[2] W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis.* Kluwer, Deventer, the Netherlands, 1981.

[3] D.W. Bushaw. *Differential Equations with Discontinuous Forcing Term.* PhD thesis, Dept. of Math., Princeton Univ., 1952.

[4] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Dynamical analysis of linear passive networks with diodes. Part II: Consistency of a time-stepping method. Technical Report 00 I/03, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted for publication.

[5] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem.* Academic Press, Inc., Boston, 1992.

[6] B. C. Eaves and C. E. Lemke. Equivalence of LCP and PLS. *Mathematics of Operations Research*, 6:475–484, 1981.

[7] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides.* Mathematics and Its Applications. Prentice-Hall, Dordrecht, The Netherlands, 1988.

[8] M. L. J. Hautus and L. M. Silverman. System structure and singular control. *Linear Algebra and Its Applications*, 50:369–402, 1983.

[9] W.P.M.H. Heemels, M.K. Çamlıbel, and J.M. Schumacher. A time-stepping method for relay systems. In *Proc. of the 39th IEEE Conference on Decision and Control*, Sydney (Australia), 2000.

[10] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. The rational complementarity problem. *Linear Algebra and Its Applications*, 294:93–135, 1999.

[11] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics*, 60(4):1234–1269, 2000.

[12] I. Kaneko. A linear complementarity problem with an n by 2n "P"-matrix. *Math. Programming Study*, 7:120–141, 1978.

[13] I. Kaneko and J. S. Pang. Some n by dn linear complementarity problems. *Linear Algebra Appl.*, 34:297–319, 1980.

[14] J. P. LaSalle. Time optimal control systems. *Proc. Natl. Acad. Sci. U.S.*, 45:573–577, 1959.

[15] D.M.W. Leenaerts and W.M.G. van Bokhoven. *Piecewise linear modelling and analysis.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[16] Y.J. Lootsma, A.J. van der Schaft, and M.K. Çamlıbel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[17] P. Lötstedt. Coulomb friction in two-dimensional rigid body systems. *ZAMM*, 61:605–615, 1981.

[18] A. Yu. Pogromsky, W. P. M. H. Heemels, and H. Nijmeijer. On the well-posedness of relay systems. *Manuscript under preparation*, 2000.

[19] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1976.

[20] A.J. van der Schaft and J.M. Schumacher. Complementarity modelling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[21] R. Sznajder and M. S. Gowda. Generalizations of P_0- and P-properties; Extended vertical and horizontal linear complementarity problems. *Linear Algebra and Its Applications*, 223/224:695–715, 1995.

[22] Ya. Z. Tsypkin. *Relay Control Systems*. Cambridge University Press, New York, 1984.

[23] V. I. Utkin. *Sliding Modes in Control Optimization*. Springer, Berlin, 1992.

[24] L. Vandenberghe, B. L. De Moor, and J. Vandewalle. The generalized linear complementarity problem applied to the complete analysis of resistive piecewise-linear circuits. *IEEE Trans. Circuits Syst.*, CAS-36:1382–1391, 1989.

Part II

Approximations

Chapter 5

From Lipschitzian to non-Lipschitzian characteristics: convergence of solutions

5.1 Introduction

Modeling process can be viewed as a mapping which assigns models to physical systems. Reasonably, *close* physical systems should be associated to *close* models. Stated differently, the modeling process should depend on physical systems continuously. We aim to address the question of continuity of linear complementarity models in this chapter.

The approach that is taken in this chapter has certain parallelism with the work that has been done in the context of singular perturbations. For an extensive survey on the subject, we refer to [3] (see also [6, Sections 4.3 and 5.5] for a quick review). Our treatment differs from this vein of research considerably since the systems under investigation are *nonsmooth* in general.

In mechanics, the smoothing methods have been extensively studied. Roughly speaking, the aim of the smoothing methods is to consider the *nonsmooth* system as the limit of (in a suitable sense) a sequence of *smooth* systems for which strong properties such as existence and uniqueness of solutions, continuous dependence on parameters etc. are known. For an encyclopedic survey, we refer to [1]. A comparison of smoothing methods, time stepping methods and event-tracking methods (in the context of mechanics again) can be found in [4]. The framework that is used in this chapter is very close the framework used in that context. However, the systems under investigation here, namely linear complementarity systems, are not completely covered by the work in nonsmooth mechanical systems.

Continuity of linear dynamical models is addressed for instance in [2, 8]. While continuity is defined via pointwise convergence of trajectories in [8], [2] considers continuity in the graph topology. Similar to what has been done in [8], we will look at convergence of trajectories.

We will mainly focus on the linear complementarity systems given by

$$\dot{x} = Ax + Bu \tag{5.1a}$$

$$y = Cx + Du \tag{5.1b}$$

$$0 \leq u \perp y \geq 0. \tag{5.1c}$$

Notice that the so-called complementarity conditions (5.1c) as depicted in Figure 5.1 do not define a function between u and y. However, a slight perturbation of the piece with

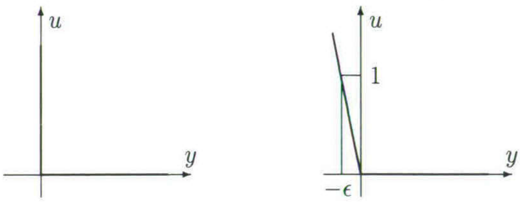


Figure 5.1: Complementarity characteristic and one of its possible approximations

infinite slope allows to express u as a piecewise-linear (and hence Lipschitz continuous) function of y. Naturally, one might expect/desire that this approximated characteristic generates trajectories *close* to ones of the complementarity system (5.1). However, if the complementarity system is ill-posed it is not hard to find examples for which this property does not hold . The main objective of the present chapter is to prove the convergence of the trajectories generated by the Lipschitzian characteristics to those generated by the (non-Lipschitzian) complementarity characteristic for a class of well-posed complementarity systems including linear passive ones.

In the sequel, we distinguish two types of approximations, namely *structured* and *unstructured* approximations. As an example of structured approximations, consider the linear system (5.1a)-(5.1b) with the approximating characteristic of Figure 5.1. It can be verified that the overall approximating system is equivalent to the complementarity system given by

$$\dot{x}^\epsilon = A_\epsilon x^\epsilon + B_\epsilon u^\epsilon \tag{5.2a}$$

$$y^\epsilon = C_\epsilon x^\epsilon + D_\epsilon u^\epsilon \tag{5.2b}$$

$$0 \leq u^\epsilon \perp y^\epsilon \geq 0 \tag{5.2c}$$

with $(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon) = (A, B, C, D + \epsilon I)$ in the sense that there is a one-to-one correspondence between the state trajectories of the two systems. We call this type of approximations structured because of the explicit dependence of $(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ on (A, B, C, D). When such an explicit dependence is absent, we call the approximations unstructured.

The outline of the chapter is as follows. In Section 5.2, we recall several facts such as Carathéodory solution of a differential equation, and the notions of passivity and passifiability by pole shifting. The well-posedness results on linear passive complementarity systems will be summarized in Section 5.3 for the sake of completeness. Section 5.4 contains the main contributions of the chapter. It consists of two subsections in which the continuity of the solutions of structured and unstructured approximations are investigated, respectively. This will be followed by Section 5.5 where convergence of approximating trajectories for nonregular initial states of the original system is considered. The chapter will be closed by conclusions in Section 5.6 and proofs in Section 5.7.

5.2 Preliminaries

Consider the continuous-time, linear and time-invariant system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{5.3a}$$

$$y(t) = Cx(t) + Du(t) \tag{5.3b}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^m$ and A, B, C, and D are matrices with appropriate sizes. We denote (5.3) by $\Sigma(A, B, C, D)$.

A triple $(u, x, y) \in \mathcal{L}_2((t_0, t_1), \mathbb{R}^{m+n+m})$ is said to be an *$\mathcal{L}_2$-solution* on (t_0, t_1) of $\Sigma(A, B, C, D)$ with the initial state x_0 if it satisfies (5.3a) in the sense of Carathéodory, i.e., for almost all $t \in [t_0, t_1]$,

$$x(t) = x_0 + \int_{t_0}^{t} [Ax(s) + Bu(s)]ds. \tag{5.4}$$

and (5.3b) holds.

Next, we recall the definition of the passivity notion.

Definition 5.2.1 [7] The system $\Sigma(A, B, C, D)$ given by (5.3) is said to be *passive (dissipative with respect to the supply rate $u^\top y$)* if there exists a function $V : \mathbb{R}^n \to \mathbb{R}_+$ (a *storage function*), such that

$$V(x(t_0)) + \int_{t_0}^{t_1} u^\top(t)y(t)dt \geq V(x(t_1)) \tag{5.5}$$

holds for all t_0 and t_1 with $t_1 \geq t_0$, and all $\mathcal{L}_2$-solutions $(u, x, y) \in \mathcal{L}_2((t_0, t_1), \mathbb{R}^{m+n+m})$ of $\Sigma(A, B, C, D)$.

We state a well-known result on passive systems which characterizes passivity in terms of linear matrix inequalities.

Lemma 5.2.2 [7] *Assume that (A, B, C) is minimal. Then $\Sigma(A, B, C, D)$ is passive if and only if the matrix inequalities*

$$K = K^\top > 0 \text{ and } \begin{bmatrix} A^\top K + KA & KB - C^\top \\ B^\top K - C & -(D + D^\top) \end{bmatrix} \leq 0$$

have a solution. Moreover, $V(x) = \frac{1}{2}x^\top Kx$ is a quadratic storage function if and only if K is a solution of the above matrix inequalities.

The notion of passifiability by pole shifting is of interest in the context of linear complementarity systems. For the sake of completeness, we recall this notion which was introduced in Chapter 3.

Definition 5.2.3 The quadruple (A, B, C, D) is said to be *passifiable by pole shifting* if there exists $\rho \in \mathbb{R}$ such that $\Sigma(A + \rho I, B, C, D)$ is passive.

The following theorem is quoted from Chapter 3. It provides necessary and sufficient conditions for passifiability by pole shifting.

Theorem 5.2.4 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and $\operatorname{col}(B, D + D^\top)$ is of full column rank. Let E be such that $\ker E = \{0\}$ and $\operatorname{im} E = \ker (D + D^\top)$. Then (A, B, C, D) is passifiable by pole shifting if and only if D is nonnegative definite and $E^\top CBE$ is symmetric positive definite.*

5.3 Linear Complementarity Systems

The main objects of study will be the linear complementarity systems, that is to say, linear systems with complementarity conditions given by

$$\dot{x} = Ax + Bu \tag{5.6a}$$

$$y = Cx + Du \tag{5.6b}$$

$$0 \leq u \perp y \geq 0. \tag{5.6c}$$

We denote the linear complementarity system (5.6) by $\mathrm{LCS}(A, B, C, D)$. Next, we shall define what is meant by a solution of a linear complementarity system by clarifying the meaning of the complementarity conditions in (5.6c).

Definition 5.3.1 The triple $(u, x, y) \in \mathcal{L}_2((0, \tau), \mathbb{R}^{m+n+m})$ is a *$\mathcal{L}_2$-solution* of $\mathrm{LCS}(A, B, C, D)$ on $[0, \tau]$ with initial state x_0 if the following conditions hold.

1. (u, x, y) is a $\mathcal{L}_2$-solution of on $[0, \tau]$ of $\Sigma(A, B, C, D)$ with the initial state x_0.

2. For almost all $t \in [0, \tau]$, $0 \leq u(t) \perp y(t) \geq 0$.

The initial state is said to be *regular* if there exists a solution with this initial state and *nonregular* otherwise.

As it is shown in Chapter 3, the passifiability of the system $\Sigma(A, B, C, D)$ guarantees the existence and uniqueness of solutions (in the sense of Definition 5.3.1) to LCS(A, B, C, D) for suitable initial conditions. Indeed, the following lemma has been proven in Chapter 3.

Lemma 5.3.2 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and* $\operatorname{col}(B, D + D^\top)$ *is of full column rank. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole-shifting. Let $\tau > 0$ be given. Then, there exists a unique $\mathcal{L}_2$-solution on $[0, \tau]$ of LCS(A, B, C, D) with the initial state x_0 if and only if $Cx_0 \in \mathcal{Q}_D^*$.*

Here the dual cone of the set $\mathcal{Q}_D = \mathrm{SOL}(0, D) = \{v \mid v \geq 0,\ Dv \geq 0, \text{ and } v^\top Dv = 0\}$ (as defined in Chapter 1) is denoted by $\mathcal{Q}_D^*$.

5.4 Continuity of Solutions

In this section we investigate continuity of behaviors of linear complementarity systems. First, some specific approximation schemes will be under consideration. These approximations will be obtained by approximating *only* the complementarity characteristic. In this respect, they are *structured* approximations. Later on, the investigation will be carried out for more general approximations. Those approximations will be obtained by approximating linear complementarity system *as a whole.* For this reason, we call them *unstructured* approximations.

5.4.1 Structured approximations

For a given LCS(A, B, C, D), we consider the following systems

$$\dot{x} = Ax + Bv \tag{5.7a}$$

$$z = Cx + Dv \tag{5.7b}$$

$$v_i(t) = \begin{cases} -(\epsilon_i^z)^{-1} z_i(t) & \text{if } z_i(t) \leq 0 \\ -\epsilon_i^v z_i(t) & \text{if } z_i(t) \geq 0 \end{cases} \quad \text{for all } t \text{ and for each } i \in \overline{m} \tag{5.7c}$$

where $\epsilon^v, \epsilon^z \in \mathbb{R}^m$ and $\mathrm{col}(\epsilon^v, \epsilon^z) > 0$. The piecewise linear relation (5.7c) between v_i and z_i is depicted in Figure 5.2. Note that these characteristics *converge* to the complementarity ones as $\mathrm{col}(\epsilon^v, \epsilon^z)$ tends to zero. We denote (5.7) by $\mathrm{App}(A, B, C, D, \epsilon^v, \epsilon^z)$.

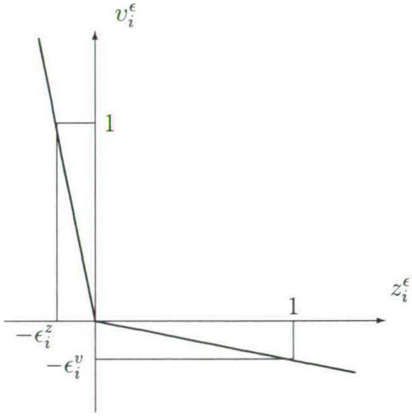


Figure 5.2: Approximation of the complementarity characteristic

We say that a triple $(v, x, z) \in \mathcal{L}_2((0,\tau), \mathbb{R}^{m+n+m})$ is a *solution* on $[0,\tau]$ of $\mathrm{App}(A, B, C, D, \epsilon^v, \epsilon^z$ with the initial state x_0 if (v, x, z) is an $\mathcal{L}_2$-solution of $\Sigma(A, B, C, D)$ with the initial state x_0 and (5.7c) holds.

With a change of variables, every $\mathrm{App}(A, B, C, D, \epsilon^v, \epsilon^z)$ can be rewritten as a linear complementarity system. Next, we state this equivalence in the following proposition. To do this, we need to introduce some nomenclature. Let two positive vectors $\epsilon^v \in \mathbb{R}^m$ and $\epsilon^z \in \mathbb{R}^m$ be given. We denote $\mathrm{col}(\epsilon^v, \epsilon^z)$, $\mathrm{diag}(\epsilon^v)$ and $\mathrm{diag}(\epsilon^z)$ by $\bar{\epsilon}$, Λ^v and Λ^z, respectively. For each matrix quadruple (A, B, C, D), we associate a matrix quadruple $(A_{\bar{\epsilon}}, B_{\bar{\epsilon}}, C_{\bar{\epsilon}}, D_{\bar{\epsilon}})$ where

$$
\begin{aligned}
A_{\bar{\epsilon}} &= A - B\Lambda^v(I + D\Lambda^v)^{-1}C \\
B_{\bar{\epsilon}} &= B(I + \Lambda^v D)^{-1} \\
C_{\bar{\epsilon}} &= (1 - \Lambda^v\Lambda^z)(I + D\Lambda^v)^{-1}C \\
D_{\bar{\epsilon}} &= (\Lambda^z + D)(I + \Lambda^v D)^{-1}.
\end{aligned}
$$

Note that $(A_{\bar{\epsilon}}, B_{\bar{\epsilon}}, C_{\bar{\epsilon}}, D_{\bar{\epsilon}})$ converges to (A, B, C, D) as $\bar{\epsilon}$ tends to zero.

Proposition 5.4.1 *Consider a matrix quadruple (A, B, C, D) and two positive m-vectors*

ϵ^v and ϵ^z . Let the pairs (v, z) and (u, y) satisfy

$$\begin{pmatrix} u \\ y \end{pmatrix} = \begin{pmatrix} I & \Lambda^v \\ \Lambda^z & I \end{pmatrix} \begin{pmatrix} v \\ z \end{pmatrix}. \tag{5.8a}$$

Then, the following two statements are equivalent for all sufficiently small ϵ^v and ϵ^z.

1. *The triple (v, x, z) is a solution of App$(A, B, C, D, \epsilon^v, \epsilon^z)$ with the initial state x_0.*
2. *The triple (u, x, y) is an $\mathcal{L}_2$-solution of LCS$(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ with the initial state x_0.*

In the study of convergence of the solutions of App$(\cdot, \cdot, \cdot, \cdot, \cdot, \cdot)$, the uniform passifiability notion, as introduced in the following definition, plays an important role.

Definition 5.4.2 The sequence of systems $\Sigma(A^\nu, B^\nu, C^\nu, D^\nu)$ is said to be *uniformly passifiable by pole shifting in a neighborhood of zero* if there exist a real number ρ and a positive definite matrix K such that for all sufficiently small ν $\Sigma(A^\nu + \rho I, B^\nu, C^\nu, D^\nu)$ is passive with the storage function $x \mapsto x^\top K x$.

For ease of reference, we state the following rather obvious fact which will be used later.

Fact 5.4.3 Suppose that the sequence of systems $\Sigma(A^\nu, B^\nu, C^\nu, D^\nu)$ is uniformly passifiable by pole shifting for all sufficiently small ν and $(A^\nu, B^\nu, C^\nu, D^\nu)$ converges to (A, B, C, D) as ν tends to zero. Then, (A, B, C, D) is passifiable by pole shifting.

Now, we can state our first convergence result.

Theorem 5.4.4 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and B is of full column rank. Suppose that $\Sigma(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ is uniformly passifiable by pole shifting for all sufficiently small $\bar\epsilon$. Let $\tau > 0$ and a regular initial state x_0 of LCS(A, B, C, D) be given. Assume that $C_{\bar\epsilon} x_0 \in \mathcal{Q}^*_{D_{\bar\epsilon}}$ for all sufficiently small $\bar\epsilon$. Let $(u_{x_0}, x_{x_0}, y_{x_0})$ and $(v^{\bar\epsilon}_{x_0}, x^{\bar\epsilon}_{x_0}, z^{\bar\epsilon}_{x_0})$ denote the unique solutions on $[0, \tau]$ of LCS(A, B, C, D) and App$(A, B, C, D, \epsilon^v, \epsilon^z)$ with the initial state x_0, respectively. Then, the following statements hold as $\bar\epsilon$ tends to zero.*

1. *The sequence of state trajectories $\{x^{\epsilon}_{x_0}\}$ converges uniformly to x_{x_0} on $[0, \tau]$.*
2. *If $\{v^{\epsilon}_{x_0}\}$ is bounded for all sufficiently small $\bar\epsilon$ then $\{(v^{\epsilon}_{x_0}, z^{\epsilon}_{x_0})\}$ converges weakly to (u_{x_0}, y_{x_0}).*

Next, we will show that the uniform passifiability hypothesis holds for two specific approximations.

Lemma 5.4.5 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and B is of full column rank. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Let ρ be such that $\Sigma(A + \rho I, B, C, D)$ is passive with the storage function $x \mapsto \frac{1}{2}x^\top Kx$. Then, the following statements hold.*

1. *The system $\Sigma(A + \rho I, B, C, D + \epsilon I)$ is passive with the storage function $x \mapsto \frac{1}{2}x^\top Kx$ for all $\epsilon > 0$, i.e., $\Sigma(A, B, C, D + \epsilon I)$ is uniformly passifiable by pole shifting for all sufficiently small ϵ.*

2. *For all $\rho' < \rho$ and for all sufficiently small ϵ, the system $\Sigma(A_\epsilon + \rho' I, B_\epsilon, C_\epsilon, D_\epsilon)$ is passive with the storage function $x \mapsto \frac{1}{2}x^\top Kx$ where*

$$\begin{aligned} A_\epsilon &= A - \epsilon B(I + \epsilon D)^{-1}C \\ B_\epsilon &= B(I + \epsilon D)^{-1} \\ C_\epsilon &= (1 - \epsilon^2)(I + \epsilon D)^{-1}C \\ D_\epsilon &= (I + \epsilon D)^{-1}(D + \epsilon I). \end{aligned}$$

   In other words, $\Sigma(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ is uniformly passifiable by pole shifting for all sufficiently small ϵ.

Therefore, the following corollary can be stated as an application of Theorem 5.4.4.

Corollary 5.4.6 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and B is of full column rank. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Let (u_{x0}, x_{x0}, y_{x0}) be the unique solution on $[0, \tau]$ of LCS(A, B, C, D) with the initial state x_0.*

1. *Let $(v^\epsilon_{x0}, x^\epsilon_{x0}, z^\epsilon_{x0})$ denote the unique solution on $[0, \tau]$ of App$(A, B, C, D, 0, \epsilon\,\iota)$. As ϵ tends to zero,*
   (a) *the sequence of state trajectories x^ϵ_{x0} converges uniformly to x_{x0} on $[0, \tau]$, and*
   (b) *if v^ϵ_{x0} is bounded for all sufficiently small ϵ then $\{(v^\epsilon_{x0}, z^\epsilon_{x0})\}$ converges weakly to (u_{x0}, y_{x0}).*

2. *Let $(v^\epsilon_{x0}, x^\epsilon_{x0}, z^\epsilon_{x0})$ denote the unique solution on $[0, \tau]$ of App$(A, B, C, D, \epsilon\,\iota, \epsilon\,\iota)$. As ϵ tends to zero,*
   (a) *the sequence of state trajectories x^ϵ_{x0} converges uniformly to x_{x0} on $[0, \tau]$, and*
   (b) *if v^ϵ_{x0} is bounded for all sufficiently small ϵ then $\{(v^\epsilon_{x0}, z^\epsilon_{x0})\}$ converges weakly to (u_{x0}, y_{x0}).*

Here ι denotes the vector of ones.

5.4.2 Unstructured approximations

Having established results on the convergence of structured approximations, we pass to the investigation of the convergence of more general approximating systems. The introduction of the class of approximating systems that will be under consideration is in order.

Definition 5.4.7 The sequence $\{(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)\}$ is said to be *an admissible approximation of* (A, B, C, D) if the following conditions hold.

1. D_ϵ is positive definite for all sufficiently small positive ϵ.

2. $\{(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)\}$ converges to (A, B, C, D) as ϵ tends to zero.

Note that the positive definiteness of D_ϵ implies passifiability by pole shifting. Therefore, for all sufficiently small ϵ the system $\Sigma(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ admits unique solutions for all initial states.

Now, we can present the main result of this subsection.

Theorem 5.4.8 *Consider a matrix quadruple (A, B, C, D) such that (A, B, C) is a minimal representation and B is of full column rank. Suppose that $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Let $\tau > 0$ and a regular initial state of LCS(A, B, C, D) x_0 be given. Also let $\{(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)\}$ be an admissible approximation of (A, B, C, D), and let (u_{x0}, x_{x0}, y_{x0}) and $(u^\epsilon_{x0}, x^\epsilon_{x0}, y^\epsilon_{x0})$ be the unique solutions on $[0, \tau]$ of LCS(A, B, C, D) and LCS$(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ with the initial state x_0, respectively. If $\{u^\epsilon_{x0}\}$ is bounded then $\{x^\epsilon_{x0}\}$ converges (strongly) to x_{x0} and $\{(u^\epsilon_{x0}, y^\epsilon_{x0})\}$ converges to (u_{x0}, y_{x0}) weakly in $\mathcal{L}_2$-sense as ϵ tends to zero.*

As illustrated in the following example, not all admissible approximations produce bounded u-trajectories.

Example 5.4.9 Consider the linear complementarity system LCS(A, B, C, D) given by

$$\begin{aligned} \dot{x}_1 &= u_1 \\ \dot{x}_2 &= u_2 \\ y_1 &= x_1 \\ y_2 &= x_2 \\ 0 \leq u &\perp y \geq 0 \end{aligned}$$

and the approximating systems LCS$(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ given by

$$\begin{aligned}
\dot{x}_1^\epsilon &= u_1^\epsilon \\
\dot{x}_2^\epsilon &= u_2^\epsilon \\
y_1^\epsilon &= x_1^\epsilon - \epsilon x_2^\epsilon + \epsilon^k u_1^\epsilon \\
y_2^\epsilon &= -\epsilon x_1^\epsilon + x_2^\epsilon + \epsilon^k u_2^\epsilon \\
0 \leq u^\epsilon &\perp y^\epsilon \geq 0.
\end{aligned}$$

It is easy to see that the above approximations are admissible. The unique solution $(u^\epsilon, x^\epsilon, y^\epsilon)$ of LCS$(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ with the initial state $x_0 = \begin{bmatrix} 0 & 1 \end{bmatrix}^\top$ can be computed as

$$(u^\epsilon, x^\epsilon) = \left(\begin{bmatrix} \epsilon^{-k-1} e^{-\epsilon^{-k} t} \\ 0 \end{bmatrix}, \begin{bmatrix} -\epsilon e^{-\epsilon^{-k} t} + \epsilon \\ 1 \end{bmatrix} \right) y^\epsilon = \begin{pmatrix} 0 \\ 1 - \epsilon^2 + \epsilon^2 e^{-\epsilon^{-k} t} \end{pmatrix}$$

One can check that $\|u_1^\epsilon\|^2 = \frac{\epsilon^{-k+2}}{2}(1 - e^{-2\epsilon^{-k}\tau})$ on a given interval $[0, \tau]$. Consequently, $\{u^\epsilon\}$ is not bounded if $k > 2$. An interesting observation is that the sequence of approximating systems is not uniformly passifiable by pole shifting in a neighborhood of zero if $k > 2$.

5.5 Nonregular Initial States

So far, what has been done is to investigate the convergence of the solutions, only those with a regular initial state of the limit system, of approximating systems. Although the limit system does not have solutions with the nonregular initial states, the admissible approximations have. Then, it is natural to raise the question if and in what sense the approximating solutions with nonregular initial states converge. By means of the following example, we will illustrate that different approximations may yield different limits in this case.

Example 5.5.1 Consider the LCS(A, B, C, D) given by

$$\begin{aligned}
\dot{x}_1 &= 2u_1 + u_2 \\
\dot{x}_2 &= u_1 + 2u_2 \\
y_1 &= x_1 \\
y_2 &= x_2 \\
0 \leq u &\perp y \geq 0,
\end{aligned}$$

the approximating systems $\mathrm{LCS}(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ given by

$$\begin{aligned}
\dot{x}_1^\epsilon &= 2u_1^\epsilon + u_2^\epsilon \\
\dot{x}_2^\epsilon &= u_1^\epsilon + 2u_2^\epsilon \\
y_1^\epsilon &= x_1^\epsilon + \epsilon u_1^\epsilon \\
y_2^\epsilon &= x_2^\epsilon + \epsilon u_2^\epsilon \\
0 \leq u^\epsilon &\perp y^\epsilon \geq 0
\end{aligned}$$

and $\mathrm{LCS}(A_\mu, B_\mu, C_\mu, D_\mu)$ given by

$$\begin{aligned}
\dot{x}_1^\mu &= 2u_1^\mu + u_2^\mu \\
\dot{x}_2^\mu &= u_1^\mu + 2u_2^\mu \\
y_1^\mu &= x_1^\mu + 2\mu u_1^\mu + \mu u_2^\mu \\
y_2^\mu &= x_2^\mu + \mu u_1^\mu + 2\mu u_2^\mu \\
0 \leq u^\mu &\perp y^\mu \geq 0.
\end{aligned}$$

Evidently, both $\{(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)\}$ and $\{(A_\mu, B_\mu, C_\mu, D_\mu)\}$ qualify as admissible approximations of (A, B, C, D). Let $(u^\epsilon, x^\epsilon, y^\epsilon)$ and (u^μ, x^μ, y^μ) denote the solutions of $\mathrm{LCS}(A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$ and $\mathrm{LCS}(A_\mu, B_\mu, C_\mu, D_\mu)$ with the initial state $x_0 = \begin{bmatrix}-5 & -1\end{bmatrix}^\top$. It can be checked that both $\{u^\epsilon\}$ and $\{u^\mu\}$ are convergent in the distributional sense. Indeed, they converge to $\begin{bmatrix}3 - \frac{4\sqrt{2}}{3\sqrt{3}} & \frac{4\sqrt{2}}{3\sqrt{3}} - 1\end{bmatrix}^\top \delta$ and $\begin{bmatrix}\frac{5}{2} & 0\end{bmatrix}^\top \delta$, respectively. The fact that these approximations converge to different limits naturally weakens the power of ideal modeling in this context. In fact, it shows that the ideal model cannot capture the *fast dynamics* of the actual system.

5.6 Conclusions

We have considered linear complementarity systems described by linear time invariant systems coupled to complementarity characteristics. It is known that these systems possess unique solutions if the underlying linear system is passifiable by pole shifting. For the uniformly passifiable structured approximations of these systems, it has been shown that the approximating state trajectories converge uniformly on each finite interval to the state trajectory of the original one. However, only weak convergence (in $\mathcal{L}_2$-sense) of approximating u-trajectories to the original u-trajectory could be established provided that they are uniformly bounded. As a side result, we proved that the uniform passifiability assumption holds for two particular approximations. Not surprisingly, stronger conditions were needed to prove convergence of approximating trajectories for unstructured approximations. Indeed, what has been shown is that the approximating state trajectories and u-trajectories converge, respectively, strongly and weakly in $\mathcal{L}_2$-sense to the corresponding original ones if u-trajectories are uniformly bounded. Moreover, by means of an example, it has been illustrated that the limit of the approximating trajectories for nonregular initial states of the original system depends in general on the approximation scheme.

We believe that the uniform passifiability property, both for structured and unstructured approximations, needs to be studied further. Another interesting research topic might be the characterization of the class of approximations for which the approximating trajectories for a given nonregular initial state converge to the same limit.

5.7 Proofs

This section contains the proofs of previously stated results in this chapter. We start with two preliminary subsections. First of them is devoted to the topological complementarity problem which will play a key role in proving the main results. The second one collects some basic facts from matrix theory. These subsections are followed by the proofs.

5.7.1 Topological complementarity problem

To prove Theorem 5.4.4 and Theorem 5.4.8, we employ a result, which was established in Chapter 6, on the convergence of solutions to the topological complementarity problem. For the sake of completeness, we quote TCP and related facts from Chapter 6.

TCP for the function space $\mathcal{L}_2([0,\tau],\mathbb{R})$ can be formulated as follows.

Problem 5.7.1 (TCP(q,T)) Given $q \in \mathcal{L}_2([0,\tau],\mathbb{R}^m)$ and $T : \mathcal{L}_2([0,\tau],\mathbb{R}^m) \to \mathcal{L}_2([0,\tau],\mathbb{R}^m)$, find $z \in \mathcal{L}_2([0,\tau],\mathbb{R}^m)$ such that

$$z(t) \geq 0 \tag{5.9a}$$

$$q(t) + (Tz)(t) \geq 0 \tag{5.9b}$$

for almost all $t \in [0,\tau]$ and

$$\langle z, q + Tz \rangle = 0. \tag{5.9c}$$

If z satisfies (5.9), we say that z solves TCP(q,T).

Note that the conditions given in item 2 of Definition 5.3.1 may be equivalently written as $u(t) \geq 0$, $y(t) \geq 0$ for almost all $t \in [0,\tau]$ and $\langle u, y \rangle = 0$. Hence, by associating the operator $T_{(A,B,C,D)}$ defined by

$$(T_{(A,B,C,D)}u)(t) = Du(t) + \int_0^t Ce^{A(t-s)}Bu(s)ds$$

to the matrix quadruple (A, B, C, D), the solutions of LCS(A, B, C, D) can be identified with the solutions of certain TCPs in the following manner.

Proposition 5.7.2 *The following statements hold.*

1. *If $(u, x, y) \in \mathcal{L}_2([0, \tau], \mathbb{R}^{m+n+m})$ is a solution on $[0, \tau]$ of LCS(A, B, C, D) with initial state x_0, then u is a solution of TCP$(Ce^{A\cdot}x_0|_{[0,\tau]}, T_{(A,B,C,D)})$.*

2. *If $u \in \mathcal{L}_2([0, \tau], \mathbb{R}^m)$ is a solution of TCP$(Ce^{A\cdot}x_0|_{[0,\tau]}, T_{(A,B,C,D)})$, then (u, x, y) is a solution on $[0, \tau]$ of LCS(A, B, C, D) with initial state x_0 where*

$$\begin{aligned} x &= e^{A\cdot}x_0|_{[0,\tau]} + T_{(A,B,I,0)}u \\ y &= Cx + Du. \end{aligned}$$

Before stating the theorem concerning convergence of solutions to TCP we need to introduce some nomenclature. Let $\mathcal{U}$ be a normed space. A sequence of operators $S_k : \mathcal{U} \to \mathcal{U}$ is said to be *uniformly convergent* to S if $\|S_k - S\|$ converges to zero where $\|\cdot\|$ denotes the norm induced by the norm defined on $\mathcal{U}$. An operator $T : \mathcal{U} \to \mathcal{U}$ will be said to be a compact operator if it maps every weakly convergent sequence of $\mathcal{U}$ to a strongly convergent one.

Theorem 5.7.3 *Let $T : \mathcal{L}_2([0, \tau], \mathbb{R}^m) \to \mathcal{L}_2([0, \tau], \mathbb{R}^m)$ be a compact operator and let $S : \mathcal{L}_2([0, \tau], \mathbb{R}^m) \to \mathcal{L}_2([0, \tau], \mathbb{R}^m)$ be a linear continuous operator. Suppose that there exist sequences $\{q_k\}$, $\{S_k\}$ and $\{T_k\}$ such that $\{q_k\}$ converges to q, S_k is linear continuous nonnegative definite (i.e. $\langle v, S_k v\rangle \geq 0$ for all $v \in \mathcal{L}_2([0, \tau], \mathbb{R}^m)$) for all sufficiently large k, and TCP$(q_k, S_k + T_k)$ is solvable for all k. Let z_k be a solution of TCP$(q_k, S_k + T_k)$. If $\{z_k\}$ converges weakly to z, S_k converges uniformly to S and $\{T_k z_k - T z_k\}$ converges to zero then z solves TCP$(q, S + T)$.*

5.7.2 Proofs for Section 5.4

Proof of Proposition 5.4.1: Note that all the inverses mentioned in the statement of the proposition and below exist for all sufficiently small ϵ^v and ϵ^z. One can check that

$$\begin{pmatrix} B_{\bar\epsilon} & -\frac{d}{dt}I + A_{\bar\epsilon} & 0 \\ C_{\bar\epsilon} & D_{\bar\epsilon} & -I \end{pmatrix} = \begin{pmatrix} I & -B\Lambda^v(I+D\Lambda^v)^{-1} \\ 0 & (I-\Lambda^v\Lambda^z)(I+D\Lambda^v)^{-1} \end{pmatrix}$$
$$\times \begin{pmatrix} B & -\frac{d}{dt}I + A & 0 \\ C & D & -I \end{pmatrix} \begin{pmatrix} I & 0 & \Lambda^v I \\ 0 & I & 0 \\ \Lambda^z I & 0 & I \end{pmatrix}^{-1} \tag{5.10}$$

by using the identities $[I - \Lambda^v(I+D\Lambda^v)^{-1}(D+\Lambda^z)](I-\Lambda^v\Lambda^z)^{-1} = (I-\Lambda^v D)^{-1}$ and $(I-\Lambda^v\Lambda^z)(I+D\Lambda^v)^{-1}(D+\Lambda^z)(I-\Lambda^v\Lambda^z)^{-1} = (D+\Lambda^z)(I+\Lambda^v D)^{-1}$. The equations (5.8a) and (5.10) imply that (v, x, z) is a solution of $\Sigma(A, B, C, D)$ if and only if (u, x, y) is a solution of $\Sigma(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$. On the other hand, the equation (5.8a) and the condition (5.7c) imply for each $i \in \overline{m}$ and for all t

i. $z_i(t) \leq 0$ if and only if $y_i(t) = 0$ and $u_i(t) \geq 0$ for all sufficiently small ϵ^v and ϵ^z.

ii. $z_i(t) \geq 0$ if and only if $u_i(t) = 0$ and $y_i(t) \geq 0$ for all sufficiently small ϵ^v and ϵ^z.

Therefore, the equivalence of the statements 1 and 2 follows from Definitions 5.3.1. ■

Proof of Theorem 5.4.4: We denote the system $\Sigma(A_{\bar\epsilon} + \rho I, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ by $\Sigma_\rho^{\bar\epsilon}$ in the rest of this proof. Let $(u_{x0}^{\bar\epsilon}, x_{x0}^{\bar\epsilon}, y_{x0}^{\bar\epsilon})$ be defined by

$$\begin{pmatrix} u_{x0}^{\bar\epsilon} \\ x_{x0}^{\bar\epsilon} \\ y_{x0}^{\bar\epsilon} \end{pmatrix} = \begin{pmatrix} I & 0 & \Lambda^v \\ 0 & I & 0 \\ \Lambda^z & 0 & I \end{pmatrix} \begin{pmatrix} v_{x0}^{\bar\epsilon} \\ x_{x0}^{\bar\epsilon} \\ z_{x0}^{\bar\epsilon} \end{pmatrix}. \tag{5.11}$$

Proposition 5.4.1 implies that $(u_{x0}^{\bar\epsilon}, x_{x0}^{\bar\epsilon}, y_{x0}^{\bar\epsilon})$ is the unique solution on $[0, \tau]$ of LCS$(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ with the initial state x_0 in light of Fact 5.4.3 and Lemma 5.3.2.

1: Note that $(u_{x0} + \Lambda^v y_{x0}, x_{x0}, y_{x0} + \Lambda^v u_{x0})$ is an $\mathcal{L}_2$-solution of $\Sigma_0^{\bar\epsilon}$ with the initial state x_0. Hence, $(u_{x0}^{\bar\epsilon} - u_{x0} - \Lambda^v y_{x0}, x_{x0}^{\bar\epsilon} - x_{x0}, y_{x0}^{\bar\epsilon} - y_{x0} - \Lambda^v u_{x0})$ is an $\mathcal{L}_2$-solution of $\Sigma_0^{\bar\epsilon}$ with the zero initial state. Since $\Sigma_0^{\bar\epsilon}$ is uniformly passifiable for all sufficiently small $\bar\epsilon$, there exist a real number ρ and a positive matrix K such that $\Sigma_\rho^{\bar\epsilon}$ is passive with the storage function $x \mapsto x^\top Kx$ for all sufficiently small $\bar\epsilon$. On the other hand, Fact 3.4.1 reveals that $e^{\rho\cdot}(u_{x0}^{\epsilon} - u_{x0} - \Lambda^v y_{x0}, x_{x0}^{\epsilon} - x_{x0}, y_{x0}^{\epsilon} - y_{x0} - \Lambda^v u_{x0})$ is an $\mathcal{L}_2$-solution of $\Sigma_\rho^{\bar\epsilon}$ with the zero initial state. Then, the dissipation inequality yields

$$\int_0^t e^{2\rho s}[u_{x_0}^{\bar\epsilon}(s)-u_{x_0}(s)-\Lambda^v y_{x_0}(s)]^\top[y_{x_0}^{\bar\epsilon}(s)-y_{x_0}(s)-\Lambda^z u_{x_0}(s)]\,ds$$
$$\geq e^{2\rho t}[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)]^\top K[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)] \tag{5.12}$$

for all $t\in[0,\tau]$. Note that

$$\int_0^t e^{2\rho s}[u_{x_0}^{\bar\epsilon}(s)]^\top y_{x_0}^{\bar\epsilon}(s)\,ds=0$$
$$\int_0^t e^{2\rho s}[u_{x_0}^{\bar\epsilon}(s)]^\top y_{x_0}(s)\,ds\geq 0$$
$$\int_0^t e^{2\rho s}[u_{x_0}^{\bar\epsilon}(s)]^\top \Lambda^z u_{x_0}(s)\,ds\geq 0$$
$$\int_0^t e^{2\rho s}[u_{x_0}(s)]^\top y_{x_0}^{\bar\epsilon}(s)\,ds\geq 0$$
$$\int_0^t e^{2\rho s}[u_{x_0}(s)]^\top y_{x_0}(s)\,ds=0$$
$$\int_0^t e^{2\rho s}[\Lambda^v y_{x_0}(s)]^\top y_{x_0}^{\bar\epsilon}(s)\,ds\geq 0$$
$$\int_0^t e^{2\rho s}[\Lambda^v y_{x_0}(s)]^\top \Lambda^z u_{x_0}(s)\,ds=0$$

for all $t\in[0,\tau]$ due to the complementarity conditions and the diagonality of Λ^v and Λ^z. As a consequence of above inequalities, (5.12) yields

$$e^{2\rho t}[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)]^\top K[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)] \tag{5.13}$$
$$\leq \int_0^t e^{2\rho s}[u_{x_0}(s)]^\top\Lambda^z u_{x_0}(s)\,ds+\int_0^t e^{2\rho s}[y_{x_0}(s)]^\top\Lambda^v y_{x_0}(s)\,ds \tag{5.14}$$

for all $t\in[0,\tau]$. Note that without any loss of generality ρ can be taken negative. Thus, $e^{2\rho\tau}\leq e^{2\rho t}\leq 1$ for all $t\in[0,\tau]$. Therefore, we get

$$[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)]^\top K[x_{x_0}^{\bar\epsilon}(t)-x_{x_0}(t)]\leq \alpha_{\tau,x_0}(\bar\epsilon)$$

for all $t\in[0,\tau]$ where $\alpha_{\tau,x_0}(\bar\epsilon)$ converges to zero as $\bar\epsilon$ tends to zero. The above inequality together with the positive definiteness of K immediately implies that $x_{x_0}^{\bar\epsilon}$ converges to x_{x_0} uniformly on $[0,\tau]$.

2: Consider the TCP$(q,S+T)$ where

- $q = Ce^{At}x_0|_{[0,\tau]}$,
- $S = D$, and
- $(Tv)(t) = \int_0^t Ce^{A(t-s)}Bv(s)\,ds$.

Consider also the TCP$(q_{\bar\epsilon}, S_{\bar\epsilon} + T_{\bar\epsilon})$ where

- $q_{\bar\epsilon} = C_{\bar\epsilon}e^{A_{\bar\epsilon}t}x_0|_{[0,\tau]}$,
- $S_{\bar\epsilon} = D_{\bar\epsilon}$, and
- $(T_{\bar\epsilon}v)(t) = \int_0^t C_{\bar\epsilon}e^{A_{\bar\epsilon}(t-s)}B_{\bar\epsilon}v(s)\,ds$.

Proposition 5.7.2 item 1 implies that u_{x_0} and $u^{\bar\epsilon}_{x_0}$ are the unique solutions of TCP$(q, S + T)$ and TCP$(q_{\bar\epsilon}, S_{\bar\epsilon} + T_{\bar\epsilon})$, respectively. Since $v^{\bar\epsilon}_{x_0}$ is bounded by hypothesis, so is $u^{\bar\epsilon}_{x_0}$. Therefore, Lemma 6.6.1 item 1 implies that there exists a subsequence, say $\bar\epsilon_k$, such that $\{u^{\bar\epsilon_k}_{x_0}\}$ converges weakly. Let $\bar u$ denote this limit. On the other hand, it can be checked that

- $T : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ is compact (see [5, Exercise 4.15])
- $S : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ is linear continuous,
- $q_{\bar\epsilon}$ converges to q,
- $S_{\bar\epsilon}$ is linear continuous and moreover is nonnegative definite for all sufficiently small $\bar\epsilon$ due to uniform passifiability,
- TCP$(q_{\bar\epsilon}, S_{\bar\epsilon} + T_{\bar\epsilon})$ is solvable for all sufficiently small $\bar\epsilon$ (by $u^{\bar\epsilon}_{x_0}$), and
- $S_{\bar\epsilon}$ and $T_{\bar\epsilon}$ converge uniformly to S and T respectively.

Theorem 5.7.3 implies, in light of the facts listed above, that $\bar u$ solves TCP$(q, S + T)$. Hence, $(\bar u, \bar x, \bar y)$ is a solution of LCS(A, B, C, D) for some $\bar x$ and $\bar y$ due to Proposition 5.7.2 item 2. We know already that LCS(A, B, C, D) admits unique solutions. Therefore, every weakly convergent subsequence of $\{u^{\bar\epsilon}_{x_0}\}$ has the same limit, namely u_{x_0}. This implies that the sequence $\{u^{\bar\epsilon}_{x_0}\}$ itself converges to u_{x_0} weakly according to Lemma 6.6.1 item 2. Evidently, $\{y^{\bar\epsilon}_{x_0}\}$ converges to y_{x_0} weakly. It follows from (5.11) that $\{(v^{\bar\epsilon}_{x_0}, z^{\bar\epsilon}_{x_0})\}$ converges to (u_{x_0}, y_{x_0}) weakly. ■

We introduce the following notation for the sake of brevity.

Notation 5.7.4 For a given matrix quadruple (A, B, C, D) and K, $\mathcal{K}\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ denotes the matrix

$$\begin{bmatrix} A^\top K + KA & KB - C^\top \\ B^\top K - C & -(D + D^\top) \end{bmatrix}.$$

Proof of Lemma 5.4.5 Note that if the matrix triple (A, B, C) is minimal then so is $(A + \rho I, B, C)$ for all ρ.

1: Since $\Sigma(A + \rho I, B, C, D)$ is passive and $(A + \rho I, B, C)$ is minimal, Lemma 5.2.2 implies that $\mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D \end{smallmatrix}\right)$ is nonpositive definite. Note that

$$\mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D+\epsilon I \end{smallmatrix}\right) = \mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D \end{smallmatrix}\right) + \begin{pmatrix} 0 & 0 \\ 0 & -2\epsilon I \end{pmatrix}.$$

As the sum of two nonpositive matrices, $\mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D+\epsilon I \end{smallmatrix}\right)$ is nonpositive too. It follows from Lemma 5.2.2 that $\Sigma(A+\rho I, B, C, D+\epsilon I)$ is passive with the storage function $x \mapsto \frac{1}{2}x^\top K x$ for all $\epsilon > 0$.

2: Note that

$$\begin{aligned} A_\epsilon &= A - \epsilon BC + \mathcal{O}(\epsilon^2) \\ B_\epsilon &= B - \epsilon BD + \mathcal{O}(\epsilon^2) \\ C_\epsilon &= C - \epsilon DC + \mathcal{O}(\epsilon^2) \\ D_\epsilon &= D + \epsilon(I - D^2) + \mathcal{O}(\epsilon^2). \end{aligned} \tag{5.15}$$

Hence, we get

$$\mathcal{K}\left(\begin{smallmatrix} A_\epsilon+\rho' I & B_\epsilon \\ C_\epsilon & D_\epsilon \end{smallmatrix}\right) = \mathcal{K}\left(\begin{smallmatrix} A+\rho' I & B \\ C & D \end{smallmatrix}\right) + \epsilon\mathcal{K}\left(\begin{smallmatrix} -BC & BD \\ DC & I-D^2 \end{smallmatrix}\right) + \mathcal{O}(\epsilon^2).$$

Let $\mathrm{col}(x, u)$ be such that

$$\begin{pmatrix} x \\ u \end{pmatrix}^\top \mathcal{K}\left(\begin{smallmatrix} A+\rho' I & B \\ C & D \end{smallmatrix}\right) \begin{pmatrix} x \\ u \end{pmatrix} = 0.$$

Note that

$$\mathcal{K}\left(\begin{smallmatrix} A+\rho' I & B \\ C & D \end{smallmatrix}\right) = \mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D \end{smallmatrix}\right) + \begin{pmatrix} 2(\rho - \rho')I & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\Sigma(A + \rho I, B, C, D)$ is passive and $\rho - \rho' < 0$, the two summands on the right hand side of above equation are both nonpositive definite. Therefore, we get

$$\begin{pmatrix} x \\ u \end{pmatrix}^\top \mathcal{K}\left(\begin{smallmatrix} A+\rho I & B \\ C & D \end{smallmatrix}\right) \begin{pmatrix} x \\ u \end{pmatrix} = 0$$

$$\begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{pmatrix} 2(\rho - \rho')I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} = 0$$

The latter equation implies $x=0$ and hence we get $u^\top Du = 0$ from the former one. Since D is nonnegative definite due to the passifiability of $\Sigma(A,B,C,D)$, $u^\top Du=0$ implies that $(D+D^\top)u=0$. Therefore, $u^\top D^2 u = -\|Du\|^2$ whenever $u^\top Du = 0$. This means that we have the following implication

$$u \neq 0,\ u^\top Du = 0 \Rightarrow u^\top(I-D^2)u > 0. \tag{5.16}$$

Hence, we can apply Lemma 3.8.3 by taking $M = -\mathcal{K}\left(\begin{smallmatrix} A+\rho' I & B \\ C & D \end{smallmatrix}\right)$ and $N = -\mathcal{K}\left(\begin{smallmatrix} -BC & BD \\ DC & I-D^2 \end{smallmatrix}\right)$. Indeed, M is nonnegative definite and $v \neq 0$, $v^\top Mv = 0$ implies $v^\top Nv > 0$. Then, we know that there exists $\mu < 0$ such that $\mathcal{K}\left(\begin{smallmatrix} A+\rho' I & B \\ C & D \end{smallmatrix}\right) + \epsilon\mathcal{K}\left(\begin{smallmatrix} -BC & BD \\ DC & I-D^2 \end{smallmatrix}\right) \leq \mu\epsilon I$ for all sufficiently small ϵ. Since $\epsilon \mapsto \mathcal{K}\left(\begin{smallmatrix} A_\epsilon+\rho' I & B_\epsilon \\ C_\epsilon & D_\epsilon \end{smallmatrix}\right)$ is continuous and $\mathcal{K}\left(\begin{smallmatrix} A_\epsilon+\rho' I & B_\epsilon \\ C_\epsilon & D_\epsilon \end{smallmatrix}\right) \leq \mu\epsilon + \mathcal{O}(\epsilon^2)$, we can conclude that $\mathcal{K}\left(\begin{smallmatrix} A_\epsilon+\rho' I & B_\epsilon \\ C_\epsilon & D_\epsilon \end{smallmatrix}\right)$ is nonpositive definite for all sufficiently large σ. Note that the set of all minimal matrix triples is open. Hence, $(A_\epsilon, B_\epsilon, C_\epsilon)$ is minimal due to the continuity of the mappings $\epsilon \mapsto A_\epsilon$, $\epsilon \mapsto B_\epsilon$, $\epsilon \mapsto C_\epsilon$ and $\epsilon \mapsto D_\epsilon$. Consequently, Lemma 5.2.2 implies that $\Sigma(A_\epsilon + \rho' I, B_\epsilon, C_\epsilon, D_\epsilon)$ is passive with the storage function $x \mapsto \frac{1}{2}x^\top Kx$. ∎

Proof of Corollary 5.4.6:

1: Note that $\bar{\epsilon} = \mathrm{col}(0, \epsilon\,\iota)$ and hence $\Lambda_v = 0$ and $\Lambda_z = \epsilon I$ for this case. Then, $(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon}) = (A, B, C, D+\epsilon I)$. It follows from Lemma 5.4.5 item 1 that the sequence of systems $\Sigma(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ is uniformly passifiable for all sufficiently small ϵ. Clearly, $D_{\bar\epsilon}$ is positive definite for all $\epsilon > 0$. Hence, $\mathcal{Q}^*_{D_{\bar\epsilon}} = \mathbb{R}^m$ for all sufficiently small ϵ. Then, the rest follows from the application of Theorem 5.4.4.

2: Note that $\bar\epsilon = \mathrm{col}(\epsilon\,\iota, \epsilon\,\iota)$ and hence $\Lambda_v = \Lambda_z = \epsilon I$ for this case. Then, $(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ $= (A_\epsilon, B_\epsilon, C_\epsilon, D_\epsilon)$. It follows from Lemma 5.4.5 item 2 that the sequence of systems $\Sigma(A_{\bar\epsilon}, B_{\bar\epsilon}, C_{\bar\epsilon}, D_{\bar\epsilon})$ is uniformly passifiable for all sufficiently small ϵ. On the other hand, it follows from (5.16) and Lemma 3.8.3 that $D + \epsilon(I - D^2) \geq \mu\epsilon I$ for some positive μ. Then, (5.15) and the fact that $\epsilon \mapsto D_{\bar\epsilon}$ is continuous imply that $D_{\bar\epsilon}$ is positive definite for all sufficiently small ϵ. Hence, $\mathcal{Q}^*_{D_{\bar\epsilon}} = \mathbb{R}^m$ for all sufficiently small ϵ. Then, the rest follows from the application of Theorem 5.4.4. ∎

Proof of Theorem 5.4.8: Consider the TCP$(q, S + CT)$ where

- $q = Ce^{At}x_0|_{[0,\tau]}$,
- $S = D$, and
- $(Tv)(t) = \int_0^t e^{A(t-s)}Bv(s)\,ds$.

Consider also the TCP$(q_\epsilon, S_\epsilon + CT_\epsilon)$ where

- $q_\epsilon = C_\epsilon e^{A_\epsilon t} x_0|_{[0,\tau]}$,
- $S_\epsilon = D_\epsilon$, and
- $(T_\epsilon v)(t) = \int_0^t e^{A_\epsilon(t-s)} B_\epsilon v(s)\, ds$.

Proposition 5.7.2 item 1 implies that u_{x_0} and $u^\epsilon_{x_0}$ are solutions of TCP$(q, S + CT)$ and TCP$(q_\epsilon, S_\epsilon + C_\epsilon T_\epsilon)$, respectively. Since $\{u^\epsilon_{x_0}\}$ is bounded, it is known from Lemma 6.6.1 item 1 that $\{u^\epsilon_{x_0}\}$ has a weakly convergent subsequence, say $\{u^{\epsilon_k}_{x_0}\}$. Let $\bar{u}$ denote the weak limit of this subsequence. Note that

- $T : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ is compact (see [5, Exercise 4.15])
- $S : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ is linear continuous,
- q_ϵ converges to q,
- S_ϵ is linear continuous and moreover is positive definite for all sufficiently small ϵ by the hypothesis,
- TCP$(q_\epsilon, S_\epsilon + C_\epsilon T_\epsilon)$ is solvable for all sufficiently small ϵ (by $u^\epsilon_{x_0}$), and
- S_ϵ and T_ϵ converge uniformly to S and T respectively.

Therefore Theorem 5.7.3 implies, in light of the facts listed above, that $\bar{u}$ solves TCP$(q, S + CT)$. Hence, $(\bar{u}, \bar{x}, \bar{y})$ is a solution of LCS(A, B, C, D) for some $\bar{x}$ and $\bar{y}$ due to Proposition 5.7.2 item 2. We know already that LCS(A, B, C, D) admits unique solutions due to Lemma 5.3.2 since $\Sigma(A, B, C, D)$ is passifiable by pole shifting. Therefore, every weakly convergent subsequence of $\{u^\epsilon_{x_0}\}$ has the same limit, namely u_{x_0}. This implies that the sequence $\{u^\epsilon_{x_0}\}$ itself converges to u_{x_0} weakly according to Lemma 6.6.1 item 2. Evidently, $\{y^\epsilon_{x_0}\}$ converges to y_{x_0} weakly. It remains to show that $\{x^\epsilon_{x_0}\}$ converges (strongly) to x_{x_0}. Indeed, we have

$$x^\epsilon_{x_0} = e^{A_\epsilon t} x_0|_{[0,\tau]} + T_\epsilon u^\epsilon_{x_0}$$

according to Proposition 5.7.2 item 2. Note that $\{e^{A_\epsilon t} x_0|_{[0,\tau]}\}$ converges to $e^{At} x_0|_{[0,\tau]}$ as ϵ tends to zero. Since $\{T_\epsilon\}$ converges uniformly to T, we know that $\{T_\epsilon u^\epsilon_{x_0} - T u^\epsilon_{x_0}\}$ converges to zero. It follows from the compactness of T and the weak convergence of $\{u^\epsilon_{x_0}\}$ that $\{T u^\epsilon_{x_0}\}$ converges (strongly) to $T u_{x_0}$. Hence, $\{x^\epsilon_{x_0}\}$ converges strongly to $e^{At} x_0|_{[0,\tau]} + T u_{x_0}$. Since u_{x_0} is a solution of TCP(q, CT), we have $x_{x_0} = e^{At} x_0|_{[0,\tau]} + T u_{x_0}$ due to Proposition 5.7.2 item 2. ■

References

[1] B. Brogliato. *Nonsmooth Impact Mechanics.* Springer-Verlag, London, 1996.

[2] J. de Does and J.M. Schumacher. Continuity of singular perturbations in the graph topology. *Linear Algebra and Its Applications*, 205:1121–1143, 1994.

[3] P. Kokotovic, H. K. Khalil, and J. O'Reilly. *Singular Perturbation Methods in Control: Analysis and Design.* SIAM, Philadephia, 1999.

[4] J.J. Moreau. Numerical aspects of the sweeping process. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):329–349, 1999.

[5] W. Rudin. *Functional Analysis.* McGraw-Hill, New York, 1977.

[6] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice-Hall, New Jersey, 1993.

[7] J. C. Willems. Dissipative dynamical systems. *Arch. Rational Mech. Anal.*, 45:321–393, 1972.

[8] J. C. Willems and J.W. Nieuwenhuis. Continuity of latent variable models. *IEEE Transactions on Automatic Control*, 36:528–538, 1991.

Chapter 6

Consistency of Backward Euler Method for Linear Networks with Ideal Diodes

6.1 Introduction

Simulation of switching networks is a problem that has been studied extensively in circuit theory [1,2,7,14,15,19,24,31]. Roughly speaking, there are two main approaches, namely event-tracking (see e.g. [1,19]) and time-stepping methods (see [2,14,15,24] for electrical networks and [17,20,21,28,30] for unilaterally constrained mechanical systems with friction phenomena). Representing a hybrid systems point of view (see for instance [27]), event-tracking methods are based on the idea of solving corresponding DAEs of the current circuit topology (called 'mode' in the hybrid systems terminology), monitoring possible changes of circuit topology (mode transition), and (if necessary) determining the exact time (event time) instant of the change of topology and the next topology. Time-stepping methods differ from this scheme by putting aside the hybrid features, and by regarding the whole system as a collection of differential equations with constraints and trying to approximate the solutions of these differential equations with constraints. As a consequence of this point of view, there is no need to locate exact event times. However, the convergence of the approximations in a suitable sense has to be guaranteed. Since the methods seem to work well in practice, the question of convergence is usually neglected in the literature. It is the objective of this chapter to provide a rigorous basis for the use of time-stepping methods in the simulation of circuits with state events.

In Chapters 2 and 3 (see also [4,5,9]) the meaning of a transient true solution to linear dynamical network models with ideal diodes has already been established. Using techniques borrowed from the theory of linear complementarity systems (LCS) [10, 11, 16, 25, 26], existence and uniqueness of solutions have been proven under mild conditions. Moreover, several regularity properties have been shown from which this chapter will benefit.

The particular time-stepping method that we will study here is based on the well-known backward Euler scheme and has been described, for instance, in [2, 14, 15] for electrical networks. Similar methods have been used in a mechanical context in [17,20,21,28,30]. The advantage of the method is that it is straightforward to implement and many algorithms (e.g. Lemke's algorithm [6], Katzenelson's algorithm [13] and others [15]) are available to solve the one-step problems consisting of linear complementarity problems (LCPs).

In [14] the use of a time-stepping method based on the backward Euler scheme (or higher order linear multistep integration methods [8] like the trapezoidal rule) has already been proposed for the class of linear complementarity systems, i.e., linear time-invariant dynamical systems coupled with ideal diode characteristics (complementarity conditions). By an example (cf. Example 6.3.3 below), it will be shown that the method is not suited for the general class of linear complementarity systems. This example indicates also that, although the method has proven itself in practice, one should not indiscriminately apply it to general discontinuous dynamical systems.

Convergence problems of time-stepping methods for mechanical systems subject to unilateral constraints or friction have been studied by Stewart [28, 29]. He shows that for a broad class of nonlinear constrained mechanical systems there always exists a *subsequence* of approximating time functions that converge to a real solution of the mechanical model. However, the convergence of the complete sequence has not been shown in [28, 29]. The conditions used in [28, 29] do not cover electrical networks containing ideal diodes, which form the subject of this paper. Specifically, we will show that for the class of discontinuous dynamical systems consisting of linear electrical passive circuits with ideal diodes the backward Euler time-stepping method is consistent. To be specific, we prove that the whole sequence (and not only a subsequence) of the approximating time functions converges to the real transient solution of the network model, when the step size decreases to zero. Although the results are written down here for networks containing ideal diodes (internally controlled switches) only, externally controlled switches can easily be included without destroying the convergence proof. The results presented here form a justification of the backward Euler time-stepping scheme in the field of switched electrical networks. Such a justification seems required considering the problems that might occur due to changing configurations of the network, the possibility of Dirac impulses and the discontinuities of the system's variables.

The outline of the chapter is as follows. In Section 6.2 preliminaries on linear complementarity systems and passivity are stated. The time-stepping method that will be studied is considered in Section 6.3. Moreover, a result on consistency of the numerical method is formulated for a general class of linear complementarity systems. In the next section, this result is applied to linear passive complementarity systems. The continuous dependence of solution trajectories on the initial states is also mentioned in section 6.4. The conclusions follow in Section 6.5. The proofs of the main results can be found in Section 6.6.

6.2 Preliminaries

As discussed in [5,9], linear networks with ideal diodes can be modeled as linear complementarity systems (see [10,11,25,26] for detailed discussion), which are dynamical versions of the linear complementarity problem. They are of the form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{6.1a}$$

$$y(t) = Cx(t) + Du(t) \tag{6.1b}$$

$$0 \leq u(t) \perp y(t) \geq 0 \tag{6.1c}$$

where $u(t) \in \mathbb{R}^m$, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^m$ and A, B, C, and D are matrices of appropriate dimensions. We denote (6.1a)-(6.1b) by $\Sigma(A, B, C, D)$ and (6.1) by LCS(A, B, C, D).

Next, we recall (see [10, 11]) the notion of *initial solution* which is of considerable importance in the analysis of linear complementarity systems. Notice that the following definition is slightly more general than the ones we have worked with in Chapters 2 and 3 in the sense that it allows the presence of Dirac distributions.

Definition 6.2.1 The triple $(\mathtt{u}, \mathtt{x}, \mathtt{y})^1 \in \mathcal{B}_\delta^{m+n+m}$ is an *initial solution* of LCS(A, B, C, D) with initial state x_0 if there exists an index set $K \subseteq \overline{m}$ such that

$$\begin{aligned}
\dot{\mathtt{x}} &= A\mathtt{x} + B\mathtt{u} + x_0\delta \\
\mathtt{y} &= C\mathtt{x} + D\mathtt{u} \\
\mathtt{u}_i &= 0 \text{ if } i \in K \\
\mathtt{y}_i &= 0 \text{ if } i \notin K
\end{aligned}$$

hold in the distributional sense, and $\mathtt{u}$ and $\mathtt{y}$ are initially nonnegative.

It can be shown (see for instance [10, 26]) that there is a one-to-one relation between the initial solutions to LCS(A, B, C, D) with initial state x_0 and the *proper* solutions of the so-called *rational complementarity problem.*

Problem 6.2.2 (RCP(x_0, A, B, C, D)) Given $x_0 \in \mathbb{R}^n$ and (A, B, C, D) with $A \in \mathbb{R}^{n \times n}$,

[1] Throughout this chapter, we use typewriter font for the distributions.

$B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$ and $D \in \mathbb{R}^{m\times m}$, find $\hat{u}(s) \in \mathbb{R}^m(s)$ and $\hat{y}(s) \in \mathbb{R}^m(s)$ such that

$$\hat{y}(s) = C(sI - A)^{-1}x_0 + [C(sI - A)^{-1}B + D]\hat{u}(s)$$
$$\hat{u}(s) \perp \hat{y}(s)$$

for all $s \in \mathbb{C}$ and $\hat{u}(\sigma) \geq 0$ and $\hat{y}(\sigma) \geq 0$ for all sufficiently large $\sigma \in \mathbb{R}$.

We say that $\hat{u}(s)$ is a solution of (or $\hat{u}(s)$ solves) RCP if it satisfies above conditions. In a similar fashion, we sometimes also write $(\hat{u}(s), \hat{y}(s))$ is a solution of (or solves) RCP.

The following proposition states the above mentioned one-to-one relation which is given by the Laplace transform and its inverse. This connection indicates the relevance of the rational complementarity problem to the study of LCS.

Proposition 6.2.3 [10] *The triple* $(\mathrm{u}, \mathrm{x}, \mathrm{y})$ *is an initial solution of LCS*(A, B, C, D) *with initial state* x_0 *if and only if its Laplace transform* $(\hat{u}(s), \hat{x}(s), \hat{y}(s))$ *is such that* $(\hat{u}(s), \hat{y}(s))$ *is a proper solution of RCP*(x_0, A, B, C, D) *and* $\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)$.

Now, we can give a precise definition of what is meant by solution of LCS(A, B, C, D). For a more detailed discussion see Chapter 3.

Definition 6.2.4 The triple $(\mathrm{u}, \mathrm{x}, \mathrm{y}) \in \mathcal{L}_2^{\delta}([0, \tau], \mathbb{R}^{m+n+m})$ is a *(global) solution* on $[0, \tau]$ of LCS(A, B, C, D) with initial state x_0 if the following conditions hold.

1. There exists an initial solution $(\bar{\mathrm{u}}, \bar{\mathrm{x}}, \bar{\mathrm{y}})$ such that

$$(\mathrm{u}_{\mathrm{imp}}, \mathrm{x}_{\mathrm{imp}}, \mathrm{y}_{\mathrm{imp}}) = (\bar{\mathrm{u}}_{\mathrm{imp}}, \bar{\mathrm{x}}_{\mathrm{imp}}, \bar{\mathrm{y}}_{\mathrm{imp}}).$$

2. The equations

$$\dot{\mathrm{x}} = A\mathrm{x} + B\mathrm{u} + x_0\delta$$
$$\mathrm{y} = C\mathrm{x} + D\mathrm{u}$$

hold in the distributional sense.

3. For almost all $t \in [0, \tau]$, $0 \leq \mathrm{u}_{\mathrm{reg}}(t) \perp \mathrm{y}_{\mathrm{reg}}(t) \geq 0$.

In the sequel, we are mainly concerned with linear passive complementarity systems. For ease of readability, we shall quickly review the notion of passivity and its characterizations in terms of the state representation and the transfer matrix of the system.

Definition 6.2.5 [32] The system $\Sigma(A, B, C, D)$ given by (6.1a)-(6.1b) is said to be *passive (dissipative with respect to the supply rate $u^\top y$)* if there exists a function $V : \mathbb{R}^n \to \mathbb{R}_+$

(called a *storage function*), such that

$$V(x(t_0)) + \int_{t_0}^{t_1} u^\top(t)y(t)dt \geq V(x(t_1))$$

holds for all t_0 and t_1 with $t_1 \geq t_0$, and all $(u, x, y) \in \mathcal{L}_2([t_0, t_1], \mathbb{R}^{m+n+m})$ satisfying (6.1a)-(6.1b).

We repeat a well-known theorem on passive systems which is sometimes called the positive real lemma.

Lemma 6.2.6 [32] *Assume that (A, B, C) is minimal. Then the following statements are equivalent:*

1. *$\Sigma(A, B, C, D)$ is passive.*
2. *The matrix inequalities*

$$K = K^\top \geq 0 \text{ and } \begin{bmatrix} A^\top K + KA & KB - C^\top \\ B^\top K - C & -(D + D^\top) \end{bmatrix} \leq 0$$

have a solution.

3. *$G(s)$ is positive real, i.e., $G(\lambda) + G^\top(\bar{\lambda}) \geq 0$ for all $\lambda \in \mathbb{C}$ with $\lambda \notin \sigma(A)$ and* $\mathrm{Re}(\lambda) > 0$.

Moreover, if $\Sigma(A, B, C, D)$ is passive all solutions of the matrix inequalities in item 2 are positive definite.

Throughout the paper, we will frequently use the following assumption.

Assumption 6.2.7 (A, B, C) is a minimal representation and B is of full column rank.

The following theorem is quoted from Chapter 3 and deals with the existence and uniqueness of solutions to linear passive complementarity systems.

Theorem 6.2.8 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Let $\tau > 0$ be given. For each x_0, there exists a unique solution $(\mathrm{u}, \mathrm{x}, \mathrm{y}) \in \mathcal{L}_2^\delta([0, \tau], \mathbb{R}^{m+n+m})$ on $[0, \tau]$ of LCS(A, B, C, D) with initial state x_0.*

6.3 The Backward Euler Time-stepping Method

For the numerical approximation of the solutions of switched electrical networks the following time-stepping scheme has frequently been used [2, 14, 15, 24]. For LCS the method consists of discretizing the system description by applying the well known backward Euler integration routine and imposing the complementarity conditions at every time step. This comes down to the computation of u^h_{k+1}, y^h_{k+1}, and x^h_{k+1} given x^h_k through the linear complementarity problem given by

$$\frac{\mathrm{x}^h_{k+1} - \mathrm{x}^h_k}{h} = A\mathrm{x}^h_{k+1} + B\mathrm{u}^h_{k+1} \tag{6.2a}$$

$$\mathrm{y}^h_{k+1} = C\mathrm{x}^h_{k+1} + D\mathrm{u}^h_{k+1} \tag{6.2b}$$

$$0 \leq \mathrm{y}^h_{k+1} \perp \mathrm{u}^h_{k+1} \geq 0. \tag{6.2c}$$

Here $\bullet^h_k$ denotes the value at the kth step of the corresponding variable for the step size $h > 0$. Based on this scheme, one can construct approximations of the transient response of a LCS by applying the algorithm below.

Algorithm 6.3.1 $(\{\mathrm{u}^h_k\}, \{\mathrm{x}^h_k\}, \{\mathrm{y}^h_k\}) = \mathrm{Approximation}(A, B, C, D, \tau, h, x_0)$

1. $N_h = \lfloor \frac{\tau}{h} \rfloor$
2. $\mathrm{x}^h_{-1} := x_0$
3. $k := -1$
4. solve the *one-step problem*

$$\mathrm{y}^h_{k+1} = C(I - hA)^{-1}\mathrm{x}^h_k + [D + hC(I - hA)^{-1}B]\mathrm{u}^h_{k+1}$$
$$0 \leq \mathrm{u}^h_{k+1} \perp \mathrm{y}^h_{k+1} \geq 0$$

5. $\mathrm{x}^h_{k+1} := (I - hA)^{-1}\mathrm{x}^h_k + h(I - hA)^{-1}B\mathrm{u}^h_{k+1}$
6. $k := k + 1$
7. if $k < N_h$ goto 4
8. stop.

The one-step problem is given by a linear complementarity problem in step 4. In general a linear complementarity problem may have multiple solutions or have no solutions at all. We shall proceed by assuming unique solvability of the problem. The assumption is introduced here for reasons of generality; later on we will prove that the assumption is implied by passivity.

Assumption 6.3.2 For all sufficiently small $h > 0$, $\mathrm{LCP}(C(I - hA)^{-1}\bar{x}, G(h^{-1}))$ has a unique solution for all $\bar{x}$, where $G(s)$ is given by $D + C(sI - A)^{-1}B$.

This assumption implies that for all sufficiently small $h > 0$, Algorithm 6.3.1 generates an output, which is unique. Hence, for a given step size $h > 0$ (sufficiently small), we can define the approximations $(\mathrm{u}^h, \mathrm{x}^h, \mathrm{y}^h) \in \mathcal{L}_2^\delta([0,\tau], \mathbb{R}^{m+n+m})$ given by

$$\mathrm{u}^h_{\mathrm{imp}} = h\mathrm{u}^h_0\delta \tag{6.3a}$$

$$\mathrm{x}^h_{\mathrm{imp}} = h\mathrm{x}^h_0\delta \tag{6.3b}$$

$$\mathrm{y}^h_{\mathrm{imp}} = h\mathrm{y}^h_0\delta \tag{6.3c}$$

$$\left.\begin{aligned} \mathrm{u}^h_{\mathrm{reg}}(t) &= \mathrm{u}^h_l \\ \mathrm{x}^h_{\mathrm{reg}}(t) &= \mathrm{x}^h_l \\ \mathrm{y}^h_{\mathrm{reg}}(t) &= \mathrm{y}^h_l \end{aligned}\right\} \quad \text{whenever } (l-1)h \leq t < lh, \tag{6.3d}$$

where u^h_k, x^h_k and y^h_k, $k = 0, 1, \dots, N_h$ have been obtained from Algorithm 6.3.1. One of the main goals of the paper is to prove that for a passive system these piecewise constant approximations converge in a suitable sense to the actual solution of the system. This property is called *consistency* of the numerical method. In the following example, we illustrate that Algorithm 6.3.1 is not always consistent even if Assumption 6.3.2 holds.

Example 6.3.3 Consider the linear complementarity system (consisting of a triple integrator with complementarity conditions)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u \\ y &= x_1 \\ 0 \leq u &\perp y \geq 0 \end{aligned}$$

with the initial state $x_0 = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix}^\top$. As we have already mentioned before, Definition 6.2.4 is a simplified version of the general one given in [11] for linear passive complementarity systems. Since the triple integrator is not a passive system, we must utilize the general definition rather than the simplified one. Indeed, it can be checked that $(\mathrm{u}, \mathrm{x}, \mathrm{y}) = (\dot{\delta}, 0, 0)$, which does not qualify as a solution in the sense of Definition 6.2.4, is the 'true' global solution of the system with the given initial state. Here $\dot{\delta}$ denotes the distributional derivative of the Dirac impulse δ. Algorithm 6.3.1 gives

$$(\mathrm{u}^h_k, \mathrm{y}^h_k) = \begin{cases} (h^{-2}, 0) \text{ if } k = 0 \\ (0, \frac{k(k+1)}{2}h) \text{ if } k \neq 0. \end{cases}$$

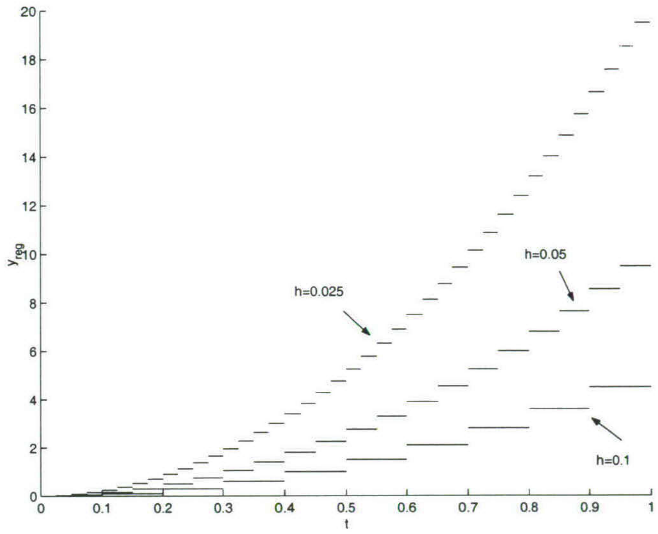


Figure 6.1: Nonconvergence of backward Euler approximations for the triple integrator with diode.

It follows from (6.3d) that

$$\|\mathrm{y}^h_{\mathrm{reg}}\| \geq \Big(\int_{(N_h-2)h}^{(N_h-1)h} \|\mathrm{y}^h_{(N_h-1)}\|^2\,dt\Big)^{1/2} = \frac{(N_h-1)N_h}{2}h^{3/2} = O(h^{-1/2})$$

whenever $N_h \geq 2$. Therefore, $\mathrm{y}^h_{\mathrm{reg}}$ is far from being convergent as it is not bounded as h converges to zero. For three different values of h, the trajectories of $\mathrm{y}^h_{\mathrm{reg}}$ on $[0,1]$ are depicted in Figure 6.1.

This example indicates that one should be cautious in applying a time-stepping method to a general LCS. As a consequence, verification of the numerical scheme in the sense of showing consistency is needed. Before passing the following theorem which states conditions that imply consistency, we need to introduce some nomenclature. We say that the sequence of distributions $\{u^k_0\delta + \mathrm{u}^k_{\mathrm{reg}}\} \subset \mathcal{L}^\delta_2(\Omega,\mathbb{R})$ converges (weakly) to $u_0\delta + \mathrm{u}_{\mathrm{reg}}$, if $\{u^k_0\}$ converges to u_0 and $\{\mathrm{u}^k_{\mathrm{reg}}\}$ converges (weakly) to $\mathrm{u}_{\mathrm{reg}}$ in $\mathcal{L}_2$-sense.

Theorem 6.3.4 *Consider LCS(A,B,C,D) such that Assumption 6.3.2 holds. Let $\tau > 0$ and $x_0 \in \mathbb{R}^n$ be given. Also let $(\mathrm{u}^h, \mathrm{x}^h, \mathrm{y}^h)$ be given by (6.3) via Algorithm 6.3.1. Suppose that there exists $\alpha > 0$ such that for all sufficiently small h*

$$\|h\mathrm{u}^h_0\| \leq \alpha \text{ and } \|\mathrm{u}^h_{\mathrm{reg}}\| \leq \alpha.$$

Then, we have the following statements:

1. *There exists a unique initial solution of LCS(A, B, C, D) with initial state x_0 in the sense of Definition 6.2.1 and the impulsive part of this solution is of the form $(u_0\delta, 0, y_0\delta)$ with $u_0, y_0 \in \mathbb{R}^m$.*

2. *As h tends to zero, $\{(\mathrm{u}^h_{\mathrm{imp}}, \mathrm{x}^h_{\mathrm{imp}}, \mathrm{y}^h_{\mathrm{imp}})\}$ converges to $(\mathrm{u}_{\mathrm{imp}}, 0, \mathrm{y}_{\mathrm{imp}})$ where $(\mathrm{u}_{\mathrm{imp}}, \mathrm{y}_{\mathrm{imp}}) = (u_0\delta, y_0\delta)$.*

3. *Let $\{h_k\}$ converge to zero. Suppose that D is nonnegative definite. Then the following holds.*

   (a) *There exists a subsequence $\{h_{k_l}\} \subseteq \{h_k\}$ such that $(\{\mathrm{u}^{h_{k_l}}\}, \{\mathrm{y}^{h_{k_l}}\})$ converges weakly to some (u, y) and $\{\mathrm{x}^{h_{k_l}}\}$ converges to some x.*

   (b) *$(\mathrm{u}, \mathrm{x}, \mathrm{y})$ is a solution of LCS(A, B, C, D) on $[0, \tau]$ with the initial state x_0.*

   (c) *If the solution $(\mathrm{u}, \mathrm{x}, \mathrm{y})$ is unique for the initial state x_0 in the sense of Definition 6.2.4, then the complete sequence $(\{\mathrm{u}^{h_k}\}, \{\mathrm{y}^{h_k}\})$ converges weakly to (u, y) and $\{\mathrm{x}^{h_k}\}$ converges to x.*

6.4 Main Results for LPCS

In Section 6, we shall show that the conditions of Theorem 6.3.4 are satisfied in the case of passive complementarity systems so that the following result holds.

Theorem 6.4.1 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Let $\tau > 0$ and $x_0 \in \mathbb{R}^n$ be given. Let $(\mathrm{u}, \mathrm{x}, \mathrm{y})$ be the solution of LCS(A, B, C, D) on $[0, \tau]$ with the initial state x_0. Also let $(\mathrm{u}^h, \mathrm{x}^h, \mathrm{y}^h)$ be given by (6.3) via Algorithm 6.3.1. Then, $(\{\mathrm{u}^h\}, \{\mathrm{y}^h\})$ converges weakly to (u, y) and $\{\mathrm{x}^h\}$ converges to x as the step size h tends to zero.*

The above theorem assumes exact computations. In implementing the backward Euler time-stepping method numerical errors will of course be introduced. To give some justification that also in the case of (small) numerical errors the method is still suitable, we study the issue of the dependence of the solution trajectories on the initial conditions. For general LCS such a property does not hold. However, in the special case of linear passive complementarity systems, the continuous dependence holds. To formulate this in a mathematically precise way, we have to introduce some nomenclature. Let $\mathcal{H}$ be a Hilbert space. We say that $T : \mathbb{R}^n \to \mathcal{H}$ is *continuous* (*weakly continuous*), if continuity is considered with respect to the strong (weak) topology on $\mathcal{H}$. In other words, T is continuous (weakly continuous), if for all convergent (weakly convergent) sequences $\{x_k\}$, $\{Tx_k\}$ converges (converges weakly) to Tx^* where $x^* = \lim_{k\to\infty} x_k$.

Theorem 6.4.2 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Let $\tau > 0$ be given. Define the operators $x_0 \mapsto (\mathtt{u}, \mathtt{y})$ and $x_0 \mapsto \mathtt{x}$, where $(\mathtt{u}, \mathtt{x}, \mathtt{y})$ is the solution of LCS(A, B, C, D) on $[0, \tau]$ with the initial state x_0. The operators $x_0 \mapsto (\mathtt{u}, \mathtt{y})$ and $x_0 \mapsto \mathtt{x}$ are weakly continuous and continuous, respectively.*

Note that Theorem 6.4.2 is not a property of the numerical scheme, but of the class of LCS satisfying a passivity assumption. Of course one may look for schemes that perform particularly well in coping with numerical errors, but this is outside the scope of the present paper.

6.5 Conclusions

In this chapter, we studied the consistency of a time-stepping method based on the backward Euler integration routine. The method has proven itself already in practice for the transient simulation of piecewise linear electrical circuits and constrained mechanical systems. However, one cannot indiscriminately apply this method for general classes of discontinuous systems as shown by an example in this paper. The main result of the chapter has presented a rigorous proof of the consistency of the backward Euler time-stepping method when applied to the class of linear passive electrical networks with ideal diodes. In spite of the mixed continuous and discrete behaviour of the circuit and the possibility of Dirac impulses occurring at the initial time, we have shown the convergence of the approximations to the actual transient solution of the network model. Using almost the same arguments, we have also proven the continuous dependence of the transient solutions on the initial state. For simulation of linear passive networks with ideal diodes, this has the important consequence that numerical errors do not have a large influence on the outcomes of the approximation method. These results provide a justification for the use of time-stepping methods.

Of course, it would be interesting to generalize these results to other systems of a mixed continuous and discrete nature. In particular, we are currently studying the consistency of the backward Euler method for dynamical systems with relays and for other linear complementarity systems. For many systems where the backward Euler time-stepping scheme does not generate proper output (like the triple integrator), it is useful to consider extensions of the time-stepping algorithm that are consistent.

6.6 Proofs

The outline of this section, in which we give the proofs of the results presented in the previous sections, is as follows. We begin with some preliminaries that will be employed in the sequel. The proofs of items 1 and 2 of Theorem 3.4 will be followed by a recall of the so-called topological complementarity problem (TCP) which is the main tool used in proving item 3 of Theorem 3.4, Theorem 4.1 and Theorem 4.2. After fitting the solution concept as well as the approximations into a TCP framework, we present a general result (Theorem 6.9) concerning the convergence of the solutions to TCPs and deduce the proof of item 3 of Theorem 3.4 from this result. Finally, the section ends with some technical lemmas on LCPs and the proofs of Theorem 4.1 and Theorem 4.2 as inferred from these lemmas and the result on the convergence of the solutions to TCPs.

6.6.1 Preliminaries

For ease of reference, we recall some standard results on weakly convergent sequences.

Lemma 6.6.1 *The following statements hold in every real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$.*

1. *Every bounded sequence has a weakly convergent subsequence.*
2. *If all weakly convergent subsequences of a bounded sequence have the same weak limit, then the sequence itself converges weakly to this limit.*
3. *Assume that $\{v_k\} \subset \mathcal{H}$ converges weakly to v and $\{w_k\} \subset \mathcal{H}$ converges to w. Then*
   (a) *There exists $\alpha > 0$ such that $\|v_k\| \leq \alpha$ for all k and $\|v\| \leq \alpha$.*
   (b) *$\{Sv_k\}$ converges weakly to Sv whenever $S : \mathcal{H} \to \mathcal{H}$ is a continuous linear operator.*
   (c) *$\{\langle v_k, w_k \rangle\}$ converges to $\langle v, w \rangle$.*

Proof: For the proofs of the statements 1, 3a, 3b, and 3c see Theorem 3.7, Exercise 3.3.10a and Proposition 3.6, Proposition 3.8, and Exercise 3.3.10b of [12], respectively. For the proof of the statement 2, let $\{v_k\} \in \mathcal{H}$ be such a sequence. Without loss of generality, we may assume that the limit of all its weakly convergent subsequences is zero. If the sequence itself is weakly convergent then its weak limit is zero since every sequence is a subsequence of itself. Suppose that the sequence does not weakly converge to zero. Then there exist $\epsilon > 0$, $w \in \mathcal{H}$ and a subsequence, say $\{v_{k_l}\}$, such that for all l

$$|\langle v_{k_l}, w \rangle| > \epsilon. \tag{6.4}$$

for a given ϵ. Since the sequence $\{v_k\}$ is bounded, this subsequence is also bounded and hence has a weakly convergent subsequence. By the hypothesis, it must converge weakly to zero. Clearly, this contradicts (6.4). ■

We recall the notion of a *compact operator* for ease of reference.

Definition 6.6.2 Let $\mathcal{H}$ be a Hilbert space. $T : \mathcal{H} \to \mathcal{H}$ is said to be a *compact operator* if for any weakly convergent sequence $\{u_k\} \subset \mathcal{H}$, $\{Tu_k\}$ is a convergent sequence.

In the following lemma, we state some results for the matrix inverse $(I - hA)^{-1}$.

Lemma 6.6.3 *Let $A \in \mathbb{R}^{n\times n}$. The following statements hold:*

1. $\|(I - hA)^{-1}\| \leq \dfrac{1}{1-\lambda h}$ *for all h with $\lambda h < 1$ where λ is the largest eigenvalue of* $\frac{1}{2}(A + A^\top)$.
2. *There exists an $\alpha > 0$ such that $\|(I - hA)^{-1}\| \leq \alpha$ for all sufficiently small h.*
3. *If $\{r_k h_k\}$ converges to t then $\{(I - h_k A)^{-r_k}\}$ converges to e^{At}. Moreover, the convergence is uniform in t on any bounded interval.*

Proof:

1: By the Wazewski inequality (see e.g. [33, Theorem 8.1]), $\|e^{At}\| \leq e^{\lambda t}$ for all t where λ is the largest eigenvalue of $\frac{1}{2}(A + A^\top)$. Theorem 1.5.3 in [22] gives now the desired inequality.

2: It can easily be verified by using item 1 that $\|(I - hA)^{-1}\| \leq \dfrac{1}{1-\beta}$ whenever $\lambda h \leq \beta < 1$.

3: This follows from [22, Theorem 3.5.3]. ■

6.6.2 Proof of Theorem 6.3.4 items 1 and 2

For proving Theorem 6.3.4, we start by considering items 1 and 2, which are concerned with the existence/uniqueness of the initial solution and the convergence of the impulsive parts of the approximations to the impulsive part of this initial solution. Note that the latter is needed to show that the limit of the approximations exists and satisfies Definition 6.2.4 item 1.

We shall use the following proposition which establishes the relation between the solutions of the one-step problem and the solutions of the rational complementarity problem.

Proposition 6.6.4 *Consider a matrix quadruple satisfying Assumption 6.3.2. We have the following statements for all $x_0 \in \mathbb{R}^n$.*

1. *RCP(x_0, A, B, C, D) has a unique solution.*

2. *For all sufficiently small h,*

$$\hat{u}(h^{-1}) = h\mathrm{u}_0^h, \;\; \hat{x}(h^{-1}) = h\mathrm{x}_0^h, \;\; \hat{y}(h^{-1}) = h\mathrm{y}_0^h$$

where $(\hat{u}(s), \hat{y}(s))$ is the solution of RCP(x_0, A, B, C, D), $\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)$ and $(\mathrm{u}_0^h, \mathrm{x}_0^h, \mathrm{y}_0^h)$ is the solution of the one-step problem of Algorithm 6.3.1 for $k = 0$.

Proof:

1: Observe the basic fact that if LCP(q, M) is solvable then LCP($\alpha q, M$) is also solvable provided that $\alpha \geq 0$. As a consequence, Assumption 6.3.2 implies together with the identity $h(I - hA)^{-1} = (h^{-1}I - A)^{-1}$ that for all sufficiently small h, LCP($C(h^{-1}I - A)^{-1}x_0, G(h^{-1})$) has a unique solution. From [10, Theorem 4.1 and Corollary 4.10], we can conclude that RCP(x_0, A, B, C, D) has a unique solution.

2: Let $(\hat{u}(s), \hat{y}(s))$ be the solution of RCP(x_0, A, B, C, D). It can be easily seen that $\hat{u}(h^{-1})$ solves LCP($C(h^{-1}I - A)^{-1}x_0, G(h^{-1})$) for all sufficiently small h. Note that if z is a solution of LCP(q, M) then αz is a solution of LCP($\alpha q, M$) provided $\alpha \geq 0$. Therefore, $h^{-1}\hat{u}(h^{-1})$ solves LCP($C(I - hA)^{-1}x_0, G(h^{-1})$) for all sufficiently small h. In other words, for all sufficiently small h

$$\hat{u}(h^{-1}) = h\mathrm{u}_0^h \tag{6.5a}$$
$$\hat{x}(h^{-1}) = h\mathrm{x}_0^h \tag{6.5b}$$
$$\hat{y}(h^{-1}) = h\mathrm{y}_0^h \tag{6.5c}$$

where $\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)$. ∎

Proof of Theorem 6.3.4 items 1 and 2:

1: From Proposition 6.6.4 item 1, it is known that RCP(x_0, A, B, C, D) is uniquely solvable. Let $(\hat{u}(s), \hat{y}(s))$ denote this unique solution and $\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)$. Since $\|h\mathrm{u}_0^h\|$ is bounded for sufficiently small h by the hypothesis of the theorem, $\hat{u}(s)$ is proper due to Proposition 6.6.4 item 2. It follows that $\hat{x}(s)$ is strictly proper and $\hat{y}(s)$ is proper. Clearly, Proposition 6.2.3 implies that the inverse Laplace transform of $(\hat{u}(s), \hat{x}(s), \hat{y}(s))$ is the unique initial solution of LCS(A, B, C, D) with initial state x_0.

The impulsive part of this solution is of the form $(u_0\delta, 0, y_0\delta)$ with $u_0 = \lim_{s\to\infty} \hat{u}(s)$ and $y_0 = \lim_{s\to\infty} \hat{y}(s)$ since $\hat{u}(s)$, $\hat{x}(s)$ and $\hat{y}(s)$ are proper, strictly proper and proper, respectively.

2: It is clear from (6.3) and Proposition 6.6.4 item 2 that $(\mathrm{u}^h_{\mathrm{imp}}, \mathrm{x}^h_{\mathrm{imp}}, \mathrm{y}^h_{\mathrm{imp}})$ converges to $(\mathrm{u}_{\mathrm{imp}}, 0, \mathrm{y}_{\mathrm{imp}})$ as h tends to zero. ∎

6.6.3 Topological complementarity problem

In this subsection, an infinite dimensional version of the LCP will be considered. This so-called *topological complementarity problem* has strong relations to (the regular parts of) the solutions of LCS. Moreover, it is possible to embed the discretizations obtained from the backward Euler time-stepping method in the TCP as well.

To be specific, we briefly recall the TCP for the function space $\mathcal{L}_2([0,\tau],\mathbb{R})$. More details on the TCP can be found in [3] and the references therein.

Problem 6.6.5 (TCP(q,T)) Given $q \in \mathcal{L}_2([0,\tau],\mathbb{R}^m)$ and $T : \mathcal{L}_2([0,\tau],\mathbb{R}^m) \to \mathcal{L}_2([0,\tau],\mathbb{R}^m)$, find $z \in \mathcal{L}_2([0,\tau],\mathbb{R}^m)$ such that

$$z(t) \geq 0 \tag{6.6a}$$

$$q(t) + (Tz)(t) \geq 0 \tag{6.6b}$$

for almost all $t \in [0,\tau]$ and

$$\langle z, q + Tz\rangle = 0. \tag{6.6c}$$

If z satisfies (6.6), we say that z solves TCP(q,T).

Note that the conditions given in item 3 of Definition 6.2.4 may be equivalently written as $\mathrm{u}_{\mathrm{reg}}(t) \geq 0$, $\mathrm{y}_{\mathrm{reg}}(t) \geq 0$ for almost all $t \in [0,\tau]$ and $\langle \mathrm{u}_{\mathrm{reg}}, \mathrm{y}_{\mathrm{reg}}\rangle = 0$. Hence, by associating the operator $T_{(A,B,C,D)}$ defined by

$$(T_{(A,B,C,D)}u)(t) = Du(t) + \int_0^t Ce^{A(t-s)}Bu(s)ds$$

to LCS(A,B,C,D), the solutions of LCS(A,B,C,D) can be identified with the solutions of certain TCPs in the following manner.

Proposition 6.6.6 *The following statements hold.*

1. *If* $(\mathrm{u}, \mathrm{x}, \mathrm{y}) \in \mathcal{L}_2^\delta([0,\tau], \mathbb{R}^{m+n+m})$ *is a solution of* $LCS(A,B,C,D)$ *on* $[0,\tau]$ *with initial state* x_0*, then* $\mathrm{u}_{\mathrm{reg}}$ *is a solution of TCP(*$Ce^{A\cdot}x_0^+|_{[0,\tau]}, T_{(A,B,C,D)}$*), where* $x_0^+ = x_0 + Bu_0$ *and* $\mathrm{u}_{\mathrm{imp}} = u_0\delta$.

2. *If* $u \in \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ *is a solution of TCP(*$Ce^{A\cdot}x_0|_{[0,\tau]}, T_{(A,B,C,D)}$*), then* (u, x, y) *is a solution of* $LCS(A,B,C,D)$ *on* $[0,\tau]$ *with initial state* x_0 *where*

$$x = e^{A\cdot}x_0|_{[0,\tau]} + T_{(A,B,I,0)}u$$
$$y = Cx + Du.$$

6.6.4 The time-stepping method in a TCP formulation

The approximations of (6.3) by the backward Euler time-stepping scheme can also be formulated as the solutions of certain TCPs. To do so, we introduce the operators $\tilde{C}_h : \mathbb{R}^{nN_h} \to \mathbb{R}^{mN_h}$, $\tilde{D}_h : \mathbb{R}^{mN_h} \to \mathbb{R}^{mN_h}$, $R_h : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathbb{R}^{mN_h}$, $Q_h : \mathbb{R}^{mN_h} \to \mathbb{R}^{nN_h}$, and $P_h^j : \mathbb{R}^{jN_h} \to \mathcal{L}_2([0,\tau], \mathbb{R}^j)$ for given $\tau > 0$ and h with $N_h = \lfloor \tau/h \rfloor$.

$$\tilde{C}_h := \begin{bmatrix} C & 0 & \cdots & 0 \\ 0 & C & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & C \end{bmatrix} \quad \tilde{D}_h := \begin{bmatrix} D & 0 & \cdots & 0 \\ 0 & D & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & D \end{bmatrix} \quad R_h u := \frac{1}{h} \begin{bmatrix} \int_0^h u(s)ds \\ \int_h^{2h} u(s)ds \\ \vdots \\ \int_{(N_h-1)h}^{\tau} u(s)ds \end{bmatrix}$$

$$Q_h := h \begin{bmatrix} (I-hA)^{-1}B & 0 & \cdots & 0 \\ (I-hA)^{-2}B & (I-hA)^{-1}B & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ (I-h_A)^{-N_h}B & (I-hA)^{-N_h+1}B & \cdots & (I-hA)^{-1}B \end{bmatrix}$$

$$(P_h^j \mathrm{w})(t) := \mathrm{w}_{\overline{lj}\setminus\overline{(l-1)j}} \text{ if } t \in [(l-1)h, lh) \text{ for } l = 1, 2, \ldots, N_h.$$

For ease of reference, we summarize some of the properties of these operators, which will be needed in the sequel. Without loss of generality, we can assume that $N_h h = \tau$.

Proposition 6.6.7 *Let* $\mathrm{v}, \mathrm{w} \in \mathbb{R}^{mN_h}$ *and* $\mathrm{x} \in \mathbb{R}^{nN_h}$*. The following statements hold.*

1. $R_h P_h^m \mathrm{v} = \mathrm{v}$.

2. $\mathrm{v} \geq 0$ *if and only if* $P_h^m \mathrm{v}(t) \geq 0$ *for (almost) all* $t \in [0,\tau]$.

3. $\langle P_h^m \mathrm{v}, P_h^m \mathrm{w} \rangle = h\, \mathrm{v}^\top \mathrm{w}$.

4. $DP_h^m \mathrm{v} = P_h^m \tilde{D}_h \mathrm{v}$.

5. $CP_h^n \mathrm{x} = P_h^m \tilde{C}_h \mathrm{x}$.

Proof: Evident from the definitions of P_h^j, R_h, $\tilde{C}_h$ and $\tilde{D}_h$. ■

It can be easily seen that $\tilde{\mathrm{u}}_h$ solves LCP$(\tilde{C}_h\tilde{\mathrm{q}}_h, \tilde{D}_h + \tilde{C}_h Q_h)$, where

$$\tilde{\mathrm{u}}_h = \begin{bmatrix} \mathrm{u}_1^h \\ \mathrm{u}_2^h \\ \vdots \\ \mathrm{u}_{N_h}^h \end{bmatrix}, \quad \text{and } \tilde{\mathrm{q}}_h = \begin{bmatrix} (I-hA)^{-1}\mathrm{x}_0^h \\ (I-hA)^{-2}\mathrm{x}_0^h \\ \vdots \\ (I-hA)^{-N_h}\mathrm{x}_0^h \end{bmatrix}.$$

Indeed, LCP$(\tilde{C}_h\tilde{\mathrm{q}}_h, \tilde{D}_h + \tilde{C}_h Q_h)$ is pieced together from N_h one-step problems of Algorithm 6.3.1 step 4. The following lemma will complete the TCP formulation of the time-stepping method by expressing the approximations as solutions of TCPs as well as by establishing the requirements of Theorem 6.6.9 below.

Lemma 6.6.8 *Let $T_h' = P_h^n Q_h R_h$ and $q_h' = P_h^n \tilde{\mathrm{q}}_h$. The following statements hold.*

1. *For all sufficiently small h, $\mathrm{u}_{\mathrm{reg}}^h$ solves TCP$(Cq_h', D + CT_h')$.*
2. *$\{q_h'(\cdot)\}$ converges to $e^{A\cdot}(x_0 + Bu_0)$ with u_0 as in item 2 of Theorem 6.3.4 as h tends to zero.*
3. *$\{T_h'\mathrm{u}_{\mathrm{reg}}^h - T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}}^h\}$ converges to 0 as h tends to zero.*

Proof:

1: Since $\tilde{\mathrm{u}}_h$ solves LCP$(\tilde{C}_h\tilde{\mathrm{q}}_h, \tilde{D}_h + \tilde{C}_h Q_h)$, we have

$$\tilde{\mathrm{u}}_h \geq 0 \tag{6.8a}$$
$$\tilde{\mathrm{y}}_h := \tilde{C}_h\tilde{\mathrm{q}}_h + (\tilde{D}_h + \tilde{C}_h Q_h)\tilde{\mathrm{u}}_h \geq 0 \tag{6.8b}$$
$$\tilde{\mathrm{u}}_h^\top \tilde{\mathrm{y}}_h = 0. \tag{6.8c}$$

Note that $\mathrm{u}_{\mathrm{reg}}^h = P_h^m \tilde{\mathrm{u}}_h$ and $\mathrm{y}_{\mathrm{reg}}^h = P_h^m \tilde{\mathrm{y}}_h$ due to (6.3) and the definition of P_h^m. Hence, (6.8a) and (6.8b) together with Proposition 6.6.7 item 2 imply that

$$\mathrm{u}_{\mathrm{reg}}^h(t) \geq 0 \text{ and } \mathrm{y}_{\mathrm{reg}}^h(t) \geq 0 \text{ for (almost) all } t \in [0, \tau]. \tag{6.9}$$

Moreover, we have

$$\begin{aligned} \langle \mathrm{u}_{\mathrm{reg}}^h, \mathrm{y}_{\mathrm{reg}}^h \rangle &= \langle P_h^m \tilde{\mathrm{u}}_h, P_h^m \tilde{\mathrm{y}}_h \rangle \\ &= h\, \tilde{\mathrm{u}}_h^\top \tilde{\mathrm{y}}_h \\ &= 0 \end{aligned} \tag{6.10}$$

from Proposition 6.6.7 item 3, and (6.8c). On the other hand, we have

$$\begin{aligned} \mathrm{y}^h_{\mathrm{reg}} &= P^m_h \tilde{\mathrm{y}}_h = P^m_h[\tilde{C}_h \tilde{\mathrm{q}}_h + (\tilde{D}_h + \tilde{C}_h Q_h)]\tilde{\mathrm{u}}_h \quad \text{(from (6.8b))} \\ &= CP^n_h \tilde{\mathrm{q}}_h + (D + CP^n_h Q_h)P^m_h \tilde{\mathrm{u}}_h \quad \text{(from items 4 and 5 of Proposition 6.6.7)} \\ &= Cq'_h + (D + CT'_h)\mathrm{u}^h_{\mathrm{reg}} \quad \text{(from item 1 of Proposition 6.6.7).} \end{aligned} \tag{6.11}$$

Clearly, (6.9), (6.10) and (6.11) imply that $\mathrm{u}^h_{\mathrm{reg}}$ solves $\mathrm{TCP}(Cq'_h, D + CT'_h)$.

2: Note that from Algorithm 6.3.1 step 5 we have

$$\mathrm{x}^h_0 := (I - hA)^{-1}x_0 + h(I - hA)^{-1}B\mathrm{u}^h_0. \tag{6.12}$$

Let $\hat{u}(s)$ be the solution of $\mathrm{RCP}(x_0, A, B, C, D)$ and $u_0 = \lim_{s\to\infty} \hat{u}(s)$. As shown in the proof of Theorem 6.3.4 item 2, $h\mathrm{u}^h_0$ converges to u_0 as h tends to zero. Then, (6.12) implies that

$$\{\mathrm{x}^h_0\} \text{ converges to } x_0 + Bu_0 \tag{6.13}$$

as h tends to zero. Note that $q'_h(t) = (I - hA)^{-\lfloor t/h \rfloor}\mathrm{x}^h_0$. Hence, from the triangle inequality we get

$$\begin{aligned} \|q'_h(\cdot) - e^{A\cdot}(x_0 + Bu_0)\| &\le \|(I - hA)^{-\lfloor \cdot/h \rfloor}\mathrm{x}^h_0 - e^{A\cdot}\mathrm{x}^h_0\| + \|e^{A\cdot}\mathrm{x}^h_0 - e^{A\cdot}(x_0 + Bu_0)\| \\ &\le \Big(\int_0^T \|(I - hA)^{-\lfloor t/h \rfloor} - e^{At}\|^2\, dt\Big)^{1/2}\|\mathrm{x}^h_0\| + \Big(\int_0^T \|e^{At}\|^2\, dt\Big)^{1/2}\|\mathrm{x}^h_0 - (x_0 + Bu_0)\|. \end{aligned}$$

Since $\{\lfloor t/h \rfloor h\}$ converges to t as h tends to zero, Lemma 6.6.3 item 3 and (6.13) reveal that the right hand side converges to zero.

3: Note that

$$(T'_h\mathrm{u}^h_{\mathrm{reg}})(t) = \sum_{p=1}^{l} h(I - hA)^{-(l-p+1)}B\mathrm{u}^h_p = \sum_{p=1}^{l} \int_{(p-1)h}^{ph} (I - hA)^{-(l-p+1)}B\mathrm{u}^h_p ds$$

and also that

$$(T_{(A,B,I,0)}\mathrm{u}^h_{\mathrm{reg}})(t) = \sum_{p=1}^{l-1} \int_{(p-1)h}^{ph} e^{A(t-s)}B\mathrm{u}^h_p ds + \int_{(l-1)h}^{t} e^{A(t-s)}B\mathrm{u}^h_l ds$$

for $l = \lfloor t/h \rfloor$. By exploiting the triangle inequality, we get

$$\|(T'_h \mathrm{u}^h_{\mathrm{reg}})(t) - (T_{(A,B,I,0)} \mathrm{u}^h_{\mathrm{reg}})(t)\| \leq \sum_{p=1}^{\lfloor t/h \rfloor} \int_{(p-1)h}^{ph} \|(I - hA)^{-(\lfloor t/h \rfloor - \lfloor s/h \rfloor + 1)} - e^{A(t-s)}\| \|B\mathrm{u}^h_p\| ds \quad (6.14)$$

since $(p-1)h < s \leq ph$ gives $p = \lfloor s/h \rfloor$. Clearly, $\{(\lfloor t/h \rfloor - \lfloor s/h \rfloor + 1)h\}$ converges to $t - s$ as h tends to zero. We already know from the hypothesis that $\|\mathrm{u}^h_p\|$ is bounded for $p \neq 0$. Therefore, from Lemma 6.6.3 item 3 we can conclude that the right hand side converges to zero uniformly in t on any bounded interval. It follows that $\{T'_h \mathrm{u}^h_{\mathrm{reg}} - T_{(A,B,I,0)} \mathrm{u}^h_{\mathrm{reg}}\}$ converges in $\mathcal{L}_2$-sense to zero as h tends to zero. ■

6.6.5 Convergence of solutions to TCPs

From the previous subsection, it is obvious that the convergence problem for the time-stepping method can be reduced to convergence of the solutions of a sequence of TCPs. The following theorem provides a general framework in which we shall prove the convergence of the regular parts of the approximation obtained by the backward Euler time-stepping method.

Let $\mathcal{U}$ be a normed space. A sequence of operators $S_k : \mathcal{U} \to \mathcal{U}$ is said to be *uniformly convergent* to S if $\|S_k - S\|$ converges to zero where $\|\cdot\|$ denotes the norm induced by the norm defined on $\mathcal{U}$.

Theorem 6.6.9 *Let $T : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ be a compact operator and let $S : \mathcal{L}_2([0,\tau], \mathbb{R}^m) \to \mathcal{L}_2([0,\tau], \mathbb{R}^m)$ be a linear continuous operator. Suppose that there exist sequences $\{q_k\}$, $\{S_k\}$ and $\{T_k\}$ such that $\{q_k\}$ converges to q, S_k is linear continuous nonnegative definite (i.e. $\langle v, S_k v\rangle \geq 0$ for all $v \in \mathcal{L}_2([0,\tau], \mathbb{R}^m)$) for all sufficiently large k, and TCP$(q_k, S_k + T_k)$ is solvable for all k. Let z_k be a solution of TCP$(q_k, S_k + T_k)$. If $\{z_k\}$ converges weakly to z, S_k converges uniformly to S and $\{T_k z_k - T z_k\}$ converges to zero then z solves TCP$(q, S+T)$.*

Proof: In order to prove the theorem, one should show that z, the weak limit of $\{z_k\}$, satisfies

$$z(t) \geq 0 \quad (6.15a)$$

$$q(t) + ((S+T)z)(t) \geq 0 \quad (6.15b)$$

for almost all $t \in [0, \tau]$ and

$$\langle z, q + (S + T)z \rangle = 0. \tag{6.15c}$$

Since z_k solves TCP$(q_k, S + T_k)$, we have

$$z_k(t) \geq 0 \tag{6.16a}$$

$$q_k(t) + ((S_k + T_k)z_k)(t) \geq 0 \tag{6.16b}$$

for almost all $t \in [0, \tau]$ and

$$\langle z_k, q_k + (S_k + T_k)z_k \rangle = 0 \tag{6.16c}$$

for all k. Let $\mathcal{K}$ be the nonnegative cone of $\mathcal{L}_2([0, \tau], \mathbb{R}^m)$, i.e., $\{v \mid v(t) \geq 0$ for almost all $t \in [0, \tau]\}$. Note that $\mathcal{K}$ is weakly closed (i.e., the weak limit of every weakly convergent sequence in $\mathcal{K}$ is in $\mathcal{K}$) by Theorem 3.12 of [23]. Then, (6.15a) follows from (6.16a) and the fact that $\mathcal{K}$ is weakly closed. Lemma 6.6.1 item 3b, the fact that $\{\|S_k - S\|\}$ converges to the zero and Definition 6.6.2 imply that

$$\{S_k z_k\} \text{ converges weakly to } Sz, \tag{6.17a}$$

$$\{T z_k\} \text{ converges to } Tz. \tag{6.17b}$$

As a consequence of (6.17b), we have

$$\{T_k z_k\} \text{ converges to } Tz \tag{6.17c}$$

since $\{T_k z_k - T z_k\}$ converges to zero by the hypothesis. The equations (6.17a), (6.17c) and the convergence of $\{q_k\}$ imply that $\{q_k + (S_k + T_k)z_k\}$ converges weakly to $q + (S + T)z$. Hence, (6.15b) follows from (6.16b) and the fact that $\mathcal{K}$ is weakly closed. Now, it remains to show that (6.15c) holds. Equation (6.16c) gives $\langle z_k, S_k z_k \rangle = -\langle z_k, q_k + T_k z_k \rangle$. The convergence of $\{q_k\}$ and the weak convergence of $\{z_k\}$, together with (6.17c) and Lemma 6.6.1 item 3c, imply that $\lim_{k \to \infty} \langle z_k, S_k z_k \rangle = -\lim_{k \to \infty} \langle z_k, q_k + T_k z_k \rangle = -\langle z, q + Tz \rangle$. We also have from (6.15a) and (6.15b) that $\langle z, q + (S + T)z \rangle \geq 0$. Thus,

$$\langle z, Sz \rangle \geq -\langle z, q + Tz \rangle = \lim_{k \to \infty} \langle z_k, S_k z_k \rangle. \tag{6.18}$$

The nonnegative definiteness of S_k for sufficiently large k implies that

$$\langle z_k - z, S_k(z_k - z) \rangle \geq 0. \tag{6.19}$$

Since $\lim_{k\to\infty}\langle z, S_k z_k\rangle = \lim_{k\to\infty}\langle z_k, S_k z\rangle = \langle z, Sz\rangle$ due to the facts that $\{z_k\}$ converges weakly to z, $\{S_k\}$ converges uniformly to S and Lemma 6.6.1 items 3b and 3c, we get

$$\lim_{k\to\infty}\langle z_k, S_k z_k\rangle \geq \langle z, Sz\rangle \tag{6.20}$$

by letting k tend to infinity in (6.19). Together with (6.18), this yields $\lim_{k\to\infty}\langle z_k, S_k z_k\rangle = \langle z, Sz\rangle$. Combining this equation, (6.17c), the convergence of $\{q_k\}$ to q and Lemma 6.6.1 item 3c results in $\lim_{k\to\infty}\langle z_k, q_k + (S_k + T_k)z_k\rangle = \langle z, q + (S+T)z\rangle$. Finally, (6.15c) follows from the last equation and (6.16c). ■

6.6.6 Completing the proof of Theorem 6.3.4

The proofs of item 1 and 2 in Theorem 6.3.4 have already been shown. The remaining items will be proven in this subsection.

Proof of item 3 of Theorem 6.3.4:

3a: The convergence of the impulsive parts has already been shown in the proof of item 2. Hence, it remains to show that the claim on the regular parts holds. By the hypothesis of the theorem, we know that $\|\mathrm{u}^h_{\mathrm{reg}}\|$ is bounded for sufficiently small h. According to Lemma 6.6.1 item 1, the existence of a weakly convergent subsequence of $\{\mathrm{u}^{h_k}_{\mathrm{reg}}\}$, say $\{\mathrm{u}^{h_{k_l}}_{\mathrm{reg}}\}$, is clear. Let $\mathrm{u}_{\mathrm{reg}}$ denote the weak limit of this subsequence, and also let q'_{h_k} and T'_{h_k} be defined as in Lemma 6.6.8. Since $T_{(A,B,I,0)}$ is a compact operator (see e.g. [23, Exercise 4.15]), it follows from Definition 6.6.2 that $\{T_{(A,B,I,0)}\mathrm{u}^{h_{k_l}}_{\mathrm{reg}}\}$ converges (strongly) to $T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}}$. Then, Lemma 6.6.8 item 3 implies that

$$\{T'_{h_{k_l}}\mathrm{u}^{h_{k_l}}_{\mathrm{reg}}\} \text{ converges to } T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}}. \tag{6.21}$$

Note that

$$\mathrm{x}^{h_{k_l}}_{\mathrm{reg}} = q'_{h_{k_l}} + T'_{h_{k_l}}\mathrm{u}^{h_{k_l}}_{\mathrm{reg}} \tag{6.22a}$$
$$\mathrm{y}^{h_{k_l}}_{\mathrm{reg}} = Cq'_{h_{k_l}} + (D + CT'_{h_{k_l}})\mathrm{u}^{h_{k_l}}_{\mathrm{reg}}. \tag{6.22b}$$

It is clear from Lemma 6.6.8 item 2, (6.21) and (6.22a) that $\{\mathrm{x}^{h_{k_l}}_{\mathrm{reg}}\}$ converges to $\mathrm{x}_{\mathrm{reg}} := e^{A\cdot}(x_0+Bu_0)|_{[0,\tau]}+T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}}$. Since $\{D\mathrm{u}^{h_{k_l}}_{\mathrm{reg}}\}$ converges weakly to $D\mathrm{u}_{\mathrm{reg}}$ due to Lemma 6.6.1 item 3b, it follows from Lemma 6.6.8 item 2, (6.21) and (6.22b) that $\{\mathrm{y}^{h_{k_l}}_{\mathrm{reg}}\}$ converges weakly to $\mathrm{y}_{\mathrm{reg}} := Ce^{A\cdot}(x_0 + Bu_0)|_{[0,\tau]} + T_{(A,B,C,D)}\mathrm{u}_{\mathrm{reg}}$.

3b: Item 2 of Theorem 6.3.4 states the convergence of $(\mathbf{u}_{\text{imp}}^{h_k}, \mathbf{x}_{\text{imp}}^{h_k}, \mathbf{y}_{\text{imp}}^{h_k})$ to

$$(\mathbf{u}_{\text{imp}}, 0, \mathbf{y}_{\text{imp}}) = (u_0\delta, 0, y_0\delta) = (\bar{\mathbf{u}}_{\text{imp}}, \bar{\mathbf{x}}_{\text{imp}}, \bar{\mathbf{y}}_{\text{imp}}), \tag{6.23}$$

where $(\bar{\mathbf{u}}, \bar{\mathbf{x}}, \bar{\mathbf{y}}) \in \mathcal{B}_\delta^{m+n+m}$ is the unique initial solution for initial state x_0. Hence, we also have $\mathbf{y}_{\text{imp}} = D\mathbf{u}_{\text{imp}}$ due to $\mathbf{x}_{\text{imp}} = 0$. Let us define in the framework of Theorem 6.6.9

- $T = T_{(A,B,C,0)}$,
- $S = D$,
- $q_l = Cq'_{h_{k_l}}$,
- $S_l = D$, and
- $T_l = CT'_{h_{k_l}}$.

It can be checked that

- T is compact ([23, Exercise 4.15]),
- S is linear and continuous,
- $\{q_l\}$ converges to $Ce^{A\cdot}(x_0 + Bu_0)|_{[0,\tau]}$ (from Lemma 6.6.8 item 2),
- S_l is linear continuous and nonnegative definite for all l (by Lemma 6.6.11 item 1),
- $\mathrm{TCP}(q_l, S + T_l)$ is solvable for all sufficiently large l (from Lemma 6.6.8 item 1),
- S_l converges uniformly to S, and
- $\{T_l\mathbf{u}_{\text{reg}}^{h_{k_l}} - T\mathbf{u}_{\text{reg}}^{h_{k_l}}\}$ converges to zero (from Lemma 6.6.8 item 3).

Then, Theorem 6.6.9 implies that $\mathbf{u}_{\text{reg}}$ solves $\mathrm{TCP}(Ce^{A\cdot}(x_0 + Bu_0)|_{[0,\tau]}, T_{(A,B,C,D)})$. Due to Proposition 6.6.6 item 2, $(\mathbf{u}_{\text{reg}}, \mathbf{x}_{\text{reg}}, \mathbf{y}_{\text{reg}})$ is a solution of $\mathrm{LCS}(A, B, C, D)$ on $[0, \tau]$ with the initial state $x_0 + Bu_0$ (with u_0 as in (6.23)), where

$$\begin{aligned} \mathbf{x}_{\text{reg}} &= e^{A\cdot}(x_0 + Bu_0)|_{[0,\tau]} + T_{(A,B,I,0)}\mathbf{u}_{\text{reg}} \\ \mathbf{y}_{\text{reg}} &= C\mathbf{x}_{\text{reg}} + D\mathbf{u}_{\text{reg}}. \end{aligned}$$

Equivalently,

$$\dot{\mathbf{x}}_{\text{reg}} = A\mathbf{x}_{\text{reg}} + B\mathbf{u}_{\text{reg}} + (x_0 + Bu_0)\delta \tag{6.24a}$$

$$\mathbf{y}_{\text{reg}} = C\mathbf{x}_{\text{reg}} + D\mathbf{u}_{\text{reg}} \tag{6.24b}$$

holds in the distributional sense and

$$0 \leq \mathrm{u}_{\mathrm{reg}}(t) \perp \mathrm{y}_{\mathrm{reg}}(t) \geq 0 \tag{6.24c}$$

for almost all $t \in [0, \tau]$. Since $\mathrm{u}_{\mathrm{imp}} = u_0\delta$, $\mathrm{y}_{\mathrm{imp}} = Du_{\mathrm{imp}}$ and $\mathrm{x}_{\mathrm{imp}} = 0$, (6.24a) and (6.24b) yield

$$\dot{\mathrm{x}}_{\mathrm{reg}} = A\mathrm{x} + B\mathrm{u} + x_0\delta \tag{6.25a}$$

$$\mathrm{y} = C\mathrm{x} + D\mathrm{u} \tag{6.25b}$$

Clearly, (6.23), (6.24c) and (6.25) imply that $(\mathrm{u}, \mathrm{x}, \mathrm{y})$ is a solution of LCS(A, B, C, D) on $[0, \tau]$ with the initial state x_0.

3c: We have already proven that the complete sequence of impulsive parts $(\mathrm{u}_{\mathrm{imp}}^{h_k}, \mathrm{x}_{\mathrm{imp}}^{h_k}, \mathrm{y}_{\mathrm{imp}}^{h_k})$ converges. Note that the sequence of regular parts $(\mathrm{u}_{\mathrm{reg}}^{h_k}, \mathrm{x}_{\mathrm{reg}}^{h_k}, \mathrm{y}_{\mathrm{reg}}^{h_k})$ is bounded by assumption. Moreover, following the proof of item 3a above, it is clear that every converging subsequence $(\mathrm{u}_{\mathrm{reg}}^{h_{k_l}}, \mathrm{x}_{\mathrm{reg}}^{h_{k_l}}, \mathrm{y}_{\mathrm{reg}}^{h_{k_l}})$ converges to a solution of the LCS(A, B, C, D) with initial state $x_0 + Bu_0$. Since this solution is unique, every converging subsequence of the bounded sequence of regular parts has the same limit. Applying Theorem 6.6.1 item 2 completes the proof. ∎

6.6.7 Some results on LCPs

We will present in this subsection some results on LCPs, that will be needed to prove the main result (Theorem 6.4.1) for linear passive complementarity systems.

Proposition 6.6.10 *Let $M \in \mathbb{R}^{n \times n}$ be a positive definite matrix and z_i the unique solution of LCP(q_i, M) for $i = 1, 2$. Then,*

$$\|z_1 - z_2\| \leq \frac{n^{3/2}}{\mu(M)} \|q_1 - q_2\|$$

where $\mu(M)$ denotes the smallest eigenvalue of the symmetric part of M, i.e., $\frac{1}{2}(M + M^\top)$.

Proof: By Lemma 7.3.10 and Proposition 5.10.10 in [6], we have

$$\|z_1 - z_2\|_\infty \leq \frac{n}{\mu(M)} \|q_1 - q_2\|_\infty. \tag{6.26}$$

Since $\|z\| \leq n^{1/2}\|z\|_\infty$ and $\|z\|_\infty \leq \|z\|$ for all $z \in \mathbb{R}^n$, (6.26) yields $\|z_1 - z_2\| \leq \dfrac{n^{3/2}}{\mu(M)}\|q_1 - q_2\|$. ■

Using the passivity property, we can compute a lower bound on $\mu(G(h^{-1}))$ with $G(s) := C(sI - A)^{-1}B + D$, that will be useful for the application of Proposition 6.6.10.

Lemma 6.6.11 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Let $\mu(N)$ denote the smallest eigenvalue of the symmetric part of a matrix N and define $G(s) = D + C(sI - A)^{-1}B$. The following statements hold.*

1. $D \geq 0$.
2. *$u \neq 0$ and $u^\top Du = 0$ imply that $u^\top CBu > 0$.*
3. *There exists $\alpha > 0$ such that $\mu(D + hCB) \geq \alpha h$ for all sufficiently small h.*
4. *There exists $\beta > 0$ such that $\mu(G(h^{-1})) \geq \beta h$ for all sufficiently small h.*

Proof:

1: It follows from Lemma 3.8.5 item 1.

2: It follows from Lemma 3.8.5 item 3.

3: It follows from the previous item and Lemma 3.8.3.

4: It is known from matrix theory (see e.g. [18, Property 9.13.4.9]) that

$$\mu(N_1 + N_2) \geq \mu(N_1) + \mu(N_2)$$

for all square matrices N_1 and N_2. Hence, we get

$$\begin{aligned} \mu(G(h^{-1})) &\geq \mu(D + hCB) + h^2\mu(CA(I - hA)^{-1}B) \\ &\geq \beta h \qquad \text{(from item 3)} \end{aligned}$$

for some $\beta > 0$ and all sufficiently small h. ■

The following auxiliary lemma will be needed in the sequel.

Lemma 6.6.12 *Let $\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ be a given nonempty polyhedron with $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^m$ and let x^* be equal to* $\arg\min_{x\in\mathcal{P}} \|x\|$. *There exists an index set $J \subseteq \overline{n}$ such that* $x^* = \arg\min_{A_{J\bullet}x=b_J} \|x\|$.

Proof: Consider the convex quadratic optimization problem

$$\min_{Ax\geq b} \tfrac{1}{2}x^\top x.$$

The well-known Kuhn-Tucker conditions are necessary and sufficient for this problem because of its convexity (see for instance [6, Chapter 1.2]), i.e., x^* is the solution of the optimization problem above if and only if there exists a $u \in \mathbb{R}^m$ such that

$$\begin{aligned} x^* &= A^\top u \\ Ax^* &\geq b \\ u &\geq 0 \\ u^\top(Ax^* - b) &= 0. \end{aligned}$$

Take such a vector u. Let $J = \{i \mid u_i > 0\}$ and $v = u_J$. Then, x^* satisfies

$$x^* = (A_{J\bullet})^\top v \tag{6.27a}$$

$$A_{J\bullet}x^* = b_J. \tag{6.27b}$$

Note that (6.27) are necessary and sufficient (Kuhn-Tucker) conditions for the convex quadratic minimization problem $\min_{A_{J\bullet}x=b_J} \frac{1}{2}x^\top x$. ∎

The next lemma establishes bounds on the solutions of linear complementarity problems with nonnegative definite matrices.

Lemma 6.6.13 *Let $M \in \mathbb{R}^{n\times n}$ be nonnegative definite and $\mathcal{Q} = SOL(0, M)$. Also let $\mathcal{Q}^*$ denote the dual cone of the set $\mathcal{Q}$ as defined in Chapter 1. We have the following statements.*

1. *$LCP(q, M)$ is solvable if and only if $q \in \mathcal{Q}^*$.*

2. *For each $q \in \mathcal{Q}^*$, there exists a unique least-norm solution $z^* \in SOL(q, M)$ such that $\|z^*\| \leq \|z\|$ for all $z \in SOL(q, M)$.*

3. *There exists $\alpha > 0$ such that for all $q \in \mathcal{Q}^*$*

$$\|z^*(q)\| \leq \alpha\|q\|,$$

where $z^*(q)$ denotes the least-norm solution (see item 2) of LCP(q, M).

Proof:

1: It follows from [6, Exercise 3.12.1] and Lemma 2.

2: This follows from the fact that SOL(q, M) is a nonempty polyhedron whenever $q \in \mathcal{Q}^*$ [6, Theorem 3.1.7(c)].

3: Define

$$\alpha(A) = \begin{cases} 0 & \text{if } A = 0 \\ \max\limits_{\substack{y \in \operatorname{im} A \\ \|y\|=1}} \min\limits_{Ax=y} \|x\| & \text{if } A \neq 0 \end{cases}$$

Note that

$$\max_{\substack{y \in \operatorname{im} A \\ \|y\|=1}} \min_{Ax=y} \|x\| = \max_{\|Ax\|=1} \min_{Ax'=0} \|x - x'\|.$$

The mapping $x \mapsto \min\limits_{Ax'=0} \|x - x'\|$ achieves its maximum on the set $\{x \mid \|Ax\| = 1\}$. Hence, the quantity $\alpha(A)$ is well-defined for all A. Take

$$\alpha = \sqrt{2} \max_{J \subseteq \overline{n}} \max_{K \subseteq \overline{3n}} \alpha\Big(\begin{bmatrix} I \\ -I_{J^c\bullet} \\ M \\ -M_{J\bullet} \end{bmatrix}_{K\bullet} \Big)$$

where $J^c = \overline{n} \setminus J$. For any $q \in \mathcal{Q}^*$, we know from the items 1 and 2 that LCP(q, M) is solvable and that there exists a unique least-norm solution $z^*(q)$. Let $J = \{i \mid z_i^*(q) > 0\}$. Clearly, $\mathcal{P} = \{v \mid v_J \geq 0,\ v_{J^c} = 0,\ q_J + M_{JJ} v_J = 0, \text{ and } q_{J^c} + M_{J^c J} v_J \geq 0\} \subseteq SOL(q, M)$ and $z^*(q) \in \mathcal{P}$. Note that $\mathcal{P}$ is a polyhedron, since $\mathcal{P} = \{v \mid Av \geq b\}$ where

$$A = \begin{bmatrix} I \\ -I_{J^c\bullet} \\ M \\ -M_{J\bullet} \end{bmatrix} \text{ and } b = \begin{bmatrix} 0 \\ 0 \\ -q \\ q_J \end{bmatrix}.$$

Moreover, it is obvious that $z^*(q) = \arg\min_{Av \geq b} \|v\|$. Then, according to Lemma 6.6.12 there exists $K \subseteq \overline{3n}$ such that $z^*(q) = \arg\min_{A_{K\bullet} v = b_K} \|v\|$. Thus, we have $\|z^*(q)\| \leq \alpha(A_{K\bullet}) \|b_K\|$. Note that $\|b_K\|^2 \leq \|b\|^2 \leq \|q\|^2 + \|q_J\|^2 \leq 2\|q\|^2$ and $\sqrt{2}\alpha(A_{K\bullet}) \leq \alpha$. Consequently, $\|z^*(q)\| \leq \alpha \|q\|$. ∎

6.6.8 Proof of Theorem 6.4.1

After these results on LCPs, the proof of the main result on linear passive complementarity systems is in order. The proof will be based on showing that the requirements of Theorem 6.3.4 are fulfilled for this class of linear complementarity systems.

Lemma 6.6.14 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Then, for all sufficiently small h, LCP$(hC(I-hA)^{-1}\bar{x}$, $G(h^{-1}))$ has a unique solution for each $\bar{x} \in \mathbb{R}^n$.*

Proof: The statement follows from the positive definiteness of $G(h^{-1})$ for all sufficiently small h (Lemma 6.6.11 item 4 together with Theorem 3.1.6 of [6]). ■

Lemma 6.6.15 *Consider a matrix quadruple (A, B, C, D) such that Assumption 6.2.7 holds and $\Sigma(A, B, C, D)$ is passive. Let $\tau > 0$ and $\mathcal{Q} = SOL(0, D)$. Also let $(\{\mathrm{u}_k^h\}, \{\mathrm{x}_k^h\}, \{\mathrm{y}_k^h\})$ be produced by Algorithm 6.3.1. The following statements hold for all sufficiently small h.*

1. *$C\mathrm{x}_k^h \in \mathcal{Q}^*$ for all $k \neq -1$.*

2. *There exists an $\alpha > 0$ independent of x_0 such that $\|\mathrm{u}_k^h\| \leq \alpha\|x_0\|$ for all $k \neq 0$.*

Proof:

1: It is evident from (6.2b) and (6.2c) that u_k^h solves LCP$(C\mathrm{x}_k^h, D)$ when $k \neq -1$. Since D is nonnegative definite (Lemma 6.6.11 item 1), $C\mathrm{x}_k^h \in \mathcal{Q}^*$ due to Lemma 6.6.13 item 1.

2: All inequalities involving h are meant to hold for all sufficiently small h, and $\alpha_1, \alpha_2, \dots, \alpha_6$ are suitably chosen positive constants in this proof. Note that LCP$(C\mathrm{x}_k^h, D)$ is solvable for all $k \neq -1$ due to item 1 and Lemma 6.6.13 item 1. Let u^* be the least-norm solution of LCP$(C\mathrm{x}_k^h, D)$. Clearly, u^* solves also LCP$(C\mathrm{x}_k^h - hC(I-hA)^{-1}Bu^*, G(h^{-1}))$. According to Proposition 6.6.10, we have

$$\|\mathrm{u}_{k+1}^h - u^*\| \leq \frac{m^{3/2}}{\mu(G(h^{-1}))}\|C(I-hA)^{-1}\mathrm{x}_k^h - C\mathrm{x}_k^h + hC(I-hA)^{-1}Bu^*\|$$

since u_{k+1}^h solves LCP$(C(I-hA)^{-1}\mathrm{x}_k^h, G(h^{-1}))$ and $G(h^{-1})$ is positive definite for all sufficiently small h. By using the triangle inequality and Lemma 6.6.11 item 4, we obtain

$$\|\mathrm{u}_{k+1}^h - u^*\| \leq \frac{\alpha_1}{h}\|C[(I-hA)^{-1} - I]\mathrm{x}_k^h\| + \alpha_1\|C(I-hA)^{-1}Bu^*\|.$$

Note that $(I-hA)^{-1} - I = hA(I-hA)^{-1}$. It can be easily verified that Lemma 6.6.3 item 2 and Lemma 6.6.13 item 3 result in

$$\|\mathrm{u}_{k+1}^h - u^*\| \leq \alpha_2\|\mathrm{x}_k^h\|. \tag{6.28}$$

Consequently, we get

$$\|\mathrm{u}_{k+1}^h\| \leq \|u^*\| + \|\mathrm{u}_{k+1}^h - u^*\| \leq \alpha_3 \|\mathrm{x}_k^h\| \tag{6.29}$$

by applying the triangle inequality and employing Lemma 6.6.13 item 3 and (6.28). It follows that

$$\begin{aligned}
\|\mathrm{x}_{k+1}^h\| &\leq \|\mathrm{x}_k^h\| + \|\mathrm{x}_{k+1}^h - \mathrm{x}_k^h\| \\
&\leq \|\mathrm{x}_k^h\| + \|[(I-hA)^{-1} - I]\mathrm{x}_k^h + h(I-hA)^{-1}B\mathrm{u}_{k+1}^h\| \quad \text{(from (6.2a))} \\
&\leq (1+\alpha_4 h)\|\mathrm{x}_k^h\|. \quad \text{(from Lemma 6.6.3 item 2)}
\end{aligned} \tag{6.30}$$

Since $\lim_{h\to 0}(1+\alpha_4 h)^{N_h} = e^{\alpha_4 \tau}$ (Lemma 6.6.3 item 3), (6.30) implies that $\|\mathrm{x}_k^h\| \leq \alpha_5 \|\mathrm{x}_0^h\|$ for some $\alpha_5 > 0$. Here $N_h = \lfloor \frac{\tau}{h} \rfloor$. Note that we have

$$\|\mathrm{x}_0^h\| = \|\mathrm{x}_{-1}^h + hB\mathrm{u}_0^h\| = \|x_0 + hB\mathrm{u}_0^h\| \leq \alpha_6 \|x_0\|$$

from Proposition 6.3. Finally, (6.29) and (6.6.8) establish the desired inequality. ∎

After all these preliminaries, we can prove Theorem 6.4.1.

Proof of Theorem 6.4.1 According to Lemma 6.6.14, Assumption 6.3.2 holds. Then, Proposition 6.6.4 item 1 implies that $\mathrm{RCP}(x_0, A, B, C, D)$ has a unique solution, say $(\hat{u}(s), \hat{y}(s))$. It is known from Lemma 3.8.13 item 2 that $\hat{u}(s)$ is proper. Therefore, boundedness of $\|h\mathrm{u}_0^h\|$ for all sufficiently small h follows from Proposition 6.6.4 item 2. On the other hand, D is nonnegative definite due to item 1 of Lemma 6.6.11 and

$$\|\mathrm{u}_{\mathrm{reg}}^h\| = \left(\int_0^\tau \|\mathrm{u}_{\mathrm{reg}}^h(t)\|\, dt\right)^{1/2} \leq \alpha \tau^{1/2} \|x_0\| \tag{6.31}$$

due to (6.3) and Lemma 6.6.15 item 2. Finally, it is known from Theorem 6.2.8 that $(\mathbf{u}, \mathbf{x}, \mathbf{y})$ is the unique solution on $[0, \tau]$ with the initial state x_0. As a consequence of Theorem 6.3.4 item 3c, $\{(\mathrm{u}^{h_k}, \mathrm{y}^{h_k})\}$ converges weakly to (u, y) and $\{\mathrm{x}^{h_k}\}$ converges to x for any sequence $\{h_k\}$ that converges to zero. In other words, $\{(\mathrm{u}^h, \mathrm{y}^h)\}$ converges weakly to $(\mathbf{u}, \mathbf{y})$ and $\{\mathrm{x}^h\}$ converges to $\mathbf{x}$ as h tends to zero. ∎

6.6.9 Proof of Theorem 6.4.2

In this subsection, the continuous dependence of solution trajectories on the initial states will be proven as formulated in Theorem 6.4.2.

Proof of Theorem 6.4.2 Let the sequence $\{\bar{x}_k\} \subset \mathbb{R}^n$ converge to $\bar{x} \in \mathbb{R}^n$. Denote the solution of LCS(A, B, C, D) on $[0, \tau]$ with the initial states $\bar{x}_k$ and $\bar{x}$ by $(\mathrm{u}^k, \mathrm{x}^k, \mathrm{y}^k)$ and $(\mathrm{u}, \mathrm{x}, \mathrm{y})$, respectively. Then, it should be shown that

1. $\{(\mathrm{u}^k_{\mathrm{imp}}, \mathrm{x}^k_{\mathrm{imp}}, \mathrm{y}^k_{\mathrm{imp}})\}$ converges to $(\mathrm{u}_{\mathrm{imp}}, \mathrm{x}_{\mathrm{imp}}, \mathrm{y}_{\mathrm{imp}})$,

2. $\{(\mathrm{u}^k_{\mathrm{reg}}, \mathrm{y}^k_{\mathrm{reg}})\}$ converges weakly to $(\mathrm{u}_{\mathrm{reg}}, \mathrm{y}_{\mathrm{reg}})$ and $\{\mathrm{x}^k_{\mathrm{reg}}\}$ converges to $\mathrm{x}_{\mathrm{reg}}$.

1: Let $(\mathrm{u}^k_{\mathrm{imp}}, \mathrm{x}^k_{\mathrm{imp}}, \mathrm{y}^k_{\mathrm{imp}}) = (u^k_0\delta, x^k_0\delta, y^k_0\delta)$. Also let $u^k_0(h)$ and $u_0(h)$ be the solutions of the one-step problems LCP$(C(I - hA)^{-1}\bar{x}_k, hC(I - hA)^{-1}B + D)$ and LCP$(C(I - hA)^{-1}\bar{x}, hC(I - hA)^{-1}B + D)$, respectively. From Proposition 6.6.10 and Lemma 6.6.11 item 4, we get

$$\|u^k_0(h) - u_0(h)\| \leq \frac{\alpha}{h}\|C(I - hA)^{-1}\| \, \|\bar{x}_k - \bar{x}\|$$

for sufficiently small h. By multiplying the inequality above by h and using Lemma 6.6.3 item 2, we obtain

$$\|hu^k_0(h) - hu_0(h)\| \leq \alpha'\|\bar{x}_k - \bar{x}\| \tag{6.32}$$

for sufficiently small h. On the other hand, it is already known from the proof of Theorem 6.3.4 item 2 that $\lim_{h\to 0} hu^k_0(h) = u^k_0$ and $\lim_{h\to 0} hu_0(h) = u_0$. Thus, (6.32) yields

$$\|u^k_0 - u_0\| \leq \alpha'\|\bar{x}_k - \bar{x}\|. \tag{6.33}$$

Clearly, $\{u^k_0\}$ converges to u_0. Consequently, $\{\mathrm{u}^k_{\mathrm{imp}}\}$ converges to $\mathrm{u}_{\mathrm{imp}}$. Since $\mathrm{x}^k_{\mathrm{imp}} = 0$ and $\mathrm{y}^k_{\mathrm{imp}} = D\mathrm{u}^k_{\mathrm{imp}}$, we can conclude that $\{(\mathrm{u}^k_{\mathrm{imp}}, \mathrm{x}^k_{\mathrm{imp}}, \mathrm{y}^k_{\mathrm{imp}})\}$ converges to $(\mathrm{u}_{\mathrm{imp}}, \mathrm{x}_{\mathrm{imp}}, \mathrm{y}_{\mathrm{imp}})$.

2: Observe that $(\mathrm{u}^k_{\mathrm{reg}}, \mathrm{x}^k_{\mathrm{reg}}, \mathrm{y}^k_{\mathrm{reg}})$ and $(\mathrm{u}_{\mathrm{reg}}, \mathrm{x}_{\mathrm{reg}}, \mathrm{y}_{\mathrm{reg}})$ are the solutions of LCS(A, B, C, D) on $[0, \tau]$ with the initial states $\bar{x}_k + Bu^k_0$ and $\bar{x} + Bu_0$, respectively. Moreover, $\{\bar{x}_k + Bu^k_0\}$ converges to $\bar{x} + Bu_0$ as shown in the proof of item 1 above. Lemma 6.6.15 item 2 together with (6.31) implies that for some $\beta > 0$ independent of $\bar{x}_k + Bu^k_0$, $\|\mathrm{u}^k_{\mathrm{reg}}\| \leq \beta\|\bar{x}_k + Bu^k_0\|$ for all k. This means that the sequence $\{\mathrm{u}^k_{\mathrm{reg}}\}$ is bounded since the sequence $\{\bar{x}_k + Bu^k_0\}$ is convergent. Hence, there exists at least one weakly convergent subsequence of $\{\mathrm{u}^k_{\mathrm{reg}}\}$ according to Lemma 6.6.1 item 3a. Take any such subsequence of $\{\mathrm{u}^k_{\mathrm{reg}}\}$, say $\{\mathrm{u}^{k_l}_{\mathrm{reg}}\}$. Define

- $T = T_{(A,B,C,0)}$,
- $S = D$,
- $q_l = Ce^{A\cdot}(\bar{x}_{k_l} + Bu^{k_l}_0)$, and

- $T_l = T$.

It can be checked that

- T is compact ([23, Exercise 4.15]),
- S is linear continuous,
- $\{q_l\}$ converges to $Ce^{A\cdot}(\bar{x} + Bu_0)|_{[0,\tau]}$ (since $\|q_l - Ce^{A\cdot}(\bar{x} + Bu_0)\| \leq \|Ce^{A\cdot}\| \, \|(\bar{x}_{k_l} + Bu_0^{k_l}) - (\bar{x} + Bu_0)\|$)
- TCP$(q_l, S + T_l)$ is solvable (from Proposition 6.6.6 item 1),
- S_l is linear continuous nonnegative definite (by Lemma 6.6.11 item 1), and
- $\{T_l \mathrm{u}_{\mathrm{reg}}^{k_l} - T\mathrm{u}_{\mathrm{reg}}^{k_l}\} = 0$.

Therefore, $\{\mathrm{u}_{\mathrm{reg}}^{k_l}\}$ converges weakly to the solution $\mathrm{u}_{\mathrm{reg}}$ of TCP$(Ce^{A\cdot}(\bar{x}+Bu_0)|_{[0,\tau]}, T_{(A,B,C,D)})$ according to Theorem 6.6.9. Since $\mathrm{u}_{\mathrm{reg}}$ is unique due to Proposition 6.6.6 item 2 and Theorem 6.2.8, the reasoning above shows that any weakly convergent subsequence of $\{u_{\mathrm{reg}}^k\}$ has the same limit. Lemma 6.6.1 item 2 implies now that the whole sequence $\{\mathrm{u}_{\mathrm{reg}}^k\}$ converges weakly to $\mathrm{u}_{\mathrm{reg}}$. Note that Proposition 6.6.6 item 2 and uniqueness of the solutions of LCS(A, B, C, D) yield that

$$\mathrm{x}_{\mathrm{reg}}^k = e^{A\cdot}(\bar{x}_k + Bu_0^k)|_{[0,\tau]} + T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}}^k \tag{6.34a}$$

$$\mathrm{y}_{\mathrm{reg}}^k = C\mathrm{x}_{\mathrm{reg}}^k + D\mathrm{u}_{\mathrm{reg}}^k \tag{6.34b}$$

and

$$\mathrm{x}_{\mathrm{reg}} = e^{A\cdot}(\bar{x} + Bu_0)|_{[0,\tau]} + T_{(A,B,I,0)}\mathrm{u}_{\mathrm{reg}} \tag{6.34c}$$

$$\mathrm{y}_{\mathrm{reg}} = C\mathrm{x}_{\mathrm{reg}} + D\mathrm{u}_{\mathrm{reg}} \tag{6.34d}$$

Then, convergence of $\{\mathrm{x}_{\mathrm{reg}}^k\}$ to $\mathrm{x}_{\mathrm{reg}}$ and weak convergence of $\{\mathrm{y}_{\mathrm{reg}}^k\}$ to $\mathrm{y}_{\mathrm{reg}}$ follow from (6.34), the convergence of $\{\bar{x}_k + Bu_0^k\}$ to $\bar{x} + Bu_0$ and the compactness of $T_{(A,B,I,0)}$. ∎

References

[1] D. Bedrosian and J. Vlach. Time-domain analysis of networks with internally controlled switches. *IEEE Trans. Circuits and Systems-I*, 39(3):199–212, 1992.

[2] W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis.* Kluwer, Deventer, the Netherlands, 1981.

[3] J. M. Borwein and M. A. H. Dempster. The linear order complementarity problem. *Mathematics of Operations Research*, 14:534–558, 1989.

[4] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. The nature of solutions to linear passive complementarity systems. In *Proc. of the 38th IEEE Conference on Decision and Control*, pages 3043–3048, Phoenix (USA), 1999.

[5] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Well-posedness of a class of linear network with ideal diodes. In *Proc. of the 14th International Symposium of Mathematical Theory of Networks and Systems*, Perpignan (France), 2000.

[6] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem.* Academic Press, Inc., Boston, 1992.

[7] J.T.J. van Eijndhoven. Solving the linear complementarity problem in circuit simulation. *SIAM Journal on Control and Optimization*, 24(5):1050–1062, 1986.

[8] C.W. Gear. *Numerical Initial Value Problems in Ordinary Differential Equations.* Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

[9] W.P.M.H. Heemels, M.K. Çamlıbel, and J.M. Schumacher. Dynamical analysis of linear passive networks with diodes. Part I: Well-posedness. Technical Report 00 I/02, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted to for publication.

[10] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. The rational complementarity problem. *Linear Algebra and its Applications*, 294:93–135, 1999.

[11] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity systems. *SIAM Journal on Applied Mathematics*, 60(4):1234–1269, 2000.

[12] F. Hirsch and G. Lacombe. *Elements of Functional Analysis.* Springer-Verlag, New York, 1999.

[13] J. Katzenelson. An algorithm for solving nonlinear resistor networks. *Bell Syst. Tech. J.*, 44:1605–1620, 1965.

[14] D.M.W. Leenaerts. On linear dynamic complementarity systems. *IEEE Transactions on Circuits and Systems-I*, 46(8):1022–1026, 1999.

[15] D.M.W. Leenaerts and W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[16] Y.J. Lootsma, A.J. van der Schaft, and M.K. Çamlıbel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[17] P. Lötstedt. Numerical simulation of time-dependent contact and friction problems in rigid body mechanics. *SIAM Journal on Scientific and Statistical Computing*, 5:370–393, 1984.

[18] H. Lütkepohl. *Handbook of Matrices*. Wiley, New York, 1996.

[19] A. Massarini, U. Reggiani, and K. Kazimierczuk. Analysis of networks with ideal switches by state equations. *IEEE Trans. Circuits and Systems-I*, 44(8):692–697, 1997.

[20] J.J. Moreau. Numerical aspects of the sweeping process. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):329–349, 1999.

[21] L. Paoli and M. Schatzman. Schéma numérique pour un modèle de vibrations avec contraintes unilatérales et perte d'énergie aux impacts, en dimension finie. *C.R. Acad. Sci. Paris Sér. I Math.*, 317:211–215, 1993.

[22] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.

[23] W. Rudin. *Functional Analysis*. McGraw-Hill, New York, 1977.

[24] I.W. Sandberg. Theorems on the computation of the transient response of nonlinear networks containing transistors and diodes. *Bell System Technical Journal*, 49:1739–1776, 1970.

[25] A.J. van der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9:266–301, 1996.

[26] A.J. van der Schaft and J.M. Schumacher. Complementarity modelling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[27] A.J. van der Schaft and J.M. Schumacher. *An Introduction to Hybrid Dynamical Systems*. Springer-Verlag, London, 2000.

[28] D.E. Stewart. Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problem. *Archive for Rational Mechanics and Analysis*, 145(3):215–260, 1998.

[29] D.E. Stewart. Time-stepping methods and the mathematics of rigid body dynamics. Chapter 1 of *Impact and Friction,* A. Guran, J.A.C. Martins and A. Klarbring (eds.), Birkhäuser, 1999.

[30] D.E. Stewart and J.C. Trinkle. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and Coulomb friction. *Int. Journal for Numerical Methods in Engineering*, 39:2673–2691, 1996.

[31] J. Vlach, J.M. Wojciechowski, and A. Opal. Analysis of nonlinear networks with inconsistent initial conditions. *IEEE Transactions on Circuits and Systems-I*, 42(4):195–200, 1995.

[32] J. C. Willems. Dissipative dynamical systems. *Arch. Rational Mech. Anal.*, 45:321–393, 1972.

[33] J. L. Willems. *Stability Theory of Dynamical Systems.* Thomas Nelson and Sons Ltd., 1970.

Chapter 7

A Time-stepping Method for Relay Systems

7.1 Introduction

Simulation is a common tool (and final escape) when analytical solutions or properties of model equations cannot be derived. It is recognized that new techniques are required for approximating the solution trajectories of hybrid systems. Simulators and languages like Chi (χ) [2], Matlab/Simulink/Stateflow, Modelica [20], Omola/Omsim [1], Psi [4] and SHIFT [8] have recently been developed or added hybrid features to their existing simulation environments. Most of the mentioned hybrid simulators can be categorized as event-driven methods according to a classification made by Moreau [21] in the context of unilaterally constrained mechanical systems.

Event-driven methods are based on considering the simulation interval as a union of disjoint subintervals on which the mode (active constraint set) remains unchanged. On each of these subintervals we are dealing in general with *differential and algebraic equations* (DAE), which can be solved by standard integration routines (*DAE simulation*). As integration proceeds, one has to monitor certain indicators (invariants) to determine when the subinterval ends (*event detection*). At this event time a mode transition occurs, which means that one has to determine what the new mode will be on the next subinterval (*mode selection*). If the state at the event time is not consistent with the selected mode, a jump is necessary (*re-initialization*). The complete numerical method is based on repetitive cycles consisting of DAE simulation, event detection, mode selection and re-initialization.

The idea of *smoothing methods* is to approximately replace the nonsmooth relationships by some regularized ones [21] (see also [13] in which the term "regularization" is used). As an example in a mechanical setting, a non-interpenetrability constraint will be replaced by some stiff repulsion laws and damping actions which are effective as soon as two bodies of the mechanical system come close to each other. The dynamics of the resulting approximate system is then governed by differential equations with sufficient smoothness to be handled through standard numerical techniques. Discrete modes do not really exist anymore, so event detection and mode selection are not necessary. Instantaneous jumps are replaced by (finitely) fast motions, so also the problem of re-initialization disappears. A drawback of this method is that an accurate simulation requires the use of very stiff approximate laws. The time-stepping procedures have to resort to very small step-length and possibly also have to enforce numerical stability by introducing artificial terms in the equations [21]. This results in long simulation times and the effect of the artificial modifications may blur the simulation results.

Time-stepping methods replace the describing equations directly by some "discretized" equivalent. Numerical integration routines are applied to approximate the system equations involving derivatives and all algebraic relations are enforced to hold at each time-step. In this way, one has to solve at each time-step an algebraic problem (sometimes called the "one-step problem") involving information obtained from previous time-steps. In contrast with event-driven methods, time-stepping methods do not determine the event times accurately, but "overstep" them, which puts the consistency of the method into question.

In this chapter we will study linear dynamical systems coupled to relay switches. Such relay systems attract a lot of attention as they are used in many control schemes and are suitable for modeling friction in mechanical systems. In relay systems one may observe chattering and even when the sliding mode is modeled explicitly (as described by Filippov [10]), the system may display an infinite number of relay switches (mode transitions) in a finite interval (see the example in Section 7.3 below). This so-called "Zeno behavior" causes difficulties for simulation methods, especially if one uses an event-driven methodology. In [13] one proposed several techniques to extend simulations beyond the Zeno time: regularization (called smoothing in the discussion above), averaging and Filippov extension (suggested in the context of relay systems also in [19]). The example in Section 7.3 will show that Filippov extension does not always yield a feasible option as Zeno behavior is still present in spite of introducing additional modes corresponding to sliding regimes. Arriving at the Zeno point still requires simulation with an infinite number of mode (relay) switches, which leads to numerical difficulties. Smoothing of the non-Lipschitzian relay characteristic may be an option. However, this route is not taken here. A related paper [5] investigates this method for electrical networks with ideal diodes. The connection to the work described in this paper lies in the fact that *linear complementarity systems* [11,23], a subclass of hybrid dynamical systems, form a superclass of both linear electrical circuits with diodes and linear relay systems.

In this chapter we will study an alternative method based on time-stepping that can handle Zeno behavior for (linear) relay systems. In particular, the question of *consistency* will be of interest: Will the approximations converge to the solution of the original relay system and in what sense? For linear complementarity systems the answer to this question is in general "no" (see Example 6.3.3). However, in the case of linear relay systems consistency can be proven under certain additional assumptions. Moreover, the example in Section 7.3 will be discussed in some detail to show how the proposed method deals with Zeno behavior.

7.2 Linear Relay Systems

Consider the systems given by

$$\dot{x}(t) = Ax(t) + B\bar{u}(t) \tag{7.1a}$$

$$\bar{y}(t) = Cx(t) + D\bar{u}(t) \tag{7.1b}$$

$$\bar{u}_i(t) = \mathrm{sgn}(-\bar{y}_i(t)) \tag{7.1c}$$

with $\bar{u} \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, $\bar{y} \in \mathbb{R}^m$ and A, B, C and D matrices of appropriate dimensions. Each pair $(-\bar{y}_i, \bar{u}_i)$ satisfies an ideal relay characteristic $\bar{u}_i = \mathrm{sgn}(-\bar{y}_i)$, where "sgn" denotes the signum relation as depicted in Figure 7.1. Sometimes, we will also write $(-\bar{y}_i, \bar{u}_i) \in F_{\mathrm{relay}}$.

7.3 Example

A time reversed version of a system studied by Filippov [10, p. 116] (also mentioned in [11, 18, 24]) is given by

$$\dot{x}_1 = -\mathrm{sgn}(x_1) + 2\mathrm{sgn}(x_2) \tag{7.2a}$$

$$\dot{x}_2 = -2\mathrm{sgn}(x_1) - \mathrm{sgn}(x_2). \tag{7.2b}$$

Solutions of this piecewise constant system are spiraling towards the origin, which is an equilibrium. Since $\frac{\mathrm{d}}{\mathrm{d}t}(|x_1(t)| + |x_2(t)|) = -2$ when $x(t) \neq 0$ along trajectories x of the system, solutions reach the origin in finite time (see Figure 7.2 for a trajectory). However, solutions cannot arrive at the origin without going through an infinite number of relay switches (mode transitions). Since these mode switches occur in a finite time interval, the event times contain an accumulation point (i.e. the time that the solution reaches the origin) after which the solution stays at zero. It may be clear that an event-driven methodology will not produce a good approximation, since the method can in principle not simulate beyond the accumulation point. Hence, one has to take recourse to some other

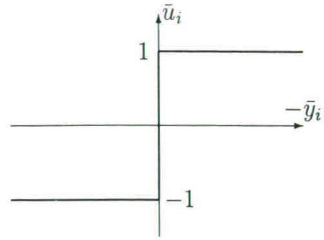


Figure 7.1: Ideal relay characteristic

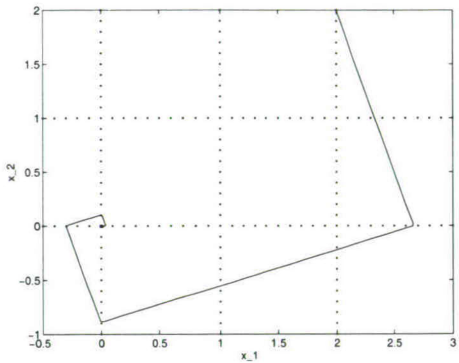


Figure 7.2: Trajectory with initial state $(2,2)^\top$.

techniques.

7.4 The Backward Euler Time-stepping Method

For the numerical simulation of a (linear) relay system we propose the use of time-stepping methods as used in a mechanical context in [22, 25] and for electrical circuits in [15, 16]. The particular method considered here is based on applying the well-known backward Euler scheme to the differential equations and imposing the relay characteristic on every time-step. This converts (7.1) into

$$\frac{x_{j+1} - x_j}{h} = Ax_{j+1} + B\bar{u}_{j+1} \tag{7.3a}$$

$$\bar{y}_{j+1} = Cx_{j+1} + D\bar{u}_{j+1} \tag{7.3b}$$

$$(-\bar{y}_{j+1,i}, \bar{u}_{j+1,i}) \in F_{\text{relay}}, \tag{7.3c}$$

where h is the chosen step-size (assumed to be constant for ease of notation) and $\bar{u}_j$, x_j and $\bar{y}_j$ denote the approximations at time instant $t_j = jh$, $j = 0, 1, 2, \ldots$. The relations (7.3) result in the following algebraic one-step problem, that must be solved for every time-step:

$$\bar{y}_{j+1} = C(I - Ah)^{-1}x_j + \overbrace{[C(\frac{1}{h}I - A)^{-1}B + D]}^{=:G(h^{-1})}\bar{u}_{j+1} \tag{7.4a}$$

$$(-\bar{y}_{j+1,i}, \bar{u}_{j+1,i}) \in F_{\text{relay}}. \tag{7.4b}$$

The update for the state variable follows now from

$$x_{j+1} = (I - Ah)^{-1}x_j + (\frac{1}{h}I - A)^{-1}B\bar{u}_{j+1}. \tag{7.5}$$

Given an initial state $x(0) = x_0$, the scheme starts by setting $x_j := x_0$ and $j := 0$. Solving the one-step problem for j as given in (7.4) results in $\bar{u}_{j+1}$ and $\bar{y}_{j+1}$. Next we can determine x_{j+1} from (7.5) as x_j and $\bar{u}_{j+1}$ are known. The counter j can be increased resulting in a new one-step problem ($j := j + 1$). This cycle is repeated till the desired end time T is reached (i.e. $jh \geq T$). For a given step-size h this procedure results in a sequence of approximations (provided the one-step problems are solvable) $\{\bar{u}_j^h\}$, $\{x_j^h\}$, $\{\bar{y}_j^h\}$ for $j = 1, 2, \dots, \lceil\frac{T}{h}\rceil$ with $\lceil\frac{T}{h}\rceil$ denoting the smallest integer larger than or equal to $\frac{T}{h}$. Hence, we can define a family of approximations as a function of the step-size h. The functions $(\bar{u}^h, x^h, \bar{y}^h)$ are defined on $[0, T]$ as the piecewise constant functions defined by

$$(\bar{u}^h(t), x^h(t), \bar{y}^h(t)) := (\bar{u}_j^h, x_j^h, \bar{y}_j^h) \text{ if } t \in [jh, (j+1)h). \tag{7.6}$$

7.5 Complementarity Framework

Next we discuss two methods to rewrite the one-step problem (7.4) as a so-called linear complementarity problem. It is well-known that a relay characteristic can be reformulated in terms of LCPs (see e.g. [14, 18, 22]). In this section we will discuss two methods. One method will be used to prove unique solvability of the one-step problem (under suitable conditions). The other will be utilized for showing consistency and for numerical implementation.

7.5.1 Solvability of the one-step problem

The first method is described in e.g. [18]. There it is stated that $(-\bar{y}_i, \bar{u}_i) \in F_{\text{relay}}$ (or $\bar{u}_i = \text{sgn}(-\bar{y}_i)$) for all i is equivalent to

$$y_a - y_b = \bar{y} \tag{7.7a}$$

$$u_a = e + \bar{u} \tag{7.7b}$$

$$u_b = e - \bar{u} \tag{7.7c}$$

$$0 \leq u_a \perp y_a \geq 0 \tag{7.7d}$$

$$0 \leq u_b \perp y_b \geq 0 \tag{7.7e}$$

where e denotes the vector (of any dimension) with all components being equal to one.

Combining (7.3) and (7.7), defining $q := C(I - Ah)^{-1}x_j$ and $G := C(h^{-1}I - A)^{-1}B + D$, and finally assuming that $G(h^{-1})$ is invertible, we obtain the LCP (omitting the subscripts for brevity)

$$\begin{pmatrix} u_a \\ u_b \end{pmatrix} = \begin{pmatrix} e - G^{-1}q \\ e + G^{-1}q \end{pmatrix} + \begin{pmatrix} G^{-1} & -G^{-1} \\ -G^{-1} & G^{-1} \end{pmatrix} \begin{pmatrix} y_a \\ y_b \end{pmatrix} \tag{7.8a}$$

$$0 \leq \begin{pmatrix} u_a \\ u_b \end{pmatrix} \perp \begin{pmatrix} y_a \\ y_b \end{pmatrix} \geq 0. \tag{7.8b}$$

Theorem 7.5.1 *If G is a $\mathcal{P}$-matrix, then the linear complementarity problem (7.8) is uniquely solvable for arbitrary q.*

The corresponding corollary for the time-stepping scheme is the following.

Corollary 7.5.2 *Consider the relay system (7.1) and suppose that $G(s) := C(sI - A)^{-1}B + D$ is a $\mathcal{P}$-matrix for all sufficiently large $s \in \mathbb{R}$. The one-step problem (7.3) (or equivalently (7.8)) resulting from applying the time-stepping method based on backward Euler is uniquely solvable for arbitrary x_j and all h sufficiently small.*

7.5.2 Numerical scheme

To approximate the solution to the relay system, one could recursively solve (7.4) by just trying all possibilities of the relay characteristic (exhaustive search) at each time-step. Since each relay has three branches, this amounts to 3^k possibilities that have to be checked. An alternative is the use of LCPs. Although the LCP is NP-hard, which indicates an exponential growth of computing time as a function of the size of the problem (k) (in worst case), the available algorithms have proven to work well in practice and are used for a wide range of applications for simulation of electrical circuits [3, 15, 16] and rigid body dynamics [22, 25].

The reformulation of the one-step problem (7.4) into an LCP of the form (7.8) is only valid under the assumption of invertibility of $G(h^{-1})$ for sufficiently small $h > 0$. In this subsection we will show an alternative modeling method due to Van der Schaft and Schumacher [24] that has two advantages. Firstly, the condition of invertibility is not needed. Secondly, it avoids inclusion of algebraic constraints in the system equations, which would complicate the use of the consistency results of Chapter 6 needed in Section 7.7. The statement $(-\bar{y}_i, \bar{u}_i) \in F_{\text{relay}}$ for all i is equivalent to

$$u^a = \tfrac{1}{2}(e - \bar{u}) \tag{7.9a}$$
$$y^b = \tfrac{1}{2}(e + \bar{u}) \tag{7.9b}$$
$$-\bar{y} = y^a - u^b \tag{7.9c}$$
$$0 \leq u^a \perp y^a \geq 0 \tag{7.9d}$$
$$0 \leq u^b \perp y^b \geq 0. \tag{7.9e}$$

Note that $u_i^a = 0$ and $y_i^b = 0$ cannot occur simultaneously because of (7.9a)-(7.9b). This implies due to the complementarity (7.9e) that either $y_i^b = 0$ or $u_i^a = 0$ must be true. As a consequence of (7.9c), we obtain that $y_i^a = \max(0, -\bar{y}_i)$ and $u_i^b = \max(0, \bar{y}_i)$. Moreover, it follows that $\bar{u} = e - 2u^a$ and $y^b = e - u^a$. The one-step problem (7.4) can thus be rewritten as

$$\begin{pmatrix} y_{j+1}^a \\ y_{j+1}^b \end{pmatrix} = \begin{pmatrix} -C(I - Ah)^{-1}x_j + G(h^{-1})e \\ e \end{pmatrix} + \begin{pmatrix} 2G(h^{-1}) & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} u_{j+1}^a \\ u_{j+1}^b \end{pmatrix} \tag{7.10a}$$

$$0 \leq \begin{pmatrix} u_{j+1}^a \\ u_{j+1}^b \end{pmatrix} \perp \begin{pmatrix} y_{j+1}^a \\ y_{j+1}^b \end{pmatrix} \geq 0. \tag{7.10b}$$

and the update of the state is given by

$$x_{j+1} = (I - Ah)^{-1}x_j + (\frac{1}{h}I - A)^{-1}B[e - 2u_{j+1}^a]. \tag{7.11}$$

Note that solvability of (7.4) and (7.10) are equivalent. Due to the relation between (7.4) and (7.8) Corollary 7.5.2 also applies to (7.10) under the conditions stated.

7.6 Linear Complementarity Systems

The modeling of (7.9) can be directly applied (before any discussion on approximation schemes) to the relay system (7.1) to obtain the following dynamical extension of the LCP:

$$\dot{x} = Ax + Be - 2Bu^a \tag{7.12a}$$
$$\begin{pmatrix} y^a \\ y^b \end{pmatrix} = \begin{pmatrix} -Cx - De \\ e \end{pmatrix} + \begin{pmatrix} 2D & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} u^a \\ u^b \end{pmatrix} \tag{7.12b}$$
$$0 \leq u^a \perp y^a \geq 0 \tag{7.12c}$$
$$0 \leq u^b \perp y^b \geq 0. \tag{7.12d}$$

The general form of such system descriptions is given by

$$\dot{x}(t) = \tilde{A}x(t) + \tilde{B}u(t) + \tilde{f} \tag{7.13a}$$

$$y(t) = \tilde{C}x(t) + \tilde{D}u(t) + \tilde{g} \tag{7.13b}$$

$$0 \leq y(t) \perp u(t) \geq 0 \tag{7.13c}$$

and called a linear complementarity system (LCS). Systems of this form have been introduced in [23] and were studied further in e.g. [6, 11, 18].

7.7 Consistency of Time-stepping for Relay Systems

By applying backward Euler to the LCS (7.13), we obtain the one-step problem

$$\frac{x_{j+1} - x_j}{h} = \tilde{A}x_{j+1} + \tilde{B}u_{j+1} + \tilde{f} \tag{7.14a}$$

$$y_{j+1} = \tilde{x}_{j+1} + \tilde{D}u_{j+1} + \tilde{g} \tag{7.14b}$$

$$0 \leq y_{j+1} \perp u_{j+1} \geq 0, \tag{7.14c}$$

which is equivalent to the LCP

$$y_{j+1} = \tilde{g} + \tilde{(}I - h\tilde{A})^{-1}\{x_j + h\tilde{f}\} + [\tilde{D} + h\tilde{(}I - h\tilde{A})^{-1}\tilde{B}]u_{j+1} \tag{7.15a}$$

$$0 \leq y_{j+1} \perp u_{j+1} \geq 0 \tag{7.15b}$$

and

$$x_{j+1} := (I - h\tilde{A})^{-1}\{x_j + h\tilde{f}\} + h(I - h\tilde{A})^{-1}\tilde{B}u_{j+1}. \tag{7.16}$$

Applying this backward Euler time-stepping scheme to the LCS (7.12) obtained from the complementarity reformulation of (7.1) yields the approximation scheme from subsection 7.5.2, i.e. the one given by (7.10) and (7.11). Hence, the order of complementarity reformulation and application of the time-stepping scheme to (7.1) is irrelevant for the resulting approximation scheme.

In Chapter 6 the *consistency* – indicating the existence of a sequence of approximations that converges to an actual solution trajectory of the original system description with the same initial condition – of time-stepping methods has been investigated for linear electrical networks with ideal diodes. One should be cautious in applying a time-stepping method to a general LCS (or other multimodal or hybrid systems). This is illustrated by an example of a triple integrator connected to complementarity conditions for which it has been shown that the approximations are not even bounded (see Example 6.3.3). As a consequence, verification of the numerical scheme in the sense of showing consistency is needed.

For the backward Euler time-stepping method the following fairly general result has been proven for LCS in Chapter 6[1].

Theorem 7.7.1 *Consider the LCS given by (7.13) such that the one-step problems given by (7.15) are uniquely solvable for all sufficiently small h. Let $T > 0$ and $x_0 \in \mathbb{R}^n$ be given. Also let $(u^h, x^h, y^h) \in \mathcal{L}_2([0,T], \mathbb{R}^{m+n+m})$ be the piecewise constant approximations obtained for step-size h and initial state x_0 as in (7.6). Suppose that there exists an $\alpha > 0$ such that $\|u^h\| \leq \alpha$ for all sufficiently small h, where $\|\cdot\|$ denotes the $\mathcal{L}_2$-norm. Suppose that $\tilde{D}$ is nonnegative definite (not necessarily symmetric). Then the following holds for any sequence $\{h_k\}$ of step-sizes that converges to zero.*

1. *There exists a subsequence $\{h_{k_l}\} \subseteq \{h_k\}$ such that $(u^{h_{k_l}}, y^{h_{k_l}})$ converges weakly in $\mathcal{L}_2$ to some (u, y) and $x^{h_{k_l}}$ converges in $\mathcal{L}_2$ to some x.*

2. *The triple (u, x, y) is a solution to the LCS* (7.13) *on $[0,T]$ for initial state x_0 in the sense that for almost all $t \in (0,T)$*

$$x(t) = x_0 + \int_0^t [\tilde{A}x(\tau) + \tilde{B}u(\tau)]d\tau + \tilde{f}t \tag{7.17a}$$

$$y(t) = \tilde{C}x(t) + \tilde{D}u(t) + \tilde{g} \tag{7.17b}$$

$$0 \leq u(t) \perp y(t) \geq 0. \tag{7.17c}$$

3. *If the solution (u, x, y) is unique for the initial state x_0 in the sense of (7.17), then the sequence (u^{h_k}, y^{h_k}) as such converges weakly to (u, y) and x^{h_k} converges to x.*

This theorem will be applied to the relay system (7.1) by converting it to the LCS in (7.12).

Theorem 7.7.2 *Consider the relay system (7.1) and suppose that $G(s) := C(sI-A)^{-1}B + D$ is a $\mathcal{P}$-matrix for all sufficiently large $s \in \mathbb{R}$ and D is nonnegative definite*[2]*. Let $T > 0$ and $x_0 \in \mathbb{R}^n$ be given. Let $\{h_k\}$ converge to zero and consider the piecewise constant approximations $(\bar{u}^{h_k}, x^{h_k}, \bar{y}^{h_k})$ given by (7.6). Then the following holds for any sequence $\{h_k\}$ of step-sizes that converges to zero.*

1. *There exists a subsequence $\{h_{k_l}\} \subseteq \{h_k\}$ such that $(\bar{u}^{h_{k_l}}, \bar{y}^{h_{k_l}})$ converges weakly in $\mathcal{L}_2$ to some $(\bar{u}, \bar{y})$ and $x^{h_{k_l}}$ converges in $\mathcal{L}_2$ to some x.*

[1] The result stated in Chapter 6 is more general in the sense that it even includes the possibility of impulsive motions, i.e. Dirac delta distributions, and the corresponding re-initializations in the solution trajectories.

[2] If D is symmetric the last condition can be dropped, since it is then implied by the first. See page 147 in [7].

2. The triple $(\bar{u}, x, \bar{y}) \in \mathcal{L}_2([0,T], \mathbb{R}^{m+n+m})$ is a solution to the relay system (7.1) on $[0,T]$ with initial state x_0 in the sense that for almost all $t \in [0,T]$

$$x(t) = x_0 + \int_0^t [Ax(\tau) + B\bar{u}(\tau)]d\tau \tag{7.18a}$$

$$\bar{y}(t) = \tilde{C}x(t) + D\bar{u}(t) \tag{7.18b}$$

$$\bar{u}_i(t) = \mathrm{sgn}(-\bar{y}_i(t)). \tag{7.18c}$$

3. If the solution $(\bar{u}, x, \bar{y}) \in \mathcal{L}_2([0,T], \mathbb{R}^{m+n+m})$ is unique for the initial state x_0 in the sense of (7.18), then the complete sequence (u^{h_k}, y^{h_k}) converges weakly to (u, y) and x^{h_k} converges to x.

Note that the theorem also guarantees the global existence of a solution to the linear relay system under the assumptions stated.

7.8 Example

The example of Section 7.3 can be written in the form (7.1) with

$$A = D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}; C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that $G(s) = C(sI - A)^{-1}B + D = CBs^{-1}$ is a $\mathcal{P}$-matrix for all $s > 0$. Hence, the theorem above guarantees the existence of a sequence of step sizes for which the corresponding approximations converge to an actual solution of the relay system given an initial state. Although in [18] the uniqueness of solutions has been proven for relay systems under the condition that $G(s)$ is a $\mathcal{P}$-matrix for sufficiently large s, the kind of uniqueness does not correspond to the $\mathcal{L}_2$-uniqueness as formulated by (7.18). The reason is that in [18] left-accumulations (see [12]) of events are excluded in the solution concept (which we will denote by "forward sense"). Hence, convergence of any arbitrary sequence of approximations cannot be concluded in general from Theorem 7.7.2 item 3.

The difference between the $\mathcal{L}_2$- and the forward sense uniqueness can be illustrated best by considering the time-reversed version of the system in Section 7.3 (which is then the original example in [10, p. 116]) given by

$$\dot{x}_1 = \mathrm{sgn}(x_1) - 2\mathrm{sgn}(x_2) \tag{7.19a}$$

$$\dot{x}_2 = 2\mathrm{sgn}(x_1) + \mathrm{sgn}(x_2). \tag{7.19b}$$

This system has (infinitely many) solutions in the sense of (7.18) corresponding to initial state $x_0 = 0$. To see this, observe that infinitely many solution trajectories in (7.2) reach the origin in finite time (e.g. the trajectory depicted in Figure 7.2). The time-reversed trajectories satisfy (7.19) in the $\mathcal{L}_2$-sense of (7.18) with the origin as initial state. These trajectories start with a left-accumulation point of events at the initial time (see [12] for more details). Note that the zero trajectory satisfies (7.19) in $\mathcal{L}_2$-sense as well. However, the solution concept in "forward sense" as used in [18] allows only the zero-trajectory for initial state $x_0 = 0$[3]. These phenomena might obstruct the uniqueness needed to apply Theorem 7.2 item 3, which would guarantee the convergence of any arbitrary sequence of approximations. However, for the example at hand (7.2) $\mathcal{L}_2$-uniqueness can be proven and consequently, convergence of any sequence of approximations is guaranteed.

We return now to the simulation of (7.2) by the backward Euler time-stepping scheme. The discretization (7.3) results for (7.2) in

$$\frac{x_{1,j+1} - x_{1,j}}{h} = -\mathrm{sgn}(x_{1,j+1}) + 2\mathrm{sgn}(x_{2,j+1}) \tag{7.20a}$$

$$\frac{x_{2,j+1} - x_{2,j}}{h} = -2\mathrm{sgn}(x_{1,j+1}) - \mathrm{sgn}(x_{2,j+1}). \tag{7.20b}$$

This problem has to be solved for given $x_{1,j}$ and $x_{2,j}$ in the unknowns $x_{1,j+1}$ and $x_{2,j+1}$. Considering the three possibilities for each relay characteristic yields nine (discrete) possibilities. Since the problem is uniquely solvable for each combination of $x_{1,j}$ and $x_{2,j}$ according to Corollary 7.5.2, the nine areas lead to a partitioning of the state space (see also [24, p. 30]). One of the nine possibilities is the case where $x_{1,j+1} = 0$ and $x_{2,j+1} = 0$ (both relays will be in the middle branches ("sliding modes")). We can derive necessary and sufficient conditions on $x_{1,j}$, $x_{2,j}$ for this being the right mode. The conditions follow from (7.20) by realizing that the values of $\bar{u}_{1,j+1} := -\mathrm{sgn}(x_{1,j+1})$ and $\bar{u}_{2,j+1} = -\mathrm{sgn}(x_{2,j+1})$ must be contained in $[-1, 1]$. These conditions correspond to the central area in Figure 7.3. Hence, if the previous state $x_{1,j}$, $x_{2,j}$ lies in this central area, the new state will be the origin. The figure shows that the discretized system behaves like the original continuous system except in the vertical and horizontal strips that do not have much influence on the solution trajectory. Only in the central area the behavior of the discretized and the original system differ considerably. The discretized solution "jumps" to the origin in one discrete step, while the continuous solution continues to go through (infinitely many) mode changes at an increasing speed. After the discretized system jumps to the origin, it stays there. Hence, the discretized system reaches the origin in finitely many steps. The theory presented above guarantees that each sequence of approximations converges to the unique solution of the original system.

[3]Interestingly, the backward Euler time-stepping scheme applied to this Filippov example generates only zero-trajectories as approximations starting from the origin. Hence, this discretization method might inherently use some "forward sense" as well.

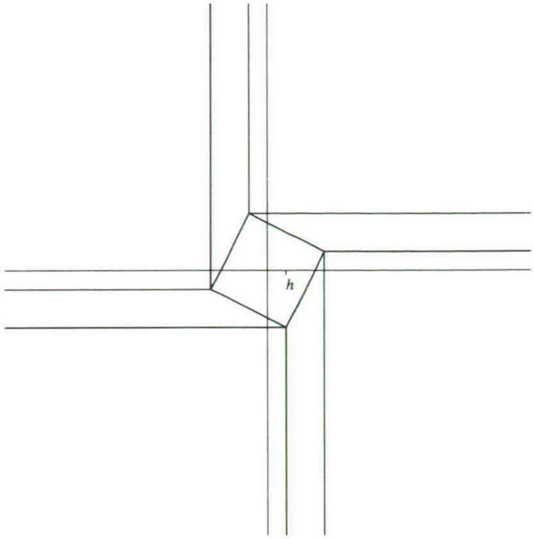


Figure 7.3: Partitioning of plane by relay system.

We simulated the above system for the step-sizes $h = 1$, $h = 0.1$ and $h = 0.01$ and initial state $(2,2)^\top$. The simulation results can be found in Figure 7.4. For $h = 1$ the origin is reached within two steps. For $h = 0.1$ the system is at time 1.8 exactly in the origin, while for $h = 0.01$ this occurs at time 1.9. For decreasing step sizes this value gets closer and closer to the exact accumulation point 2 for the original system. Note that the simulation is exact beyond time 2 for all step sizes. The time-stepping method is able to deal with the Zeno behavior in this example satisfactorily. Moreover, the convergence of the approximations has been guaranteed.

7.9 Lemke's Method

The most well-known algorithm to solve an LCP is the complementary pivoting scheme due to C.E. Lemke [17]. Under the condition that M is a $\mathcal{L}$-matrix, Eaves has proven that Lemke's method produce a solution to LCP(q, M), provided a solution exists [9].

Definition 7.9.1 A matrix $M \in \mathbb{R}^{k\times k}$ is an $\mathcal{L}$-matrix, if

1. For all $w \geq 0$, $w \neq 0$ there exists a j such that $w_k > 0$ and $(Mw)_k \geq 0$.

2. There exist diagonal matrices $\Lambda \geq 0$ and $\Omega \geq 0$ satisfying $\Omega w \neq 0$ and $(\Lambda M + M^\top \Omega)w = 0$ for all nonzero $w \in \mathrm{SOL}(0, M)$.

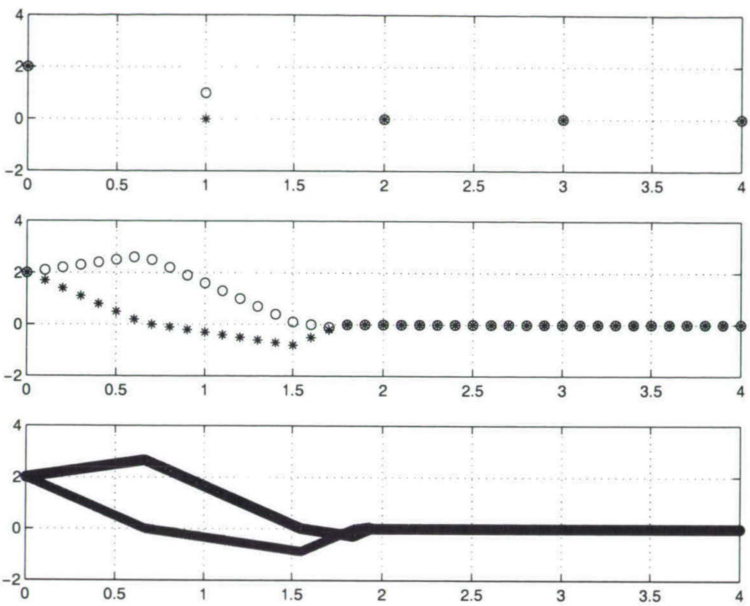


Figure 7.4: Simulation of the x_1 and x_2 trajectories for $h = 1$ (top), $h = 0.1$ (middle) and $h = 0.01$ (bottom).

The following result claims that Lemke's algorithm will process the one-step problem (7.10) successfully under the condition that $G(h^{-1})$ is a $\mathcal{P}$-matrix for all sufficiently small h.

Theorem 7.9.2 *If G is a $\mathcal{P}$-matrix, then the matrix* $M := \begin{pmatrix} G & I \\ -I & 0 \end{pmatrix}$ *is an $\mathcal{L}$-matrix.*

7.10 Conclusions

In this paper we proposed and analyzed a time-stepping method for simulating a class of hybrid dynamical systems, to wit linear relay systems. One motivation for considering a time-stepping method instead of an event-driven method, as is more usual in the context of hybrid systems, is the possible occurrence of Zeno behavior. A relay system exhibiting this kind of phenomena was presented in Section 3. In spite of the possible presence of Zeno trajectories and the fact that event times are overstepped, a formal proof of the convergence of the approximations to an actual solution of the linear relay systems was given under certain additional assumptions (which guarantee well-posedness in "forward sense"). This justifies the use of the method and shows that it is an alternative technique for simulating systems exhibiting Zeno behavior. This has been demonstrated by an example as well.

The consistency that we showed in the paper guaranteed the existence of a sequence of step-sizes such that the corresponding approximations converge to the actual solution trajectory. To obtain that any arbitrary sequence of approximations converges, it is sufficient to prove $\mathcal{L}_2$-uniqueness of the solutions to the relay system. Unfortunately, uniqueness in the sense of $\mathcal{L}_2$ does not necessarily hold when uniqueness in the "forward sense" is true as shown by Filippov's example (7.19). Under conditions related to passivity, $\mathcal{L}_2$-uniqueness of solutions to linear complementarity systems has been proven in Chapter 3.

7.11 Proofs

7.11.1 Proof of Theorem 7.5.1

In order to prove Theorem 7.5.1, we need some preparations. The following two lemmas, which have a rather technical nature, will be employed later.

Lemma 7.11.1 *If $M \in \mathbb{R}^{m\times m}$ is a $\mathcal{P}$-matrix then so is M^{-1}.*

Proof: According to [7, Theorem 3.3.4], the following two statements are equivalent:

1. M is a $\mathcal{P}$-matrix.

2. $(z_i(Mz)_i \leq 0 \text{ for all } i \in \overline{m}) \Rightarrow z = 0$.

Let $z \in \mathbb{R}^m$ and define y by $y = M^{-1}z$ then

$$z_i(M^{-1}z)_i = (My)_i(M^{-1}My)_i = y_i(My)_i. \tag{7.21}$$

Since M is a $\mathcal{P}$-matrix, the implication $(y_i(My)_i \leq 0 \text{ for all } i \in \overline{m}) \Rightarrow y = 0$ holds. It follows from (7.21) that M^{-1} is a $\mathcal{P}$-matrix. ■

Lemma 7.11.2 *Let $M \in \mathbb{R}^{m\times m}$ be $\mathcal{P}$-matrix and*

$$N = \begin{pmatrix} I & -I \end{pmatrix}^\top M \begin{pmatrix} I & -I \end{pmatrix}.$$

Then, the following implication holds:

$$(z_i(Nz)_i \leq 0 \textit{ for all } i \in \overline{n}) \Rightarrow Nz = 0.$$

Proof: Assume that $z_i(Nz)_i \leq 0$ for all $i \in \overline{2m}$ for some $z \in \mathbb{R}^{2m}$. Let $z = \mathrm{col}(u, v)$ where $u,\, v \in \mathbb{R}^m$. then, we get

$$Nz = \begin{pmatrix} I & -I \end{pmatrix}^\top M \begin{pmatrix} I & -I \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} M(u-v) \\ -M(u-v) \end{pmatrix}.$$

Hence, we have

$$z_i(Nz)_i = \begin{cases} u_i(M(u-v))_i & \text{if } i = 1, 2, \ldots, m, \\ -v_{i-k}(M(u-v))_{i-k} & \text{if } i = m+1, m+2, \ldots, 2m. \end{cases}$$

So, for all $i \in \overline{m}$

$$u_i(M(u-v))_i \leq 0 \tag{7.22}$$

$$-v_i(M(u-v))_i \leq 0. \tag{7.23}$$

Therefore, we get $(u_i - v_i)(M(u-v))_i \leq 0$ for all $i \in \overline{m}$. Since M is a $\mathcal{P}$-matrix, it follows from [7, Theorem 3.3.4] that $u - v = 0$. Consequently, $Nz = 0$. ■

Next, we recall the notion of row sufficient matrix from [7].

Definition 7.11.3 A matrix $N \in \mathbb{R}^{n\times n}$ is said to be *row sufficient* if the implication $(z_i(N^\top z)_i \leq 0$ for all $i \in \overline{n}) \Rightarrow (z_i(N^\top z)_i = 0$ for all $i \in \overline{n})$ holds.

Lemma 7.11.4 *If G is a $\mathcal{P}$-matrix then the matrix*

$$\begin{pmatrix} G^{-1} & -G^{-1} \\ -G^{-1} & G^{-1} \end{pmatrix}$$

is row sufficient.

Proof: It follows from a direct application of Lemma 7.11.2 with $M = G^{-\top}$. ■

As shown in the following lemma, the LCP (7.8) is feasible. This fact will be used to show its solvability later.

Lemma 7.11.5 *If G is a $\mathcal{P}$-matrix then the linear complementarity problem* (7.8) *is feasible for all q.*

Proof: Let $y^a = q^+$ and $y^b = q^-$ where q^+ and q^- are the nonnegative and nonpositive parts of the vector q. It is clear that both y^a and y^b are nonnegative. Substituting the vector $\mathrm{col}(y^a, y^b)$ into (7.8a), we get

$$u^a = e - G^{-1}q + G^{-1}q^+ - G^{-1}q^- = e \geq 0 \tag{7.24}$$

$$u^b = e + G^{-1}q - G^{-1}q^+ + G^{-1}q^- = e \geq 0. \tag{7.25}$$

So the LCP (7.8) is feasible. ■

Now, we are in a position that we can prove Theorem 7.5.1.

Proof of Theorem 7.5.1:

solvability: It follows from [7, Corollary 3.5.5] and Lemmas 7.11.4 and 7.11.5 that LCP (7.8) is solvable.

uniqueness: Suppose that $\mathrm{col}(u^{a,i}, u^{b,i})$ and $\mathrm{col}(y^{a,i}, y^{b,i})$ for $i = 1, 2$ are two solutions of the LCP (7.8). It follows from Lemma 7.11.2 and [7, Theorem 3.4.4] that $\mathrm{col}(y^{a,i}, y^{b,i})$ for $i = 1, 2$ satisfies

$$\begin{pmatrix} G^{-1} & -G^{-1} \\ -G^{-1} & G^{-1} \end{pmatrix} \begin{pmatrix} y^{a,1} - y^{a,2} \\ y^{b,1} - y^{b,2} \end{pmatrix} = 0.$$

So, $G^{-1}((y^{a,1} - y^{a,2}) - (y^{b,1} - y^{b,2})) = 0$. Since G is a $\mathcal{P}$-matrix, so is G^{-1} due to Lemma 7.11.1. Therefore, $y^{a,1} - y^{a,2} = y^{b,1} - y^{b,2}$. Now, define the vector $\Delta \in \mathbb{R}^m$ by $\Delta = y^{a,1} - y^{a,2} = y^{b,1} - y^{b,2}$. This results in

$$y_i^{a,1} y_i^{b,1} = y_i^{a,2} y_i^{b,2} + (y^{a,2} + y^{b,2})_i \Delta_i + \Delta_i^2 \tag{7.26}$$

for all $i \in \overline{m}$. It follows from (7.7a) that $y^{a,i}$ and $y^{b,i}$ are both nonnegative and $(y^{a,i})^\top y^{b,i} = 0$ for $i = 1, 2$. Then, from (7.26) we get $(y_i^{a,2} + y_i^{b,2})\Delta_i + \Delta_i^2$, so $\Delta_i = -(y^{a,2} + y^{b,2})_i$ or $\Delta_i = 0$. Both possibilities result in $\Delta_i = 0$ for $i \in \overline{m}$. Consequently, we can conclude that $u^{a,1} = u^{a,2}$, $u^{b,1} = u^{b,2}$, $y^{a,1} = y^{a,2}$ and $y^{b,1} = y^{b,2}$. ■

7.11.2 The remaining proofs

Proof of Theorem 7.7.1: It follows from Theorem 6.3.4 by considering the extended state $\bar{x} := \mathrm{col}(x, \tilde{f}, \tilde{g})$. ■

Proof of Theorem 7.7.2: From Corollary 7.5.2 we obtain that the one-step problems are uniquely solvable for sufficiently small h. Moreover, if D is nonnegative definite, then the $\tilde{D}$-matrix of the corresponding LCS (7.12) is also nonnegative definite. Hence, we only have to show the uniform boundedness of the approximations $u^{a,h}$ and $u^{b,h}$ as appearing in (7.10) and (7.12) in the sense of $\mathcal{L}_2$. It is clear that there exists an M_1 such that $\|\bar{u}^h\|_\infty \leq M_1$ for all h as all components of $\bar{u}^h$ are contained in $[-1, 1]$. Here $\bar{u}^h$ denotes the approximation of $\bar{u}$ for step-size h and $\|\cdot\|_\infty$ is the $\mathcal{L}_\infty$-norm. The equations (7.9a)-(7.9b) yield now that also $\|u^{a,h}\|_\infty \leq M_2$ and $\|y^{b,h}\|_\infty \leq M_2$ for all h. Since there exist $\alpha > 0$ and $\beta > 0$ such that $\|(I - Ah)^{-1}\| \leq 1 + \alpha h \leq \beta$ for all sufficiently small h, we obtain from (7.11) that there exists a constant $\gamma > 0$ such that $\|x^h_{j+1}\| \leq (1 + \gamma h)\|x^h_j\|$ for all sufficiently small h. This yields the existence of an M_3 such that $\|x^h_j\| \leq M_3$ for all sufficiently small h and all $j = 1, 2, \ldots, \lceil \frac{T}{h} \rceil$ (see the proof of Lemma 6.6.15 item 2). Since $\bar{y}^h_j = Cx^h_j + D\bar{u}^h_j$ and x^h, $\bar{u}^h$ are uniformly bounded in h, it follows that the approximations satisfy $\|\bar{y}^h_j\| \leq M_4$ for all sufficiently small h and all $j = 1, 2, \ldots, \lceil \frac{T}{h} \rceil$. From the discussion after (7.9) it follows that $y^a = \max(0, -\bar{y})$ and $u^b = \max(0, \bar{y})$ (interpret "max" componentwise), so these quantities are uniformly bounded in h. Hence, we showed boundedness of $\|u^{a,h}\|_\infty$ and $\|u^{b,h}\|_\infty$ for all sufficiently small h and consequently the required $\mathcal{L}_2$-boundedness. This completes the proof, since we can apply Theorem 7.7.1 and immediately translate all results from the LCS to the original relay system. ■

Proof of Theorem 7.9.2: Note that

$$M \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{=w} = \begin{pmatrix} Gu + v \\ -u \end{pmatrix}.$$

In case $u = 0$, it is clear that there exists a j such that $w_{j+k} = v_j > 0$ and $(Mw)_{j+k} = 0$. So, for this case statement *1* holds. In case $u \neq 0$ and $u \geq 0$, there must exist a j such that $u_j > 0$ and $(Gu)_j > 0$. The reason is that a $\mathcal{P}$-matrix G does not reverse the sign of any nonzero vector u [7, Thm. 3.3.4]. Hence, $w_j = u_j > 0$ and $(Mw)_j = (Gu)_j + v_j > 0$, because $vg \geq 0$. It can easily be verified that $\text{SOL}(0, M) = \{\begin{pmatrix} 0 \\ v \end{pmatrix} \mid v \in \mathbb{R}^k\}$. Take

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

The first part of statement *2* follows from $\Omega u = u$ for all $u \in \text{SOL}(0, M)$. The second part is obtained from computing $\Lambda M + M^\top \Omega$ which is equal to

$$\begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}.$$

If we combine this with the special form that elements in $\text{SOL}(0, M)$ have, the proof is complete. ■

References

[1] M. Andersson, S.E. Mattsson, D. Brück, and T. Schöntal. Omsim - An integrated evironment for object-oriented modelling and simulation. In *Proceedings of the IEEE/IFAC joint symposium on Computer-Aided Control System Design*, Tucson, Arizona, pages 285–290, 1994.

[2] D.A. van Beek, S.H.F. Gordijn, and J.E. Rooda. Integrating continuous-time and discrete-event concepts in modelling and simulation of manufacturing machines. *Journal of simulation practice and theory*, 5:653–669, 1997.

[3] W.M.G. van Bokhoven. *Piecewise Linear Modelling and Analysis*. Kluwer, Deventer, the Netherlands, 1981.

[4] P.P.J. van den Bosch, H. Butler, A.R.M. Soeterboek, and M.M.W.G. Zaat. *Modelling and simulation with PSI/c*. BOZA Automatisering BV, Nuenen, The Netherlands, 1995.

[5] M.K. Çamlıbel, M.K.K. Cevik, W.P.M.H. Heemels, and J.M. Schumacher. From Lipschitzian to non-Lipschitzian characteristics: continuity of behaviors. In *Proc. of the 39th IEEE Conference on Decision and Control*, Sydney (Australia), 2000.

[6] M.K. Çamlıbel, W.P.M.H. Heemels, and J.M. Schumacher. Dynamical analysis of linear passive networks with diodes. Part II: Consistency of a time-stepping method. Technical Report 00 I/03, Eindhoven University of Technology, Dept. of Electrical Engineering, Measurement and Control Systems, Eindhoven, The Netherlands, 2000, submitted for publication.

[7] R.W. Cottle, J.-S. Pang, and R.E. Stone. *The Linear Complementarity Problem*. Academic Press, Boston, 1992.

[8] A. Deshpande, A. Göllü, and L. Semenzato. The SHIFT programming language for dynamic networks of hybrid automata. *IEEE TAC*, 43(4):584–587, 1998.

[9] B. C. Eaves. The linear complementarity problem. *Management Science*, 17:612–634, 1971.

[10] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides*. Mathematics and Its Applications. Kluwer, Dordrecht, The Netherlands, 1988.

[11] W.P.M.H. Heemels. *Linear complementarity systems: a study in hybrid dynamics*. Ph.D. Thesis of the Eindhoven University of Technology, Dept. of Electrical Engineering, Eindhoven, The Netherlands, 1999.

[12] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Well-posedness of linear complementarity systems. In *38-th IEEE Conference on Decision and Control,* Phoenix (USA), pages 3037–3042, 1999.

[13] K.H. Johansson, J. Lygeros, S. Sastry, and M. Egerstedt. Simulation of Zeno hybrid automata. In *38-th IEEE Conference on Decision and Control,* Phoenix (USA), pages 3538–3543, 1999.

[14] A. Klarbring. A mathematical programming approach to contact problems with friction and varying contact surface. *Computers & Structures*, 30(5):1185–1198, 1986.

[15] D.M.W. Leenaerts. On linear dynamic complementary systems. *IEEE Transactions on Circuits and Systems-I*, 46(8):1022–1026, 1999.

[16] D.M.W. Leenaerts and W.M.G. van Bokhoven. *Piecewise linear modelling and analysis.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[17] C. E. Lemke. On complementary pivot theory. In G.B. Dantzig and A. F. Veinott, JR., editors, *Mathematics of the Decision Sciences, Part 1*, pages 95–114. American Mathematical Society, Rhode Island, 1968.

[18] Y.J. Lootsma, A.J. van der Schaft, and M.K. Çamlıbel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[19] S.E. Mattson. On object-oriented modeling of relays and sliding mode behaviour. In *Preprints 13th IFAC World Congress,* volume F, pages 259–264, 1996.

[20] S.E. Mattsson, H. Elmqvist, and J.F. Broenink. Modelica: an international effort to design the next generation modelling language. *Journal A*, 38(3):16–19, 1997.

[21] J.J. Moreau. Numerical aspects of the sweeping process. *Comput. Methods Appl. Mech. Engrg.*, 177(3-4):329–349, 1999.

[22] F. Pfeiffer and C. Glocker. *Multibody Dynamics with Unilateral Contacts.* Wiley, Chichester, 1996.

[23] A.J. van der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9:266–301, 1996.

[24] A.J. van der Schaft and J.M. Schumacher. *An introduction to hybrid dynamical systems*, volume 251 of *Lecture Notes in Control and Information Sciences.* Springer, London, 1999.

[25] D.E. Stewart. Convergence of a time-stepping scheme for rigid body dynamics and resolution of Painlevé's problem. *Archive for Rational Mechanics and Analysis*, 145(3):215–260, 1998.

Chapter 8

Conclusions

A class of piecewise linear systems that we call complementarity systems has been under consideration. The two main themes were well-posedness and continuity problems.

8.1 Contributions

The contributions of the thesis can be summarized as follows.

- Sufficient conditions for the existence and uniqueness of solutions, in $\mathcal{PB}$-sense, to (low index) linear complementarity systems (LCS), which are subjected to $\mathcal{PB}$-like inputs, have been established.

- It has been shown that the solutions of linear passive complementarity systems (LPCS) do exist and are unique in $\mathcal{L}_2$-sense which is more general than $\mathcal{PB}$-sense.

- As a generalization of linear passive systems, a new class of systems that are passifiable by pole shifting (PPS) has been introduced. After providing necessary and sufficient conditions for a system to be PPS, it has been proven that all the results that were presented for LPCS hold for linear PPS complementarity system.

- For both senses of solutions, the regular initial states, the initial states for which there is a solution, have been characterized in terms of solvability of some linear complementarity problems. In fact, it has been shown that the set of regular initial states is the same for both cases.

- As a side result of the study of linear PPS complementarity systems, we have shown that Zeno behavior cannot occur for certain LCSs. For a larger class of LCSs, we also proved that the zero state (which is the most natural candidate) cannot be a Zeno state.

- Based on a motivation from mechanical systems, a jump rule in terms of stored energy for nonregular initial states of LPCS has been introduced. Several different characterizations of this jump rule have been established. Particularly, the one with the so-called rational complementarity problem has been used to define a distributional solution concept which treats nonregular initial states as well. Previously presented well-posedness results have been reformulated in this distributional framework.

- By following the footsteps of the work has been done for LCS, a class of piecewise linear systems, which can be formulated with the help of complementarity methods, has been considered with an eye towards the existence and uniqueness of solutions. It has been shown that similar sign conditions as in the case of LCS are sufficient for well-posedness in $\mathcal{PB}$-sense.

- We presented some sufficient conditions for the convergence of the approximations that are obtained by replacing the non-Lipschitzian complementarity characteristic with *close* Lipschitzian characteristics. More precisely, it has been shown that for a given regular initial state of a linear PPS complementarity system the state trajectories of such approximations converge in suitable senses to the state trajectory of the original system if the approximating systems satisfy uniform passifiability condition. The convergence of (u, y)-trajectories in $\mathcal{L}_2$-weak-sense has also been proven under the additional condition of uniform $\mathcal{L}_2$-boundedness of u-trajectories of approximating systems. For more general approximations, what we could reach was to show $\mathcal{L}_2$-strong-sense convergence of the state trajectories and $\mathcal{L}_2$-weak-sense convergence of (u, y)-trajectories if the u-trajectories of approximating systems is uniformly $\mathcal{L}_2$-bounded.

- For LPCS and linear relay systems, the Backward Euler method was proven to be consistent as a time-stepping method in the sense that the approximating trajectories obtained by the Backward Euler method do convergence to the actual solution of the original system in a suitable sense. Another achievement for LPCS was the continuous dependence of the solutions to the initial states.

8.2 Further Research Topics

As the completion/extension of our work, the following points can be considered for further study.

- $\mathcal{L}_2$-uniqueness of the solutions to general LCS, with its potential implications for consistency of time-stepping methods, deserves to be studied further.

- The study of Zeno behavior is of considerable importance not only from a hybrid systems point of view but also for continuity problems. Indeed, absence of Zeno would yield stronger convergence results.

- We believe that uniformly passifiable families of systems can be exploited further than we have done in Chapter 5. So, study of uniformly passifiable systems is on our near-future research programme.

- Generalization of time-stepping methods to piecewise linear systems can be done by following the footsteps of Chapter 6 and 7.

- In the treatment of well-posedness problems, one of the standing assumptions was *the low index assumption.* Systems with higher index may arise from the context of optimal control problems with state constraints in particular. Only this particular application area provides enough motivation for future work in this direction.

- In Chapter 2, only $\mathcal{PB}$-like inputs have been allowed. Although from a modularity point of view this class of inputs is enough to cover interconnections of LCSs, such a restriction on the inputs is not a quite natural one. Therefore, the possible generalization to larger classes of inputs is of interest.

As items of a long-term *programme*, we might think of the following research directions/subjects.

- The well-posedness of nonlinear complementarity systems can be dealt with by using time-stepping methods. Indeed, these methods have been employed for showing existence of solutions in mechanical systems context.

- Having studied well-posedness, stability issues are naturally in order. One may take the extensions of Lyapunov theory to hybrid systems as a starting point. For linear passive complementarity systems, we know that there exists a common quadratic Lyapunov function for all modes. However, our impression is that even the additional structure offered by complementarity systems does not help too much and stability issues are very far from being trivial for general LCSs.

- With an aim to develop control theory for complementarity systems, we should first address controllability and observability issues. Again existing literature for hybrid systems gives some hints.

Summary

The main object of this thesis is a class of piecewise linear dynamical systems that lie in the realm of the intersection of system theory and mathematical programming. We call them *complementarity systems*. For these *nonlinear* and *nonsmooth* dynamical systems, our research is concentrated on two themes: *well-posedness* and *approximations*.

The well-posedness issue, in the sense of existence and uniqueness of solutions, is of considerable importance from a model validation point of view. If the physical system that is being modeled is deterministic in the sense that it shows identical behavior under identical circumstances, then the mathematical model should have the same property. Model validity would be put into serious doubt if it would turn out that the equations of the mathematical model allow multiple solutions for some initial data. With any model formulation for a deterministic physical system it is therefore important to establish well-posedness of the model. The first part of the thesis is devoted to well-posedness issue. We provide sufficient conditions in order for the solutions of a complementarity system do exist and are unique. Comparisons of several solution concepts are made, regularity of solutions is studied and characterizations of the initial states that yield regular solutions (in the sense that they do not contain impulses) are established.

If one considers the modeling process as a mapping which assigns models to physical systems, it is rather natural to ask whether this mapping is continuous or not. Stated differently, one may ask whether the modeling process associates *close* physical systems to *close* models. The second part of the thesis opens with an investigation on the behavior of the complementarity systems subject to small parameter changes. It is shown that the modeling process is continuous for a class of complementarity systems.

Simulation of complementarity systems is another source of motivation to consider approximations. Since they are members of the family of nonsmooth systems, the classical numerical methods cannot be indiscriminately applied to complementarity systems. We illustrate this fact by means of examples in which the well-known backward Euler method fails to approximate the actual solution when it is applied to general complementarity systems. We say that a numerical method is consistent if the approximating trajectories that are produced by the method converge to the actual ones. The last two chapters of the sec-

ond part are concerned with the consistency of time-stepping methods for complementarity systems. It is shown that under certain conditions (such as passivity of the underlying system) the consistency of the backward Euler time-stepping method for complementarity systems can be guaranteed.

Samenvatting

Het voornaamste object van studie in dit proefschrift is een klasse van stuksgewijs lineaire dynamische systemen die gerelateerd zijn zowel aan de systeemtheorie als aan de mathematische programmering. De systemen in deze klasse worden aangeduid met de term *complementariteitssystemen.* Het onderzoek met betrekking tot deze *niet-lineaire* en *niet-gladde* dynamische systemen dat in dit proefschrift wordt beschreven is toegespitst op twee thema's: *goedgesteldheid* en *benaderingen.*

Goedgesteldheid, opgevat als existentie en uniciteit van oplossingen, is belangrijk onder meer vanuit het oogpunt van modelvalidatie. Fysische systemen kunnen dikwijls als deterministisch worden opgevat, in de zin dat identieke omstandigheden steeds hetzelfde gedrag zullen teweegbrengen, en als regel wordt dan van een wiskundig model van een dergelijk systeem dezelfde eigenschap verwacht. De geldigheid van een model zou ernstig in twijfel worden getrokken als dan zou blijken dat er voor sommige beginvoorwaarden verschillende oplossingen mogelijk zijn. Het verifiëren van goedgesteldheid is daarom een wezenlijk onderdeel van modelformulering. Het eerste deel van het proefschrift heeft betrekking op goedgesteldheid. Voldoende voorwaarden worden gegeven waaronder de oplossingen van een complementariteitssysteem bestaan en eenduidig zijn bepaald door de beginvoorwaarden. Verschillende oplossingsconcepten worden met elkaar vergeleken; de mate van regulariteit van oplossingen wordt beschreven, en karakteriseringen worden gegeven van begintoestanden die leiden tot reguliere oplossingen (in de zin dat geen impulsen voorkomen).

Als we een modelleringsproces beschouwen als een afbeelding die modellen toevoegt aan fysische systemen, dan ligt het voor de hand te vragen of deze afbeelding continu is of niet. In andere woorden, de vraag is of *dichtbijzijnde* fysische systemen aanleiding geven tot *dichtbijzijnde* modellen. Het tweede deel van het proefschrift begint met een onderzoek van het gedrag van complementariteitssystemen die aan kleine parameterveranderingen worden onderworpen. Aangetoond wordt dat het modelleringsproces inderdaad continu is voor een klasse van complementariteitssystemen.

Een tweede motivering voor het onderzoek van benaderingen ligt in de simulatie van complementariteitssystemen. Aangezien deze systemen niet glad zijn kunnen klassieke numerieke methoden niet zonder meer worden toegepast. In het proefschrift wordt getoond dat bij sommige complementariteitssystemen de welbekende impliciete Euler-methode geen benadering oplevert van de werkelijke oplossing. We noemen een numerieke methode "consistent" als de benaderende trajekten die worden geproduceerd door de methode convergeren naar de werkelijke oplossing. De laatste twee hoofdstukken van het tweede deel hebben betrekking op de consistentie van zogenaamde time-stepping methoden voor complementariteitssystemen. Aangetoond wordt dat onder bepaalde voorwaarden (zoals passiviteit van het onderliggende systeem) de consistentie van de impliciete Euler time-stepping methode voor complementariteitssystemen kan worden gegarandeerd.

Center for Economic Research, Tilburg University, The Netherlands
Dissertation Series

No.	Author	Title	Published
1	P.J.J. Herings	Static and Dynamic Aspects of General Disequilibrium Theory; ISBN 90 5668 001 3	June 1995
2*	Erwin van der Krabben	Urban Dynamics: A Real Estate Perspective - An institutional analysis of the production of the built environment; ISBN 90 5170 390 2	August 1995
3	Arjan Lejour	Integrating or Desintegrating Welfare States? - a qualitative study to the consequences of economic integration on social insurance; ISBN 90 5668 003 X	September 1995
4	Bas J.M. Werker	Statistical Methods in Financial Econometrics; ISBN 90 5668 002 1	September 1995
5	Rudy Douven	Policy Coordination and Convergence in the EU; ISBN 90 5668 004 8	September 1995
6	Arie J.T.M. Weeren	Coordination in Hierarchical Control; ISBN 90 5668 006 4	September 1995
7	Herbert Hamers	Sequencing and Delivery Situations: a Game Theoretic Approach; ISBN 90 5668 005 6	September 1995
8	Annemarie ter Veer	Strategic Decision Making in Politics; ISBN 90 5668 007 2	October 1995
9	Zaifu Yang	Simplicial Fixed Point Algorithms and Applications; ISBN 90 5668 008 0	January 1996
10	William Verkooijen	Neural Networks in Economic Modelling - An Empirical Study; ISBN 90 5668 010 2	February 1996
11	Henny Romijn	Acquisition of Technological Capability in Small Firms in Developing Countries; ISBN 90 5668 009 9	March 1996
12	W.B. van den Hout	The Power-Series Algorithm - A Numerical Approach to Markov Processes; ISBN 90 5668 011 0	March 1996
13	Paul W.J. de Bijl	Essays in Industrial Organization and Management Strategy; ISBN 90 5668 012 9	April 1996

No.	Author	Title	Published
14	Martijn van de Ven	Intergenerational Redistribution in Representative Democracies; ISBN 90 5668 013 7	May 1996
15	Eline van der Heijden	Altruism, Fairness and Public Pensions: An Investigation of Survey and Experimental Data; ISBN 90 5668 014 5	May 1996
16	H.M. Webers	Competition in Spatial Location Models; ISBN 90 5668 015 3	June 1996
17	Jan Bouckaert	Essays in Competition with Product Differentiation and Bargaining in Markets; ISBN 90 5668 016 1	June 1996
18	Zafar Iqbal	Three-Gap Analysis of Structural Adjustment in Pakistan; ISBN 90 5668 017 X	September 1996
19	Jimmy Miller	A Treatise on Labour: A Matching-Model Analysis of Labour-Market Programmes; ISBN 90 5668 018 8	September 1996
20	Edwin van Dam	Graphs with Few Eigenvalues - An interplay between combinatorics and algebra; ISBN 90 5668 019 6	October 1996
21	Henk Oosterhout	Takeover Barriers: the good, the bad, and the ugly; ISBN 90 5668 020 X	November 1996
22	Jan Lemmen	Financial Integration in the European Union: Measurement and Determination; ISBN 90 5668 021 8	December 1996
23	Chris van Raalte	Market Formation and Market Selection; ISBN 90 5668 022 6	December 1996
24	Bas van Aarle	Essays on Monetary and Fiscal Policy Interaction: Applications to EMU and Eastern Europe; ISBN 90 5668 023 4	December 1996
25	Francis Y. Kumah	Common Stochastic Trends and Policy Shocks in the Open Economy: Empirical Essays in International Finance and Monetary Policy; ISBN 90 5668 024 2	May 1997
26	Erik Canton	Economic Growth and Business Cycles; ISBN 90 5668 025 0	September 1997

No.	Author	Title	Published
27	Jeroen Hoppenbrouwers	Conceptual Modeling and the Lexicon; ISBN 90 5668 027 7	October 1997
28	Paul Smit	Numerical Analysis of Eigenvalue Algorithms Based on Subspace Iterations; ISBN 90 5668 026 9	October 1997
29	Uri Gneezy	Essays in Behavioral Economics; ISBN 90 5668 028 5	October 1997
30	Erwin Charlier	Limited Dependent Variable Models for Panel Data; ISBN 90 5668 029 3	November 1997
31	Rob Euwals	Empirical Studies on Individual Labour Market Behaviour; ISBN 90 5668 030 7	December 1997
32	Anurag N. Banerjee	The Sensitivity of Estimates, Inferences, and Forecasts of Linear Models; ISBN 90 5668 031 5	December 1997
33	Frans A. de Roon	Essays on Testing for Spanning and on Modeling Futures Risk Premia; ISBN 90 5668 032 3	December 1997
34	Xiangzhu Han	Product Differentiation, Collusion and Standardization; ISBN 90 5668 033 1	January 1998
35	Marcel Das	On Income Expectations and Other Subjective Data: A Micro-Econometric Analysis; ISBN 90 5668 034 X	January 1998
36	Jeroen Suijs	Cooperative Decision Making in a Stochastic Environment; ISBN 90 5668 035 8	March 1998
37	Talitha Feenstra	Environmental Policy Instruments and International Rivalry: A Dynamic Analysis; ISBN 90 5668 036 6	May 1998
38	Jan Bouwens	The Use of Management Accounting Systems in Functionally Differentiated Organizations; ISBN 90 5668 037 4	June 1998
39	Stefan Hochguertel	Households' Portfolio Choices; ISBN 90 5668 038 2	June 1998
40	Henk van Houtum	The Development of Cross-Border Economic Relations; ISBN 90 5668 039 0	July 1998
41	Jorg Jansen	Service and Inventory Models subject to a Delay-Limit; ISBN 90 5668 040 4	September 1998

No.	Author	Title	Published
42	F.B.S.L.P. Janssen	Inventory Management Systems: control and information issues; ISBN 90 5668 041 2	September 1998
43	Henri L.F. de Groot	Economic Growth, Sectoral Structure and Unemployment; ISBN 90 5668 042 0	October 1998
44	Jenke R. ter Horst	Longitudinal Analysis of Mutual Fund Performance; ISBN 90 5668 043 9	November 1998
45	Marco Hoeberichts	The Design of Monetary Institutions; ISBN 90 5668 044 7	December 1998
46	Adriaan Kalwij	Household Consumption, Female Employment and Fertility Decisions: A microeconometric analysis; ISBN 90 5668 045 5	February 1999
47	Ursula Glunk	Realizing High Performance on Multiple Stakeholder Domains: A Resource-based Analysis of Professional Service Firms in the Netherlands and Germany; ISBN 90 5668 046 3	April 1999
48	Freek Vermeulen	Shifting Ground: Studies on the Intersection of Organizational Expansion, Internationalization, and Learning; ISBN 90 5668 047 1	February 1999
49	Haoran Pan	Competitive Pressures on Income Distribution in China; ISBN 90 5668 048 X	March 1999
50	Laurence van Lent	Incomplete Contracting Theory in Empirical Accounting Research; ISBN 90 5668 049 8	April 1999
51	Rob Aalbers	On the Implications of Thresholds for Economic Science and Environmental Policy; ISBN 90 5668 050 1	May 1999
52	Abe de Jong	An Empirical Analysis of Capital Structure Decisions in Dutch Firms; ISBN 90 5668 051 X	June 1999
53	Trea Aldershof	Female Labor Supply and Housing Decisions; ISBN 90 5668 052 8	June 1999
54	Jan Fidrmuc	The Political Economy of Reforms in Central and Eastern Europe; ISBN 90 5668 053 6	June 1999
55	Michael Kosfeld	Individual Decision-Making and Social Interaction; ISBN 90 5668 054 4	June 1999

No.	Author	Title	Published
56	Aldo de Moor	Empowering Communities - A method for the legitimate user-driven specification of network information systems; ISBN 90 5668 055 2	October 1999
57	Yohane A. Khamfula	Essays on Exchange Rate Policy in Developing Countries; ISBN 90 5668 056 0	September 1999
58	Maurice Koster	Cost Sharing in Production Situations and Network Exploitation; ISBN 90 5668 057 9	December 1999
59	Miguel Rosellón Cifuentes	Essays on Financial Policy, Liquidation Values and Product Markets; ISBN 90 5668 058 7	December 1999
60	Sharon Schalk	Equilibrium Theory: A salient approach; ISBN 90 5668 059 5	December 1999
61	Mark Voorneveld	Potential Games and Interactive Decisions with Multiple Criteria; ISBN 90 5668 060 9	December 1999
62	Edward Droste	Adaptive Behavior in Economic and Social Environments; ISBN 90 5668 061 7	December 1999
63	Jos Jansen	Essays on Incentives in Regulation and Innovation; ISBN 90 5668 062 5	January 2000
64	Franc J.G.M. Klaassen	Exchange Rates and their Effects on International Trade; ISBN 90 5668 063 3	January 2000
65	Radislav Semenov	Cross-country Differences in Economic Governance: Culture as a major explanatory factor; ISBN 90 5668 065 X	February 2000
66	Alexandre Possajennikov	Learning and Evolution in Games and Oligopoly Models; ISBN 90 5668 064 1	March 2000
67	Eric Gaury	Designing Pull Production Control Systems: Customization and Robustness; ISBN 90 5668 066 8	March 2000
68	Marco Slikker	Decision Making and Cooperation Restrictions; ISBN 90 5668 067 6	April 2000
69	Ruud Brekelmans	Stochastic Models in Risk Theory and Management Accounting; ISBN 90 5668 068 4	June 2000
70	Flip Klijn	A Game Theoretic Approach to Assignment Problems; ISBN 90 5668 069 2	June 2000

No.	Author	Title	Published
71	Maria Montero	Endogenous Coalition Formation and Bargaining; ISBN 90 5668 070 6	June 2000
72	Bas Donkers	Subjective Information in Economic Decision Making; ISBN 90 5668 071 4	June 2000
73	Richard Nahuis	Knowledge and Economic Growth; ISBN 90 5668 072 2	October 2000
74	Kuno Huisman	Technology Investment: A Game Theoretic Real Options Approach; ISBN 90 5668 073 0	October 2000
75	Gijsbert van Lomwel	Essays on Labour Economics; ISBN 90 5668 074 9	November 2000
76	Tina Girndt	Cultural Diversity and Work-Group Performance: Detecting the Rules; ISBN 90 5668 075 7	December 2000
77	Teye A. Marra	The Influence of Proprietary Disclosure Costs on the Decision to Go Public: Theory and Evidence; ISBN 90 5668 076 5	January 2001
78	Xiaodong Gong	Empirical Studies on the Labour Market and on Consumer Demand; ISBN 90 5668 078 1	February 2001
79	Kanat Camlibel	Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems; ISBN 90 5668 079 X	May 2001

* Copies can be ordered from Thela Thesis, Prinseneiland 305, 1013 LP Amsterdam, The Netherlands, phone: + 31 20 6255429; fax: +31 20 6203395; e-mail: office@thelathesis.nl

M. KANAT ÇAMLIBEL was born on May 17 of 1970 in Istanbul. Before coming to the Netherlands to pursue his Ph. D. degree, he was a teaching assistant at Istanbul Technical University where he got the B. Sc. and M. Sc. degrees in Control and Computer Engineering in 1991 and in 1994, respectively. In 1997, he started his Ph. D. studies under the supervision of Hans Schumacher.

The main object of this thesis is a class of piecewise linear dynamical systems that are related both to system theory and to mathematical programming. The dynamical systems in this class are known as complementarity systems. With regard to these nonlinear and nonsmooth dynamical systems, the research in the thesis concentrates on two themes: well-posedness and approximations. The well-posedness issue, in the sense of existence and uniqueness of solutions, is of considerable importance from a model validation point of view. In the thesis, sufficient conditions are established for the well-posedness of complementarity systems. Furthermore, an investigation is made of the convergence of approximations of these systems with an eye towards simulation.

ISBN: 90 5668 079 X