Solution of Ordinary Differential Equations

Antony Jameson

Department of Aeronautics and Astronautics, Stanford University, Stanford, California, 94305

General linear multistep multiderivative method for ODE's

To solve

$$\frac{dy}{dt} = f(t, y) \tag{0.1}$$

Let

$$f^{(1)} = f = \frac{dy}{dt} \tag{0.2}$$

$$f^{(2)} = \frac{d^2y}{dt^2} = \frac{\partial f}{\partial t} + f\frac{\partial f}{\partial y} \tag{0.3}$$

$$\cdots \tag{0.4}$$

Then set

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \sum_{j=1}^{l_i} h^j \beta_{ij} f_{n+i}^{(j)} \tag{0.5}$$

where h is the time step.

Examples:

Taylor Method

$$y_{n+1} = y_n + hf^{(1)} + \frac{h^2}{2} f^{(2)} \cdots \frac{h^n}{n!} f^{(n)} \tag{0.6}$$

(1 stage, n derivatives)

Linear multistep

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} h\beta_i f_{n+i} \tag{0.7}$$

where $\alpha_k = 1$

Convergence

$$\lim_{h\to 0,\ nh\to T} y_n = y(x_n) \tag{0.8}$$

whenever starting conditions

$$y_\mu = \eta_\mu(h) \to \eta \tag{0.9}$$

as

$$h \to 0, \mu = 1, 2, \cdots k-1 \tag{0.10}$$

where $y(0) = \eta$

Order

Define

$$L(y(x), h) = \sum_{j=0}^{k} \{\alpha_j y(x+jh) - h\beta_j y'(x+jh)\} \tag{0.11}$$

where $y(x)$ is an arbitrary function in C'. Expanding in Taylor series

$$L(y(x), h) = C_0 y(x) + C_1 h y^{(1)}(x) + \cdots \tag{0.12}$$

where

$$C_0 = \alpha_0 + \alpha_1 + \cdots + \alpha_k \tag{0.13}$$

$$C_1 = \alpha_1 + 2\alpha_2 + \cdots + k\alpha_k - (\beta_0 + \beta_1 + \cdots + \beta_k) \tag{0.14}$$

$$C_q = \frac{1}{q!}(\alpha_1 + 2^q\alpha_2 + \cdots + k^q\alpha_k) - \frac{1}{(q-1)!}(\beta_1 + 2^{(q-1)}\beta_2 + \cdots + k^{q-1}\beta_k) \tag{0.15}$$

L is the local discretization error. If $L = O(h^{q+1})$, the scheme is of order q.

Consistency

The scheme is consistent if it is of order $p \geq 1$. This requires

$$\sum_{j=0}^{k} \alpha_j = 0, \quad \sum_{j=0}^{k} j\alpha_j = \sum_{j=0}^{k} \beta_j. \tag{0.16}$$

The first follows because as $h \to 0$ $y_{n+j} \to y(x)$ and the right hand side $\to 0$. The second follows because

$$y_{n+j} - y_n = jhy'(x) + jh\phi_{j,n}(h) \tag{0.17}$$

where $\phi_{j,n}(h) \to 0$ as $h \to 0$. Thus

$$\sum \alpha_j y_{n+j} - \sum \alpha_j y_n = h \sum j\alpha_j y'(x) + h \sum j\alpha_j \phi_{j,n}(h) \tag{0.18}$$

where

$$h \sum \beta_j f_{n+j} - y_n \sum \alpha_j = h \sum j\alpha_j y'(x) + h \sum j\alpha_j \phi_{j,n}(h). \tag{0.19}$$

Also $\sum \alpha_j = 0$, and $f_{n+j} \to y'(x)$. On dividing by h and taking the limit we get the 2nd condition.

Linear multistep methods

The schemes are implicit if $\beta_k \neq 0$, explicit if $\beta_k = 0$. The most common are Adams methods for which $\alpha_k = 1$, $\alpha_{k-1} = -1$, and $\alpha_{k-j} = 0, j > 1$. Explicit Adam's methods are called Adams Bashforth methods and implicit Adams methods are called Adams Moulton methods. They can be constructed starting from

$$y_{n+k} - y_{n+k-1} = \int_{(n+k-1)h}^{(n+k)h} f(t,y)dt \tag{0.20}$$

and fitting a polynomial through different numbers of points up to $n+k-1$ for explicit methods and $n+k$ for implicit methods and integrating the polynomial.

Explicit schemes with $\alpha_k = 1$, $\alpha_{k-2} = -1$ and $\alpha_{k-j} = 0$, $j \neq 0, 2$ are called Nystrom methods. The simplest is leap frog

$$y_{n+2} = y_n + 2hf_{n+1}. \tag{0.21}$$

Implicit methods with these coefficients on the left are called Milne methods.

4th order Adams methods

4 step Adams Bashforth

$$y_{n+1} = y_n + \frac{h}{24}\left(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}\right) \tag{0.22}$$

Stability intervals:

β_{real} $\quad -0.3$

β_{imag} 0.42

3 step Adams Bashforth

$$y_{n+1} = y_n + \frac{h}{24}\left(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}\right) \tag{0.23}$$

Stability intervals:

β_{real} -0.3

β_{imag} 1

Adams Bashforth Moulton, PECE

$$P: y_{n+1} = y_n + \frac{h}{24}\left(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}\right) \tag{0.24}$$

$$C: y_{n+1} = y_n + \frac{h}{24}\left(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}\right) \tag{0.25}$$

Stability intervals:

β_{real} -1.25

β_{imag} 0.9

4th order Nystrom scheme

$$y_{n+1} = y_{n-1} + \frac{h}{3}\left(8f_n - 5f_{n-1} + 4f_{n-2} - f_{n-3}\right) \tag{0.26}$$

4th order Milne scheme (Simpson's rule)

$$y_{n+1} = y_{n-1} + \frac{h}{3}\left(f_{n+1} + 4f_n + f_{n-1}\right) \tag{0.27}$$

Stability intervals:

β_{real} 0

Solution of recurrence equations

Let

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = 0 \tag{0.28}$$

where

$$y_0, y_1, ..., y_{k-1} \tag{0.29}$$

are given.

Consider a tried solution $y_n = \zeta^n$. This satisfies 0.28 if

$$P(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j = 0. \tag{0.30}$$

If $P(\zeta)$ has distinct roots ζ_i, then the general solution is

$$y_n = \sum_{i=1}^{k} A_i \zeta_i^n \tag{0.31}$$

where the k constants A_i can be chosen to satisfy the k conditions 0.29 by Lagrange interpolation. If $P(\zeta)$ has a root ζ_i of multiplicity q then there are not enough independent solutions ζ_i^n to satisfy an arbitrary choice of initial conditions. However, the root is also a root of multiplicity q of $f(\zeta) = \zeta^n P(\zeta)$ where n is an arbitrary integer.

This means that the first $q-1$ derivatives of $f(\zeta)$ vanish at this root, or

$$\alpha_k \zeta_i^{n+k} \cdots + \alpha_0 \zeta_i^n = 0 \tag{0.32}$$

$$\alpha_k (n+k) \zeta_i^{n+k-1} \cdots + \alpha_0 n \zeta_i^{n-1} = 0 \tag{0.33}$$

$$\cdots$$

$$\begin{aligned} \alpha_k (n+k)(n+k-1) \cdots + (n+k-q+2) \zeta_i^{n+k-q+1} \\ \cdots + \alpha_0 n(n-1) \cdots (n-q+2) \zeta_i^{n-q+1} = 0. \end{aligned} \tag{0.34}$$

It follows that

$$y_n = \zeta_i^n \tag{0.35}$$

$$y_n = n \zeta_i^n \tag{0.36}$$

$$\cdots$$

$$y_n = n(n-1) \cdots (n-q+2) \zeta_i^n. \tag{0.37}$$

are q alternative solutions of the recurrence equations. When these are included it may be verified that the equations for the constants A_i have a non singular determinant, so that these are independent. It follows also that there are no growing solutions of 0.28 if

$$|\zeta_i| \leq 1, \quad i = 1, 2, ..., k \tag{0.38}$$

and if ζ_q is a multiple root, then

$$|\zeta_q| < 1 \tag{0.39}$$

Characteristic polynomials

The 1st and 2nd characteristic polynomials are

$$p(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j$$

and

$$\sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j.$$

The scheme is consistent if and only if

$$p(1) = 0, \quad p'(1) = \sigma(1).$$

In fact is is of order p if and only if

$$\phi(\zeta) = \frac{p(\zeta)}{log\zeta} - \sigma(\zeta)$$

has a p fold zero at $\zeta = 1$.

Proof:

$$\begin{aligned} L(e^x, h) &= e^x \left(p(e^h) - h\sigma(e^h) \right) \\ &= e^x C_{p+1} h^{p+1} + O(h^{p+2}). \end{aligned}$$

Thus

$$\frac{p(e^h)}{h} - \sigma(e^h) = C_{p+1}h^p + \cdots$$

has a zero of order p at $h = 0$. Setting $\zeta = e^h$ yields the condition. Conversely if $\phi(\zeta)$ has a zero of order p at $\zeta = 1$, then $h\phi(e^h)$ has a zero of order $p+1$ at $h = 0$, and the order of $L(y, h)$ is p when $y = e^x$. Since the order depends only on the coefficients α_j, β_j, the order is therefore p.

Stability

Consider the model problem

$$\frac{dy}{dx} = \alpha y. \tag{0.40}$$

The scheme is said to be zero stable if there are no growing solutions for $\alpha = 0$. This will be the case if $p(\zeta)$ satisfies the root condition that all its roots ζ_i are in the unit disk

$$|\zeta_i| \leq 1, \quad i = 1, 2, ..., k$$

and if ζ_q is a multiple root,

$$|\zeta_q| < 1.$$

Consistency requires $p(1) = 0$, so 1 must be a simple root.

Let

$$\overline{h} = \alpha h$$

. If there exists a region of $\overline{h}$ in the complex plane for which there are no growing solutions of the difference approximation to (0.40), then the approximation is said to be weakly stable in this region, which is called the stability region. From the solution of the recurrence equations the stability region can be seen to be the region in which the polynomial

$$p(\zeta) - \overline{h}\sigma(\zeta)$$

satisfies the root condition.

Fundamental theorem of convergence

Dahlquist 1956 Math Scand 4, pp 33-53, Convergence and stability in the numerical integration of ODE's. (Henrici Section 5.3-3, pp 244-246)

The necessary and sufficient conditions for a linear multistep method to be convergent are that it be consistent and zero stable, that is that it should satisfy the root condition that the roots ζ_s of $p(\zeta)$ satisfy

$$|\zeta_s| \leq 1,$$

$|\zeta_s| < 1$ if the root is not simple.

Also

$$p(1) = 0$$

so 1 is a simple root (the principal root) and

$$p'(1) = \sigma(1).$$

The theorem was proved by Dahlquist for the general nonlinear ODE. The corresponding theorem for linear partial differential operators was proved by Lax.

Stable and consistent schemes are unstable for $\lambda h > 0$

Since

$$\begin{aligned}
\pi(\zeta, \overline{h}) &= p(\zeta) - \overline{h}\sigma(\zeta) && (0.41)\\
&= (\alpha_k - \overline{h}\beta_k)(\zeta - \zeta_1)\cdots(\zeta - \zeta_k) && (0.42)\\
&= O(\overline{h})^{p+1} && (0.43)
\end{aligned}$$

for a p^{th} order accurate scheme, if we set $\zeta = \exp(\overline{h})$ we find

$$(\exp(\overline{h}) - \zeta_1)(\quad)\cdots(\quad) = O(\overline{h})^{p+1}$$

and since $\zeta_1 = 1$ is a simple zero of $p(\zeta)$ only the first factor can be $O(\overline{h})^{p+1}$. Thus

$$\zeta_1 = \exp(\overline{h}) + O(\overline{h})^{p+1}$$

and must have a positive real part if $\overline{h}$ has a positive real part.

Regions of absolute and relative stability

For the model problems

$$y' = \lambda y$$

with step length h, the region (or interval) of absolute stability is the region containing values of

$$\overline{h} = \lambda h$$

for which all roots ζ_s of

$$\pi(\zeta, \overline{h}) = p(\zeta) - \overline{h}\sigma(\zeta)$$

satisfy $|\zeta_s| < 1$.

Defining ζ_1 as the principal root so that $\zeta_1 \to 1$ as $h \to 0$ and is a simple root, the region of relative stability is the region of containing values of $\overline{h}$ such that

$$|\zeta_s| < |\zeta_1|, \quad s = 2, 3, \cdots k$$

Example of relative and absolute stability

Consider the Milne scheme

$$y_{n+2} - y_n = \frac{h}{3}(f_{n+2} + 4f_{n+1} + f_n).$$

Then

$$p(\zeta) = \zeta^2 - 1$$

$$\sigma(\zeta) = \frac{1}{3}(\zeta^2 + 4\zeta + 1)$$

$$\pi(\zeta, \overline{h}) = \zeta^2 - 1 - \frac{\overline{h}}{3}(\zeta^2 + 4\zeta + 1)$$

From consistency we have

$$\zeta_1 = 1 + \overline{h} + O(\overline{h}^2).$$

Set

$$\zeta_2 = -1 + \gamma\overline{h} + O(\overline{h}^2).$$

Then

$$\begin{aligned}
\pi(\zeta, \overline{h}) &= \left(\zeta - 1 - \overline{h} + O(\overline{h}^2)\right)\left(\zeta + 1 - \gamma\overline{h} + O(\overline{h}^2)\right) & (0.44) \\
&= \zeta^2 - (1+\gamma)\overline{h}\zeta - (1+\overline{h})(1-\gamma\overline{h}) + O(\overline{h}^2) & (0.45)
\end{aligned}$$

Since the coefficient of ζ is $-\frac{4\overline{h}}{3}/(1-\frac{\overline{h}}{3})$ it follows that

$$\gamma = \frac{1}{3}$$

and

$$\zeta_1 \to 1 + \overline{h}, \quad \zeta_2 \to -1 + \frac{1}{3}\overline{h}$$

as

$$\overline{h} \to 0$$

.

There is no interval of absolute stability, but the scheme is relatively stable for $\overline{h} > 0$. Using the exact expressions for ζ_1, ζ_2 one finds that the interval of relative stability is $(0, \infty)$.

Conditions on the attainable order of a multistep scheme

A general linear multistep scheme has $2k+2$ coefficients of which $\alpha_k = 1$, and if the scheme is explicit $\beta_k = 0$. For a scheme of order p the first $p+1$ terms in the Taylor expansion of $L(y(x), h)$ must vanish, placing $p+1$ conditions on the coefficients. Thus the highest order that could be attained by a k step method is $2k$ if it is implicit, $2k-1$ if it is explicit. Dahlquist showed that the attainable order of a stable scheme is much lower.

Maximum order of a stable operator

The maximum order of a stable linear k step scheme is $k+1$ if k is odd, or $k+2$ if k is even (Dahlquist, 1956, Math Scand 4, pp. 33-53) (Henrici pp. 229-233). This result is obtained by studying the properties of the associated function

$$\phi(\zeta) = \frac{p(\zeta)}{log\zeta} - \sigma(\zeta)$$

which has a zero of order p at $\zeta = 1$ if the scheme is of order p.

To analyze the properties we map the unit disk in the ζ plane to the left half of the z plane by the transformation

$$z = \frac{\zeta - 1}{\zeta + 1}, \quad \zeta = \frac{1+z}{1-z}.$$

Now instead of the polynomials $p(\zeta)$ and $\sigma(\zeta)$ we consider the functions

$$r(z) = \left(\frac{1-z}{2}\right)^k p\left(\frac{1+z}{1-z}\right), \quad s(z) = \left(\frac{1-z}{2}\right)^k \sigma\left(\frac{1+z}{1-z}\right)$$

which are also polynomials of degree $\leq k$. If $p(\zeta)$ has a root of multiplicity p at $\zeta = \zeta_0$, then $r(z)$ has a root if the same multiplicity at the point $z = (\zeta_0 - 1)/(\zeta_0 + 1)$ unless $\zeta_0 = -1$, in which case the degree of $r(z)$ is reduced to $k - p$. Since $\zeta = 1$ is a simple root of $p(\zeta)$, $zeta = 0$ is a simple root of $p(\zeta)$, $z = 0$ is a simple root of $r(z)$, so

$$r(z) = q_1 z + q_2 z^2 \cdots + q_k z^k$$

where $q_1 \neq 0$, and we can take $q_1 > 0$ by multiplying $p(\zeta)$ by a suitable function. Then

$$a_\mu \geq 0, \quad \mu = 1, 2, \cdots, k$$

because

$$r(z) = a_1 z \Pi(z - x_\lambda) \Pi\left((z - x_\mu)^2 + y\mu^2\right)$$

where λ and μ range over the real and complex roots, and for stability $x_v \leq 0$ for all roots.

Now consider

$$p(z) = \left(\frac{1-z}{2}\right)^k \phi\left(\frac{1+z}{1-z}\right) = \frac{r(z)}{log\frac{1+z}{1-z}} - s(z).$$

$p(z)$ has a zero of order p at $z = 0$ if and only if $\phi(\zeta)$ has a zero of order p at $\zeta = 1$.

Consequently, if the difference scheme has order p, then if we expand

$$\frac{z}{log\frac{1+z}{1-z}}\frac{r(z)}{z} = b_0 + b_1 z + b_2 z^2 \cdots$$

the first p terms must be cancelled by $s(z)$, or

$$s(z) = b_0 + b_1 z \cdots + b_{p+1} z^{p-1}.$$

Since $s(z)$ has degree $\leq k$, the existence of a stable scheme of order $p > k+1$ requires

$$b_{k+1} = b_{k+2} \cdots = b_{p-1} = 0.$$

Now

$$\frac{z}{log\frac{1+z}{1-z}} = c_0 + c_2 z^2 + c_4 z^4 \cdots$$

where

$$c_0 = \frac{1}{2}$$

$$c_2 = -\frac{1}{6}$$

$$c_4 = -\frac{2}{4\zeta}$$

$$c_6 = -\frac{22}{94\zeta}$$

$$c_8 = -\frac{214}{1417\zeta}$$

$$\cdots$$

and it can be proved that $c_{2v} < 0$ when $v \geq 1$. Also, defining $a_v = 0$ for $v > k$, we find that

$$b_0 = c_0 a_1$$

$$b_1 = c_0 a_2$$

$$\cdots$$

$$b_{2v} = c_0 a_{2v+1} + c_2 a_{2v-1} \cdots + c_{2v} a_1$$

$$b_{2v+1} = c_0 a_{2v+2} + c_2 a_{2v} \cdots + c_{2v} a_2$$

If k is odd then

$$b_{k+1} = c_2 a_k + c_4 a_{k-2} \cdots + c_{k+1} a_1$$

and since $a_1 > 0$ and no $a_v < 0$ it follows that

$$b_{k+1} < 0$$

so that the maximum order of a stable scheme cannot exceed $k+1$.

If k is even then

$$b_{k+1} = c_2 a_k + c_4 a_{k-2} \cdots + c_k a_2.$$

A necessary and sufficient condition for $b_{k+1} = 0$ is then

$$a_2 = a_4 \cdots = a_k = 0.$$

Then

$$r(z) = -r(-z)$$

so the roots of $r(z)$ must all lie on the imaginary axis, and consequently all roots of $p(\zeta)$ lie on $|\zeta| = 1$. Also

$$b_{k+2} = c_4 a_{k-1} + c_6 a_{k-3} \cdots + c_{k+2} a_1$$

is negative since $a_1 > 0$ and $a_v \geq 0$. Thus the order cannot exceed $k+2$.

Optimal method for $k = 2$

Then we have

$$p(\zeta) = \zeta^2 - 1$$

since 1 is a simple root and the other root must also lie on the unit circle and must be real. This leads to the Milne scheme

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

Stability on the imaginary axis and A-stability

Jeltsch showed (BIT 18, 1978, pp. 170-174) that a linear multistep scheme which is stable on the imaginary axis is A-stable.

Proof: consider the characteristic polynomials

$$p(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j, \quad \sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j$$

and the characteristic equation

$$p(\zeta) - \overline{h}\sigma(\zeta) = 0 \tag{0.46}$$

where $\overline{h} = \alpha h$. This implies

$$\phi(\zeta) = Re\frac{p(\zeta)}{\sigma(\zeta)} = Re(\overline{h}) \tag{0.47}$$

Thus stability on the imaginary axis implies

$$\phi(\zeta) \neq 0 \tag{0.48}$$

for ζ in Ω, $|\zeta| > 1$

From the continuity of the roots of 0.46, and stability on the imaginary axis, $\sigma(\zeta)$ has no roots in Ω. Otherwise 0.47 would have a root for large imaginary $\overline{h}$.

Hence $\phi(\zeta)$ has no poles in Ω, and therefore it is a continuous real function in Ω, which according to 0.48 is not zero in Ω. Thus either $\phi(\zeta) < 0$ or $\phi(\zeta) > 0$ everywhere in Ω. On the positive real axis

$$\phi(\zeta) = \phi_r(x) = \frac{p(x)}{\sigma(x)}$$

Consistency requires

$$\phi_r(1) = \frac{p(1)}{\sigma(1)} = 0, \quad \phi_r'(1) = \frac{p'(1)}{\sigma(1)} = 1 > 0$$

Hence $\phi(\sigma) > 0$ everywhere in Ω, and since ϕ is continuous,

$$\phi(\zeta) \geq 0 \text{ or } \phi(\zeta) = \infty \text{ when } |\zeta| \geq 1.$$

But then if $Re(\overline{h}) = \phi(\zeta) < 0$ it follows that $|\zeta| < 1$. Thus the scheme is A-stable.

Stability barrier for a second order ODE

Consider the equation

$$\frac{d^2y}{dt^2} = f(t, y) \tag{0.49}$$

A linear multistep method now has the form

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h^2 \sum_{j=0}^{k} \beta_j f_{n+j}$$

with characteristic polynomials

$$p(\zeta) = \sum_{j=0}^{k} \alpha_j \zeta^j, \quad \sigma(\zeta) = \sum_{j=0}^{k} \beta_j \zeta^j$$

Consider the test equation

$$\frac{d^2y}{dt^2} = -\omega^2 y \tag{0.50}$$

The scheme is stable on the imaginary axis if solutions of the difference equations remain bounded for any real ωh. This requires that the characteristic equation

$$p(\zeta) + (\omega h)^2 \sigma(\zeta) = 0$$

has no roots in Ω, $|\zeta| > 1$. At a root

$$\frac{p(\zeta)}{\sigma(\zeta)} = -(\omega h)^2$$

Thus a necessary requirement for stability is that $\frac{p(\zeta)}{\sigma(\zeta)}$ must not be real and non-positive in Ω. It follows that the branch of $\left(\frac{p(\zeta)}{\sigma(\zeta)}\right)^{\frac{1}{2}}$ which takes positive values for $\zeta > 1$ exists everywhere in Ω and satisfies the condition

$$Re\left(\frac{p(\zeta)}{\sigma(\zeta)}\right)^{\frac{1}{2}} > 0 \quad \text{in } \Omega \tag{0.51}$$

If the order of accuracy is p the function

$$p(e^h) - h^2\sigma(e^h) = O(h^{p+2})$$

and setting $h = \log\zeta$,

$$p(\zeta) - (\log\zeta)^2\sigma(\zeta) \approx c(\zeta-1)^{p+2}.$$

Also

$$p(\zeta) \approx \frac{1}{2}p''(1)(\zeta-1)^2$$

Set $c' = 2c/p''(1)$. Then

$$\frac{\sigma(\zeta)}{p(\zeta)} \approx \frac{1}{(\log\zeta)^2}\left(1 - c'(\zeta-1)^p + \cdots\right)$$

and

$$\left(\frac{\sigma(\zeta)}{p(\zeta)}\right)^{\frac{1}{2}} \approx \frac{1}{\log\zeta}\left(1-\frac{1}{2}c'(\zeta-1)^p+\cdots\right) \tag{0.52}$$

where $\log\zeta$ is the branch which is real when ζ is positive and positive when $\zeta > 1$. Transform the unit disk to the left half plane by setting

$$\zeta = \frac{z+1}{z-1}, \quad z = \frac{\zeta+1}{\zeta-1}$$

Also set

$$r(z) = (z+1)^k p\left(\frac{z+1}{z-1}\right), \quad s(z) = (z+1)^k \sigma\left(\frac{z+1}{z-1}\right).$$

Then 0.51 and 0.52 become

$$Re\left(\frac{s(z)}{r(z)}\right)^{\frac{1}{2}} > 0 \text{ for } Re(z) > 0$$

and for $z \to \infty$

$$\left(\frac{s(z)}{r(z)}\right)^{\frac{1}{2}} = \frac{1}{\log\left(\frac{z+1}{z-1}\right)}\left(1-\frac{1}{2}c'\left(\frac{z}{2}\right)^{-p}+\cdots\right).$$

Since

$$\frac{1}{\log\left(\frac{z+1}{z-1}\right)} = \frac{1}{2}z - \frac{1}{6}z^{-1}+\cdots$$

it follows that

$$\left(\frac{s(z)}{r(z)}\right)^{\frac{1}{2}} = \frac{1}{2}z\left(1-\frac{1}{3}z^{-2}+\cdots-\frac{1}{2}c'\left(\frac{z}{2}\right)^{-p}+\cdots\right).$$

It follows from the minimum principle for harmonic functions on a half plane that if $f(z)$ has a positive real part for $Rez > 0$ and a pole at ∞ with residue a, then either

$$Re\,(f(z) - az) > 0 \text{ for } Rez > 0$$

or $f(z) - az$ is an imaginary constant which in this case must be zero. Thus

$$Re\left[\left(\frac{s(z)}{r(z)}\right)^{\frac{1}{2}} - \frac{1}{2}z\right] = Re\left[-\frac{1}{6}z^{-1} - \frac{1}{2}c'\left(\frac{z}{2}\right)^{1-p} + \cdots\right] \geq 0$$

when $Rez > 0$ and $|z|$ is large.

This is impossible if $p > 2$ since $Re\left(-\frac{1}{6}z^{-1}\right) < 0$. Hence $p \leq 2$. If $p = 2$

$$-\frac{1}{6} - c' \geq 0 \text{ or } c' \leq -\frac{1}{6}.$$

If the left side were not identically zero it would behave like a z^{-q} for large z with $q > 1$. But then the real part could not be positive whenever $Rez > 0$. Thus the left side is zero, or

$$\frac{s(z)}{r(z)} = \frac{1}{4}z^2$$

and

$$\frac{p(\zeta)}{\sigma(\zeta)} = \frac{1}{4}\frac{(\zeta+1)^2}{(\zeta-1)^2}$$

It follows that apart from a common factor

$$p(\zeta) = (\zeta - 1)^2, \quad \sigma(\zeta) = \frac{1}{4}(\zeta+1)^2.$$

Then the characteristic equation has the solution

$$\zeta_1 = \frac{1 + \frac{1}{2} i\omega h}{1 - \frac{1}{2} i\omega h}, \quad \zeta_2 = \overline{\zeta_1}$$

and the scheme is stable for all h. This is the most accurate scheme which is stable for all h since it minimizes c'.

This result is due to Dahlquist (BIT 18, 1978, pp. 133-136), Jeltsch (BIT 18, 1978, pp. 170-174) pointed out that it also leads to the result that schemes for first order systems which are stable on the imaginary axis are at most second order accurate, and hence that A-stable schemes are at most second order accurate.

Write 0.50 as a first order system

$$\frac{dy}{dt} = v$$

$$\frac{dv}{dt} = -\omega^2 y.$$

Now a scheme with characteristic polynomials $p(\zeta)$ and $\sigma(\zeta)$ produces a scheme for 0.50 with characteristic polynomials

$$\tilde{p}(\zeta) = p(\zeta), \quad \tilde{\sigma(\zeta)} = \sigma^2(\zeta)$$

so an A-stable scheme with order of accuracy $p > 2$ would generate a scheme for 0.50 with order of accuracy p which is stable for all real ωh, violating the previous theorem. Also the scheme with minimum error constant is generated by the trapezoidal rule

$$p(\zeta) = \zeta - 1, \quad \sigma(\zeta) = \frac{\zeta + 1}{2}.$$

Runge Kutta method of order 2 for ODE's

To solve

$$\frac{dy}{dt} = f(t, y)$$

let y_n be the approximation at $t = nh$ and set

$$y^{(1)} = y_n + hf(t, y_n)$$

$$y_{n+1} = y_n + \frac{h}{2}\left\{f(t, y_n) + f(t + h, y^{(1)})\right\}$$

approximating the trapezoidal rule. Alternatively set

$$y^{(1)} = y_n + hf(t, y_n)$$

$$y_{n+1} = y_n + hf(t + \frac{h}{2}, y^{(1)})$$

approximating the rectangle rule. These are both special cases of

$$y_{n+1} = y_n + h(A_1 g_1 + A_2 g_2)$$

where

$$g_1 = f(t, y_n)$$

$$g_2 = f(t + \alpha h, y_n + \beta h g_1).$$

Now if

$$\frac{dy}{dt} = f(t, y)$$

then

$$\frac{d^2y}{dt^2} = \frac{d}{dt}f(t,y) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \frac{\partial f}{\partial t} + f\frac{\partial f}{\partial y}$$

For second order accuracy we want to approximate

$$\begin{aligned} y(t+h) &= y + h\frac{dy}{dt} + \frac{h^2}{2}\frac{d^2y}{dt^2} + O(h^3) \\ &= y + hf + \frac{h^2}{2}\left(\frac{\partial f}{\partial t} + f\frac{\partial f}{\partial y}\right) + O(h^3) \end{aligned}$$

Substituting g_1 and g_2 we find that

$$\begin{aligned} y_{n+1} &= y_n + hA_1 f(t,y_n) + hA_2\left\{f(t,y_n) + \frac{\partial f}{\partial t}\alpha h + \frac{\partial f}{\partial y}\beta h f(t,y_n)\right\} \\ &= y_n + h(A_1+A_2)f(t,y_n) + h^2A_2\left(\alpha\frac{\partial f}{\partial t} + \beta f(t,y_n)\frac{\partial f}{\partial y}\right) + O(h^3) \end{aligned}$$

Thus we obtain second order accuracy if

$$A_1 + A_2 = 1$$

$$A_2\alpha = \frac{1}{2}$$

$$A_2\beta = \frac{1}{2}$$

The trapezoidal and rectangle rules correspond to

$$A_1 = A_2 = \frac{1}{2}, \quad \alpha = \beta = 1$$

and

$$A_1 = 0, \quad A_2 = 1, \quad \alpha = \beta = \frac{1}{2}$$

Rectangle rule

$$y^{(1)} = y_n + \frac{h}{2} f(t, y_n) = y_n + \frac{h}{2} g_1$$

$$y^{(2)} = y_n + hf\left(t + \frac{h}{2}, y_n + \frac{1}{2} h g_1\right) = y_n + h g_2$$

Trapezoidal rule

$$y^{(1)} = y_n + hf(t, y_n) = y_n + h g_1$$

$$y^{(2)} = y_n + \frac{h}{2}\left(f(t, y_n) + f(t + h, y_n + h g_1)\right) = y_n + \frac{h}{2}(g_1 + g_2)$$

Runge Kutta schemes of higher order

To solve

$$\frac{dy}{dt} = f(t, y)$$

we wish to represent the leading terms of the Taylor expansion

$$y(t + h) = y(t) + h\frac{dy}{dt} + \frac{h^2}{2}\frac{d^2y}{dt^2} + \frac{h^3}{6}\frac{d^3y}{dt^3} + \cdots$$

With the notation

$$\frac{\partial}{\partial t} f(y, t) = f_t, \text{ etc.}$$

one finds that

$$\frac{\partial d^2 y}{dt^2} = \frac{d}{dt} f(t, y) = f_t + f_y \frac{dy}{dt} = f_t + f f_y$$

and

$$\begin{aligned}\frac{d^3y}{dt^3} &= \frac{d}{dt}(f_t + ff_y) = \frac{\partial}{\partial t}(f_t + ff_y) + \frac{\partial}{\partial y}(f_t + ff_y)\frac{dy}{dt} \\ &= f_{tt} + f_t f_y + f f_{yt} + (f_{yt} + f_y^2 + f f_{yy})f \\ &= f_{tt} + f_t f_y + 2f f_{yt} + f f_y^2 + f^2 f_{yy}\end{aligned}$$

Thus

$$\frac{d^2y}{dt^2} = F, \qquad \frac{d^3y}{dt^3} = F f_y + G$$

where

$$F = f_t + f f_y = \left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right) f$$

and

$$G = f_{tt} + 2f f_{yt} + f^2 f_{yy} = \left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right)^2 f$$

Consider

$$y_{n+1} = y_n + h(A_1 g_1 + A_2 g_2 + A_3 g_3)$$

where

$$g_1 = f(t, y_n)$$

$$g_2 = f(t + \alpha_2 h, y_n + \beta_{21} h g_1)$$

$$g_3 = f(t + \alpha_3 h, y_n + \beta_{31} h g_1 + \beta_{32} h g_2)$$

and

$$\alpha_2 = \beta_{21}$$

$$\alpha_3 = \beta_{31} + \beta_{32}$$

Then

$$\begin{aligned} g_2 &= f + \alpha_2 h f_t + \beta_{21} h f f_y + \alpha_2^2 \frac{h^2}{2} f_{tt} + \alpha_2 \beta_{21} h^2 f f_{ty} + \beta_{21}^2 \frac{h^2}{2} f^2 f_{yy} + \cdots \\ &= f + \alpha_2 h F + \frac{1}{2} \alpha_2^2 h^2 G + \cdots \end{aligned}$$

Also

$$\begin{aligned} g_3 &= f + \alpha_3 h f_t + \beta_{31} h f f_y + \beta_{32} h g_2 f_y + \alpha_3^2 \frac{h^2}{2} f_{tt} + \alpha_3 (\beta_{31} + \beta_{32}) h^2 f f_{ty} + (\beta_{31} + \beta_{32})^2 \frac{h^2}{2} f^2 f_{yy} + \cdots \\ &= f + \alpha_3 h (f_t + f f_y) - \beta_{32} h f f_y + \beta_{32} h f f_y + \alpha_2 \beta_{32} h^2 F f_y + \frac{1}{2} \alpha_3^2 h^2 (f_{tt} + 2 f f_{ty} + f^2 f_{yy}) + \cdots \\ &= f + h \alpha_3 F + h^2 (\alpha_2 \beta_{32} F f_y + \frac{1}{2} \alpha_3^2 G) + \cdots \end{aligned}$$

Thus

$$y_{n+1} = y_n + h(A_1 + A_2 + A_3) f + h^2 (A_2 \alpha_2 + A_3 \alpha_3) F + \frac{h^3}{2} \left[2 A_3 \alpha_2 \beta_{32} F f_y + (A_2 \alpha_2^2 + A_3 \alpha_3^2) G \right] + O(h^4)$$

Therefore one needs

$$A_1 + A_2 + A_3 = 1$$

$$A_2 \alpha_2 + A_3 \alpha_3 = \frac{1}{2}$$

$$A_2 \alpha_2^2 + A_3 \alpha_3^2 = \frac{1}{3}$$

Two solutions are

$$A_1 = \frac{2}{9}, A_2 = \frac{3}{9}, A_3 = \frac{4}{9}$$

$$\alpha_2 = \beta_{21} = \frac{1}{2}, \beta_{31} = 0, \alpha_3 = \beta_{32} = \frac{3}{4}$$

and

$$A_1 = \frac{1}{6}, A_2 = \frac{4}{6}, A_3 = \frac{1}{6}$$

$$\alpha_2 = \beta_{21} = \frac{1}{2}, \alpha_3 = 1, \beta_{31} = -1, \beta_{32} = 1$$

If

$$A_1 = A_2 = 0$$

then

$$A_3 = 1, \quad \alpha_3 = \frac{1}{2}, \quad \alpha_3^2 = \frac{1}{3}$$

which is impossible.

Runge Kutta method of order four

To solve

$$\frac{dy}{dt} = f(t, y)$$

the fourth order method is

$$y_{n+1} = y_n + h(A_1 g_1 + A_2 g_2 + A_3 g_3 + A4 g_4)$$

where

$$g_1 = f(t, y_n)$$

$$g_2 = f(t + \alpha_2 h, y_n + \beta_{21} h g_1)$$

$$g_3 = f(t + \alpha_3 h, y_n + \beta_{31} h g_1 + \beta_{32} h g_2)$$

$$g_4 = f(t + \alpha_4 h, y_n + \beta_{41} h g_1 + \beta_{42} h g_2 + \beta_{43} h g_3)$$

and

$$\alpha_r = \sum_{s=1}^{r-1} \beta_{rs}$$

Then we need

$$A_1 + A_2 + A_3 + A_4 = 1$$

$$A_2 \alpha_2 + A_3 \alpha_3 + A_4 \alpha_4 = \frac{1}{2}$$

$$A_2 \alpha_2^2 + A_3 \alpha_3^2 + A_4 \alpha_4^2 = \frac{1}{3}$$

$$A_2 \alpha_2^3 + A_3 \alpha_3^3 + A_4 \alpha_4^3 = \frac{1}{4}$$

The classical solution is

$$y_{n+1} = y_n + \frac{h}{6}(g_1 + 2g_2 + 2g_3 + g_4)$$

where

$$g_1 = f(t, y_n)$$

$$g_2 = f(t + \frac{h}{2}, y_n + \frac{h}{2} g_1)$$

$$g_3 = f(t + \frac{h}{2}, y_n + \frac{h}{2} g_2)$$

$$g_4 = f(t + h, y_n + h g_3)$$

Attainable order of accuracy of Runge Kutta methods

Butcher has established following results (Math Comp 19, 1965, pp. 408-417):

Number of stages	Attainable order of accuracy
$p \leq 4$	p
5	4
6	5
7	6
$p \geq 8$	$\leq p - 2$

Conclusions

Acknowledgments

References