Dynamical analysis of linear passive networks with ideal diodes

Citation for published version (APA):

Document status and date:
Published: 01/01/2000

Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Dynamical Analysis of Linear Passive Networks with Ideal Diodes
Part II: Consistency of a Time-stepping Method

M.K. Çamhbel* W.P.M.H. Heemels† J.M. Schumacher‡

Abstract

In a companion paper, a definition has been proposed for what is to be understood by a "solution" of a linear network with ideal diodes, and conditions have been given under which such a solution exists and is unique. This framework provides the basis for a rigorous study of the convergence of numerical methods for computing transient solutions. In particular, we will deal with a time-stepping method based on the well-known backward Euler scheme. After demonstrating, by means of an example, that this widely applied time-stepping method does not necessarily produce useful output for arbitrary linear dynamical systems with ideal diode characteristics, the consistency of the method will be proven for linear passive circuits coupled with ideal diodes by showing that the approximations generated by the method converge to the true solution of the system in a suitable sense.

1 Introduction

Simulation of switching networks is a problem that has been studied extensively in circuit theory [1,2,6,13,14,18,23,30]. Roughly speaking, there are two main approaches, namely event-tracking (see e.g. [1,18]) and time-stepping methods (see [2,13,14,23] for electrical networks and [16,19,20,27,29] for unilaterally constrained mechanical systems with friction phenomena). Having a hybrid systems point of view (see for instance [26]), 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 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 paper to provide a rigorous basis for the use of time-stepping methods in the simulation of circuits with state events.

In the companion paper [8] (see also [4]) the meaning of a transient true solution to the dynamical network model with ideal diodes has already been established. Using techniques borrowed from the theory of linear complementarity systems (LCS) [9,10,15,24,25], existence and uniqueness of solutions have been proven under mild conditions. Moreover, several regularity properties have been shown from which this paper will benefit.

*Dept. of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, and Dept. of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. On leave from Dept. of Electrical Eng., Istanbul Technical University, 06560 Maslak Istanbul, Turkey. E-mail: K.Camhbel@tue.nl
†Dept. of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, and Dept. of Economics, Tilburg University, E-mail: w.p.m.h.heemels@tue.nl
‡Dept. of Econometrics, Tilburg University, P.O. Box 90153 5000 LE Tilburg, The Netherlands, E-mail: jms@kub.nl
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, 13, 14] for electrical networks. Similar methods have been used in a mechanical context in [16, 19, 20, 27, 29]. The advantage of the method is that it is straightforward to implement and many algorithms (e.g., Lemke’s algorithm [5], Katzenelson’s algorithm [12] and others [14]) are available to solve the one-step problems consisting of linear complementarity problems (LCPs).

In [13] the use of a time-stepping method based on the backward Euler scheme (or higher order linear multistep integration methods [7] 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 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 [27, 28]. 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 [27, 28]. The conditions used in [27, 28] 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 converge 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 paper is as follows. In section 2 preliminaries on linear complementarity systems and passivity are stated. The time-stepping method that will be studied is considered in section 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 4. The conclusions follow in section 5. The proofs of the main results can be found in section 6.

Throughout the paper, \( \mathbb{R} (\mathbb{R}^n) \) denotes the set of \((n\)-tuples of\) real numbers. \( \mathbb{R}_+ \) denotes the set of nonnegative real numbers, i.e., \( \mathbb{R}_+ = [0, \infty) \). For the real part of the complex number \( \lambda \), we write \( \text{Re} \lambda \). For any \( x, y \in \mathbb{R}^n \), \( x \perp y \) means that \( x^T y = 0 \). Inequalities for vectors are always meant to hold componentwise. The Euclidean and maximum norm of a vector \( x \in \mathbb{R}^n \) will be denoted by \( \|x\| := \sqrt{\sum_{i=1}^n x_i^2} \) and \( \|x\|_\infty := \max_{i \in \mathbb{N}} |x_i| \), respectively. Given an integer \( n \), \( \overline{n} \) denotes the set \( \{1, 2, \ldots, n\} \). For a real number \( r \in \mathbb{R} \), we use the notation \( \lfloor r \rfloor \) to denote the smallest integer larger than or equal to \( r \). The set of real matrices with \( n \) rows and \( m \) columns is denoted by \( \mathbb{R}^{n \times m} \). For any \( A \in \mathbb{R}^{n \times m} \), \( J \subseteq \overline{n} \), and \( K \subseteq \overline{m} \), \( A_{JK} \) denotes the submatrix \( \{A_{ij}\}_{i \in J, j \in K} \). If \( J = \overline{n} \) (\( K = \overline{m} \)), we also write \( A_{\bullet K} \) (\( A_{J \bullet} \)). For any \( A \in \mathbb{R}^{m \times m} \), \( \|A\| := \sup_{\|x\|=1} \|Ax\| \) denotes the matrix norm induced by the Euclidean vector norm. A square matrix \( A \in \mathbb{R}^{m \times m} \) is said to be nonnegative (positive) definite if \( x^T Ax \geq 0 \) (\( x^T Ax > 0 \)) for all \( 0 \neq x \in \mathbb{R}^m \). We write \( \sigma(A) \) for the set of eigenvalues of \( A \) and \( \rho(A) := \max_{\lambda \in \sigma(A)} |\lambda| \) for the spectral radius of \( A \). By the symmetric part of \( A \), we mean the matrix \( \frac{1}{2} (A + A^T) \). The identity matrix is denoted by \( I \). The set of \( n \)-tuples of square integrable functions on \((t_0, t_1)\) is denoted by \( L^2((t_0, t_1)) \). The notation \( (x, y) \) denotes the inner product of \( x, y \in L^2((t_0, t_1)) \), i.e., \( (x, y) = \int_{t_0}^{t_1} x(t)y(t)dt \). The norm on \( L^2((t_0, t_1)) \) is defined by \( \|x\| = \sqrt{(x, x)} \). Moreover, the time function \( x(t) \) denotes the restriction of the time function \( x \) to the interval \( \Omega \). We say that the sequence \( \{x_k\} \subset L^2((t_0, t_1)) \) converges (converges weakly) to \( x \) if \( \lim_{k \to \infty} \|x_k - x\| = 0 \).
A ∈ ℝ^{nx1}, B ∈ ℝ^{nxm} and C ∈ ℝ^{mxn} is said to be minimal, when rank \([B \ AB \cdots A^{n-1}B]\) = n and rank \([C^\top \ C^\top A^\top \cdots C^\top (A^\top)^{n-1}]\) = n.

2 Preliminaries

We begin by briefly recalling the linear complementarity problem (LCP) of mathematical programming. For an extensive survey on the problem, the reader is referred to [5].

**Problem 2.1** (LCP(\(q, M\))) Given \(q ∈ ℝ^n\) and \(M ∈ ℝ^{nxn}\), find \(z ∈ ℝ^n\) such that

\[
\begin{align*}
  z & ≥ 0 \\
  q + Mz & ≥ 0 \\
  z^\top (q + Mz) & = 0
\end{align*}
\]

We say that \(z\) solves LCP(\(q, M\)) if \(z\) satisfies (1). The set of all solutions of LCP(\(q, M\)) will be denoted by SOL(\(q, M\)). Note that the so-called complementarity (1) conditions also appear in the ideal diode characteristic \(v ≤ 0, i ≥ 0,\) and \(iv = 0\). Not surprisingly, the linear complementarity problem plays a major role in the analysis of networks with ideal diodes. Indeed, as discussed in [8], linear networks with ideal diodes can be modelled as linear complementarity systems (see [9, 10, 24, 25] for detailed discussion), which are dynamical versions of the linear complementarity problem. They are of the form

\[
\begin{align*}
  \dot{x}(t) &= Ax(t) + Bu(t) \\
  y(t) &= Cx(t) + Du(t) \\
  0 ≤ u(t) &⊥ y(t) ≥ 0
\end{align*}
\]

where \(u(t) ∈ ℝ^m\), \(x(t) ∈ ℝ^n\), \(y(t) ∈ ℝ^m\) and \(A, B, C,\) and \(D\) are matrices of appropriate dimensions. We denote (2) by LCS(\(A, B, C, D\)) and associate to (\(A, B, C, D\)) the transfer matrix \(G(s) = D + C(sI - A)^{-1}B\).

Before precisely defining the solution concept of LCS(\(A, B, C, D\)), we need to mention several spaces of functions and distributions, which play a crucial role in the sequel. The space \(B\) denotes the space of Bohl functions, i.e., functions having rational Laplace transforms. The space \(B_k\) consists of the distributions of the form \(u = u_{imp} + u_{reg}\), where \(u_{imp} = u_0δ\) is called the **impulsive part** with \(u_0 ∈ ℝ\) and \(u_{reg} ∈ B\) is called the **regular part**. A distribution \(u ∈ B_k\) is said to be **initially nonnegative**, if its Laplace transform \(\hat{u}(s)\) satisfies \(\hat{u}(σ) ≥ 0\) for all sufficiently large \(σ ∈ ℝ\). In a similar fashion, the space \(L_2(0, τ)\) consists of the distributions of the form \(u = u_{imp} + u_{reg}\) where \(u_{imp} = u_0δ\) is called the **impulsive part** with \(u_0 ∈ ℝ\) and \(u_{reg} ∈ L_2(0, τ)\) is called the **regular part**. We say that the sequence of distributions \(\{u_0δ + u_{reg}\} ⊂ L_2(0, τ)\) converges (weakly) to \(u_0δ + u_{reg}\), if \(\{u_0δ\}\) converges to \(u_0\) and \(\{u_{reg}\}\) converges (weakly) to \(u_{reg}\) in \(L_2\) sense.

Next, we recall the notion of **initial solution** which is of considerable importance in the analysis of linear complementarity systems.

**Definition 2.2** The triple \((u, x, y) ∈ B_k^{n+n+m}\) is an **initial solution** of LCS(\(A, B, C, D\)) with initial state \(x_0\) if there exists an index set \(I ⊆ m\) such that

\[
\begin{align*}
  \dot{x} &= Ax + Bu + x_0δ \\
  y &= Cx + Du \\
  u_i &= 0 \text{ if } i ∈ I \\
  y_i &= 0 \text{ if } i ∉ I
\end{align*}
\]

hold in the distributional sense (for more details see the companion paper [8]), and \(u\) and \(y\) are initially nonnegative.
It can be shown 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 (see for instance [10, 19]).

**Problem 2.3** (RCP\((x_0, A, B, C, D)\)) Given \(x_0 \in \mathbb{R}^n\) and \((A, B, C, D)\) with \(A \in \mathbb{R}^{nxn}\), \(B \in \mathbb{R}^{nxm}\), \(C \in \mathbb{R}^{mxn}\) and \(D \in \mathbb{R}^{mxm}\), 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 \(\sigma \in \mathbb{R}\).

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 2.4** [10] The triple \((u, x, y)\) 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\((A, B, C, D)\). Actually, the (global) solution concept for general linear complementarity systems (see [9]) is more complicated than the one we will present. In the case of linear passive complementarity systems, it can be trimmed as shown in the companion paper [8].

**Definition 2.5** \((u, x, y) \in L^{n+m+n+m}_t(0, \tau)\) is a (global) solution of LCS\((A, B, C, D)\) on \([0, \tau]\) with initial state \(x_0\) if the following conditions hold.

1. There exists an initial solution \((\tilde{u}, \tilde{x}, \tilde{y})\) such that

\[
(u_{imp}, x_{imp}, y_{imp}) = (\tilde{u}_{imp}, \tilde{x}_{imp}, \tilde{y}_{imp}).
\]

2. The equations

\[
\dot{x} = Ax + Bu + x_0\delta
\]

\[
y = Cx + Du
\]

hold in the distributional sense.

3. For almost all \(t \in [0, \tau]\), \(0 \leq u_{reg}(t) \perp y_{reg}(t) \geq 0\).

Notice that the above definition is equivalent to the integral form given in the companion paper [8, Definition 5.1].

The first item in the Definition 2.5 imposes a relation between the impulsive part and the rest of the solution. In the following example, we illustrate the necessity of such a connection.

**Example 2.6** Consider the simple circuit shown in Figure 1. By denoting the voltage across the capacitor and the diode by \(v_c\) and \(v_d\), respectively and the current through the diode by \(i_d\), one can obtain the circuit equations as follows:

\[
\dot{v}_c = -i_d
\]

\[
v_d = v_c
\]

\[
0 \geq v_d \perp i_d \geq 0.
\]
They can be rewritten in the form of a linear complementarity system

\[
\begin{align*}
\dot{x} &= u \\
y &= x \\
0 &\leq u \perp y \geq 0
\end{align*}
\]

with the definitions \( u = i_d, x = -i_v, \) and \( y = -v_d \). For the initial state \( x_0 = -1 \), the triple \( (u, x, y) = (a\delta, a - 1, a - 1) \) with \( a \geq 1 \) satisfies the last two items of the Definition 2.5. However, \( (a\delta, a - 1, a - 1) \) is only a solution for initial state \( x_0 = -1 \), if \( a = 1 \), since this is the only solution complying with the circuit from a physical point of view. Its interpretation is that there is an instantaneous discharge of the capacitor. Note that \( (u, x, y) = (\delta, 0, 0) \) is indeed the unique initial solution.

In the sequel, we are mainly concerned with linear passive complementarity systems. To be reasonably self-contained, 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 2.7** [31] The system \((A, B, C, D)\) given by (2a)-(2b) is said to be passive (dissipative with respect to the supply rate \( u^T y \)) if there exists a function \( V : \mathbb{R}^n \to \mathbb{R}_+ \) (a storage function), such that

\[
V(x(t_0)) + \int_{t_0}^{t_1} u^T(t)y(t)dt \geq V(x(t_1))
\]

holds for all \( t_0 \) and \( t_1 \) with \( t_1 \geq t_0 \), and all \((u, x, y) \in L_2^{m+n+m}(t_0, t_1)\) satisfying (2a)-(2b).

We state a well-known theorem on passive systems which is sometimes called the positive real lemma.

**Lemma 2.8** [31] Assume that \((A, B, C)\) is minimal. Then the following statements are equivalent:

1. \((A, B, C, D)\) is passive.
2. The matrix inequalities

\[
K = K^T \geq 0 \quad \text{and} \quad \begin{bmatrix}
A^T K + KA & KB - C^T \\
B^T K - C & -(D + D^T)
\end{bmatrix} \leq 0
\]

have a solution.
3. \( G(s) \) is positive real, i.e., \( G(\lambda) + G^T(\lambda) \geq 0 \) for all \( \lambda \in \mathbb{C} \) with \( \lambda \notin \sigma(A) \) and \( \text{Re} \lambda > 0 \).

Moreover, if \((A, B, C, D)\) 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 2.9** \((A, B, C)\) is a minimal representation and \( B \) is of full column rank.

The proof of the following theorem can be found in the companion paper [8] and deals with the existence and uniqueness of solutions to linear passive complementarity systems.

**Theorem 2.10** Suppose that \((A, B, C, D)\) is such that Assumption 2.9 holds and \((A, B, C, D)\) is passive. Let \( \tau > 0 \) be given. For each \( x_0 \), there exists a unique solution \((u, x, y) \in L_2^{m+n+m}(0, \tau)\) of \( LCS(A, B, C, D) \) on \([0, \tau]\) with initial state \( x_0 \).
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 [2,13,14, 23]. 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 \( u_{k+1}^h \), \( y_{k+1}^h \), and \( x_{k+1}^h \) given \( x_k^h \) through the linear complementarity problem given by

\[
\begin{align*}
\frac{x_{k+1}^h - x_k^h}{h} &= Ax_{k+1}^h + Bu_{k+1}^h \quad \text{(4a)} \\
y_{k+1}^h &= Cx_{k+1}^h + Du_{k+1}^h \quad \text{(4b)} \\
0 &\leq y_{k+1}^h \perp u_{k+1}^h \geq 0. \quad \text{(4c)}
\end{align*}
\]

Here \( x_k^h \) denotes the value at the \( k \)th 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 3.1 (\( \{u_k^h\}, \{x_k^h\}, \{y_k^h\} \) = Approximation(\( A, B, C, D, \tau, h, x_0 \))

1. \( N_h = \lceil \frac{\tau}{h} \rceil \)
2. \( x_{-1}^h := x_0 \)
3. \( k := -1 \)
4. solve the one-step problem

\[
y_{k+1}^h = (I - hA)^{-1}x_k^h + [D + hC(I - hA)^{-1}B]u_{k+1}^h \\
0 \leq u_{k+1}^h \perp y_{k+1}^h \geq 0
\]
5. \( x_{k+1}^h := (I - hA)^{-1}x_k^h + h(I - hA)^{-1}Bu_{k+1}^h \)
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 3.2 For all sufficiently small \( h > 0 \), \( \text{LCP}(C(I - hA)^{-1}x, G(h^{-1})) \) has a unique solution for all \( x \), where \( G(h^{-1}) \) is given by \( D + hC(I - hA)^{-1}B \).

This assumption implies that for all sufficiently small \( h > 0 \), Algorithm 3.1 generates an output, which is unique. Hence, for a given step size \( h > 0 \) (sufficiently small), we can define the approximations \((u^h, x^h, y^h) \in \mathcal{C}(0, \tau)\) given by

\[
\begin{align*}
u_{\text{imp}}^h &= hu_{\text{imp}}^h \delta \\
x_{\text{imp}}^h &= hx_{\text{imp}}^h \delta \\
y_{\text{imp}}^h &= hy_{\text{imp}}^h \delta \\
u_{\text{reg}}^h(t) &= u_{\text{imp}}^h \quad \text{whenever } (l - 1)h \leq t < lh, \\
x_{\text{reg}}^h(t) &= x_{\text{imp}}^h \\
y_{\text{reg}}^h(t) &= y_{\text{imp}}^h
\end{align*}
\]

\( \text{whenever } (l - 1)h \leq t < lh, \)
Figure 2: Nonconvergence of backward Euler approximations for the triple integrator with diode.

where $u^h_k$, $x^h_k$, and $y^h_k$, $k = 0, 1, \ldots, N_h$ have been obtained from Algorithm 3.1. One of the main goals of the paper is to prove that for a passive system these 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 3.1 is not always consistent even if Assumption 3.2 holds.

**Example 3.3** Consider the linear complementarity system (consisting of a triple integrator with complementarity conditions)

$$
\begin{align*}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 \\
\dot{x}_3 &= u \\
y &= x_1 \\
0 &\leq u \perp y \geq 0
\end{align*}
$$

with the initial state $x_0 = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix}^T$. As we have already mentioned before, Definition 2.5 is a simplified version of the general one given in [9] 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 $(u, x, y) = (\delta, 0, 0)$, which does not qualify as a solution in the sense of Definition 2.5, is the 'true' global solution of the system with the given initial state. Here $\delta$ denotes the distributional derivative of the Dirac impulse $\delta$. Algorithm 3.1 gives

$$
(u^h_k, y^h_k) = \begin{cases} 
(h^{-2}, 0) & \text{if } k = 0 \\
(0, \frac{k(k+1)}{2} h) & \text{if } k \neq 0.
\end{cases}
$$

It follows from (5d) that

$$
\|y^h_{\text{reg}}\| \geq \int_{(N_h-2)h}^{(N_h-1)h} \|y^h_{(N_h-1)}\|^2 \, dt)^{1/2} = \frac{(N_h - 1)N_h}{2} h^{3/2} = O(h^{-1/2})
$$

whenever $N_h \geq 2$. Therefore, $y^h_{\text{reg}}$ is far from being convergent as it is not bounded as $h$ converges to zero. For three different values of $h$, the trajectories of $y^h_{\text{reg}}$ on $[0, 1]$ are depicted in Figure 2.

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. The following theorem states conditions that imply consistency.
Theorem 3.4 Consider \( \text{LCS}(A,B,C,D) \) such that Assumption 3.2 holds. Let \( \tau > 0 \) and \( x_0 \in \mathbb{R}^n \) be given. Also let \((u^h, x^h, y^h)\) be given by (5) via Algorithm 3.1. Suppose that there exists \( \alpha > 0 \) such that for all sufficiently small \( h \)
\[
\|h u^h_0\| \leq \alpha \quad \text{and} \quad \|u^h_0\| \leq \alpha.
\]
Then, we have the following statements:

1. There exists a unique initial solution of \( \text{LCS}(A,B,C,D) \) with initial state \( x_0 \) in the sense of Definition 2.2 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, \((u^h_{\text{imp}}, x^h_{\text{imp}}, y^h_{\text{imp}})\) converges to \((u_{\text{imp}}, 0, y_{\text{imp}})\) where \((u_{\text{imp}}, y_{\text{imp}}) = (u_0\delta, y_0\delta)\).

3. Let \( \{h_k\} \) converge to zero. Suppose that \( D \) is nonnegative definite. Then the following holds.
   
   (a) There exists a subsequence \( \{h_{k_i}\} \subseteq \{h_k\} \) such that \((\{u^{h_{k_i}}\}, \{y^{h_{k_i}}\})\) converges weakly to some \((u,y)\) and \((x^{h_{k_i}})\) converges weakly to some \(x\).

   (b) \((u,x,y)\) is a solution of \( \text{LCS}(A,B,C,D) \) on \([0,T]\) with the initial state \( x_0 \).

   (c) If the solution \((u,x,y)\) is unique for the initial state \( x_0 \) in the sense of Definition 2.5, then the complete sequence \((\{u^{h_k}\}, \{y^{h_k}\})\) converges weakly to \((u,y)\) and \((x^{h_k})\) converges to \(x\).

Proof See section 6. □

4 Main results for passive LCS

In section 6, we shall show that the conditions of Theorem 3.4 are satisfied in the case of passive complementarity systems so that the following result holds.

Theorem 4.1 Consider \( \text{LCS}(A,B,C,D) \) such that Assumption 2.9 holds and \((A,B,C,D)\) is passive. Let \( \tau > 0 \) and \( x_0 \in \mathbb{R}^n \) be given. Let \((u,x,y)\) be the solution of \( \text{LCS}(A,B,C,D) \) on \([0,T]\) with the initial state \( x_0 \). Also let \((u^h,x^h,y^h)\) be given by (5) via Algorithm 3.1. Then, \((\{u^h\}, \{y^h\})\) converges weakly to \((u,y)\) and \((x^h)\) converges to \(x\) as the step size \( h \) tends to zero.

Proof See section 6. □

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 \( T : \mathbb{R}^m \to \mathcal{H} \) 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 (weakly converges) to \( Tx^* \) where \( x^* = \lim_{k \to \infty} x_k \).

Theorem 4.2 Consider the \( \text{LCS}(A,B,C,D) \) such that Assumption 2.9 holds and \((A,B,C,D)\) is passive. Let \( \tau > 0 \) be given. Define the operators \( x_0 \mapsto (u,y) \) and \( x_0 \mapsto x \), where \((u,x,y)\) is the solution of \( \text{LCS}(A,B,C,D) \) on \([0,T]\) with the initial state \( x_0 \). The operators \( x_0 \mapsto (u,y) \) and \( x_0 \mapsto x \) are weakly continuous and continuous, respectively.

Proof See section 6. □

Note that Theorem 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.
5 Conclusions

In this paper, 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 paper 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 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.

6.1 Preliminaries

For ease of reference, we recall some standard results on weakly convergent sequences.

Lemma 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) $\{(v_k, w_k)\}$ 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 [11], 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 \( l \)
\[
|\langle v_{k_l}, w \rangle| > \epsilon. \tag{6}
\]
for a given \( \cdot \). 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). \( \blacksquare \)

We recall the notion of a compact operator for ease of reference.

**Definition 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 6.3** Let \( A \in \mathbb{R}^{n \times n} \). The following statements hold:

1. \[ ||(I - hA)^{-1}|| \leq \frac{1}{1 - \lambda h} \text{ for all } h \text{ with } \lambda h < 1 \text{ where } \lambda \text{ is the largest eigenvalue of } \frac{1}{2}(A + A^T). \]
2. There exists an \( \alpha > 0 \) such that \[ ||(I - hA)^{-1}|| \leq \alpha \text{ for all sufficiently small } h. \]
3. If \( \{v_k, h_k\} \) converges to \( t \) then \( \{(I - h_kA)^{-1}\} \) converges to \( e^{At} \). Moreover, the convergence is uniform in \( t \) on any bounded interval.

**Proof**: 1: By the Wazewski inequality (see e.g. [32, Theorem 8.1]), \( ||e^{At}|| \leq e^{\lambda t} \) for all \( t \) where \( \lambda \) is the largest eigenvalue of \( \frac{1}{2}(A + A^T) \). Theorem 1.5.3 in [21] gives now the desired inequality.

2: It can easily be verified by using item 1 that
\[ ||(I - hA)^{-1}|| \leq \frac{1}{1 - \beta} \]
whenever \( \lambda h \leq \beta < 1 \).

3: This follows from [21, Theorem 3.5.3]. \( \blacksquare \)

### 6.2 Proof of Theorem 3.4 items 1 and 2

For proving Theorem 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 2.5 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 6.4** Consider the matrices \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{m \times n} \), \( C \in \mathbb{R}^{m \times n} \) and \( D \in \mathbb{R}^{m \times m} \) such that Assumption 3.2 holds. We have the following statements for all \( x_0 \in \mathbb{R}^n \).

1. \( \text{RCP}(x_0, A, B, C, D) \) has a unique solution.
2. For all sufficiently small \( h \),
\[
\begin{align*}
\hat{u}(h^{-1}) &= hu_0^h \\
\hat{x}(h^{-1}) &= hx_0^h \\
\hat{y}(h^{-1}) &= hy_0^h
\end{align*}
\]
where \( (\hat{u}(s), \hat{y}(s)) \) is the solution of \( \text{RCP}(x_0, A, B, C, D) \), \( \hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s) \) and \( (u_0^h, x_0^h, y_0^h) \) is the solution of the one-step problem of Algorithm 3.1 for \( k = 0 \).
Proof 1: Observe the basic fact that if \( LCP(q, M) \) is solvable then \( LCP(aq, M) \) is also solvable provided that \( \alpha \geq 0 \). As a consequence, Assumption 3.2 implies together with the identity \( h(I - hA)^{-1} = (h^{-1}I - hA)^{-1} \) 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 \( \alpha z \) is a solution of \( LCP(aq, M) \) 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 \( h \)

\[
\begin{align*}
\hat{u}(h^{-1}) &= hu^0_h \\
\hat{x}(h^{-1}) &= hx^0_h \\
\hat{y}(h^{-1}) &= hy^0_h
\end{align*}
\]

where \( \hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s) \).

Proof of Theorem 3.4 items 1 and 2: From Proposition 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 \( ||hu^0_h|| \) is bounded for sufficiently small \( h \) by the hypothesis of the theorem, \( \hat{u}(s) \) is proper due to Proposition 6.4 item 2. It follows that \( \hat{x}(s) \) is strictly proper and \( \hat{y}(s) \) is proper. Clearly, Proposition 2.4 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 (5) and Proposition 6.4 item 2 that \((u_{\text{imp}}, x_{\text{imp}}, y_{\text{imp}})\) converges to \((u_{\text{imp}}, 0, y_{\text{imp}})\) as \( h \) tends to zero.

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 \( L_2(0, \tau) \). More details on the TCP can be found in [3] and the references therein.

Problem 6.5 (TCP\((q, T)\)) Given \( q \in L_2^{\text{mu}}(0, \tau) \) and \( T : L_2^{\text{mu}}(0, \tau) \to L_2^{\text{mu}}(0, \tau) \), find \( z \in L_2^{\text{mu}}(0, \tau) \) such that

\[
\begin{align*}
z(t) &\geq 0 \\
qu(t) + (Tz)(t) &\geq 0 \\
\langle z, q + Tz \rangle &= 0.
\end{align*}
\]

If \( z \) satisfies (8), we say that \( z \) solves TCP\((q, T)\).

Note that the conditions given in item 3 of Definition 2.5 may be equivalently written as

\[
\begin{align*}
u_{\text{reg}}(t) &\geq 0 \\
y_{\text{reg}}(t) &\geq 0
\end{align*}
\]
for almost all \( t \in [0, \tau] \) and
\[
\langle u_{\text{reg}}, y_{\text{reg}} \rangle = 0.
\] (9c)

Hence, by associating the operator \( T_{i(A,B,C,D)} \) defined by
\[
(T_{i(A,B,C,D)} u)(t) = Du(t) + \int_0^t Ce^{A(t-s)} Bu(s) \, ds
\]
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** The following statements hold.

1. If \((u, x, y) \in L^2_0([0, \tau])\) is a solution of LCS\((A, B, C, D)\) on \([0, \tau]\) with initial state \( x_0 \), then \( u_{\text{reg}} \) is a solution of TCP\((Ce^A x_0^+|[0, \tau], T_{i(A,B,C,D)})\), where \( x_0^+ = x_0 + Bu_0 \) and \( u_{\text{imp}} = u_{\text{reg}} \).

2. If \( u \in L^m_0([0, \tau]) \) is a solution of TCP\((Ce^A x_0^+|[0, \tau], T_{i(A,B,C,D)})\), then \((u, x, y)\) is a solution of LCS\((A, B, C, D)\) on \([0, \tau]\) with initial state \( x_0 \) where
\[
x = e^A x_0^+|[0, \tau] + T_{i(A,B,C,D)} u
\]
\[
y = C x + Du.
\]

### 6.4 The time-stepping method in a TCP formulation

The approximations of (5) by the backward Euler time-stepping scheme can also be formulated as the solutions of certain TCPs. To do so, we introduce the operators \( \tilde{C}_h : \mathbb{R}^{mN_h} \to \mathbb{R}^{mN_h} \), \( \tilde{D}_h : \mathbb{R}^{mN_h} \to \mathbb{R}^{mN_h} \), \( R_h : L^2_0([0, \tau]) \to \mathbb{R}^{mN_h} \), \( P_h^l : \mathbb{R}^{mN_h} \to \mathbb{R}^{mN_h} \), and \( P_h : \mathbb{R}^{mN_h} \to \mathbb{R}^{mN_h} \) for given \( \tau > 0 \) and \( h \) with \( N_h = \lceil \tau/h \rceil \).

\[
\tilde{C}_h := \begin{bmatrix} C & 0 & \cdots & 0 \\ 0 & C & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C \end{bmatrix}, \quad \tilde{D}_h := \begin{bmatrix} D & 0 & \cdots & 0 \\ 0 & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D \end{bmatrix}, \quad R_h u := \begin{bmatrix} \int_0^h u(s) \, ds \\ \int_h^\tau u(s) \, ds \\ \vdots \end{bmatrix}
\]

\[
Q_h := \begin{bmatrix} (I - hA)^{-1} B & 0 & \cdots & 0 \\ (I - hA)^{-2} B & (I - hA)^{-1} B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (I - hA)^{-N_h} B & (I - hA)^{-N_h+1} B & \cdots & (I - hA)^{-1} B \end{bmatrix}
\]

\[
(P_h^m w)(t) := w_{[(l-1)h]} \text{ if } t \in [(l-1)h, lh) \text{ for } l = 1, 2, \ldots, N_h.
\]

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 6.7** Let \( v, w \in \mathbb{R}^{mN_h} \) and \( x \in \mathbb{R}^{nN_h} \). The following statements hold.

1. \( R_h P_h^m v = v \).
2. \( v \geq 0 \) if and only if \( P_h^m v(t) \geq 0 \) for (almost) all \( t \in [0, \tau] \).
3. \( (P_h^m v, P_h^m w) = h v^T w \).
4. \( DP_h^m v = P_h^m \tilde{D}_h v \).
5. \( CP_h^m x = P_h^m \tilde{C}_h x \).
Proof  Evident from the definitions of $P^h$, $R_h$, $C_h$ and $D_h$. ■

It can be easily seen that $\bar{u}_h$ solves \( \text{LCP}(\bar{C}_h\bar{q}_h, \bar{D}_h + \bar{C}_hQ_h) \), where
\[
\bar{u}_h = \begin{bmatrix} u_1^h \\ u_2^h \\ \vdots \\ u_N^h \end{bmatrix}, \quad \text{and} \quad \bar{q}_h = \begin{bmatrix} (I - hA)^{-1}x_0^h \\ (I - hA)^{-2}x_0^h \\ \vdots \\ (I - hA)^{-N_h}x_0^h \end{bmatrix}.
\]

Indeed, \( \text{LCP}(\bar{C}_h\bar{q}_h, \bar{D}_h + \bar{C}_hQ_h) \) is pieced together from $N_h$ one-step problems of Algorithm 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 6.9 below.

Lemma 6.8  Let $T_h^* = P_h^*Q_hR_h$ and $q_h^* = P_h^*\bar{q}_h$. The following statements hold.

1. For all sufficiently small $h$, $u_{\text{reg}}^h$ solves \( \text{TCP}(Cq_h^*, D + CT_h^*) \).

2. $\{q_h^*(\cdot)\}$ converges to $e^A(x_0 + Bu_0)$ with $u_0$ as in item 2 of Theorem 3.4 as $h$ tends to zero.

3. $\{T_h^*u_{\text{reg}}^h - T(A,B,T,I,0)u_{\text{reg}}^h\}$ converges to 0 as $h$ tends to zero.

Proof  1: Since $\bar{u}_h$ solves \( \text{LCP}(\bar{C}_h\bar{q}_h, \bar{D}_h + \bar{C}_hQ_h) \), we have
\[
\bar{u}_h \geq 0, \\
\bar{y}_h := \bar{C}_h\bar{q}_h + (\bar{D}_h + \bar{C}_hQ_h)\bar{u}_h \geq 0, \\
\bar{u}_h^T\bar{y}_h = 0.
\]

Note that $u_{\text{reg}}^h = P_h^m\bar{u}_h$ and $y_{\text{reg}}^h = P_h^m\bar{y}_h$ due to (5) and the definition of $P_h^m$. Hence, (11a) and (11b) together with Proposition 6.7 item 2 imply that
\[
(u_{\text{reg}}^h(t)) \geq 0 \quad \text{and} \quad (y_{\text{reg}}^h(t)) \geq 0 \quad \text{for (almost) all} \quad t \in [0, \tau].
\]

Moreover, we have
\[
(u_{\text{reg}}^h, y_{\text{reg}}^h) = (P_h^m\bar{u}_h, P_h^m\bar{y}_h) = h\bar{u}_h^T\bar{y}_h = 0
\]
from Proposition 6.7 item 3, and (11c). On the other hand, we have
\[
y_{\text{reg}}^h = P_h^m\bar{y}_h = P_h^m[(\bar{C}_h\bar{q}_h + (\bar{D}_h + \bar{C}_hQ_h)\bar{u}_h] \quad \text{(from (11b))}
\]
\[
= CP_h^m\bar{q}_h + (D + CT_h^*)u_{\text{reg}}^h \quad \text{(from items 4 and 5 of Proposition 6.7)}
\]
\[
= Cq_h^* + (D + CT_h^*)u_{\text{reg}}^h \quad \text{(from item 1 of Proposition 6.7).} \tag{14}
\]

Clearly, (12), (13) and (14) imply that $u_{\text{reg}}^h$ solves \( \text{TCP}(Cq_h^*, D + CT_h^*) \).

2: Note that from Algorithm 3.1 step 5 we have
\[
x_0^h := (I - hA)^{-1}x_0 + h(I - hA)^{-1}Bu_0^h. \tag{15}
\]

Let $\hat{u}(s)$ be the solution of \( \text{RCP}(x_0, A, B, C, D) \) and $u_0 = \lim_{s \to \infty} \hat{u}(s)$. As shown in the proof of Theorem 3.4 item 2, $hu_0^h$ converges to $u_0$ as $h$ tends to zero. Then, (15) implies that
\[
\{x_0^h\} \text{ converges to } x_0 + Bu_0 \tag{16}
\]
as $h$ tends to zero. Note that
\[
q_h^*(t) = (I - hA)^{-\lfloor t/h \rfloor}x_0^h.
\]

13
Hence, from the triangle inequality we get
\[
\|q^h(t) - e^A(x_0 + Bu_0)\| \leq \|(I - hA)^{-1/2}x_0^h - e^A x_0^h\| + \\
+\|e^A x_0^h - e^A(x_0 + Bu_0)\| \\
\leq (\int_0^T \|(I - hA)^{-[1/h]} - e^A t\| dt)^{1/2}\|x_0^h\| + \\
+\|e^A\|\|t\|^{1/2}\|x_0^h - (x_0 + Bu_0)\|.
\]

Since \{\lfloor t/h \rfloor h \} converges to \( t \) as \( h \) tends to zero, Lemma 6.3 item 3 and (16) reveal that the right hand side converges to zero.

3: Note that
\[
(T_h u^h_{\text{tr}})(t) = \sum_{p=1}^l h(I - hA)^{-(l-p+1)} Bu_p^h \\
= \sum_{p=1}^l \int_{(p-1)h}^{ph} (I - hA)^{-(l-p+1)} Bu_p^h ds
\]
and also that
\[
(T_{(A,B,Z,O)} u^h_{\text{tr}})(t) = \sum_{p=1}^{l-1} \int_{(p-1)h}^{ph} e^{A(s-t)} Bu_p^h ds + \int_{(l-1)h}^t e^{A(t-s)} Bu_p^h ds
\]
for \( l = \lfloor t/h \rfloor \). By exploiting the triangle inequality, we get
\[
\|(T_h u^h_{\text{tr}})(t) - (T_{(A,B,Z,O)} u^h_{\text{tr}})(t)\| \leq \\
\sum_{p=1}^{\lfloor t/h \rfloor} \int_{(p-1)h}^{ph} \|(I - hA)^{-(l/h) - 1} - e^{A(t-s)}\||Bu_p^h||ds \tag{17}
\]
since \((p-1)h < s \leq ph\) gives \( p = \lfloor s/h \rfloor \). Clearly, \{\lfloor t/h \rfloor - \lfloor s/h \rfloor\} converges to \( t - s \) as \( h \) tends to zero. We already know from the hypothesis that \( \|u_p^h\| \) is bounded for \( p \neq 0 \). Therefore, from Lemma 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 \( T_h u^h_{\text{tr}} - T_{(A,B,Z,O)} u^h_{\text{tr}} \) converges in \( L^2 \) sense to zero as \( h \) tends to zero.

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.

**Theorem 6.9** Let \( T : L^\infty_2(0, \tau) \to L^\infty_2(0, \tau) \) be a compact operator and let \( S : L^\infty_2(0, \tau) \to L^\infty_2(0, \tau) \) be a linear continuous nonnegative definite (i.e. \( \langle v, Sv \rangle \geq 0 \) for all \( v \in L^\infty_2(0, \tau) \)) operator. Suppose that there exist sequences \( \{q_k\} \) and \( \{T_k\} \) such that \( \{q_k\} \) converges to \( q \) and TCP\((q_k, S + T_k)\) is solvable for all \( k \). Let \( z_k \) be a solution of TCP\((q_k, S + T_k)\). If \( \{z_k\} \) converges weakly to \( z \) 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 \( z \), which is the weak limit of \( \{z_k\} \), satisfies
\[
z(t) \geq 0 \tag{18a}
\]
\[
q(t) + ((S + T)z)(t) \geq 0 \tag{18b}
\]
for almost all $t \in [0, \tau]$ and
\[ \langle z, q + (S + T)z \rangle = 0. \] (18c)

Since $z_k$ solves TCP($q_k, S + T_k$), we have
\[ z_k(t) \geq 0 \] (19a)
\[ q_k(t) + ((S + T_k)z_k)(t) \geq 0 \] (19b)
for almost all $t \in [0, \tau]$ and
\[ \langle z_k, q_k + (S + T_k)z_k \rangle = 0 \] (19c)
for all $k$. Let $\mathcal{K}$ be the nonnegative cone of $L^0_{\alpha}(0, \tau)$, i.e.,
\[ \mathcal{K} = \{ v \mid v(t) \geq 0 \text{ for almost all } t \in [0, \tau] \}. \]

Note that $\mathcal{K}$ is weakly closed (i.e., the weak limit of every weakly converging sequence in $\mathcal{K}$ is in $\mathcal{K}$) by Theorem 3.12 of [22]. Then, (18a) follows from (19a) and the fact that $\mathcal{K}$ is weakly closed. Lemma 6.1 item 3b and Definition 6.2 imply that
\[ \{ S z_k \} \text{ converges weakly to } S z \] (20a)
and
\[ \{ T z_k \} \text{ converges to } T z. \] (20b)

As a consequence of (20b), we have
\[ \{ T_k z_k \} \text{ converges to } T z \] (20c)
since $\{ T_k z_k - T z_k \}$ converges to zero by assumption. The equations (20a), (20c) and the convergence of $\{ q_k \}$ imply that $\{ q_k + (S + T_k)z_k \}$ converges weakly to $q + (S + T)z$. Hence, (18b) follows from (19b) and the fact that $\mathcal{K}$ is weakly closed. Now, it remains to show that (18c) holds. Equation (19c) gives
\[ \langle z_k, S z_k \rangle = -\langle z_k, q_k + T_k z_k \rangle. \]
The convergence of $\{ q_k \}$ and the weak convergence of $\{ z_k \}$, together with (20c) and Lemma 6.1 item 3c, imply that
\[ \lim_{k \to \infty} \langle z_k, S z_k \rangle = -\lim_{k \to \infty} \langle z_k, q_k + T_k z_k \rangle = -\langle z, q + T z \rangle. \]
We also have from (18a) and (18b) that
\[ \langle z, q + (S + T)z \rangle \geq 0. \]
Thus,
\[ \langle z, S z \rangle \geq -\langle z, q + T z \rangle = \lim_{k \to \infty} \langle z_k, S z_k \rangle. \] (21)
The nonnegative definiteness of $S$ implies
\[ \langle z_k - z, S(z_k - z) \rangle \geq 0. \] (22)
Since $\lim_{k \to \infty} \langle z, S z_k \rangle = \lim_{k \to \infty} \langle z_k, S z \rangle = \langle z, S z \rangle$ due to the fact that $\{ z_k \}$ converges weakly to $z$ and Lemma 6.1 items 3b and 3c, we get
\[ \lim_{k \to \infty} \langle z_k, S z_k \rangle \geq \langle z, S z \rangle \] (23)
by letting $k$ tend to infinity in (22). Together with (21), this yields
\[ \lim_{k \to \infty} \langle z_k, S z_k \rangle = \langle z, S z \rangle. \] (24)
Combining (20c), (24), the convergence of $\{ q_k \}$ to $q$ and Lemma 6.1 item 3c results in
\[ \lim_{k \to \infty} \langle z_k, q_k + (S + T_k)z_k \rangle = \langle z, q + (S + T)z \rangle. \] (25)
Finally, (18c) follows from (19c) and (25). ■
6.6 Completing the proof of Theorem 3.4

The proofs of item 1 and 2 in Theorem 3.4 have already been shown. The remaining items will be proven in this subsection.

**Proof of item 3 of Theorem 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 $\|u_{reg}\|$ is bounded for sufficiently small $h$. According to Lemma 6.1 item 1, the existence of a weakly convergent subsequence of $\{u_{reg}^{h_k}\}$, say $\{u_{reg}^{h_k}\}$, is clear. Let $u_{reg}$ denote the weak limit of this subsequence, and also let $q_{h_k}$ and $T_{h_k}'$ be defined as in Lemma 6.8. Since $T_{(A,B,I,0)}$ is a compact operator (see e.g. [22, Exercise 4.15]), it follows from Definition 6.2 that $\{T_{(A,B,I,0)}u_{reg}\}$ converges (strongly) to $T_{(A,B,I,0)}u_{reg}$. Then, Lemma 6.8 item 3 implies that

$$\{T_{h_k}'u_{reg}^{h_k}\} \text{ converges to } T_{(A,B,I,0)}u_{reg}.$$  

Note that

$$x_{reg}^{h_k} = q_{h_k} + T_{h_k}'u_{reg}^{h_k} \quad (27a)$$

and

$$y_{reg}^{h_k} = Cq_{h_k} + (D + CT_{h_k}')u_{reg}^{h_k}.$$  

It is clear from Lemma 6.8 item 2, (26) and (27a) that $\{x_{reg}^{h_k}\}$ converges to $x_{reg} := e^A(x_0 + Bu_0)|_{[0,\tau]} + T_{(A,B,I,0)}u_{reg}$. Since $\{Du_{reg}^{h_k}\}$ converges weakly to $Du_{reg}$ due to Lemma 6.1 item 3b, it follows from Lemma 6.8 item 2, (26) and (27b) that $\{y_{reg}^{h_k}\}$ converges weakly to $y_{reg} := Ce^A(x_0 + Bu_0)|_{[0,\tau]} + T_{(A,B,C,D)}u_{reg}$.

3b: Item 2 of Theorem 3.4 (see also the proof) states the convergence of $(u_{imp}^{h_k}, y_{imp}^{h_k}, \tilde{y}_{imp}^{h_k})$ to

$$(u_{imp}, 0, y_{imp}) = (u_0\delta, 0, y_0\delta) = (\bar{u}_{imp}, \bar{x}_{imp}, \bar{y}_{imp}),$$

where $(\bar{u}, \bar{x}, \bar{y}) \in B_h^{m+n+m}$ is the unique initial solution for initial state $x_0$. Hence, we also have

$$y_{imp} = Du_{imp}$$

due to $x_{imp} = 0$. Let us define in the framework of Theorem 6.9

- $T = T_{(A,B,C,0)}$,
- $S = D$,
- $q_l = Cq_{h_k}$,
- $T_l = CT_{h_k}'$.

It can be checked that

- $T$ is compact ( [22, Exercise 4.15]),
- $S$ is nonnegative definite (by the hypothesis $D \geq 0$),
- $\{q_l\}$ converges to $Ce^A(x_0 + Bu_0)|_{[0,\tau]}$ (from Lemma 6.8 item 2),
- TCP($q_l, S + T_l$) is solvable for all sufficiently large $l$ (from Lemma 6.8 item 1), and
- $\{T_l u_{reg}^{h_k} - T u_{reg}^{h_k}\}$ converges to zero (from Lemma 6.8 item 3).
Then, Theorem 6.3 implies that $u_{\text{reg}}$ solves \( TCP(Ce^{A}(x_0 + Bu_0)T_{(A,B,C,D)}). \) Due to Proposition 6.6 item 2, \((u_{\text{reg}}, x_{\text{reg}}, y_{\text{reg}})\) is a solution of \( LCS(A, B, C, D) \) on \([0, \tau]\) with the initial state \( x_0 + Bu_0 \) (with \( u_0 \) as in (28)), where

\[
x_{\text{reg}} = e^{A}(x_0 + Bu_0) + T_{(A,B,T,0)}u_{\text{reg}} \]
\[
y_{\text{reg}} = Cx_{\text{reg}} + Du_{\text{reg}}.
\]

Equivalently,

\[
\dot{x}_{\text{reg}} = Ax_{\text{reg}} + Bu_{\text{reg}} + (x_0 + Bu_0)\delta \quad (29a)
\]
\[
\dot{y}_{\text{reg}} = Cx_{\text{reg}} + Du_{\text{reg}} \quad (29b)
\]

holds in the distributional sense and

\[
0 \leq u_{\text{reg}}(t) \perp y_{\text{reg}}(t) \geq 0 \quad (29c)
\]

for almost all \( t \in [0, \tau] \). Since \( u_{\text{imp}} = u_0\delta, y_{\text{imp}} = Du_{\text{imp}} \) and \( x_{\text{imp}} = 0 \), (29a) and (29b) yield

\[
\dot{x}_{\text{reg}} = Ax + Bu + x_0\delta \quad (30a)
\]
\[
y = Cx + Du \quad (30b)
\]

Clearly, (28), (29c) and (30) imply that \((u, x, y)\) 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 \((u_{\text{imp}}, x_{\text{imp}}, y_{\text{imp}})\) converges. Note that the sequence of regular parts \((u_{\text{reg}}, x_{\text{reg}}, y_{\text{reg}})\) is bounded by assumption. Moreover, following the proof of item 3a above, it is clear that every converging subsequence \((u_{\text{reg}}, x_{\text{reg}}, y_{\text{reg}})\) 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 6.1 item 2 completes the proof.

### 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 4.1) for linear passive complementarity systems.

**Proposition 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., \( \frac{1}{2}(M + M^T) \).

**Proof** By Lemma 7.3.10 and Proposition 5.10.10 in [5], we have

\[
\|z_1 - z_2\| \leq \frac{n}{\mu(M)}\|q_1 - q_2\| \quad (31)
\]

Since \( \|z\| \leq n^{1/2}\|z\|_\infty \) and \( \|z\|_\infty \leq \|z\| \) for all \( z \in \mathbb{R}^n \), (31) yields

\[
\|z_1 - z_2\| \leq \frac{n^{3/2}}{\mu(M)}\|q_1 - q_2\|.
\]

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 6.10.
Lemma 6.11 Consider the matrices \( A \in \mathbb{R}^{m \times n}, \ B \in \mathbb{R}^{m \times m}, \ C \in \mathbb{R}^{m \times n} \) and \( D \in \mathbb{R}^{m \times m} \) such that Assumption 2.9 holds and \((A,B,C,D)\) 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 \( u^T Du = 0 \) imply that \( u^T C Bu > 0 \).
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 is clear from Lemma 2.8 item 2.

2: Let \( V(x) = \frac{1}{2} x^T K x \) be a storage function for the passive system \((A,B,C,D)\). Assume that \( u \neq 0 \) and \( u^T Du = 0 \). Then, Lemma 2.4 of [8] implies that \( C^T u = K Bu \). Hence, we get

\[
u^T C Bu = u^T B^T C^T u = u^T B^T K Bu > 0
\]

since \( K \) is positive definite and \( B \) has full column rank.

3: Note that \( a_1 + a_2 h \geq b_1 + b_2 h \) for all sufficiently small \( h > 0 \) if and only if \((a_1 > b_1) \) or \((a_1 = b_1 \) and \( a_2 \geq b_2)\). Let \( \alpha = \min_{\|u\|=1} u^T C Bu \). Since

\[
u^T D u \geq 0 \) or \( u^T D u = 0 \) and \( u^T C Bu \geq \alpha\)

holds for all \( u \) with \( \|u\| = 1 \) due to the items 1 and 2, it is clear that for all \( u \) with \( \|u\| = 1 \)

\[
u^T D u + h u^T C Bu \geq \alpha h \quad \text{for all sufficiently small } h > 0.
\]

Thus, we get from [17, Property 5.2.2.1 (Rayleigh-Ritz theorem)]

\[
\mu(D + hCB) = \min_{\|u\|=1} u^T(D + hCB)u \\
\geq \alpha h
\]

for all sufficiently small \( h > 0 \). Since \( \alpha > 0 \) according to item 2, the proof of this part is complete.

4: It is known from matrix theory (see e.g. [17, 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

\[
\mu(G(h^{-1})) \geq \mu(D + hCB) + h^2 \mu(CA(I - hA)^{-1}B) \\
\geq \beta h \quad \text{(from item 3)}
\]

for some \( \beta > 0 \) and all sufficiently small \( h \). \( \blacksquare \)

The following auxiliary lemma will be needed in the sequel.

Lemma 6.12 Let \( \mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \geq b\} \) be a given nonempty polyhedron with \( A \in \mathbb{R}^{m \times n} \) and \( b \in \mathbb{R}^m \) and let \( x^* \) be equal to \( \arg \min_{x \in \mathcal{P}} \|x\| \). There exists an index set \( I \subseteq \mathbb{N} \) such that \( x^* = \arg \min_{A x = b, x \in \mathcal{P}} \|x\| \).

Proof Consider the convex quadratic optimization problem

\[
\min_{Ax \geq b} \frac{1}{2} x^T x.
\]
The well-known Kuhn-Tucker conditions are necessary and sufficient for this problem because of its convexity (see for instance [5, 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

\[
x^* = A^T u
\]

\[
Ax^* \geq b
\]

\[
u \geq 0
\]

\[
u^T (Ax^* - b) = 0.
\]

Take such a vector $u$. Let $I = \{i \mid u_i > 0\}$ and $v = u_I$. Then, $x^*$ satisfies

\[
x^* = (A_J^*)^T v
\]

\[
A_J x^* = b_J.
\]

Note that (32) are necessary and sufficient (Kuhn-Tucker) conditions for the convex quadratic minimization problem

\[
\min_{A_J x^* = b_J} \frac{1}{2} x^T x.
\]

To formulate the next lemma, we need to recall the concept of a dual cone.

**Definition 6.13** For any nonempty set $Q \subset \mathbb{R}^n$, the set

\[
\{w \in \mathbb{R}^n \mid w^T v \geq 0 \text{ for all } v \in Q\}
\]

is called the **dual cone** of $Q$ and is denoted by $Q^*$.

**Lemma 6.14** Let $M \in \mathbb{R}^{nxn}$ be nonnegative definite and $Q = SOL(0, M)$. We have the following statements.

1. $LCP(q, M)$ is solvable if and only if $q \in Q^*$.

2. For each $q \in Q^*$, 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 $q \in Q^*$

\[
\|z^*(q)\| \leq \alpha \|q\|
\]

where $z^*(q)$ denotes the least-norm solution (see item 2) of $LCP(q, M)$.

**Proof**

1: It follows from [5, Exercise 3.12.1 and Corollary 3.8.10].

2: This follows from the fact that $SOL(q, M)$ is a nonempty polyhedron whenever $q \in Q^*$ [5, Theorem 3.1.7(c)].

3: Define

\[
\alpha(A) = \begin{cases} 
0 & \text{if } A = 0 \\
\max_{y \in \text{im } A} \min_{Ax = y} \|x\| & \text{if } A \neq 0
\end{cases}
\]

Note that

\[
\max_{\|y\| = 1} \min_{Ax = y} \|x\| = \max_{\|Ax\| = 1} \min_{A^{x'} = 0} \|x - x'\|.
\]
The mapping \( x \mapsto \min_{x' : \|A x'\| = 1} \|x - x'\| \) achieves its maximum on the set \( \{ x : \|A x\| = 1 \} \). Hence, the quantity \( \alpha(A) \) is well-defined for all \( A \). Take

\[
\alpha = \sqrt{2} \max_{l \leq n} \max_{j \leq n} \alpha \left( \begin{bmatrix} l & 0 \\ -M_{l*} \\ -M_{j*} \end{bmatrix} \right).
\]

For any \( q \in Q^* \), 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 \( I = \{ i : z^*_i(q) > 0 \} \). Clearly, \( P = \{ v : v_{j*} = 0, q_I + M_I v_I = 0, \text{ and } q_{j*} + M_{j*} v_{j*} \supseteq SOL(q, M) \} \) and \( z^*(q) \in P \). Note that \( P \) is a polyhedron, since

\[
P = \{ v : v_I = 0, v_{j*} = 0, (I-I)x_I = 0, (j-I)v_{j*} = 0 \}.
\]

Moreover, it is obvious that \( z^*(q) = \arg\min_{v_{j*} = 0} \|v\| \). Then, according to Lemma 6.12 there exists \( J \subseteq 3n \) such that \( z^*(q) = \arg\min_{A_{j*} = b_j} \|v\| \). Thus, we have \( \|z^*(q)\| \leq \alpha(A_{j*}) \|b_j\| \). Note that

\[
\|b_j\|^2 \leq \|b\|^2 \leq \|q\|^2 + \|q_I\|^2 \leq 2\|q\|^2 \text{ and } \sqrt{2\alpha(A_{j*})} \leq \alpha.
\]

Consequently,

\[
\|z^*(q)\| \leq \alpha \|q\|.
\]

### 6.8 Proof of Theorem 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 3.4 are fulfilled for this class of linear complementarity systems.

**Lemma 6.15** Consider LCS(A, B, C, D) such that Assumption 2.9 holds and \( (A, B, C, D) \) is passive. Then, for all sufficiently small \( h, LCP(hC(I - hA)^{-1}\tilde{x}, G(h^{-1})) \) has a unique solution for each \( \tilde{x} \in \mathbb{R}^n \).

**Proof**

The statement follows from the positive definiteness of \( G(h^{-1}) \) for all sufficiently small \( h \) (Lemma 6.11 item 4 together with Theorem 3.1.6 of [5]).

**Lemma 6.16** Consider LCS(A, B, C, D) such that Assumption 2.9 holds and \( (A, B, C, D) \) is passive. Let \( T > 0 \) and \( Q = SOL(0, D_I) \), i.e., \( Q = \{ x \in \mathbb{R}^n : x \geq 0, D_I x \geq 0 \text{ and } x^T D_I x = 0 \} \) be given. Also let \((\{x_k^h\}, \{x_k^l\}, \{y_k^h\})\) be produced by Algorithm 3.1. The following statements hold for all sufficiently small \( h \).

1. \( Cx_k^h \in Q^* \) for all \( k \neq -1 \).

2. There exists an \( \alpha > 0 \) independent of \( x_0 \) such that \( \|u_k^h\| \leq \alpha \|x_0\| \) for all \( k \neq 0 \).

**Proof**

1: It is evident from (4b) and (4c) that \( u_k^h \) solves LCP(\( Cx_k^h, D \)) when \( k \neq -1 \). Since \( D \) is nonnegative definite (Lemma 6.11 item 1), \( Cx_k^h \in Q^* \) due to [5, Corollary 3.8.10].

2: All inequalities involving \( h \) are meant to hold for all sufficiently small \( h \), and \( \alpha_1, \alpha_2, \ldots, \alpha_6 \) are suitably chosen positive constants in this proof. Note that LCP(\( Cx_k^h, D \)) is solvable for all \( k \neq -1 \) due to item 1 and [5, Corollary 3.8.10]. Let \( u^* \) be the least-norm solution of LCP(\( Cx_k^h, D \)). Clearly, \( u^* \) solves also LCP\((C, hC(I - hA)^{-1}Bu^*, G(h^{-1}))\). According to Proposition 6.10, we have

\[
\|u_{k+1}^h - u^*\| \leq \frac{m^{1/2}}{\mu(G(h^{-1}))} \|C(I - hA)^{-1}x_k^h - Cx_k^h + hC(I - hA)^{-1}Bu^*\|,
\]

20
since \( u_{k+1}^h \) solves \( \text{LCP}(C(I - hA)^{-1}x_k^h, G(h^{-1})) \) and \( G(h^{-1}) > 0 \) for all sufficiently small \( h \). By using the triangle inequality and Lemma 6.11 item 4, we obtain

\[
\|u_{k+1}^h - u^*\| \leq \frac{\alpha_1}{h} \|C((I - hA)^{-1} - I)x_k^h\| + \alpha_1 \|C(I - hA)^{-1}Bu^*\|.
\]

Note that \((I - hA)^{-1} - I = hA(I - hA)^{-1}\). It can be easily verified that Lemma 6.3 item 2 and Lemma 6.14 item 3 result in

\[
\|u_{k+1}^h - u^*\| \leq \alpha_2 \|x_k^h\|.
\]

Consequently, we get

\[
\|u_k^h\| \leq \|u^*\| + \|u_{k+1}^h - u^*\| \leq \alpha_3 \|x_k^h\|
\]

by applying the triangle inequality and employing Lemma 6.14 item 3 and (33). It follows that

\[
\|x_{k+1}^h\| \leq \|x_k^h\| + \|u_{k+1}^h - x_k^h\|
\]

\[
\leq \|x_k^h\| + \|((I - hA)^{-1} - I)x_k^h + h(I - hA)^{-1}Bu_{k+1}^h\| \quad \text{(from (4a))}
\]

\[
\leq (1 + \alpha_4 h)\|x_k^h\| \quad \text{(from Lemma 6.3 item 2)}
\]

(35)

Since \( \lim_{h \to 0} (1 + \alpha_4 h)^N_h = e^{\alpha_4 \tau} \) (Lemma 6.3 item 3), (35) implies that

\[
\|x_k^h\| \leq \alpha_5 \|x_0^h\|
\]

(36)

for some \( \alpha_5 > 0 \). Here \( N_h = [\frac{\tau}{h}] \). Note that we have

\[
\|x_0^h\| = \|x_{-1}^h + hBu_0^h\|
\]

\[
= \|z_0 + hBu_0^h\| \leq \alpha_6 \|z_0\|
\]

(37)

from Proposition 6.3 item 2. Finally, (34), (36) and (37) establish the desired inequality. ■

After all these preliminaries, we can prove Theorem 4.1.

**Proof of Theorem 4.1** According to Lemma 6.15, Assumption 3.2 holds. Then, Proposition 6.4 item 1 implies that \( \text{RCP}(z_0, A, B, C, D) \) has a unique solution, say \((u(s), y(s))\). It is known from [8, Theorem 3.6] that \( u(s) \) is proper. Therefore, boundedness of \( \|u_{k+1}^h\| \) for all sufficiently small \( h \) follows from Proposition 6.4 item 2. On the other hand, \( D \) is nonnegative definite due to item 1 of Lemma 6.11 and

\[
\|u_{reg}^h\| = \left( \int_0^\tau \|u_{reg}(t)\| \, dt \right)^{1/2} \leq \alpha_2^{1/2} \|x_0\|
\]

(38)

due to (5) and Lemma 6.16 item 2. Finally, it is known from Theorem 2.10 that \((u, x, y)\) is the unique solution on \([0, \tau]\) with the initial state \( z_0 \). As a consequence of Theorem 3.4 item 3c, \( \{(u^{h_k}, y^{h_k})\} \) converges weakly to \((u, y)\) and \( \{x^{h_k}\} \) converges to \( x \) for any sequence \( \{h_k\} \) that converges to zero. In other words, \( \{(u^{h_k}, y^{h_k})\} \) converges weakly to \((u, y)\) and \( \{x^{h_k}\} \) converges to \( x \) as \( h \) tends to zero. ■

6.9 **Proof of Theorem 4.2**

In this subsection, the continuous dependence of solution trajectories on the initial states will be proven as formulated in Theorem 4.2.

**Proof of Theorem 4.2** Let the sequence \( \{\tilde{z}_k\} \subset \mathbb{R}^n \) converge to \( \tilde{x} \in \mathbb{R}^n \). Denote the solution of \( \text{LCS}(A, B, C, D) \) on \([0, \tau]\) with the initial states \( \tilde{z}_k \) and \( \tilde{x} \) by \((u^{k}, x^k, y^k)\) and \((u, x, y)\), respectively. Then, it should be shown that

1. \( \{(u^{h_k}, x^{h_k}, y^{h_k})\} \) converges to \((u_{imp}, x_{imp}, y_{imp})\).
2. \( \{ (u_{\text{reg}}^k, y_{\text{reg}}^k) \} \) converges weakly to \( (u_{\text{reg}}, y_{\text{reg}}) \) and \( \{ x_{\text{reg}}^k \} \) converges to \( x_{\text{reg}} \).

1. Let \( (u_{\text{imp}}^k, x_{\text{imp}}^k, y_{\text{imp}}^k) = (u_0^k, x_0^k, y_0^k) \). Also let \( u_0^k(h) \text{ and } u_0(h) \) be the solutions of the one-step problems \( \text{LCP}(C(I-hA)^{-1} \tilde{x}_k, hC(I-hA)^{-1} B + D) \) and \( \text{LCP}(C(I-hA)^{-1} \tilde{x}, hC(I-hA)^{-1} B + D) \), respectively. From Proposition 6.10 and Lemma 6.11 item 4, we get
\[
\| u_0^k(h) - u_0(h) \| \leq \frac{\alpha}{h} \| C(I-hA)^{-1} \| \| \tilde{x}_k - \tilde{x} \|
\]
for sufficiently small \( h \). By multiplying the inequality above by \( h \) and using Lemma 6.3 item 2, we obtain
\[
\| hu_0^k(h) - hu_0(h) \| \leq \alpha' \| \tilde{x}_k - \tilde{x} \|
\]
(39)
for sufficiently small \( h \). On the other hand, it is already known from the proof of Theorem 3.4 item 2 that \( \lim_{h \to 0} hu_0^k(h) = u_0^k \) and \( \lim_{h \to 0} hu_0(h) = u_0 \). Thus, (39) yields
\[
\| u_0^k - u_0 \| \leq \alpha' \| \tilde{x}_k - \tilde{x} \|
\]
(40)
Clearly, \( \{ u_0^k \} \) converges to \( u_0 \). Consequently, \( \{ u_{\text{imp}}^k \} \) converges to \( u_{\text{imp}} \). Since \( x_{\text{imp}}^k = 0 \) and \( y_{\text{imp}}^k = Du_{\text{imp}}^k \), we can conclude that \( \{ (u_{\text{imp}}^k, x_{\text{imp}}^k, y_{\text{imp}}^k) \} \) converges to \( (u_{\text{imp}}, x_{\text{imp}}, y_{\text{imp}}) \).

2. Observe that \( (u_{\text{reg}}^k, x_{\text{reg}}^k, y_{\text{reg}}^k) \) and \( (u_{\text{reg}}^k, x_{\text{reg}}^k, y_{\text{reg}}^k) \) are the solutions of \( \text{LCS}(A, B, C, D) \) on \([0, r]\) with the initial states \( \tilde{x}_k + Bu_0^k \) and \( \tilde{x} + Bu_0 \) respectively. Moreover, \( \tilde{x}_k + Bu_0^k \) converges to \( \tilde{x} + Bu_0 \) as shown in the proof of item 1 above. Lemma 6.16 item 2 together with (38) implies that for some \( \beta > 0 \) independent of \( \tilde{x}_k + Bu_0^k \), \( \| u_{\text{reg}}^k \| \leq \beta \| \tilde{x}_k + Bu_0^k \| \) for all \( k \). This means that the sequence \( \{ u_{\text{reg}}^k \} \) is bounded since the sequence \( \{ \tilde{x}_k + Bu_0^k \} \) is convergent. Hence, there exists at least one weakly convergent subsequence of \( \{ u_{\text{reg}}^k \} \) according to Lemma 6.1 item 3a. Take any such subsequence of \( \{ u_{\text{reg}}^k \} \), say \( \{ u_{\text{reg}}^k \} \).

Define
- \( T = T(A, B, C, D) \),
- \( S = D \),
- \( q_l = Ce^A (\tilde{x}_k + Bu_0^k) \), and
- \( T_l = T \).

It can be checked that
- \( T \) is compact ( [22, Exercise 4.15]),
- \( S \) is nonnegative definite (by Lemma 6.11 item 1),
- \( \{ q_l \} \) converges to \( Ce^A (\tilde{x} + Bu_0) \) on \([0, r]\) (since \( \| q_l - Ce^A (\tilde{x} + Bu_0) \| \leq \| Ce^A \| \| (\tilde{x}_k + Bu_0^k) - (\tilde{x} + Bu_0) \| \))
- TCP(\( q_l, S + T_l \)) is solvable (from Proposition 6.6 item 1), and
- \( \{ T_l u_{\text{reg}}^k - Tu_{\text{reg}}^k \} = 0 \).

Therefore, \( \{ u_{\text{reg}}^k \} \) converges weakly to the solution \( u_{\text{reg}} \) of \( TCP(Ce^A (\tilde{x} + Bu_0) \) on \([0, r], T(A, B, C, D)) \) according to Theorem 6.9. Since \( u_{\text{reg}} \) is unique due to Proposition 6.6 item 2 and Theorem 2.10, the reasoning above shows that any weakly convergent subsequence of \( \{ u_{\text{reg}}^k \} \) has the same limit. Lemma 6.1 item 2 implies now that the whole sequence \( \{ u_{\text{reg}}^k \} \) converges weakly to \( u_{\text{reg}} \). Note that Proposition 6.6 item 2 and uniqueness of the solutions of \( LCS(A, B, C, D) \) yield that
\[
x_{\text{reg}}^k = Ae^A (\tilde{x}_k + Bu_0^k) + T_l(A, B, C, D) u_{\text{reg}}^k
\]
(41a)
\[
y_{\text{reg}}^k = Ce^A (\tilde{x}_k + Bu_0^k) + Du_{\text{reg}}^k
\]
(41b)
and

\[ x_{\text{reg}} = e^{At}(\bar{x} + Bu_0)|_{0,T} + T_{(A,B,I,0)}u_{\text{reg}} \]  
\[ y_{\text{reg}} = Cx_{\text{reg}} + Du_{\text{reg}} \]  

Then, convergence of \( \{x^k_{\text{reg}}\} \) to \( x_{\text{reg}} \) and weak convergence of \( \{y^k_{\text{reg}}\} \) to \( y_{\text{reg}} \) follow from (41), the convergence of \( \{\bar{x}_k + Bu_k\} \) to \( \bar{x} + Bu_0 \) and the compactness of \( T_{(A,B,I,0)} \).

Acknowledgement

We thank NWO (Dutch Organization for Scientific Research) and The Scientific and Technical Research Council of Turkey (TÜBİTAK) for making the participation of M.K. Çamlibel in this study possible. The Co-operation Centre Tilburg and Eindhoven Universities (SOBU) is acknowledged for making the participation of W.P.M.H. Heemels possible.

References


