Erratum Numer. Math. 76, 441-462 (1997)

The life-span of backward error analysis for numerical integrators

E. Hairer, Ch. Lubich

Some terms have been forgotten in Eq. (4.26) on page 460. The correct estimate is

$$\|E_N\| \le \frac{R}{\mathrm{e}} \sum_{j\ge 2} \frac{1}{(2j-1)} \binom{2j}{j} \Big(\frac{0.2\,\mathrm{e}}{4}\Big)^j \sum_{\substack{1\le\rho_1,\ldots,\rho_j\le N\\ \rho_1+\ldots+\rho_j\ge N+1}} (B\rho_1 h)^{\rho_1}\cdot\ldots\cdot(B\rho_j h)^{\rho_j}.$$

Because of $B = 2\mathrm{e}\nu M/R$ and $hN \le R/(2\mathrm{e}^2\nu M)$ we have $Bh \le (\mathrm{e}N)^{-1}$, so that we must estimate the expression

$$C(j,N) = \sum_{\substack{1\le\rho_1,\ldots,\rho_j\le N\\ \rho_1+\ldots+\rho_j\ge N+1}} \Big(\frac{\rho_1}{\mathrm{e}N}\Big)^{\rho_1}\cdot\ldots\cdot\Big(\frac{\rho_j}{\mathrm{e}N}\Big)^{\rho_j}. \tag{E.1}$$

Lemma. For all $j\ge 2$ and for all $N\ge 2$ it holds

$$C(j,N) \le c\,\mathrm{e}^{-j}\,\mathrm{e}^{-(N+1)} \qquad \text{with} \quad c = 24.27\,. \tag{E.2}$$

Remark. Numerical computations revealed that the estimate holds with $c = 13.13$ for all $N\ge 2$, with $c = 11.78$ for $N\ge 3$, and with $c = 8.55$ for $N\ge 4$.

As in the proof of the publication we thus obtain

$$\|E_N\| \le \frac{R}{\mathrm{e}} \sum_{j\ge 2} \frac{1}{(2j-1)} \binom{2j}{j} \Big(\frac{0.2}{4}\Big)^j c\,\mathrm{e}^{-(N+1)} \le \frac{R}{\mathrm{e}}\Big(1-\sqrt{1-0.2}-0.1\Big)c\,\mathrm{e}^{-(N+1)},$$

and consequently $\|E_N\| \le 0.0021\,c\,R\,\mathrm{e}^{-(N+1)}$. Using the values of the above remark, we see that Lemma 12 of the article remains true for $N\ge 4$. For $N=2$, the constant 0.02 has to be replaced by 0.028.

Proof of the Lemma. We start by proving two useful estimates (parts (a) and (b)), and we prove the estimate of the lemma in parts (c) to (f).

a) For all $N\ge 2$ it holds

$$\sum_{\rho=1}^{N} \Big(\frac{\rho}{\mathrm{e}N}\Big)^{\rho} \le \frac{0.7}{N}. \tag{E.3}$$

This estimate can be proved as follows: The function $f(x) = (x/\mathrm{e}N)^x$ is monotonically decreasing on the interval $0 < x \le N$, because $f'(x) = (x/\mathrm{e}N)^x\big(\ln(x/\mathrm{e}N)+1\big)$. Consequently, $(\rho/\mathrm{e}N)^\rho \le (3/\mathrm{e}N)^3$ for $3\le\rho\le N$ and we have

$$\sum_{\rho=1}^{N} \Big(\frac{\rho}{\mathrm{e}N}\Big)^{\rho} \le \frac{1}{\mathrm{e}N} + \frac{4}{\mathrm{e}^2N^2} + (N-2)\frac{27}{\mathrm{e}^3N^3} \le \frac{1}{N}\Big(\frac{1}{\mathrm{e}} + \frac{4}{\mathrm{e}^2 N} + \frac{27(N-2)}{\mathrm{e}^3N^2}\Big) \le \frac{0.7}{N}$$

(the expression in brackets is maximal for $N=3$).

b) For a j-tupel $(\rho_1,\dots,\rho_j)$ satisfying $1\le\rho_1,\dots,\rho_j\le N$ and $\rho_1+\ldots+\rho_j=N+\ell$ (with $\ell\ge 1$) it holds

$$\Big(\frac{\rho_1}{N}\Big)^{\rho_1}\cdot\ldots\cdot\Big(\frac{\rho_j}{N}\Big)^{\rho_j}\le\frac{1}{N^{j-s-1}}\Big(\frac{k}{N}\Big)^{k}\tag{E.4}$$

where s,k are the unique integers ($s\ge 0$, $1\le k\le N-1$) such that

$$N+\ell=sN+j-s-1+k.\tag{E.5}$$

This means that the maximum of the left-hand expression in (E.4) is attained at the border of the considered domain, i.e., s values among $\rho_1,\dots,\rho_j$ equal N, $j-s-1$ values equal 1, and the remaining one (determined by $\rho_1+\ldots+\rho_j=N+\ell$) is denoted by k.

The statement (E.4) follows from the fact that for $2\le\rho\le\sigma\le N-1$

$$\Big(\frac{\rho}{N}\Big)^{\rho}\Big(\frac{\sigma}{N}\Big)^{\sigma}\le\Big(\frac{\rho-1}{N}\Big)^{\rho-1}\Big(\frac{\sigma+1}{N}\Big)^{\sigma+1}.\tag{E.6}$$

Indeed, repeated application of this estimate pushes the values ρ_i to the border of the domain. The estimate (E.6), which is equivalent to $\rho\ln\rho+\sigma\ln\sigma\le(\rho-1)\ln(\rho-1)+(\sigma+1)\ln(\sigma+1)$, follows from the fact that $f(x)=x\ln x-(x-1)\ln(x-1)$ is monotonically increasing for $x>1$ ($f'(x)=\ln x-\ln(x-1)=\ln(x/(x-1))>0$ for $x>1$).

c) We now prove the statement of the lemma for small values of N, say, for $N\le 3$. It holds

$$C(j,N)=\Big(\sum_{\rho=1}^{N}\Big(\frac{\rho}{\mathrm{e}N}\Big)^{\rho}\Big)^{j}\qquad\text{for}\quad j\ge N+1,$$

because the constraint $\rho_1+\ldots+\rho_j\ge N+1$ is automatically satisfied for $j\ge N+1$. Therefore, we obtain

$$\begin{aligned}C(j,2)&=\mathrm{e}^{-(N+1)}\,\mathrm{e}^{3}\,\mathrm{e}^{-j}\Big(\frac12+\frac1{\mathrm{e}}\Big)^{j}\le\mathrm{e}^{-(N+1)}\,\mathrm{e}^{-j}\Big(1+\frac{\mathrm{e}}{2}\Big)^{3}\le 13.13\,\mathrm{e}^{-j}\,\mathrm{e}^{-(N+1)}\\ C(j,3)&=\mathrm{e}^{-(N+1)}\,\mathrm{e}^{4}\,\mathrm{e}^{-j}\Big(\frac13+\frac{4}{9\mathrm{e}}+\frac{1}{\mathrm{e}^2}\Big)^{j}\le\ldots\le 8.72\,\mathrm{e}^{-j}\,\mathrm{e}^{-(N+1)}\end{aligned}$$

for $j\ge 3$ and $j\ge 4$, respectively. The remaining cases $j=2$ (for $N=2$) and $j\le 3$ (for $N=3$) can be treated directly. This yields the values of c given in the remark on page 1.

d) From now on we suppose $N\ge 4$. Using the estimate of part (b) we get

$$C(j,N)\le\sum_{\ell\ge 1}\mathrm{e}^{-(N+\ell)}\frac{1}{N^{j-s-1}}\Big(\frac{k}{N}\Big)^{k}M(j,N,\ell).\tag{E.7}$$

Here, s and k are given by (E.5), and $M(j,N,\ell)$ counts the number of j-tupels $(\rho_1,\dots,\rho_j)$ satisfying $1\le\rho_1,\dots,\rho_j\le N$ and $\rho_1+\ldots+\rho_j=N+\ell$. The desired estimate is then obtained from parts (e) and (f) below, where the terms with $s=0$ and those with $s>0$ are estimated separately.

e) We first consider those $\ell \geq 1$, for which $s = 0$ in the representation (E.5), i.e., $\ell \leq j-2$ (observe that for $j = 2$ this set is empty and nothing has to be proved). In this situation, the condition $\rho_1 + \ldots + \rho_j = N + \ell$ automatically implies $\rho_i < N$, so that the number $M(j,N,\ell)$ is given by the coefficient of $z^{N+\ell}$ in the series for $(z + z^2 + z^3 + \ldots)^j$. This coefficient is the binomial coefficient $(-1)^{N+\ell-j}\binom{-j}{N+\ell-j} = \binom{N+\ell-1}{j-1}$. Consequently, the first $j-2$ terms of the sum in (E.7) are given by

$$\mathrm{e}^{-(N+1)} \sum_{\ell=1}^{j-2} \mathrm{e}^{-\ell+1} \frac{1}{N^{j-1}} \Big(\frac{N+\ell-j+1}{N}\Big)^{N+\ell-j+1} \binom{N+\ell-1}{j-1}. \tag{E.8}$$

The expression $(\frac{N+\ell-j+1}{N})^{N+\ell-j+1}$ is a convex function of $\ell - j + 1$ (in fact, the second derivative of $f(x) = (\frac{N+x}{N})^{N+x}$ is $f''(x) = (\frac{N+x}{N})^{N+x}\big((\ln(\frac{N+x}{N}) + 1)^2 + \frac{1}{N+x}\big) > 0$ for $x > -N$). Hence, it is bounded by $(\frac{N-1}{N})^{N-1} \leq (3/4)^3$ for $N \geq 4$. Furthermore,

$$\frac{1}{N^{j-1}} \binom{N+\ell-1}{j-1} \leq \frac{1}{N^{j-1}} \binom{N+j-3}{j-1} = \prod_{i=1}^{j-1} \Big(\frac{N+i-2}{N \cdot i}\Big) \leq 10.05\, \mathrm{e}^{-j}.$$

Here, the first estimate follows from $\ell \leq j-2$. The second one can be checked by hand for $j = 3, 4, 5$ and for larger j it follows from the fact that the factor $(N+i-2)/(N \cdot i)$ is bounded by e^{-1} for $i \geq 5$ and $N \geq 4$. Inserting these estimates into (E.8) yields the bound

$$\mathrm{e}^{-(N+1)} \sum_{\ell=1}^{j-2} \mathrm{e}^{-\ell+1} \Big(\frac{3}{4}\Big)^3 10.05\, \mathrm{e}^{-j} \leq 6.71\, \mathrm{e}^{-j}\, \mathrm{e}^{-(N+1)}.$$

f) We finally consider in the sum (E.7) those $\ell \geq 1$ for which $s > 0$. Estimating $M(j,N,\ell)$ by N^{j-1}, this sum is seen to be bounded by

$$\mathrm{e}^{-(N+1)} \sum_{\ell \geq \max(1,j-3)} \mathrm{e}^{-\ell+1} \frac{1}{N^{j-s-1}} \Big(\frac{k}{N}\Big)^k N^{j-1},$$

where again s and k are given by (E.5). Writing the sum over ℓ as two sums (over s and over k) we get for this expression

$$\begin{aligned}
&\mathrm{e}^{-(N+1)} \sum_{s\geq 1} \sum_{k=1}^{N-1} \mathrm{e}^{-(s-1)(N-1)-j-k+3}\, N^s \Big(\frac{k}{N}\Big)^k \\
&= \mathrm{e}^{-(N+1)} \sum_{s \geq 1} \Big(\frac{N}{\mathrm{e}^{N-1}}\Big)^{s-1} \mathrm{e}^{-j+3}\, N \sum_{k=1}^{N-1} \Big(\frac{k}{\mathrm{e}N}\Big)^k \\
&\leq \mathrm{e}^{-(N+1)}\, \mathrm{e}^{-j}\, \mathrm{e}^3\, 0.7 \sum_{s\geq 1} \Big(\frac{N}{\mathrm{e}^{N-1}}\Big)^{s-1} \leq 17.56\, \mathrm{e}^{-j}\, \mathrm{e}^{-(N+1)}.
\end{aligned}$$

The estimate of the last line follows from (E.3) and from the fact that $N/\mathrm{e}^{N-1} \leq 4/\mathrm{e}^3$ for $N \geq 4$. The estimates of parts (e) and (f) yield the statement of the lemma for $N \geq 4$.

Acknowledgement. We thank Pierre Leone for drawing our attention to the error.