Rend. Istit. Mat. Univ. Trieste
Vol. 53 (2021), Art. No. 3, pp. 1–14 (electronic)
DOI: 10.13137/2464-8728/31216

Propagation of perturbations in the initial value along solutions of linear ODEs: a componentwise relative error analysis

Asma Farooq and Stefano Maset

Abstract. *This paper addresses how a perturbation in the initial value propagates along the solution of an n-dimensional linear ordinary differential equation, by considering the relative errors in the components of the solution. We are particularly interested in the long-time behavior of this propagation.*

Keywords: Linear ordinary differential equation, conditioning, componentwise relative errors.
MS Classification 2020: 65F35, 65F60, 34D10, 34D20.

1. Introduction

Consider the n-dimensional linear Ordinary Differential Equation (ODE)

$$\begin{cases} y'(t) = Ay(t), \ t \geq 0, \\ y(0) = y_0, \end{cases} \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $y(t), y_0 \in \mathbb{R}^n$. Suppose that the initial value y_0 is perturbed to $\widetilde{y}_0$ and, as a consequence, the solution y is perturbed to $\widetilde{y}$. In this paper, we are interested in the componentwise relative errors of $\widetilde{y}$ with respect to y. In particular, we are interested in studying how in the long-time the perturbation in y_0 propagates componentwise along y. In other words, for $l = 1, \ldots, n$, we study the (relative) conditioning of the problem

$$y_0 \mapsto y_l(t) = e_l^T \mathrm{e}^{tA} y_0, \tag{2}$$

where e_l^T is the l-th vector of the canonical basis of $\mathbb{R}^n$, with emphasis on the asymptotic behavior as $t \to +\infty$.

The conditioning of the problem $A \mapsto \mathrm{e}^{tA}$ has been widely studied in literature: see [9], [8], [10], [12], [1], [14], and [5]. Less attention has received the conditioning of the action of the matrix exponential e^{tA} on a vector. The conditioning of the problem $(A, y_0) \mapsto \mathrm{e}^{tA} y_0$ was considered in [2] and [4], but

these papers were interested in computational aspects. A qualitative analysis of the conditioning of the problem $y_0 \mapsto \mathrm{e}^{tA}y_0$, for A normal, was accomplished in [11], where it was studied how this conditioning depends on t.

The present paper provides a qualitative analysis of the conditioning of the problem (2).

2. The condition numbers

Suppose that the initial value $y_0 \neq 0$ of the ODE (1) is perturbed to $\widetilde{y}_0$. Due to this perturbation, the solution $y(t) = \mathrm{e}^{tA}y_0$, $t \geq 0$, of (1) is perturbed to $\widetilde{y}(t) = \mathrm{e}^{tA}\widetilde{y}_0$, $t \geq 0$. Let

$$\varepsilon = \frac{\|\widetilde{y}_0 - y_0\|}{\|y_0\|},$$

be the normwise relative error on y_0, where $\|\cdot\|$ is a norm on $\mathbb{R}^n$. For $l = 1, \ldots, n$, let

$$\delta_l(t) = \frac{|\widetilde{y}_l(t) - y_l(t)|}{|y_l(t)|},\ t \geq 0,$$

be the relative error on the l-th component of y. We study how the error $\delta_l(t)$ is related to the error ε.

Remark 2.1:

1. Since $y_0 \neq 0$, the error ε is well-defined. On the other hand, it could happen to have $y_l(t_1) = 0$ for some $t_1 \geq 0$. In this case, we could consider $\delta_l(t_1)$ not defined, or equal to $+\infty$ and indeterminate for $\widetilde{y}_l(t_1) \neq 0$ and $\widetilde{y}_l(t_1) = 0$, respectively. However, we are not interested to $\delta_l(t)$ for $t = t_1$, because when $y_l(t)$ becomes zero the absolute error is more important than the relative error.

2. The normwise relative error ε on y_0 and the componentwise relative errors on y_0 are related, as well as the normwise relative error on y and the componentwise relative errors $\delta_l(t)$, $l = 1, \ldots, n$, on y. In fact, if $\|\cdot\|$ is a p-norm, then

$$\varepsilon \leq \max_{l=1,\ldots,n} \frac{|\widetilde{y}_{0l} - y_{0l}|}{|y_{0l}|} \quad \text{and} \quad \frac{\|\widetilde{y}(t) - y(t)\|}{\|y(t)\|} \leq \max_{l=1,\ldots,n} \delta_l(t),\ t \geq 0.$$

By writing the perturbation in y_0 as $\widetilde{y}_0 = y_0 + \varepsilon \|y_0\| \widehat{z}_0$, where $\widehat{z}_0 \in \mathbb{R}^n$ is a unit vector, i.e. $\|\widehat{z}_0\| = 1$, we obtain, for $l = 1, \ldots, n$,

$$\delta_l(t) = K_l(t, A, y_0, \widehat{z}_0)\varepsilon,\ t \geq 0, \tag{3}$$

where

$$K_l(t, A, y_0, \widehat{z}_0) = \frac{|e_l^T \mathrm{e}^{tA}\widehat{z}_0|}{|e_l^T \mathrm{e}^{tA}\widehat{y}_0|} \tag{4}$$

with $\widehat{y}_0 = \frac{y_0}{\|y_0\|}$. We define $K_l(t, A, y_0, \widehat{z}_0)$ as the *condition number with direction of perturbation* $\widehat{z}_0$ of the problem (2). The formula (3) is of theoretical interest. From a practical point of view, when there is no information about the direction of perturbation, we can write

$$\delta_l(t) \leq K_l(t, A, y_0)\varepsilon, \ t \geq 0,$$

where

$$K_l\left(t, A, y_0\right) := \sup_{\substack{\widehat{z}_0 \in \mathbb{R}^n \\ \|\widehat{z}_0\|=1}} K_l\left(t, A, y_0, \widehat{z}_0\right) = \frac{\left\|e_l^T \mathrm{e}^{tA}\right\|}{|e_l^T \mathrm{e}^{tA}\widehat{y}_0|}, \tag{5}$$

with $\left\|e_l^T \mathrm{e}^{tA}\right\|$ the matrix norm of the row vector $e_l^T \mathrm{e}^{tA}$ relevant to the vector norm $\|\cdot\|$. We define $K_l\left(t, A, y_0\right)$ as the *condition number* of the problem (2) (see [3] for the definition of the condition number of a general problem).

In the next section, we analyze the conditions numbers (4) and (5) by assuming that A is diagonalizable. This is a generic situation for the matrix A. Here and in the following, a generic situation for A, or y_0 or $\widehat{z}_0$, means that A satisfies a property which is generic according to the measure theory definition (the complementary property corresponds to a zero measure subset) or the topological definition (the property corresponds to a dense open subset) given for example in [7]. Roughly speaking, a generic situation considers "typical", not "exceptional", cases.

3. Analysis of the condition numbers

Let A be diagonalizable. We partition the spectrum $\Lambda = \{\lambda_1, \ldots, \lambda_p\}$ of A, where $\lambda_1, \ldots, \lambda_p$ are the distinct complex eigenvalues of A, by decreasing real parts into the subsets

$$\Lambda_j := \{\lambda_{i_{j-1}+1}, \lambda_{i_{j-1}+2}, \ldots, \lambda_{i_j}\}, \ j = 1, \ldots, q,$$

where

$$\mathrm{Re}(\lambda_{i_{j-1}+1}) = \mathrm{Re}(\lambda_{i_{j-1}+2}) = \cdots = \mathrm{Re}(\lambda_{i_j}) = r_j$$

with $r_1 > r_2 > \cdots > r_q$.

For $i = 1, \ldots, p$, let $v^{(i,1)}, \ldots, v^{(i,\nu_i)}$ be a basis for the eigenspace of the eigenvalue λ_i, where ν_i is the multiplicity of λ_i. Let $V \in \mathbb{C}^{n\times n}$ be the matrix of columns

$$v^{(1,1)}, \ldots, v^{(1,\nu_1)}, \ldots, v^{(p,1)}, \ldots, v^{(p,\nu_p)}$$

and let $W = V^{-1}$. We denote the rows of W by

$$w^{(1,1)}, \ldots, w^{(1,\nu_1)}, \ldots, w^{(p,1)}, \ldots, w^{(p,\nu_p)},$$

correspondly to the columns of V. Observe that the transposed of such rows are eigenvectors of A^T.

For $i = 1, \ldots, p$, the projection $P_i \in \mathbb{C}^{n\times n}$ on the eigenspace of λ_i is given by

$$P_i = V^{(i)} W^{(i)}, \tag{6}$$

where $V^{(i)} \in \mathbb{C}^{n\times \nu_i}$ is the matrix of columns $v^{(i,1)}, \ldots, v^{(i,\nu_i)}$ and $W^{(i)} \in \mathbb{C}^{\nu_i \times n}$ is the matrix of rows $w^{(i,1)}, \ldots, w^{(i,\nu_i)}$.

Finally, for $i = 1, \ldots, p$, let ω_i be the imaginary part of the eigenvalue λ_i.

The next theorem gives expressions for the condition numbers $K_l(t, A, y_0, \widehat{z}_0)$ and $K_l(t, A, y_0)$, $l = 1, \ldots, p$. Here and in the following, $\sqrt{-1}$ denotes the imaginary unit.

THEOREM 3.1. *Assume A diagonalizable. We have*

$$K_l(t, A, y_0, \widehat{z}_0) = \frac{\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{z}_0 \right|}{\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|}, \tag{7}$$

$$K_l(t, A, y_0) = \frac{\left\| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \right\|}{\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|}. \tag{8}$$

Proof. Since A is diagonalizable, we have

$$\mathrm{e}^{tA} = \sum_{\lambda_i \in \Lambda} \mathrm{e}^{\lambda_i t} P_i = \sum_{j=1}^{q} \mathrm{e}^{r_j t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} P_i$$

and then we obtain

$$\begin{aligned} K_l(t, A, y_0, \widehat{z}_0) &= \frac{\left| \sum_{j=1}^{q} \mathrm{e}^{r_j t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{z}_0 \right|}{\left| \sum_{j=1}^{q} \mathrm{e}^{r_j t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|} \\ &= \frac{\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{z}_0 \right|}{\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|}. \end{aligned}$$

Similarly, we obtain (8). □

REMARK 3.2:

1. In (7) and (8) all the exponentials $\mathrm{e}^{(r_j-r_1)t}$, $j=2,\ldots,q$, have $r_j-r_1<0$.
2. For a pair of complex conjugate eigenvalues λ_i and $\lambda_k=\overline{\lambda_i}$, we obtain (since $P_k=\overline{P_i}$)

$$\begin{aligned}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i+\mathrm{e}^{\sqrt{-1}\omega_k t}e_l^TP_k&=\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i+\overline{\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i}\\&=2\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\right).\end{aligned}$$

Then, in (7) and (8) we have, for $j=1,\ldots,q$ and $\lambda_i\in\Lambda_j$,

$$\sum_{\lambda_i\in\Lambda_j}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i=\sum_{\substack{\lambda_i\in\Lambda_j\\ \lambda_i\in\mathbb{R}}}e_l^TP_i+2\sum_{\substack{\lambda_i\in\Lambda_j\\ \omega_i>0}}\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\right).$$

3.1. Asymptotic behaviour

The next theorem describes the asymptotic behavior, as $t\to+\infty$, of the condition numbers $K_l(t,A,y_0,\widehat{z}_0)$ and $K_l(t,A,y_0)$, $l=1,\ldots,p$. We use the notation

$$f(t)\sim g(t),\ t\to+\infty,\ \text{for}\ \lim_{t\to+\infty}\frac{f(t)}{g(t)}=1.$$

THEOREM 3.3. *Assume A diagonalizable. We have*

$$K_l(t,A,y_0,\widehat{z}_0)\sim\mathrm{e}^{\left(r_{j^{**}}-r_{j^*}\right)t}\frac{\left|\sum_{\lambda_i\in\Lambda_{j^{**}}}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\widehat{z}_0\right|}{\left|\sum_{\lambda_i\in\Lambda_{j^*}}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\widehat{y}_0\right|},\ t\to+\infty,\quad(9)$$

$$K_l(t,A,y_0)\sim\mathrm{e}^{\left(r_{\overline{j}^*}-r_{j^*}\right)t}\frac{\left\|\sum_{\lambda_i\in\Lambda_{\overline{j}^*}}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\right\|}{\left|\sum_{\lambda_i\in\Lambda_{j^*}}\mathrm{e}^{\sqrt{-1}\omega_i t}e_l^TP_i\widehat{y}_0\right|},\ t\to+\infty,\quad(10)$$

where

$$\begin{aligned}j^*&:=\min\left\{j\in\{1,\ldots,q\}:e_l^TP_i\widehat{y}_0\neq0\ \textit{for some}\ \lambda_i\in\Lambda_j\right\},\\ j^{**}&:=\min\left\{j\in\{1,\ldots,q\}:e_l^TP_i\widehat{z}_0\neq0\ \textit{for some}\ \lambda_i\in\Lambda_j\right\},\\ \overline{j}^*&:=\min\left\{j\in\{1,\ldots,q\}:e_l^TP_i\neq0\ \textit{for some}\ \lambda_i\in\Lambda_j\right\}.\end{aligned}$$

Proof. For the numerator or denominator in (7) we have, with $u = \widehat{z}_0$ and $j(u) = j^{**}$ or $u = \widehat{y}_0$ and $j(u) = j^*$,

$$\left| \sum_{j=1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right| = \left| \sum_{\lambda_i \in \Lambda_{j(u)}} \mathrm{e}^{\left(r_{j(u)} - r_1\right)t} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right| (1 + E)$$

with

$$\begin{aligned} |E| &\leq \frac{\left| \sum_{j=j(u)+1}^{q} \mathrm{e}^{(r_j - r_1)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right|}{\left| \mathrm{e}^{\left(r_{j(u)} - r_1\right)t} \sum_{\lambda_i \in \Lambda_{j(u)}} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right|} \\ &= \frac{\left| \sum_{j=j(u)+1}^{q} \mathrm{e}^{\left(r_j - r_{j(u)}\right)t} \sum_{\lambda_i \in \Lambda_j} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right|}{\left| \sum_{\lambda_i \in \Lambda_{j(u)}} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i u \right|}. \end{aligned}$$

Now, by letting $t \to +\infty$ (see point 1 in Remark 3.4 below), we obtain (9). Similarly, we obtain (10). □

Remark 3.4:

1. In (9) we assume there exists $\sigma > 0$ such that

$$\mathcal{A}_\sigma = \left\{ t \geq 0 : \left| \sum_{\lambda_i \in \Lambda_{j^{**}}} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{z}_0 \right| \geq \sigma \right. \\ \left. \text{and} \left| \sum_{\lambda_i \in \Lambda_{j^*}} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right| \geq \sigma \right\} \quad (11)$$

has $+\infty$ as an accumulation point. In (9), we consider $t \to +\infty$ with $t \in \mathcal{A}_\sigma$. Analogously, in (10) we assume there exists $\sigma > 0$ such that

$$\mathcal{B}_\sigma = \left\{ t \geq 0 : \left| \sum_{\lambda_i \in \Lambda_{j^*}} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right| \geq \sigma \right\} \quad (12)$$

has $+\infty$ as an accumulation point. In (10), we consider $t \to +\infty$ with $t \in \mathcal{B}_\sigma$.

2. We have $\overline{j}^* \leq j^*$ and then $r_{\overline{j}^*} - r_{j^*} \geq 0$ in the exponential $\mathrm{e}^{(r_{\overline{j}^*} - r_{j^*})t}$ in (10).

3. A generic situation for A, y_0 and $\widehat{z}_0$ is $j^* = j^{**} = \overline{j}^* = 1$, where we have, as $t \to +\infty$,

$$K_l(t, A, y_0, z_0) \sim \frac{\left| \sum_{\lambda_i \in \Lambda_1} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{z}_0 \right|}{\left| \sum_{\lambda_i \in \Lambda_1} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|},$$

$$K_l(t, A, y_0) \sim \frac{\left\| \sum_{\lambda_i \in \Lambda_1} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \right\|}{\left| \sum_{\lambda_i \in \Lambda_1} \mathrm{e}^{\sqrt{-1}\omega_i t} e_l^T P_i \widehat{y}_0 \right|}.$$

The next theorem considers the generic situation for A, y_0 and $\widehat{z}_0$ described in point 3 in the previous remark, namely $j^* = j^{**} = \overline{j}^* = 1$.

THEOREM 3.5. *Suppose that A is diagonalizable and it has a unique real eigenvalue λ_1 of multiplicity one, or a unique pair λ_1 and $\lambda_2 = \overline{\lambda_1}$ of complex conjugate eigenvalues of multiplicity one, as rightmost eigenvalues. Let v be an eigenvector of λ_1 and let w be the first row of $W = V^{-1}$, V being the matrix of the eigenvectors with v as first column (see page 3). Let $l = 1, \ldots, n$ such that $v_l \neq 0$. If $w\widehat{y}_0 \neq 0$ and $w\widehat{z}_0 \neq 0$, then, as $t \to +\infty$,*

$$K_l(t, A, y_0, \widehat{z}_0) \to \frac{|w\widehat{z}_0|}{|w\widehat{y}_0|}, \quad K_l(t, A, y_0) \to \frac{\|w\|}{|w\widehat{y}_0|} \tag{13}$$

when the rightmost eigenvalue is the real eigenvalue and

$$K_l(t, A, y_0, \widehat{z}_0) \sim \frac{\left| \mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w \right) \widehat{z}_0 \right|}{\left| \mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w \right) \widehat{y}_0 \right|}, \tag{14}$$

$$K_l(t, A, y_0) \sim \frac{\left\| \mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w \right) \right\|}{\left| \mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w \right) \widehat{y}_0 \right|} \tag{15}$$

when the rightmost eigenvalues are the complex conjugate pair.

Proof. We have (see (6)) $P_1 = vw$ and then $e_l^T P_1 = v_l w \neq 0$, $e_l^T P_1 \widehat{y}_0 = v_l w \widehat{y}_0 \neq 0$ and $e_l^T P_1 \widehat{z}_0 = v_l w \widehat{z}_0 \neq 0$. So $j^* = j^{**} = \overline{j}^* = 1$ and then (see point 3 in

Remark 3.4), as $t \to +\infty$,

$$K_l(t, A, y_0, z_0) \sim \frac{\left|e_l^T P_1 \widehat{z}_0\right|}{\left|e_l^T P_1 \widehat{y}_0\right|} = \frac{|w\widehat{z}_0|}{|w\widehat{y}_0|},$$
$$K_l(t, A, y_0) \sim \frac{\left\|e_l^T P_1\right\|}{\left|e_l^T P_1 \widehat{y}_0\right|} = \frac{\|w\|}{|w\widehat{y}_0|}$$

when the rightmost eigenvalue is the real eigenvalue and

$$\begin{aligned} K_l(t, A, y_0, z_0) &\sim \frac{\left|\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} e_l^T P_1 + \mathrm{e}^{\sqrt{-1}\omega_2 t} e_l^T P_2\right) \widehat{z}_0\right|}{\left|\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} e_l^T P_1 + \mathrm{e}^{\sqrt{-1}\omega_2 t} e_l^T P_2\right) \widehat{y}_0\right|} \\ &= \frac{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right) \widehat{z}_0\right|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right) \widehat{y}_0\right|}, \\ K_l(t, A, y_0) &\sim \frac{\left\|\mathrm{e}^{\sqrt{-1}\omega_1 t} e_l^T P_1 + \mathrm{e}^{\sqrt{-1}\omega_2 t} e_l^T P_2\right\|}{\left|\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} e_l^T P_1 + \mathrm{e}^{\sqrt{-1}\omega_2 t} e_l^T P_2\right) \widehat{y}_0\right|} \\ &= \frac{\left\|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\right\|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right) \widehat{y}_0\right|} \end{aligned}$$

when the rightmost eigenvalues are the complex conjugate pair (recall point 2 in Remark 3.2). □

Remark 3.6:

1. The assumption that A is diagonalizable and it has a unique real eigenvalue of multiplicity one, or a unique pair of complex conjugate eigenvalues of multiplicity one, as rightmost eigenvalues is a generic situation for A. Moreover, also $v_l \neq 0$ for any $l = 1, \ldots, n$ is a generic situation for A. Finally, $w\widehat{y}_0 \neq 0$ and $w\widehat{z}_0 \neq 0$ are generic situations for y_0 and $\widehat{z}_0$.

2. When the rightmost eigenvalue is the real eigenvalue, there exists $\sigma > 0$ such that $\mathcal{A}_\sigma = \mathbb{R}^+$ and there exists $\sigma > 0$ such that $\mathcal{B}_\sigma = \mathbb{R}^+$ (remind (11) and (12)) . So in (13) we can consider $t \to +\infty$ without restrictions on t. Moreover, observe that the limits in (13) are independent of l (independent of the particular component).

3. In (14) and (15), by setting

$$v_l = |v_l| \mathrm{e}^{\sqrt{-1}\alpha_l}, \ \widehat{w} = \frac{w}{\|w\|} = \left(|\widehat{w}_k| \, \mathrm{e}^{\sqrt{-1}\beta_k}\right)_{k=1,\ldots,n}$$

(observe that $\widehat{w}$ is a unit vector and, if $\|\cdot\|$ is a p-norm, $|\widehat{w}_k| \leq 1$, $k = 1, \ldots, n$) and

$$\widehat{w}\widehat{y}_0 = |\widehat{w}\widehat{y}_0|\, \mathrm{e}^{\sqrt{-1}\gamma(\widehat{y}_0)},\ \widehat{w}\widehat{z}_0 = |\widehat{w}\widehat{z}_0|\mathrm{e}^{\sqrt{-1}\gamma(\widehat{z}_0)},$$

we can write

$$\frac{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{z}_0\right|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{y}_0\right|} = \frac{|\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{z}_0\right)\right)|}{|\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{y}_0\right)\right)|} \cdot \frac{|\widehat{w}\widehat{z}_0|}{|\widehat{w}\widehat{y}_0|},$$

and

$$\frac{\left\|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\right\|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{y}_0\right|} = \frac{\left\|\left(|\widehat{w}_k|\cos\left(\omega_1 t + \alpha_l + \beta_k\right)\right)_{k=1,\ldots,n}\right\|}{|\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{z}_0\right)\right)|} \cdot \frac{1}{|\widehat{w}\widehat{y}_0|},$$

So, the long-time oscillations of $K_l(t, A, y_0, \widehat{z}_0)$ and $K_l(t, A, y_0)$ are scaled by the factors $\frac{|\widehat{w}\widehat{z}_0|}{|\widehat{w}\widehat{y}_0|} = \frac{|w\widehat{z}_0|}{|w\widehat{y}_0|}$ and $\frac{1}{|\widehat{w}\widehat{y}_0|} = \frac{\|w\|}{|w\widehat{y}_0|}$, respectively, independent of l (independent of the particular component). Moreover, observe that

$$\begin{aligned}\mathcal{A}_\sigma \quad = \quad & \{t \geq 0 : |\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{z}_0\right)\right)| \cdot |v_l| \cdot |w\widehat{z}_0| \geq \sigma \\ & \qquad \text{and } |\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{y}_0\right)\right)| \cdot |v_l| \cdot |w\widehat{y}_0| \geq \sigma\}\end{aligned}$$

and

$$\mathcal{B}_\sigma = \{t \geq 0 : |\cos\left(\omega_1 t + \alpha_l + \gamma\left(\widehat{y}_0\right)\right)| \cdot |v_l| \cdot |w\widehat{y}_0| \geq \sigma\}$$

Thus, there exists $\sigma > 0$ such that $\mathcal{A}_\sigma$ and $\mathcal{B}_\sigma$ are countable unions of intervals (whose lengths are uniformly away from zero) with $+\infty$ as an accumulation point.

4. Examples

We show two examples with the matrix A in (1) taken from the MATLAB gallery test. We use the euclidean norm $\|\cdot\| = \|\cdot\|_2$ for measuring the relative error of the perturbation of y_0.

In the first example, we consider $A = gallery('lesp', n)$ with dimension $n = 10$. The matrix has ten real eigenvalues: the rightmost is -4.5491. In Figures 1 and 2, for two different initial values y_0, the graphs of $t \mapsto K_l(t, A, y_0) = \frac{\left\|e_l^T \mathrm{e}^{tA}\right\|}{|e_l^T \mathrm{e}^{tA}\widehat{y}_0|}$ (blue line), $l = 1, 2, 3, 4$, are plotted along with the constant value $\frac{\|w\|}{|w\widehat{y}_0|}$ (red line). Just for a comparison, in Figure 3 we see the graph of $t \mapsto \left\|e_l^T \mathrm{e}^{tA}\right\|$, where $\left\|e_l^T \mathrm{e}^{tA}\right\|$ is the worst magnification factor of the absolute error at the time t.

In the second example, we consider as $A = -gallery('parter', n)$ with dimension $n = 10$. The matrix has five complex conjugate pairs of eigenvalues:

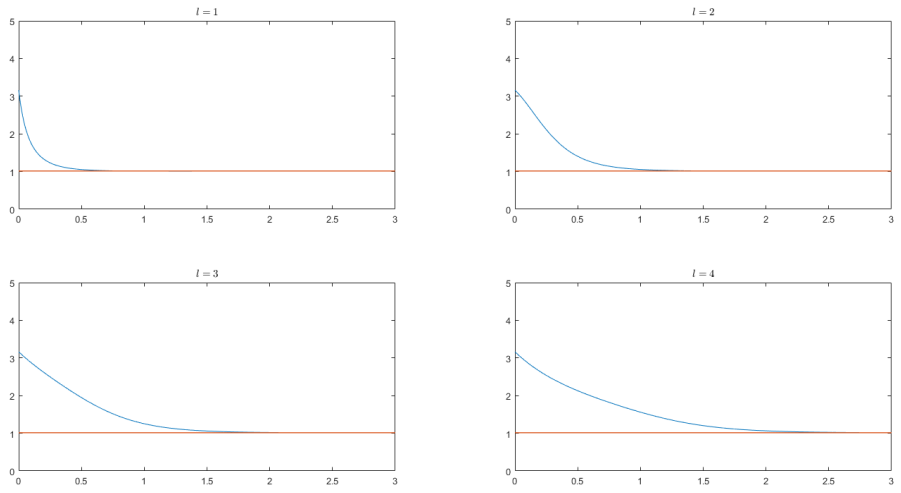

Figure 1: $K_l(t, A, y_0)$ (blue line), $l = 1, 2, 3, 4$, along with the constant value $\frac{\|w\|}{|w\widehat{y}_0|} = 1.0143$ (red line) for $y_0 = (1, \ldots, 1)$. The abscissas are the times $t \in [0, 3]$.

the rightmost pair is $-0.9066 \pm \sqrt{-1} \cdot 2.7709$. In Figures 4 and 5, for two different initial values y_0, we see the graphs of $t \mapsto K_l(t, A, y_0)$ (blue line), $l = 1, 2, 3, 4$, along with the graph of $t \mapsto \frac{\left\|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\right\|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{y}_0\right|}$ (red line) and the constant value $\frac{\|w\|}{|w\widehat{y}_0|}$ (yellow line). In Figure 6, we see the graph of $t \mapsto \left\|e_l^T \mathrm{e}^{tA}\right\|$ for the absolute error.

In both examples, the asymptotic behavior described in Theorem 3.5 is confirmed. Observe that the peaks of $K_l(t, A, y_0)$ not shown in Figures 2 and 4 are not of interest, because the components of the solution become zero at the peaks and the important error becomes the absolute error (recall point 1 in Remark 2.1). As remarked in point 3 of Remark 3.6, the quantity of interest in Figures 2 and 4 is the scale factor $\frac{\|w\|}{|w\widehat{y}_0|}$ of the oscillations (the constant yellow line).

5. Conclusions

In this paper we have studied the propagation of a perturbation in the initial value along the solution of a linear ODE. A normwise relative error is used for the perturbed initial value and componentwise relative errors are used for the perturbed solution.

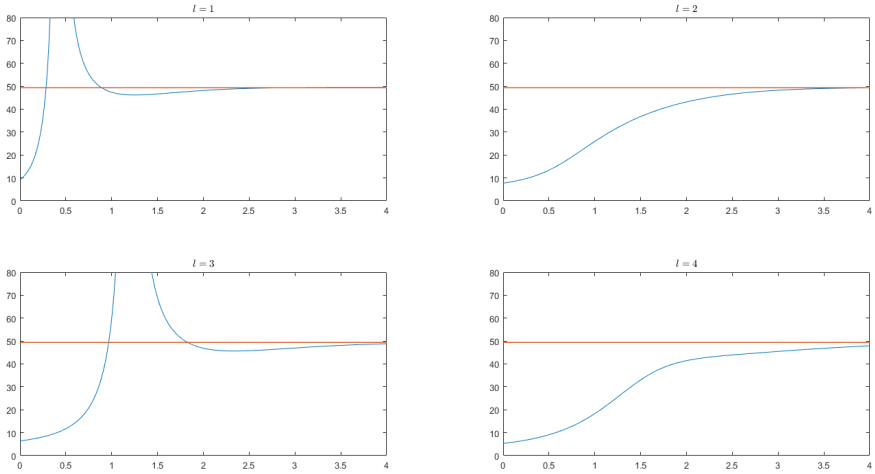

Figure 2: $K_l(t, A, y_0)$ (blue line), for $l = 1, 2, 3, 4$, along with the constant value $\frac{\|w\|}{|w\widehat{y}_0|} = 49.3891$ (red line) for $y_0 = ((-1.2)^l)_{l=1,\ldots,10}$. The abscissas are the times $t \in [0, 4]$.

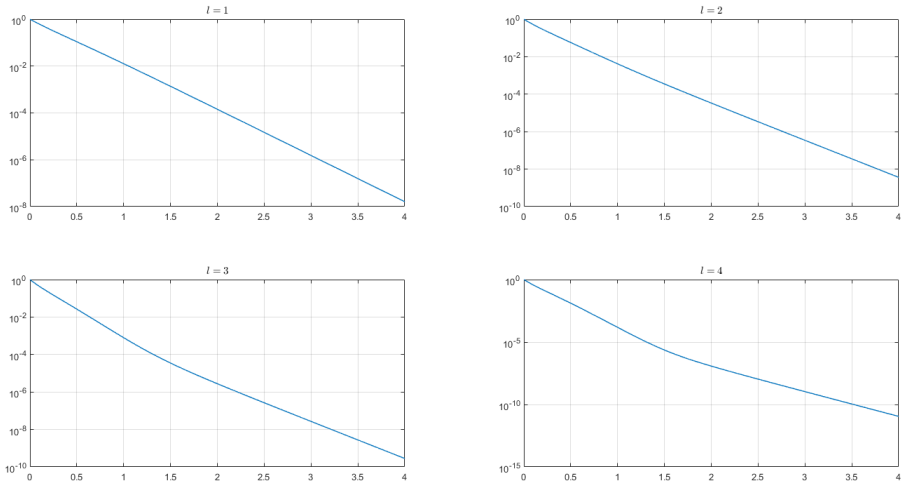

Figure 3: $\left\|e_l^T \mathrm{e}^{tA}\right\|$ (blue line) for $l = 1, 2, 3, 4$. The abscissas are the times $t \in [0, 4]$.

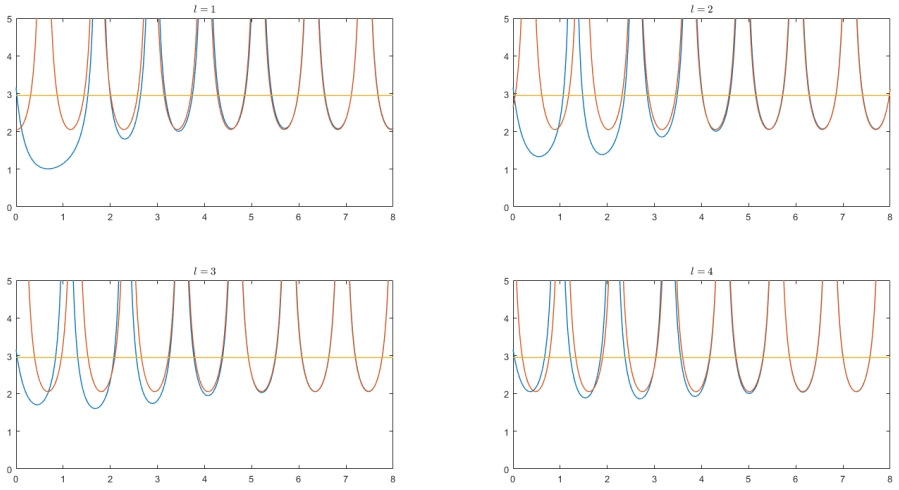

Figure 4: $K_l(t, A, y_0)$ (blue line) for $l = 1, 2, 3, 4$, along with $\frac{\left\|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\right\|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{y}_0\right|}$ (red line) and the constant value $\frac{\|w\|}{|w\widehat{y}_0|} = 2.9509$ (yellow line) for $y_0 = (1, \ldots, 1)$. The abscissas are the times $t \in [0, 8]$.

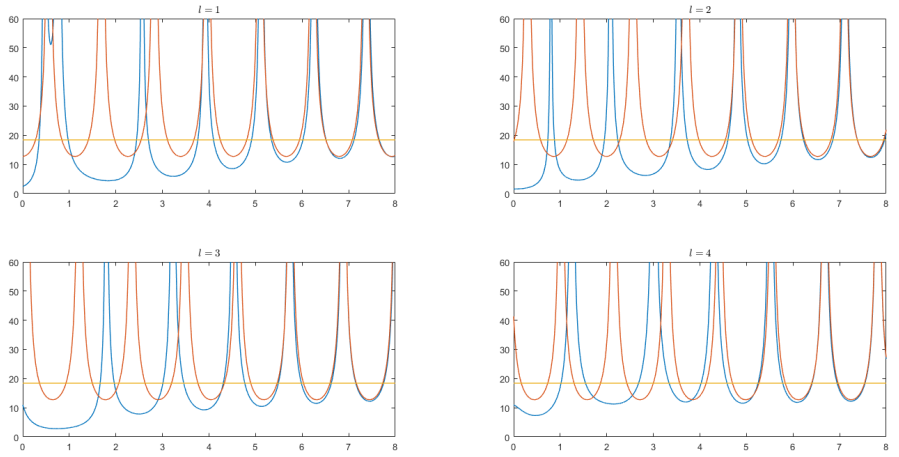

Figure 5: $K_l(t, A, y_0)$ (blue line), for $l = 1, 2, 3, 4$, along with $\frac{\left\|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\right\|}{\left|\mathrm{Re}\left(\mathrm{e}^{\sqrt{-1}\omega_1 t} v_l w\right)\widehat{y}_0\right|}$ (red line) and the constant value $\frac{\|w\|}{|w\widehat{y}_0|} = 18.4079$ (yellow line) for $y_0 = (0.9, -1.4, 0.2, 0.2, -0.2, 0.9, -0.4, -0.8, 0.3, 0.5)$. The abscissas are the times $t \in [0, 8]$.

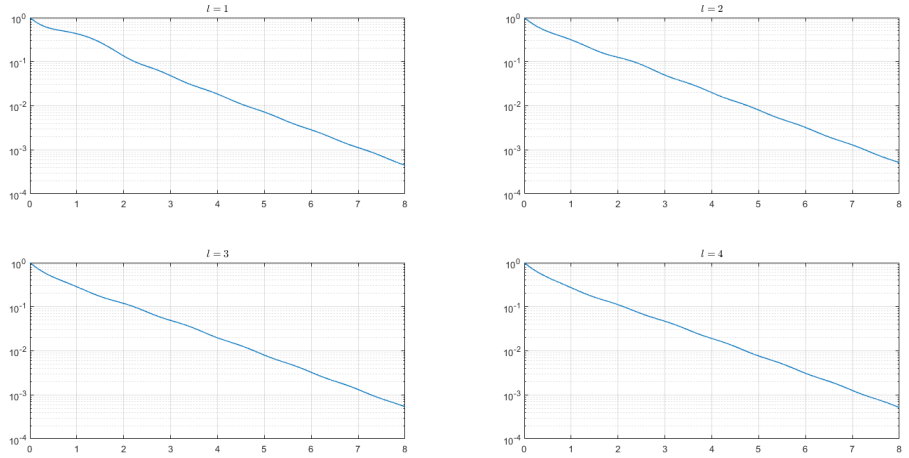

Figure 6: $\left\|e_l^T \mathrm{e}^{tA}\right\|$ (blue line) for $l = 1, 2, 3, 4$. The abscissas are the times $t \in [0, 8]$.

The main result of the paper says that in the generic situation of a linear ODE with a diagonalizable matrix having a real eigenvalue of multiplicity one, or a complex conjugate pair of eigenvalues of multiplicity one, as rightmost eigenvalues, the error in the initial value is magnified in the components of the solution, in the worst case, by the factor $\frac{\|w\| \, \|y_0\|}{|w y_0|}$ over a long-time, where y_0 is the initial value and w is the first row of the inverse of the eigenvectors matrix (i.e. w^T is an eigenvector of A^T relevant to the rightmost real eigenvalue or to the rightmost complex conjugate pair). The magnification factor is the same for all the components.

References

[1] A. Al-Mohy and N. Higham, *Computing the Frechét derivative of the matrix exponential, with an application to condition number estimation*, SIAM Journal on Matrix Analysis and Applications **30** (2008/2009) no. 4, 1639-1657.

[2] A. Al-Mohy and N. Higham, *Computing the action of the matrix exponential, with an application to exponential integrator*s, SIAM Journal on Scientific Computing **33** (2011) no. 2, 488-511.

[3] P. Bürgisser and F. Cucker, *Condition: the geometry of numerical algorithms*, Springer, 2013.

[4] E. Deadman, *Estimating the condition number of $f(A)b$*, Numerical Algorithms **70** (2015), 287–308.

[5] A. Al-Mohy, *An efficient bound for the condition number of the matrix exponential*, Journal of Taibah University for Science **11** (2017) no. 2, 280-289.

[6] N. Higham, *Functions of matrices - theory and computation*, SIAM, 2008.
[7] B. Hunt and V. Kaloshin, *Prevalence. Handbook of Dynamical Systems*, Elsevier, 2010.
[8] B. Kågström, *Bounds and perturbation bounds for the matrix exponential*, BIT **17** (1977) no. 1, 39-57.
[9] A. Levis, *Some computational aspects of the matrix exponential*, IEEE Transactions on Automatic Control AC-14 1969, 410–411.
[10] C. Van Loan, *The sensitivity of the matrix exponential*, SIAM Journal on Numerical Analysis **14** (1977) no. 6, 971-981.
[11] S. Maset, *Conditioning and relative error propagation in linear autonomous ordinary differential equations*, Discrete and Continuous Dynamical Systems Series B **23** (2018) no. 7, 2879-2909.
[12] C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Review **45** (2003) no. 1, 3-49.
[13] C. Van Loan, *A study of the matrix exponential*, Numerical Analysis Report No. 10 (1975, reprinted 2006), University of Manchester.
[14] W. Zhu, J. Xue and W. Gao, T*he sensitivity of the exponential of an essentially nonnegative matrix*, Journal of Computational Mathematics **26** (2008) no. 2, 250-258.

Authors' addresses:

A. Farooq
Dipartimento di Matematica e Geoscienze
Università di Trieste, Trieste, Italy
E-mail: asma.farooq@phd.units.it

S. Maset
Dipartimento di Matematica e Geoscienze
Università di Trieste, Trieste, Italy
E-mail: maset@units.it

Received December 23, 2020
Revised January 18, 2021
Accepted January 20, 2021