(1) $u'(t) = Lu$ $\qquad L \in \mathbb{C}^{m\times m}$

$u(t): \mathbb{R} \to \mathbb{C}^m$

$L = R\Lambda R^{-1}$

R: eigenvectors (lin. indep.)

Λ: diagonal matrix of eigenvalues

$$u'(t) = R\Lambda R^{-1} u(t)$$

$$R^{-1}u'(t) = \Lambda R^{-1} u(t)$$

Let $w(t) = \Lambda R^{-1} u(t)$.

Then $w'(t) = \Lambda w(t)$

This is a system of m decoupled ODEs:

$$w_i'(t) = \lambda_i w_i(t)$$

$$1 \le i \le m$$

So discretizations of (1) will be absolutely if

$$k\lambda_i \in S \qquad 1 \le i \le m.$$

region of absolute stability

$$S = \{z \in \mathbb{C} : |R(z)| \le 1\}$$

What is the stabilit function $R(z)$ for a Runge-Kutta method?

$$Y_i = U^n + K\sum_{j=1}^{s} a_{ij} f(Y_j) \qquad 1 \le i \le s$$

$$U^{n+1} = U^n + K\sum_{j=1}^{s} b_j f(Y_j)$$

The coefficients a_{ij}, b_j determine the properties of the method.

Y_i is referred to as a "stage."

We often present the method using a "Butcher tableau":

$$\begin{array}{c|c} & A \\ \hline & b^T \end{array}$$

For example:

$$Y_1 = U^n$$

$$Y_2 = U^n + Kf(Y_1)$$

$$U^{n+1} = U^n + \frac{K}{2}\left(f(Y_1) + f(Y_2)\right)$$

$$\begin{array}{c|cc} & 0 & 0 \\ & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array}$$

A is strictly lower-triangular for every explicit RK method.

Let

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_s \end{bmatrix} \qquad \mathbb{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^s \qquad z = K\lambda$$

and consider the ODE

$$U'(t) = \lambda U$$

We have
$$Y = \mathbb{1}U^n + zAY \quad (2)$$

$$U^{n+1} = U^n + zb^TY \quad (3)$$

Solve (2) for Y:

$$(I - zA)Y = \mathbb{1}U^n$$

$$Y = (I - zA)^{-1}\mathbb{1}U^n$$

Substitute in (3):

$$U^{n+1} = \underbrace{\left(1 + zb^T(I - zA)^{-1}\mathbb{1}\right)}_{R(z)}U^n$$

$$U^{n+1} = R(z)U^n$$

What kind of function is $R(z)$?

- In general, it is rational

$$R(z) = \frac{P(z)}{Q(z)}$$ where P, Q are polynomials of degree at most s.

Recall:

Explicit Euler: $R(z) = 1 + z$

Implicit Euler: $R(z) = \frac{1}{1-z}$

These can be viewed as RK methods with $s = 1$ stage.

For explicit methods:

A is strictly lower-triangular

So $A^s = 0 \in \mathbb{R}^{s \times s}$.

We can write (Neumann)

$$(I - zA)^{-1} = I + zA + z^2A^2 + \cdots + z^{s-1}A^{s-1}.$$

We get

$$R(z) = \underbrace{1 + z b^T \left(\sum_{j=0}^{s-1} z^j A^j \right) \mathbb{1}}_{\text{Polynomial of degree } s.}$$

What is $R(z)$ for the method

$$\begin{array}{c|cc} & 0 & 0 \\ & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array} \quad ?$$

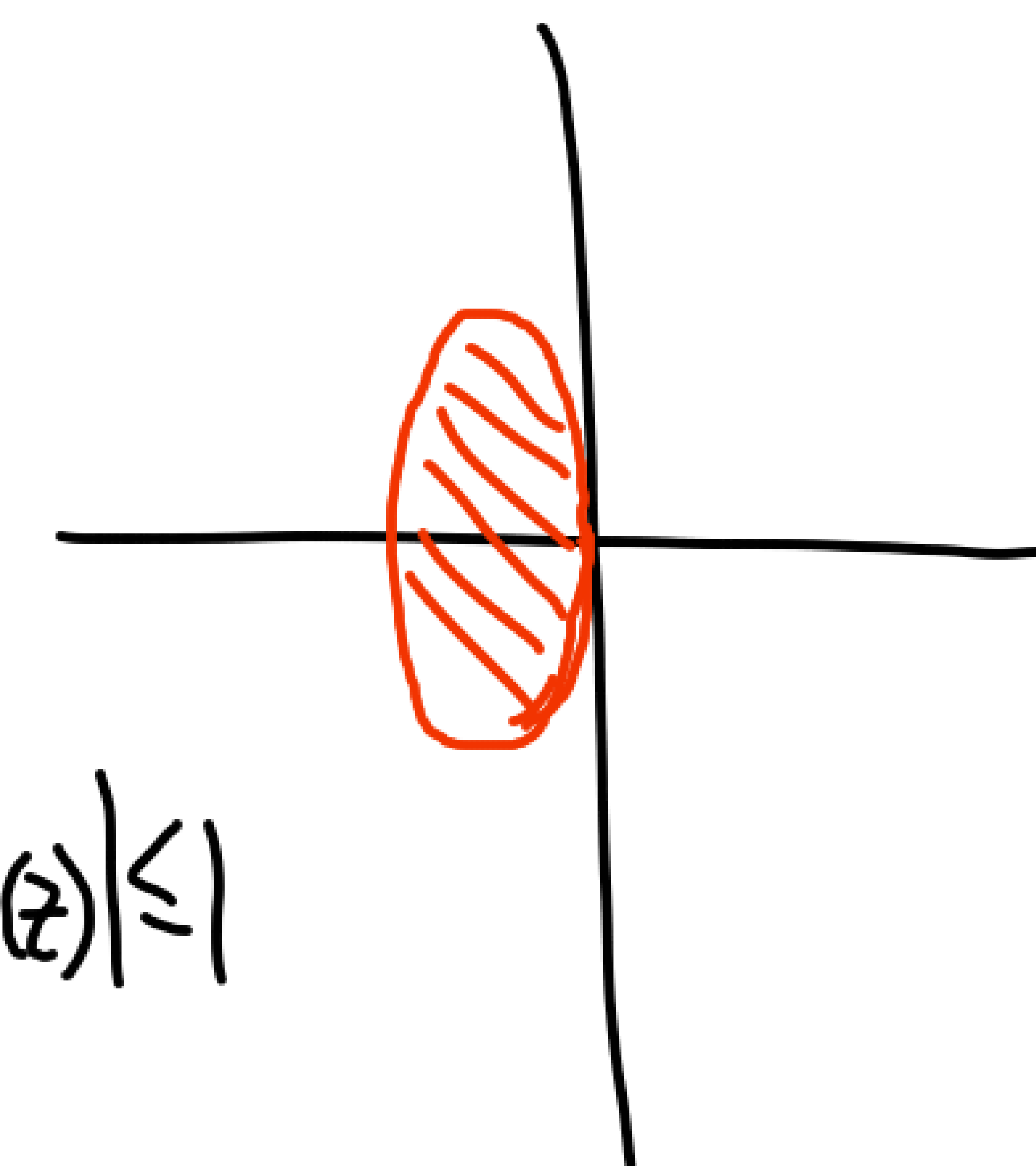
$$R(z) = 1 + z \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} (I + zA) \mathbb{1}$$

$$= 1 + z \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ z & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= 1 + z \begin{bmatrix} 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 1+z \end{bmatrix}$$

$$= 1 + z \left(\frac{1}{2} + \frac{1}{2} + \frac{z}{2} \right)$$

$$= 1 + z + \frac{z^2}{2}$$



$|R(z)| \le 1$

$$u'(t) = \lambda u \qquad u(0) = u_0$$

$$u(t) = e^{\lambda t} u_0$$

$$u(t_n + k) = e^{\lambda k} u(t_n)$$

$$u(t_n + k) = e^{z} u(t_n)$$

Compare: $U^{n+1} = R(z) U^n$

So we should have $R(z) \approx e^z$

$$e^z = 1 + z + \frac{z^2}{2} + \cdots$$

$$= \sum_{j=0}^{\infty} \frac{z^j}{j!}$$

In general we need

$$b^T A^{j-1} \mathbb{1} = \frac{1}{j!} \qquad 1 \le j \le p$$

for a method of order p.

These conditions are sufficient for the method to be accurate when applied to linear IVPs

$$u'(t) = Lu.$$

For accuracy on general IVPs

$$u'(t) = f(u)$$

additional conditions are required.

Error Estimation (local)

We want to bound

$$E^n = \|U^n - u(t_n)\|$$

but we can more efficiently bound the local error

$$\mathcal{L}^n = k\tau^n$$

chosen error tolerance

We want

$$\|\mathcal{L}^n\| < \varepsilon$$

Richardson Estimation

① Compulte $U_k^{n+1} \approx u(t_n+k)$ using step size k.

② Compute $U_{\frac{k}{2}}^{n+1} \approx u(t_n+k)$ using 2 steps of size $\frac{k}{2}$.

Since $\mathcal{L}_k^n = \mathcal{O}(k^{p+1}) \approx Ck^{p+1}$

where p is the order of accuracy of the method and

$$\mathcal{L}_{\frac{k}{2}}^n \approx C\left(\frac{k}{2}\right)^{p+1} \approx \frac{1}{2^{p+1}}\mathcal{L}_k^n$$

So $U_k^{n+1} - U_{\frac{k}{2}}^{n+1} = u(t_{n+1}) + \delta_k^n - \left(u(t_{n+1}) + \delta_{\frac{k}{2}}^n\right)$

$$= \delta_k^n - \delta_{\frac{k}{2}}^n \approx \delta_k^n\left(1 - \frac{1}{2^{p+1}}\right) \leftarrow$$ approximation of the local error.

Cost is a 50% increase in # of steps.

Embedded Runge-Kutta pairs

$$Y_1 = U^n$$
$$Y_2 = U^n + kf(Y_1)$$
$$U^{n+1} = U^n + \frac{k}{2}\left(f(Y_1) + f(Y_2)\right)$$

Y_2 is just the solution given by Euler's method.

$$Y_2 = u(t_{n+1}) + O(k^2)$$
$$U^{n+1} = u(t_{n+1}) + O(k^3)$$

So $\|U^{n+1} - Y_2\| = O(k^2)$

gives an estimate of the error in Y_2.

This estimate is "free"

We can write

$$\begin{array}{c|cc} & 0 & 0 \\ & 1 & 0 \\ \hline & \tfrac{1}{2} & \tfrac{1}{2} \\ & 1 & 0 \end{array}$$

Same A

Different b

More generally we can use an embedded pair

$$\begin{array}{c|c} & A \\ \hline & b \\ & \hat{b} \end{array}$$

$\hat{b}$ gives a lower order method used to estimate the error.

If method (A,b) has order p and method $(A,\hat{b})$ has order $p-1$ then

$$\delta = \|U^{n+1} - \hat{U}^{n+1}\| = \mathcal{O}(k^p)$$

estimates the error in the lower-order method.

Adapting the step size

Given an error estimate δ
and tolerance ε
if

$$\delta > \varepsilon$$

We reduce K and retake
the step. If

$$\delta < \varepsilon$$

We accept the step (U^{n+1}) and
proceed, possibly increasing the
step size.