Journal of Computational and Applied Mathematics 235 (2011) 2790–2804

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

On convergence conditions of waveform relaxation methods for linear differential-algebraic equations☆

Zhong-Zhi Bai [a,b,*], Xi Yang [b]

[a] School of Mathematics and Computer Science, Guizhou Normal University, Guiyang 550001, PR China

[b] State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, PR China

ARTICLE INFO

Article history:
Received 24 March 2010
Received in revised form 2 September 2010

MSC:
44A10
65L05
65L20
65L80
65N15
CR: G1.3

Keywords:
Differential-algebraic equations
Laplace transform
Waveform relaxation method
Convergence theory

ABSTRACT

For linear constant-coefficient differential-algebraic equations, we study the waveform relaxation methods without demanding the boundedness of the solutions based on infinite time interval. In particular, we derive explicit expression and obtain asymptotic convergence rate of this class of iteration schemes under weaker assumptions, which may have wider and more useful application extent. Numerical simulations demonstrate the validity of the theory.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Differential-algebraic equations (**DAEs**) can be generally formulated as implicit differential equations of the form

$$F(\dot{x}(t), x(t), t) = 0, \tag{1.1}$$

where F and x are vector-valued functions, and $\dot{x}(t)$ denotes the derivative of $x(t)$ with respect to t. If the Jacobian of F with respect to $\dot{x} := \dot{x}(t)$, denoted as $\frac{\partial F}{\partial \dot{x}}$, is nonsingular, then the system (1.1) defines implicit *ordinary differential equations* (**ODEs**), and by the implicit function theorem we can solve $\dot{x}$ from the system (1.1) to obtain the explicit ordinary differential equations

$$\dot{x}(t) = G(x(t), t), \tag{1.2}$$

where G is a properly defined functional such that

$$F(G(x(t), t), x(t), t) = 0.$$

☆ Supported by The National Basic Research Program (No. 2005CB321702, No. 2011CB309703), The Hundred Talent Project of Chinese Academy of Sciences, and The China Outstanding Young Scientist Foundation (No. 10525102), PR China.

* Corresponding author at: State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, PR China.
E-mail addresses: bzz@lsec.cc.ac.cn (Z.-Z. Bai), yangxi@lsec.cc.ac.cn (X. Yang).

0377-0427/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2010.11.031

Through this approach, we can solve implicit ODEs via solving explicit ODEs. If the Jacobian $\frac{\partial F}{\partial \dot{x}}$ is singular, then the system (1.1) defines differential-algebraic equations, as now the function $\dot{x}(t)$ cannot be expressed in the explicit form (1.2).

The waveform relaxation methods can be regarded as natural extensions of the classical relaxation methods for solving systems of algebraic equations with iterating space changing from $\mathbb{R}^n$ to the time-dependent functional or the waveform space. They are powerful solvers for numerically computing the solution of the DAEs (2.1) on both sequential and parallel computers; see, e.g., [1–4]. The basic idea of this class of iteration methods is to apply the relaxation technique directly to the DAEs. As a result, the system of DAEs is first decomposed into decoupled sub-systems, and each sub-system is then analyzed for the entire simulation time interval by means of standard simulation techniques; see, for example, [5,6]. Appropriately designed waveform relaxation method can be efficiently implemented on a parallel computer by distributing the sub-systems to different processors. Each processor is then responsible for computing the update of a sub-system. A task of large enough scale often leads to very satisfactory parallel performance.

The waveform relaxation method was first introduced by Lelarasmee in [5] for simulating the behavior of very large-scale electrical networks. Lelarasmee proved that the waveform relaxation method is convergent as long as the splitting function of the nonlinear system is Lipschitz continuous. Later, there are lots of expansions and applications of this theory; see [4]. This theory can sufficiently guarantee the convergence of the waveform relaxation method. There is, however, no sufficient and necessary convergence condition, neither precise description about the convergence rate. Miekkala is the first who studied the waveform relaxation method for solving linear DAEs in a more precise point of view, and obtained convergence rate of the waveform relaxation method; see [7,8]. As is known, the convergence of the waveform relaxation methods demands a strict restriction that the solution of the DAEs (2.1) is uniformly bounded; see [7–12] for details.

The main purpose of this paper is to establish the convergence theory for the waveform relaxation methods without the boundedness assumption on the solution of the DAEs (2.1). The organization of the paper is as follows. We briefly review the solvability and the Kronecker canonical form of the DAEs (2.1) in Section 2. In Section 3, we present the basic theory of the Laplace transform, establish a few useful lemmas and discuss the choice of the function space. The waveform relaxation method is precisely described and the explicit expression of the iteration is carefully derived in Section 4. In Section 5 we demonstrate the convergence and the asymptotic convergence rate of the waveform relaxation method. Numerical results are shown in Section 6 to verify our theory. Finally, we give a few conclusions and remarks in Section 7.

2. Solvability of DAEs

In what follows, we assume that the solution of the DAEs (1.1) exists and is uniquely defined on an interval of interest. We first introduce the concept of solvability proposed by Petzold in [13].

Definition 2.1. Let $\mathcal{I}$ be an open sub-interval of $\mathbb{R}^1$, Ω be a connected open subset of $\mathbb{R}^{2m+1}$, and F a differentiable function from Ω to $\mathbb{R}^m$. Then the DAEs (1.1) is solvable on $\mathcal{I}$ in Ω if there is an r-dimensional family of solutions $\phi(t, c)$ defined on a connected open set $\mathcal{I} \times \tilde{\Omega}$, with $\tilde{\Omega} \subset \mathbb{R}^r$, such that

(a) $\phi(t, c)$ is defined on all of $\mathcal{I}$ for each $c \in \tilde{\Omega}$;
(b) $(\dot{\phi}(t, c), \phi(t, c), t) \in \Omega$ for $(t, c) \in \mathcal{I} \times \tilde{\Omega}$;
(c) if $\psi(t)$ is any other solution with $(\dot{\phi}(t, c), \phi(t, c), t) \in \Omega$, then $\psi(t) = \phi(t, c)$ for some $c \in \tilde{\Omega}$;
(d) the graph of $\phi(t, c)$ as a function of (t, c) is an $(r + 1)$-dimensional manifold.

Here, $\dot{\phi}(t, c)$ represents the derivative of $\phi(t, c)$ with respect to t.

Consider the initial-value problem of linear constant-coefficient DAEs

$$B\dot{x}(t) + Ax(t) = f(t), \quad x(0) = x_0, \tag{2.1}$$

where B and A are square matrices of suitable dimensions. In looking for solutions of the form $x(t) = e^{\lambda t}x_0$ (for $f(t) = 0$), we are led to consider the matrix pencil $\lambda B + A$. If the determinant of $\lambda B + A$, denoted as $\det(\lambda B + A)$, is not identically zero as a function of λ, then the pencil is said to be regular. Solvability as defined in Definition 2.1 can be difficult to determine for a general DAEs, but for the special case (2.1) there is a nice characterization.

Theorem 2.1 (*[13,14]*). *The linear constant-coefficient DAEs (2.1) is solvable if and only if the matrix pencil $\lambda B + A$ is regular.*

If the determinant of a matrix pencil $\lambda B + A$ is identically zero as a function of λ, the DAEs (2.1) has either no solution or infinitely many solutions for a given initial value x_0. We shall therefore deal with regular matrix pencils. Recall that a matrix N is said to have nilpotency k if $N^k = 0$ and $N^{k-1} \neq 0$. An essential approach for computing the solution of (2.1) is to simultaneously transform the matrices B and A into Kronecker canonical forms. This property is precisely described in the following theorem.

Theorem 2.2 (*[13,15]*). *Let B and A be given square matrices of the same dimension. Suppose that $\lambda B + A$ is a regular matrix pencil. Then there exist nonsingular matrices P and Q such that*

$$PBQ = \begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix} \quad \text{and} \quad PAQ = \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix}, \tag{2.2}$$

where N is a matrix of nilpotency k and I is the identity matrix. Here, for $N = 0$ we define $k = 1$. For B being particularly nonsingular, we take $PBQ = I$, $PAQ = C$ and $k = 0$. Moreover, when $\det(\lambda B + A)$ *is identically constant, (2.2) can be simplified to $PBQ = N$ and $PAQ = I$. Here and in what follows, I is used to denote the identity matrix of suitable dimension.*

Remark 2.1. According to the proof of Theorem 2.2 in [15], when both A and B are real matrices, we can choose real matrices P and Q to transform the matrix pencil $\lambda B + A$ into the Kronecker canonical form (2.2).

3. Laplace transform and function space

Laplace transform is one of the most useful and important tools in mathematical analysis. In order to make this paper more self-contained, we first review the basic definitions related to the Laplace transform [16]. Furthermore, some fundamental theorems of the Laplace transform can be found in the Appendix.

Definition 3.1 (*Causal Function*)**.** A continuous time-dependent function $f : \mathbb{R}^1 \to \mathbb{C}^1$ is called causal if $f(t) = 0$, for $t < 0$.

Note that if $f(t)$ is a causal function, then $f(0) = 0$ holds true.

Definition 3.2 (*Laplace Transform*)**.** Let $f : \mathbb{R}^1 \to \mathbb{C}^1$ be a causal function. Then the Laplace transform $F(s)$ of $f(t)$ is a complex function defined, for all $s \in \mathbb{C}^1$, by

$$F(s) = \int_0^\infty f(t)\mathrm{e}^{-st}\,\mathrm{d}t \triangleq (\mathcal{L}f)(s),$$

provided the integral exists.

Definition 3.3 (*Abscissa of Absolute Convergence*)**.** For a given causal function $f : \mathbb{R}^1 \to \mathbb{C}^1$, there exists a number $\sigma_a \in \mathbb{R}^1$ with $-\infty \le \sigma_a \le \infty$ such that the integral $\int_0^\infty f(t)\mathrm{e}^{-st}\,\mathrm{d}t$ is absolutely convergent for all $s \in \mathbb{C}^1$ with $\mathrm{Re}(s) > \sigma_a$, and not absolutely convergent for all $s \in \mathbb{C}^1$ with $\mathrm{Re}(s) < \sigma_a$. Here, by $\sigma_a = -\infty$ we mean that the integral is absolutely convergent for all $s \in \mathbb{C}^1$, and by $\sigma_a = \infty$ we mean that the integral is absolutely convergent for no $s \in \mathbb{C}^1$. The number σ_a is called the abscissa of absolute convergence.

Definition 3.4 (*Exponential Order*)**.** The causal function $f : \mathbb{R}^1 \to \mathbb{C}^1$ is of exponential order if there are constants $\alpha \in \mathbb{R}^1$ and $æ > 0$ such that $|f(t)| \le æ\mathrm{e}^{\alpha t}$ for all $t \ge 0$.

Remark 3.1. The set of all functions of exponential orders is denoted as

$$\mathcal{E} = \{f(t) : f(t) \text{ is a causal function of exponential order}\}.$$

Note that $\mathcal{E}$ is a vector space.

Definition 3.5 (*Abscissa of Function Space*)**.** Let $\mathcal{E}$ be the vector space of all functions of exponential orders and $X \subset \mathcal{E}$. Then there exists a number $\sigma_X \in \mathbb{R}^1$ such that for any $x := x(t) \in X$ the functional $\mathcal{L}x$ is absolutely convergent for all $s \in \mathbb{C}^1$ satisfying $\mathrm{Re}(s) > \sigma_X + \epsilon$, with $\epsilon > 0$ being arbitrary, and not all $\mathcal{L}x$ is absolutely convergent for all $s \in \mathbb{C}^1$ satisfying $\mathrm{Re}(s) < \sigma$, with $\sigma < \sigma_X$. That is to say,

$$\sigma_X = \inf\{\sigma \in \mathbb{R}^1 : \mathcal{L}x \text{ is absolutely convergent}, \forall\, x \in X,\ \forall\, \mathrm{Re}(s) > \sigma\}.$$

Such a real number σ_X is called the abscissa of the function space X.

We have already known the convolution product of two functions. When, moreover, f and g are both causal functions, we have, for $t > 0$, that

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,\mathrm{d}\tau = \int_0^t f(\tau)g(t-\tau)\,\mathrm{d}\tau,$$

since the integrand is zero for both $\tau < 0$ and $t - \tau < 0$. Moreover, we can easily prove the existence of the convolution product.

Lemma 3.1. *Let $f, g : \mathbb{R}^1 \to \mathbb{C}^1$ be two causal functions of exponential orders. Suppose that $\alpha, \beta \in \mathbb{R}^1$ and $æ > 0$ are constants such that*

$$|f(t)| \le æ\mathrm{e}^{\alpha t} \quad \text{and} \quad |g(t)| \le æ\mathrm{e}^{\beta t}.$$

Then

(a) *if $\alpha \ne \beta$, there exists a constant $\widetilde{æ} > 0$ satisfying*

$$|(f * g)(t)| \le \widetilde{æ}\mathrm{e}^{\max\{\alpha,\beta\}t};$$

(b) *if $\alpha = \beta$, for any $\epsilon > 0$ there exists a constant $\widetilde{æ}_\epsilon > 0$ satisfying*

$$|(f * g)(t)| \le \widetilde{æ}_\epsilon \mathrm{e}^{(\alpha+\epsilon)t}.$$

Proof. Because f and g are causal functions, we know that the convolution of f and g exists.

For $\alpha \neq \beta$, without loss of generality, we may assume that $\alpha < \beta$. Then direct computations lead to the estimate

$$\begin{aligned}
|(f * g)(t)| &= \left| \int_0^t f(\tau) g(t-\tau)\, \mathrm{d}\tau \right| \\
&\leq \int_0^t |f(\tau)||g(t-\tau)| \mathrm{d}\tau \\
&\leq æ^2 \int_0^t \mathrm{e}^{\alpha\tau} \mathrm{e}^{\beta(t-\tau)} \mathrm{d}\tau \\
&= æ^2 \mathrm{e}^{\beta t} \int_0^t \mathrm{e}^{(\alpha-\beta)\tau}\, \mathrm{d}\tau \\
&= \frac{1}{\beta-\alpha} æ^2 (1 - \mathrm{e}^{-(\beta-\alpha)t}) \mathrm{e}^{\beta t} \\
&\leq \frac{1}{\beta-\alpha} æ^2 \mathrm{e}^{\beta t} \\
&= \widetilde{æ} \mathrm{e}^{\max\{\alpha,\beta\} t}.
\end{aligned}$$

For $\alpha = \beta$, we can similarly obtain the estimate

$$\begin{aligned}
|(f * g)(t)| &= \left| \int_0^t f(\tau) g(t-\tau)\, \mathrm{d}\tau \right| \\
&\leq \int_0^t |f(\tau)||g(t-\tau)|\, \mathrm{d}\tau \\
&\leq æ^2 \int_0^t \mathrm{e}^{\alpha\tau} \mathrm{e}^{\beta(t-\tau)}\, \mathrm{d}\tau \\
&= æ^2 t \mathrm{e}^{\beta t}.
\end{aligned}$$

It follows from the above derivations that for any $\epsilon > 0$ there exists a constant $æ_\epsilon$ such that $t \leq æ_\epsilon \mathrm{e}^{\epsilon t}$. Therefore, it holds that

$$|(f * g)(t)| \leq æ^2 t \mathrm{e}^{\beta t} \leq æ^2 æ_\epsilon \mathrm{e}^{(\beta+\epsilon)t} = \widetilde{æ}_\epsilon \mathrm{e}^{(\beta+\epsilon)t},$$

where $\widetilde{æ}_\epsilon = æ^2 æ_\epsilon$. This completes the proof. □

For the initial-value problem of linear constant-coefficient DAEs (2.1), according to [17] we assume that its solution $x(t)$ satisfies $x(t) \in L^1_{\mathrm{loc}}(\bar{\mathbb{R}}^1_+)$, where $\bar{\mathbb{R}}^1_+ = [0, +\infty)$ and

$$L^1_{\mathrm{loc}}(\bar{\mathbb{R}}^1_+) := \{x(t) : x(t) \in L^1(\Omega) \text{ is locally integrable for every open } \Omega \subset \bar{\mathbb{R}}^1_+\}.$$

Let $\{t_n\}_{n=0}^{\infty}$ be an increasing series of positive reals satisfying $\lim_{n\to\infty} t_n = \infty$, and denote by $\Omega_n = [0, t_n]$. Then we have $\bar{\mathbb{R}}^1_+ = \cup_{n=0}^{\infty} \Omega_n$. In accordance with the theory of Sobolev space [17,18], we know that the set of restrictions to Ω_n of the functions in $\mathbb{C}^\infty_0(\mathbb{R}^1) \subset \mathbb{C}^\infty(\mathbb{R}^1)$ is dense in $L^1(\Omega_n)$. This shows that the solution $x(t)$ of the DAEs (2.1), when restricted to Ω_n, can be well approximated by the functions in $\mathbb{C}^\infty_0(\mathbb{R}^1) \subset \mathbb{C}^\infty(\mathbb{R}^1)$. Noticing that $\lim_{n\to\infty} \Omega_n = \bar{\mathbb{R}}^1_+$, we can then conclude that the above-mentioned approximation can be done on the domain $\bar{\mathbb{R}}^1_+$.

For a given small positive real number γ, we define a function space on $\bar{\mathbb{R}}_\gamma = [\gamma, +\infty)$ as

$$\mathbb{C}^\infty_\gamma(\bar{\mathbb{R}}_\gamma) = \left\{x(t) : \|x(t)\|_{\gamma,\infty} < \infty\right\},$$

where the norm $\|\cdot\|_{\gamma,\infty}$ is defined as

$$\|x(t)\|_{\gamma,\infty} = \sup_{m\geq 0} \max_{t\in\bar{\mathbb{R}}_\gamma} \left| \frac{\mathrm{d}^m x(t)}{\mathrm{d}t^m} \right|.$$

Then the following conclusion holds true.

Theorem 3.1. *The function space $\mathbb{C}^\infty_\gamma(\bar{\mathbb{R}}_\gamma)$ is complete under the norm $\|\cdot\|_{\gamma,\infty}$.*

Proof. We denote by

$$\|x(t)\|_{\gamma,k} = \sup_{0\leq m\leq k} \max_{t\in\bar{\mathbb{R}}_\gamma} \left| \frac{\mathrm{d}^m x(t)}{\mathrm{d}t^m} \right|.$$

Let $\{x_k(t)\}_{k=0}^{\infty} \subset \mathbb{C}_{\gamma}^{\infty}(\bar{\mathbb{R}}_{\gamma})$ be a Cauchy sequence. Then for any $\varepsilon > 0$, there exists a positive integer n_o such that

$$\|x_n(t) - x_m(t)\|_{\gamma,\infty} < \varepsilon, \quad \forall n, m \geq n_o.$$

It follows immediately that

$$\|x_n(t) - x_m(t)\|_{\gamma,k} < \varepsilon, \quad \forall k \geq 0.$$

Hence, there exist continuous functions $x^{(k)}$ and $x^{(k+1)}$ such that

$$\frac{\mathrm{d}^k x_n(t)}{\mathrm{d}t^k} \longrightarrow x^{(k)} \quad \text{and} \quad \frac{\mathrm{d}^{k+1} x_n(t)}{\mathrm{d}t^{k+1}} \longrightarrow x^{(k+1)}, \quad \forall k \geq 0,$$

where the symbol "$\longrightarrow$" means that a sequence of functions converges uniformly to a limiting function. Moreover, we have

$$x^{(k+1)} = \frac{\mathrm{d}x^{(k)}}{\mathrm{d}t}.$$

So, we see that there exists a function $x(t)$ satisfying

$$\frac{\mathrm{d}^k x_n(t)}{\mathrm{d}t^k} \longrightarrow \frac{\mathrm{d}^k x(t)}{\mathrm{d}t^k}, \quad \forall k \geq 0.$$

Obviously, for any $k \geq 0$, there exists a positive integer $\bar{n}_k > n_o$ such that

$$\|x_{\bar{n}_k}(t) - x(t)\|_{\gamma,k} < \varepsilon.$$

Therefore, it holds that

$$\begin{aligned} \|x(t)\|_{\gamma,k} &\leq \|x(t) - x_{\bar{n}_k}(t)\|_{\gamma,k} + \|x_{\bar{n}_k}(t) - x_{n_o+1}(t)\|_{\gamma,k} + \|x_{n_o+1}(t)\|_{\gamma,k} \\ &\leq 2\varepsilon + \|x_{n_o+1}(t)\|_{\gamma,k} \\ &\leq 2\varepsilon + \|x_{n_o+1}(t)\|_{\gamma,\infty}, \quad \forall k \geq 0. \end{aligned}$$

This clearly shows that

$$\|x(t)\|_{\gamma,\infty} \leq 2\varepsilon + \|x_{n_o+1}(t)\|_{\gamma,\infty} < \infty$$

and, hence, $x(t) \in \mathbb{C}_{\gamma}^{\infty}(\bar{\mathbb{R}}_{\gamma})$.

In addition, because for any $k \geq 0$ it holds that

$$\|x_n(t) - x_m(t)\|_{\gamma,k} < \varepsilon, \quad \forall n, m \geq n_o,$$

and there exists a positive integer $\bar{m}_k > n_o$ such that

$$\|x_{\bar{m}_k}(t) - x(t)\|_{\gamma,k} < \varepsilon,$$

we have

$$\|x_n(t) - x(t)\|_{\gamma,k} \leq \|x(t) - x_{\bar{m}_k}(t)\|_{\gamma,k} + \|x_{\bar{m}_k}(t) - x_n(t)\|_{\gamma,k} < 2\varepsilon, \quad \forall k \geq 0.$$

This straightforwardly implies that

$$\|x_n(t) - x(t)\|_{\gamma,\infty} \leq 2\varepsilon, \quad \forall n \geq n_o.$$

Therefore, $\mathbb{C}_{\gamma}^{\infty}(\bar{\mathbb{R}}_{\gamma})$ is complete. □

Define a function space

$$\mathbb{T}_{\sigma,\gamma} = \left\{x(t) : \mathrm{e}^{-\sigma t} x(t) \in \mathbb{C}_{\gamma}^{\infty}(\bar{\mathbb{R}}_{\gamma})\right\}.$$

The norm equipped in $\mathbb{T}_{\sigma,\gamma}$ is

$$\|x(t)\|_{\sigma} = \|\mathrm{e}^{-\sigma t} x(t)\|_{\gamma,\infty}.$$

Since $\mathbb{C}_{\gamma}^{\infty}(\bar{\mathbb{R}}_{\gamma})$ is complete, $\mathbb{T}_{\sigma,\gamma}$ is also complete under the norm $\|\cdot\|_{\sigma}$. We smoothly prolong $x(t) \in \mathbb{T}_{\sigma,\gamma}$ as a causal function and consider different prolongations of $x(t)$ as the same function. Denote by $\mathbb{T}_{\sigma}$ the set of all the above prolonged functions, and define the norm in $\mathbb{T}_{\sigma}$ as $\|\cdot\|_{\sigma}$ again. Then we conclude that $\mathbb{T}_{\sigma}$ is complete under the norm $\|\cdot\|_{\sigma}$. In the remainder, we will concentrate on the space $\mathbb{T}_{\sigma}$. It is straightforward that the abscissa of $\mathbb{T}_{\sigma}$ is $\sigma_{\mathbb{T}_{\sigma}} = \sigma$.

Furthermore, we can demonstrate that the solution of the DAEs (2.1) is of exponential order. This result is precisely described in the following lemma.

Lemma 3.2. *For the initial-value problem of linear constant-coefficient DAEs (2.1) with the regular matrix pencil $\lambda B + A$, let $f(t) \in \mathbb{T}_\sigma^n$. Then its solution $x(t)$ is of exponential order, where σ is a real number and $\mathbb{T}_\sigma^n := \mathbb{T}_\sigma \times \mathbb{T}_\sigma^{n-1}$, $n = 2, 3, \ldots$, with $\mathbb{T}_\sigma^1 = \mathbb{T}_\sigma$.*

Proof. According to Theorem 2.2, there exist nonsingular matrices P and Q such that

$$PBQ = \begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix} \quad \text{and} \quad PAQ = \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix}.$$

Let $x(t) = Qy(t)$ and post-multiply P on both sides of (2.1). Then we obtain

$$PBQ\,\dot{y}(t) + PAQ\,y(t) = Pf(t), \quad y(0) = y_0 \triangleq Q^{-1}x_0,$$

or equivalently,

$$\begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} + \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} \widehat{f}_1(t) \\ \widehat{f}_2(t) \end{pmatrix}, \qquad y(0) = \begin{pmatrix} y_{10} \\ y_{20} \end{pmatrix},$$

with $\widehat{f}(t) = Pf(t)$. In block vector form, the above equations can be rewritten as

$$\begin{cases} \dot{y}_1(t) + Cy_1(t) = \widehat{f}_1(t), \\ N\dot{y}_2(t) + y_2(t) = \widehat{f}_2(t), \end{cases}$$

which has the explicit solution

$$\begin{cases} y_1(t) = \mathrm{e}^{-tC}y_{10} + \displaystyle\int_0^t \mathrm{e}^{-(t-\tau)C}\widehat{f}_1(\tau)\,\mathrm{d}\tau, \\ y_2(t) = \left(I + N\dfrac{\mathrm{d}}{\mathrm{d}t}\right)^{-1}\widehat{f}_2(t). \end{cases}$$

It follows immediately from $f(t) \in \mathbb{T}_\sigma^n$ that $y(t)$ is of exponential order and, hence, $x(t) = Qy(t)$ is of exponential order, too. □

Remark 3.2. Since the solution of the DAEs (2.1) is of exponential order, it is reasonable to discuss the waveform relaxation method for (2.1) in the space $\mathbb{T}_\sigma^n$ of functionals of exponential orders.

4. Explicit expression of waveform relaxation method

In this section, we are going to propose a new explicit expression of the waveform relaxation method for the DAEs (2.1). The basic assumption on the DAEs (2.1) is that the corresponding matrix pencil $\lambda B + A$ is regular. By splitting the matrices $B, A \in \mathbb{C}^{n\times n}$ into

$$B = M_B - N_B \quad \text{and} \quad A = M_A - N_A,$$

we can precisely describe the waveform relaxation method for the DAEs (2.1) as follows:

$$\begin{cases} M_B\dot{x}^{(k)}(t) + M_Ax^{(k)}(t) = N_B\dot{x}^{(k-1)}(t) + N_Ax^{(k-1)}(t) + f(t), \\ x^{(k)}(0) = x_0. \end{cases} \tag{4.1}$$

According to Theorem 2.2, for each k the problem in (4.1) is solvable if the matrix pencil $\lambda M_B + M_A$ is regular. Moreover, x_0 should be chosen as the consistent initial condition or the admissible initial value, which admits the solution be smooth; see, e.g., [13,7].

Miekkala introduced an explicit expression of the waveform relaxation method (4.1) for the DAEs (2.1) in [7]. Since the matrix pencil $\lambda M_B + M_A$ is regular, according to Theorem 2.1 there exist nonsingular matrices $P, Q \in \mathbb{R}^{n\times n}$ such that the scheme (4.1) can be transformed into

$$\begin{cases} \dot{\tilde{x}}_1^{(k)}(t) + J\tilde{x}_1^{(k)}(t) = N_{B_1}\dot{\tilde{x}}^{(k-1)}(t) + N_{A_1}\tilde{x}^{(k-1)}(t) + \widehat{f}_1(t), \\ H\dot{\tilde{x}}_2^{(k)}(t) + \tilde{x}_2^{(k)}(t) = N_{B_2}\dot{\tilde{x}}^{(k-1)}(t) + N_{A_2}\tilde{x}^{(k-1)}(t) + \widehat{f}_2(t), \end{cases} \tag{4.2}$$

with

$$PM_BQ = \begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix}, \qquad PM_AQ = \begin{pmatrix} J & 0 \\ 0 & I \end{pmatrix},$$

$$PN_BQ = \begin{pmatrix} N_{B_1} \\ N_{B_2} \end{pmatrix}, \qquad PN_AQ = \begin{pmatrix} N_{A_1} \\ N_{A_2} \end{pmatrix},$$

$$Q^{-1}x(t) = \tilde{x}(t) = \begin{pmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \end{pmatrix}, \qquad Q^{-1}x_0 = \tilde{x}_0 = \begin{pmatrix} \tilde{x}_{10} \\ \tilde{x}_{20} \end{pmatrix},$$

and

$$\widehat{f}(t) = Pf(t) = \begin{pmatrix} \widehat{f}_1(t) \\ \widehat{f}_2(t) \end{pmatrix},$$

where H is a matrix with nilpotency ν. Then the scheme (4.2) can be rewritten into the operator form

$$\tilde{x}^{(k+1)}(t) = (\mathcal{H}\tilde{x}^{(k)})(t) + \Phi(t), \tag{4.3}$$

where

$$\mathcal{H} = \begin{pmatrix} \mathcal{H}_1 \\ \mathcal{H}_2 \end{pmatrix}, \qquad (\mathcal{H}\tilde{x})(t) = \begin{pmatrix} (\mathcal{H}_1\tilde{x})(t) \\ (\mathcal{H}_2\tilde{x})(t) \end{pmatrix},$$

with

$$\begin{cases} (\mathcal{H}_1\tilde{x})(t) = N_{B_1}\tilde{x}(t) + \displaystyle\int_0^t \mathrm{e}^{(s-t)J}(N_{A_1} - JN_{B_1})\tilde{x}(s)\mathrm{d}s, \\ (\mathcal{H}_2\tilde{x})(t) = \displaystyle\sum_{i=0}^{\nu-1}(-1)^i H^i \left(N_{B_2}\frac{\mathrm{d}}{\mathrm{d}t} + N_{A_2}\right)\frac{\mathrm{d}^i\tilde{x}(t)}{\mathrm{d}t^i}, \end{cases}$$

and

$$\Phi(t) = \begin{pmatrix} \mathrm{e}^{-tJ}\tilde{x}_{10} - \mathrm{e}^{-tJ}N_{B_1}\tilde{x}_0 + \displaystyle\int_0^t \mathrm{e}^{(s-t)J}\widehat{f}_1(s)\mathrm{d}s \\ \displaystyle\sum_{i=0}^{\nu-1}(-1)^i H^i \frac{\mathrm{d}^i\widehat{f}_2(t)}{\mathrm{d}t^i} \end{pmatrix}.$$

This operator form seems very complicated. In order to analyze its convergence property, Miekkala made several assumptions that are difficult to be verified; see [7]. We will review these assumptions in the next section.

Therefore, it is necessary to find a better operator form for the waveform relaxation method (4.1). In fact, we can take advantage of the Laplace transform to obtain a simpler and more useful operator form, which is precisely described in the following theorem.

Theorem 4.1. *Consider the initial-value problem of linear constant-coefficient DAEs (2.1) with the regular matrix pencil $\lambda B + A$. Let*

$$B = M_B - N_B \quad \text{and} \quad A = M_A - N_A$$

be splittings of the matrices $B, A \in \mathbb{R}^{n\times n}$ such that $\lambda M_B + M_A$ is a regular matrix pencil. Assume that x_0 is a consistent initial condition, and both $x^{(k-1)}(t)$ and $f(t)$ belong to $\mathbb{T}^n_\sigma$, with σ a real number. Then the waveform relaxation method (4.1) can be rewritten into the operator form

$$x^{(k)}(t) = \mathcal{K}(x^{(k-1)}(t)) + \Phi(f(t)), \tag{4.4}$$

where

$$\begin{cases} \mathcal{K}(x(t)) = (\mathcal{L}^{-1}(sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L})(x(t)), \\ \Phi(f(t)) = (\mathcal{L}^{-1}(sM_B + M_A)^{-1}\mathcal{L})(f(t)). \end{cases}$$

Moreover, if $\mathrm{Re}(s) \le \sigma$ holds for $\forall s \in \mathrm{sp}(M_A, -M_B) := \{s \in \mathbb{C}^1 : \det(M_A + sM_B) = 0\}$, then $x^{(k)}(t)$ also belongs to $\mathbb{T}^n_\sigma$, which means that the iteration (4.1) is closed in $\mathbb{T}^n_\sigma$.

Proof. Without loss of generality, we simply assume that $x_0 = 0$. Because $x^{(k-1)}(t), f(t) \in \mathbb{T}^n_\sigma$, according to Lemma 3.2 we know that the solution $x^{(k)}(t)$ of the problem (4.1) is of exponential order. Hence, we can apply the Laplace transform to both sides of (4.1), i.e.,

$$\mathcal{L}(M_B\dot{x}^{(k)}(t) + M_A x^{(k)}(t)) = \mathcal{L}(N_B\dot{x}^{(k-1)}(t) + N_A x^{(k-1)}(t) + f(t)),$$

obtaining the equality

$$sM_B\mathcal{L}(x^{(k)}(t)) + M_A\mathcal{L}(x^{(k)}(t)) = sN_B\mathcal{L}(x^{(k-1)}(t)) + N_A\mathcal{L}(x^{(k-1)}(t)) + \mathcal{L}(f(t))$$

due to the linearity of $\mathcal{L}$, or equivalently,

$$\mathcal{L}(x^{(k)}(t)) = (sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}(x^{(k-1)}(t)) + (sM_B + M_A)^{-1}\mathcal{L}(f(t)). \tag{4.5}$$

Assume that the inverse Laplace transform $\mathcal{L}^{-1}$ can be applied to

$$(sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}(x^{(k-1)}(t)) \quad \text{and} \quad (sM_B + M_A)^{-1}\mathcal{L}(f(t)).$$

As Theorem A.3 shows that $\mathcal{L}^{-1}$ is one-to-one, we can act $\mathcal{L}^{-1}$ on both sides of (4.5), i.e.,

$$\begin{aligned} x^{(k)}(t) &= \left(\mathcal{L}^{-1}(sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}\right)(x^{(k-1)}(t)) + \left(\mathcal{L}^{-1}(sM_B + M_A)^{-1}\mathcal{L}\right)(f(t)) \\ &= \mathcal{K}(x^{(k-1)}(t)) + \Phi(f(t)), \end{aligned}$$

obtaining the expression (4.4) due to the linearity of $\mathcal{L}^{-1}$.

Now we turn to prove the existence of $\mathcal{L}^{-1}$ with respect to

$$(sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}(x^{(k-1)}(t)) \quad \text{and} \quad (sM_B + M_A)^{-1}\mathcal{L}(f(t)).$$

To this end, we let P and Q be real matrices transforming the matrix pencil $sM_B + M_A$ into the Kronecker canonical form

$$P(sM_B + M_A)Q = s\begin{pmatrix} I & 0 \\ 0 & N \end{pmatrix} + \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix},$$

or

$$(sM_B + M_A)^{-1} = Q\begin{pmatrix} (sI + C)^{-1} & 0 \\ 0 & (sN + I)^{-1} \end{pmatrix} P,$$

where

$$N = \begin{pmatrix} N_1 & & \\ & \ddots & \\ & & N_\ell \end{pmatrix}, \qquad N_j = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{R}^{n_j \times n_j},$$

and n_j $(j = 1, 2, \ldots, \ell)$ are positive integers such that $\sum_{j=1}^{\ell} n_j = n$. Evidently, the block matrices C and N correspond to the nonzero and the zero generalized eigenvalues in $\mathrm{sp}(M_A, -M_B)$, respectively. Here, for $\ell = 0$, we have stipulated that the matrix block N does not exist or the matrix block C is of size $n \times n$.

The inverse of any nonsingular matrix $G = (g_{i,j}) \in \mathbb{R}^{n\times n}$ can be expressed as

$$G^{-1} = \frac{1}{\det(G)}\begin{pmatrix} G_{1,1} & G_{2,1} & \cdots & G_{n,1} \\ G_{1,2} & G_{2,2} & \cdots & G_{n,2} \\ \cdots & \cdots & \cdots & \cdots \\ G_{1,n} & G_{2,n} & \cdots & G_{n,n} \end{pmatrix},$$

where $G_{i,j}$ denotes the algebraic complement of the element $g_{i,j}$, $i, j = 1, 2, \ldots, n$. Based on this fact, we know that the elements of $(sI + C)^{-1}$ are of the form $\frac{p(s)}{q(s)}$, with $p(s)$ and $q(s)$ being real-coefficient polynomials with respect to s and satisfying:

(i) $\deg(q(s)) - \deg(p(s)) \geq 1$, if the elements of $(sI + C)^{-1}$ are on the diagonal;
(ii) $\deg(q(s)) - \deg(p(s)) \geq 2$, if the elements of $(sI + C)^{-1}$ are on the off-diagonal.

On the other hand, the elements of $(sN_j + I)^{-1}$, the jth diagonal block of $(sN + I)^{-1}$, are of the form

$$(sN_j + I)^{-1} = \sum_{k=0}^{n_j-1}(-1)^k s^k N_j^k = \begin{pmatrix} 1 & -s & \cdots & \cdots & (-s)^{(n_j-1)} \\ & 1 & -s & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & -s \\ & & & & 1 \end{pmatrix}.$$

It then follows that the elements of $(sM_B + M_A)^{-1}(sN_B + N_A)$ and $(sM_B + M_A)^{-1}$ are of the form

$$u(s) := r(s) + \frac{p(s)}{q(s)}, \quad \text{with } \deg(r(s)) \geq 0 \text{ and } \deg(q(s)) - \deg(p(s)) \geq 1,$$

where $r(s)$, $p(s)$ and $q(s)$ are real-coefficient polynomials with respect to s. Hence, to prove the existence of $\mathcal{L}^{-1}$ for $(sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}(x^{(k-1)}(t))$ and $(sM_B + M_A)^{-1}\mathcal{L}(f(t))$, we only need to demonstrate the existence of $\mathcal{L}^{-1}$ for $u(s)\mathcal{L}(x)$, where $x := x(t)$ denotes a scalar-valued causal function in $\mathbb{T}_\sigma$. Because $u(s) = r(s) + \frac{p(s)}{q(s)}$, it is sufficient to discuss the existence of $\mathcal{L}^{-1}(r(s)\mathcal{L}(x))$ and $\mathcal{L}^{-1}(\frac{p(s)}{q(s)}\mathcal{L}(x))$.

For $x \in \mathbb{T}_\sigma$, by applying Theorem A.1 to $s^n\mathcal{L}(x)$ we immediately obtain

$$s^n\mathcal{L}(x) = \mathcal{L}\left(\frac{\mathrm{d}^n x}{\mathrm{d}t^n}\right), \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \sigma.$$

It then follows from Theorem A.3 that

$$\mathcal{L}^{-1}\left(r(s)\mathcal{L}(x)\right) = \mathcal{L}^{-1}\mathcal{L}\left(r\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)(x)\right) = r\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)(x)$$

holds for $\forall s \in \mathbb{C}^1$ satisfying $\mathrm{Re}(s) > \sigma$. As $x \in \mathbb{T}_\sigma$, $r(\frac{\mathrm{d}}{\mathrm{d}t})(x)$ belongs to $\mathbb{T}_\sigma$, too. Therefore, $\mathcal{L}^{-1}\left(r(s)\mathcal{L}(x)\right)$ exists and belongs to $\mathbb{T}_\sigma$.

In addition, by applying the partial fractional expansion technique we can rewrite the rational function $\frac{p(s)}{q(s)}$, with $\deg(q(s)) - \deg(p(s)) \geq 1$, as the sum of terms like

$$\frac{\xi}{(s+a)^k} \quad \text{and} \quad \frac{\eta s + \mu}{((s+b)^2 + c^2)^l},$$

where k, l are nonnegative integers, and ξ, η, μ, a, b and c are real numbers. Because

$$\left(\mathcal{L}(t^{k-1}\mathrm{e}^{-at})\right)(s) = \frac{(k-1)!}{(s+a)^k}, \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > -a,$$

$$\left(\mathcal{L}(\mathrm{e}^{-bt}\sin(ct))\right)(s) = \frac{c}{(s+b)^2 + c^2}, \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > -b, \quad \text{and}$$

$$\left(\mathcal{L}(\mathrm{e}^{-bt}\cos(ct))\right)(s) = \frac{s+b}{(s+b)^2 + c^2}, \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > -b,$$

by repeatedly applying Theorem A.2 to the terms like $\frac{\eta s+\mu}{((s+b)^2+c^2)^l}$, with l an integer not less than 2, based on Theorem A.3 we can easily determine

$$h(t) := \mathcal{L}^{-1}\left(\frac{p(s)}{q(s)}\right), \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \widetilde{\sigma},$$

where $\widetilde{\sigma}$ is the least upper bound of $\mathrm{Re}(s)$, $\forall s \in \mathrm{sp}(M_A, -M_B)$. Define $\bar{\sigma} = \max\{\sigma, \widetilde{\sigma}\}$. Then both $\mathcal{L}(h(t))$ and $\mathcal{L}(x(t))$ exist as absolutely convergent integrals in the half plane $\mathrm{Re}(s) > \bar{\sigma}$. It then follows from Theorem A.4 that

$$\mathcal{L}(h * x)(t) = \frac{p(s)}{q(s)}\mathcal{L}(x), \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \bar{\sigma}.$$

Thereby, by making use of Theorem A.3 we have

$$\mathcal{L}^{-1}\left(\frac{p(s)}{q(s)}\mathcal{L}(x)\right) = (h * x)(t), \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \bar{\sigma}.$$

In summary, we have obtained the equality

$$\begin{aligned} H(t) := {} & \mathcal{L}^{-1}\left((sM_B + M_A)^{-1}(sN_B + N_A)\mathcal{L}(x^{(k-1)}(t))\right) \\ & + \mathcal{L}^{-1}\left((sM_B + M_A)^{-1}\mathcal{L}(f(t))\right), \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \bar{\sigma}. \end{aligned}$$

Moreover, if $\mathrm{Re}(s) \leq \sigma$, $\forall s \in \mathrm{sp}(M_A, -M_B)$, we have $\bar{\sigma} = \sigma$.

Hence, we have demonstrated that $x^{(k)}(t) = \mathcal{K}(x^{(k-1)}(t)) + \Phi(f(t))$ belongs to $\mathbb{T}_\sigma^n$, which shows that the waveform relaxation iteration (4.1) is closed in $\mathbb{T}_\sigma^n$. □

Remark 4.1. $\mathcal{K}$ is a linear operator due to the linearity of the Laplace transform $\mathcal{L}$. For the initial-value problem of linear constant-coefficient DAEs (2.1) satisfying $f(t) \in \mathbb{T}_\beta^n$, with β the abscissa of $f(t)$, we know that $x(t)$ belongs to $\mathbb{T}_\sigma^n$, where $\sigma = \max\{\alpha, \beta\}$ and α is the least upper bound of $\mathrm{Re}(s)$ such that $s \in \mathrm{sp}(A, -B)$. In fact, from Lemma 3.2 we know that the solution $x(t)$ of the DAEs (2.1) is of exponential order. Thus we can apply the Laplace transform $\mathcal{L}$ to (2.1), obtaining

$$\mathcal{L}(x(t)) = (sB + A)^{-1}\mathcal{L}(f(t)) - (sB + A)^{-1}Bx_0$$

in a similar fashion to Theorem 4.1. By denoting

$$\mathcal{L}(H(t)) = (sB + A)^{-1}, \quad \text{for } s \in \mathbb{C}^1 \text{ with } \mathrm{Re}(s) > \alpha,$$

from Theorem A.4 and Lemma 3.1 we have the estimate

$$|(x(t))_i| \leq æ\mathrm{e}^{\max\{\alpha,\beta\}t}, \quad i = 1, 2, \ldots, n,$$

where æ is a positive constant and $(x(t))_i$ represents the ith element of $x(t)$. It is deserved to mention that α is also the least upper bound of the abscissas of elements in $H(t)$. Therefore, the solution $x(t)$ of the DAEs (2.1) can be approximated in the function space $\mathbb{T}^n_\sigma$.

5. Convergence theory

From Theorem 4.1 we see that the convergence property of the waveform relaxation method (4.1) closely depends on the linear operator $\mathcal{K}$. In this section, we will discuss the asymptotic convergence rate of the waveform relaxation method (4.1).

To this end, let $x_\star(t)$ be the solution of the DAEs (2.1) and define

$$\varepsilon^{(k)}(t) = x^{(k)}(t) - x_\star(t).$$

Then by Theorem 4.1 we have

$$\begin{aligned}\varepsilon^{(k)}(t) &= \left(\mathcal{K}(x^{(k-1)}(t)) + \Phi(f(t))\right) - (\mathcal{K}(x_\star(t)) + \Phi(f(t)))\\ &= \mathcal{K}(x^{(k-1)}(t)) - \mathcal{K}(x_\star(t))\\ &= \mathcal{K}(x^{(k-1)}(t) - x_\star(t))\\ &= \mathcal{K}(\varepsilon^{(k-1)}(t))\\ &= \mathcal{K}^{k-1}(\varepsilon^{(1)}(t)).\end{aligned}$$

Hence, to analyze the convergence of the functional sequence $\{\varepsilon^{(k)}(t)\}_{k=0}^{\infty}$, we need to determine the spectral radius of the operator $\mathcal{K}$, which is denoted by $\rho(\mathcal{K})$.

In [7], Miekkala determined the spectral radius of the operator $\mathcal{H}$ in (4.3) under the following assumptions:

(a) when $\mathrm{Re}(\lambda) \geq 0$, it holds that $\det(\lambda B + A) \neq 0$;
(b) when $\mathrm{Re}(\lambda) \geq 0$, it holds that $\det(\lambda M_B + M_A) \neq 0$;
(c) if $\nu \geq 1$, then $N_{B_2} = 0$;
(d) if $\nu \geq 2$, then $N_{B_2} = 0$ and $N_{A_2} = 0$;
(e) $f(t)$ in the DAEs (2.1) is bounded.

Here N_{B_2} and N_{A_2} are matrices in (4.2), and ν is the nilpotency of the matrix H in (4.2). The result can be described in the following theorem.

Theorem 5.1 ([7]). *If the above assumptions* (a)–(e) *are satisfied, then the spectral radius of the operator $\mathcal{H}$ in (4.3) is given by*

$$\rho(\mathcal{H}) = \sup_{\mathrm{Re}(s)=0} \rho(\mathbf{K}(s)),$$

where $\mathbf{K}(s) = (sM_B + M_A)^{-1}(sN_B + N_A)$ is the dynamic iteration matrix.

This result has been well known and widely used in the literature; see, e.g., [8,10–12].

In fact, two of the above assumptions are not necessary when we study the convergence property of the waveform relaxation method (4.1) based on the new operator $\mathcal{K}$ in (4.4), and the other three assumptions can be much weakened. Specifically, we only need to impose the following conditions:

(a) there exists $\lambda \in \mathbb{C}^1$ such that $\det(\lambda B + A) \neq 0$;
(b) there exists $\lambda \in \mathbb{C}^1$ such that $\det(\lambda M_B + M_A) \neq 0$;
(e) $f(t)$ in the DAEs (2.1) is of exponential order.

Under the above weaker assumptions, we can obtain a more general theorem.

Theorem 5.2. *Consider the initial-value problem of linear constant-coefficient DAEs (2.1) with the regular matrix pencil $\lambda B + A$. Let*

$$B = M_B - N_B \quad \text{and} \quad A = M_A - N_A$$

be splittings of the matrices $B, A \in \mathbb{R}^{n\times n}$ such that $\lambda M_B + M_A$ is a regular matrix pencil. Assume that x_0 is a consistent initial condition, and $x^{(k-1)}(t)$ and $f(t)$ belong to $\mathbb{T}^n_\beta$, with β a real number. Define $\sigma = \max\{\alpha, \beta\}$, with α the least upper bound of $\mathrm{Re}(s)$ such that $s \in \mathrm{sp}(A, -B)$. Then, for $\forall s \in \mathrm{sp}(M_A, -M_B)$ with $\mathrm{Re}(s) \leq \sigma$, the spectral radius of the iteration operator $\mathcal{K}$ of the waveform relaxation method (4.1) is given by

$$\rho(\mathcal{K}) = \sup_{\mathrm{Re}(s)=\sigma} \rho(\mathbf{K}(s)), \tag{5.1}$$

where $\mathbf{K}(s) = (sM_B + M_A)^{-1}(sN_B + N_A)$ is the dynamic iteration matrix.

Proof. Denote by $\mathrm{sp}(\mathcal{K})$ the spectral set of the operator $\mathcal{K}$ defined on the function space $\mathbb{T}_\sigma^n$. Then we have

$$\mathrm{sp}(\mathcal{K}) = \mathrm{sp}_p(\mathcal{K}) \cup \mathrm{sp}_r(\mathcal{K}) \cup \mathrm{sp}_c(\mathcal{K}),$$

where $\mathrm{sp}_p(\mathcal{K})$ is the point spectrum, $\mathrm{sp}_r(\mathcal{K})$ is the residual spectrum, and $\mathrm{sp}_c(\mathcal{K})$ is the continuous spectrum.

Now, we are going to prove $\mathrm{sp}_r(\mathcal{K}) = \mathrm{sp}_c(\mathcal{K}) = \emptyset$, which immediately implies $\mathbb{C}^1 \setminus \mathrm{sp}(\mathcal{K}) = \mathbb{C}^1 \setminus \mathrm{sp}_p(\mathcal{K})$. We remark that complex numbers in the complementary set $\mathbb{C}^1 \setminus \mathrm{sp}(\mathcal{K})$ are called regular numbers of the operator $\mathcal{K}$ defined on the function space $\mathbb{T}_\sigma^n$.

In fact, for $\lambda \in \mathbb{C}^1 \setminus \mathrm{sp}_p(\mathcal{K})$, the operator $\lambda I - \mathcal{K}$ is invertible. By direct computations we have

$$(\lambda I - \mathcal{K})(x) = \mathcal{L}^{-1}\left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)\mathcal{L}(x), \quad \forall x := x(t) \in \mathbb{T}_\sigma^n.$$

For $\forall y := y(t) \in \mathbb{T}_\sigma^n$, the solution $x \in \mathbb{T}_\sigma^n$ of the operator equation

$$(\lambda I - \mathcal{K})(x) = y,$$

or equivalently,

$$\mathcal{L}^{-1}\left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)\mathcal{L}(x) = y,$$

can be computed in the following way. First, by applying the Laplace transform operator $\mathcal{L}$ on both sides we get

$$\left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)\mathcal{L}(x) = \mathcal{L}(y).$$

As $\lambda I - \mathcal{K}$ is invertible, for $\mathrm{Re}(s) \geq \sigma$,

$$\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A) = (sM_B + M_A)^{-1}\left[s(\lambda M_B - N_B) + (\lambda M_A - N_A)\right]$$

is invertible, too. This readily implies

$$\mathrm{Re}(s) < \sigma, \quad \forall s \in \mathrm{sp}(\lambda M_A - N_A, -(\lambda M_B - N_B)).$$

Hence, it holds that

$$\mathcal{L}(x) = \left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)^{-1}\mathcal{L}(y).$$

Now, similar to the proof of Theorem 4.1, we apply $\mathcal{L}^{-1}$ to both sides of the above equality, obtaining

$$\begin{aligned} x &= \mathcal{L}^{-1}\left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)^{-1}\mathcal{L}(y) \\ &= \mathcal{L}^{-1}\left\{\left[s(\lambda M_B - N_B) + (\lambda M_A - N_A)\right]^{-1}(sM_B + M_A)\right\}\mathcal{L}(y). \end{aligned}$$

It follows immediately that $x \in \mathbb{T}_\sigma^n$ and $\mathrm{range}(\lambda I - \mathcal{K}) = \mathbb{T}_\sigma^n, \forall \lambda \in \mathbb{C}^1 \setminus \mathrm{sp}_p(\mathcal{K})$, where $\mathrm{range}(\lambda I - \mathcal{K})$ represents the range space of $\lambda I - \mathcal{K}$. This fact clearly shows that $\mathrm{sp}_r(\mathcal{K}) = \mathrm{sp}_c(\mathcal{K}) = \emptyset$, or in other words, $\mathbb{C}^1 \setminus \mathrm{sp}(\mathcal{K}) = \mathbb{C}^1 \setminus \mathrm{sp}_p(\mathcal{K})$ or $\mathrm{sp}(\mathcal{K}) = \mathrm{sp}_p(\mathcal{K})$. Furthermore, we have

$$\begin{aligned} \rho(\mathcal{K}) &= \sup\{|\lambda| : \lambda \in \mathrm{sp}(\mathcal{K})\} \\ &= \sup\{|\lambda| : \lambda \in \mathrm{sp}_p(\mathcal{K})\} \\ &= \sup\{|\lambda| : (\lambda I - \mathcal{K})(x) = 0, \exists\, x \in \mathbb{T}_\sigma^n \setminus \{0\}\} \\ &= \inf\{\varrho : (\lambda I - \mathcal{K})(x) \neq 0, 0 \neq x \in \mathbb{T}_\sigma^n, |\lambda| > \varrho\} \\ &= \inf\{\varrho : \mathcal{L}^{-1}\left(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)\right)\mathcal{L}(x) \neq 0,\ 0 \neq x \in \mathbb{T}_\sigma^n, |\lambda| > \varrho\}. \end{aligned}$$

Because $\mathrm{Re}(s) \leq \sigma, \forall s \in \mathrm{sp}(M_A, -M_B)$, according to Theorem 4.1 we know that there exists a matrix-valued function $H(t)$ such that

$$\mathcal{L}(H(t)) = (sM_B + M_A)^{-1}(sN_B + N_A)$$

exists for any s, satisfying $\mathrm{Re}(s) > \sigma$, as absolutely convergent integrals in the componentwise fashion. Therefore,

$$\begin{aligned} \rho(\mathcal{K}) &= \inf\{\varrho : \det(\lambda I - (sM_B + M_A)^{-1}(sN_B + N_A)) \neq 0, \mathrm{Re}(s) > \sigma, |\lambda| > \varrho\} \\ &= \sup_{\mathrm{Re}(s)>\sigma} \rho((sM_B + M_A)^{-1}(sN_B + N_A)) \\ &= \sup_{\mathrm{Re}(s)=\sigma} \rho((sM_B + M_A)^{-1}(sN_B + N_A)). \end{aligned}$$

This completes the proof. □

Remark 5.1. Theorem 5.1 is just a special case of Theorem 5.2 when we consider the bounded case, i.e., $\sigma = 0$. In general, the waveform relaxation method (4.1) is convergent under the condition

$$\rho(\mathcal{K}) := \sup_{\mathrm{Re}(s)=\sigma} \rho((sM_B + M_A)^{-1}(sN_B + N_A)) < 1,$$

with $\rho(\mathcal{K})$ being the corresponding asymptotic convergence factor. In fact, it is difficult to understand the convergence of the waveform relaxation method for the linear DAEs (2.1) when it displays infinite time blow up, but it will truly happen. According to Lemma 3.2, the solution of the linear DAEs (2.1) is a function of exponential order, which can be simply considered as the product of a bounded function and an exponential function. Therefore, the waveform relaxation iteration will make the exponential part of the iterate remain the exponential order and the bounded part converge to a certain bounded function as long as the convergence factor $\rho(\mathcal{K})$ is less than one.

6. Numerical results

Consider the initial-value problem

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} (\sin(t) + 2.02)\mathrm{e}^{0.02t} \\ (2\sin(t) + 1)\mathrm{e}^{0.02t} \end{pmatrix}, \tag{6.1}$$

with the initial conditions $x_1(0) = 1$ and $x_2(0) = 0$. The exact solution of (6.1) is

$$x_\star(t) = \begin{pmatrix} \mathrm{e}^{0.02t} \\ \sin(t)\mathrm{e}^{0.02t} \end{pmatrix}.$$

We solve this problem by the *SOR waveform relaxation* (**SORWR**) method, for which the splittings are chosen as

$$M_B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad N_B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$M_A = \frac{1}{\omega}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \qquad N_A = \frac{1-\omega}{\omega}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},$$

where ω is a relaxation parameter. The corresponding dynamic iteration matrix is given by

$$\mathbf{K}^{\mathrm{SOR}}(s) = \begin{pmatrix} s + \dfrac{2}{\omega} & 0 \\ 1 & \dfrac{2}{\omega} \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{2(1-\omega)}{\omega} & -1 \\ 0 & \dfrac{2(1-\omega)}{\omega} \end{pmatrix},$$

where $s = \sigma + \imath\xi$, with $\xi \in \mathbb{R}^1$ and $\imath$ the imaginary unit. According to Remark 4.1, we have $\sigma = 0.02$.

With respect to different ω, Fig. 1 shows the spectral radius of the dynamic iteration matrix of the SORWR method evaluated as a function of the frequency ξ. We see that the spectral radius always keeps less than one, with the maximal spectral radius appearing around $\xi = 0$. When ξ goes away from 0, the spectral radius keeps small, which implies that the SORWR method for the initial-value problem (6.1) is convergent in an infinite time interval.

In actual computations, to perform an infinite time simulation on a computer is often infeasible. Instead we take a long time simulation to illustrate the behavior of the waveform relaxation method. So, we set $\omega = 1$ and choose the time interval to be $\Omega_t = [0, \Delta t \times \ell_t]$, where $\Delta t = 0.01$ is the time stepsize and ℓ_t is the number of time steps. In Figs. 2 and 3 we plot the error $\epsilon_2^{(k)}(t)$ of the second component of the iterate after $k = 0, 1, 2$ and 3 iteration steps for $\ell_t = 5000$ and 80 000, respectively. In both figures, we see that the errors decrease rapidly along the whole time interval Ω_t during the iteration process. Moreover, in Fig. 3 we find that each iteration lengthens the convergence time interval for the larger time step case $\ell_t = 80\,000$, and the upper bound of the error decreases monotonically after every iteration. We remark that the situation for the error $\epsilon_1^{(k)}(t)$ of the first component of the iterate is very similar. Hence, we can expect that the waveform relaxation method is convergent in an infinite time interval.

In Table 1 we show the numbers of iteration steps of the SORWR method for different ω and time intervals. Here, the stopping criterion is set to be

$$\max\left\{\sup_{t\in\Omega_t} |\epsilon_1^{(k)}(t)|, \sup_{t\in\Omega_t} |\epsilon_2^{(k)}(t)|\right\} < 10^{-6}.$$

Evidently, the SORWR method is convergent for all tested cases. We see that the number of iteration steps of the SORWR method increases only one time, while the length of time interval increases 15 times for different ω, i.e., the length increases from 50 to 800. Furthermore, choosing ω around one will lead to less number of iteration steps.

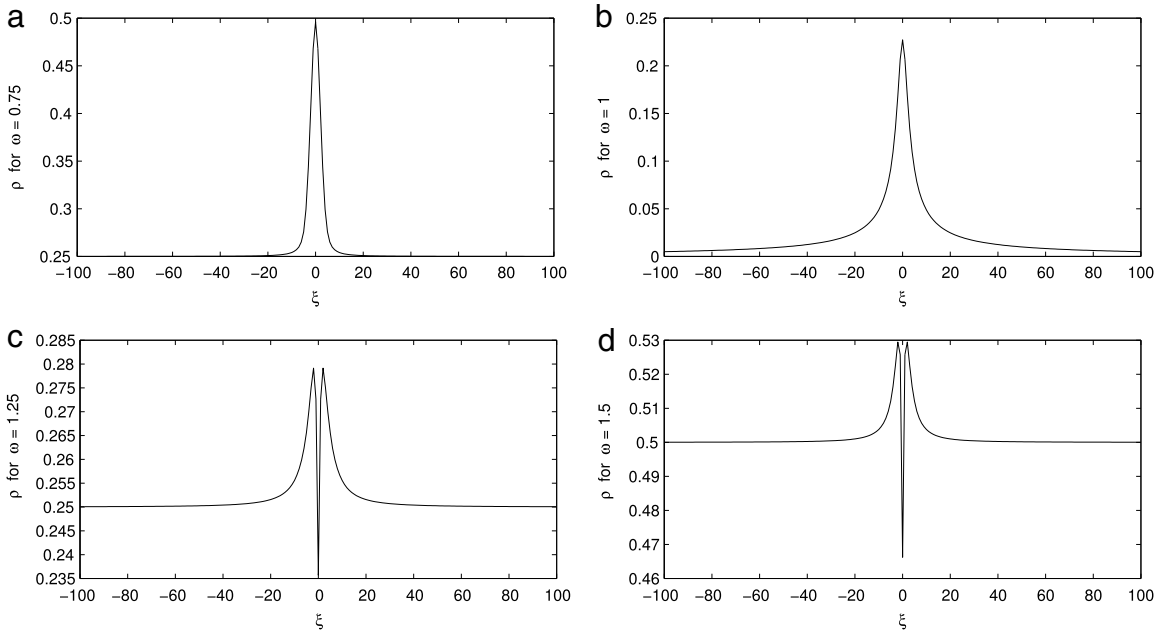

Fig. 1. Spectral radius of the dynamic iteration matrix of the SORWR method evaluated as a function of the frequency ξ: (a) $\omega = 0.75$; (b) $\omega = 1$; (c) $\omega = 1.25$; (d) $\omega = 1.5$.

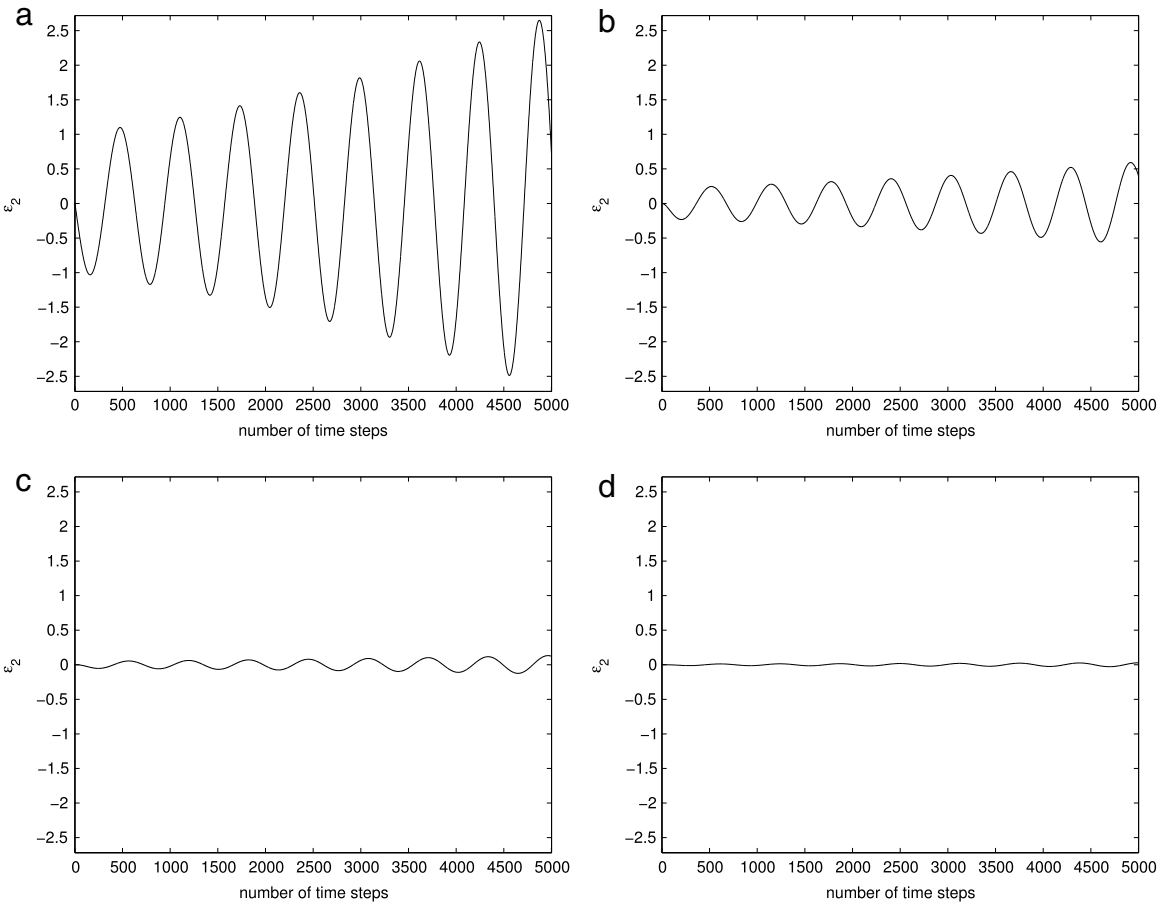

Fig. 2. The error $\epsilon_2^{(k)}(t)$ after the kth iterate for 5000 time steps and $\omega = 1$: (a) $k = 0$; (b) $k = 1$; (c) $k = 2$; (d) $k = 3$.

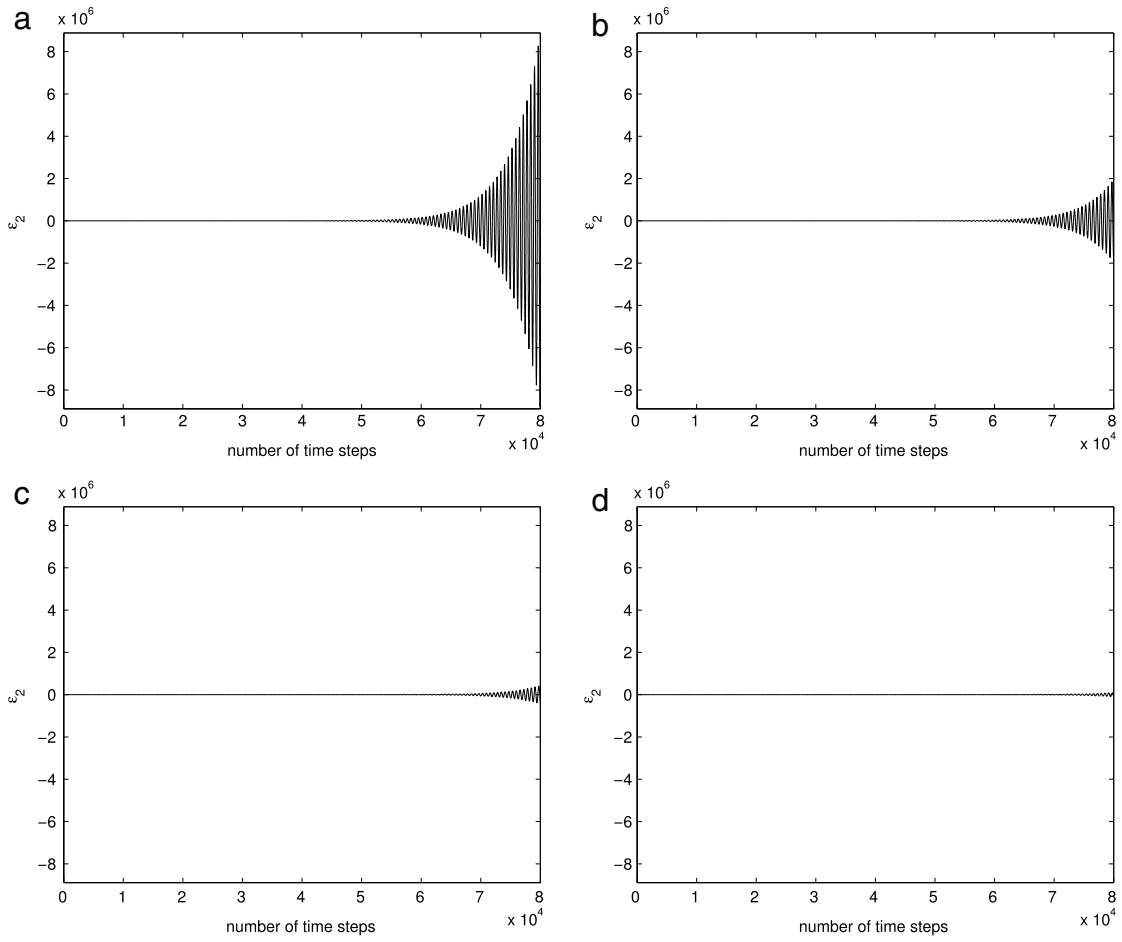

Fig. 3. The error $\epsilon_2^{(k)}(t)$ after the kth iterate for 80 000 time steps and $\omega = 1$: (a) $k = 0$; (b) $k = 1$; (c) $k = 2$; (d) $k = 3$.

Table 1
Numbers of iterations of the SORWR method for different ω and time intervals.

ℓ_t	5000	10 000	20 000	40 000	80 000
Ω_t	[0, 50]	[0, 100]	[0, 200]	[0, 400]	[0, 800]
$\omega = 0.75$	22	23	26	32	44
$\omega = 1.00$	11	11	13	15	21
$\omega = 1.25$	12	13	15	18	24
$\omega = 1.50$	25	27	30	37	50

7. Concluding remarks

We have studied the waveform relaxation methods for solving the linear constant-coefficient differential-algebraic equations by applying the Laplace transform, obtaining explicit expression and asymptotic convergence rate for this class of iteration schemes without demanding the boundedness of the solutions. This theory extends the existing one to problems of wider and more useful backgrounds, and provides a fundamental and powerful tool for analyzing the convergence of the waveform relaxation methods for solving the initial-value problems of linear constant-coefficient differential-algebraic equations.

Acknowledgements

The authors are very much indebted to Profs. Feng-Shan Bai, Xing-Ping Liu and Shu-Fang Xu for valuable discussions and constructive suggestions. The referees provided very useful comments and suggestions, which greatly improved the original manuscript of this paper.

Appendix. Basic properties about Laplace transform

Theorem A.1 (*Differentiation in Time Domain*). *Let $f : \mathbb{R}^1 \to \mathbb{C}^1$ be a causal function which is, in addition, differentiable in $\mathbb{R}^1$. In a half plane where $\mathcal{L}f$ and $\mathcal{L}f'$ both exist one has*

$$(\mathcal{L}f')(s) = s(\mathcal{L}f)(s) - f(0_+) = s(\mathcal{L}f)(s),$$

where $f' := f'(t)$ denotes the first-order derivative of $f(t)$ with respect to t.

Theorem A.2 (*Differentiation in Frequency Domain*). *Let $f : \mathbb{R}^1 \to \mathbb{C}^1$ be a function with the Laplace transform $F(s) = \mathcal{L}f(s)$, and let σ_a be the abscissa of absolute convergence. Then $F(s)$ is an analytic function of s for $\mathrm{Re}(s) > \sigma_a$, and*

$$\frac{\mathrm{d}}{\mathrm{d}s}F(s) = -(\mathcal{L}(tf(t)))(s).$$

Theorem A.3 (*The Laplace Transform is One-to-One*). *Let $f(t)$ and $g(t)$ be two causal functions of exponential orders, and $F(s) = (\mathcal{L}f)(s)$ and $G(s) = (\mathcal{L}g)(s)$ be the Laplace transforms of $f(t)$ and $g(t)$, respectively. When $F(s) = G(s)$ holds in a half plane $\mathrm{Re}(s) > \sigma$, $f(t) = g(t)$ is satisfied at all points where both f and g are continuous. Here, σ is a given real number.*

Theorem A.4 (*Convolution Theorem for $\mathcal{L}$*). *Let $f, g : \mathbb{R}^1 \to \mathbb{C}^1$ be two causal functions, and the Laplace transforms $F = \mathcal{L}f$ and $G = \mathcal{L}g$ exist as absolutely convergent integrals in a half plane $\mathrm{Re}(s) > \sigma$, with σ a given real number. Then $\mathcal{L}(f * g)$ exists for $\forall s \in \mathbb{C}^1$ such that $\mathrm{Re}(s) > \sigma$, and satisfies*

$$\mathcal{L}(f * g)(s) = F(s)G(s) = (\mathcal{L}f)(s) \cdot (\mathcal{L}g)(s).$$

References

[1] S.L. Campbell, Singular Systems of Differential Equations, Pitman Advanced Publishing Program, London, Melbourne, 1980.
[2] F.L. Lewis, A survey of linear singular systems, Circuits Systems Signal Processing 5 (1986) 3–36.
[3] J. Wang, Z.-Z. Bai, Convergence analysis of two-stage waveform relaxation method for the initial value problems, Appl. Math. Comput. 172 (2006) 797–808.
[4] Y.-L. Jiang, Waveform Relaxation Methods, Science Press, Beijing, 2009.
[5] E. Lelarasmee, A. Ruheli, A.L. Sangiovanni-Vincentelli, The waveform relaxation method for time-domain analysis of large scale integrated circuits, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. CAD-1 (1982) 131–145.
[6] E. Lelarasmee, A. Ruheli, A.L. Sangiovanni-Vincentelli, Relaxation-based electrical simulation, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. CAD-3 (1984) 308–331.
[7] U. Miekkala, Dynamic iteration methods applied to linear DAE systems, J. Comput. Appl. Math. 25 (1989) 133–151.
[8] U. Miekkala, O. Nevanlinna, Convergence of dynamic iteration methods for initial value problem, SIAM J. Sci. Stat. Comput. 8 (1987) 459–482.
[9] Y.-L. Jiang, O. Wing, A note on the spectra and pseudospectra of waveform relaxation operators for linear differential–algebraic equations, SIAM J. Numer. Anal. 38 (2000) 186–201.
[10] J.-Y. Pan, Z.-Z. Bai, On the convergence of waveform relaxation methods for linear initial value problems, J. Comput. Math. 22 (2004) 681–698.
[11] Z.-Z. Bai, M.K. Ng, J.-Y. Pan, Alternating splitting waveform relaxation method and its successive overrelaxation acceleration, Comput. Math. Appl. 49 (2005) 157–170.
[12] J.-Y. Pan, Z.-Z. Bai, M.K. Ng, Two-step waveform relaxation methods for implicit linear initial value problems, Numer. Linear Algebra Appl. 12 (2005) 293–304.
[13] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in Differential–Algebraic Equations, Elsevier B.V., North-Holland, New York, London, 1989.
[14] R.F. Sincovec, B. Dembart, M.A. Epton, A.M. Erisman, S.W. Manke, E.L. Yip, Solvability of Large Scale Descriptor Systems, Boeing Computer Services Company, Seattle, WA, 1979.
[15] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems, Springer-Verlag, Berlin, Heidelberg, 1996.
[16] R.J. Beerends, H.G. ter Morsche, J.C. van den Berg, E.M. van de Vrie, Fourier and Laplace Transforms, Cambridge University Press, Cambridge, 2003.
[17] R.A. Adams, J.J.F. Fournier, Sobolev Spaces, second ed., Academic Press, London, New York, 2003.
[18] L.C. Evans, Partial Differential Equations, American Mathematical Society, Providence, Rhode Island, 2002.