Waveform Relaxation: A Convergence Criterion for Differential-Algebraic Equations

Jonas Pade Caren Tischendorf

June 30, 2018

Abstract

While waveform relaxation (also known as dynamic iteration or co-simulation) methods are known to converge for coupled systems of ordinary differential equations (ODEs), they may suffer from instabilities for coupled differential-algebraic equations (DAEs). Several convergence criteria have been developed for index-1 DAEs. We present here a convergence criterion for a coupled system of an index-2 DAE with an ODE. The analysis is motivated by the wish to combine electromagnetic field simulation with circuit simulation in a stable manner. The spatially discretized Maxwell equations in vector potential formulation with Lorenz gauging represent an ODE system. A lumped circuit model via the established modified nodal analysis is known to be a DAE system of index ≤ 2. Finally, we present sufficient network topological criteria to the coupling that are easy to check and that guarantee convergence.

1 Introduction

The modeling of a large number of today's applications is increasingly complex and in many cases calls for a multiphysical approach resulting in coupled systems. Waveform relaxation (WR) methods are well-established for coupled problems. They allow for each subsystem to be solved by a dedicated numerical solver taking into account the different structure and time scales of the subsystems. WR methods are known to be convergent on bounded intervals for coupled ordinary differential equations (ODEs) [9]. This is not necessarily true in the case of coupled differential-algebraic equations (DAEs) [1, 10, 11]. Therefore, a number of studies were dedicated to finding convergence criteria for different classes of coupled DAEs, e.g. [1, 2, 5, 6, 10, 11, 13, 14] to name only some of them. Here, we investigate a novel class of systems: coupled systems of implicit quasilinear DAEs of (tractability) index ≤ 2 [8, p.485] and ODEs. The choice of this class of equations is motivated by the interest in finding sufficient criteria for convergence of co-simulation approaches for coupled electromagnetic field (EM)/circuit systems. Convergence results are well-known for coupled EM/circuit systems for index-1 DAEs, see e.g. [2, 13, 14]. Here, we provide convergence results for the Gauss-Seidel WR for EM couplings with index-2 circuit DAEs. The Gauss-Seidel WR method is chosen as one prototype example of many different WR methods.

In Section 2, we provide a general convergence criterion for the Gauss-Seidel method applied to index-2 DAEs coupled with an ODE. Section 3 discusses

the consequences for circuit systems coupled to an ODE (e.g. spatially discretized Maxwell equations in potential formulation including Lorenz gauging) and provides a network topological interpretation of the convergence criterion. Furthermore, the criterion is illustrated by two simple examples.

2 An abstract convergence criterion

This work presents convergence criteria for an iterative method applied to initial value problems (IVPs) of the form

$$\dot{u} + b(t,u) = c_1(x), \qquad u(t_0) = u_0 \in \mathbb{R}^{n_u}, \tag{1}$$
$$E(x)\dot{x} + f(x) = q(t) + c_2(u), \qquad x(t_0) = x_0 \in \mathbb{R}^{n_x}. \tag{2}$$

on a finite time interval $\mathcal{I}$, where $E(x) \in \mathbb{R}^{n_x \times n_x}$ is square. The right hand side functions c_1 and c_2 describe the coupling of both systems.
We assume that the IVP of the form

$$E(x)\dot{x} + f(x) = s(t), \qquad x(t_0) = x_0, \qquad t \in \mathcal{I}, \tag{3}$$

whith E, f as in (2), can be equivalently transformed into equations of the form

$$\dot{y} = f_0(y, z_1, z_2, s), \qquad y(t_0) = y_0, \qquad t \in \mathcal{I} \tag{4a}$$
$$z_1 = M_1(y, z_2)\dot{s} + f_1(y, s) \tag{4b}$$
$$z_2 = f_2(y, s), \tag{4c}$$

where the functions f_0, f_1, f_2, M_1 satisfy some smoothness conditions, i.e.

Assumption 2.1 (decoupled form) *There exists a nonsingular transformation matrix $T = (T_0\ T_1\ T_2)$ and functions f_0, f_1, f_2, M_1, defined on euclidean spaces $\mathbb{R}^k$, with the properties*

- $M_1, f_0, f_1, f_2 \in C^1$
- $f_1(y,s)$ *and* $f_2(y,s)$ *are Lipschitz continuous w.r.t.* y
- $f_0(y, z_1, z_2, s)$ *is Lipschitz continuous w.r.t.* y, z_1, z_2
- M_1 *is Lipschitz continuous*

such that for any $s \in C^2(\mathcal{I})$, the function $(y, z_1, z_2) \in C^1(\mathcal{I})$, uniquely defined by $T_0 y + T_1 z_1 + T_2 z_2 = x$, solves IVP (4) if and only if $x \in C^1(\mathcal{I})$ solves the IVP (3).

Assumption 2.1 implies some essential properties regarding the nature of Equation (3):
The differential equation in (3) is a *differential-algebraic equation* (DAE), that is, $E(x)$ is singular for all $x \in \mathbb{R}^n$. It is implicitly composed of a dynamic part (4a), sometimes called *inherent ODE* in the literature [8], and *algebraic constraints* (4b),(4c). This mixed differential-algebraic nature is the basic feature of DAEs. The *tractability index* of the DAE in (3) is 2, since in equation (4b) the first derivative but no higher derivatives are involved in the algebraic constraints (see [8]).

There exists a *consistent initial value* for the problem (3), that is, an initial value such that there exists a solution x. This can be seen if we successively insert z_1 and z_2 into Equation (4a). That way, we obtain an ODE in y, with a vector field which is Lipschitz continuous w.r.t. y as it is composed of Lipschitz continuous functions. Hence, a global version of the Picard Lindelöf Theorem [15] yields global unique solvability of this ODE. Inserting this unique solution y successively into the algebraic Equations (4c),(4b) yields unique solutions z_1, z_2, and thence $x = T(y^\top, z_1^\top, z_2^\top)^\top$ solves the problem (3). Consequently, the initial value $x_0 = T(y_0^\top, z_{1,0}^\top, z_{2,0}^\top)^\top$ is consistent if and only if the initial value y_0 is chosen arbitrarily and $z_{1,0}, z_{2,0}$ are the resulting fixed initial values of the algebraic constraints.
Given a consistent inital value x_0, the corresponding solution x of problem (3) is unique and global on the finite time interval $\mathcal{I}$.

Assumption 2.2 *We consider the coupled system (1)-(2). The vector field $b(t,\cdot)$ is Lipschitz continuous for all $t \in \mathbb{R}$. The coupling function c_1 is continuously differentiable, and c_2 and the source function q are twice continuously differentiable.*

2.1 Convergence for Gauß-Seidel waveform relaxation

When applying the Gauss-Seidel WR method on (1)-(2), we obtain the iterative scheme

$$\dot{u}^k + b(t, u^k) = c_1(x^{k-1}), \qquad u^k(t_0) = u_0 \tag{5}$$

$$E(x^k)\dot{x}^k + f(t, x^k) = q(t) + c_2(u^k), \qquad x^k(t_0) = x_0^k \tag{6}$$

for $t \in \mathcal{I}$ and the iteration parameter $k \in \mathbb{N} \setminus \{0\}$. The initial guess function u^0 can be found through extrapolation of the initial values u_0. The initial value x_0^k is defined by

$$x_0^k := T_0 y_0 + T_1 M_1(y_0, z_{2,0})[\dot{q} + C_2(u^k)\dot{u}^k] + T_2 z_{2,0},$$

with C_2 the Jacobian of c_2. This definition seems cumbersome, but is necessary to obtain consistent initial values as can be seen from the transformed equations (4). For a given u^k and $s(t) := q(t) + c_2(u^k(t))$, Equations (3) and (6) are identical. If Assumption 2.1 holds for Equation (3), it consequently holds for Equation (6) as well, with $x^k = T(y^k, z_1^k, z_2^k)^\top$. Since Equation (4b) involves the derivative of u^k, the algebraically fixed initial value $z_1^k(t_0)$ depends on k.

The following theorem about the convergence of solutions of a sequence of ODEs is the groundwork for our convergence results of waveform relaxation for coupled ODE-DAE systems. The investigated sequence therein resembles the Picard iteration or waveform relaxation methods for ODEs, which are described in [3,9]. In contrast to these works, we deal with an ODE that depends not only on the solution of the ODE in the previous iteration step but also on its derivative.

Theorem 2.3 *For $[t_0, T] = \mathcal{I} \subset \mathbb{R}$, we consider the function*

$$f : \mathcal{I} \times \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m, \quad (t, x, y, z) \mapsto f(t, x, y, z).$$

Let f be continuously differentiable, contractive w.r.t. z with the contraction constant c for all $t \in \mathcal{I}$ and Lipschitz continuous w.r.t. x, y with the Lipschitz constants L_1, L_2 for all $t \in \mathcal{I}$. Let furthermore the IVP

$$\dot{x} = f(t, x, x, \dot{x}), \quad x(t_0) = x_0, \quad t \in \mathcal{I} \tag{7}$$

have a unique solution on $\mathcal{I}$. If the time interval $\mathcal{I}$ is sufficiently small to satisfy

$$T - t_0 < \frac{1-c}{L_1 + L_2},$$

then the sequence of IVP-solutions

$$\dot{x}^k = f(t, x^k, x^{k-1}, \dot{x}^{k-1}), \quad x^k(t_0) = x_0 \ \forall k \in \mathbb{N}, \quad x^0 \in C^1, \quad t \in \mathcal{I} \tag{8}$$

converges in $(C^1(\mathcal{I}), \|\cdot\|_{C^1(\mathcal{I})})$ to the solution of (7), i.e. $x^k \to x$, for all initial functions $x^0 \in C^1$.

Proof: Rewriting the difference of the equations (8), (7) in the integral form, we get

$$x^k(t) - x(t) = \int_{t_0}^{t} f(s, x^k(s), x^{k-1}(s), \dot{x}^{k-1}(s))ds - \int_{t_0}^{t} f(s, x(s), x(s), \dot{x}(s))ds.$$

We define

$$v_k := \|x^k - x\|_{C^0(\mathcal{I})}, \quad w_k := \|\dot{x}^k - \dot{x}\|_{C^0(\mathcal{I})}$$

based on an arbitrary vector norm $\|\cdot\|$. We will show that v_k and w_k converge to zero. For clarity, we write down the argument t whenever it occurs in the following. Using that for arbitrary $z \in C^0[a,b]$ it holds

$$\| \int_a^b z(t)dt\| \leq (b-a)\|z\|_{C^0[a,b]},$$

we obtain with $H := T - t_0$

$$\begin{aligned}
v_k &\leq H \sup_{t\in\mathcal{I}} \|f(t, x^k(t), x^{k-1}(t), \dot{x}^{k-1}(t)) - f(t, x(t), x(t), \dot{x}(t))\| \\
&\leq H \sup_{t\in\mathcal{I}} \|f(t, x^k(t), x^{k-1}(t), \dot{x}^{k-1}(t)) - f(t, x(t), x^{k-1}(t), \dot{x}^{k-1}(t))\| \\
&\quad + H \sup_{t\in\mathcal{I}} \|f(t, x(t), x^{k-1}(t), \dot{x}^{k-1}(t)) - f(t, x(t), x(t), \dot{x}^{k-1}(t))\| \\
&\quad + H \sup_{t\in\mathcal{I}} \|f(t, x(t), x(t), \dot{x}^{k-1}(t)) - f(t, x(t), x(t), \dot{x}(t))\| \\
&\leq HL_1 \sup_{t\in\mathcal{I}} \|x^k(t) - x(t)\| + HL_2 \sup_{t\in\mathcal{I}} \|x^{k-1}(t) - x(t)\| \\
&\quad + Hc \sup_{t\in\mathcal{I}} \|(\dot{x}^{k-1}(t) - \dot{x}(t))\| \\
&= HL_1 v_k + HL_2 v_{k-1} + Hcw_{k-1}.
\end{aligned}$$

Forming the difference of the equations (8) and (7) yields analogously

$$\begin{aligned}
w_k &= \sup_{t\in\mathcal{I}} \|f(t, x^k(t), x^{k-1}(t), \dot{x}^{k-1}(t)) - f(t, x(t), x(t), \dot{x}(t))\| \\
&\leq L_1 v_k + L_2 v_{k-1} + cw_{k-1}.
\end{aligned}$$

In matrix form, we obtain

$$\underbrace{\begin{pmatrix} 1-HL_1 & 0 \\ -L_1 & 1 \end{pmatrix}}_{T} \begin{pmatrix} v_k \\ w_k \end{pmatrix} \leq \begin{pmatrix} HL_2 & Hc \\ L_2 & c \end{pmatrix} \begin{pmatrix} v_{k-1} \\ w_{k-1} \end{pmatrix}$$

with the inequality understood component-wise. For $H < \frac{1-c}{L_1+L_2}$ and $0 < c < 1$, it follows $H < \frac{1}{L_1}$. Thus, the matrix T is invertible. Since all components of T^{-1} are non-negative we obtain

$$\begin{aligned} 0 \leq \begin{pmatrix} v_k \\ w_k \end{pmatrix} &\leq T^{-1} \begin{pmatrix} HL_2 & Hc \\ L_2 & c \end{pmatrix} \begin{pmatrix} v_{k-1} \\ w_{k-1} \end{pmatrix} \\ &= \underbrace{\frac{1}{1-HL_1} \begin{pmatrix} HL_2 & Hc \\ L_2 & c \end{pmatrix}}_{=:K} \begin{pmatrix} v_{k-1} \\ w_{k-1} \end{pmatrix} \leq K^k \begin{pmatrix} v_0 \\ w_0 \end{pmatrix}. \end{aligned}$$

For the spectral radius $\rho(K) = \frac{HL_2+c}{1-HL_1}$ of K it holds

$$\frac{HL_2+c}{1-HL_1} < 1 \iff H < \frac{1-c}{L_1+L_2}.$$

Thus v_k and w_k converge to zero if $0 < c < 1$ and $H < \frac{1-c}{L_1+L_2}$. This yields the conclusion

$$\|x^k - x\|_{C^1(\mathcal{I})} = \|x^k - x\|_{C^0(\mathcal{I})} + \|\dot{x}^k - \dot{x}\|_{C^0(\mathcal{I})} = v_k + w_k \overset{k\to\infty}{\to} 0.$$

□

The following theorem represents the main convergence results for waveform relaxation methods applied to coupled ODE-DAE systems with index-2 DAEs that can be decoupled into a form described by Assumption 2.1.

Theorem 2.4 *Let the Assumptions 2.1 and 2.2 be satisfied. Then, the sequence (u^k, x^k) of iterative solutions of the Gauss-Seidel method (5)-(6) converges in $(C^0(\mathcal{T}), \|\cdot\|_{C^0(\mathcal{T})})$ to the solution (u, x) of (1)-(2), if the time interval $\mathcal{T} \subseteq \mathcal{I}$ is sufficiently small and it holds*

$$\rho\left[C_1(x) T_1 M_1(y, z_2) C_2(u)\right] < 1 \quad \forall x, y, z_2, u \tag{9}$$

with $C_1(x) := c_1'(x)$ and $C_2(u) := c_2'(u)$.

Proof: The proof consists of two steps. First, we use Theorem 2.3 in order to show that the criterion (9) implies that the solutions of the inherent ODE of the iterated scheme (5)-(6) converge in $(C^1(\mathcal{T}), \|\cdot\|_{C^1(\mathcal{T})})$ to the solution of the inherent ODE of the original system (1)-(2).We write $C^1 := (C^1(\mathcal{T}), \|\cdot\|_{C^1(\mathcal{T})})$ and $C^0 := (C^0(\mathcal{T}), \|\cdot\|_{C^0(\mathcal{T})})$ throughout this proof. The second step proves that C^1 convergence of the inherent ODE implies C^0 convergence of the full system.

First step: Exploiting Assumption 2.1 with $s(t) = q(t) + u^k(t)$, we can substitute the equivalent equations

$$\dot{y}^k = f_0(y^k, z_1^k, z_2^k, q + c_2(u^k)), \qquad y^k(t_0) = y_0 \tag{10}$$

$$z_1^k = M_1(y^k, z_2^k) \frac{d}{dt}(q + c_2(u^k)) + f_1(y^k, z_2^k, q + c_2(u^k)) \tag{11}$$

$$z_2^k = f_2(y^k, q + c_2(u^k)) =: \bar{f}_2(t, y^k, u^k) \tag{12}$$

for Equation (6). Notice that $x^k = T_0 y^k + T_1 z_1^k + T_2 z_2^k$. Defining

$$\begin{aligned}\bar{M}_1(t, y^k, u^k) &:= M_1(y^k, \bar{f}_2(t, y^k, u^k)),\\ \bar{f}_1(t, y^k, u^k) &:= \bar{M}_1(t, y^k, u^k))\dot{q} + f_1(y^k, z_2^k, q + c_2(u^k))\end{aligned}$$

and inserting Equation (12) into (11) yields

$$z_1^k = \bar{M}_1(t, y^k, u^k) C_2(u^k)\dot{u}^k + \bar{f}_1(t, y^k, u^k). \tag{13}$$

We replace x^k in (5) by $T(y^k, z_1^k, z_2^k)^\top$ and substitute the right hand sides of (13),(12) for the algebraic components of z_1^k, z_2^k in (5) and (10).

$$\begin{aligned}\dot{u}^k &= -b(t, u^k) + c_1(x^{k-1})\\ &= -b(t, u^k) + c_1(T_0 y^{k-1} + T_1 z_1^{k-1} + T_2 z_2^{k-1})\\ &= -b(t, u^k) + c_1(T_0 y^{k-1} + T_1 M_1(t, y^{k-1}, u^{k-1}) C_2(u^{k-1})\dot{u}^{k-1}\\ &\qquad + \bar{f}_1(t, y^{k-1}, u^{k-1}) + T_2 \bar{f}_2(t, y^{k-1}, u^{k-1}))\\ &=: \theta_1(t, u^k, u^{k-1}, y^{k-1}, \dot{u}^{k-1})\end{aligned} \tag{14}$$

$$\begin{aligned}\dot{y}^k &= f_0(y^k, z_1^k, z_2^k, q(t) + c_2(u^k))\\ &= f_0(y^k, \bar{M}_1(t, y^k, u^k) C_2(u^k)\dot{u}^k + \bar{f}_1(t, y^k, u^k), \bar{f}_2(t, y^k, u^k), q + c_2(u^k))\end{aligned}$$

In the above equations, $\dot{u}^k$ depends only on u^k and (derivatives of) previous solutions, whereas $\dot{y}^k$ depends additionally on $\dot{u}^k$. To obtain an explicit ODE in $(u^k, y^k)^\top$, we eliminate $\dot{u}^k$ in the right hand side by simply substituting θ_1 for $\dot{u}^k$.

$$\begin{aligned}\dot{y}^k &= f_0(y^k, \bar{M}_1(t, y^k, u^k) C_2(u^k)\theta_1(t, u^k, u^{k-1}, y^{k-1}, \dot{u}^{k-1})\\ &\qquad + \bar{f}_1(t, y^k, u^k), \bar{f}_2(t, y^k, u^k), q + c_2(u^k))\\ &=: \theta_2(t, u^k, y^k, u^{k-1}, y^{k-1}, \dot{u}^{k-1})\end{aligned} \tag{15}$$

Together, we have the ODE system

$$\begin{aligned}\dot{u}^k &= \theta_1(t, u^k, u^{k-1}, y^{k-1}, \dot{u}^{k-1}),\\ \dot{y}^k &= \theta_2(t, u^k, y^k, u^{k-1}, y^{k-1}, \dot{u}^{k-1}).\end{aligned} \tag{16}$$

Next, we perform the analogous transformations on equations (1),(2), using Assumption 2.1 once more. We obtain

$$z_2 = f_2(t, u, y), \quad z_1 = M_1(y, \varphi_3(t, u)) M_2 C_2(u)\dot{u} + \varphi_1(t, u, y).$$

Analogously to the inherent ODE (16) of the system (5)-(6), this leads to the inherent ODE of the overall system (1)-(2)

$$\begin{aligned}\dot{u} &= \theta_1(t, u, u, y, \dot{u}),\\ \dot{y} &= \theta_2(t, u, y, u, y, \dot{u}).\end{aligned} \tag{17}$$

For all $t \in \mathcal{T}$, the right hand side functions $\theta_1(t, \cdot)$ and $\theta_2(t, \cdot)$ defined in (14) and (15) are compositions of Lipschitz continuous functions due to Assumption

2.1. Consequently, they are Lipschitz continuous. Applying Theorem 2.3 to the systems (17) and (16), we only need to consider the spectral radius of

$$M := \begin{pmatrix} \frac{\partial \theta_1}{\partial \dot{u}^{k-1}} & \frac{\partial \theta_1}{\partial \dot{y}^{k-1}} \\ \frac{\partial \theta_2}{\partial \dot{u}^{k-1}} & \frac{\partial \theta_2}{\partial \dot{y}^{k-1}} \end{pmatrix} = \begin{pmatrix} \frac{\partial \theta_1}{\partial \dot{u}^{k-1}} & 0 \\ \frac{\partial \theta_2}{\partial \dot{u}^{k-1}} & 0 \end{pmatrix}.$$

Finally, C^1 convergence of $(u^k, y^k) \to (u, y)$ is given if if the time interval $\mathcal{T}$ is sufficiently small and

$$\rho(M) = \rho\left(\tfrac{\partial \theta_1}{\partial \dot{u}^{k-1}}\right) = \rho\left[C_1(x)T_1M_1(y, z_3))M_3C_2(u)\right] < 1 \quad \forall x \, \forall y \, \forall z_3 \, \forall u. \tag{18}$$

Second step: By inequality (18) we know that $(u^k, y^k) \to (u, y)$ in C^1. We recall that $\bar{M}_1, \bar{f}_1, \bar{f}_2$, defined in (12), (13), are continuous as compositions of (Lipschitz) continuous functions. Hence we can exploit the continuity to switch functions and limits in C^0 as follows

$$\begin{aligned} \lim_{k\to\infty} z_2^k &= \lim_{k\to\infty} \bar{f}_2(t, u^k, y^k) = \bar{f}_2(t, u, y) =: z_2 \\ \lim_{k\to\infty} z_1^k &= \lim_{k\to\infty} [\bar{M}_1(t, y^k, u^k)C_2(u^k)\dot{u}^k + \bar{f}_1(t, y^k, u^k)] \\ &= \bar{M}_1(t, y, u)C_2(u)\dot{u} + \bar{f}_1(t, y, u) =: z_1. \end{aligned}$$

The last equation holds since $(u^k, y^k) \to (u, y)$ in C^1, which implies $\dot{u}^k \to \dot{u}$ in C^0. Hence, additionally to $u^k \to u$, we obtain in C^0

$$\lim_{k\to\infty} x^k = \lim_{k\to\infty} (T_0 y^k + T_1 z_1^k + T_2 z_2^k) = T_0 y + T_1 z_1 + T_2 z_2 = x.$$

□

3 Convergence criteria for circuit coupled systems

In this section, we consider a coupled system (1),(2) whose DAE arises from an electrical circuit, modeled by the modified nodal analysis. That is,

$$\begin{aligned} E(x) :&= \begin{bmatrix} A_C C(A_C^\top e) A_C^\top & 0 & 0 \\ 0 & -L(i_l) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad x = \begin{pmatrix} e \\ i_l \\ i_v \end{pmatrix}, \quad q = \begin{pmatrix} q_i \\ 0 \\ q_v \end{pmatrix}, \\ f(t, x) :&= \begin{bmatrix} A_R g(A_R^\top e) + A_L i_l + A_V i_v + q_i(t) \\ A_L^\top e \\ A_V^\top e - q_v(t), \end{bmatrix} \end{aligned} \tag{19}$$

where A_C, A_R, A_L, A_V are incidence matrices (see Definition 3.7) of the capacitances, resistances, inductances and voltage sources in the circuit [4]. $L(\cdot) \in \mathbb{R}^{n_l \times n_l}$, $C(\cdot) \in \mathbb{R}^{n_c \times n_c}$ are state-dependent square matrices describing the inductances and capacitances, respectively. The function $g : \mathbb{R}^{n_r} \to \mathbb{R}^{n_r}$ describes the voltage-current relation of resistive elements.

Motivated by physical reasons we make some additional assumptions for the system (3),(19), including strong monotonicity, which allow us to decouple the circuit equations and to derive sufficient convergence criteria.

Definition 3.1 (Strong Monotonicity) *Let a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be given. Then, f is strongly monotone w.r.t. the second argument if there exists a constant $\mu_f > 0$ such that*

$$(y_2 - y_1)^\top (f(x, y_2) - f(x, y_1)) \geq \mu_f \|y_2 - y_1\|^2, \ \forall x \in \mathbb{R}^n, y_1, y_2 \in \mathbb{R}^m. \quad (20)$$

Note that for functions $\tilde{f}(x, y) := A(x)y$ which are linear w.r.t. y, the strong monotonicity criterion (20) reduces to the existence of a μ_A such that

$$y^\top A(x)y \geq \mu_A \|y\|^2, \ \forall x \in \mathbb{R}^n, y \in \mathbb{R}^m. \quad (21)$$

Assumption 3.2 *The system (3),(19) has the following properties:*

(i) g is strongly monotone, and the functions $\tilde{C}, \tilde{L}$, defined by $\tilde{C}(x, z) := C(x)z$, $\tilde{L}(x, z) := L(x)z$, are strongly monotone w.r.t. z.

(ii) $g(\cdot)$, $C(\cdot)$ and $L(\cdot)$ are C^1 and Lipschitz continuous, and q is C^2.

(iii) A_V has full column rank, and $\begin{pmatrix} A_C & A_V & A_R & A_L \end{pmatrix}$ has full row rank.

(iv) For any matrix Q_c with full column rank satisfying $\operatorname{im} Q_c = \ker A_C^\top$, the product $Q_c^\top A_V$ has full column rank.

Assumption (i) reflects the global passivity of the respective elements. Section 3.3 provides topological equivalents to the rank assumptions. Namely, Assumption (iii) excludes the electrically forbidden configurations of loops of voltage sources and cutsets of current sources. Assumption (iv) excludes loops of capacitances and voltage sources with at least one voltage source. The latter assumption is not necessary for the decoupling, but allows for a simpler derivation of convergence criteria.

3.1 Decoupling

We introduce the auxillary matrices Q_{cvr} and $\tilde{A}_L$ which allow us to describe the decoupling in a simple manner.

Definition 3.3 *We denote by Q_{cvr} a matrix with full column rank satisfying $\operatorname{im} Q_{cvr} = \ker \begin{pmatrix} A_C & A_V & A_R \end{pmatrix}^\top$. Furthermore, we denote $\tilde{A}_L := Q_{cvr}^\top A_L$.*

Lemma 3.4 *Let $\begin{pmatrix} A_C & A_V & A_R & A_L \end{pmatrix}$ have full row rank. Then, $\tilde{A}_L$ has full row rank.*

Proof: For any $y \in \ker \tilde{A}_L^\top$ and $x := Q_{cvr}y$ we get

$$\begin{aligned} \tilde{A}_L^\top y = 0 \implies A_L^\top Q_{cvr} y = 0 \implies A_L^\top x = 0 \ \wedge \ (A_C \ A_V \ A_R)^\top x = 0 \\ \implies (A_C \ A_V \ A_R \ A_L)^\top x = 0 \implies x = 0 \implies y = 0. \end{aligned}$$

□

As a consequence of Lemma 3.4, the matrix $\tilde{A}_L(L(i_l))^{-1}\tilde{A}_L^\top$ is invertible if the Assumption 3.2 is satisfied.

Theorem 3.5 *Let the Assumption 3.2 be satisfied and let M_1 be a matrix-valued function, defined by*

$$M_1(x) := (\tilde{A}_L(L(x))^{-1}\tilde{A}_L^\top)^{-1}(Q_{cvr}^\top\ 0).$$

Then, there exists a nonsingular transformation matrix $T = [T_0\ T_1\ T_2]$ with $T_1 = (Q_{cvr}^\top\ 0)^\top$ and Lipschitz continuous functions $f_0, f_1, f_2, h \in C^1$ with $h(y, z_2) = i_l$, such that the IVP

$$\dot{y} = f_0(y, z_1, z_2, s), \qquad y(t_0) = y_0 \tag{22}$$
$$z_1 = M_1(h(y, z_2))\dot{s} + f_1(y, s) \tag{23}$$
$$z_2 = f_2(y, s), \tag{24}$$

is equivalent to the IVP (3) with $x = [T_0\ T_1\ T_2](y^\top, z_1^\top, z_2^\top)^\top$.

A proof following the dissection concept presented in [7] is given in [12].

3.2 Convergence results

Combining Theorem (2.4) and Theorem (3.5), we obtain the following convergence theorem for coupled circuit systems.

Theorem 3.6 *Let the Assumptions 2.2 and 3.2 be satisfied. The Gauss-Seidel iteration (5)-(6),(19) converges in $C^0(\mathcal{I})$ to the exact solution of (1)-(2), (19) if the time interval $\mathcal{I}$ is sufficiently small and the spectral radius satisfies*

$$\rho\left[\hat{C}_1(x)Q_{cvr}(\tilde{A}_L(L(i_l))^{-1}\tilde{A}_L^\top)^{-1}Q_{cvr}^\top\hat{C}_2(u)\right] < 1 \quad \forall x \in \mathbb{R}^{n_x}, u \in \mathbb{R}^{n_u}, i_l \in \mathbb{R}^{n_l} \tag{25}$$

where $\hat{C}_1(x) \in \mathbb{R}^{n_u \times n_e}$ are the first n_u columns of $C_1(x) = c_1'(x)$ and $\hat{C}_2 \in \mathbb{R}^{n_e \times n_u}$ are the first n_e rows of $C_2(x) = c_2'(x)$.

Proof: Provided Assumptions 3.2 and 2.2 hold for a coupled circuit system (1)-(2), (19), the Decoupling Theorem 3.5 shows that Assumption 2.1 also holds. Hence, we can apply the Convergence Theorem 2.4. The Decoupling Theorem 3.5 provides a specific structure of the matrix T_1 and the matrix valued function $M_1(h(\cdot))$. Insertion into the critical matrix product (9) of the 2.4 yields

$$\begin{aligned} C_1(x)T_1M_1(h(y,z_2))C_2(u) &= C_1(x)\begin{pmatrix} Q_{cvr} \\ 0 \end{pmatrix}(\tilde{A}_L(L(x))^{-1}\tilde{A}_L^\top)^{-1}(Q_{cvr}^\top\ 0)C_2(u) \\ &= \hat{C}_1Q_{cvr}(\tilde{A}_L(L(x))^{-1}\tilde{A}_L^\top)^{-1}Q_{cvr}^\top\hat{C}_2(u). \end{aligned}$$

□

3.3 Topological interpretation

The next results show how the assumptions and the convergence criterion (25) can be interpreted topologically.

Definition 3.7 ((reduced) incidence matrix) *Let $\mathcal{G}$ be a graph with $N+1$ nodes and B oriented branches. Then, $\bar{A}$ is called the incidence matrix of $\mathcal{G}$ if $\bar{A} = (a_{ij}) \in \mathbb{R}^{(N+1)\times B}$ with*

$$a_{ij} = \begin{cases} +1 & \text{if node } i \text{ is start node of branch } j, \\ -1 & \text{if node } i \text{ is end node of branch } j, \\ 0 & \text{else.} \end{cases} \tag{26}$$

A matrix $A \in \mathbb{R}^{N\times B}$ obtained by deleting the row which corresponds to a freely chosen reference node is called reduced incidence matrix of $\mathcal{G}$.

Lemma 3.8 *Let $\mathcal{G}$ be a graph with $N+1$ nodes and incidence matrix $\bar{A}$. Then, $\mathcal{G}$ is connected if and only if $\operatorname{rk} \bar{A} = \operatorname{rk} A = N$.*

Proof: It is a well known result in circuit theory [4] that $\operatorname{rk} \bar{A} = \operatorname{rk} A = N$ if $\mathcal{G}$ is connected. If $\mathcal{G}$ is not connected then we find a connected component $\mathcal{G}_c$ of $\mathcal{G}$ that does not contain the reference node. Consequently, the sum of all rows of A belonging to nodes of $\mathcal{G}_c$ equals zero, i.e., A does not have full row rank. □

Remark. With Lemma 3.8 we see that $\begin{pmatrix} A_C & A_V & A_R & A_L \end{pmatrix}$ has full row rank as assumed in 3.2(*iii*) if and only if the circuit has no cutset of remaining branches, i.e., it has no cutset of current sources.

Lemma 3.9 *Let $\mathcal{G}$ be a graph with reduced incidence matrix A. Then, $\mathcal{G}$ is a forest if and only if A has full column rank.*

Proof: The columns of A are linearly independent if and only if the columns of $\bar{A}$ are linearly independent since the sum of all rows of $\bar{A}$ equals zero. Due to Lemma 3.8 the rank of each connected component $\mathcal{G}_c$ of $\mathcal{G}$ equals $n_c - 1$ when n_c denotes the number of nodes of $\mathcal{G}_c$. Consequently, the columns of the incidence matrix $\bar{A}_c$ of each connected component $\mathcal{G}_c$ are linearly independent if and only if $\bar{A}_c$ has $n_c - 1$ columns, i.e., if and only if $\mathcal{G}_c$ is a tree. Hence, $\bar{A}$ has full column rank if and only if $\mathcal{G}$ is a forest. □

Remark. With Lemma 3.9 we see that A_V has full column rank as assumed in 3.2(*iii*) if and only if the circuit has no loop of voltage sources.

Lemma 3.10 *Let $A : \mathbb{R}^n \to \mathbb{R}^m$, $B : \mathbb{R}^k \to \mathbb{R}^m$ be two linear mappings. It holds*

$$im\ B = \ker A^\top \quad \Leftrightarrow \quad \ker B^\top = im\ A.$$

Proof: Obviously,

$$\operatorname{im}\ B \subseteq \ker A^\top \quad \Leftrightarrow \quad A^\top B = 0 \quad \Leftrightarrow \quad B^\top A = 0 \quad \Leftrightarrow \quad \operatorname{im}\ A \subseteq \ker B^\top.$$

Considering the rank nullity theorem, we obtain

$$\operatorname{rk}\ B = \dim \ker A^\top \ \Leftrightarrow\ m - \operatorname{rk}\ B^\top = m - \dim \ker A^\top \ \Leftrightarrow\ \dim \ker B^\top = \operatorname{rk}\ A.$$

□

Lemma 3.11 *Let $\mathcal{G}$ be a graph, and let $\mathcal{G}_l$ be a subgraph with B branches and reduced incidence matrix A_l such that the branches of $\mathcal{G}_l$ form a loop $\mathcal{L}$ in $\mathcal{G}$. Then, for each column a_l^i of A_l, it holds*

$$\exists\ \lambda_1, \dots, \lambda_B \in \{1, -1\} : a_l^i = \sum_{j \neq i} \lambda_j a_l^j.$$

Proof: W.l.o.g., let the nodes be labeled such that $n_1, \dots, n_m \in \mathcal{L}$, and $n_{m+1}, \dots, n_k \notin \mathcal{L}$. Since a node $n \notin \mathcal{L}$ can not be incident with a branch $b \in \mathcal{L}$, it holds $A_l = (A_{l_*}^\top \ \ 0)^\top$, with A_{l_*} the (possibly reduced) incidence matrix of $\mathcal{L}$. We set an arbitrary orientation for the loop $\mathcal{L}$, and we define

$$\lambda_j = \begin{cases} 1, \text{ if branch } b_j \text{ is oriented in the sense of the loop } \mathcal{L}, \\ -1 \text{ if branch } b_j \text{ is oriented against the sense of the loop } \mathcal{L} \end{cases}$$

for $j = 1, \dots, B$. We denote the columns of A_{l_*} by $a_{l_*}^i$. Since each node of $\mathcal{L}$ connects precisely two branches in $\mathcal{L}$, each row of the matrix $(\lambda_1 a_{l_*}^1 \ \dots \ \lambda_B a_{l_*}^B)$, has one $+1$, one -1 and otherwise zeros as entries. It follows $\sum_{j=1}^B \lambda_j a_{l_*}^j = 0$, and hence $\sum_{j=1}^B \lambda_j a_l^j = 0$. □

Lemma 3.12 *Let $\mathcal{G}_1, \mathcal{G}_2$ be two graphs with identical node sets and disjoint branch sets, and let A_1, A_2 be their reduced incidence matrices. Let furthermore Q_1 be a matrix with full column rank such that im $Q_1 = \ker A_1^\top$. Then, $Q_1^\top A_2$ has full column rank if and only if there exists no loop in $\mathcal{G}_1 \cup \mathcal{G}_2$ with at least one branch $b \in \mathcal{G}_2$.*

Proof: " $\Longrightarrow$ " First, we note that full column rank of $Q_1 A_2$ implies full column rank of A_2. Furthermore, we obtain

$$\ker Q_1^\top A_2 = \{0\} \implies \text{im } A_2 \cap \ker Q_1^\top \{0\} \overset{3.10}{\Longrightarrow} \text{im } A_2 \cap \text{im } A_1 = \{0\}.$$

Hence, no column of A_2 can be represented as a linear combination of columns of A_1. Lemma 3.11 implies that if a column of the reduced incidence matrix can not be represented as a linear combination of other columns of the reduced incidence matrix, then the corresponding branch is not element of a loop, which yields the desired result.
" $\Longleftarrow$ " Let $\mathcal{F}_1$ be an arbitrary maximal spanning forest of $\mathcal{G}_1$ with incidence matrix $A_{\mathcal{F}_1}$. Then, $\mathcal{F}_1 \cup \mathcal{G}_2$ is a maximal spanning forest of $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ since loops exist only in A_1, but not in A_2. Thus, the incidence matrix $(A_{\mathcal{F}_1} \ A_2)$ has full column rank with Lemma 3.9. This implies im $A_{\mathcal{F}_1} \cap$ im $A_2 = \{0\}$. Together with Lemma 3.10 and $\ker A_2 = \{0\}$, we obtain the desired result. □

Remark. Lemma 3.12 implies that $Q_c^\top A_V$ has full column rank as assumed in 3.2(*iv*) if and only if the circuit has no loops of capacitances and voltage sources with at least one voltage source.
In the following, we choose a specific matrix for Q_{cvr}. The definition of it requires so called CVR-components.

Definition 3.13 *A CVR-component of a graph is a maximal connected subgraph which consists of only capacitances, voltage sources and resistances and their incident nodes.*

Remark. A CVR-component can consist of only one node.
We consider an electrical circuit with $c+1$ CVR-components $S_0, S_1, \dots, S_c$, where w.l.o.g. the reference node belongs to S_0. We choose

$$(Q_{cvr})_{ij} := \begin{cases} 1, & \text{if node } i \in S_j, j \geq 1 \\ 0, & \text{else.} \end{cases}$$

Clearly, this choice of Q_{cvr} complies with Definition 3.3. For linear coupling functions $c_1 = C_1, c_2 = C_2$, we obtain the following corollary.

Corollary 3.14 *Let an electrical circuit with the nodes $n_0, \dots, n_N$ and $c+1$ CVR-components $S_0, S_1, \dots, S_c$ be given, and w.l.o.g. let the reference node belong to S_0. Let the Assumptions 2.2, 3.2 and 3.2 be satisfied. The Gauss-Seidel iteration (5)-(6),(19) converges with $\rho = 0$ in $C^0(\mathcal{I})$ to the exact solution of (1)-(2),(19),if one of the following criteria is satisfied.*

(1) $\hat{C}_1 = 0$, i.e. the node potentials e do not contribute to the ODE.

(2) $\hat{C}_2 = 0$, i.e. the ODE-variable u does not contribute to the Kirchhoff node equations of the circuit.

(3) $\sum_{n_j \in S_k, k \geq 1} (\hat{C}_1)_{ij} = 0 \ \forall i$, i.e. for each CVR-component $S_k, k \geq 1$, the sum $\sum_{n_i \in S_k} \lambda_i e_i$ of node potentials coupled into the ODE can be written as a sum of differences $\sum_{n_i, n_j \in S_k} \mu_{ij}(e_i - e_j)$.

(4) $\sum_{n_i \in S_k, k \geq 1} (\hat{C}_2)_{ij} = 0 \ \forall j$, i.e. in each CVR-component $S_k, k \geq 1$, the external current inflow equals the external current outflow.

The first two criteria follow trivially from Theorem 3.6. Considering the all-ones structure of Q_{cvr} within CVR-components, the latter two criteria lead to the products $\hat{C}_1 Q_{cvr} = 0$ and $Q_{cvr}^\top \hat{C}_2 = 0$. Then, $\rho = 0$ follows again with 3.6.
Note that the first condition implies the third one, and the second condition implies the fourth one.

3.4 Examples

The example of a linear circuit with a two-terminal coupling as shown in Figure 3.4 illustrates Corollary 3.14. The blue inductance is a simple representative of the ODE in (1), whereas the red part of the circuit represents the circuit subsystem (2),(19). The element equation for the inductance of the ODE-subsystem is $i'_{l_1} = \frac{e_1 - e_2}{L_1}$, where e_1, e_2 are node potentials of the red cirucit subsystem which are coupled into the blue ODE-subsystem. Hence, $\hat{C}_1 = \frac{1}{L}(1 \ -1 \ 0)$ and Criterion 3.14(3) is satisfied. Indeed, the coupling nodes n_1, n_2 belong to the same CVR-component which contains n_1, n_2, n_3, and the circuit contributes the difference $\frac{1}{L_1}(e_1 - e_2)$ to the ODE. Since the external current inflow into and outflow of this CVR-component are equally given by i_{l_1}, Criterion 3.14(4) is also satisfied.
In the similar circuit in Figure 3.4, only the choice of one coupling node has changed, which is now n_r instead of n_2. Since these nodes belong to different CVR-components, $\frac{1}{L_1} e_1$ is the only node potential of the corresponding

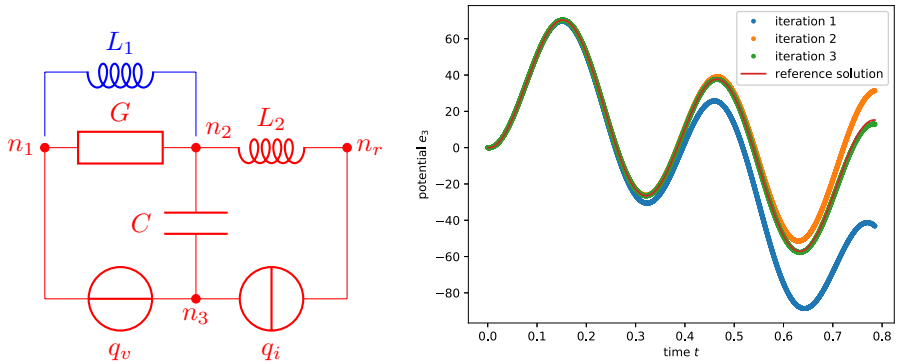

Figure 1: *The left side shows an electrical circuit with inductances L_1, L_2, conductance G, capacitance C, and independent sources providing current q_i and voltage q_v. The coupling nodes n_1, n_2 belong to the same CVR-component. The right side shows simulation plots of the potential e_3 at n_3. The reference solution is monolithical. The graphs "Iteration $1, 2, 3$" are the respective solutions of the Gauss-Seidel WR method. We observe convergent behaviour w.r.t. the iteration parameter k, which is in accordance with Corollary 3.14.*

CVR-component which is coupled into the ODE, and we cannot write this as a potential difference of the CVR-component. Since also the other criteria of Corollary 3.14 are not satisfied in this case, we cannot guarantee convergence of the Gauss-Seidel WR method. Indeed, it is divergent as Figure 3.4 shows.

4 Conclusions

With Theorem 3.6, a general convergence criterion for the Gauss-Seidel WR method applied to an index-2 DAE coupled with an ODE has been presented. It requires that the DAE can be decoupled as given in Assumption 2.1. For circuit systems, such a decoupling is presented in Theorem 3.5. Theorem 3.6 describes the convergence criterion for circuit DAEs in terms of the coupling and particular transformation matrices. Certain topological properties of the circuit are maintained in the transformed system, which is shown in Section 3.3. The topological considerations allowed us to provide simple network topological criteria leading to convergence of the Gauss-Seidel waveform relaxation, see Corollary 3.14. Finally, we demonstrated the relevance of the presented convergence criteria by the examples presented in Figure 3.4 (convergence) and Figure 3.4 (divergence).

Acknowledgement

This work is supported by the German Ministry of Economic Affairs and Energy and (BMWi) within the project MathEnergy, 0324019E.

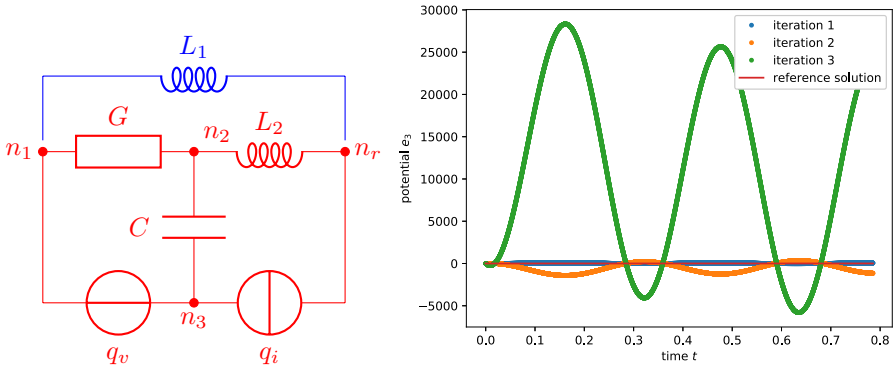

Figure 2: *The electrical circuit on the left differs from the circuit presented in Figure 3.4 only in the coupling nodes, i.e. n_r instead of n_2 is the second coupling node here. All element parameters and the topology of the red circuit remain the same. The coupling nodes n_1, n_r now belong to two different CVR-components. In contrast to Figure 3.4, the simulation plots on the right show a highly divergent behaviour of the Gauss-Seidel WR method w.r.t. the iteration parameter k. This is possible since 3.14 does not apply here.*

References

[1] Arnold, M., Günther, M.: Preconditioned Dynamic Iteration for Coupled Differential-Algebraic Systems. BIT, **41**, 1–25 (2001)

[2] Bartel, A., Brunk, M., Günther, M., Schöps, S.: Dynamic iteration for coupled problems of electric circuits and distributed devices SIAM Journal on Scientific Computing, **35(2)**, B315–B335 (2013)

[3] Burrage, K.: Parallel and Sequential Methods for Ordinary Differential Equations. Oxford University Press (1995)

[4] Chua, L., Desoer, C., Kuh, E.: Linear and Nonlinear Circuits. McGraw-Hill (1987)

[5] Fan, X.-G., Zou, J.-H., Sun, W.: Convergence of Parallel Dynamic Iteration Methods for Nonlinear DAEs of Index-2. Proceeding of the IEEE International Conference on Automation Science and Engineering (2006).

[6] Jackiewicz, Z., Kwapisz, M.: Convergence of waveform relaxation methods for differential-algebraic systems. SIAM Journal on Numerical Analysis, Vol. 33(6), 2303-2317, 1996.

[7] Jansen, L.: A Dissection concept for DAEs - structural decoupling, unique solvability, convergence theory and half-explicit methods. Ph.D. Thesis, Humboldt-Universität zu Berlin (2015).

[8] Lamour, R., März, R., Tischendorf, C.: Differential-Algebraic Equations: A Projector Based Analysis. Springer (2013)

[9] Lelarasmee, E.: The Waveform Relaxation Method for Time Domain Analysis of Large Scale Integrated Circuits: Theory and Applications. Ph.D. thesis, University of California, Berkeley (1982). URL https://www2.eecs.berkeley.edu/Pubs/TechRpts/1982/9614.html

[10] Lelarasmee, E., Ruehli, A.E., Sangionvanni-Vincentelli, A.L.: The Waveform Relaxation Method for Time Domain Analysis of Large Scale. Integrated Circuits (Vol.81), Memorandum No. UCB/ERL, 1982.

[11] Miekkala, U.: Dynamic iteration methods applied to linear DAE systems. International Journal of Circuit Theory and Applications, **28**, 131–162 (2000).

[12] Pade, J.: Convergence criteria for waveform relaxation on differential-algebraic systems: a topological approach for circuits. Ph.D. Thesis, Humboldt-Universität zu Berlin (In preparation, 2018).

[13] Schöps, S.: Multiscale Modeling and Multirate Time-Integration of Field/Circuit Coupled Problems. Ph.D. Thesis, Bergische Universität Wuppertal (2011). URL http://elpub.bib.uni-wuppertal.de/edocs/dokumente/fbc/mathematik/diss2011/schoeps

[14] Schöps, S., De Gersem, H., Bartel, A.: A cosimulation framework for multirate time integration of field/circuit coupled problems. IEEE Transactions on Magnetics, **46(8)**, 3233–3236 (2010)

[15] Walter, W.: Ordinary Differential Equations. Springer (1998)