ELSEVIER

Applied Numerical Mathematics 18 (1995) 57-67

APPLIED NUMERICAL MATHEMATICS

# Kreiss resolvent conditions and strengthened Cauchy-Schwarz inequalities

W. Auzinger*, G. Kirlinger

*Institut für Angewandte und Numerische Mathematik, Technische Universität Wien, Wiedner Hauptstrasse 8-10/1152, A-1040 Wien, Austria*

Received 1 December 1994; revised 7 March 1995; accepted 9 April 1995

**Abstract**

In the present paper it is shown how the resolvent conditions in the Kreiss Matrix Theorem for $e^{tA}$ and $A^\nu$, respectively, can be reformulated as certain Strengthened Cauchy-Schwarz Inequalities (SCSIs) to be satisfied for all pairs $w, Aw$ ($w \in \mathbb{C}^n$). This yields certain generalizations of the notions "logarithmic norm" and "numerical radius", respectively.

## 1. Introduction and background

The present paper is concerned with the resolvent conditions appearing in the Kreiss Matrix Theorem (cf. Theorems 1 and 2 below). In the original version of this theorem (cf. [9]), necessary and sufficient conditions were given for the uniform power-boundedness of a family of complex $(n \times n)$-matrices $A$. One of these conditions involves a certain bound for the norm of the resolvent $(\mu I - A)^{-1}$, for all $\mu \in \mathbb{C}$ with $|\mu| > 1$. In Section 2 we show how such a condition can be reformulated more directly as a condition on $A$, in form of an SCSI (Strengthened Cauchy-Schwarz Inequality) which has to be satisfied for all pairs $w, Aw$ ($w \in \mathbb{C}^n$). An analogous result for matrix exponentials is also presented. The proofs of these assertions are given in Section 3. In Section 4 our results are discussed and illustrated by an example.

In the following, $\mathcal{F}$ denotes a family of complex $(n \times n)$-matrices $A$. $\langle \cdot, \cdot \rangle$ denotes an inner product in $\mathbb{C}^n$, and $\|\cdot\|$ denotes the vector or matrix norm associated with it.

Kreiss' Theorem(s) for matrix exponentials respectively matrix powers may be formulated as follows:

* Corresponding author. E-mail: winfried@uranus.tuwien.ac.at.


*SSDI* 0168-9274(95)00043-7

**Theorem 1** (Kreiss Matrix Theorem (matrix exponential case)). *If the resolvent condition*

$$\|(\mu I - A)^{-1}\| \leqslant \frac{K}{\operatorname{Re}\mu - \theta} \qquad \forall \mu \in \mathbb{C}\colon \quad \operatorname{Re}\mu > \theta \tag{1.1}$$

*(with $\theta \in \mathbb{R}$) is uniformly satisfied for all $A \in \mathcal{F}$, then the estimate*

$$\|e^{tA}\| \leqslant enKe^{t\theta} \quad \forall t \geqslant 0 \tag{1.2}$$

*holds for all $A \in \mathcal{F}$.*

**Theorem 2** (Kreiss Matrix Theorem (matrix power case)). *If the resolvent condition*

$$\|(\mu I - A)^{-1}\| \leqslant \frac{K}{|\mu| - \theta} \qquad \forall \mu \in \mathbb{C}\colon \quad |\mu| > \theta \tag{1.3}$$

*(with $\theta > 0$) is uniformly satisfied for all $A \in \mathcal{F}$, then the estimate*

$$\|A^{\nu}\| \leqslant enK\theta^{\nu} \quad \forall \nu \geqslant 0 \tag{1.4}$$

*holds for all $A \in \mathcal{F}$.*

**Remarks.**

- Actually, we have only stated the "interesting" part of the theorem (sufficiency of the resolvent condition). It is not difficult to show that, conversely, the resolvent estimate (1.1) respectively (1.3) follows if (1.2) respectively (1.4) holds without a factor $en$ (cf., e.g., [10]).
- Originally, the theorem was formulated by Kreiss for the matrix power case (cf. [9]). Kreiss' bound for $\|A^{\nu}\|$ was weaker than in (1.4), with a highly nonlinear dependence on $n$.
- Subsequently, Kreiss' proof was polished and sharpened by several authors (cf. [10,13,14,16], and others). In [10], LeVeque and Trefethen refined an elegant argument based on Cauchy integrals due to Tadmor (cf. [16]) and showed that the bounds (1.2) and (1.4) are valid with a factor $2enK$ instead of $enK$. Their conjecture that the factor 2 may be omitted was proved by Spijker (cf. [14]). The resulting bound is optimal, as can be seen from counterexamples presented in [10] and [8].
- For various extensions of the Kreiss Matrix Theorem cf. e.g. [4,11,15] and references therein.

The above theorems include important special cases, namely those where a resolvent condition with $K = 1$ is satisfied. The following facts are easy to verify: [1]

- *Cf. Theorem* 1: The resolvent condition (1.1) with $K = 1$ is equivalent to

  $$\operatorname{Re}\langle Aw, w\rangle \leqslant \theta\,\langle w, w\rangle \quad \text{for all } w \in \mathbb{C}^n, \tag{1.5}$$

  i.e., the *logarithmic norm* of $A$ is $\leqslant \theta$. In this case, $\|e^{tA}\| \leqslant e^{t\theta}$ readily follows.

  If (1.5) holds with $\theta = 0$, $A$ is called dissipative. Actually, relation (1.1) with $K = 1$ and $\theta = 0$ is sometimes used to define dissipativity (cf. e.g. [7]).

  Eq. (1.5) is the basis of many stability investigations for ordinary and partial differential equations.

[1] Cf. for instance [7,13]. The case $K = 1$ is also included as a special case ("Case (i)") in the proofs given in Section 3 of the present paper.

- *Cf. Theorem* 2: The resolvent condition (1.3) with $K = 1$ is equivalent to

$$|\langle Aw, w\rangle| \leqslant \theta \langle w, w\rangle \quad \text{for all } w \in \mathbb{C}^n, \tag{1.6}$$

i.e., the *numerical radius* of $A$ is $\leqslant \theta$. In this case, it can be shown that $\|A^\nu\| \leqslant 2\theta^\nu$ (cf. e.g. [5]).

Numerical radius techniques have been successfully used in the stability analysis for discretizations of differential equations; cf. e.g. [5,13].

In contrast to the original resolvent conditions, the special relations (1.5) and (1.6) are formulated as conditions directly on $A$ (and not on a continuum of resolvents $(\mu I - A)^{-1}$). In Section 2 we show how the general resolvent conditions (1.1) and (1.3)—with arbitrary $K$—can be similarly interpreted as conditions on the behavior of $A$ applied to arbitrary $w \in \mathbb{C}^n$. This leads to certain SCSIs including (1.5) and (1.6) as special cases (cf. (2.2) and (2.4) below).

## 2. Results

Our reformulation of the Kreiss resolvent conditions reads as follows; see Section 3 for the proofs of these assertions.

**Proposition 3.** *Let $K$ and $\theta$ be real constants, with $K \geqslant 1$. Then the resolvent condition* (1.1) *in the Kreiss Matrix Theorem for* $e^{tA}$ *is equivalent to the following condition: For each $w \in \mathbb{C}^n$, with $Aw \neq \theta w$, the quantity*

$$\gamma(w) := \frac{\langle w, Aw - \theta w\rangle}{\|w\|\,\|Aw - \theta w\|} \tag{2.1}$$

*satisfies at least one of the inequalities* (2.2a) *and* (2.2b):

$$\operatorname{Re}\gamma(w) \leqslant 0, \tag{2.2a}$$

*or*

$$[\operatorname{Re}\gamma(w)]^2 \leqslant (1 - K^{-2})\left(1 - [\operatorname{Im}\gamma(w)]^2\right). \tag{2.2b}$$

Condition (2.2) means that for all $w \in \mathbb{C}^n$, $\gamma(w)$ must be contained in a moon-shaped domain (hatched area in Fig. 1), while the Cauchy-Schwarz inequality says that $\gamma(w)$ is always contained in the unit disc (full moon). In this sense, (2.2) constitutes an SCSI which has to be satisfied for all pairs $w, Aw - \theta w$.

**Proposition 4.** *Let $K$ and $\theta$ be real constants, with $K \geqslant 1$ and $\theta > 0$. Then the resolvent condition* (1.3) *in the Kreiss Matrix Theorem for $A^\nu$ is equivalent to the following condition: For each $w \in \mathbb{C}^n$, with $w \neq 0$, the quantities*

$$\rho(w) := \left|\frac{\langle w, \frac{1}{\theta}Aw\rangle}{\langle w, w\rangle}\right|, \qquad \sigma(w) := \frac{\|\frac{1}{\theta}Aw\|}{\|w\|} \tag{2.3}$$

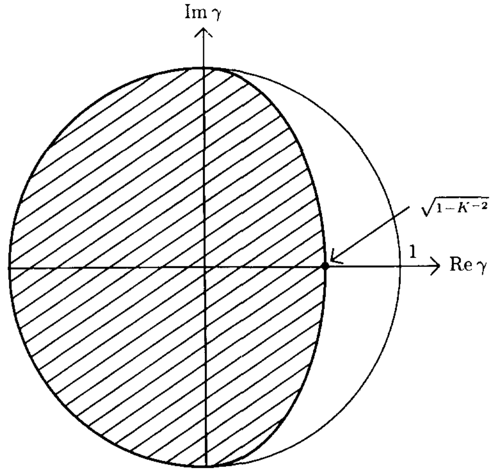


Fig. 1. Proposition 3 visualized.

*satisfy at least one of the inequalities* (2.4a) *and* (2.4b):

$$\rho(w) \leqslant 1, \tag{2.4a}$$

*or*

$$[\rho(w)-1]^2 \leqslant (1-K^{-2})\left(\sigma^2(w)-2\rho(w)+1\right). \tag{2.4b}$$

Similarly as for Proposition 3, condition (2.4) can be interpreted as an SCSI. For a visualization it is convenient to consider

$$\tilde{\rho}(w) := \frac{\rho(w)}{\sigma(w)} = \frac{|\langle w, Aw\rangle|}{\|w\|\,\|Aw\|}, \qquad \tilde{\sigma}(w) := \frac{1}{\sigma(w)} = \theta\frac{\|w\|}{\|Aw\|}. \tag{2.5}$$

In these terms, (2.4) is easily verified to be equivalent to the following condition: Unless $\tilde{\sigma}(w) \geqslant 1$, i.e., unless $\|Aw\| \leqslant \theta\,\|w\|$, $\tilde{\rho}(w)$ must lie below or on the elliptic arc

$$\tilde{\rho} = K^{-2}\tilde{\sigma} + \sqrt{1-K^{-2}}\sqrt{1-K^{-2}\tilde{\sigma}^2} \tag{2.6}$$

in the $(\tilde{\sigma}, \tilde{\rho})$-unit square (hatched area in Fig. 2).[2]

Note that the special relations (1.5) and (1.6) mentioned in Section 1 correspond to the limiting case $K = 1$, which means that condition (2.2a) respectively (2.4a) must hold for all $w$, (2.2b) respectively (2.4b) being immaterial. Thus, Propositions 1 and 2 are natural generalizations of (1.5) respectively (1.6) in form of certain SCSIs equivalent to the original resolvent conditions.

## 3. Proofs

**Proof of Proposition 3** (*Matrix exponential case*). The Kreiss resolvent condition (1.1) (cf. Theorem 1) is equivalent to the requirement that

[2] Note that $\rho \leqslant 1 \iff \tilde{\rho} \leqslant \tilde{\sigma}$.

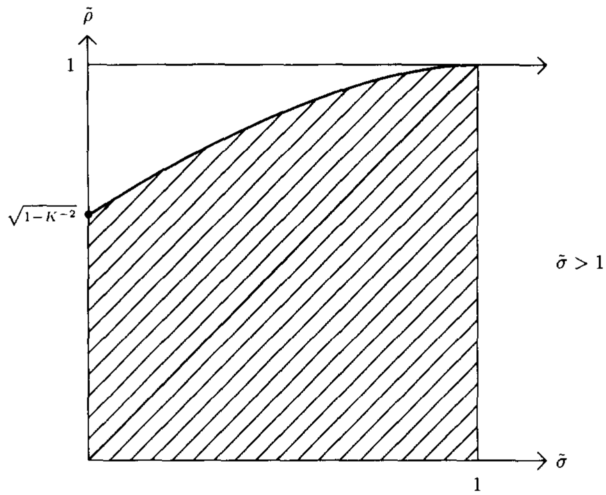


Fig. 2. Proposition 4 visualized.

$$\|(\mu-\theta)w-(Aw-\theta w)\|^2 \geqslant \frac{[\mathrm{Re}(\mu-\theta)]^2}{K^2}\|w\|^2 \tag{3.1}$$

has to be satisfied for all $w \in \mathbb{C}^n$ and for all $\mu \in \mathbb{C}$ with $\mathrm{Re}(\mu-\theta) > 0$ $(\theta \in \mathbb{R})$. For a more convenient formulation of this condition we denote

$$Aw-\theta w =: \hat{w}, \qquad \mu-\theta =: \xi + \mathrm{i}\eta, \tag{3.2}$$

and write the left-hand side of (3.1) as

$$\begin{aligned} &\|(\mu-\theta)w-(Aw-\theta w)\|^2 \\ &\quad = |\mu-\theta|^2\|w\|^2 - 2\,\mathrm{Re}[(\mu-\theta)\langle w,\hat{w}\rangle] + \|\hat{w}\|^2 \\ &\quad = (\xi^2+\eta^2)\|w\|^2 - 2\xi\,\mathrm{Re}\langle w,\hat{w}\rangle + 2\eta\,\mathrm{Im}\langle w,\hat{w}\rangle + \|\hat{w}\|^2. \end{aligned} \tag{3.3}$$

Thus, resolvent condition (1.1) is equivalent to

$$\begin{aligned} &(\xi^2+\eta^2)\|w\|^2 - 2\xi\,\mathrm{Re}\langle w,\hat{w}\rangle + 2\eta\,\mathrm{Im}\langle w,\hat{w}\rangle + \|\hat{w}\|^2 \geqslant \xi^2K^{-2}\|w\|^2 \\ &\quad \text{for all } w\in\mathbb{C}^n \text{ and for all } \xi>0,\ \eta\in\mathbb{R}. \end{aligned} \tag{3.4}$$

For the following we note that

$$[\mathrm{Re}\langle w,\hat{w}\rangle]^2 + [\mathrm{Im}\langle w,\hat{w}\rangle]^2 \leqslant \|w\|^2\|\hat{w}\|^2 \tag{3.5}$$

holds (Cauchy–Schwarz inequality).

Now we investigate condition (3.4) for arbitrary but fixed $w \in \mathbb{C}^n$. Concerning $w$, we consider two cases:

*Case* (i): $\mathrm{Re}\langle w,\hat{w}\rangle \leqslant 0$. In this case we have $-2\xi\,\mathrm{Re}\langle w,\hat{w}\rangle \geqslant 0$ for all $\xi > 0$. Bounding $2\eta\,\mathrm{Im}\langle w,\hat{w}\rangle$ $(\eta\in\mathbb{R})$ from below using the inequality $|\mathrm{Im}\langle w,\hat{w}\rangle| \leqslant \|w\|\,\|\hat{w}\|$ (cf. (3.5)), we thus can estimate the left-hand side of (3.4) in the following way:

$$\begin{aligned} &(\xi^2+\eta^2)\|w\|^2 - 2\xi\,\mathrm{Re}\langle w,\hat{w}\rangle + 2\eta\,\mathrm{Im}\langle w,\hat{w}\rangle + \|\hat{w}\|^2 \\ &\quad \geqslant (\xi^2+\eta^2)\|w\|^2 - 2|\eta|\,\|w\|\,\|\hat{w}\| + \|\hat{w}\|^2 \\ &\quad = \xi^2\|w\|^2 + [\|\hat{w}\| - |\eta|\,\|w\|]^2 \ \geqslant\ \xi^2\|w\|^2 \end{aligned} \tag{3.6}$$

for all $\xi > 0$ and $\eta \in \mathbb{R}$.

Hence for case (i), condition (3.4) is satisfied with arbitrary $K \geqslant 1$.

*Case* (ii): $\mathrm{Re}\langle w, \hat{w}\rangle > 0$. In this case we denote

$$\alpha := 1/\xi, \qquad \beta := \eta/\xi, \tag{3.7}$$

and multiply (3.4) by $\alpha^2 = 1/\xi^2$. This takes condition (3.4) into the equivalent form

$$\varphi(\alpha, \beta) \geqslant K^{-2}\,\|w\|^2 \quad \text{for all } w \in \mathbb{C}^n \text{ and for all } \alpha > 0,\ \beta \in \mathbb{R}, \tag{3.8}$$

where $\varphi(\alpha, \beta)$ is the quadratic form defined by

$$\varphi(\alpha, \beta) := \alpha^2\,\|\hat{w}\|^2 - 2\alpha\,\mathrm{Re}\langle w, \hat{w}\rangle + 2\alpha\beta\,\mathrm{Im}\langle w, \hat{w}\rangle + (1+\beta^2)\,\|w\|^2. \tag{3.9}$$

In order to minimize $\varphi(\alpha, \beta)$ we consider the linear system

$$\begin{aligned} \frac{1}{2}\frac{\partial\varphi}{\partial\alpha} &= \alpha\,\|\hat{w}\|^2 + \beta\,\mathrm{Im}\langle w, \hat{w}\rangle - \mathrm{Re}\langle w, \hat{w}\rangle = 0,\\ \frac{1}{2}\frac{\partial\varphi}{\partial\beta} &= \alpha\,\mathrm{Im}\langle w, \hat{w}\rangle + \beta\,\|w\|^2 = 0. \end{aligned} \tag{3.10}$$

Its determinant $D$ satisfies

$$D = \|w\|^2\,\|\hat{w}\|^2 - [\mathrm{Im}\langle w, \hat{w}\rangle]^2 \geqslant [\mathrm{Re}\langle w, \hat{w}\rangle]^2 > 0 \tag{3.11}$$

due to (3.5) and by assumption $\mathrm{Re}\langle w, \hat{w}\rangle > 0$, and the system is positive definite, i.e., the Hessian of $\varphi$ is positive definite. Hence the unique solution of (3.10),

$$\begin{aligned} \alpha_{\min} &= \frac{1}{D}\,\|w\|^2\,\mathrm{Re}\langle w, \hat{w}\rangle > 0,\\ \beta_{\min} &= -\frac{1}{D}\,\mathrm{Re}\langle w, \hat{w}\rangle\,\mathrm{Im}\langle w, \hat{w}\rangle, \end{aligned} \tag{3.12}$$

minimizes $\varphi(\alpha, \beta)$, with $\alpha_{\min} > 0$ as required in (3.8). A simple calculation shows that the minimal value $\varphi(\alpha_{\min}, \beta_{\min})$ evaluates to

$$\varphi(\alpha_{\min}, \beta_{\min}) = \left[1 - \frac{[\mathrm{Re}\langle w, \hat{w}\rangle]^2}{D}\right]\,\|w\|^2. \tag{3.13}$$

Hence for case (ii), condition (3.4) is equivalent to (cf. (3.8))

$$1 - \frac{[\mathrm{Re}\langle w, \hat{w}\rangle]^2}{\|w\|^2\,\|\hat{w}\|^2 - [\mathrm{Im}\langle w, \hat{w}\rangle]^2} \geqslant K^{-2}, \tag{3.14}$$

which can be rewritten as

$$[\mathrm{Re}\langle w, \hat{w}\rangle]^2 \leqslant (1 - K^{-2})\left(\|w\|^2\,\|\hat{w}\|^2 - [\mathrm{Im}\langle w, \hat{w}\rangle]^2\right). \tag{3.15}$$

Summarizing, we see that the Kreiss resolvent condition is equivalent to (2.2): With $\gamma(w)$ as defined in (2.1), (2.2a) corresponds to case (i) above, while for the alternative case (ii), (2.2b) is nothing but (3.15). □

**Proof of Proposition 4** (*Matrix power case*). The Kreiss resolvent condition (1.3) (cf. Theorem 2) is equivalent to the requirement that

$$\|\mu w - Aw\|^2 \geqslant \frac{(|\mu| - \theta)^2}{K^2} \|w\|^2 \tag{3.16}$$

has to be satisfied for all $w \in \mathbb{C}^n$ and for all $\mu \in \mathbb{C}$ with $|\mu| - \theta > 0$ $(\theta > 0)$. For a more convenient formulation of this condition we denote

$$\begin{aligned} &\frac{1}{\theta} Aw =: \hat{w}, \\ &\mu =: \varrho\, \mathrm{e}^{\mathrm{i}\psi}, \end{aligned} \tag{3.17}$$

and write the left-hand side of (3.16) as

$$\begin{aligned} \|\mu w - Aw\|^2 &= |\mu|^2 \|w\|^2 - 2\theta \operatorname{Re}(\mu\langle w, \hat{w}\rangle) + \theta^2 \|\hat{w}\|^2 \\ &= \varrho^2 \|w\|^2 - 2\varrho\theta \operatorname{Re}(\mathrm{e}^{\mathrm{i}\psi}\langle w, \hat{w}\rangle) + \theta^2 \|\hat{w}\|^2. \end{aligned} \tag{3.18}$$

Thus, resolvent condition (1.3) is equivalent to

$$\begin{aligned} &\varrho^2 \|w\|^2 - 2\varrho\theta \operatorname{Re}(\mathrm{e}^{\mathrm{i}\psi}\langle w, \hat{w}\rangle) + \theta^2 \|\hat{w}\|^2 \geqslant (\varrho - \theta)^2 K^{-2} \|w\|^2 \\ &\quad \text{for all } w \in \mathbb{C}^n \text{ and for all } \varrho > \theta,\ \psi \in (-\pi, \pi]. \end{aligned} \tag{3.19}$$

The factor $(\varrho - \theta)^2 K^{-2}$ on the right-hand side of (3.19) does not depend on the parameter $\psi$, and minimization with respect to $\psi$ of the left-hand side is straightforward: We have

$$-\operatorname{Re}(\mathrm{e}^{\mathrm{i}\psi}\langle w, \hat{w}\rangle) \geqslant -|\langle w, \hat{w}\rangle|, \tag{3.20}$$

with equality for $\psi = -\operatorname{Arg}\langle w, \hat{w}\rangle$. Using (3.20), we obtain another equivalent formulation of the resolvent condition (3.16):

$$\begin{aligned} &\varrho^2 \|w\|^2 - 2\varrho\theta |\langle w, \hat{w}\rangle| + \theta^2 \|\hat{w}\|^2 \geqslant (\varrho - \theta)^2 K^{-2} \|w\|^2 \\ &\text{for all } w \in \mathbb{C}^n \text{ and for all } \varrho > \theta. \end{aligned} \tag{3.21}$$

For the following we note that

$$|\langle w, \hat{w}\rangle| \leqslant \|w\|\, \|\hat{w}\| \tag{3.22}$$

holds (Cauchy-Schwarz inequality).

Now we investigate condition (3.21) for arbitrary but fixed $w \in \mathbb{C}^n$. Concerning $w$, we consider two cases:

*Case* (i): $|\langle w, \hat{w}\rangle| \leqslant \|w\|^2$. In this case we use (3.22) to estimate the left-hand side of (3.21) in the following way:

$$\begin{aligned} \varrho^2 \|w\|^2 - 2\varrho\theta |\langle w, \hat{w}\rangle| + \theta^2 \|\hat{w}\|^2 &\geqslant \varrho^2 \|w\|^2 - 2\varrho\theta \frac{|\langle w, \hat{w}\rangle|}{\|w\|} \|w\| + \theta^2 \left(\frac{|\langle w, \hat{w}\rangle|}{\|w\|}\right)^2 \\ &= \left(\varrho - \theta \frac{|\langle w, \hat{w}\rangle|}{\|w\|^2}\right)^2 \|w\|^2 \geqslant (\varrho - \theta)^2 \|w\|^2 \end{aligned} \tag{3.23}$$

for all $\varrho > \theta$, due to assumption $|\langle w, \hat{w}\rangle| / \|w\|^2 \leqslant 1$.

Hence for case (i), condition (3.21) is satisfied with arbitrary $K \geqslant 1$.

*Case* (ii): $|\langle w, \hat{w}\rangle| > \|w\|^2$. In this case we denote

$$\delta := \frac{\theta}{\varrho - \theta} \quad \left(\frac{\varrho}{\varrho - \theta} = 1 + \delta\right) \tag{3.24}$$

and multiply (3.21) by $\delta^2 = (\theta/(\varrho - \theta))^2$. This takes condition (3.21) into the equivalent form

$$\varphi(\delta) \geqslant K^{-2}\|w\|^2 \quad \text{for all } w \in \mathbb{C}^n \text{ and for all } \delta > 0, \tag{3.25}$$

where $\varphi(\delta)$ is the quadratic function defined by

$$\varphi(\delta) := \delta^2\|\hat{w}\|^2 - 2\delta(1+\delta)\,|\langle w, \hat{w}\rangle| + (1+\delta)^2\|w\|^2. \tag{3.26}$$

In order to minimize $\varphi(\delta)$ we consider the linear equation

$$\frac{1}{2}\frac{\mathrm{d}\varphi}{\mathrm{d}\delta} = \delta\|\hat{w}\|^2 - (1+2\delta)\,|\langle w, \hat{w}\rangle| + (1+\delta)\,\|w\|^2 = 0. \tag{3.27}$$

The second derivative of $\varphi(\delta)$ satisfies

$$\frac{1}{2}\frac{\mathrm{d}^2\varphi}{\mathrm{d}^2\delta} = \|\hat{w}\|^2 - 2\,|\langle w, \hat{w}\rangle| + \|w\|^2 \geqslant (\|\hat{w}\| - \|w\|)^2 > 0 \tag{3.28}$$

due to (3.22) and by assumption $|\langle w, \hat{w}\rangle| > \|w\|^2$, implying $\|\hat{w}\| > \|w\|$. Hence the unique solution of (3.27),

$$\delta_{\min} = \frac{|\langle w, \hat{w}\rangle| - \|w\|^2}{\|\hat{w}\|^2 - 2\,|\langle w, \hat{w}\rangle| + \|w\|^2}, \tag{3.29}$$

minimizes $\varphi(\delta)$, with $\delta_{\min} > 0$ as required in (3.25). A simple calculation shows that the minimal value $\varphi(\delta_{\min})$ evaluates to

$$\varphi(\delta_{\min}) = \left[1 - \frac{\frac{1}{\|w\|^2}\left[\langle w, \hat{w}\rangle - \|w\|^2\right]^2}{\|\hat{w}\|^2 - 2\,|\langle w, \hat{w}\rangle| + \|w\|^2}\right]\|w\|^2. \tag{3.30}$$

Hence for case (ii), condition (3.21) is equivalent to (cf. (3.25))

$$1 - \frac{\frac{1}{\|w\|^2}\left[\langle w, \hat{w}\rangle - \|w\|^2\right]^2}{\|\hat{w}\|^2 - 2\,|\langle w, \hat{w}\rangle| + \|w\|^2} \geqslant K^{-2}, \tag{3.31}$$

which can be rewritten as

$$\left[\frac{|\langle w, \hat{w}\rangle|}{\|w\|^2} - 1\right]^2 \leqslant (1 - K^{-2})\left(\left(\frac{\|\hat{w}\|}{\|w\|}\right)^2 - 2\frac{|\langle w, \hat{w}\rangle|}{\|w\|^2} + 1\right). \tag{3.32}$$

Summarizing, we see that the Kreiss resolvent condition is equivalent to (2.4): With $\rho(w)$, $\sigma(w)$ as defined in (2.3), (2.4a) corresponds to case (i) above, while for the alternative case (ii), (2.4b) is nothing but (3.32). □

## 4. Discussion and example

Although natural and not difficult to prove, our interpretation of resolvent conditions as SCSIs seems not to have appeared in the literature so far. Therefore the main motivation for the present paper was simply to make it known. Concerning its usefulness for theoretical or practical purposes we have no definitive opinion at the moment. Let us mention some aspects:

- The SCSI formulation has some advantages with regard to numerical examinations. Plotting the distribution of the relevant parameters (in the spirit of Fig. 1 or Fig. 2) using an appropriate set of samples $w \in \mathbb{C}^n$, $\|w\| = 1$, may be more practical than plotting the distribution of resolvent norms. An example is presented below (cf. Fig. 3).
- Its usefulness for theoretical purposes, e.g. for concrete stability investigations, is hard to judge. For the matrix power case, the simplest condition involving numerical radii only has been successfully applied for such purposes; cf. e.g. [5], where the stability of the Lax–Wendroff scheme is proved by investigating the numerical radius of its amplification matrix. The general SCSI (2.4), however, has a more complicated structure, and the extension of numerical radius techniques is not straightforward.
- Potentially, the Kreiss Matrix Theorem is a powerful tool in the stability theory of linear numerical processes, especially in highly non-normal situations where more elementary techniques are too weak or may lead to misleading results. However, applications to concrete problems are rare in the literature, probably due to the fact that the conditions involved are technically difficult to handle. The same is true for the SCSI formulation.
- For an example of a concrete application of the Kreiss Matrix Theorem cf. e.g. [6] where, following [2], the stability of multistep schemes applied to linear stiff IVPs is studied on the basis of a resolvent estimate; cf. Lemma 7.3 in [6]. Note, however, that the corresponding stability constant is not quantitatively specified.

Due to the above reasons, it is not clear at the moment how relevant the SCSI formulation is for the applicability of techniques related to the Kreiss Theorem. This is certainly an interesting question to be further discussed.

Many other types of resolvent conditions play a role in linear stability theory; cf. e.g. [11,12,15]. Corresponding SCSI interpretations are certainly possible and should be studied together with its theoretical respectively practical usefulness.

**Example.** For a visualization of the SCSI condition in a concrete, simple example we consider the two-step BDF scheme, with fixed stepsize $h$, applied to the stiff model problem $y' = \lambda y$ ($\operatorname{Re}\lambda \leqslant 0$). Its companion matrix is

$$C(\zeta) = \begin{pmatrix} \dfrac{4}{3-2\zeta} & \dfrac{-1}{3-2\zeta} \\ 1 & 0 \end{pmatrix} \tag{4.1}$$

($\zeta = h\lambda$). The two-step BDF scheme is $A$-stable, and therefore Lemma 7.3 in [6] is applicable and implies that there exists a constant $M$ such that $\|C(\zeta)^\nu\| \leqslant M$ for all $\nu$, uniformly with respect to $\zeta$ in the left complex halfplane.

An appropriate numerical sampling of the corresponding quantities $\tilde{\rho}(w;\zeta)$ and $\tilde{\sigma}(w;\zeta)$ (cf. (2.5), with $\theta = 1$) with respect to $w \in \mathbb{C}^2$ and $\operatorname{Re}\zeta \leqslant 0$ yields the plot shown in Fig. 3, in the spirit of

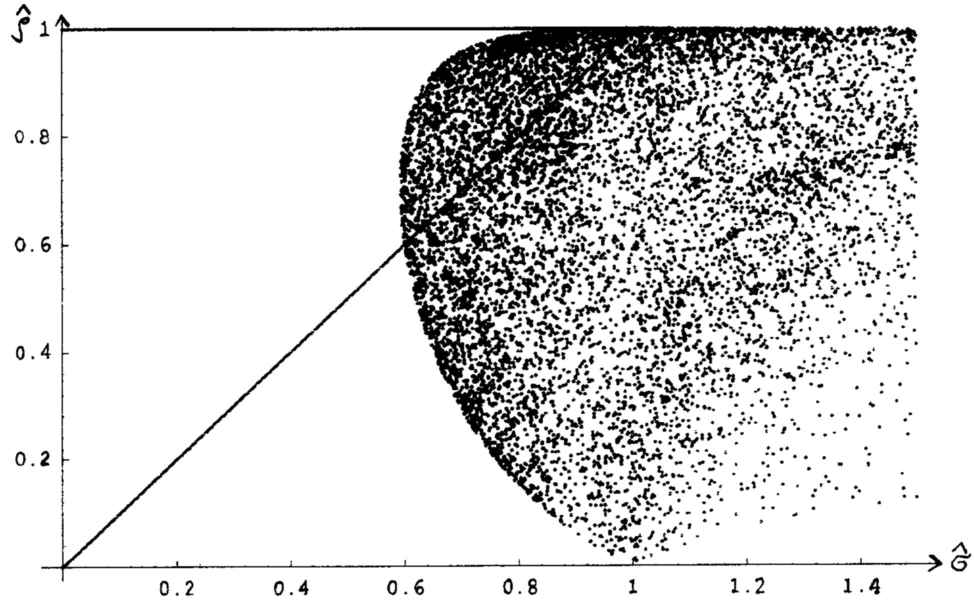


Fig. 3. Stability portrait for the two-step BDF scheme applied to $y' = \lambda y$.

Fig. 2. This yields the numerical estimate $\max_{\mathrm{Re}\,\zeta\leqslant 0} \|C(\zeta)\| \approx 1/0.6 \approx 1.7$ ($\|\cdot\|$ = Euclidean norm). The global stability constant $M$ (such that $\|C(\zeta)^\nu\| \leqslant M$) is harder to estimate directly from this picture. A reasonable guess seems to be $\sqrt{1-K^{-2}} \approx 0.75$, hence $K \approx 1.4$. Together with (1.4) this leads us to the estimate $M \leqslant \mathrm{e}\,2K \approx 7.6$.

Note that another estimate for $\|C(\zeta)^\nu\|$ can be derived using A-stability: For the two-step BDF scheme, A-stability is equivalent to its G-stability (cf. [6]). The latter implies $\|GC(\zeta)G^{-1}\| \leqslant 1$ whenever $\mathrm{Re}\,\zeta \leqslant 0$, with the matrix $G$ as given in [6, Section V.6]:

$$G = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}. \tag{4.2}$$

For the Euclidean norm we obtain $\|C(\zeta)^\nu\| \leqslant \|G\|\,\|G^{-1}\|$ for all $\mathrm{Re}\,\zeta \leqslant 0$, where $\|G\|\,\|G^{-1}\| = (3+2\sqrt{2})/(3-2\sqrt{2}) \approx 34$. This bound is rigorous but appears to be quite unrealistic in view of the above numerical estimate.

Let us conclude with a remark on nonlinear problems. As mentioned in Section 1, estimates for $\|\mathrm{e}^{tA}\|$ based on the concept of logarithmic norms are a special case of the Kreiss estimate ($K = 1$ in Theorem 1). On the other hand, the concept of logarithmic norms can easily be transplanted to nonlinear problems. This leads to the notion of one-sided Lipschitz constants, which has been extensively used in the analysis of nonlinear stiff differential equations (cf. e.g. [3,6]). For problems with real data, it is easy to see that a complex embedding is not necessary in this context, and the one-sided Lipschitz constant $m$ has a straightforward geometrical interpretation. It is, however, important to note that for stiff IVPs the usefulness of this approach heavily relies on the existence of a *moderate* $m$, which means that the problem is locally well-conditioned throughout. As has been shown in [1], however, the assumption of a moderate $m$ is actually extremely restrictive because—unless in very special cases (normal Jacobian)—a stiff problem is locally very ill-conditioned in general—despite its good global condition. (Cf. [1] for a more detailed discussion.)

In view of this observation, the question arises whether general resolvent conditions ($K > 1$) respectively their SCSI equivalents can be adapted for the analysis of certain nonlinear problems to overcome this drawback. This appears not to be straightforward because the relevant parameters in the SCSI conditions have no useful geometrical meaning in a real context which would enable such a transplantation. In other words: Stability estimates based on resolvent calculus are deep and elegant, but essentially linear. Nevertheless, they also can serve as an important *tool* for the analysis of nonlinear problems.

## References

[1] W. Auzinger, R. Frank and G. Kirlinger, A note on convergence concepts for stiff problems, *Computing* 44 (1990) 197–208.

[2] G. Dahlquist, H. Mingyou and R.J. LeVeque, On the uniform power-boundedness of a family of matrices and the applications to one-leg and linear multistep methods, *Numer. Math.* 42 (1983) 1–13.

[3] K. Dekker and J.G. Verwer, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations* (North-Holland, Amsterdam, 1984).

[4] J.L.M. van Dorsselaer, J.F.B.M. Kraaijevanger and M.N. Spijker, Linear stability analysis in the numerical solution of initial value problems, *Acta Numer.* (1993) 199–237.

[5] M. Goldberg and E. Tadmor, On the numerical radius and its applications, *Linear Algebra Appl.* 42 (1982) 263–284.

[6] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems* (Springer-Verlag, Berlin, 1991).

[7] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1982).

[8] J.F.B.M. Kraaijevanger, Two counterexamples related to the Kreiss matrix theorem, *BIT* 34 (1994) 113–119.

[9] H.O. Kreiss, Über die Stabilitätsdefinition für Differenzengleichungen die Differentialgleichungen approximieren, *BIT* 2 (1962) 153–181.

[10] R.J. LeVeque and L.N. Trefethen, On the resolvent condition in the Kreiss matrix theorem, *BIT* 24 (1984) 584–591.

[11] C. Lubich and O. Nevanlinna, On resolvent conditions and stability estimates, *BIT* 31 (1991) 293–313.

[12] S.C. Reddy and L.N. Trefethen, Stability of the method of lines, *Numer. Math.* 62 (1992) 235–267.

[13] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial-Value Problems* (Wiley, New York, 2nd ed., 1967).

[14] M.N. Spijker, On a conjecture by LeVeque and Trefethen related to the Kreiss matrix theorem, *BIT* 31 (1991) 551–555.

[15] M.N. Spijker and F.A.J. Straetemans, Stability estimates for families of matrices of nonuniformly bounded order, Report, University of Leiden, Leiden, Netherlands (1993); also: *Linear Algebra Appl.* (to appear).

[16] E. Tadmor, The equivalence of $L_2$-stability, the resolvent condition, and strict $H$-stability, *Linear Algebra Appl.* 41 (1981) 151–159.