

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## The properties of differential-algebraic equations representing optimal control problems

### Journal Item

How to cite:

England, Roland; Gómez, Susana and Lamour, René (2009). The properties of differential-algebraic equations representing optimal control problems. *Applied Numerical Mathematics*, 59(10) pp. 2357–2373.

For guidance on citations see [FAQs](#).

© 2009 IMACS

Version: [\[not recorded\]](#)

Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.1016/j.apnum.2009.04.003>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

## Accepted Manuscript

The properties of differential-algebraic equations representing optimal control problems

Roland England, Susana Gómez, René Lamour

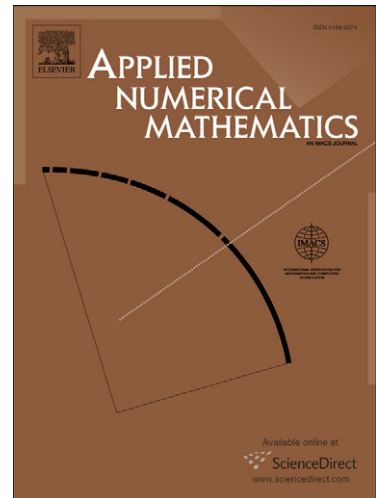
PII: S0168-9274(09)00054-3  
DOI: [10.1016/j.apnum.2009.04.003](https://doi.org/10.1016/j.apnum.2009.04.003)  
Reference: APNUM 2262

To appear in: *Applied Numerical Mathematics*

Received date: 9 June 2006  
Revised date: 17 March 2009  
Accepted date: 10 April 2009

Please cite this article as: R. England, S. Gómez, R. Lamour, The properties of differential-algebraic equations representing optimal control problems, *Applied Numerical Mathematics* (2009), doi: 10.1016/j.apnum.2009.04.003

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# The properties of differential-algebraic equations representing optimal control problems

Roland ENGLAND<sup>1</sup>  
Susana GÓMEZ<sup>2</sup>  
René LAMOUR<sup>3</sup>

## Abstract

A procedure is described for transforming a general optimal control problem to a system of Differential-Algebraic Equations (DAEs). The Kuhn-Tucker conditions consist of differential equations, complementarity conditions and corresponding inequalities. The latter are converted to equalities by adding a new variable combining the slack variable and the corresponding Lagrange multiplier.

We investigate the properties of the resulting DAEs. The index of a system of DAEs determines the well-conditioning of the problem. The concept of the *tractability index* is used to investigate the index in a systematic way, and during this process, it indicates which components of the system of equations must be differentiated to reduce the index. For an index-3 problem, the index is reduced without increasing the number of equations, and a numerical procedure is used to determine the index.

In the examples used here, the DAEs can be solved analytically. The examples are tested by the numerical determination of the index, and the results confirm the previously known properties of these examples.

The reformulation proposed here, as well as the index determination, might be used in the future, to develop a methodology to solve optimal control problems.

**Keywords:** Optimal control; Differential-algebraic equations; Tractability index; Differentiation index; Kuhn-Tucker conditions

## 1 Introduction

The purpose of this paper is to express an optimal control problem in terms of a system of Differential-Algebraic Equations (DAEs) and to investigate their properties. This system is

---

<sup>1</sup>Dept. of Mathematics, The Open University, Milton Keynes, MK7 6AA Great Britain.  
E-mail, r.england@open.ac.uk

<sup>2</sup>IIMAS, National University of Mexico, Apdo. Postal 20-726, Mexico D.F..  
E-mail, susanag@servidor.unam.mx

<sup>3</sup>Humboldt-University of Berlin, I. of Mathematics, D-10099, Berlin, Germany.  
E-mail, lamour@math.hu-berlin.de

obtained using calculus of variations to get the Kuhn-Tucker conditions. The inequalities associated with the complementarity conditions are converted to equalities by the addition of a new variable, combining the slack variable and the corresponding Lagrange multiplier. The sign of this variable indicates whether the constraint is active or not.

The well-conditioning of the problem can be expressed in terms of the index of the resulting system of DAEs, which is a measure of the difficulty involved in obtaining a numerical solution. The concept of the *tractability index* is introduced as a general purpose way of determining the index even when some components are not sufficiently differentiable. A projector related to the *tractability index* makes it possible, in the case of higher index, to determine exactly which components of the system of equations must be differentiated in order to reduce the index. This is important in the formulation of boundary value problems (BVP) arising from optimal control problems, because there is as yet no general purpose code to solve directly higher index BVPs (index  $> 2$ ) and for numerical reasons, it seems unlikely that a practical method will be available in the near future.

Methods based on the concept of the differentiation index transform the DAE to a system of index 1 or index 0 (ODE) by the differentiation of the equations involving algebraic components. However, it is not always clear which and how many equations should be differentiated. The process that we are presenting here clearly indicates which components of the system of equations must be differentiated to reduce the index. Also, this is the first time that a methodology is proved to deal with an index reduction of a nonlinear index-3 problem, without increasing the number of equations as in the work done in [10].

As has been stressed before, the purpose of this paper is to study the properties of optimal control problems, through their transformation to DAEs. The *tractability index* concept is applied to the class of DAEs thus obtained, and so provides a more complete analysis than that given in our report [6], which contained a partial analysis of three specific examples. We hope that the reformulation developed here, as well as our study of the *tractability index*, can be used in the future as an approach for the numerical solution of optimal control problems and as an alternative to the usual constrained optimization formulation, such as in [1, 2, 4, 5, 9, 15, 19, 20, 21, 22].

The examples used here are the minimization of the time to travel a fixed distance, subject to bounds on the acceleration and on the velocity, and the maximization of the yield of a component on a packed bed reactor. These problems have index varying from 1 to 3 and the theoretical investigation of the index shows the potential advantage of the concept of the *tractability index*, but also the difficulties, which necessitate a robust numerical index monitor, such as the one we present here.

In Section 2 we give an outline of the methodology to transform an optimal control problem into a system of DAEs (for a detailed presentation see [6, 7]). In Section 3, the *tractability index* is introduced and applied to the DAEs obtained before. Theorem 3.2 is a generalization of a theorem given in [6].

In Section 4, we present three examples which are transformed to DAEs, and their properties

as well as their indices are determined. In Section 5, we present a numerical method to obtain the index, even in cases when it is not possible to do so analytically. We give some conclusions in Section 6.

## 2 General transformation process

### 2.1 Formulation of an optimal control problem

Consider an optimal control problem, expressed as a dynamical system of ordinary differential equations subject to a number of initial and terminal conditions, and to a number of inequalities on the state and control variables, and with some unknown constant parameters. The objective function has the form of an integral of some function of the same state and control variables and parameters.

$$\text{minimize } J(\mathbf{u}) = \int_0^b h(\mathbf{y}, \mathbf{u}, \mathbf{c}) ds \quad (1)$$

$$\text{subject to: } \mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{u}, \mathbf{c}), \quad y_i(0) = y_{i0} \quad (i \in \mathcal{I}), \quad y_j(b) = y_{j1} \quad (j \in \mathcal{F}), \quad (2)$$

$$\mathbf{0} \leq \mathbf{g}(\mathbf{y}, \mathbf{u}, \mathbf{c}). \quad (3)$$

Here,  $\mathcal{I}$  and  $\mathcal{F}$  are subsets of the indices  $i$  of the state variables  $y_i$  for which initial and terminal values, respectively, are specified.

### 2.2 Calculus of variations

As we wish to transform this problem to a system of DAEs, we use the variational formulation to obtain the first-order necessary conditions. Most of this derivation has been presented elsewhere ([6, 7]), but the outline is given here for completeness.

Introducing small perturbations  $\delta \mathbf{y}(s)$ ,  $\delta \mathbf{u}(s)$ ,  $\delta \mathbf{c}$  constant, and the Lagrange multipliers  $\mathbf{v}(s)$  for the differential equations (equality constraints), and  $\mathbf{w}(s)$  for the inequality constraints, then the perturbation of the objective function  $J(\mathbf{u})$  is given by

$$\begin{aligned} \int_0^b (h_y^T \delta \mathbf{y} + h_u^T \delta \mathbf{u} + h_c^T \delta \mathbf{c}) ds &= \int_0^b \mathbf{v}^T (\delta \mathbf{y}' - \mathbf{f}_y \delta \mathbf{y} - \mathbf{f}_u \delta \mathbf{u} - \mathbf{f}_c \delta \mathbf{c}) ds \\ &+ \int_0^b \mathbf{w}^T (\mathbf{g}_y \delta \mathbf{y} + \mathbf{g}_u \delta \mathbf{u} + \mathbf{g}_c \delta \mathbf{c}) ds, \end{aligned} \quad (4)$$

the perturbations of  $\mathbf{y}(s)$  must satisfy the zero boundary conditions:

$$\delta y_i(0) = 0 \quad (i \in \mathcal{I}), \quad \delta y_j(b) = 0 \quad (j \in \mathcal{F}), \quad (5)$$

and the Lagrange multipliers  $\mathbf{w}(s)$  must satisfy the complementarity conditions:

$$w_i g_i(\mathbf{y}(s), \mathbf{u}(s), \mathbf{c}) = 0 \quad (\forall s, \forall i), \quad \mathbf{0} \leq \mathbf{w}(s). \quad (6)$$

To eliminate the term  $\delta \mathbf{y}'$  using integration by parts, under the assumption that both  $\delta \mathbf{y}(s)$  and  $\mathbf{v}(s)$  are continuous and piecewise differentiable,

$$\int_0^b \mathbf{v}^T \delta \mathbf{y}' ds = - \int_0^b \mathbf{v}^T \delta \mathbf{y} ds,$$

where

$$v_i(0) = 0 \ (i \notin \mathcal{I}), \quad v_j(b) = 0 \ (j \notin \mathcal{F}). \quad (7)$$

The perturbations  $\delta \mathbf{y}(s)$ ,  $\delta \mathbf{u}(s)$ ,  $\delta \mathbf{c}$  are independent, and apart from the continuity condition on  $\delta \mathbf{y}(s)$ , they are also arbitrary, and so their coefficients must each match separately in equation (4), giving

$$\mathbf{v}^T = -\mathbf{v}^T \mathbf{f}_y + \mathbf{w}^T \mathbf{g}_y - h_y^T, \quad (8)$$

$$\mathbf{0}^T = -\mathbf{v}^T \mathbf{f}_u + \mathbf{w}^T \mathbf{g}_u - h_u^T, \quad (9)$$

$$\mathbf{0}^T = \int_0^b (-\mathbf{v}^T \mathbf{f}_c + \mathbf{w}^T \mathbf{g}_c - h_c^T) ds. \quad (10)$$

The original differential equations (and boundary conditions) (2), together with the adjoint equations (8–10), boundary conditions (7), inequality constraints (3), and complementarity conditions (6), form the Kuhn-Tucker necessary conditions for  $(\mathbf{y}, \mathbf{u}, \mathbf{c})$  to be a minimizer of the functional  $J(\mathbf{u})$  in equation (1) subject to constraints (2) and (3).

In order to express the integral equation (10) as a differential equation, new variables  $\mathbf{r}(s)$  may be introduced, corresponding to the constants  $\mathbf{c}$ , and satisfying

$$\mathbf{r}'^T = -\mathbf{v}^T \mathbf{f}_c + \mathbf{w}^T \mathbf{g}_c - h_c^T, \quad \mathbf{r}(0) = \mathbf{0}, \quad \mathbf{r}(b) = \mathbf{0}. \quad (11)$$

A Hamiltonian function may be introduced in the form

$$H(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{w}) := -\mathbf{f}^T(\mathbf{y}, \mathbf{u}, \mathbf{c})\mathbf{v} + \mathbf{g}^T(\mathbf{y}, \mathbf{u}, \mathbf{c})\mathbf{w} - h(\mathbf{y}, \mathbf{u}, \mathbf{c})$$

enabling the right-hand sides of (8–11) to be expressed in terms of  $H$ .

## 2.3 Elimination of inequalities

In order to eliminate the inequalities on  $\mathbf{g}$  and  $\mathbf{w}$  in the complementarity conditions (3) and (6), new variables  $\mathbf{p} = \mathbf{g} - \mathbf{w}$  may be introduced, such that

$$\begin{aligned} \mathbf{g} &= \max(\mathbf{0}, \mathbf{p}) := \mathbf{p}^+ \quad (\text{the positive part of } \mathbf{p}), \\ \mathbf{w} &= \max(-\mathbf{p}, \mathbf{0}) := \mathbf{p}^- \quad (\text{the negative part of } \mathbf{p}, \text{ with positive sign}). \end{aligned}$$

Then the Kuhn-Tucker necessary conditions (2),(3),(6–9) and (11) may be expressed in the form of the following system of DAEs subject to initial and terminal conditions:

$$\begin{aligned} \text{From (2):} \quad & \mathbf{y}' = \mathbf{f}(\mathbf{y}, \mathbf{u}, \mathbf{c}), & y_i(0) &= y_{i_0} \ (i \in \mathcal{I}), & y_j(b) &= y_{j_1} \ (j \in \mathcal{F}); \\ \text{From (8), (7):} \quad & \mathbf{v}' = H_y(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-), & v_i(0) &= 0 \ (i \notin \mathcal{I}), & v_j(b) &= 0 \ (j \notin \mathcal{F}); \\ \text{From (11):} \quad & \mathbf{r}' = H_c(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-), & \mathbf{r}(0) &= \mathbf{0}, & \mathbf{r}(b) &= \mathbf{0}; \\ \text{From (9):} \quad & \mathbf{0} = H_u(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-); \\ \text{From (3), (6):} \quad & \mathbf{0} = \mathbf{p}^+ - \mathbf{g}(\mathbf{y}, \mathbf{u}, \mathbf{c}). \end{aligned} \quad (12)$$

### 3 The tractability index Concept

#### 3.1 Short Introduction

In the case of linear DAEs, the index indicates how often we have to differentiate parts of the right-hand side of the DAE to obtain an expression for the solution. Therefore the index describes the difficulty involved in solving a system numerically.

A way of determining the index of a DAE is given by the *tractability index* concept (see also [16]). The motivation for the *tractability index* comes from an equivalent reformulation of a DAE without differentiation. This is important e.g. if the data of the DAE have low smoothness properties.

The definition of the *tractability index* is based on a matrix chain  $G_i, i \geq 0$  in the following way. Consider a DAE in quasilinear form

$$\mathbf{F}((D(t)\mathbf{x}(t))', \mathbf{x}(t), t) := A(\mathbf{x}, t)(D(t)\mathbf{x})' + \mathbf{b}(\mathbf{x}, t) = \mathbf{0}, \quad (13)$$

where  $\mathbf{F}(\mathbf{z}, \mathbf{x}, t) : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $A(\mathbf{x}, t) \in \mathbb{R}^{n \times m}$ ,  $D(t) \in \mathbb{R}^{m \times n}$  and  $\mathbf{b}(\mathbf{x}, t) \in \mathbb{R}^n$ .  $\mathbf{F}$  and  $\mathbf{b}$  must be sufficiently smooth.

We prefer DAEs with properly stated leading term, because of their clearer description and their better numerical properties (see [11], [12]). Properly stated leading term means that  $\ker A(\mathbf{x}, t) \oplus \text{im} D(t) = \mathbb{R}^m$  and the projector realizing this splitting is continuously differentiable (see [17]). With

$$B(\mathbf{z}, \mathbf{x}, t) := \mathbf{F}_x(\mathbf{z}, \mathbf{x}, t)$$

(we will drop the arguments) a matrix chain is defined by

$$\begin{aligned} G_0 &:= AD, & B_0 &:= B, \\ G_{i+1} &:= G_i + B_i Q_i, \\ B_{i+1} &:= (B_i - G_{i+1} D^- (DP_0 \dots P_{i+1} D^-)' DP_0 \dots P_{i-1}) P_i, \end{aligned} \quad (14)$$

where  $Q_i$  denotes a projector onto  $N_i := \ker G_i$ ,  $P_i := I - Q_i$  and  $D^-$  describes a reflexive generalized inverse of  $D$ , i.e.  $D = DD^-D$ ,  $D^- = D^-DD^-$  and additionally  $D^-D = P_0$ .

**Definition 3.1** (See [17]) *An equation (13) with properly stated leading term is said to be a DAE with tractability index  $\mu$  on the interval  $I$ , if there is a continuous matrix function sequence (14) such that*

- $$\left. \begin{aligned} &\text{(a) } G_i \text{ has constant rank } r_i \text{ on } I, \\ &\text{(b) } N_0 \oplus N_1 \oplus \dots \oplus N_{i-1} \subseteq \ker Q_i, \\ &\text{(c) } Q_i \in C(I, \mathbb{R}^{n \times n}), DP_0 \dots P_i D^- \in C^1(I, \mathbb{R}^{m \times m}) \\ &\text{(d) } 0 \leq r_0 \leq \dots \leq r_{\mu-1} < r_\mu = n. \end{aligned} \right\} 0 \leq i \leq \mu,$$

To check the index of a DAE we have to check the ranks of the matrices  $G_i, 0 \leq i \leq \mu$ .

**Remark:** The ranks, and therefore the index, are independent of linear transformations of the variables, the scaling of the equations, and the choice of projectors.

By means of the *tractability index* concept, it is also possible to get a cheap way to reduce the index of a higher index system of DAEs.

If we consider a system of DAEs of semiexplicit structure (13)

$$A(\mathbf{x}, t)(D(t)\mathbf{x})' + \mathbf{b}(\mathbf{x}, t) = \mathbf{0}$$

of index  $k$  (i.e.  $G_k$  remains nonsingular) the system of DAEs

$$A(\mathbf{x}, t)(D(t)\mathbf{x})' + (I - W_{k-1})\mathbf{b}(\mathbf{x}, t) + W_{k-1}\frac{d}{dt}(W_{k-1}\mathbf{b}(\mathbf{x}, t)) = \mathbf{0} \quad (15)$$

has, for a wide class of DAEs, index  $k - 1$ , where  $W_{k-1}$  denotes a projector along  $\text{im } G_{k-1}$ .

This is proved for linear equations and for index-2 equations of structure (13) (see [8]).

A theorem has been proved for index-3 equations, under certain conditions, in [6]. An extension of that theorem is given below.

**Theorem 3.2** *Let*

$$A(D\mathbf{x})' + \mathbf{b}(\mathbf{x}, t) = \mathbf{0} \quad (16)$$

*be a DAE with constant matrices  $A$  and  $D$ . Let  $W_2$  be a constant projector along  $\text{im } G_2$  and  $(W_2\mathbf{b})(\mathbf{x}, t) = (W_2\mathbf{b})(P_0\mathbf{x}, t)$ .*

1. *Let (16) have index 3 and  $I + Q_2G_3^{-1}[(W_2BD^-z)_x - (W_2B)_xD^-z]P_0Q_1Q_2$  be nonsingular for arbitrary  $\mathbf{z}$ . Then the system of DAEs*

$$A(D\mathbf{x})' + (I - W_2)\mathbf{b}(\mathbf{x}, t) + W_2\frac{d}{dt}(W_2\mathbf{b}(\mathbf{x}, t)) = \mathbf{0} \quad (17)$$

*has index 2.*

2. *Let DAE (17) have index 2 and  $W_2[(W_2BD^-z)_x - (W_2B)_xD^-z]P_0Q_1Q_2 = 0$ . Then (16) has index 3.*

Before we prove the theorem we collect together some useful properties of the projectors. Let  $W_{i+1}$  be a projector along  $\text{im } G_{i+1}$ . Using (14) we obtain

$$0 = W_{i+1}G_{i+1} = W_{i+1}(G_i + B_iQ_i). \quad (18)$$

Multiplying (18) from the right by  $P_i$  gives  $0 = W_{i+1}G_i$  and using the definition of  $G_i$  we derive that

$$0 = W_{i+1}G_j \text{ for } j = 0, \dots, i + 1, \quad (19)$$

which means that  $W_{i+1}(I - W_j) = 0$ ,  $j = 0, \dots, i + 1$ .

On the other hand multiplying (18) by  $Q_i$  we obtain  $0 = W_{i+1}B_iQ_i$  and taking the structure



of  $B_i$  into account (cf. (14)) we have  $W_{i+1}B_0Q_i = 0$ .

With Definition 3.1(b) we obtain  $Q_iQ_j = 0$ ,  $0 \leq j < i$  and multiplying (18) by  $Q_j$  from the right

$$W_{i+1}B_0Q_j = 0, \quad 0 \leq j \leq i. \quad (20)$$

**Proof** of Theorem 3.2:

1. Using

$$\frac{d}{dt}(W_2\mathbf{b}(\mathbf{x}, t)) = (W_2\mathbf{b})_x(\mathbf{x}, t)(P_0\mathbf{x})' + (W_2\mathbf{b})_t(\mathbf{x}, t)$$

equation (17) can be written in greater detail as

$$A(D\mathbf{x})' + (I - W_2)\mathbf{b}(\mathbf{x}, t) + W_2(W_2BD^-(D\mathbf{x})' + (W_2\mathbf{b})_t(\mathbf{x}, t)) = \mathbf{0} \quad (21)$$

where  $B = \mathbf{b}_x(\mathbf{x}, t)$ . The matrix chain of (17) with matrices linearized in  $(\mathbf{z}, \mathbf{x})$  is given by the following

$$\begin{aligned} \tilde{A} &= A + W_2BD^-, \quad \tilde{D} = D, \\ \tilde{G}_0 &= \tilde{A}\tilde{D} = G_0 + W_2BP_0, \\ \tilde{B}_0 &= (I - W_2)B + W_2(W_2BD^-\mathbf{z} + (W_2\mathbf{b})_t)_xP_0. \end{aligned}$$

We have to look for a nullspace projector of  $\tilde{G}_0$ .

We can write  $\tilde{G}_0 = (I - W_2)G_0P_0 + W_2BP_0$ , which shows that  $\ker \tilde{G}_0 = \ker(I - W_2)G_0P_0 \cap \ker W_2BP_0$ , i.e. that we can choose the same nullspace projector  $\tilde{Q}_0 = Q_0$ .

The next chain element is given by

$$\begin{aligned} \tilde{G}_1 = \tilde{G}_0 + \tilde{B}_0\tilde{Q}_0 &= G_0 + W_2BD^-D + (I - W_2)BQ_0 \\ &= G_1 + W_2BP_0. \end{aligned}$$

From (20) we derive  $W_2BP_0 = W_2BP_0P_1$ , and with the same arguments as before  $\tilde{G}_1$  and  $G_1$  have the same nullspace, i.e. we can choose  $\tilde{Q}_1 = Q_1$ . Then

$$\begin{aligned} \tilde{B}_1 &= \tilde{B}_0P_0 - \tilde{G}_1D^-(DP_1D^-)'D \\ &= ((I - W_2)B + W_2(W_2BD^-\mathbf{z} + (W_2\mathbf{b})_t)_x)P_0 - \tilde{G}_1D^-(DP_1D^-)'D. \end{aligned}$$

The next step gives

$$\begin{aligned} \tilde{G}_2 &= \tilde{G}_1 + \tilde{B}_1\tilde{Q}_1, \\ &= G_1 + W_2BP_0 + (I - W_2)BP_0Q_1 \\ &\quad + W_2(W_2BD^-\mathbf{z})_xP_0Q_1 + \underbrace{W_2(W_2\mathbf{b})_t)_xP_0Q_1}_{=W_2(W_2B)_tP_0Q_1} - \tilde{G}_1D^-(DP_1D^-)'DQ_1, \\ &= (G_1 + BP_0Q_1)(I - P_1D^-(DP_1D^-)'DQ_1) + W_2BP_0P_1 \\ &\quad + W_2(W_2BD^-\mathbf{z})_xP_0Q_1 + W_2(W_2B)_tP_0Q_1 - \underbrace{W_2BD^-(DP_1D^-)'DQ_1}_{W_2[(W_2BP_0P_1)' - (W_2B)'P_0P_1]P_0Q_1}, \\ &= G_2 + W_2BP_0P_1 \\ &\quad + W_2(W_2BD^-\mathbf{z})_xP_0Q_1 + W_2(W_2B)_tP_0Q_1 - W_2[(W_2B)_xD^-\mathbf{z} + (W_2B)_t]P_0Q_1, \\ &= G_2 + W_2B_2 + W_2[(W_2BD^-\mathbf{z})_x - (W_2B)_xD^-\mathbf{z}]P_0Q_1. \end{aligned}$$

Consider  $\tilde{G}_2 \mathbf{w} = \mathbf{0}$ . Multiplying

$$(G_2 + W_2 B_2 + W_2 [(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}] P_0 Q_1) \mathbf{w} = \mathbf{0}$$

by  $(I - W_2)$  we get  $G_2 \mathbf{w} = \mathbf{0}$ , which leads to  $\mathbf{w} = Q_2 \mathbf{w}$ . Using this we obtain

$$(W_2 (B_2 + [(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}] P_0 Q_1) Q_2) \mathbf{w} = \mathbf{0}. \quad (22)$$

If we take into account that  $\ker W_2 = \text{im } G_2 = \ker Q_2 G_3^{-1}$  then the left side of (22) is

$$Q_2 \underbrace{G_3^{-1} (B_2 Q_2 + [(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}] P_0 Q_1 Q_2)}_{=Q_2} \mathbf{w} = \mathbf{0}. \quad (23)$$

Equation (23) leads to

$$(I + Q_2 G_3^{-1} [(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}] P_0 Q_1 Q_2) Q_2 \mathbf{w} = \mathbf{0}$$

and hence  $Q_2 \mathbf{w} = \mathbf{0}$ .

This means that  $\tilde{G}_2$  is nonsingular and (17) has index 2.

2. To prove the second part we assume that (17) has index 2. The related matrix chain element is nonsingular. We obtain

$$\begin{aligned} \tilde{G}_2 &= G_2 + W_2 B_2 + W_2 \underbrace{[(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}] P_0 Q_1}_{=:\Delta}, \\ &= (I + W_2 (B_2 + \Delta P_0 Q_1) G_2^-) (G_2 + W_2 (B_2 + \Delta P_0 Q_1) Q_2), \\ &= (I + W_2 (B_2 + \Delta P_0 Q_1) G_2^-) (G_2 + (B_2 + W_2 \Delta P_0 Q_1) Q_2) (I + G_2^- B_2 Q_2), \\ &= \underbrace{(I + W_2 (B_2 + \Delta P_0 Q_1) G_2^-)}_{\text{nonsingular}} \underbrace{(G_2 + B_2 Q_2)}_{=G_3} \underbrace{+ W_2 \Delta P_0 Q_1 Q_2}_{=0} \underbrace{(I + G_2^- B_2 Q_2)}_{\text{nonsingular}}, \end{aligned}$$

i.e. that if  $\tilde{G}_2$  is nonsingular then so is  $G_3$ , and so (16) has index 3.  $\square$

#### Remarks:

1. The nonsingularity condition  $W_2 \Delta P_0 Q_1 Q_2 = 0$  is also sufficient for the first statement of Theorem 3.2. To check that condition is not trivial, because it requires the computation of  $[(W_2 B D^- \mathbf{z})_x - (W_2 B)_x D^- \mathbf{z}]$ . But it can be seen immediately that the condition is fulfilled for linear DAEs.
2. The second part of this theorem is new, and makes it possible to show that the original system of DAEs had index 3. This will be used in Problem 3 of Section 4.

### 3.2 The tractability index of the DAEs

We will investigate the index of the system of DAEs (12) in general form, applying the *tractability index* concept. To get a system of DAEs, which has as many equations as unknowns, we introduce an extra ODE for  $\mathbf{c}$ . The system of DAEs is given by

$$\begin{aligned} \mathbf{y}' &= \mathbf{f}(\mathbf{y}, \mathbf{u}, \mathbf{c}), \\ \mathbf{c}' &= \mathbf{0}, \\ \mathbf{v}' &= H_y(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-), \\ \mathbf{r}' &= H_c(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-), \\ \mathbf{0} &= H_u(\mathbf{y}, \mathbf{v}, \mathbf{c}, \mathbf{u}, \mathbf{p}^-), \\ \mathbf{0} &= \mathbf{p}^+ - \mathbf{g}(\mathbf{y}, \mathbf{u}, \mathbf{c}). \end{aligned} \tag{24}$$

We have to stress here that (24) is not Hessenberg. Therefore an extra investigation is necessary. The matrices  $A$ ,  $D$  and  $B$  are

$$A = \begin{pmatrix} I & & & & & \\ & I & & & & \\ & & I & & & \\ & & & I & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix}, \quad D = \begin{pmatrix} I & & & & & \\ & I & & & & \\ & & I & & & \\ & & & I & & \\ & & & & I & 0 \\ & & & & & 0 \end{pmatrix}$$

and with the unknown vector  $\mathbf{x} = (\mathbf{y}^T, \mathbf{c}^T, \mathbf{v}^T, \mathbf{r}^T, \mathbf{u}^T, \mathbf{p}^T)^T$  we obtain

$$B = \mathbf{b}'_x = \left( \begin{array}{cccc|cc} -\mathbf{f}_y & -\mathbf{f}_c & 0 & 0 & -\mathbf{f}_u & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -H_{yy} & -H_{yc} & -H_{yv} & 0 & -H_{yu} & -H_{yp} \\ -H_{cy} & -H_{cc} & -H_{cv} & 0 & -H_{cu} & -H_{cp} \\ \hline H_{uy} & H_{uc} & H_{uv} & 0 & H_{uu} & H_{up} \\ -\mathbf{g}_y & -\mathbf{g}_c & 0 & 0 & -\mathbf{g}_u & \mathbf{p}_p^+ \end{array} \right) =: \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

where  $\mathbf{p}_p^+ := \frac{\partial \mathbf{p}^+}{\partial \mathbf{p}} = \left( \frac{\partial p_i^+}{\partial p_i} \right)$ . The first matrix chain element is

$$G_0 = AD = \begin{pmatrix} I & & & & & \\ & I & & & & \\ & & I & & & \\ & & & I & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} \quad \text{and} \quad Q_0 = \begin{pmatrix} 0 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & I & \\ & & & & & I \end{pmatrix}$$

is a nullspace projector of  $G_0$ . The next chain element  $G_1$  will be calculated as  $G_1 = G_0 + BQ_0$ . We find

$$G_1 = \begin{pmatrix} I & B_{12} \\ & B_{22} \end{pmatrix}.$$

It is easy to see that  $G_1$  is nonsingular iff  $B_{22}$  remains nonsingular, i.e. we have an index-1 system of DAEs. If  $B_{22}$  is singular and we know a nullspace projector of  $B_{22}$ , we can construct

a nullspace projector  $Q_1$  of  $G_1$ . For DAEs with *tractability index* we know that  $N_k \cap N_{k+1} = \{0\}$  (see [17]).

In particular for  $k = 0$ ,  $\{0\} = N_0 \cap N_1 = \ker G_0 \cap (\ker G_0 \cap \ker BQ_0) = \ker G_0 \cap \ker BQ_0$ . Therefore  $\begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix}$  must have full rank.

Let  $\bar{Q}_1$  be a nullspace projector of  $B_{22}$ ; then if  $R = \bar{Q}_1 S_{B_2}^{-1} B_{12}^T$  and  $S_{B_2} := \begin{pmatrix} B_{12}^T & B_{22}^T \end{pmatrix} \begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix}$ ,

$$Q_1 = \begin{pmatrix} B_{12}R & 0 \\ -R & 0 \end{pmatrix} \quad (25)$$

represents a nullspace projector of  $G_1$  with  $Q_1 Q_0 = 0$ . If we know  $Q_1$  we can calculate the next matrix chain element

$$G_2 = G_1 + B_1 Q_1 = \underbrace{(G_1 + B_0 P_0 Q_1)}_{=: \mathcal{G}_2} (I - P_1 D^- (D P_1 D^-)' D Q_1). \quad (26)$$

To investigate the singularity of  $G_2$  it is sufficient to investigate the singularity of  $\mathcal{G}_2$ , because the second factor in the representation (26) of  $G_2$  remains nonsingular.

In order to construct a nullspace projector  $\bar{Q}_1$  of  $B_{22}$  the structure of the given problem is sometimes useful. Very often the objective function and the right hand sides  $\mathbf{f}$  of the ODEs and  $\mathbf{g}$  of the inequalities depend only linearly on the control  $\mathbf{u}$ . In that case  $H_{uu} \equiv 0$ . If additionally  $\mathbf{g}_u$  has full rank the following lemma is valid.

**Lemma 3.3** 1. The matrix  $M = \begin{pmatrix} 0 & \mathbf{g}_u^T \mathbf{p}_p^- \\ -\mathbf{g}_u & \mathbf{p}_p^+ \end{pmatrix}$  with full rank  $\mathbf{g}_u$  is nonsingular iff  $Z = (\mathbf{p}_p^+ - \mathbf{g}_u (\mathbf{g}_u^T \mathbf{g}_u)^{-1} \mathbf{g}_u^T)$  is nonsingular and

2. if  $M$  is singular, a nullspace projector onto  $\ker M$  is given by  $\bar{Q} = \begin{pmatrix} 0 & (\mathbf{g}_u^T \mathbf{g}_u)^{-1} \mathbf{g}_u^T \tilde{Q} \\ 0 & \tilde{Q} \end{pmatrix}$ , where  $\tilde{Q}$  describes a nullspace projector onto  $\ker Z$ .

**Proof:** 1. From  $\mathbf{p} = \mathbf{p}^+ - \mathbf{p}^-$  we obtain  $I = \mathbf{p}_p^+ - \mathbf{p}_p^-$ . Using  $\mathbf{p}_p^- = \mathbf{p}_p^+ - I$  we obtain that  $\mathbf{g}_u^T Z = \mathbf{g}_u^T \mathbf{p}_p^+ - \mathbf{g}_u^T = \mathbf{g}_u^T \mathbf{p}_p^-$ . Multiplying  $M$  with a nonsingular matrix

$$M \begin{pmatrix} I & (\mathbf{g}_u^T \mathbf{g}_u)^{-1} \mathbf{g}_u^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{g}_u^T \\ -\mathbf{g}_u & I \end{pmatrix} \begin{pmatrix} I \\ Z \end{pmatrix}.$$

We obtain a factorization into two matrices. The first factor remains nonsingular for full rank  $\mathbf{g}_u$  and it is shown that  $M$  is nonsingular iff  $Z$  is nonsingular.

2. Let  $\tilde{Q}$  be a nullspace projector onto  $\ker Z$ . From  $\mathbf{g}_u^T Z \tilde{Q} = 0$  we obtain, using  $I = \mathbf{p}_p^+ - \mathbf{p}_p^-$ , that  $\mathbf{g}_u^T \mathbf{p}_p^- \tilde{Q} = 0$ . Then it is easy to see that  $M \bar{Q} = 0$ .  $\square$

## 4 Examples

To illustrate the preceding theoretical developments, we apply them to three known examples.

## 4.1 Problem 1 (see [5, 15])

A simple example of such a problem is that of *Minimum time to cover a fixed distance*.

A vehicle has to travel a fixed distance (300 units) in the shortest possible time, starting from rest, finishing at rest, and subject to limits (1 and -2) on the acceleration and deceleration.

### 4.1.1 Problem statement

Let the time taken to cover the distance be  $t_f > 0$ . Then the problem is to

$$\begin{aligned} &\text{minimize} && t_f \\ &\text{subject to:} && \frac{dx_1}{dt} = x_2, \quad x_1(0) = 0, \quad x_1(t_f) = 300, \\ &&& \frac{dx_2}{dt} = a, \quad x_2(0) = 0, \quad x_2(t_f) = 0, \\ &&& -2 \leq a \leq 1, \end{aligned}$$

where  $a$  is the acceleration.

### 4.1.2 Conversion to the general formulation

In order to express this problem in the form given in (1–3), we define variables and constants as follows:

$$x_1 = y_1, \quad x_2 = y_2, \quad t = t_f s, \quad a = u, \quad t_f = c > 0,$$

and so obtain

$$\begin{aligned} &\text{minimize} && t_f = \int_0^1 c \, ds \\ &\text{subject to:} && \frac{dy_1}{ds} = y_2 c, \quad y_1(0) = 0, \quad y_1(1) = 300, \\ &&& \frac{dy_2}{ds} = uc, \quad y_2(0) = 0, \quad y_2(1) = 0, \\ &&& 0 \leq 2 + u, \\ &&& 0 \leq 1 - u. \end{aligned}$$

The exact solution of this problem is given in [6, 15]. However it can be checked that initially, with maximum acceleration we have

$$\begin{aligned} u &= 1, \\ y_2 &= cs, \\ y_1 &= \frac{1}{2}c^2s^2; \end{aligned} \tag{27}$$

and finally, with greatest deceleration we obtain

$$\begin{aligned} u &= -2, \\ y_2 &= 2c(1-s), \\ y_1 &= 300 - c^2(1-s)^2; \end{aligned} \tag{28}$$

and that these two solutions match when  $s = \frac{2}{3}$  and  $c = 30$ .

The Hamiltonian function is

$$H = -y_2cv_1 - ucv_2 + (2+u)p_1^- + (1-u)p_2^- - c.$$

#### 4.1.3 System of DAEs

Using the procedure outlined above, this gives rise to the following system of DAEs (12) without inequalities:

$$\begin{aligned} \frac{dy_1}{ds} &= y_2c, & y_1(0) &= 0, & y_1(1) &= 300, \\ \frac{dy_2}{ds} &= uc, & y_2(0) &= 0, & y_2(1) &= 0, \\ \frac{dv_1}{ds} &= 0, \\ \frac{dv_2}{ds} &= -cv_1, \\ \frac{dr}{ds} &= -y_2v_1 - uv_2 - 1, & r(0) &= 0, & r(1) &= 0, \\ 0 &= -cv_2 + p_1^- - p_2^-, \\ 0 &= p_1^+ - 2 - u, \\ 0 &= p_2^+ - 1 + u. \end{aligned}$$

This system has 8 variables ( $y_1, y_2, v_1, v_2, r, u, p_1, p_2$ ) and 1 unknown constant ( $c$ ) which must satisfy 5 equations with derivatives and 3 algebraic equations. It therefore requires 6 boundary conditions corresponding to the 5 differential equations and the unknown constant. It appears to have exactly the correct number of boundary conditions to determine a unique solution.

#### 4.1.4 The matrix chain

The vector of dependent variables is given by  $\mathbf{x} = (y_1, y_2, c, v_1, v_2, r, u, p_1, p_2)$ .

$$G_0 = \begin{pmatrix} I_6 & \\ & 0_3 \end{pmatrix}, \quad Q_0 = \begin{pmatrix} 0_6 & \\ & I_3 \end{pmatrix}.$$

The matrix  $B$  is given by

$$B = \left( \begin{array}{cccccc|ccc} 0 & -c & -y_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -u & 0 & 0 & 0 & -c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v_1 & c & 0 & 0 & 0 & 0 & 0 \\ 0 & v_1 & 0 & y_2 & u & 0 & v_2 & 0 & 0 \\ \hline 0 & 0 & -v_2 & 0 & -c & 0 & 0 & p_{1p_1}^- & -p_{2p_2}^- \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ \end{array} \right),$$

and the next chain matrix is calculated as

$$G_1 = \left( \begin{array}{cccccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -c & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & v_2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{1p_1}^- & -p_{2p_2}^- \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ \end{array} \right).$$

The nonsingularity of  $G_1$  depends on the nonsingularity of  $\begin{pmatrix} 0 & p_{1p_1}^- & -p_{2p_2}^- \\ -1 & p_{1p_1}^+ & 0 \\ 1 & 0 & p_{2p_2}^+ \end{pmatrix}$ . This matrix has exactly the structure of matrix  $M$  of Lemma 3.3. The relevant matrix  $Z$  is given by

$$Z = \begin{pmatrix} p_{1p_1}^+ - \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & p_{2p_2}^+ - \frac{1}{2} \end{pmatrix}.$$

Here two cases are possible: either  $p_1$  and  $p_2$  have the same sign or they do not.

For different signs of  $p_1$  and  $p_2$ ,  $\det Z = -\frac{1}{2}$ , which means that  $G_1$  is nonsingular and the DAE has index 1. If both  $p_1$  and  $p_2$  are negative the last two equations create a contradiction, because each of them gives a fixed value of  $z$ , but they are different (-2 and 1); in terms of the original problem statement both constraints are active simultaneously. The DAE has no *tractability index* in that case, because  $\begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix}$  does not have full rank.

If both  $p_1$  and  $p_2$  are positive the last two equations do not determine the control  $u$ . The algebraic equation  $0 = -cv_2$  implies that  $v_2 = 0$  and from the fourth equation  $v_1 = 0$ . The acceleration  $u$  appears in the second equation but, together with the first equation, that provides insufficient information to determine  $u$ . The DAE has no *tractability index* in that case and this can be confirmed in the numerical results obtained in Section 5.

## 4.2 Problem 2 (see [15])

A slightly more complicated problem is given by imposing a *Speed limit*.

### 4.2.1 Problem statement

Let the speed limit be  $k$ , where the other variables have the same meaning as before. The problem is to

$$\begin{aligned} & \text{minimize} && t_f \\ & \text{subject to :} && \frac{dx_1}{dt} = x_2, \quad x_1(0) = 0, \quad x_1(t_f) = 300, \\ & && \frac{dx_2}{dt} = a, \quad x_2(0) = 0, \quad x_2(t_f) = 0, \\ & && -2 \leq a \leq 1, \quad x_2 \leq k. \end{aligned}$$

### 4.2.2 Conversion to the general formulation

We define variables and constants as before

$$x_1 = y_1, \quad x_2 = y_2, \quad t = t_f s, \quad a = u, \quad t_f = c > 0,$$

and so obtain

$$\begin{aligned} & \text{minimize} && t_f = \int_0^1 c \, ds \\ & \text{subject to :} && \frac{dy_1}{ds} = y_2 c, \quad y_1(0) = 0, \quad y_1(1) = 300, \\ & && \frac{dy_2}{ds} = uc, \quad y_2(0) = 0, \quad y_2(1) = 0, \\ & && 0 \leq 2 + u, \\ & && 0 \leq 1 - u, \\ & && 0 \leq k - y_2. \end{aligned}$$

Again, the exact solution of this problem is given in [6, 15].

If  $k \geq 20$  the solution is identical to that of Problem 1.

If  $k \leq 20$  the solution is:

for  $0 \leq s \leq \frac{k}{c}$

$$\begin{aligned} u &= 1, \\ y_2 &= cs, \\ y_1 &= \frac{1}{2}c^2 s^2; \end{aligned} \tag{29}$$



for  $\frac{k}{c} \leq s \leq 1 - \frac{k}{2c}$

$$\begin{aligned} u &= 0, \\ y_2 &= k, \\ y_1 &= kcs - \frac{1}{2}k^2; \end{aligned} \tag{30}$$

for  $1 - \frac{k}{2c} \leq s \leq 1$

$$\begin{aligned} u &= -2, \\ y_2 &= 2c(1-s), \\ y_1 &= 300 - c^2(1-s)^2; \end{aligned} \tag{31}$$

and these match if  $c = \frac{3}{4k}(400 + k^2)$ .

The Hamiltonian function is

$$H = -y_2cv_1 - ucv_2 + (2+u)p_1^- + (1-u)p_2^- + (k-y_2)p_3^- - c.$$

### 4.2.3 System of DAEs

This gives rise to the system of DAEs (12) without inequalities:

$$\begin{aligned} \frac{dy_1}{ds} &= y_2c, & y_1(0) &= 0, & y_1(1) &= 300, \\ \frac{dy_2}{ds} &= uc, & y_2(0) &= 0, & y_2(1) &= 0, \\ \frac{dv_1}{ds} &= 0, \\ \frac{dv_2}{ds} &= -cv_1 - p_3^-, \\ \frac{dr}{ds} &= -y_2v_1 - uv_2 - 1, & r(0) &= 0, & r(1) &= 0, \\ 0 &= -cv_2 + p_1^- - p_2^-, \\ 0 &= p_1^+ - 2 - u, \\ 0 &= p_2^+ - 1 + u, \\ 0 &= p_3^+ - k + y_2. \end{aligned}$$

This system has 9 variables ( $y_1, y_2, v_1, v_2, r, u, p_1, p_2, p_3$ ) and 1 unknown constant ( $c$ ) which must satisfy 5 equations with derivatives and 4 algebraic equations. It therefore requires 6 boundary conditions corresponding to the 5 differential equations and the unknown constant. It appears to have exactly the correct number of boundary conditions to determine a unique solution.

#### 4.2.4 The matrix chain

The vector of dependent variables is given by  $\mathbf{x} = (y_1, y_2, c, v_1, v_2, r, u, p_1, p_2, p_3)$ .

$$G_0 = \begin{pmatrix} I_6 & \\ & 0_4 \end{pmatrix}, \quad Q_0 = \begin{pmatrix} 0_6 & \\ & I_4 \end{pmatrix}.$$

The matrix  $B$  is given by

$$B = \left( \begin{array}{cccccc|cccc} 0 & -c & -y_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -u & 0 & 0 & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v_1 & c & 0 & 0 & 0 & 0 & 0 & p_{3p_3}^- \\ 0 & v_1 & 0 & y_2 & u & 0 & v_2 & 0 & 0 & 0 \\ \hline 0 & 0 & -v_2 & 0 & -c & 0 & 0 & p_{1p_1}^- & -p_{2p_2}^- & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{3p_3}^+ \end{array} \right),$$

and the next chain matrix is calculated as

$$G_1 = \left( \begin{array}{cccccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & p_{3p_3}^- \\ 0 & 0 & 0 & 0 & 0 & 1 & v_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{1p_1}^- & -p_{2p_2}^- & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{3p_3}^+ \end{array} \right).$$

The matrix in the lower right-hand corner, which determines the singularity of  $G_1$ , has the structure of  $M$  in Lemma 3.3 and the associated matrix  $Z$  has the structure

$$Z = \begin{pmatrix} -p_{1p_1}^+ - \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & p_{2p_2}^+ - \frac{1}{2} & 0 \\ 0 & 0 & p_{3p_3}^+ \end{pmatrix}.$$

If  $p_3 > 0$  we discover the same cases as in Problem 1: if  $p_1$  and  $p_2$  have different signs then  $Z$  and therefore  $M$  is nonsingular and the DAE has index 1; if  $p_1$  and  $p_2$  have the same sign no index is defined.

If  $p_3 < 0$  and both  $p_1$  and  $p_2$  are negative,  $\begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix}$  does not have full rank (all three constraints are active). If  $p_3 < 0$  and  $p_1$  and  $p_2$  have different signs, the DAE has no *tractability index*.

If  $p_3 < 0$ ,  $p_1 > 0$  and  $p_2 > 0$ ,  $Z$  is given by  $Z = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  and a nullspace projector is

$\tilde{Q} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . Using the nullspace projector of  $M$  constructed in Lemma 3.3 we obtain a nullspace projector of  $G_1$  by (25) as

$$Q_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c^2\mu & 0 & 0 & 0 & -cv_2\mu & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -cv_2\mu & 0 & 0 & 0 & v_2^2\mu & 0 & 0 & 0 & 0 \\ 0 & c\mu & 0 & 0 & 0 & -v_2\mu & 0 & 0 & 0 & 0 \\ 0 & c\mu & 0 & 0 & 0 & -v_2\mu & 0 & 0 & 0 & 0 \\ 0 & -c\mu & 0 & 0 & 0 & v_2\mu & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

with  $\mu = \frac{1}{c^2+v_2^2}$  and we obtain  $\mathcal{G}_2$  by (26) as

$$\mathcal{G}_2 = \begin{pmatrix} 1 & -c^3\mu & 0 & 0 & 0 & c^2v_2\mu & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -c & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & c^2v_1\mu & 0 & 0 & u & 1 - cv_1v_2\mu & v_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -c & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & c^2\mu & 0 & 0 & 0 & -cv_2\mu & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix  $\mathcal{G}_2$  remains nonsingular ( $\det(\mathcal{G}_2) = c^2$ ) and we have an index-2 DAE. The projector

$$W_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

shows that we have to differentiate the seventh and the tenth equations to reduce the index. This index-2 DAE corresponds to the solution (30).

### 4.3 Problem 3 (see [15], [13])

A problem concerning *Catalyst mixing for a packed bed reactor*.

A chemical  $A$  is fed into one end of the reactor. One catalyst applies to a reversible reaction which converts  $A$  to a chemical  $B$ . A second catalyst converts  $B$  to a product  $C$ . The aim is to mix the catalysts, with a proportion  $F$  of the first and  $1 - F$  of the second, in such a way as to maximize the final concentration of  $C$ .

#### 4.3.1 Problem statement

In [15], this problem is given as one of maximizing the concentration  $(1 - z^a(x_f) - z^b(x_f))$ . The statement of the problem is as follows:

$$\begin{aligned} \max_F \quad & (1 - z^a(x_f) - z^b(x_f)) \\ \text{subject to: } \quad & \frac{dz^a}{dx} = F(10z^b - z^a), \quad z^a(0) = 1, \\ & \frac{dz^b}{dx} = F(z^a - 10z^b) - (1 - F)z^b, \quad z^b(0) = 0, \\ & 0 \leq F \leq 1. \end{aligned}$$

#### 4.3.2 Conversion to the general formulation

In order to express the problem in the form given in (1-3), we define variables and constants as follows:

$$z^a = y_1, \quad z^b = y_2, \quad x = s, \quad F = u, \quad x_f = b,$$

rewrite the objective function as

$$\min_F (z^a(x_f) + z^b(x_f) - 1) = \min_u \int_0^b (u - 1)y_2 \, ds$$

and the inequality constraints as  $0 \leq u, 0 \leq 1 - u$ .

In [6], we produced the exact solution that had also been reported in [13].

If  $b \leq \frac{1}{11} \ln(1 + \sqrt{12.1}) + \ln(1 + \sqrt{0.1}) =: b_c \approx 0.4111$  and  $s_c$  satisfies

$$e^{s_c}(e^{11s_c} + 10) = 11e^b,$$

the solution is:

for  $0 \leq s \leq s_c$

$$\begin{aligned} u &= 1, \\ y_1 &= \frac{1}{11}(10 + e^{-11s}), \\ y_2 &= \frac{1}{11}(1 - e^{-11s}); \end{aligned} \tag{32}$$

for  $s_c \leq s \leq b$

$$\begin{aligned} u &= 0, \\ y_1 &= \frac{1}{11}(10 + e^{-11s_c}), \\ y_2 &= \frac{1}{11}(1 - e^{-11s_c})e^{-(s-s_c)}. \end{aligned} \quad (33)$$

If  $b \geq b_c$  the solution consists of three parts.

Let  $s_1 := \frac{1}{11} \ln(1 + \sqrt{12.1}) \approx 0.1363$  and  $s_2 := b - \ln(1 + \sqrt{0.1}) \approx b - 0.2748$ . Then for  $0 \leq s \leq s_1$

$$\begin{aligned} u &= 1, \\ y_1 &= \frac{1}{11}(10 + e^{-11s}), \\ y_2 &= \frac{1}{11}(1 - e^{-11s}); \end{aligned} \quad (34)$$

for  $s_1 \leq s \leq s_2$

$$\begin{aligned} u &= \frac{5\sqrt{10} - 4}{52} \approx 0.2271, \\ y_1 &= \frac{1}{111}(100 + \sqrt{10})e^{\frac{1}{52}(-6+\sqrt{10})(s-s_1)}, \\ y_2 &= \frac{1}{111}(11 - \sqrt{10})e^{\frac{1}{52}(-6+\sqrt{10})(s-s_1)}; \end{aligned} \quad (35)$$

for  $s_2 \leq s \leq b$

$$\begin{aligned} u &= 0, \\ y_1 &= \frac{1}{111}(100 + \sqrt{10})e^{\frac{1}{52}(-6+\sqrt{10})(s_2-s_1)}, \\ y_2 &= \frac{1}{111}(11 - \sqrt{10})e^{\frac{1}{52}(-6+\sqrt{10})(s_2-s_1)-(s-s_2)}. \end{aligned} \quad (36)$$

The Hamiltonian is

$$H = -u(10y_2 - y_1)(v_1 - v_2) + (1 - u)y_2(v_2 + 1) + up_1^- + (1 - u)p_2^-.$$

### 4.3.3 System of DAEs

The problem gives rise to the system of DAEs (12) without inequalities:

$$\begin{aligned} \frac{dy_1}{ds} &= (10y_2 - y_1)u, & y_1(0) &= 1, \\ \frac{dy_2}{ds} &= (y_1 - 10y_2)u - (1 - u)y_2, & y_2(0) &= 0, \\ \frac{dv_1}{ds} &= (v_1 - v_2)u, & v_1(b) &= 0, \\ \frac{dv_2}{ds} &= 10(v_2 - v_1)u + (v_2 + 1)(1 - u), & v_2(b) &= 0, \\ 0 &= (y_1 - 10y_2)(v_1 - v_2) - y_2(v_2 + 1) + p_1^- - p_2^-, \\ 0 &= p_1^+ - u, \\ 0 &= p_2^+ - 1 + u. \end{aligned}$$

We now have a system with 7 variables  $(y_1, y_2, v_1, v_2, u, p_1, p_2)$  which must satisfy 4 equations with derivatives and 3 algebraic equations. It therefore requires 4 boundary conditions corresponding to the 4 differential equations, and appears to have exactly the correct number of boundary conditions to determine a unique solution.

#### 4.3.4 The matrix chain

The vector of dependent variables is given by  $\mathbf{x} = (y_1, y_2, v_1, v_2, u, p_1, p_2)$ .

$$G_0 = \begin{pmatrix} I_4 & \\ & 0_3 \end{pmatrix}, \quad Q_0 = \begin{pmatrix} 0_4 & \\ & I_3 \end{pmatrix}.$$

For the matrix  $B$  we have

$$B = \left( \begin{array}{cccc|ccc} u & -10u & 0 & 0 & y_1 - 10y_2 & 0 & 0 \\ -u & 1 + 9u & 0 & 0 & -y_1 + 9y_2 & 0 & 0 \\ 0 & 0 & -u & u & -v_1 + v_2 & 0 & 0 \\ 0 & 0 & 10u & -1 - 9u & 1 + 10v_1 - 9v_2 & 0 & 0 \\ \hline v_1 - v_2 & -(1 + 10v_1 - 9v_2) & y_1 - 10y_2 & -y_1 + 9y_2 & 0 & p_{1p_1}^- & p_{2p_2}^- \\ 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ \end{array} \right),$$

and the first chain element is given by

$$G_1 = \left( \begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & y_1 - 10y_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & -y_1 + 9y_2 & 0 & 0 \\ 0 & 0 & 1 & 0 & -v_1 + v_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 + 10v_1 - 9v_2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & p_{1p_1}^- & p_{2p_2}^- \\ 0 & 0 & 0 & 0 & -1 & p_{1p_1}^+ & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & p_{2p_2}^+ \end{array} \right) =: \begin{pmatrix} G_{11}^1 & G_{12}^1 \\ G_{21}^1 & G_{22}^1 \end{pmatrix}.$$

The submatrix  $G_{22}^1$  of  $G_1$ , which determines the singularity is exactly the same as in Problem 1. We have a nonsingular matrix  $G_1$  if the signs of  $p_1$  and  $p_2$  are different. If  $p_1$  and  $p_2$  are both negative we have a nonregular DAE (see Problem 1). Lastly we have to investigate the case where both  $p_1$  and  $p_2$  are positive.

If  $p_1 > 0$  and  $p_2 > 0$  then  $Z = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$  and a nullspace projector  $\tilde{Q} = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}$ .

The gradient of the constraint vector  $\mathbf{g}$  is  $\mathbf{g}_u = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  and from Lemma 3.3 the projector

$$\bar{Q}_1 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}. \text{ From (26)}$$

$$G_2 = G_1 + BP_0Q_1 = \begin{pmatrix} I & B_{12} \\ 0 & B_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} I & \\ & 0 \end{pmatrix} \begin{pmatrix} B_{12}R & 0 \\ -R & 0 \end{pmatrix} = \begin{pmatrix} I + B_{11}B_{12}R & B_{12} \\ B_{21}B_{12}R & B_{22} \end{pmatrix}.$$

By examination of  $B$ , it may be seen that  $B_{21}B_{12} \equiv 0$  and  $B_{22} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$  therefore  $\mathcal{G}_2$  has a zero row and is singular. The singularity of  $\mathcal{G}_2$  leads to a singular matrix  $G_2$ , which means that the DAE has index at least 3 (if it exists). To investigate that theoretically, we apply Theorem 3.2. To check the assumptions we have used a formula manipulation system.

The image of  $\mathcal{G}_2$  is given by  $\{\mathbf{y} : \mathbf{y} = \mathcal{G}_2\mathbf{z}, \mathbf{z} \in \mathbb{R}^n\}$ . Setting  $\mathbf{z} = \begin{pmatrix} (I - B_{12}R)\mathbf{u} \\ Ru + (I - \bar{Q}_1) \begin{pmatrix} 0 \\ \mathbf{v} \end{pmatrix} \end{pmatrix}$  with

arbitrary vectors  $\mathbf{u} \in \mathbb{R}^4$  and  $\mathbf{v} \in \mathbb{R}^2$  we get  $\mathcal{G}_2\mathbf{z} = \begin{pmatrix} \mathbf{u} \\ 0 \\ \mathbf{v} \end{pmatrix}$ . This shows that the projector along  $\text{im } G_2 = \text{im } \mathcal{G}_2$  is given by

$$W_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have to differentiate the fifth equation. The resulting equation (17) which may be written

$$(A + W_2BD^-)(D\mathbf{x})' + (I - W_2)\mathbf{b}(\mathbf{x}) = \mathbf{0} \quad (37)$$

has index 2.

We check this index-2 property by computing  $\tilde{\mathcal{G}}_2 = \tilde{G}_1 + \tilde{B}_1Q_1$  with  $\tilde{G}_1 = G_0 + W_2BP_0$  and  $\tilde{B}_1 = (I - W_2)BP_0$  using  $Q_1$  given by (25). We obtain  $|\tilde{\mathcal{G}}_2| = 2y_1(1 + 10v_1 - 9v_2 + u(10 + 111v_1 - 101v_2)) - y_2(9 + 90v_1 - 81v_2 + u(91 + 1010v_1 - 919v_2))$ . To check the assumption of the second part of Theorem 3.2 that  $W_2[(W_2BD^- \mathbf{z})_x - (W_2B)_x D^- \mathbf{z}]P_0Q_1Q_2 = 0$ , we compute

$$W_2B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ v_1 - v_2 & -10v_1 + 9v_2 - 1 & y_1 - 10y_2 & -y_1 + 9y_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which we can write with constant vectors  $\mathbf{a}_i$ ,  $i = 1, \dots, 4$  and  $\mathbf{x} = (y_1, y_2, v_1, v_2, u, p_1, p_2)$  as

$$W_2B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{a}_1^T \mathbf{x} & \mathbf{a}_2^T \mathbf{x} - 1 & \mathbf{a}_3^T \mathbf{x} & \mathbf{a}_4^T \mathbf{x} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Because of the symmetry of the matrix  $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4)$  both expressions  $(W_2BD^{-\mathbf{z}})_x$  and  $(W_2B)_xD^{-\mathbf{z}}$  are equal for arbitrary  $\mathbf{z}$ . From Theorem 3.2 we obtain that, for the case  $p_1, p_2 > 0$  the index of the DAE is 3. The index-3 DAE for that case corresponds to the solution (35).

## 5 The numerical tests

The matrix chain (14) is calculated numerically (see [14]) by means of a Matlab code. We will use it to test the problems numerically. Numerically means that we check the properties of the DAEs which we have studied, pointwise for particular values of the variables  $(t(=\bar{t}), \mathbf{x})$ . The *tractability index* works with linearizations of the DAE along a function, with appropriate smoothness properties. Here in our experiments we linearize the DAE along a linear function through  $\mathbf{x}(\bar{t})$  with the derivative of every component equal to 1.

A hyphen (-) in the “index” column indicates that the code does not determine an index because the matrix chain could not be constructed (*tractability index* not defined). Under “remarks” we put the result of the algorithm and the last calculated matrix chain element, or we indicate to which part of the solution it corresponds.

Problem 1:

We use  $\mathbf{x}(\bar{t}) = (y_1, y_2, c, v_1, v_2, r, u, p_1, p_2)^T = (0, 0, 1, 1, 1, 0, 1, \pm 1, \pm 1)^T$ .

$p_1$	$p_2$	index	remarks
$> 0$	$> 0$	-	index not defined - $G_4$
$> 0$	$< 0$	1	refers to (27)
$< 0$	$> 0$	1	refers to (28)
$< 0$	$< 0$	-	index not defined - $G_1$

Problem 2:

We use  $\mathbf{x}(\bar{t}) = (y_1, y_2, c, v_1, v_2, r, u, p_1, p_2, p_3)^T = (0, 0, 1, 1, 1, 0, 1, \pm 1, \pm 1, \pm 1)^T$ .

$p_1$	$p_2$	$p_3$	index	remarks
$> 0$	$> 0$	$> 0$	-	index not defined - $G_4$
$> 0$	$< 0$	$> 0$	1	refers to (29)
$< 0$	$> 0$	$> 0$	1	refers to (31)
$< 0$	$< 0$	$> 0$	-	index not defined - $G_1$
$> 0$	$> 0$	$< 0$	2	refers to (30)
$> 0$	$< 0$	$< 0$	-	index not defined - $G_3$
$< 0$	$> 0$	$< 0$	-	index not defined - $G_3$
$< 0$	$< 0$	$< 0$	-	index not defined - $G_1$

Problem 3:

We use  $\mathbf{x}(\bar{t}) = (z_a, z_b, v_a, v_b, u, p_1, p_2)^T = (0, 0, 1, 1, 1, \pm 1, \pm 1)^T$ .

$p_1$	$p_2$	index	remarks
$> 0$	$> 0$	3	refers to (35)
$> 0$	$< 0$	1	refers to (34)
$< 0$	$> 0$	1	refers to (36)
$< 0$	$< 0$	-	index not defined - $G_0$



## 6 Conclusions

We have outlined here a procedure for transforming a general optimal control problem to a system of DAEs.

The *tractability index* concept presented here, provides an automatic tool for determining the index of a general system of DAEs theoretically and numerically. In this paper, the application to DAEs obtained from optimal control problems shows its potentiality to determine the index, and also the image projector  $W_{\mu-1}$  of an index- $\mu$  system of DAEs provides information on which equations need to be differentiated in order to reduce the index.

Furthermore, this procedure is an improvement over existing methods [3], since it does not increase the number of equations in the system.

The numerical algorithm used here to determine the index can be used for solving problems without the knowledge of an analytic solution.

This opens the door for solving optimal control problems in an alternative way.

In more applications, with a large number of variables and higher nonlinearities, as in the simulation of electrical networks, an investigation of the properties of the problem, such as has been done here in Section 4 is not yet practicable. However, the development of a numerical index monitor, such as the one presented in Section 5, is a first step.

**Acknowledgment:** We thank both the unknown referees for their helpful comments.

## References

- [1] L.T. Biegler, "Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation", *Comput. Chem. Eng.*, Vol. 8, pp 243-248, 1984.
- [2] T. Binder, A. Cruse, C.A. Cruz Villar and W. Marquardt, "Dynamic optimization using a wavelet based adaptive control vector parametrization strategy", *Comput. Chem. Eng.*, Vol. 24, pp 1201-1207, 2000.
- [3] K.E. Brenan, S.I. Campbell and L. R. Petzold, "Numerical solution of initial-value problems in differential-algebraic equations", SIAM, NY, 1989.
- [4] M.D. Canon, C.D. Cullum and E. Polak, "Theory of optimal control and mathematical programming", Mc.Graw-Hill, NY, 1970.
- [5] J.E. Cuthrell and L.T. Biegler, "On the optimization of differential-algebraic process systems", *AICHE J.* Vol.33, No. 8, pp 1257-1270, 1987.
- [6] R. England, S. Gómez, R. Lamour, "A study of the index of differential-algebraic equations for optimal control problems", *Reportes de Investigación, IIMAS-UNAM*, Vol. 12, no. 62, 2007.

- [7] R. England, S. Gómez, R. Lamour, "Expressing optimal control problems as differential algebraic equations", *Comput. Chem. Eng.*, Vol. 29, no. 8, pp 1720-1730, 2005.
- [8] D. Estévez Schwarz and R. Lamour, "The computation of consistent initial values for nonlinear index-2 differential-algebraic equations", *Numer. Algorithms*, Vol. 26, No. 1, pp 49-75, 2001.
- [9] W.F. Feehery and P.I. Barton, "Dynamic optimization with equality path constraints", *Ind. Eng. Chem. Res.*, Vol. 38, no. 6, pp 2350-2363, 1999.
- [10] C.W. Gear, B. Leimkuhler and G.K. Gupta, "Automatic integration of Euler-Lagrange equations with constraints", *J. Comp. Appl. Math.* Vol.12/13, pp 77-90, 1985.
- [11] I. Higuera, R. März, C. Tischendorf, "Stability preserving integration of index-1 DAE's", *Appl. Numer. Math.*, Vol. 45, pp 175-200, 2003.
- [12] I. Higuera, R. März, C. Tischendorf, "Stability preserving integration of index-2 DAEs", *Appl. Numer. Math.*, Vol. 45, pp 201-229, 2003.
- [13] R. Jackson, "Optimal use of mixed catalyst for two successive chemical reactions", *J. Optim. Theory Appl.*, Vol. 2, no. 1, 1968.
- [14] R. Lamour, "Index determination and calculation of consistent initial values for DAEs", *Comp. Math. with Appl.*, Vol. 50, pp 1125-1140, 2005.
- [15] J.S. Logsdon and L.T. Biegler, "Accurate solution of differential-algebraic optimization problems", *Int. Eng. Chem. Res.* Vol.28, No. 11, pp 1628-1639, 1989.
- [16] R. März, "Numerical methods for differential-algebraic equations", *Acta Numer.* 1992, pp. 141-198.
- [17] R. März, "The index of differential algebraic systems with properly stated leading term", *Results Math.*, Vol. 42, no 3/4, pp 308-338, 2002.
- [18] C.C. Pantelides, R.W.H. Sargent and V.S. Vassiliadis, "Optimal control of multistage systems described by high-index differential-algebraic equations", *Computational Optimal Control*, Eds. R. Bulirsch and D. Kraft, International Series of Numerical Mathematics, Birkhäuser-Verlag, Basel, Vol. 115, pp 177-191, 1994.
- [19] J.G. Rentro, A.M. Moshedi and O.A. Asbjornsen, "Simultaneous optimization and solution of systems described by differential/algebraic equations", *Comput. Chem. Eng.*, Vol. 11, pp 503-517, 1987.
- [20] S. Vasantharajan and L.T. Biegler, "Simultaneous strategies for the optimization of differential-algebraic systems with enforcement of error criteria", *Comput. Chem. Eng.*, Vol. 14, pp 1083-1100, 1990.
- [21] V.S. Vassiliadis, R.W.H. Sargent and C.C. Pantelides, "Solution of a class of multistage dynamic optimization problems. I. Problems without path constraints", *Ind. Eng. Chem. Res.*, Vol. 33, no. 9, pp 2111-2122, 1994.

- [22] V.S. Vassiliadis, R.W.H. Sargent and C.C. Pantelides, "Solution of a class of multistage dynamic optimization problems. II. Problems with path constraints", *Ind. Eng. Chem. Res.*, Vol. 33, no. 9, pp 2123-2133, 1994.

Mailing address:

Roland England, Dept. of Mathematics, The Open University,  
Milton Keynes, MK7 6AA Great Britain.

E-mail: [r.England@open.ac.uk](mailto:r.England@open.ac.uk)