

Mathematical Tripos Part IB: Lent Term 2024

Numerical Analysis – Lecture 10

10 Stiff ODEs: linear stability and A-stability

Phenomenon 10.1 Consider the linear system

$$\mathbf{y}' = A\mathbf{y}, \quad \text{where} \quad A = \begin{bmatrix} -100 & 0 \\ 0 & -1 \end{bmatrix}.$$

The exact solution has two components e^{-t} and e^{-100t} : the first decays gently, whereas the second becomes practically zero almost at once. Suppose that we solve this ODE with the *forward Euler method*. As will be shown soon, the requirement that $\mathbf{y}_n \rightarrow \mathbf{0}$ (for fixed $h > 0$) leads to an unacceptable restriction on the size of h .

With greater generality, let us solve with the Euler method $\mathbf{y}' = A\mathbf{y}$ for a general $N \times N$ constant matrix. Then $\mathbf{y}_{n+1} = (I + hA)\mathbf{y}_n$, therefore $\mathbf{y}_n = (I + hA)^n \mathbf{y}_0$. Let the eigenvalues of A be (λ_i) , with corresponding linearly independent eigenvectors $(\mathbf{w}_i)$, and let $D = \text{diag}(\lambda_i)$ and $W = [\mathbf{w}_1, \dots, \mathbf{w}_N]$. Then $AW = WD$, hence $A = WDW^{-1}$, and we may represent the exact solution of the ODE explicitly as

$$\mathbf{y}(t) = e^{tA} \mathbf{y}_0, \quad \text{where} \quad e^{tA} := \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k = W e^{tD} W^{-1}.$$

Now assume that $\text{Re } \lambda_i < 0$ for all i . In that case, $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \mathbf{0}$, however, $\mathbf{y}_n = W(I + hD)^n W^{-1} \mathbf{y}_0$, therefore $\mathbf{y}_n \rightarrow \mathbf{0}$ for all initial values $\mathbf{y}_0$ if and only if $|1 + h\lambda_i| < 1$ for all i . In our example we thus require $|1 - h|, |1 - 100h| < 1$, hence $h < \frac{1}{50}$.

It is important to realise that this restriction, necessary to recovery of correct asymptotic behaviour, has *nothing* to do with local accuracy, since, for large n , the genuine ‘unstable’ component is exceedingly small. So, this restriction is solely to prevent this component from leading to an unbounded growth (instability) in the numerical solution.

Definition 10.2 (Stiff ODEs) We say that the ODE $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$ is *stiff* if (for some methods) we need to depress h to maintain *stability* well beyond requirements of accuracy.

Definition 10.3 (Linear stability domain) Suppose that a numerical method, applied to $y' = \lambda y$, $y(0) = 1$, with constant h , produces the solution sequence $\{y_n\}$. We call the set

$$\mathcal{D} = \{z := \lambda h \in \mathbb{C} : \lim_{n \rightarrow \infty} y_n = 0\}$$

the *linear stability domain* of the method.

Noting that the set of $\lambda \in \mathbb{C}$ for which $y(t) = e^{\lambda t} \rightarrow 0$ as $t \rightarrow \infty$ is the left half-plane $\mathbb{C}^- = \{z \in \mathbb{C} : \text{Re } z < 0\}$, we say that the method is *A-stable* if $\mathbb{C}^- \subseteq \mathcal{D}$.

Example 10.4 1) For the *forward Euler method*, as we have already seen, $y_n \rightarrow 0$ iff $|1 + \lambda h| < 1$, therefore its linear stability domain is $\mathcal{D} = \{z \in \mathbb{C} : |1 + z| < 1\}$.

2) For the *backward Euler*: $y_{n+1} = y_n + \lambda h y_{n+1}$, we get $y_{n+1} = (1 - \lambda h)^{-1} y_n$, whence $\mathcal{D} = \{z \in \mathbb{C} : |1 - z| > 1\}$, and the method is A-stable.

3) Solving $y' = \lambda y$ with the *trapezoidal rule*: $y_{n+1} = y_n + h(\frac{1}{2}f_{n+1} + \frac{1}{2}f_n)$, we obtain $y_{n+1} = [(1 + \frac{1}{2}\lambda h)/(1 - \frac{1}{2}\lambda h)] y_n$ thus, by induction, $y_n = [(1 + \frac{1}{2}\lambda h)/(1 - \frac{1}{2}\lambda h)]^n y_0$. Therefore

$$z \in \mathcal{D} \quad \Leftrightarrow \quad \left| \frac{1 + \frac{1}{2}z}{1 - \frac{1}{2}z} \right| < 1 \quad \Leftrightarrow \quad \text{Re } z < 0,$$

and we deduce that $\mathcal{D} = \mathbb{C}^-$. Hence, the method is A-stable as well.

Note that A-stability does not mean that *any* step size will do. We need to choose h small enough to ensure the right accuracy, but we don’t want to depress it much further to prevent instability.

10.1 Stability of multistep methods

By definition, the number $z = \lambda h$ is in the linear stability domain $\mathcal{D}$ of a multistep method

$$\sum_{m=0}^s a_m y_{n+m} = h \sum_{m=0}^s b_m f_{n+m} \quad (10.1)$$

if the sequence (y_n) which is a solution to the recurrence relation arising from (10.1) with $f_n = \lambda y_n$ satisfies $y_n \rightarrow 0$. The latter is true if and only if the roots of the (characteristic) equation

$$p(x) := \rho(x) - z\sigma(x) = \sum_{m=0}^s a_m x^m - z \sum_{m=0}^s b_m x^m = 0 \quad (10.2)$$

are less than one in absolute values.

If $z = \lambda h$ is on the boundary of the linear stability domain $\mathcal{D}$, then the characteristic polynomial p has a root x of modulus 1, i.e. $x = e^{it}$. Substituting such an x into (10.2), we get that the boundary of $\mathcal{D}$ is the curve $z(t) = \frac{\rho(e^{it})}{\sigma(e^{it})}$.

If we need to determine the *real stability interval* $I = \mathcal{D} \cap \mathbb{R}$, then it is highly likely that $z_1 = \frac{\rho(1)}{\sigma(1)}$ and $z_2 = \frac{\rho(-1)}{\sigma(-1)}$ are the end-points of I . Further, since $\rho(1) = 0$, this interval for many methods is $I = (\frac{\rho(-1)}{\sigma(-1)}, 0)$. Still some further work is needed to prove that this is indeed the case.

Example 10.5 Consider the 3-step 3-rd-order Adams-Bashforth method

$$y_{n+3} - y_{n+2} = h \left(\frac{23}{12} f_{n+2} - \frac{4}{3} f_{n+1} + \frac{5}{12} f_n \right).$$

With $f = \lambda y$, the characteristic polynomial is

$$p(x) = (x^3 - x^2) - z \left(\frac{23}{12} x^2 - \frac{4}{3} x + \frac{5}{12} \right).$$

We have then

$$\rho(-1) = -2, \quad \sigma(-1) = \frac{11}{3},$$

so we guess that $(\frac{\rho(-1)}{\sigma(-1)}, 0) = (-\frac{6}{11}, 0)$ is the real stability interval $I = \mathcal{D} \cap \mathbb{R}$.

1) If $z > 0$, then

$$p(1) = -z < 0, \quad p(+\infty) = +\infty,$$

hence p has a root $x_* > 1$, thus no stability.

2) If $z < -\frac{6}{11}$, then

$$p(-1) = -2 - \frac{11}{3} z > 0, \quad p(-\infty) = -\infty,$$

hence p has a root $x_* < -1$, and there is no stability either.

3) For $z \in (-\frac{6}{11}, 0)$, we have

$$p(-1) = -2 - \frac{11}{3} z < 0, \quad p(0) = -\frac{5}{12} z > 0, \quad p\left(\frac{1}{2}\right) = -\frac{1}{8} - \frac{11}{48} z < 0, \quad p(1) = -z > 0,$$

hence all three roots of p are real and reside within $(-1, 1)$, therefore the method is stable for $z = \lambda h \in (-\frac{6}{11}, 0)$.

Theorem 10.6 (The second Dahlquist barrier) *No multistep method of order $p \geq 3$ may be A-stable. (The $p = 2$ barrier for A-stability is attained by the trapezoidal rule.)*

However, some high-order multistep methods are still of good use. Stability properties of BDF, say, are satisfactory for most stiff equations. The point is that, in many stiff linear systems in applications, the eigenvalues are not just in $\mathbb{C}^-$ but also well away from $i\mathbb{R}$. All BDF s -step methods of order $p = s \leq 6$ (i.e., all convergent BDF methods) share the feature that their linear stability domain $\mathcal{D}$ includes a wedge about $(-\infty, 0)$: such methods are said to be A_0 -stable.

10.2 Stability of Runge-Kutta methods

Lemma 10.7 No explicit Runge-Kutta method may be A-stable (or even A_0 -stable).

Proof. See Exercise 6 on Examples Sheet 2. □

Unlike multistep or explicit Runge-Kutta methods, implicit high-order RK may be A-stable. For an s -stage method, its linear stability domain $\mathcal{D}$ is determined from the relation

$$y_{n+1} = r(z)y_n, \quad r(z) = \frac{p(z)}{q(z)}, \quad p, q \in \mathcal{P}_s, \quad (10.3)$$

where $r(z)$ is a rational function of degree s , and the inequality $|r(z)| < 1$ for $z \in \mathbb{C}^-$ is possible.

Example 10.8 Consider the 2-stage 3rd-order method from Example 9.5

$$\begin{aligned} k_1 &= f\left(t_n, y_n + \frac{1}{4}h(k_1 - k_2)\right), \\ k_2 &= f\left(t_n + \frac{2}{3}h, y_n + \frac{1}{12}h(3k_1 + 5k_2)\right), \\ y_{n+1} &= y_n + \frac{1}{4}h(k_1 + 3k_2). \end{aligned}$$

Applying it to $y' = \lambda y$ (and multiplying both sides with h), we have

$$\begin{aligned} hk_1 &= h\lambda\left(y_n + \frac{1}{4}hk_1 - \frac{1}{4}hk_2\right), \\ hk_2 &= h\lambda\left(y_n + \frac{1}{4}hk_1 + \frac{5}{12}hk_2\right). \end{aligned}$$

This is a linear system for hk_1 and hk_2 , whose solution (with $z = \lambda h$) is

$$\begin{bmatrix} hk_1 \\ hk_2 \end{bmatrix} = \begin{bmatrix} 1 - \frac{1}{4}z & \frac{1}{4}z \\ -\frac{1}{4}z & 1 - \frac{5}{12}z \end{bmatrix}^{-1} \begin{bmatrix} zy_n \\ zy_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} 1 - \frac{5}{12}z & -\frac{1}{4}z \\ \frac{1}{4}z & 1 - \frac{1}{4}z \end{bmatrix} \begin{bmatrix} zy_n \\ zy_n \end{bmatrix}.$$

To get y_{n+1} we need $\frac{1}{4}(hk_1 + 3hk_2) = \frac{z(1 - \frac{1}{6}z)}{\det(A)} y_n$, therefore (with $\det(A) = 1 - \frac{2}{3}z + \frac{1}{6}z^2$)

$$y_{n+1} = y_n + \frac{1}{4}h(k_1 + 3k_2) = \left(1 + \frac{z(1 - \frac{1}{6}z)}{1 - \frac{2}{3}z + \frac{1}{6}z^2}\right) y_n = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2} y_n.$$

Let

$$r(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}. \quad (10.4)$$

Then $y_{n+1} = r(z)y_n$, therefore, by induction, $y_n = r(z)^n y_0$ and we deduce that

$$\mathcal{D} = \{z \in \mathbb{C} : |r(z)| < 1\}.$$

We wish to prove that $|r(z)| < 1$ for every $z \in \mathbb{C}^-$, since this is equivalent to A-stability. This is done by a technique that can be applied to other RK methods (because of (10.3)). According to the *maximum modulus principle* from Complex Methods/Analysis, if $g(z)$ is analytic in a complex domain $\Omega \subset \mathbb{C}$, then $|g(z)|$ attains its maximum on the boundary $\partial\Omega$.

We apply this principle to $r(z)$ in (10.4). This is a rational function, hence its only singularities are the poles $2 \pm i\sqrt{2}$, thus r is analytic in $\Omega = \mathbb{C}^- = \{z \in \mathbb{C} : \operatorname{Re} z < 0\}$. Therefore, it attains its maximum modulus on $\partial\Omega = i\mathbb{R}$. Hence

$$\text{A-stability} \quad \Leftrightarrow \quad |r(z)| < 1, \quad z \in \mathbb{C}^- \quad \Leftrightarrow \quad |r(it)| \leq 1, \quad t \in \mathbb{R}.$$

In turn,

$$|r(it)|^2 \leq 1 \quad \Leftrightarrow \quad \left|1 - \frac{2}{3}it - \frac{1}{6}t^2\right|^2 - \left|1 + \frac{1}{3}it\right|^2 \geq 0.$$

But the last expression is $(1 - \frac{1}{6}t^2)^2 + \frac{4}{9}t^2 - (1 + \frac{1}{9}t^2) = \frac{1}{36}t^4 \geq 0$, and it follows that the method is A-stable.