Z angew Math Phys 43 (1992)
(ZAMP)

0044-2275/92/020319-09 \$ 1.50 + 0.20

On differential algebraic equations with discontinuities

By L. Brüll and U. Pallaske, Bayer AG, Abt. AV-IM-AM,
D-5090 Leverkusen, Germany

1. Introduction

In general, most of the mathematical models in chemical engineering are given in the form of a linear implicit differential algebraic equation (DAE) system:

$$
\begin{aligned}
&A(z, u) \cdot \dot{z} = f_1(z, u) \\
&0 = f_2(z, u) \qquad\qquad (1.1)\\
&z(0) = z_0 \\
&f_1\colon R^n \times R^l \to R^k \\
&f_1\colon R^n \times R^l \to R^{n-k}; \quad f_1, f_2 \text{ smooth} \\
&\operatorname{rg}(A(z, u)) = k.
\end{aligned}
$$

Here, z denotes the vector of state variables, while u contains the control or input variables which are used to control the chemical plant described by (1.1). In this paper we study the influence of discontinuous control variables on the solution of system (1.1). In chemical engineering the "Sprung-Antwort-Verhalten" of a chemical plant is often investigated in order to study the dynamic behaviour of the system. Further, in the simulation of chemical plants with controls, discontinuous control variables (i.e. bang-bang controls) occur. Let us, for example, have a look at the case of a distillation column which is controlled by the reflux ratio, i.e. the quotient of the outlet stream and the reflux stream at the top of the column. This ratio is closely connected with the purity of the outlet stream, which itself is often connected with the temperature of certain stages of the column. Hence, at discrete times, the temperature is measured and the new value of the reflux ratio is calculated.

Problems of this kind can usually be formulated by (1.1) with control variables u which are piecewise continuous. For simplicity, we will assume in the following that the input variables $u(t)$ are continuous (or even differentiable with continuous derivative) for $t < t_0$ and $t > t_0$, but with a jump at $t = t_0 > 0$, i.e. $u^+ := u(t_0 + 0) \neq u^- := u(t_0 - 0)$.

Now, what is the response of the state variable vector z to the jump in u at time t_0? For physical reasons the following definition seems to be appropriate. This definition is given for full implicit DAE systems of the form

$$f(\dot{z}, z, t, u) = 0, \quad z(0) = z_0. \tag{1.2}$$

Consider the set of regularizations on the interval $[t_0, t_0 + 1]$:

$$\mathscr{F} := \{\tilde{u} \in C([t_0, t_0 + 1], R^l) \cap C^1((t_0, t_0 + 1), R^l) \mid \tilde{u}(t_0) = u^-, \tilde{u}(t_0 + 1) = u^+\}.$$

Let $\tilde{u}_\varepsilon(t) = \tilde{u}(t_0 + (t - t_0)/\varepsilon)$, $\varepsilon > 0$. Further, denote by $z_\varepsilon(t, \tilde{u}_\varepsilon)$ the unique solution of the initial value problem

$$f(\dot{z}, z, t, \tilde{u}_\varepsilon) = 0$$

$$z(0) = z_0.$$

Definition. z^+ is called the *genuine initial value* (corresponding to $z^- := z(t_0 - 0)$ and the jump of u at t_0) iff:

the limit

$$z^+ = \lim_{\varepsilon \to 0} z_\varepsilon(t_0 + \varepsilon, \tilde{u}_\varepsilon)$$

exists for all regularizations $\tilde{u} \in \mathscr{F}$ of u and does not depend on $\tilde{u} \in \mathscr{F}$.

In many applications $z^+ = z^-$ is assumed. This assumption is correct in the treatment of ordinary differential equations with discontinuous control variables (see [2]), but not in the case of DAE systems (even in the index-1 case). This is evident because of the example of a pure system of algebraic equations. Moreover, for the index-1 example (see [2])

$$\dot{z}_1 + z_3 \cdot \dot{z}_2 = 0$$

$$z_2 = u_1(t)$$

$$z_3 = u_2(t).$$

the genuine initial value generally does not even exist. Take, for example, $z_1(t) = z_2(t) = z_3(t) = u_1(t) = u_2(t) = 0$ for $t < 0$ and $u_1(t) = u_2(t) = 1$ for $t \geq 0$. Suppose $\tilde{u}(t)$ is an arbitrary function (differentiable on $(0, 1)$, $\tilde{u}(0) = 0$, $\tilde{u}(1) = 1$) and consider the regularizations $\tilde{u}_1(t) = \tilde{u}(t)$, $\tilde{u}_2(t) = [\tilde{u}(t)]^n$, $n \in N$. Then:

$$\lim_{\varepsilon \to 0} z_1(\varepsilon) = -1/(n + 1).$$

In this paper we will characterize those semi-linear DAE systems (1.1) for which the genuine initial value z^+ always exists. After some introductory remarks in Section II, we will transform the DAE system (1.1) into an ordinary differential equation which is easier to handle. Necessary and sufficient conditions (Section III) for the existence of z^+ are then determined. We will see that these conditions are a natural generalization of the well known integrability conditions in the context with curve integrals. Finally, in Section IV, we will apply the results to DAE systems and state a sufficient criterion which is easier to check for special problems. Last, but not least, a simple numerical method is provided to calculate z^+.

We will assume in the whole paper that (1.1) has global (or differential) index 1 (see [3]). Further it is assumed that the solution $z(t)$ exists globally for bounded continuous control variables u and depends at least twice continuously differentiable on u. It will become clear that these assumptions, as well as the assumptions about the region of definition of $A(z, u)$, f_1 and f_2, can be generalized (i.e. Theorem 1 for star-shaped regions etc.) in the well known manner. Finally, we put $t_0 = 0$ for simplicity.

II. Transformation into an ODE system

First we note that the existence of the genuine initial value z^+ of system (1.1) is not influenced by the right-hand side $f_1(z, u)$ of the differential part of (1.1) (except for the fact that the global solvability and continuous dependence upon u has to be guaranteed). To see this, consider an arbitrary regularization $\tilde{u} \in \mathscr{F}$, $\varepsilon > 0$ and the corresponding solution $z_\varepsilon(t, \tilde{u}_\varepsilon)$, $t \in [0, \varepsilon]$ of (1.1). Then

$$z(t, \varepsilon) := z_\varepsilon(\varepsilon t, \tilde{u}_\varepsilon)$$

solves the system

$$\begin{aligned} A(z, \tilde{u}) \cdot \dot{z} &= \varepsilon \cdot f_1(z, \tilde{u}), \\ 0 &= f_2(z, \tilde{u}), \end{aligned} \qquad z(0) = z^-$$

for $t \in [0, 1]$ and

$$z^+ = \lim_{\varepsilon \to 0} z(1, \varepsilon)$$

holds. Therefore, the assertion follows from the theorem on continuous dependence of the solution on parameters.

It is well known that the transformation of a DAE system with the Gear-Petzold algorithm (see [4]) into a pure ODE system leads to tremendous numerical problems, due to the fact that the number of degrees of freedom is increased (see [1]). However, for consistent initial values the new system is analytically equivalent to the original one and therefore it can be used to describe analytical conditions for the existence of the genuine initial value z^+.

Using this transformation (i.e. differentiating the algebraic part of (1.1) and solving the resulting set of equations for z), the original DAE system (1.1) with $f_1(z, u) = 0$ is transformed into

$$\dot{z} = F(z, u) \cdot \dot{u}, \quad z(0) = z^-$$

where

$$F(z, u) := M^{-1} \cdot \begin{pmatrix} 0 \\ -\frac{\partial f_2}{\partial u}(z, u) \end{pmatrix} \tag{2.1}$$

and

$$M := \begin{pmatrix} A(z, u) \\ \frac{\partial f_2}{\partial z}(z, u) \end{pmatrix}. \tag{2.2}$$

We have proved:

Lemma 1. The following conditions are equivalent:

i) the genuine initial value z^+ exists for

$$\begin{aligned} A(z, u) \cdot \dot{z} &= f_1(z, u) \\ 0 &= f_2(z, u), \quad z(0) = z^- \end{aligned} \tag{2.3}$$

ii) the genuine initial value z^+ exists for

$$\dot{z} = F(z, u) \cdot \dot{u}, \quad z(0) = z^-. \tag{2.4}$$

III. Necessary and sufficient conditions

In the following, $F_{i,j}(z, u)$ denotes the element in row i and column j of the matrix $F(z, u)$, $i = 1, \ldots, n$; $j = 1, \ldots, l$. The main result in this paper is as follows.

Theorem 1. For the differential equation

$$\dot{z} = F(z, u) \cdot \dot{u} \tag{3.1}$$

with fixed $z(0)$ and $u(0)$, the solution $z(1)$ only depends on $u(1)$ if and only if the integrability conditions

$$\sum_{\nu=1}^{n} \frac{\partial F_{i,j}}{\partial z_\nu} \cdot F_{\nu,\mu} + \frac{\partial F_{i,j}}{\partial u_\mu} = \sum_{\nu=1}^{n} \frac{\partial F_{i,\mu}}{\partial z_\nu} \cdot F_{\nu,j} + \frac{\partial F_{i,\mu}}{\partial u_j} \tag{3.2}$$

are fulfilled for

$$i = 1, \ldots, n; \quad j, \mu = 1, \ldots, l.$$

In particular, the genuine initial value z^+ of (3.1) always exists if and only if (3.2) holds.

Proof.

I. Necessity:

Consider fixed values for z^- and u^-.

Let us assume that for each u^+ a unique value of z^+ exists. Then there is a function H with

$$z^+ = H(u^+).$$

Because of the invariance of the solution of $\dot{z} = F(z, u) \cdot \dot{u}$ with respect to the transformation $t \to \xi(t)$, $\dot{\xi}(t) \neq 0$ we have

$$z(t) = H(u(t)),$$

where $z(t)$ is the solution of (3.1) with arbitrarily given $u(t)$, $u(0) = u^-$. Hence, for differentiable $u(t)$ we have

$$\dot{z}(t) = \frac{\partial H}{\partial u} \cdot \dot{u}(t).$$

Since $\dot{u}$ can be chosen arbitrarily, we find

$$\frac{\partial H}{\partial u} = F(z, u) = F(H(u), u). \tag{3.3}$$

Let us look at the components H_i of $H = (H_1, \ldots, H_n)^T$. From the general assumptions we conclude:

$$\frac{\partial^2 H_i}{\partial u_j \cdot \partial u_k} = \frac{\partial^2 H_i}{\partial u_k \cdot \partial u_j} \tag{3.4}$$

for $i = 1, \ldots, n$; $k, j = 1, \ldots, l$. Applying (3.4) to (3.3) yields the condition (3.2).

II. Sufficiency:

The proof of the sufficiency is somewhat more complicated. To begin with, consider for arbitrarily given $u \in R^l$ the special continuous regularization

$$u(t) := u^- + t \cdot (u - u^-).$$

$z(t)$ denotes the corresponding solution of problem (3.1), with $z(0) = z^-$. Further, assume that (3.2) is fulfilled. Then we have

Lemma 2.

$$\frac{\partial z_\nu}{\partial u_k}(t) = F_{\nu,k}(z(t), u(t)) \cdot t \tag{3.4}$$

$\nu = 1, \ldots, n; \quad k = 1, \ldots, l.$

Proof of the lemma. Consider the linear ordinary differential equation

$$\frac{d}{dt} w_{\nu,k}(t) = \sum_{\kappa=1}^{l} \left(\sum_{\mu=1}^{n} \frac{\partial F_{\nu,\kappa}}{\partial z_\mu} \cdot w_{\mu,k}(t) + \frac{\partial F_{\nu,\kappa}}{\partial u_k} \cdot t \right) \cdot (u_\kappa - u_{0,\kappa}) + F_{\nu,k}$$

with $\nu = 1, \ldots, n$; $k = 1, \ldots, l$. It is easy to show that both $\partial z_\nu / \partial u_k(t)$ and $F_{\nu,k}(z(t), u(t)) \cdot t$ are solutions of the linear ODE system with initial condition $w_{\nu,k}(0) = 0$ (where the intergrability conditions for the second function have to be used). Hence, Lemma 2 is proved.

Now consider for arbitrary $u = (u_1, \ldots, u_l)^T \in R^l$

$$H_i(u) := z_i^- + \int_0^1 \sum_{\kappa=1}^{l} F_{i,\kappa}(z(t), u(t)) \cdot (u_\kappa - u_{0,\kappa})\, dt$$

$$H := (H_1, \ldots, H_n)^T \colon R^l \to R^n.$$

Then we have:

i) $z(1) = H(u)$

ii) $H(u^-) = z^-$

iii) $$\frac{\partial H_i}{\partial u_k} = \int_0^1 \left[\sum_{\kappa=1}^{l} \left(\sum_{\mu=1}^{n} \frac{\partial F_{i,\kappa}}{\partial z_\mu} \cdot \frac{\partial z_\mu}{\partial u_k} + \frac{\partial F_{i,\kappa}}{\partial u_k} \cdot t \right) \cdot (u_\kappa - u_{0,\kappa}) + F_{i,k} \right] dt$$
$$= \int_0^1 \left[\sum_{\kappa=1}^{l} \left(\sum_{\mu=1}^{n} \frac{\partial F_{i,\kappa}}{\partial z_\mu} \cdot F_{\mu,k} + \frac{\partial F_{i,\kappa}}{\partial u_k} \right) \cdot t \cdot (u_\kappa - u_{0,\kappa}) + F_{i,k} \right] dt.$$

Using the integrability condition (3.2) we get

$$\frac{\partial H_i}{\partial u_k} = \int_0^1 \left[\sum_{\kappa=1}^{l} \left(\sum_{\mu=1}^{n} \frac{\partial F_{i,k}}{\partial z_\mu} \cdot F_{\mu,\kappa} + \frac{\partial F_{i,k}}{\partial u_\kappa} \right) \cdot t \cdot (u_\kappa - u_{0,\kappa}) + F_{i,k} \right] dt$$
$$= \int_0^1 \frac{d}{dt} [F_{i,k}(z(t), u(t)) \cdot t]\, dt$$
$$= F_{i,k}(z(1), u(1)) = F_{i,k}(H(u), u).$$

Hence there is a differentiable function $H: R^l \to R^n$ such that

$$\frac{\partial H}{\partial u}(u) = F(H(u), u)$$

holds.

Now consider an arbitrary function $u(t): R_0^+ \to R^l$, with $u(0) = u^-$. Look at the function

$$z(t) := H(u(t)).$$

Because of ii) we have $z(0) = z^-$.

Further,

$$\dot{z}(t) = \frac{\partial H}{\partial u} \cdot \dot{u} = F(H(u), u) \cdot \dot{u} = F(z, u) \cdot \dot{u}$$

shows that $z(t)$ is a solution of the initial value problem (3.1). Since our DAE system is uniquely solvable, the theorem is proved.

IV. The genuine initial value and DAE systems

Using the transformation (2.1), (2.2), it is evident that the described necessary and sufficient conditions can be translated for linear implicit DAE systems

$$\begin{aligned} A(z, u) \cdot \dot{z} &= f_1(z, u) \\ 0 &= f_2(z, u) \qquad (4.1) \\ z(0) &= z_0. \end{aligned}$$

However, these conditions involve the inverse of the matrices A and $\partial f_2/\partial z$ and hence will not be useful for many applications. Therefore, we will state a sufficient condition for the existence of z^+, which is often much easier to verify. For example, this condition is fulfilled in most applications in chemical engineering. This result is proved in [2].

Theorem 2. Consider the linear implicit DAE system (4.1) and assume that f_1, f_2 are bounded, differentiable functions with continuous derivatives which are Lipschitz-continuous with respect to z. Assume further that for each row $a_i(z)$, $i = 1, \ldots, k$ there exists a non-zero continuously differentiable function $g_i(z)$: $R^n \to R$, such that $g_i(z) \cdot a_i(z)$ has a potential $p_i(z)$ (i.e. $d/dt(p_i(z)) = g_i(z) \cdot a_i(z) \cdot (d/dt)z)$ $(i = 1, \ldots, k)$. Then z^+ exists and fulfills

$$\begin{aligned} &p_i(z^+) = p_i(z^-), \quad i = 1, \ldots, k \\ &f_2(z^+, u^+) = 0. \end{aligned}$$

Besides the existence of z^+ for systems (4.1) with integrating factors, the theorem gives a simple method of evaluating z^+. Unfortunately, since the potentials $p_i(z)$ have to be known, this method will not be of use in many applications. In the following, a much easier way to calculate z^+ is given by Theorem 3. However, this theorem is not only applicable in the case when integrating factors exists. Further, it can be generalized for a lot of nonlinear DAE systems of the form (1.2).

Theorem 3. Consider the initial value problem (4.1) and assume that $u(t)$: $R \to R^l$ is an arbitrary function, continuous for $t < 0$ and $t > 0$ with a discontinuity at $t = 0$. If the genuine initial value z^+ exists, then we have

$$z^+ = z(1)$$

where $z(t)$ is the solution of the initial value problem

$$\begin{aligned} A(z, \tilde{u}) \cdot \dot{z} &= 0 \\ 0 &= f_2(z, \tilde{u}) \\ z(0) &= z^-, \quad t \in [0, 1] \end{aligned} \tag{1.1}$$

and $\tilde{u}(t)$ is a continuous function for $t \in [0, 1]$ with $\tilde{u}(0) = u^-$, $\tilde{u}(1) = u^+$.

The proof can be carried out by analogy with the first remarks of Section II.

Acknowledgement

We thank the referee for several useful remarks, which have improved the presentation of this paper.

References

[1] R. Bachmann, L. Brüll, Th. Mrziglod, and U. Pallaske, *A contribution to the numerical treatment of differential algebraic equations arising in chemical engineering*. Dechema-Monographs, Vol. 116 (1989).
[2] L. Brüll and U. Pallaske, *On Consistent Initialization of Differential-Algebraic Equations with Discontinuities*. Proc. Fourth Conf. on Math. in Ind. (1991) (Hj. Wacker and W. Zulehner eds.).
[3] C. W. Gear, *Differential algebraic equations, indices, and integral algebraic equations*. Report No. UIUCDCS-R-89-1505, Uni. Illinois (1989).
[4] C. W. Gear and L. Petzold, *ODE methods for the solution of differential/algebraic systems*. SIAM J. Numer. Anal. *21* (1984).

Summary

The dynamic features of systems in mathematical process simulation are often studied by looking at the response of the system to discontinuities in the input variables (Sprung-Antwort-Verhalten). A more detailed analysis ([2]) shows that for systems in the form of differential algebraic equations (even with index 1), which frequently occur e.g. in chemical engineering, a solution to the problem need not exist.

In this paper we derive necessary and sufficient conditions for such systems to guarantee the solvability of the problem (Theorem 1). Further, a simple algorithm is stated (Theorem 3), which is suitable for numerical computation.

Zusammenfassung

Die Untersuchung der dynamischen Eigenschaften von Systemen mit Hilfe der mathematischen Prozeßsimulation geschieht häufig durch die Betrachtung des Systemverhaltens bei sprungartiger Veränderung der Eingangsgrößen (Sprung-Antwort-Verhalten). Eine genauere Analyse ([2]) zeigt, daß für Systeme in Form von Differential-Algebraischen Gleichungssystemen (sogar mit Index 1), welche häufig auftreten z.B. in der chemischen Verfahrenstechnik, eine Lösung dieser Aufgabe nicht existieren muß.

In dieser Arbeit leiten wir für Systeme vom Index 1 notwendige und hinreichende Bedingungen her, welche die Lösbarkeit des Problems garantieren (Theorem 1). Ferner wird ein einfacher Algorithmus zur numerischen Berechnung beschrieben (Theorem 3).

(Received: April 17, 1991; revised: July 8, 1991)