Logarithmic Norms and Nonlinear DAE Stability

Inmaculada Higueras[†]
Departamento de Matemática e Informática
Universidad Pública de Navarra
31006 Pamplona, Spain.
e-mail: higueras@unavarra.es

Gustaf Söderlind[‡]
Numerical Analysis
Centre for Mathematical Sciences
Lund University, Box 118
SE-221 00 Lund, Sweden.
e-mail: Gustaf.Soderlind@na.lu.se

October 9, 2003

Abstract.

Logarithmic norms are often used to estimate stability and perturbation bounds in linear ODEs. Extensions to other classes of problems such as nonlinear dynamics, DAEs and PDEs require careful modifications of the logarithmic norm. With a conceptual focus, we combine the extension to nonlinear ODEs [15] with that of matrix pencils [10] in order to treat nonlinear DAEs with a view to cover certain unbounded operators, i.e. partial differential algebraic equations. Perturbation bounds are obtained from differential inequalities for any given norm by using the relation between Dini derivatives and semi-inner products. Simple discretizations are also considered.

AMS subject classifications: 65L05.

Key words: Logarithmic norm, logarithmic Lipschitz constant, monotonicity, nonlinear stability, error bounds, differential-algebraic equations, differential inequalities, difference inequalities.

1 Introduction

Logarithmic norms play an important role in the estimation of solutions of ODEs. For a linear ODE

$$x' = Ax + p(t), \tag{1.1}$$

with $A \in \mathbb{R}^{d\times d}$, one obtains the differential inequality

$$\mathrm{D}_t^+ \|x\| \le \mu[A] \cdot \|x\| + \|p(t)\|\,, \tag{1.2}$$

[†]Supported by a grant from the “Secretaría de Estado de Universidades, Investigación y Desarrollo”, during a stay at the Center for Mathematical Sciences of Lund University.

[‡]Supported by Swedish Research Council for Engineering Sciences TFR grant 222/95-546.

where D_t^+ is the upper right Dini derivative with respect to time t. The upper left and right Dini derivatives are defined for a function $\psi(t)$ by

$$\left(\mathrm{D}_t^{\pm}\psi\right)(t) = \limsup_{h\to 0\pm} \frac{\psi(t+h)-\psi(t)}{h},$$

and the *logarithmic norm* is commonly defined as

$$\mu[A] = \lim_{h\to 0+} \frac{\|I+hA\|-1}{h}. \tag{1.3}$$

The matrix and logarithmic norms are subordinate to the vector norm $\|\cdot\|$ on $\mathbb{R}^d$. The differential inequality (1.2) still holds if A is time dependent, and by integration one obtains the perturbation bound

$$\|x(t)\| \le \mathrm{e}^{a(t)}\|x(0)\| + \mathrm{e}^{a(t)}\int_0^t \mathrm{e}^{-a(s)}\|p(s)\|\,\mathrm{d}s, \tag{1.4}$$

where $a(t) = \int_0^t \mu[A(\tau)]\,\mathrm{d}\tau$ and $\mu[A(\tau)]$ is defined by (1.3). A solution of the problem (1.1) is therefore stable on $[0,\infty)$ if $a(t)$ is bounded, and asymptotically stable if $a(t) \to -\infty$ as $t \to \infty$. Thus the condition $\mu[A] \le C < 0$ implies asymptotic stability (in fact, exponential stability).

The logarithmic norm has been extended to wider classes of maps than matrices, e.g. to *nonlinear systems*, [15], for studying stability issues in nonlinear dynamics (ODEs). By showing that the least upper bound Lipschitz constant $L[\cdot]$, defined for u and v in $\mathrm{Dom}(f)$ by

$$L[f] = \sup_{u\neq v} \frac{\|f(u)-f(v)\|}{\|u-v\|}, \tag{1.5}$$

is an operator seminorm, one defines the *lub logarithmic Lipschitz constant*

$$M[f] = \lim_{h\to 0+} \frac{L[I+hf]-1}{h} \tag{1.6}$$

in analogy with (1.3); this limit exists on account of every norm being left and right Gateaux differentiable. For a dynamical system

$$\begin{aligned} u' &= f(u) + p(t) \\ v' &= f(v) \end{aligned}$$

one obtains the differential inequality $\mathrm{D}_t^+\|u-v\| \le M[f]\cdot\|u-v\| + \|p(t)\|$, from which perturbation bounds are again derived in a standard manner.

For linear *differential-algebraic equations* the basic approach is developed in [10] by deriving bounds for the vector seminorm $\|Ax\|$ of the solution to

$$Ax' = Bx. \tag{1.7}$$

Its upper right Dini derivative is

$$\begin{aligned} \mathrm{D}_t^+\|Ax\| &= \limsup_{h\to 0+} \frac{\|Ax(t+h)\| - \|Ax(t)\|}{h} \\ &= \lim_{h\to 0+} \frac{\|Ax + hAx'\| - \|Ax\|}{h} \\ &= \lim_{h\to 0+} \frac{\|(A+hB)x\|/\|Ax\| - 1}{h}\,\|Ax\|. \end{aligned}$$

In order to properly define a logarithmic norm for the matrix pair (A, B) one defines a matrix seminorm, corresponding to the vector seminorm $\|A\cdot\|$, through

$$\|A, B\|_V = \max_{0\neq x\in V} \frac{\|Bx\|}{\|Ax\|}$$

where the admissible subspace V is required to satisfy $V \cap \mathrm{Ker}A \subseteq \mathrm{Ker}B$. For regular pencils, however, $\mathrm{Ker}A \cap \mathrm{Ker}B = \{0\}$ and therefore this condition may be written $V \cap \mathrm{Ker}A = \{0\}$. A logarithmic norm for matrix pencils may now be defined as

$$\mu_V[A, B] = \lim_{h\to 0+} \frac{\|A, A + hB\| - 1}{h}\,, \tag{1.8}$$

where V is a suitable admissible subspace; we now have the differential inequality $\mathrm{D}_t^+\|Ax\| \leq \mu_V[A, B]\cdot\|Ax\|$. A similar inequality is obtained when A is time-dependent, [10], by rewriting $A(t)x' = B(t)x$ as $(A(t)x)' = (B(t)+A'(t))x$, but the admissible subspace $V(t)$ will now generally depend on t, as does the seminorm $\|A\cdot\|$. For index 1 and index 2 problems, under some conditions on the matrices involved, one may derive asymptotic stability from the exponential decay of $\|A(t)x(t)\|$. This approach to time-dependent problems is also connected to the notion of *B-stability* for Runge–Kutta methods; a B-stable method will be contractive in the seminorm $\|Ax\|$ provided that $\mu_{V(t)}[A, B + A'] \leq 0$ and the internal stages are contained in the admissible subspaces.

Further, some results for unbounded operators make it possible to study evolution equations in partial differential equations (PDEs). For an abstract initial value problem $u' = f(u)$, one then needs to replace the standard definitions (1.3) and (1.6), which require A and f to be bounded (i.e. Lipschitz continuous) operators. In Hilbert space one defines, for u and v in $\mathrm{Dom}(f)$,

$$M[f] = \sup_{u\neq v} \frac{\langle u - v, f(u) - f(v)\rangle}{\langle u - v, u - v\rangle}\,. \tag{1.9}$$

The map or operator f is then no longer required to be Lipschitz for the *lub* logarithmic Lipschitz constant to be well defined. The condition $M[f] < 0$ is often referred to as a monotonicity condition. An extension to Banach space is made possible through the use of semi-inner products; this approach will be developed below.

The purpose of this paper is to unify these approaches and develop the *conceptual framework* needed for analyzing nonlinear DAEs. We shall avoid working with linearizations and logarithmic norm notions that are only applicable to bounded operators; the aim is a theory that can be applied to elementary problems in nonlinear partial differential-algebraic equations (PDAEs).

2 Semi-inner products

While the *lub* logarithmic Lipschitz constant $M[\cdot]$ can be directly extended to unbounded operators in Hilbert space via (1.9), the general case requires the use of semi-inner products, [3, p. 123].

DEFINITION 2.1. *Let X be a Banach space. For $u, v \in X$, the left $(\cdot\,,\cdot)_-$ and right $(\cdot\,,\cdot)_+$ semi-inner products are defined by*

$$(u,v)_{\pm} = \|u\| \lim_{\varepsilon\to 0\pm} \frac{\|u+\varepsilon v\| - \|u\|}{\varepsilon}. \tag{2.1}$$

The limits in (2.1) exist as they are G–differentials of the vector norm $\|\cdot\|$. If the norm is induced by a true inner product, then $(u,v)_{\pm} = \langle u, v\rangle$. The relation between a general norm and the semi-inner products is also similar to the case of a true inner product, as $(u,u)_{\pm} = \|u\|^2$. Likewise, the semi-inner products satisfy the Cauchy–Schwarz inequalities

$$-\|u\|\|v\| \le (u,v)_{\pm} \le \|u\|\|v\|. \tag{2.2}$$

These as well as the following elementary properties of the semi-inner products are consequences of the basic vector norm properties:

PROPOSITION 2.1. *For $u, v, w \in X$,*

1. $(u,-v)_{\pm} = -(u,v)_{\mp}$
2. $(u,\alpha v)_{\pm} = \alpha(u,v)_{\pm}; \qquad \alpha \ge 0$
3. $(u,v)_- + (u,w)_{\pm} \le (u,v+w)_{\pm} \le (u,w)_{\pm} + (u,v)_+$.

Each equality and inequality of Proposition 2.1 is two-fold; in each case one combines the left and right semi-inner products correctly by choosing either the upper or lower subscripts. Proposition 2.1.3 shows that the two semi-inner products are almost the same. Indeed, $(u,v)_+ = (u,v)_-$ if and only if the dual X^* is strictly convex. Although that condition excludes $\|\cdot\|_\infty$, the two semi-inner products are nevertheless equal a.e. in this topology, [3, p. 124].

An advantage of analyzing initial value problems with semi-inner products is that the latter are directly related to the upper Dini derivatives, as

$$\mathrm{D}_t^{\pm}\|u\| = \frac{(u,u')_{\pm}}{\|u\|^2}\,\|u\|, \tag{2.3}$$

see [3, p. 124]. Differential inequalities for estimating the solutions of ODEs (as well as DAEs) are then obtained by a matching definition of the extended logarithmic Lipschitz constants. Thus we now extend definitions (1.6) and (1.9) as follows.

DEFINITION 2.2. *Let $f : X \to X$. The least upper bound logarithmic Lipschitz constants with respect to the semi-inner products $(\cdot\,,\cdot)_\pm$ are defined by*

$$M^\pm[f] = \sup_{u \neq v} \frac{\big(u - v, f(u) - f(v)\big)_\pm}{\|u - v\|^2}\,; \tag{2.4}$$

The standard definition (1.6) is a G–differential of a supremum, the Lipschitz constant. The extended definition (2.4) is essentially the converse, the supremum of an expression containing a G–differential. This is known to be the key to handling unbounded operators, [14]. Comparing (2.4) to (1.6), one should note that for any map f, it holds that $M^-[f] \le M^+[f] \le M[f]$. If the norm is induced by an inner product, however, equality holds: $M^-[f] = M^+[f] = M[f]$.

The basic properties of the logarithmic Lipschitz constant are:

PROPOSITION 2.2.

1. $-l[f] \le M^\pm[f] \le L[f]$
2. $M^\pm[f + zI] = M^\pm[f] + \operatorname{Re} z$
3. $M^\pm[\alpha f] = \alpha M^\pm[f], \qquad \alpha \ge 0$
4. $M^\pm[f + g] \le M^\pm[f] + M^+[g]$.

Here $l[f]$ denotes the *glb Lipschitz constant*, defined by

$$l[f] = \inf_{u \neq v} \frac{\|f(u) - f(v)\|}{\|u - v\|}. \tag{2.5}$$

This extends the greatest lower bound of a matrix, as $l[A] = \operatorname{glb}[A] = \|A^{-1}\|^{-1}$, [15]. The left inequality of Proposition 2.2.1 corresponds to the uniform monotonicity theorem, [4, p. 147], [12, p. 167], [15], which is of importance for bounding inverses. The inequalities of Proposition 2.2.4 combine the left and right semi-inner products (as in Proposition 2.1, one chooses either the plus or the minus signs). Since subadditivity is important when deriving stability conditions and perturbation bounds, we prefer the right semi-inner product and $M^+[f]$, even if marginally stronger results could sometimes be obtained with the left.

3 Nonlinear index 1 DAEs

Consider a nonlinear DAE

$$Ax' = g(x); \qquad x(0) = x_0 \tag{3.1}$$

where A is a singular matrix which can be written $A = UV^{\mathrm{T}}$ with $V^{\mathrm{T}}U$ nonsingular. Let T be a nonsingular matrix with invariant subspace U such that $TU = U(V^{\mathrm{T}}U)^{-1}$. Then $TA = TUV^{\mathrm{T}} = U(V^{\mathrm{T}}U)^{-1}V^{\mathrm{T}}$ is a projector onto U along $V_\perp$. Hence (3.1) is equivalent to $TAx' = Tg(x)$ or $Px' = Tg(x)$, i.e., we may without loss of generality assume that the matrix multiplying x' is a *projector*. Therefore we assume that the nonlinear DAE is given in the form

$$Px' = f(x); \qquad x(0) = x_0 \tag{3.2}$$

where x_0 is a consistent initial value and P is a projector which may depend on t. In addition to P we introduce the complementary projector $Q = I - P$; then $PQ = QP = 0$. Obviously, the solution of (3.2) is in the manifold

$$\mathcal{M}_0 = \{x \,:\, Qf(x) = 0\}\,. \tag{3.3}$$

If P is continuously differentiable, the DAE (3.2) may be written

$$(Px)' = \big(P' + f\big)(x). \tag{3.4}$$

or in the equivalent projected form

$$(Px)' \;=\; \big(P' + Pf\big)(x) \tag{3.5a}$$
$$0 \;=\; Qf(x). \tag{3.5b}$$

We shall bound the projections $\|P\delta x\|$ and $\|Q\delta x\|$ of the difference δx between the solutions of (3.5) and its perturbation $P(x + \delta x)' = f(x + \delta x) + p$, i.e.,

$$\big(P(x + \delta x)\big)' \;=\; \big(P' + Pf\big)(x + \delta x) + Pp \tag{3.6a}$$
$$0 \;=\; Qf(x + \delta x) + Qp, \tag{3.6b}$$

with initial value $(x + \delta x)(0) = x_0 + \delta x_0$. Here, the perturbed solution is in the perturbed manifold

$$\mathcal{M}_{Qp} = \{x \,:\, Qf(x) = -Qp\}\,. \tag{3.7}$$

We shall analyze the index 1 case and exemplify some of the crucial points in a simple example.

Example 3.1 Consider a partitioned DAE of index 1 and its perturbation:

$$y' = F(y, z)\,; \qquad (y + \delta y)' = F(y + \delta y, z + \delta z) + r \tag{3.8}$$
$$0 = G(y, z)\,; \qquad 0 = G(y + \delta y, z + \delta z) + s. \tag{3.9}$$

Assume that G is Lipschitz with respect to its first argument. Index 1 problems allow us to use the implicit function theorem to solve the algebraic equation, and we may write

$$z = \eta(y, 0)\,; \qquad z + \delta z = \eta(y + \delta y, s)\,, \tag{3.10}$$

where $\eta(\cdot, \cdot)$ is Lipschitz with respect to both arguments, [12, p. 128]. This yields $\delta z = \eta(y + \delta y, s) - \eta(y, 0) = \eta(y + \delta y, s) - \eta(y, s) + \eta(y, s) - \eta(y, 0)$, which implies the bound

$$\|\delta z\| \le L[\eta(\cdot, s)] \cdot \|\delta y\| + L[\eta(y, \cdot)] \cdot \|s\|\,, \tag{3.11}$$

where the two Lipschitz constants refer to the first and second arguments of $\eta(\cdot,\cdot)$, evaluated at s and y, respectively. By inserting the solutions (3.10) into (3.8), we obtain

$$\begin{aligned}\delta y' &= F\big(y+\delta y, \eta(y+\delta y, s)\big) - F\big(y, \eta(y,0)\big) + r\\ &= F\big(y+\delta y, \eta(y+\delta y, s)\big) - F\big(y, \eta(y,s)\big)\\ &\quad + F\big(y, \eta(y,s)\big) - F\big(y, \eta(y,0)\big) + r.\end{aligned}$$

From this equation we therefore obtain the differential inequality

$$\mathrm{D}_t^+\|\delta y\| \le M^+[F\big(\cdot,\eta(\cdot,s)\big)]\cdot\|\delta y\| + L[F\big(y,\eta(y,\cdot)\big)]\cdot\|s\| + \|r\|. \tag{3.12}$$

By integration, $\|\delta y\|$ is bounded in terms of $\|r\|$ and $\|s\|$; the bound for $\|\delta z\|$ then follows from (3.11). The bound is similar to (1.4) and holds on compact intervals under the assumption $M^+[F\big(\cdot,\eta(\cdot,s)\big)] < \infty$ and on infinite intervals if $M^+[F\big(\cdot,\eta(\cdot,s)\big)] < 0$. In both cases we need to require that $L[F\big(y,\eta(y,\cdot)\big)]$ remains bounded, *unless* $s = 0$. It is important to note that, *in the nonlinear case*, $M^+[F\big(\cdot,\eta(\cdot,s)\big)]$ depends on s; a perturbation of the manifold then affects the system's stability. However, as the index 1 condition implies that η is Lipschitz with respect to both arguments, the additional assumption that F is Lipschitz with respect to its second argument implies that $M^+[F\big(\cdot,\eta(\cdot,s)\big)]$ *is a continuous function of* s. More precisely, $|M^+[F\big(\cdot,\eta(\cdot,s)\big)] - M^+[F\big(\cdot,\eta(\cdot,0)\big)]| = O(\|s\|)$, and consequently a small perturbation s will only have a small (degrading) effect on stability. Finally, note that it is also possible to derive the inequality $\mathrm{D}_t^+\|\delta y\| \le M^+[F\big(\cdot,\eta(\cdot,0)\big)]\cdot\|\delta y\| + L[F\big(y+\delta y,\eta(y+\delta y,\cdot)\big)]\cdot\|s\| + \|r\|$. In that case, however, the condition $M^+[F\big(\cdot,\eta(\cdot,0)\big)] < 0$ does *not* guarantee the stability of the solution y, as the forcing term $L[F\big(y+\delta y,\eta(y+\delta y,\cdot)\big)]\|s\|$ then depends on δy.

Let us now return to (3.2), where there is no explicit partitioning. We shall bound the projections $P\delta x$ and $Q\delta x$. For simplicity we let P be constant and start by defining *restricted (logarithmic) Lipschitz constants of projected maps.*

DEFINITION 3.1. *Let* $f : \mathrm{Dom}(f) \subset X \to X$ *be a given map, where* $\mathrm{Dom}(f)$ *is open. The* P*–restricted Lipschitz constant of* Qf *is defined by*

$$L_P[Qf] = \sup_{\substack{u,\, u+P\delta u \in \mathrm{Dom}(f)\\ P\delta u \neq 0}} \frac{\|Qf(u+P\delta u) - Qf(u)\|}{\|P\delta u\|}, \tag{3.13}$$

and the Q*–restricted logarithmic Lipschitz constant of* Qf *is defined by*

$$M_Q^+[Qf] = \sup_{\substack{v,\, v+Q\delta v \in \mathrm{Dom}(f)\\ Q\delta v \neq 0}} \frac{\big(Q\delta v, Qf(v+Q\delta v) - Qf(v)\big)_+}{\|Q\delta v\|^2}. \tag{3.14}$$

The restricted constants $M_P^+[Pf]$ *and* $L_Q[Pf]$ *are defined analogously.*

REMARK. Note that for a linear map $f = B$, the logarithmic Lipschitz constant $M_Q^+[QB] = \mu_{\mathrm{Im}(Q)}[Q, QB]$, where the latter expression is the logarithmic norm of the matrix pencil (Q, QB) as developed in [10].

LEMMA 3.1. *Let $f : \mathrm{Dom}(f) \subset X \to X$ be a given map and let P and Q be two complementary projectors on X. Then*

$$\big(Q\delta x, Qf(x+\delta x) - Qf(x)\big)_+ \le L_P[Qf]\cdot\|P\delta x\|\,\|Q\delta x\| + M_Q^+[Qf]\cdot\|Q\delta x\|^2.$$

PROOF. By the Cauchy–Schwarz inequality, Proposition 2.1.3 and the definitions of $L_P[Qf]$ and $M_Q[Pf]$, we have

$$\begin{aligned} &\big(Q\delta x, Qf(x+\delta x) - Qf(x)\big)_+ \le \\ &\big(Q\delta x, Qf(x+\delta x) - Qf(x+P\delta x)\big)_+ + \big(Q\delta x, Qf(x+P\delta x) - Qf(x)\big)_+ \\ &\le M_Q^+[Qf]\cdot\|Q\delta x\|^2 + L_P[Qf]\cdot\|P\delta x\|\|Q\delta x\|, \end{aligned}$$

where $\delta x = P\delta x + Q\delta x$ allows the interpretation $v := x + P\delta x$, $Q\delta v := Q\delta x$ in the first term, and $u := x$, $P\delta u := P\delta x$ in the second. ◻

THEOREM 3.2. *For the problem (3.5) and (3.6) the following differential-algebraic inequality holds:*

$$\mathrm{D}_t^+\|P\delta x\| \le M_P^+[Pf]\cdot\|P\delta x\| + L_Q[Pf]\cdot\|Q\delta x\| + \|Pp\| \tag{3.15a}$$
$$0 \le L_P[Qf]\cdot\|P\delta x\| + M_Q^+[Qf]\cdot\|Q\delta x\| + \|Qp\|. \tag{3.15b}$$

PROOF. Since $0 = Qf(x)$ and $0 = Qf(x+\delta x) + Qp$, the Cauchy–Schwarz inequality yields

$$-\|Q\delta x\|\|Qp\| \le \big(Q\delta x, Qf(x+\delta x) - Qf(x)\big)_+$$

from which the algebraic inequality (3.15b) follows by application of Lemma 3.1. For the differential part, by (3.5a), (3.6a), and the basic relation (2.3) between Dini derivatives and semi-inner products, we have

$$\mathrm{D}_t^+\|P\delta x\| = \big(P\delta x, Pf(x+\delta x) - Pf(x)\big)_+/\|P\delta x\| \tag{3.16a}$$
$$\le M_P^+[Pf]\cdot\|P\delta x\| + L_Q[Pf]\cdot\|Q\delta x\| + \|Pp\|, \tag{3.16b}$$

by a derivation analogous to that of Lemma 3.1. ◻

The algebraic inequality (3.15b) is of course trivial, unless one can establish that $M_Q^+[Qf] < 0$. This monotonicity condition plays a central role; it guarantees that the DAE is index 1, it is necessary in order to obtain bounds from the differential-algebraic inequality, and last but not least it leads to the following version of the implicit function theorem concerning the solvability of the two algebraic equations (3.5b) and (3.6b) on certain subspaces.

THEOREM 3.3. (Implicit function theorem) *Let P and Q be complementary projectors and assume that $M_Q^+[Qf] < 0$. Given $-Qp \in \mathrm{Im}(Qf)$, the equations $0 = Qf(x)$ and $0 = Qf(x + \delta x) + Qp$ are solvable for*

$$Qx = \varphi(Px, 0) \tag{3.17a}$$

$$Q(x + \delta x) = \varphi\big(P(x + \delta x), Qp\big), \tag{3.17b}$$

where the function $\varphi(\cdot, \cdot)$ is Lipschitz with respect to its second argument, and with respect to its first argument provided that $L_P[Qf] < \infty$.

PROOF. Since $M_Q^+[Qf] < 0$, the uniform monotonicity theorem, [4, p. 147], [12, p. 167], [15], is applicable and implies existence and uniqueness of the solutions. As for the Lipschitz continuity of $\varphi(\cdot, \cdot)$, subtracting (3.17a) from (3.17b) implies the formal bound

$$\|Q\delta x\| \le L[\varphi(\cdot, Qp)]\, \|P\delta x\| + L[\varphi(Px, \cdot)]\, \|Qp\|. \tag{3.18}$$

By inequality (3.15b), we have the actual bound

$$\|Q\delta x\| \le -\frac{L_P[Qf]}{M_Q^+[Qf]}\, \|P\delta x\| - \frac{\|Qp\|}{M_Q^+[Qf]}\,, \tag{3.19}$$

showing that $Q\delta x$ indeed depends continuously on $P\delta x$ and Qp. ∎

REMARK. Note that $\varphi(\cdot, \cdot) = Q\varphi(\cdot, \cdot)$.

In reality, $P\delta x$ and $Q\delta x$ are interdependent; the theorem merely states an index 1 condition necessary for bounding these perturbations. The simplest bound is obtained by assuming that $M_Q^+[Qf] < 0$ and inserting (3.19) into (3.15a). We then have:

THEOREM 3.4. *Let P and Q be complementary projectors and assume that $M_Q^+[Qf] < 0$. Then the following differential inequality holds for (3.5) and (3.6),*

$$\mathrm{D}_t^+ \|P\delta x\| \le S[P, f]\, \|P\delta x\| - \frac{L_Q[Pf]}{M_Q^+[Qf]}\, \|Qp\| + \|Pp\|, \tag{3.20}$$

where the "Schur complement" logarithmic Lipschitz constant is

$$S[P, f] = M_P^+[Pf] - \frac{L_Q[Pf] L_P[Qf]}{M_Q^+[Qf]}. \tag{3.21}$$

From (3.19) and (3.20) it is clear that x is an exponentially stable solution if $S[P, f] < 0$, or equivalently, if $M_P^+[Pf] M_Q^+[Qf] > L_Q[Pf] L_P[Qf]$. Bounds on $P\delta x$ and $Q\delta x$ can, however, be obtained already when $M_Q^+[Qf] < 0$ and the other three constants in (3.15) are finite. However, bounds derived from (3.20) do not reflect the interdependence between $P\delta x$ and $Q\delta x$. This is accounted for by inserting $Q(x + \delta x) = \varphi\big(P(x + \delta x), Qp\big) = Q\varphi\big(P(x + \delta x), Qp\big)$ into (3.16a). We define the map

$$g(\cdot, Qp) :\, u \mapsto u + Q\varphi(u, Qp). \tag{3.22}$$

Then (3.16a) yields

$$\begin{aligned} Pf(x+\delta x) &= Pf\big(P(x+\delta x)+Q(x+\delta x)\big) \\ &= Pf\big(P(x+\delta x)+Q\varphi(P(x+\delta x),Qp)\big) \\ &= Pfg\big(P(x+\delta x),Qp\big). \end{aligned}$$

In a similar way we find that $Pf(x) = Pfg(Px, 0)$. Therefore,

$$\begin{aligned} Pf(x+\delta x) - Pf(x) &= Pfg(P(x+\delta x),Qp) - Pfg(Px,Qp) \\ &\quad + Pfg(Px,Qp) - Pfg(Px,0), \end{aligned}$$

whose first and second parts give rise to the first and second terms, respectively, of the differential inequality of the following theorem:

THEOREM 3.5. *Let P and Q be complementary projectors and assume that $M_Q^+[Qf] < 0$. Further, let g be defined by (3.22), where φ is the function provided by the Implicit Function Theorem 3.3. Then, for (3.5) and (3.6) it holds that*

$$\mathrm{D}_t^+\|P\delta x\| \;\le\; M_P^+[Pfg(\cdot,Qp)]\cdot\|P\delta x\| + L_Q[Pfg(Px,\cdot)]\cdot\|Qp\| + \|Pp\|.$$

This result, which is analogous to (3.12), is directly obtained from (3.16a). We note that $Q\delta x$ has been eliminated from the differential inequality, which yields a bound for $P\delta x$ directly in terms of Pp and Qp. The bound for $Q\delta x$ is then obtained from (3.19). Bounds obtained in this way are stronger than those obtained from the differential-algebraic inequality (3.15), but require an explicit knowledge of φ.

REMARKS. 1) As expected, Theorem 3.5 results in bounds where the stability condition $M_P^+[Pfg(\cdot,Qp)] \le 0$ depends on the perturbation Qp of the manifold. The case $Qp = 0$ is however a theoretically important special case, in which $Pfg(Px,\cdot)$ is not required to be Lipschitz. Both x and $x+\delta x$ are then in the manifold $\mathcal{M}_0$. This case occurs e.g. in the error analysis of algebraically accurate discretization methods *as long as the nonlinear equations in each step are solved exactly.* Iteration errors in the equation solving process, however, will in practice introduce a residual $Qp \neq 0$, as will discretization methods that are not algebraically accurate. Then $x \in \mathcal{M}_0$ but $x+\delta x \in \mathcal{M}_{Qp}$.

2) Unlike the Lipschitz and logarithmic Lipschitz constants of Definition 3.1, which are computed by taking the supremum over all of Dom(f) and are therefore of *global* type (compare (2.4)), the constants of Theorem 3.5 (as well as those of Example 3.1) may also be defined along a reference solution x. For a given $x(t)$ one may thus use the less restrictive *local* alternative

$$M_P^+[Pfg(\cdot,Qp)] = \sup_{P\delta x\neq 0} \frac{\big(P\delta x, Pfg(x+P\delta x,Qp)-Pfg(x,Qp)\big)_+}{\|P\delta x\|^2}, \tag{3.23}$$

in lieu of (2.4). Then $M_P^+[Pfg(\cdot,Qp)]$ depends on $x(t)$, and perturbation bounds can be obtained from the differential inequality locally, in a "tube" around the reference solution; such bounds are stronger than if one needs to consider the entire Dom(f) like in Theorem 3.2.

3) Let $E(t) = \exp \int_0^t M_P^+[Pfg(\cdot, Qp(\tau))]\, d\tau$ and let us for simplicity assume that $P\delta x(0) = 0$. Then

$$\|P\delta x(t)\| \leq \int_0^t \big(L_Q[Pfg] \cdot \|Qp(s)\| + \|Pp(s)\|\big) E(t)/E(s)\, \mathrm{d}s. \tag{3.24}$$

Since $\|\delta x\| \leq \|P\delta x\| + \|Q\delta x\|$, the bound (3.19) now yields

$$\|\delta x\| \leq \left(1 - \frac{L_P[Qf]}{M_Q^+[Qf]}\right) \|P\delta x\| - \frac{\|Qp\|}{M_Q^+[Qf]}\,. \tag{3.25}$$

This bound is of "perturbation index" type (see [9]): the deviation $\|\delta x\|$ is in part directly proportional to the residual $\|Qp\|$, and in part proportional to an integral of $\|Pp\|$ and $\|Qp\|$. Thus if $p \in C^0$ then $P\delta x \in C^1$, but $Q\delta x$ is only C^0. One needs in addition that $Qp \in C^1$ for the entire solution to be C^1, i.e., Qp needs to be smoother than Pp. The bound also displays what the relevant proportionality constants are; this provides information on stability as well as conditioning of an index 1 DAE. One notes that the value of $M_Q^+[Qf]$ plays a central role; as $M_Q^+[Qf] \to 0-$, the bound (3.25) breaks down. This is to be expected as the same condition essentially implies that the DAE index can no longer be guaranteed to be 1.

4) The case when P is time dependent is a straightforward generalization and the corresponding results are obtained by the substitution $Pf \to P' + Pf$. We further note that in the ODE case $P = I$, $Q = 0$ and $g \equiv I$; the bounds then reduce to the standard bounds for ODEs.

4 Discretization: The implicit Euler method

For a sequence $\{u_n\}$ we introduce the backward difference operator ∇, defined by $\nabla : \{u_n\} \mapsto \{u_n - u_{n-1}\}$ and use the shorthand notation $\nabla u_n = u_n - u_{n-1}$. For a sequence $\{t_n\}$ of points in time, we let $h = \nabla t_n$ denote the stepsize. The implicit Euler discretization of the index 1 DAE (3.2) is then

$$Ph^{-1}\nabla x_n = f(x_n), \tag{4.1}$$

or, in the equivalent projected form,

$$Ph^{-1}\nabla x_n = Pf(x_n) \tag{4.2a}$$
$$0 = Qf(x_n). \tag{4.2b}$$

The question of existence and uniqueness of solutions to (4.2) will be deferred until the necessary difference-algebraic inequalities have been established.

In actual computations, iteration errors will be present; the exact numerical solution $\{x_n\}$ of (4.1) is then approximated by $\{\tilde{x}_n\}$, satisfying

$$Ph^{-1}\nabla \tilde{x}_n = Pf(\tilde{x}_n) + Pr_n \tag{4.3a}$$
$$0 = Qf(\tilde{x}_n) + Qr_n, \tag{4.3b}$$

where $\{r_n\}$ are the residuals in solving the nonlinear algebraic equation (4.1). For the exact solution to the projected DAE (3.5) there holds

$$P(h^{-1}\nabla x(t_n) - d_n) = Pf(x(t_n)) \tag{4.4a}$$
$$0 = Qf(x(t_n)), \tag{4.4b}$$

where the differentiation error $d_n = h^{-1}\nabla x(t_n) - x'(t_n)$. We define the *global error* by $e_n = \tilde{x}_n - x(t_n)$. Then we have the following global error equation,

$$Ph^{-1}\nabla e_n = Pf(x(t_n)+e_n) - Pf(x(t_n)) + P(r_n - d_n) \tag{4.5a}$$
$$0 = Qf(x(t_n)+e_n) - Qf(x(t_n)) + Qr_n, \tag{4.5b}$$

where taking $r_n \equiv 0$ turns e_n into the theoretical global error $x_n - x(t_n)$. If one in addition puts $d_n \equiv 0$, then e_n represents the difference between two neighboring (exact) numerical solutions, with different initial data, for the study of numerical stability or for proving existence and uniqueness of solutions to (4.2).

We shall derive global error bounds and link them to those of the continuous case by using semi-inner products. In analogy with (2.3) we therefore make the following definition.

DEFINITION 4.1. *The right Dini difference quotient $h^{-1}\nabla_n^+\|u_n\|$ is defined by*

$$h^{-1}\nabla_n^+\|u_n\| = \frac{(u_n, h^{-1}\nabla u_n)_+}{\|u_n\|^2}\,\|u_n\|. \tag{4.6}$$

The right Dini difference quotient is related to $\nabla\|u_n\|$ by the following bound:

LEMMA 4.1. *For every $\{u_n\}$ there holds $\nabla\|u_n\| \le \nabla_n^+\|u_n\| \le \|\nabla u_n\|$.*

PROOF. For the lower bound, the triangle inequality yields

$$\begin{aligned}\nabla_n^+\|u_n\| &= \lim_{\varepsilon\to 0+}\frac{\|u_n+\varepsilon\nabla u_n\| - \|u_n\|}{\varepsilon} = \lim_{\varepsilon\to 0+}\frac{\|(1+\varepsilon)u_n - \varepsilon u_{n-1}\| - \|u_n\|}{\varepsilon}\\ &\ge \lim_{\varepsilon\to 0+}\frac{\|(1+\varepsilon)u_n\| - \varepsilon\|u_{n-1}\| - \|u_n\|}{\varepsilon} = \|u_n\| - \|u_{n-1}\| = \nabla\|u_n\|.\end{aligned}$$

Finally, from the Cauchy–Schwarz inequality $(u_n, \nabla u_n)_+ \le \|u_n\|\,\|\nabla u_n\|$ we obtain the upper bound. ∎

The lower bound admits the study of contractivity as $h^{-1}\nabla_n^+\|u_n\| \le 0$ implies that $\|u_n\|$ is a non-increasing sequence. Hence the monotonicity conditions of the previous section will also provide conditions for numerical stability.

As a discrete analogue of Theorem 3.2 we now have:

THEOREM 4.2. *For the global error equation (4.5) the following difference-algebraic inequality holds:*

$$h^{-1}\nabla_n^+\|Pe_n\| \le M_P^+[Pf]\cdot\|Pe_n\| + L_Q[Pf]\cdot\|Qe_n\| + \|P(r_n - d_n)\| \tag{4.7a}$$
$$0 \le L_P[Qf]\cdot\|Pe_n\| + M_Q^+[Qf]\cdot\|Qe_n\| + \|Qr_n\|, \tag{4.7b}$$

where the four constants are given in Definition 3.1.

PROOF. The result is obtained by using Definition 4.1, the properties of the semi-inner product, and the proof techniques of Lemma 3.1. ◻

By combining Lemma 4.1 and Theorem 4.2 we now formally obtain

$$\begin{bmatrix} 1 - hM_P^+[Pf] & -hL_Q[Pf] \\ -L_P[Qf] & -M_Q^+[Qf] \end{bmatrix} \begin{bmatrix} \|Pe_n\| \\ \|Qe_n\| \end{bmatrix} \leq \begin{bmatrix} \|Pe_{n-1}\| + h\|P(r_n - d_n)\| \\ \|Qr_n\| \end{bmatrix}, \tag{4.8}$$

where the vector inequality is interpreted component-wise. Note that we have re-scaled (4.7a), but not (4.7b), by a factor of h. If the matrix on the left-hand side of (4.8) satisfies $hM_P^+[Pf] < 1$, $M_Q^+[Qf] < 0$ and is diagonally dominant, then it is an M–matrix and hence has a positive inverse (i.e., with all matrix elements positive), which is needed to maintain the vector inequality after inversion. The condition of diagonal dominance may be replaced by the simpler condition that the matrix has a positive determinant, which requires that the Schur complement logarithmic Lipschitz constant fulfills $hS[P, f] < 1$, which in turn implies $hM_P^+[Pf] < 1$.

We are now in a position to prove existence and uniqueness of (4.1) as well as deriving global error bounds. Starting with the former task, we write (4.1) in the form

$$\big(P - hf\big)(x_n) = Px_{n-1}, \tag{4.9}$$

and assume, for a fixed n, that $Px_{n-1} \in \mathrm{Im}(P - hf)$; this implies the existence of a pre-image of Px_{n-1} in $\mathrm{Dom}(P - hf)$. Uniqueness requires showing that $P-hf$ is an injection, and a sufficient condition is provided by the uniform monotonicity theorem in the following form. Let $r_n = d_n = 0$ in (4.8); e_n then represents the difference between two solutions to (4.9). Assume the monotonicity conditions $M_Q^+[Qf] < 0$ and $hS[P, f] < 1$. By (4.8), if the data $Pe_{n-1} = 0$ then necessarily $e_n = 0$, and it follows that $P - hf$ is a one-to-one map with Lipschitz inverse. Global existence for all n, however, depends crucially on $\mathrm{Dom}(P - hf)$, which is determined by the particulars of the individual problem, and we shall leave that question aside by simply assuming that $\mathrm{Dom}(P - hf) = X$; the monotonicity conditions then imply that $P - hf$ is coercive, hence $\mathrm{Im}(P - hf) = X$. By induction it then follows that $Px_{n-1} \in \mathrm{Im}(P - hf)$ for all n. Thus, we have:

THEOREM 4.3. (Existence and uniqueness) *If* $Px_{n-1} \in \mathrm{Im}(P - hf) = X$, *and if* $M_Q^+[Qf] < 0$ *and* $hS[P, f] < 1$ *on* $\mathrm{Dom}(f) = X$, *then (4.1) – (4.2) have unique solutions. Moreover, if the residuals under consideration are in* $\mathrm{Im}(P - hf)$, *then the global error equation (4.5) possesses a unique solution* $\{e_n\}$.

REMARKS. 1) By (4.9), $(P - hf)^{-1}$ needs only to be defined on $\mathrm{Im}(P)$ for the *exact* numerical solution. However, in order to account for a residual r_n, and hence the global error e_n, the inverse is also needed outside $\mathrm{Im}(P)$.

2) A sharper existence and uniqueness theorem can be obtained by using the implicit function theorem and making use of the function $\varphi(\cdot, \cdot)$. The assumption

$M_Q^+[Qf] < 0$ allows us to use the map g, defined by (3.22), to obtain instead of (4.7) the difference inequality

$$\begin{aligned} h^{-1}\nabla_n^+\|Pe_n\| \quad &\le \quad M_P^+[Pfg(\cdot, Qr_n)]\|Pe_n\| \\ &+ L_Q[Pfg(Px(t_n), \cdot)]\|Qr_n\| + \|P(r_n - d_n)\|, \end{aligned}$$

and for any fixed n it follows that if in addition $hM_P^+[Pfg(\cdot, Qr_n)] < 1$, then $Pe_{n-1} = r_n = d_n = 0$ implies $Pe_n = 0$. This provides a result of local type in a neighborhood of $x(t_n)$.

3) In the ODE case, where $P = I$ and $Q = 0$, the condition for existence and uniqueness reduces to the usual condition $hM^+[f] < 1$, which implies that $(I - hf)^{-1}$ exists and is Lipschitz with constant $L[(I-hf)^{-1}] \le 1/(1-hM^+[f])$.

We now return to the derivation of global error bounds. Let the monotonicity conditions $M_Q^+[Qf] < 0$ and $hS[P, f] < 1$ be fulfilled and note that the error recursion (4.8) is of the form

$$Mw_n \le Kw_{n-1} + \psi_n \qquad \Rightarrow \qquad w_n \le M^{-1}Kw_{n-1} + M^{-1}\psi_n,$$

with $w_n = (\|Pe_n\|, \|Qe_n\|)^{\mathrm{T}}$, and where the positivity of M^{-1} follows from the two monotonicity conditions . This leads to the generic global error bound

$$w_n \le (M^{-1}K)^n w_0 + \big(I - (M^{-1}K)^n\big)\,(M - K)^{-1}\psi, \tag{4.10}$$

where $\psi = (h \sup_n \|P(r_n - d_n)\|, \sup_n \|Qr_n\|)^{\mathrm{T}}$ and, in the implicit Euler case,

$$M^{-1}K \;=\; \frac{1}{1 - hS[P, f]} \begin{bmatrix} 1 & 0 \\ -L_P[Qf]/M_Q^+[Qf] & 0 \end{bmatrix}. \tag{4.11}$$

It follows that $(M^{-1}K)^n = (1 - hS[P, f])^{1-n}M^{-1}K$ and hence $M^{-1}K$ is power-bounded if $S[P, f] \le 0$, which is a sufficient condition for numerical stability in the implicit Euler method. This reflects the stability condition for the continuous case as found in Theorem 3.4.

It is, however, simpler to derive the global error bound component-wise. As

$$\|Qe_n\| \le -\frac{L_P[Qf]}{M_Q^+[Qf]}\,\|Pe_n\| - \frac{\|Qr_n\|}{M_Q^+[Qf]} \tag{4.12}$$

we have

$$\big(1 - hS[P, f]\big)\|Pe_n\| \le \|Pe_{n-1}\| + h\big(\|P(r_n - d_n)\| - \frac{L_Q[Pf]}{M_Q^+[Qf]}\,\|Qr_n\|\big), \tag{4.13}$$

from which the global error bound is obtained in a standard way. Hence:

THEOREM 4.4. *Assume that $Pe_0 = 0$ and $M_Q^+[Qf] < 0$. If $hS[P, f] < 1$, then the projected global error Pe_n for the implicit Euler method is bounded by*

$$\|Pe_n\| \;\le\; \frac{(1 - hS[P, f])^{-n} - 1}{S[P, f]}\,\sup_{n>0}\big(\|P(r_n - d_n)\| - \frac{L_Q[Pf]}{M_Q^+[Qf]}\,\|Qr_n\|\big), \tag{4.14}$$

and the projection $\|Qe_n\|$ is bounded by (4.12).

REMARKS. 1) If $hS[P,f] \to 0$ for a fixed $t_n = nh$ in a compact interval $[0,T]$, then, by (4.14), the error satisfies the estimate

$$\|Pe_n\| \lessapprox \frac{\mathrm{e}^{t_n S[P,f]}-1}{S[P,f]} \sup_{n>0} \big(\|P(r_n-d_n)\| - \frac{L_Q[Pf]}{M_Q^+[Qf]}\,\|Qr_n\|\big). \tag{4.15}$$

Further, if $S[P,f] < 0$, then for all n, $h > 0$ as well as $t \to \infty$, there holds

$$\|Pe_n\| \le -\frac{1}{S[P,f]} \sup_{n>0} \big(\|P(r_n-d_n)\| - \frac{L_Q[Pf]}{M_Q^+[Qf]}\,\|Qr_n\|\big). \tag{4.16}$$

2) If the solution of the DAE is C^2, then the differentiation error $d_n = O(h)$ and (4.14) – (4.16) show, for the theoretical case $r_n = 0$, that the method is first order convergent. As for iteration errors, the bounds show that we must have $r_n = O(h)$ to obtain convergence. Note that, unlike in the continuous case, there is—from the formal point of view—not a more stringent requirement for the projection Qr_n than for Pr_n. In practice, however, we must nevertheless impose qualifications on the statement $r_n = O(h)$. Thus, the proportionality constants in the bounds show that if $L_P[Qf] \ll 1$, then $\|Qe_n\|$ may be dominated by the residual (see (4.12)) from the iterative equation-solving process. These residuals may also lack the smoothness of the differentiation error d_n. Likewise, (4.14) – (4.16) show that $\|Pr_n\|$ should be small compared to $\|Pd_n\|$ and that $L_Q[Pf]/M_Q^+[Qf]$, if large, may cause $\|Qr_n\|$ to have a considerable influence on the projected error $\|Pe_n\|$. In practical computations the control of the residual Qr_n therefore plays a much more crucial and subtle role than that of Pr_n; it is important to solve the algebraic equation to great accuracy.

EXAMPLE 4.1 To illustrate the above theory, we shall give an example of how it may be applied to a very simple PDAE problem. We thus consider the following nonlinear heat conduction problem [13, p. 220ff]

$$\rho\, c\, \vartheta_t = -\mathrm{div}\, q + h \tag{4.17}$$

$$q = -\mathrm{grad}\, \mathcal{H}(x, \mathrm{grad}\, \vartheta), \tag{4.18}$$

with suitable boundary and initial data. Here ρ, c are material constants, ϑ is the temperature and h is a source term. Further, q is the heat flux vector associated with the potential $\mathcal{H}$, which is assumed to satisfy the convexity condition

$$\langle \mathrm{grad}\, \mathcal{H}(\cdot, v), u - v\rangle \le \mathcal{H}(\cdot, u) - \mathcal{H}(\cdot, v). \tag{4.19}$$

This is an index 1 PDAE. For the difference $\delta\vartheta$ between two solutions ϑ_1 and ϑ_2 with different initial conditions (but the same forcing/source term) we then obtain the differential inequality

$$\mathrm{D}_t^+ \|\delta\vartheta\| \le M_P^+[Pfg(\cdot,0)] \cdot \|\delta\vartheta\|, \tag{4.20}$$

where we have $Pfg(\cdot,0) = \Delta\, \mathcal{H}(x, \mathrm{grad}\, \cdot)/(\rho\, c)$, where Δ is the Laplacian and the convexity condition on $\mathcal{H}$ implies that $M_P^+[Pfg(\cdot,0)] \le 0$. Therefore the system is contractive, i.e., $\|\delta\vartheta(\cdot,t)\| \le \|\delta\vartheta(\cdot,0)\|$.

5 Runge–Kutta methods and the relation to B-stability

In order to study some aspects of Runge–Kutta methods applied to index 1 DAEs, we let $(\mathcal{A}, b)$, with $\mathcal{A}$ regular, denote the coefficients of a given s–stage Runge–Kutta method. Applying this method to the DAE (3.2) yields, in standard Kronecker notation,

$$X_n = \mathbb{1} \otimes x_n + (\mathcal{A} \otimes I_d) h \dot{X}_n \tag{5.1a}$$

$$P \dot{X}_n^i = f(X_n^i); \qquad i = 1 : s \tag{5.1b}$$

$$x_{n+1} = x_n + (b^{\mathrm{T}} \otimes I_d) h \dot{X}_n, \tag{5.1c}$$

where X_n and $\dot{X}_n$ contain the s stage vectors X_n^i and stage derivative vectors $\dot{X}_n^i$, respectively. The Runge–Kutta step can therefore be written

$$x_{n+1} = R(\infty) x_n + (b^{\mathrm{T}} \mathcal{A}^{-1} \otimes I_d) X_n, \tag{5.2}$$

where $R(\infty) = 1 - b^{\mathrm{T}} \mathcal{A}^{-1} \mathbb{1}$ is recognized as the limit of the Runge–Kutta method's stability function $R(z) = 1 + z b^{\mathrm{T}} (I_s - z\mathcal{A})^{-1} \mathbb{1}$, and where the stage vectors satisfy $(I_s \otimes P) X_n = \mathbb{1} \otimes P x_n + (\mathcal{A} \otimes I_d) h F(X_n)$, in which the map F is defined through $(e_i^{\mathrm{T}} \otimes I_d) F = f \circ (e_i^{\mathrm{T}} \otimes I_d)$, i.e., the ith component of $F(X_n)$ equals $f(X_n^i)$. Equivalently, the equation for the stage vectors may be written

$$(\mathcal{A}^{-1} \otimes P - hF)(X_n) = \mathcal{A}^{-1} \mathbb{1} \otimes P x_n; \tag{5.3}$$

solvability then essentially follows if $(\mathcal{A}^{-1} \otimes P - hF)^{-1}$ exists on a suitably restricted subspace. In the classical ODE case, when $P = I_d$, this follows from the uniform monotonicity theorem if $M^+[-\mathcal{A}^{-1} \otimes I_d + hF] < 0$; a sufficient condition is $M^+[hF] < m^-[\mathcal{A}^{-1} \otimes I_d]$, where $m^-[\mathcal{A}^{-1} \otimes I_d] = -M^+[-\mathcal{A}^{-1} \otimes I_d]$, see [15, p. 31]. These conditions can for appropriate norms be reduced to the conditions $M^+[f] \le 0$ and $m^-[\mathcal{A}^{-1}] > 0$, compare theorem 5.3.9 in [4, p. 147]; the latter condition concerns the positive definiteness of $\mathcal{A}$ and corresponds to the condition $\psi_D(\mathcal{A}^{-1} \otimes I_d) > 0$ used by Dekker and Verwer, *op. cit.*, p. 134.

In the DAE case, we consider the projected version of (5.3), which, in the presence of computation errors, yields approximate stage vectors $\tilde{X}_n$ satisfying

$$\big(\mathcal{A}^{-1} \otimes P - (I_s \otimes P) hF\big)(\tilde{X}_n) = \mathcal{A}^{-1} \mathbb{1} \otimes P \tilde{x}_n + (I_s \otimes P) h R_n \tag{5.4a}$$

$$-(I_s \otimes Q) hF(\tilde{X}_n) = (I_s \otimes Q) h R_n, \tag{5.4b}$$

where R_n contains the residuals in solving the nonlinear equation (5.3). The data, or the starting point, $\tilde{x}_n$, equals x_n of the exact numerical equation (5.3) if $R_n \equiv 0$; the difference $\tilde{x}_n - x_n$ is the error caused by not solving the nonlinear equations exactly. For a Runge–Kutta step starting at the exact solution $x(t)$, we let $\hat{X}_n$ denote the exact stage vector, satisfying

$$\big(\mathcal{A}^{-1} \otimes P - (I_s \otimes P) hF\big)(\hat{X}_n) = \mathcal{A}^{-1} \mathbb{1} \otimes P x(t_n) \tag{5.5a}$$

$$-(I_s \otimes Q) hF(\hat{X}_n) = 0. \tag{5.5b}$$

For the approximation $\tilde{x}_n$ and the exact solution $x(t)$ there now holds

$$\tilde{x}_{n+1} = R(\infty)\tilde{x}_n + (b^{\mathrm{T}}\mathcal{A}^{-1} \otimes I_d)\tilde{X}_n \tag{5.6a}$$

$$x(t_{n+1}) = R(\infty)x(t_n) + (b^{\mathrm{T}}\mathcal{A}^{-1} \otimes I_d)\hat{X}_n - l_n, \tag{5.6b}$$

where l_n represents the local error computed along the exact solution. We introduce the global error $e_n = \tilde{x}_n - x(t_n)$, and the error in the internal stages, $E_n = \tilde{X}_n - \hat{X}_n$. In projected form, the following global error propagation equation is then obtained,

$$Pe_{n+1} = R(\infty)Pe_n + (b^{\mathrm{T}}\mathcal{A}^{-1} \otimes P)E_n + Pl_n, \tag{5.7a}$$

$$Qe_{n+1} = R(\infty)Qe_n + (b^{\mathrm{T}}\mathcal{A}^{-1} \otimes Q)E_n + Ql_n, \tag{5.7b}$$

where the stage error E_n satisfies

$$(\mathcal{A}^{-1} \otimes P)E_n - (I_s \otimes P)(h\tilde{Y}_n - h\hat{Y}_n) = \mathcal{A}^{-1}\mathbb{1} \otimes Pe_n + (I_s \otimes P)hR_n \tag{5.8a}$$

$$-(I_s \otimes Q)(h\tilde{Y}_n - h\hat{Y}_n) = (I_s \otimes Q)hR_n, \tag{5.8b}$$

and where we have introduced $h\tilde{Y}_n - h\hat{Y}_n = hF(\hat{X}_n + E_n) - hF(\hat{X}_n)$.

We now need to extend the semi-inner product in $\mathbb{R}^m$ to one in $\mathbb{R}^{ms}$, see e.g. [4, p. 133]. Thus, given a positive diagonal matrix $D = \mathrm{diag}(d_1 \ldots, d_s)$ and the semi-inner product $(\cdot,\cdot)_+$ on $\mathbb{R}^m$, we can define a semi-inner product on $\mathbb{R}^{ms}$,

$$[\![U, V]\!]_{D+} = \sum_{j=1}^{s} d_i(u_j, v_j)_+ \tag{5.9}$$

retaining all the properties of Proposition 2.1. Moreover, this semi-inner product induces the norm $\|U\|_D^2 = [\![U, U]\!]_{D+} = \sum_{j=1}^{s} d_i\|u_i\|^2$. With this construction, Lemma 5.3.6 of [4, p. 145] applies, and we therefore obtain

$$[\![(I_s \otimes P)E_n\, ,\, (\mathcal{A}^{-1} \otimes P)E_n]\!]_{D+} \geq m_D[\mathcal{A}^{-1}] \cdot \|(I_s \otimes P)E_n\|_D^2, \tag{5.10}$$

where $m_D[\mathcal{A}^{-1}] = \min_{y\neq 0}(y^{\mathrm{T}}D\mathcal{A}^{-1}y)/(y^{\mathrm{T}}Dy)$ is the D-weighted Euclidean *glb* logarithmic norm on $\mathbb{R}^s$, i.e., $m_D[\mathcal{A}^{-1}] = -\mu_D[-\mathcal{A}^{-1}]$ for the usual *lub* logarithmic norm. Further, we also have

$$\begin{gathered}
[\![(I_s \otimes P)E_n\, ,\, (I_s \otimes P)\big(hF(\hat{X}_n + E_n) - hF(\hat{X}_n)\big)]\!]_{D+} \leq \\
[\![(I_s \otimes P)E_n\, ,\, (I_s \otimes P)\big(hF(\hat{X}_n + (I_s \otimes P)E_n) - hF(\hat{X}_n)\big)]\!]_{D+} + \\
+ [\![(I_s \otimes P)E_n\, ,\, (I_s \otimes P)\big(hF(\hat{X}_n + E_n) - hF(\hat{X}_n + (I_s \otimes P)E_n)\big)]\!]_{D+} \leq \\
hM_P^+[Pf] \cdot \|(I_s \otimes P)E_n\|_D^2 + hL_Q[Pf] \cdot \|(I_s \otimes Q)E_n\|_D \cdot \|(I_s \otimes P)E_n\|_D.
\end{gathered}$$

Thus we will obtain an inequality similar to (4.8),

$$\begin{bmatrix} m_D[\mathcal{A}^{-1}] - hM_P^+[Pf] & -hL_Q[Pf] \\ -L_P[Qf] & -M_Q^+[Qf] \end{bmatrix} \begin{bmatrix} \|(I_s \otimes P)E_n\|_D \\ \|(I_s \otimes Q)E_n\|_D \end{bmatrix} \leq$$

$$\leq \begin{bmatrix} \|\mathcal{A}^{-1}\mathbb{1} \otimes Pe_n\|_D + h\|(I_s \otimes P)R_n\|_D \\ \|(I_s \otimes Q)R_n\|_D \end{bmatrix}.$$

Hence if $M_Q^+[Qf] < 0$ and the Schur complement logarithmic Lipschitz constant fulfills $hS[P,f] < m_D[\mathcal{A}^{-1}]$, then the matrix of this inequality is an M–matrix with a positive inverse, and the inequality admits a bound on E_n in terms of e_n and R_n. In particular, the condition provides existence and uniqueness of solutions to (5.4) and (5.5).

For $\|(I_s \otimes P)E_n\|_D$ we obtain the bound

$$\|(I_s \otimes P)E_n\|_D \leq \frac{\|\mathcal{A}^{-1}\mathbb{1} \otimes Pe_n\|_D + h\|(I_s \otimes P)R_n\|_D - \frac{L_Q[Pf]}{M_Q^+[Qf]}\|(I_s \otimes Q)R_n\|_D}{m_D[\mathcal{A}^{-1}] - hS[P,f]},$$

which clearly shows that the Q- and P-projections of R_n enter in a different ways. As $h \to 0$ the influence of the P-projection diminishes, while the Q-projection remains constant. A corresponding bound for $\|(I_s \otimes Q)E_n\|_D$ can be obtained in a straightforward way. Both bounds depend on Pe_n, but not on Qe_n. Hence the global error recursion (5.7) shows that for methods with $R(\infty) = 0$, only the P projection of the global error, i.e. Pe_n, is propagated.

Global error bounds can in principle be obtained from the global error recursion, and stability can also be studied. However, in order to study contractivity in terms of B-stability, the semi-inner product must be replaced by a true inner product. The above approach remains valid and then provides a partitioned analysis of the error components and how they influence the method's error propagation.

6 Conclusions

The question of how to apply the notion of logarithmic norms to DAEs has long been an open question. In [10] the problem was solved for linear DAEs and matrix pencils, but that approach does not extend to nonlinear DAEs and/or PDAEs. In this paper we have shown that semi-inner products and the new notion of the restricted logarithmic Lipschitz constant can be used to cover nonlinear (P)DAEs of index 1; this analysis is made possible by relaxing some of the Lipschitz continuity requirements. This theory is, although complex, useful for deriving differential-algebraic majorants for error propagation in the continuous system, and corresponding, difference-algebraic inequalities may be derived to cover also discretizations, while linking the properties of the numerical problem to those of the continuous problem. The theory also reveals details about how perturbations in the algebraic equations affect the stability of the entire system in the nonlinear case and provides quantitative bounds of perturbation index type, i.e., insight into the problem's condition is gained.

REFERENCES

1. K.E. Brenan, S.L. Campbell and L.R. Petzold (1989). *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations.* New York: North-Holland.

2. G. Dahlquist (1959). *Stability and Error Bounds in the Numerical Solution of Ordinary Differential Equations*. Trans. of the Royal Inst. of Tech., No. 130.
3. K. Deimling (1985). *Nonlinear Functional Analysis*. Berlin: Springer-Verlag.
4. K. Dekker and J.G. Verwer (1984). *Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations*. New York: North Holland.
5. R. Frank, J. Schneid and C.W. Ueberhuber (1981). *The concept of B-convergence*. SIAM J. Num. Anal. 18, 753–780.
6. R. Frank, J. Schneid and C.W. Ueberhuber (1985). *Stability properties of implicit Runge-Kutta methods*. SIAM J. Num. Anal. 22, 497–514.
7. E. Hairer, Ch. Lubich and M. Roche (1989). *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Springer Lecture notes in mathemtaics.
8. E. Hairer, S.P. Nørsett and G. Wanner (1993). *Solving Ordinary Differential Equations I*, 2nd ed., Berlin: Springer-Verlag.
9. E. Hairer and G. Wanner (1996). *Solving Ordinary Differential Equations II*, 2nd ed., Berlin: Springer-Verlag.
10. I. Higueras and B. García–Celayeta (1997). *Logarithmic norms of matrix pencils*.
11. J.F.B.M. Kraaijevanger (1985). *B-convergence of the implicit midpoint rule and the trapezoidal rule*. BIT 25, 652–666.
12. J.M. Ortega and W.C. Rheinboldt (1970). *Iterative Solution of Nonlinear Equations in Several Variables*. New York: Academic Press.
13. J.C. Simo and T.J.R. Hughes (1999). *Computational inelasticity*. Berlin: Springer-Verlag
14. T. Ström (1975). *On logarithmic norms*. SIAM J. Num. Anal. 2, 741–753.
15. G. Söderlind (1986). *Bounds on Nonlinear Operators in Finite-dimensional Banach Spaces*. Num. Math. 50, 27–44.