Some new aspects of the convergence of dynamic iteration methods for coupled systems of DAEs

Falk Ebert
Berlin, March 20, 2007

Abstract The purpose of this paper is the analysis of dynamic iteration methods for the numerical integration of coupled systems of ODEs and DAEs. We will investigate convergence of these methods and put special emphasis on the Jacobi- and Gauss-Seidel methods. Furthermore, the fundamental difference in the convergence behaviour of coupled ODEs and DAEs is pointed out. This difference is used to explain why certain relaxation methods for coupled DAEs may fail. Finally, a remedy to this undesirable effect is proposed that makes use of a so-called preconditioned dynamic iteration strategy. This regularization also allows significant reduction of dynamic iteration steps.

AMS Subject Classification: 65L20, 65L80
Keywords: coupled systems, differential-algebraic equations, dynamical iteration, waveform relaxation methods

1 Introduction

Today, a wide range of physical effects, technical systems and even social and economical interactions are well understood and often modelled mathematically. In many cases this description of dynamics leads to differential equations. Taking eventual constraints into account, one ends up with differential-algebraic equations. Over time, many methods have been developed to cope with the properties of the equations from the above mentioned areas. Sometimes they are highly specialized in order to meet the requirements for a rather small set of equations. What happens, when two types of equations are interconnected? This is exactly the case of coupled systems. Most of today’s technical systems are composed of numerous different components, a fact that quite naturally divides them into inter-coupled subsystems. Each of these represents some aspects of the behaviour of the whole system and is thus best treated separately. Every subsystem usually comes with its own theory, equations and solvers. Instead of solving all systems at once with one solver, it has become a popular approach to solve coupled systems with coupled simulators, [2,9,10,27,31,32]. The main idea is to solve every subsystem on a small time window with an adequate solver. The solutions of the other subsystems have to be approximated. This may be realized with the help of low order reduced models for the other systems, see [28] or by extrapolating the solution of the last time window. Once every component has been dealt with, naturally the solution of one subsystem differs from the approximation that all the others have been using in their calculation. Hence, the whole process is iterated and every system now uses the most recent solutions of all other systems. This dynamic iteration process also known as waveform relaxation is repeated until some sort of convergence has been achieved, cf. [6,21,26]. Then, one proceeds to the next time-window and the whole process is repeated.

For the involved DAEs, we focus on coupled systems of quasi-linear structure, i.e., linear with respect to
the differentiated variables, interconnected in a rather general way.

\[ M_i(t)x_i(t) = f_i(t, x), \quad i = 1, \ldots, s \]
\[ x(t_0) = x^{(0)} \]
\[ x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_s(t) \end{bmatrix} \in \mathbb{R}^{n_1 + \ldots + n_s}, \quad t \in [t_0, t_0 + T] \]
\[ M_i(t) \in \mathbb{R}^{n_i}, \quad t \in [t_0, t_0 + T] \]
\[ f_i : \mathbb{R}^{n_i} \times [t_0, t_0 + T] \mapsto \mathbb{R}^{n_i} \]

This formulation is not the most general one, but it appears sufficient to treat some of the most common types of coupled differential equations such as multi-body systems, [3, 13, 15, 29], and lumped circuit equations, [4, 14, 17]. More precisely, for the convergence analysis, we will only consider semi-explicit DAEs. This greatly simplifies the notation and is not restrictive, as every DAE of the form (1) can be transformed into an analytically equivalent semi-explicit DAE.

A characteristic quantity for DAEs is the so-called index. While there are many different index definitions, see for instance [5, 16, 18, 22, 24], in this paper we will be using the differentiation index or d-index introduced in [8]. The d-index is the least number of times a DAE has to be differentiated in order to extract an equivalent ordinary differential equation, the so-called underlying ODE by algebraic means only. If one such differentiation is not sufficient, then the system contains so called hidden constraints and is said to be of higher index. Here, the term index will always correspond to the differentiation index and we only consider DAEs of d-index 1.

In this paper, we will develop criteria for the convergence of coupled systems of ODEs and DAEs. In order to facilitate the convergence analysis we will study error propagation with the help of a linearized error equation. This will reduce the investigation to the consideration of linear time-variant ODEs and DAEs. According to [7], this time-variant linearization of the original ODE or DAE proves to be appropriate in most applications. Let

\[ y' = f(t, y), \quad y(t_0) = y_0 \]
\[ y(t) \in \mathbb{R}^N, \quad t \in [t_0, t_0 + T] \]

with \( f \) sufficiently smooth, be an ordinary differential equation to be solved on the interval \([t_0, t_0 + T]\). Assume that \( \tilde{y} \) is some reference trajectory close to \( y \). Then, if we set \( y = \tilde{y} + e \), we can linearize (2) as

\[ (\tilde{y} + e)' = f(t, \tilde{y}) + f_y(t, \tilde{y})e + \mathcal{O}(e^2). \]

If we choose this reference trajectory \( \tilde{y} \) such that it is the correct solution of (2) and omit the higher order terms, then we have

\[ e' = f'(t, y)e. \]

If at the starting point \( t_0 \), the solution \( \tilde{y} \) of the nonlinear ODE (2) and the solution \( y \) of the linearized ODE (3) coincide, then the difference \( e(t_0) = y(t_0) - \tilde{y}(t_0) = 0 \). As a consequence, according to the error ODE (4), the predicted error is zero for all times.

In the DAE case the same approach for

\[ M(t)x' = f(t, x), \quad x(t_0) = x_0, \quad x(t) \in \mathbb{R}^N, \quad t \in [t_0, t_0 + T], \quad M(t) \in \mathbb{R}^{N,N}, \quad t \in [t_0, t_0 + T]. \]

yields

\[ M(t)e' = f_x(t, x)e. \]
With the consistent initial value \( e(t_0) = 0 \), the DAE (6) yields \( e(t) = 0 \) for all times. Accordingly, we can expect the actual difference \( x - \bar{x} \) to be small.

Both linearizations (4) and (6) can be regarded as approximations for the error development of the respective ODE (2) or DAE (5). In Section 2 we will use these linearizations for the study of the error development in coupled systems, where the respective reference trajectory will be the correct solution of the coupled system. Assuming a Lipschitz condition and the boundedness of \( f_y \) and \( f_x \), respectively, and starting the dynamic iteration process with a sufficiently close initial approximation to the correct solution, we can use the linearized error behaviour to approximately predict the actual error.

Another important tool for the analysis of DAEs is the so-called underlying ordinary differential equation. This ODE can be obtained by differentiating the DAE \( \nu \) times, where \( \nu \) is the differentiation index of the DAE, and then algebraically extracting an explicit ODE from the DAE and all derivatives of level up to \( \nu \). While this equation is not suited for numerical integration, see [5, 19], it can be used for the analysis of the DAE. For a quasi-linear DAE (5) of d-index 1, the underlying ODE can easily be determined. Let \( P^T \) and \( Q^T \) be matrices spanning the range and the co-range of \( M \) respectively. Then

\[
\begin{bmatrix}
P \\
Q
\end{bmatrix} M(t)x' = \begin{bmatrix}
\hat{M} \\
0
\end{bmatrix} x' = \begin{bmatrix}
P \\
Q
\end{bmatrix} f(t, x).
\]

separates differential and algebraic equations. Differentiation of the algebraic part yields

\[
\begin{bmatrix}
\hat{M} \\
-Qf_x(t, x)
\end{bmatrix} x' = \begin{bmatrix}
P f(t, x) \\
Q' f(t, x) + Qf_x(t, x)
\end{bmatrix}.
\]

In the d-index 1 case, the matrix \([\hat{M}^T -(Qf_x(t, x))^T] \) is nonsingular, see [30], and the implicit ODE can readily be transformed into an explicit one.

This paper is organized as follows. In Section 2 we will first give a brief introduction on Jacobi- and Gauss-Seidel methods for dynamic iteration. Known convergence results will be presented. We will show that these results do not hold anymore in the case of coupled DAEs. In Section 3 we will derive a regularization method to overcome convergence problems for DAEs in the Gauss-Seidel- and Jacobi dynamic iteration methods.

2 The Jacobi and Gauss-Seidel methods

Together with other splitting methods, the Gauss-Seidel and the Jacobi method have first been used for the iterative solution of linear systems. Later, this splitting technique was applied to large scale dynamical systems, see [23]. Together with similar methods, such as the Picard-Lindelöf iteration, they have also been studied in [6, 25, 26]. Later, they have been applied to differential-algebraic systems, cf., [2, 21]. When talking about the Gauss-Seidel and the Jacobi methods, in the coupled system context, we always mean block Gauss-Seidel or block Jacobi methods, with splitting between different systems.

We will subsequently only consider autonomous systems, i.e., systems that do not explicitly depend on time. This can be justified by the fact that \( t' = 1 \) can always be added to the differential equations, making \( t \) a system variable and the system itself autonomous.

Consider two systems of ordinary differential equations

\[
y_1' = f_1(y_1, y_2), \\
y_2' = f_2(y_1, y_2), \\
y_1(t_0) = y_{10}, \quad y_2(t_0) = y_{20},
\]

(7) \hspace{2cm} (8)

where \( y_i(t) \in \mathbb{R}^{N_i}, t \in [t_0, t_0 + T], i = 1, 2 \) and \( f_1: \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \mapsto \mathbb{R}^{N_1}, f_2: \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \mapsto \mathbb{R}^{N_2} \).
If the subsystems are not integrated together, then data from the other systems has to come as a control into each subsystem. The two most commonly used ways to implement this are the Jacobi method, (9) and the Gauss-Seidel method (10).

\[
\begin{align*}
(y_1^{(n)})' &= f_1(y_1^{(n)}, y_2^{(n-1)}), \\
(y_2^{(n)})' &= f_2(y_1^{(n-1)}, y_2^{(n)}).
\end{align*}
\]

\[
\begin{align*}
(y_1^{(n)})' &= f_1(y_1^{(n)}, y_2^{(n-1)}), \\
(y_2^{(n)})' &= f_2(y_1^{(n)}, y_2^{(n)}).
\end{align*}
\]

Figure 1: block diagram of the Jacobi method (left) and the Gauss-Seidel method (right)

In the dynamic iteration method, in every subsystem, only a subset of variables is determined on a time interval. After a subsystem has been solved, data is exchanged between the subsystems. The Jacobi-method has the advantage that all subsystems can be solved in parallel, while for the Gauss-Seidel method the systems are computed sequentially, always using the most recent solutions of the other subsystems. This approach makes parallelization more difficult, but tends to converge faster than the Jacobi method.

There are numerous ways of choosing and combining previous iterates, see [2, 26]. As long as only ODEs are involved, it has been proven in [26] that convergence can always be achieved, provided that the time window \([t_0, t_0 + T]\) is small enough. Going further, in [12] it has been shown that analytically, the method converges to the correct solution, provided that the solution of the complete system exists on the interval in question.

**Lemma 2.1.** Suppose that the differential equation

\[
\begin{align*}
y' &= f(y), \\
y(t_0) &= y_0
\end{align*}
\]

is uniquely solvable on \([t_0, t_0 + T]\).

Let an iteration scheme be defined by

\[
\begin{align*}
(y^{(n)})' &= \sum_{j=1}^{l} B_{j} f \left( \sum_{i=0}^{k} A_{ij}(t) y^{(n-i)} \right), \\
y^{(n)}(t_0) &= y_0,
\end{align*}
\]

where

\[
y^{(n)} \in \mathcal{C}^1([t_0, t_0 + T]; \mathbb{R}^N) \quad \forall n = 0, 1, 2, \ldots,
\]
2.1 Error propagation in dynamic iteration methods

with starting iterates \( y^{(0)}_1, \ldots, y^{(k-1)}_1 \).

Suppose that the matrix functions \( A_{ij}, i = 1, \ldots, l; j = 0, \ldots, k \) are continuous and satisfy

\[
\sum_{i=0}^{k} A_{ij}(t) = I \quad \forall t \in [t_0, t_0 + T], \forall j = 1, \ldots, l,
\]

and that the matrices \( B_j \) are constant and satisfy \( \sum_{j=1}^{l} B_j = I \). Then, the limit \( y = \lim_{n \to \infty} y^{(n)} \) exists and fulfills \( y' = f(y) \).

**Proof:** For a proof, we refer to [11, 12]. \( \square \)

2.1 Error propagation in dynamic iteration methods

In this section we will study the propagation of errors between two steps of a dynamical iteration method. As the solution of one system also depends on the solutions of the other subsystems, errors in these subsystems may affect all others. In this paper we will always assume that the exact solution of a given ODE or DAE is available and, thus, we will only study errors arising from the dynamic iteration methods.

We will derive differential and differential-algebraic difference equations that are similar in structure to the original dynamical iteration method. These difference equations produce sequences of functions representing the error of one stage of the dynamic iteration method with respect to the solution of the coupled system. In the case of a convergent method, this sequence of errors will converge to zero, uniformly. We will see, that this is not necessarily the case for DAEs.

2.1.1 The ODE case

We consider a system of two coupled ODEs as in (7)

\[
y'_1 = f_1(y_1, y_2), \quad y'_2 = f_2(y_1, y_2),
\]

\[
y_1(t_0) = y_{10}, \quad y_2(t_0) = y_{20}.
\]

As before, we can derive a linearized error equation as

\[
(e_1^{(n)})' = f_1 y_1 y_1^{(n)} + f_1 y_2 y_2^{(n)} e_1^{(n)}, \quad (14a)
\]

\[
(e_2^{(n)})' = f_2 y_1 y_1^{(n)} + f_2 y_2 y_2^{(n)} e_2^{(n)}. \quad (14b)
\]

As we have already seen, if the coupled system is embedded into a dynamic iteration method then it is possible that system 1 does not have access to the most recent solution of system 2. The same can be the case for system 2. Thus, if the two systems were coupled by the GAUSS-SEIDEL scheme and we assumed system 1 to be computed first, in (14a) we only had access to the last iterate \( y_2^{(n-1)} \) of the second system and thus (14) would become

\[
(e_1^{(n)})' = f_1 y_1 y_1^{(n)} + f_1 y_2 y_2^{(n)} e_1^{(n)}, \quad (14a)
\]

\[
(e_2^{(n)})' = f_2 y_1 y_1^{(n)} + f_2 y_2 y_2^{(n)} e_2^{(n)}. \quad (14b)
\]

For a shorter notation, we define a vector \( E^{(n)} = \begin{bmatrix} e_1^{(n)} \\ e_2^{(n)} \end{bmatrix} \). Additionally, we assume that \( f_1 \) and \( f_2 \) are sufficiently smooth such that \( f_1(y_1^{(n)}, y_2^{(n-1)}) \approx f_1(y_1^{(n)}, y_2^{(n)}) \) and \( f_2(y_1^{(n-1)}, y_2^{(n)}) \approx f_1(y_1^{(n)}, y_2^{(n)}) \). This
can be guaranteed, as long as the interval \([t_0, t_0 + T]\) is chosen sufficiently small. Hence, the errors in a Gauss-Seidel scheme can be approximated by a differential difference equation

\[
\left( E^{(n)} \right)' = \left[ \begin{array}{cccc} f_{1,y_1} & 0 & 0 & f_{1,y_2} \\ f_{2,y_1} & f_{2,y_2} & 0 & 0 \\ 0 & f_{1,y_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] E^{(n)} + \left[ \begin{array}{c} 0 \\ f_{2,y_2} \end{array} \right] E^{(n-1)},
\]

(15a)

where

\[
\left[ \begin{array}{cccc} f_{1,y_1} & 0 & 0 & f_{1,y_2} \\ f_{2,y_1} & f_{2,y_2} & 0 & 0 \\ 0 & f_{1,y_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] = \left[ \begin{array}{cccc} f_{1,y_1}(y_1^{(n)}, y_2^{(n)}) & 0 & 0 & f_{1,y_2}(y_1^{(n)}, y_2^{(n)}) \\ f_{2,y_1}(y_1^{(n)}, y_2^{(n)}) & f_{2,y_2}(y_1^{(n)}, y_2^{(n)}) & 0 & 0 \\ 0 & 0 & f_{1,y_2}(y_1^{(n)}, y_2^{(n)}) & 0 \\ 0 & 0 & 0 & 0 \end{array} \right].
\]

For a Jacobi type coupling method, the error propagation differential difference equation has the form

\[
\left( E^{(n)} \right)' = \left[ \begin{array}{cccc} f_{1,y_1} & 0 & 0 & f_{1,y_2} \\ 0 & f_{2,y_2} & 0 & 0 \\ 0 & f_{1,y_2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] E^{(n)} + \left[ \begin{array}{c} 0 \\ f_{2,y_2} \end{array} \right] E^{(n-1)}.
\]

(15b)

From Theorem 2.1, we know that both iteration methods (15a) and (15b) converge towards the solution of the differential equation

\[
E' = \left[ \begin{array}{cccc} f_{1,y_1} & f_{1,y_2} \\ f_{2,y_1} & f_{2,y_2} \end{array} \right] E,
\]

which approximates the error of the complete coupled system without any dynamic iteration method, and which in case of \(E(t_0) = 0\) is zero for all times.

All of the mentioned systems have the form of a differential difference equation

\[
\left( E^{(n)} \right)' = F^{(0)} E^{(n)} + F^{(1)} E^{(n-1)}.
\]

(16)

### 2.1.2 Semi-explicit DAEs of d-index 1

Even when dealing with differential-algebraic equations of d-index 1, compared to the ODE case, unexpected problems may arise. Usually, a problem of d-index 1 can be treated similarly to a stiff ODE and handled by many implicit solvers such as DASSL or RADAU5, see [5], [19]. The difficulties arise when two or more systems of this type are coupled by a dynamic iteration method. Consider the case of two semi-explicit systems

\[
\begin{align*}
y_1' &= f_1(y_1, z_1, y_2, z_2), \\
0 &= g_1(y_1, z_1, y_2, z_2),
\end{align*}
\]

\[
\begin{align*}
y_2' &= f_2(y_1, z_1, y_2, z_2), \\
0 &= g_2(y_1, z_1, y_2, z_2),
\end{align*}
\]

(17)

with consistent initial values

\[
\begin{align*}
y_1(t_0) &= y_{10}, \\
z_1(t_0) &= z_{10}, \\
y_2(t_0) &= y_{20}, \\
z_2(t_0) &= z_{20}.
\end{align*}
\]

Let us assume that the functions \(f_j, g_j, j = 1, 2\) are sufficiently smooth and that the solutions \(y_j, z_j\) are uniquely defined by (17). We define by \(u_j\) and \(v_j\) the numerical approximations to \(y_j\) and \(z_j\), respectively.
2.1 Error propagation in dynamic iteration methods

The errors in the differential components $e_y$ are defined by $e_{y_j} = u_j - y_j$ and for the algebraic variables
by $e_{z_j} = v_j - z_j$. For the coupled system, with (6) we obtain the truncated error equation

$$\begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{z_1} \\
e_{y_2} \\
e_{z_2}
\end{bmatrix}
= 
\begin{bmatrix}
f_{1,y_1} & f_{1,y_2} & f_{1,z_1} & f_{1,z_2} \\
0 & g_{2,y_1} & g_{2,z_1} & g_{2,z_2}
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{z_1} \\
e_{y_2} \\
e_{z_2}
\end{bmatrix}.
$$

Changing the order of the equations and collecting the variables, one can write this DAE as the semi-explicit system

$$\begin{bmatrix}
I & I \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}
= 
\begin{bmatrix}
f_{1,y_1} & f_{1,y_2} & f_{1,z_1} & f_{1,z_2} \\
0 & g_{2,y_1} & g_{2,z_1} & g_{2,z_2}
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}.
$$

By introducing

$$E_y = \begin{bmatrix} e_{y_1} \\ e_{y_2} \end{bmatrix},$$

$$E_z = \begin{bmatrix} e_{z_1} \\ e_{z_2} \end{bmatrix},$$

and dividing the Jacobian into four sub-matrices

$$\begin{bmatrix}
f_{1,y_1} & f_{1,y_2} \\
f_{2,y_1} & f_{2,y_2}
\end{bmatrix}
\begin{bmatrix}
f_{1,z_1} & f_{1,z_2} \\
f_{2,z_1} & f_{2,z_2}
\end{bmatrix}
= 
\begin{bmatrix}
F_y & F_z \\
G_y & G_z
\end{bmatrix},$$

we may write the error equation in the more compact form

$$\begin{bmatrix}
I & 0
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}
= 
\begin{bmatrix}
F_y & F_z \\
G_y & G_z
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}.
$$

Both differential-algebraic systems in (17) were assumed to be of d-index 1, which means that both $g_{1,z_1}$ and $g_{2,z_2}$ are non-singular. But this does not automatically imply that $G_z$ is invertible as well. Actually, the coupled system and consequently also the error equation for the coupled system can be of arbitrary d-index, while all sub-systems are of d-index 1. For examples, we refer to [11,12]. In Example 3.7 below, for the choice of the parameter $\alpha = \pm 1$, the d-index of the coupled system is 2 while both subsystems are of d-index 1. Here, we are only considering the case that all subsystems and the whole system are of index 1.

If system (17) is embedded into a GAUSS-SEIDEL method, then we obtain the following differential algebraic recurrence equation

$$\begin{bmatrix}
I & I \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
e_{y_1}^{(n)} \\
e_{y_2}^{(n)} \\
e_{z_1}^{(n)} \\
e_{z_2}^{(n)}
\end{bmatrix}
= 
\begin{bmatrix}
f_{1,y_1} & 0 & f_{1,z_1} & 0 \\
f_{2,y_1} & f_{2,y_2} & f_{2,z_1} & f_{2,z_2}
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}^{(n-1)}
+ 
\begin{bmatrix}
0 & f_{1,y_1} & 0 & f_{1,z_1} \\
0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}^{(n)}
+ 
\begin{bmatrix}
0 & 0 & g_{1,y_1} & g_{1,z_2} \\
0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2}
\end{bmatrix}^{(n-1)}
,$$
Remark 2.2. Both cases can be expressed more compactly as a differential-algebraic difference equation

\[
\begin{bmatrix}
I & I \\
0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2} \\
\end{bmatrix}
^{(n)} =
\begin{bmatrix}
f_{1,y_1} & 0 & f_{1,z_1} & 0 \\
t_1 & t_2 & 0 & \psi \\
0 & \tilde{g}_{2,y_1} & \tilde{g}_{2,z_1} & 0 \\
0 & \tilde{g}_{2,z_1} & \tilde{g}_{2,z_2} & 0 \\
\end{bmatrix}
\begin{bmatrix}
e_{y_1} \\
e_{y_2} \\
e_{z_1} \\
e_{z_2} \\
\end{bmatrix}
^{(n)}
\]

Both cases can be expressed more compactly as a differential-algebraic difference equation

\[
\begin{bmatrix}
I & 0 \\
\end{bmatrix}
\begin{bmatrix}
E_y \\
E_z \\
\end{bmatrix}
^{(n)} =
\begin{bmatrix}
F_y & F_z \\
G_y & G_z \\
\end{bmatrix}
^{(0)}
\begin{bmatrix}
E_y \\
E_z \\
\end{bmatrix}
^{(n)} +
\begin{bmatrix}
F_y & F_z \\
G_y & G_z \\
\end{bmatrix}
^{(1)}
\begin{bmatrix}
E_y \\
E_z \\
\end{bmatrix}
^{(n-1)}
\]

with matrices

\[
\begin{bmatrix}
F_y & F_z \\
G_y & G_z \\
\end{bmatrix}
^{(i)}
\]

partitioned as

\[
\begin{bmatrix}
F_y & F_z \\
G_y & G_z \\
\end{bmatrix}
^{(i)}
\]

Remark 2.3. It is obvious, that for the GAUSS-SEIDEL- and the JACOBI methods the index of the coupled system only depends on the index of the subsystems. This is due to the fact that in (19) the matrices \(G_y^{(0)}\) are always nonsingular as long as \(g_{1,z_1}\) and \(g_{2,z_2}\) are nonsingular.

It is a well known fact from numerical linear algebra, that convergence of the JACOBI- and GAUSS-SEIDEL methods for linear systems depends on a criterion involving spectral radii. We will have to expect a similar behaviour for all coupled systems of DAEs, as long as the algebraic variables of at least 2 systems are mutually interconnected. Whether the method converges depends on the choice of the iteration method (JACOBI, GAUSS-SEIDEL, PICARD, etc.) and how the systems are ordered in this scheme, cf. [2]. Even for a convergent iteration method, in general, the convergence will be linear at best, cf. [21].

In order to obtain a criterion, whether a dynamic iteration method is convergent, we can use the following lemma.

Lemma 2.3. Let a sequence of differentiable functions \(\{y^{(i)}\}, i = 0, 1, 2, \ldots\) on \([t_0, t_0 + T]\) be defined by

\[
\begin{align*}
\left(y^{(n)}(t)\right)' &= F^{(0)}(t)y^{(n)}(t) + F^{(1)}(t)y^{(n-1)}(t) + H(t)\left(y^{(n-1)}(t)\right)' + \psi(t), \\
y^{(n)}(t_0) &= 0
\end{align*}
\]

and a given starting function \(y^{(0)}\). Furthermore, we assume that \(F^{(0)}\) and \(F^{(1)}\) are continuous matrix functions, with \(F^{(0)}\) LIPSCHITZ continuous, i.e., a finite constant \(L > 0\) exists such that for all \(t_1, t_2 \in [t_0, t_0 + T]\) the inequality

\[
\|F^{(0)}(t_1) - F^{(0)}(t_2)\| \leq L|t_1 - t_2|
\]

is satisfied. Assume furthermore that the matrix function \(H\) is continuously differentiable on \([t_0, t_0 + T]\). Finally, the inhomogeneity \(\psi\) is assumed to be LIPSCHITZ continuous as well. Then, the sequence (21) converges to a unique fixed point if \(\|H\|_\infty < 1\).

Proof: For a proof of this lemma, we refer to [11, 12].

Corollary 2.4. The criterion \(\|H\|_\infty < 1\) in Lemma 2.3 can be weakened in such a way that it is sufficient to request \(\|H^\nu\|_\infty < 1\) for some integer \(\nu \leq 1\). This criterion is related to quasi-nilpotency of dynamic iteration operators, cf. [25]. Going even further, for the limit case \(\nu \rightarrow \infty\), this translates to \(\rho(H) < 1\), where \(\rho(H)\) denotes the maximum spectral radius of \(H\) for all \(t \in [t_0, t_0 + T]\).
Proof: A proof for the case $\nu = 2$ is given in [12], while the proof for $\nu > 2$ is straightforward. The equivalence to the spectral radius can be found in [20].

**Remark 2.5.** For the case of the Gauss-Seidel- and Jacobi-iteration schemes for semi-explicit systems, the matrix
\[
\begin{bmatrix}
I & 0 \\
-G_y^{(0)} & -G_z^{(0)}
\end{bmatrix}^{-1}
\begin{bmatrix}
0 & 0 \\
G_y^{(1)} & G_z^{(1)}
\end{bmatrix},
\] takes the role of $H$ in Lemma 2.3 and is thus responsible for the convergence behaviour of the dynamic iteration method. With Corollary 2.4 this translates to $\rho((G_z^{(0)})^{-1}G_z^{(1)}) < 1$.

### 3 Enforcing convergence of the dynamic iteration method

As we have seen in the previous section, even for the simple case of semi-explicit DAEs of d-index 1, the convergence behaviour of the dynamic iteration method heavily depends on several parameters, i.e., on the systems themselves and the dynamic iteration scheme. In some cases even the order in which the systems are computed is of importance, see [2]. It is, thus, necessary to have a regularization method, that one can apply to the decoupled system. In this section we will develop criteria that give sufficient conditions such that a dynamic iteration scheme can be regularized analytically.

For a special type of coupled semi-explicit equations, the divergence phenomenon from section 2.1.2 has been studied in [1] and [2]. The basic idea of the proposed preconditioned dynamic iteration method is to substitute the algebraic variables of the current iterate by a linear combination of those in the actual and the last iterate. We will subsequently assume that the given DAEs are in semi-explicit form

\[
y'_1 = f_1(y_1, z_1, y_2, z_2), \\
0 = g_1(y_1, z_1, y_2, z_2),
\]

\[
y'_2 = f_2(y_1, z_1, y_2, z_2), \\
0 = g_2(y_1, z_1, y_2, z_2).
\] (22a)

(22b)

With correctly reordered variables and equations, this has the linearized error equation

\[
\begin{bmatrix}
I \\
0
\end{bmatrix}
\begin{bmatrix}
E_y \\
E_z
\end{bmatrix}' =
\begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}
\begin{bmatrix}
E_y \\
E_z
\end{bmatrix}.
\] (23)

The semi-explicit formulation is a necessary requirement for the correct determination of the regularization parameter. The procedure could also be applied to a more general class of DAEs, such as

\[
M_i x'_i = f_i(x_1, ..., x_s), \quad i = 1, ..., s,
\]

but since by adequate choice of bases and transformations from the left, the semi-explicit form can always be achieved, for the analysis, we use the semi-explicit form for its simpler notation.

We introduce a regularization parameter $\Xi$ in such a way that, whenever $E^{(n)}$ has to be evaluated on the right hand side, the sum $(I - \Xi)E^{(n)} + \Xi E^{(n-1)}$ will be evaluated instead. This yields

\[
\begin{bmatrix}
I \\
0
\end{bmatrix}
\begin{bmatrix}
E_y^{(n)} \\
E_z^{(n)}
\end{bmatrix}' =
\begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}^{(0)} (I - \Xi) \begin{bmatrix}
E_y^{(n)} \\
E_z^{(n)}
\end{bmatrix} + \Xi \begin{bmatrix}
E_y^{(n-1)} \\
E_z^{(n-1)}
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}^{(1)} (I - \Xi) \begin{bmatrix}
E_y^{(n-1)} \\
E_z^{(n-1)}
\end{bmatrix},
\]

\[
= \begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}^{(0)} (I - \Xi) \begin{bmatrix}
E_y^{(n)} \\
E_z^{(n)}
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}^{(1)} \Xi + \begin{bmatrix}
F_y \\
G_y
\end{bmatrix}
\begin{bmatrix}
F_z \\
G_z
\end{bmatrix}^{(1)} \begin{bmatrix}
E_y^{(n-1)} \\
E_z^{(n-1)}
\end{bmatrix}.
\] (24)
The following theorem will show how (24) can be regularized.

**Theorem 3.1.** A coupled system of two DAEs (17), where the coupled system as well as both subsystems are of d-index 1 and that is embedded into a GAUSS-SEIDEL dynamic iteration scheme can be regularized to guarantee convergence using the substitution \( \Xi \rightarrow (I - \Xi)E^{(n)} + \Xi E^{(n-1)} \) with

\[
\Xi = \begin{bmatrix} 0 & 0 \\ 0 & \Xi_{22} \end{bmatrix} \quad \text{and} \quad \Xi_{22} = \begin{bmatrix} 0 & 0 \\ g_{2,z_2}^{-1}g_{1,z_1}^{-1} & g_{1,z_1}^{-1} \\ g_{2,z_2}^{-1}g_{1,z_1}^{-1} \end{bmatrix}.
\]  

(25)

**Proof:** Assuming the structure \( \Xi = \begin{bmatrix} 0 & 0 \\ 0 & \Xi_{22} \end{bmatrix} \), we determine an ODE that is analytically equivalent to (24) by differentiating the algebraic equations and rearranging the new differential variables on the left.

\[
\begin{align*}
\begin{bmatrix} I \\ -G_y^{(0)}(I - \Xi_{22}) & -G_z^{(0)}(I - \Xi_{22}) \end{bmatrix} \begin{bmatrix} E_y^{(n)} \\ E_z^{(n)} \end{bmatrix}' &= \begin{bmatrix} F_y^{(0)} \\ (G_y^{(0)}(I - \Xi_{22}))' \\ F_z^{(0)} \\ (G_z^{(0)}(I - \Xi_{22}))' \end{bmatrix} \begin{bmatrix} E_y^{(n)} \\ E_z^{(n)} \\ F_y^{(1)} \\ (G_y^{(1)}(I - \Xi_{22} + (G_y^{(1)})^{-1}G_z^{(1)})' \end{bmatrix} \begin{bmatrix} E_y^{(n-1)} \\ E_z^{(n-1)} \\ 0 \\ 0 \end{bmatrix} \\
&+ \begin{bmatrix} G_y^{(0)}(I - \Xi_{22}) + G_y^{(1)} \\ G_z^{(0)}(I - \Xi_{22}) + G_z^{(1)} \end{bmatrix} \begin{bmatrix} E_y^{(n-1)} \\ E_z^{(n-1)} \end{bmatrix}'
\end{align*}
\]  

(26)

This implicit ODE can be transformed into an explicit one by inverting the matrix on the left hand side. Then, by application of Corollary 2.4 we have that convergence is guaranteed if \( \rho(H) < 1 \), where

\[
H = \begin{bmatrix} I \\ -G_y^{(0)}(I - \Xi_{22}) & -G_z^{(0)}(I - \Xi_{22}) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ G_y^{(0)}(I - \Xi_{22}) + G_y^{(1)} \\ G_z^{(0)}(I - \Xi_{22}) + G_z^{(1)} \end{bmatrix}.
\]

Indeed, we will show, that \( \rho(H) = 0 \).

The first row of \( H \) contains only zeros, hence, only the lower right block \( H_{22} \) is of interest to show the nilpotency of \( H \). This block can be evaluated as

\[
H_{22} = -(I - \Xi_{22})^{-1}(G_y^{(0)})^{-1}(G_y^{(0)}(I - \Xi_{22}) + G_y^{(1)}),
\]

\[
= -(I - \Xi_{22})^{-1}(\Xi_{22} + (G_y^{(0)})^{-1}G_z^{(1)}).
\]

For the GAUSS-SEIDEL iteration scheme, the matrices \( G_y^{(0)} \) and \( G_z^{(1)} \) have the form

\[
G_y^{(0)} = \begin{bmatrix} g_{1,z_1} & 0 \\ g_{2,z_1} & g_{2,z_2} \end{bmatrix} \quad \text{and} \quad G_z^{(1)} = \begin{bmatrix} 0 & g_{1,z_1} \\ 0 & 0 \end{bmatrix},
\]

so

\[
(G_y^{(0)})^{-1}G_z^{(1)} = \begin{bmatrix} 0 & g_{1,z_1}^{-1}g_{1,z_2} \\ 0 & -g_{2,z_2}^{-1}g_{2,z_1}g_{1,z_1}^{-1}g_{1,z_2} \end{bmatrix}.
\]

With the special choice (25) for \( \Xi \), the matrix \( \Xi_{22} + (G_y^{(0)})^{-1}G_z^{(1)} \) becomes nilpotent. The matrix \( (I - \Xi_{22})^{-1} \) only affects the lower row of \( \Xi_{22} + (G_y^{(0)})^{-1}G_z^{(1)} \) which is zero and, hence, remains unchanged. It remains to show, that this inverse exists. This is the case if and only if in

\[
I - \Xi_{22} = \begin{bmatrix} I \\ 0 \\ I - g_{2,z_2}^{-1}g_{2,z_1}g_{1,z_1}^{-1}
\end{bmatrix}
\]
the submatrix \( I - g_{2,z_2}^{-1}g_{2,z_2}g_{1,z_1}^{-1}g_{1,z_1} \) or equivalently \( g_{2,z_2}^{-1}g_{2,z_2}g_{1,z_1}^{-1}g_{1,z_1} - g_{2,z_2}g_{1,z_2}^{-1}g_{1,z_2} \) is nonsingular. Here, \( g_{2,z_2} \) is nonsingular, since all subsystems were assumed to be of d-index one. The second factor is the Schur complement of \( \begin{bmatrix} g_{1,z_1} & g_{1,z_2} \\ g_{2,z_1} & g_{2,z_2} \end{bmatrix} \) which is nonsingular or (22a) and (23) would not be of d-index one. Hence, \( \Xi \) exists and the resulting \( \rho(H) = 0 \).

**Remark 3.2.** Instead of evaluating the expression \( g_{2,z_2}^{-1}g_{2,z_1}g_{1,z_1}^{-1} \) in every step, it is possible to change the constraint equations for the second system in (23) to

\[
\begin{align*}
0 &= g_1(y^{(n)}_1, z^{(n)}_1, y^{(n)}_2, z^{(n)}_2), \\
0 &= g_2(y^{(n)}_1, z^{(n)}_1, y^{(n)}_2, z^{(n)}_2),
\end{align*}
\]

where \( z^{(n)}_1 \) acts as a slack variable. A simple computation shows, that the changed set of constraints performs the regularization.

This approach is particularly interesting if the number of constraints added in this way is small. Also, a deeper investigation of the coupled system may yield that not all constraints of \( g_1 \) are necessary for the inter-coupling of systems and may, thus, be neglected.

**Remark 3.3.** The discussed approach is not applicable for the Jacobi method, since then, the product \( (G_z^{(0)})^{-1}G_z^{(1)} = \begin{bmatrix} 0 & g_{1,z}^{-1}g_{1,z} \\ g_{2,z}^{-1}g_{2,z} & 0 \end{bmatrix} \). As in the Jacobi method only the variables of the subsystem under consideration are updated, only the diagonal blocks of \( \Xi_{22} \) may be non-zero, which, in general, makes it very difficult or even impossible to find a \( \Xi_{22} \) such that \( \Xi_{22} + (G_z^{(0)})^{-1}G_z^{(1)} \) is nilpotent. For the Jacobi method we can use the following theorem instead.

**Theorem 3.4.** A coupled system of d-index one consisting of two DAEs (17) each of which have again d-index one, embedded into a Jacobi type dynamic iteration scheme can be regularized to guarantee convergence using the substitution \( \Xi \rightarrow (I - \Xi)E^{(n)} + \Xi E^{(n-2)} \) in every other step, e.g. in every even step, with

\[
\Xi = \begin{bmatrix} 0 & 0 \\ 0 & \Xi_{22} \end{bmatrix} \quad \text{and} \quad \Xi_{22} = \begin{bmatrix} g_{1,z}^{-1}g_{1,z} & 0 \\ 0 & g_{2,z}^{-1}g_{2,z} \end{bmatrix}.
\]

**Proof:** We consider two iterations at once and define vectors consisting of two consecutive steps as

\[
\begin{align*}
\tilde{y}^{(n)} &= \begin{bmatrix} y^{(n-1)} \\ y^{(n)} \end{bmatrix}, \\
\tilde{z}^{(n)} &= \begin{bmatrix} z^{(n-1)} \\ z^{(n)} \end{bmatrix},
\end{align*}
\]

and of the errors of two consecutive iterates as

\[
\begin{align*}
\tilde{E}_y^{(n)} &= \begin{bmatrix} E_y^{(n)} \\ E_y^{(n)} \end{bmatrix}, \\
\tilde{E}_z^{(n)} &= \begin{bmatrix} E_z^{(n-1)} \\ E_z^{(n)} \end{bmatrix}.
\end{align*}
\]

Then, the differential difference equation for these errors has the form

\[
\begin{bmatrix} I & 0 \\ 0 & \end{bmatrix} \begin{bmatrix} \tilde{E}_y^{(n)} \\ \tilde{E}_z^{(n)} \end{bmatrix} = \begin{bmatrix} F_y & F_z \end{bmatrix} \begin{bmatrix} \tilde{E}_y^{(n)} \\ \tilde{E}_z^{(n)} \end{bmatrix} + \begin{bmatrix} F_y & F_z \end{bmatrix} \begin{bmatrix} \tilde{E}_y^{(n-2)} \\ \tilde{E}_z^{(n-2)} \end{bmatrix},
\]

with

\[
\begin{align*}
F_y^{(0)} &= \begin{bmatrix} F_y^{(0)} & 0 \\ F_y^{(1)} & F_y^{(0)} \end{bmatrix}, \\
F_z^{(0)} &= \begin{bmatrix} F_z^{(0)} & 0 \\ F_z^{(1)} & F_z^{(0)} \end{bmatrix}, \\
G_y^{(0)} &= \begin{bmatrix} G_y^{(0)} \\ G_y^{(1)} \end{bmatrix}, \\
G_z^{(0)} &= \begin{bmatrix} G_z^{(0)} \\ G_z^{(1)} \end{bmatrix}.
\end{align*}
\]
Remark 3.5. The parameter $\Xi$ in (28) is block diagonal. This means that for the substitution of $z_{1(n)}$ and $z_{2(n)}$ only previous iterates of $z_{1(n)}$ and $z_{2(n)}$, respectively, will be used.

Remark 3.6. As in the case of the Gauss-Seidel iteration method, the evaluation of the regularization parameter (28) can be replaced by the solution of extended constraint equations

$$
0 = g_1(y_1(n), z_1(n), y_2(n-1), z_2(n-1)), \\
0 = g_2(y_1(n-1), z_1(n), y_2(n-2), z_2(n-1))
$$

for (22a) and

$$
0 = g_2(y_1(n-1), z_1(n-1), y_2(n), z_2(n)), \\
0 = g_1(y_1(n), z_1(n), y_2(n-1), z_2(n))
$$

for (22b) in every other step. The variables $z_{i(n-1)}$, $i = 1, 2$ again act as slack variables and it is possible that not all constraints have an effect on the coupling and thus, some might be neglected. In order to determine the relevant constraints, the coupled systems have to be investigated in detail.

Example 3.7. Consider the following system of two coupled semi-explicit DAEs, to be solved using the Gauss-Seidel iteration method

$$
y'_1 = -y_2, \\
0 = y_1 - z_1 + \alpha z_2, \\
y'_2 = y_1, \\
0 = y_2 - z_2 + \alpha z_1,
$$
on $[0, \pi]$ for $y_1(0) = 1$, $y_2(0) = 0$, $z_1(0) = 1/(1 - \alpha^2)$, $z_2(0) = \alpha/(1 - \alpha^2)$. The starting functions $y_1^{(0)}$, $y_2^{(0)}$, $z_1^{(0)}$, $z_2^{(0)}$ were chosen to be constant extrapolations of the initial values. The parameter $\alpha$ can be chosen arbitrarily but different from $\pm 1$ or the index of the coupled problem would rise to 2. In any case, the analytic rate of convergence for the problem is equal to $\alpha^2$.

Using Theorem 3.1, we obtain $\Xi = \begin{bmatrix} 0 \\ \alpha^2 \end{bmatrix}$. Hence, with the simple substitution

$$
z_{2(n)} \rightarrow (1 - \alpha^2)z_{2(n)} + \alpha^2 z_{2(n-1)}
$$

convergence can be forced and a significant decrease of computation time can be achieved. We show the results of a computation in Table 1. We display the number of dynamic iteration steps, i.e., the number of iterations until the numerical solution and the exact solution differ by less than $10^{-6}$.
Table 1: number of iteration steps

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>without regularization</th>
<th>with regularization</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>0.5</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>0.7</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>0.9</td>
<td>81</td>
<td>10</td>
</tr>
<tr>
<td>1.1</td>
<td>divergence</td>
<td>11</td>
</tr>
</tbody>
</table>

4 Conclusions and future work

We have investigated the simulation of coupled systems of ODEs and DAEs using dynamic iteration methods. These methods allow the simulation of several subsystems with different solvers, not necessarily on the same machine. These methods always converge, at least analytically, in the case of coupled ODEs. The situation may change completely for coupled DAEs. We have developed a criterion whether and at what rate convergence may be expected. Furthermore, we have introduced regularization methods that for the Gauss-Seidel and the Jacobi methods will make a previously divergent dynamic iteration scheme convergent.

Future work will include a perturbation analysis of the dynamic iteration process. Also investigations for special large scale systems such as fluidic systems or circuit equations will be performed in order to determine the regularization as efficiently as possible.

Acknowledgements The author would like to thank V. Mehrmann, A. Steinbrecher and S. Bächle for numerous suggestions and fruitful discussions.

References


