§5.10 Stability (this comes from K & C. B & F is more general)

Recall the general form of a multistep method:

(*) $a_k y_n + a_{k-1} y_{n-1} + \cdots + a_0 y_{n-k} = h\left[b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k}\right]$

where $f_n = f(t_n, y_n)$.

$b_k \neq 0 \Rightarrow$ implicit method (new y_n appears on both sides)

$b_k = 0 \Rightarrow$ explicit method

We associate two polynomials with (*):

$p(z) = a_k z^k + a_{k-1} z^{k-1} + \cdots + a_0$

$q(z) = b_k z^k + b_{k-1} z^{k-1} + \cdots + b_0$

Def (Convergent method): Let $y(h,t)$ be the approx sol obtained by using a numerical method with step size h. The method is said to be convergent if:

$\forall t \in [t_0, t_m]$:

$$\lim_{h\to 0} y(h,t) = y(t)$$

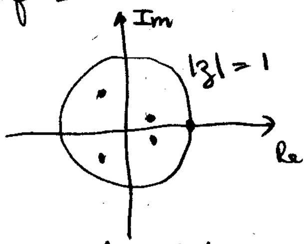
provided the starting values obey same eq, i.e.:

for all n s.t. $0 \le n \le k-1$:

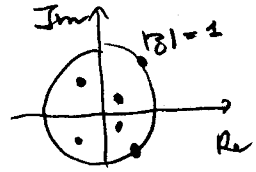
$$\lim_{h \to 0} y(h, t_0 + nh) = y(t_0 + nh).$$

and f satisfies conditions for the problem $\begin{cases} y' = f \\ y(t_0) = \alpha \end{cases}$ to be well posed

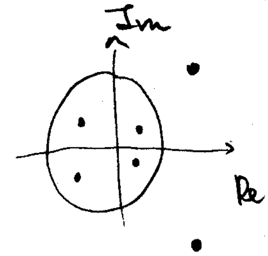
Def (Stability) A multistep method is said to be *stable* if all the roots of $p(z)$ lie in the disk $|z| \leq 1$ and if each root $|z| = 1$ is *simple* (= multiplicity 1)



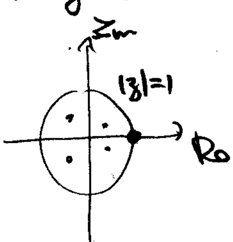
(strongly) stable
only root with $|z| = 1$ is $z = 1$.



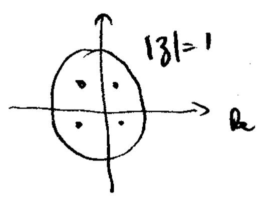
(weakly) stable
more than one root with $|z| = 1$.



unstable



unstable
$p(z) = (z-1)^2 r(z)$



stable
All roots λ, $|\lambda| < 1$

etc...

Def (consistency) A multistep method is said to be *consistent* if:

$$p(1) = 0$$

and $p'(1) = q(1)$ (we will see in a moment where this comes from)

Theorem For multistep methods of general form (*):

$$\boxed{\text{convergent} \iff (\text{stable and consistent})}$$

proof: stable and consistent $\Rightarrow$ convergent is very involved.

- $\boxed{\text{convergent} \Rightarrow \text{stable}}$ (stability is a necessary cdt° for convergence)

Assume method is not stable, we will give a simple problem where method is not convergent.

method not stable $\Rightarrow$ (A) $\exists$ root λ of $p(z)$ with $|\lambda|>1$

or (B) $\exists$ ———— with $|\lambda|=1$ and $p'(\lambda)=0$

(note: $p(\lambda)=0 \Rightarrow p(z)=(z-\lambda)r(z)$

$$p'(z)=r(z)+(z-\lambda)r'(z)$$
$$p'(\lambda)=r(\lambda).$$

thus $p'(\lambda)=0 \iff \lambda$ is a multiple root of p)

Consider the simple IVP:

$$(P1)\begin{cases} y'=0 \\ y(0)=0 \end{cases} \qquad (\text{exact sol is } y(t)=0)$$

Applying (*):

$$a_k y_n + a_{k-1}y_{n-1}+\dots+a_0 y_{n-k}=0 \qquad (1)$$

This is a difference eq and it is relatively easy to come up with sequences satisfying it (see below for a refresher on difference eq.). In particular any sequence of the form:

$$y_n = h\lambda^n, \quad \lambda \text{ root of } p.$$

satisfies the difference eq.

(A) If $|\lambda|>1$:

$$|y(h,nh)| = h|\lambda|^n < h|\lambda|^k \quad \text{for } 0 \leq n \leq k-1$$

thus $|y(h,nh)| \to 0$ (method is convergent in first few steps)

however, if we let $t=nh$ (or $h=t/n$):

$$|y(h,t)|=|y(h,nh)|=h|\lambda|^n=\frac{t}{n}|\lambda|^n \to \infty \text{ as } n\to\infty$$

(method blows up for such a simple problem!)

(3) if $|\lambda| = 1$ and $p'(\lambda) = 0$ a sol to difference eq is:

$$y_n = h\, n\, \lambda^n$$

method is convergent for first few steps since:

$$|y(h, nh)| = h\, n \underbrace{|\lambda|^n}_{=1} = h\, n < h\, k \quad (\text{for } 0 \leq n \leq k-1)$$

$hk \downarrow 0$ as $h \to 0$.

method does not converge after a few steps ($t = nh$, $h = t/n$):

$$|y(h, \underset{nh}{\overset{\shortparallel}{t}})| = \underbrace{h\, n}_{=t} |\lambda|^n = t \neq 0 \quad \text{as } h \to 0.$$

- $\boxed{\text{Convergent} \Rightarrow \text{consistent}}$

Assume method (*) is convergent.

$$(P2) \begin{cases} y' = 0 \\ y(0) = 1 \end{cases}$$

$\rightsquigarrow$ same difference eq $a_k y_n + a_{k-1} y_{n-1} + \dots + a_0 y_{n-k} = 0$ (1)

a sol to (1) is to set $y_0 = y_1 = \dots = y_{k-1} = 1$ and use (1) to find y_n, $n \geq k$.

Since method is convergent:

$\lim_{n \to \infty} y_n = 1$, plugging into (1):

$$\Rightarrow a_k + a_{k-1} + \dots + a_0 = 0$$

$$\Leftrightarrow \boxed{p(1) = 0}$$

Now consider the problem

$$(P3) \begin{cases} y' = 1 \\ y(0) = 0 \end{cases} \qquad (\text{sol is } y(t) = t)$$

we get a new eq:

$$a_k y_n + a_{k-1} y_{n-1} + \dots + a_0 y_{n-k} = h\,[b_k + b_{k-1} + \dots + b_0] \quad (2)$$

Convergent $\Rightarrow$ stable $\Rightarrow$ $p(1) = 0$, $p'(1) \neq 0$ (1 is a <u>simple root</u>)

A solution to (2) is given by:

$$y_n = (n+k)h\gamma, \quad \text{where } \gamma = \frac{q(1)}{p'(1)}.$$

Checking by substitution in LHS of (2):

$$h\gamma(a_k(n+k) + a_{k-1}(n+k-1) + \dots + a_0 n)$$

$$= n h\gamma\underbrace{(a_k + a_{k-1} + \dots + a_0)}_{= p(1) = 0} + h\gamma\underbrace{[k a_k + (k-1)a_{k-1} + \dots + a_1]}_{= p'(1)}$$

$$= h\, q(1) = h\,[b_k + b_{k-1} + \dots + b_0].$$

Now the first few steps are consistent with initial value: $y(0) = 0$:

$$|y(h, nh)| = (n+k)h\gamma \to 0 \text{ as } h \to 0.$$

(since $0 \leqslant n \leqslant k-1$)

Since the method is <u>convergent</u> we must have:

$$\lim_{n\to\infty} y_n = t \quad \text{when } nh = t$$

$$\Leftrightarrow \lim_{n\to\infty}(n+k)h\gamma \underset{\uparrow}{=} \lim_{n\to\infty} \overbrace{nh}^{t}\gamma = t \Rightarrow \gamma = 1$$

$$\Leftrightarrow \boxed{p'(1) = q(1)}$$

since $\lim_{n\to 0} kh\gamma = 0$

<u>Example</u>: Milne's method $y_n - y_{n-2} = h\left[\frac{1}{3}f_n + \frac{4}{3}f_{n-1} + \frac{1}{3}f_{n-2}\right]$

$p(z) = z^2 - 1$ roots: $+1, -1$ (simple) $\Rightarrow$ <u>stable</u>

$q(z) = \frac{1}{3}z^2 + \frac{4}{3}z + \frac{1}{3}$

$p'(z) = 2z$

$\left.\begin{array}{l} q(1) = 2 = p'(1) \\ p(1) = 0 \end{array}\right\}$ <u>consistent</u>

$\Rightarrow$ <u>method is convergent</u>

Difference equation fundamentals (optional)

$x = (x_1, x_2, x_3, \dots)$ are sequences
$y = (y_1, y_2, y_3, \dots)$

A difference eq can be written using the <u>shift operator</u>

$$Ex = (x_2, x_3, x_4, \dots), \text{ where } x = (x_1, x_2, x_3, \dots)$$

It is not hard to see that:

$$(Ex)_n = x_{n+1}$$
$$(E^k x)_n = x_{n+k}$$
$$E^0 x = x$$

A linear difference operator is:

$$L = \sum_{i=1}^{m} a_i E^i$$

A difference eq is of the form:

$$Lx = 0.$$

<u>example:</u> $x_{n+2} - 3x_{n+1} + 2x_n = 0$

$$\Leftrightarrow (E^2 - 3E^1 + 2E^0)x = 0$$
$$\Leftrightarrow p(E)x = 0 \quad \text{where } p(\lambda) = \lambda^2 - 3\lambda + 2$$

<u>Theorem</u> (Simple roots) If p is a poly and λ a root then a sol to $p(E)x = 0$ is $(\lambda, \lambda^2, \lambda^3, \dots)$. If all roots of p are simple $\neq 0$ then all solutions to $p(E)x = 0$ are in the span of all such solutions.

<u>Theorem</u> (Multiple roots) Let p be a poly with $p(0) \neq 0$. Then a basis for nullspace of $p(E)$ is:
with each root λ of p with multiplicity k, associate k solutions:

$$x(\lambda), x'(\lambda), x''(\lambda), \dots, x^{(k-1)}(\lambda), \text{ where } x(\lambda) = (\lambda, \lambda^2, \lambda^3, \dots)$$
$$x'(\lambda) = (1, 2\lambda, 3\lambda^2, \dots)$$
$$x''(\lambda) = (0, 2, 6\lambda, \dots)$$
$$\vdots$$

Local truncation error for Multistep methods

$$\begin{cases} y' = f(t,y) \\ y(a) = \alpha \end{cases}$$

Recall the general form of a k-step method

(*) $a_m y_n + a_{n-1} y_{n-1} + \cdots + a_0 y_{n-k} = h\,[b_k f_n + b_{k-1} f_{n-1} + \cdots + b_0 f_{n-k}]$

here $f_i = f(t_i, y_i)$.

To analyze this method we introduced:

$$Ly = \sum_{i=0}^{k} a_i\, y(ih) - h\, b_i\, y'(ih) = \sum_{j=0}^{\infty} d_j\, y^{(j)}(0)$$

Using Taylor: $\left.\begin{aligned} y(ih) &= \sum_{j=0}^{\infty} \frac{(ih)^j}{j!}\, y^{(j)}(0) \\ y'(ih) &= \sum_{j=0}^{\infty} \frac{(ih)^j}{j!}\, y^{(j+1)}(0) \end{aligned}\right\}$ with $d_j = \sum_{i=0}^{k} \left[\frac{i^j}{j!} a_i - \frac{i^{j-1}}{(j-1)!} b_i \right]$

The following theorem shows that a method of order m has a local truncation error $\mathcal{O}(h^{m+1})$:

Theorem If $y \in C^{m+2}$ and $\frac{\partial f}{\partial y}$ continuous, then assuming $y_i = y(t_i)$ for $i \leq n-1$:

$$y(t_n) - y_n = \left(\frac{d_{m+1}}{a_k}\right) h^{m+1}\, y^{(m+1)}(t_{n-k}) + \mathcal{O}(h^{m+2})$$

where m is the order of the method.

Proof:

$$Ly = \sum_{i=0}^{k} a_i\, y(t_i) - h\, b_i \overbrace{f(t_i, y(t_i))}^{=y'(t_i)} \qquad \text{(true sol)}$$

$$0 = \sum_{i=0}^{k} a_i y_i - h\, b_i f(t_i, y_i) \qquad \text{(approx)}$$

$$Ly = \sum_{i=0}^{k} a_i (y(t_i) - y_i) - h\, b_i (f(t_i, y(t_i)) - f(t_i, y_i)) \quad \leftarrow \text{method exact for all previous steps}$$

$$= a_k (y(t_k) - y_k) - h\, b_k\, [f(t_k, y(t_k)) - f(t_k, y_k)]$$

Apply MVT:

$$Ly = a_k\left(y(t_k)-y_k\right) - h b_k \frac{\partial f}{\partial x}(t_k,\xi)\left(y(t_k)-y_k\right)$$

where $\xi \in (y(t_k); y_k)$ (or other way around)

$$\Rightarrow Ly = \left(a_k - h b_k C\right)\left(y(t_k)-y_k\right) = d_{m+1} h^{m+1} y^{(m+1)}(t_0) + O(h^{m+2})$$

$$\Rightarrow \boxed{y(t_k) - y_k = \frac{d_{m+1} h^{m+1} y^{(m+1)}(t_0)}{a_k - h b_k C} + O(h^{m+2}) = \frac{d_{m+1}}{a_k} h^{m+1} y^{(m+1)}(t_0) + O(h^{m+2})}$$

using taylor $\frac{1}{1+x}$.

Global truncation error: not really sum of all local truncation errors because everytime we apply method we start with (slightly) wrong values of preceding steps. To understand GTE we can look at how much does the solution of a IVP depend on the initial value:

Let $y(t;s)$ be the sol to $\begin{cases} y' = f(t,y) \\ y(0) = s \end{cases}$

$u(t) = \frac{\partial y(t;s)}{\partial s}$ satisfies the IVP:

$$\begin{cases} u' = u\, f_y(t,y) \\ u(0) = 1 \end{cases} \qquad \text{(variational equation)}$$

Theorem (variational eq):

If $|f_y| \leq \lambda$ then the sol to the variational eq satisfies:

$$|u(t)| \leq e^{\lambda t} \qquad (t \geq 0)$$

proof:

$$\frac{u'}{u} = f_y = \lambda - \overbrace{\alpha(t)}^{\geq 0}$$

$\int \downarrow$

$$\ln|u| = \lambda t - \underbrace{\int_0^t \alpha(t)\,dt}_{\geq 0} \Rightarrow \ln|u| \leq \lambda t$$

$$\Rightarrow \boxed{|u| \leq e^{\lambda t}}$$

Theorem (on solution curves for IVP)

If IVP is solved with s and $s+\delta$ then the sol curves differ by at most $|\delta|e^{\lambda t}$:

$\theta \in (0,1)$

proof:

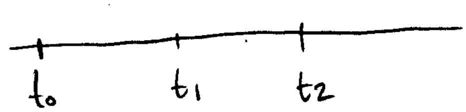
$$|y(t;s) - y(t;s+\delta)| \overset{\text{MVT}}{=} \left| \frac{\partial}{\partial s} y(t, s+\theta\delta) \right| |\delta|$$

$$= |u(t)||\delta| \leq |\delta| e^{\lambda t}$$

Theorem on Global truncation error

If the local truncation errors at $t_0, t_1 \dots t_n$ are $\leq \delta$, then the global truncation error at $t_n \leq \delta \dfrac{e^{n\lambda h} - 1}{e^{\lambda h} - 1}$

proof: let δ_i be the local truncation error at t_i $(\delta_0 = 0)$



at t_1: LTE $\leq |\delta_1|$

at t_2: LTE $\leq |\delta_2|$

error in t_1 can affect sol at most $|\delta_1| e^{\lambda h}$

$\Rightarrow$ Global truncation error at t_2 is $\leq |\delta_2| + |\delta_1| e^{\lambda h}$

at t_3: Global truncation error $\leq \underbrace{|\delta_3|}_{\text{LTE}} + (\quad) e^{\lambda h}$

$$= |\delta_1| e^{2\lambda h} + |\delta_2| e^{\lambda h} + |\delta_3|$$

at t_n:

$$\leq \sum_{k=1}^{n} |\delta_k| e^{(n-k)\lambda h}$$

$$\leq \delta \sum_{k=0}^{n-1} e^{k\lambda h} = \delta \frac{1 - e^{n\lambda h}}{1 - e^{\lambda h}}$$

Theorem: If the LTE is $\mathcal{O}(h^{m+1})$ then the GTE is $\mathcal{O}(h^m)$.

proof: $LTE = \mathcal{O}(h^{m+1})$ means $\delta = \mathcal{O}(h^{m+1})$ in preceding theorem

$$\Rightarrow \quad GTE = \mathcal{O}(h^{m+1}) \underbrace{\frac{1-e^{\lambda nh}}{1-e^{\lambda h}}}_{= \frac{\mathcal{O}(nh)}{\mathcal{O}(h)} = \mathcal{O}(h^{-1})} = \mathcal{O}(h^m)$$

5.9 Higher Order equations and systems

The general form of an m-th order system is:

$$\begin{cases} \frac{dy_1}{dt}(t) = f_1(t, y_1, \dots, y_m) \\ \frac{dy_2}{dt}(t) = f_2(t, y_1, \dots, y_m) \\ \vdots \\ \frac{dy_m}{dt}(t) = f_m(t, y_1, \dots, y_m) \\ \left.\begin{array}{l} y_1(a) = \alpha_1 \\ y_2(a) = \alpha_2 \\ \vdots \\ y_m(a) = \alpha_m \end{array}\right\} \text{I.C.} \end{cases} \quad \text{for } t \in [a,b]$$

Or in vector form:

$$(*) \begin{cases} \frac{d\underline{y}}{dt} = \underline{f}(t, \underline{y}) & \text{for } t \in [a,b] \\ \underline{y}(a) = \underline{\alpha} \end{cases}$$

where $\underline{y} = (y_1, y_2, \dots, y_m)^T$

$\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$

$$\underline{f} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$$
$$(t, \underline{y}) \to \begin{pmatrix} f_1(t, \underline{y}) \\ f_m(t, \underline{y}) \end{pmatrix}$$

Def (Lipschitz): $f(t, \underline{y})$ satisfies a Lipschitz condition on the variable y on the set:

$$D = \{(t, \underline{y}) \mid a \le t \le b, \quad \underline{y} \in \mathbb{R}^m\}$$

if $\exists L > 0$ s.t. $\forall (t, \underline{y})$; $(t, \underline{z}) \in D$

$$|f(t, \underline{y}) - f(t, \underline{z})| \le L \|\underline{y} - \underline{z}\|_1 = L \sum_{i=1}^{m} |y_i - z_i|$$

Note: If $\left|\frac{\partial f(t, \underline{y})}{\partial y_i}\right| \le L$ for $i = 1..m$ then f satisfies a Lipschitz condition with constant L.

Theorem Let $D = \{(t, \underline{y}) \mid t \in [a,b], \ \underline{y} \in \mathbb{R}^m\}$ and let $f_i(t, \underline{y})$ be continuous on D and satisfy a Lipschitz cond there. (for $i = 1, .., m$). Then the system $(*)$ has a unique sol for $t \in [a,b]$.

Numerical methods to solve systems are generalizations of methods for ODEs. So all methods we've seen can be used with systems. The generalization is straightforward if we stick to vector notation:

For example RK4 for systems:

$h = (b-a)/N$
for $i = 0, \dots, N-1$

$$\underline{F_1} = h \underline{f}(t_i, \underline{y}_i)$$
$$\underline{F_2} = h \underline{f}\left(t_i + \frac{h}{2}, \underline{y}_i + \frac{1}{2}\underline{F_1}\right)$$
$$\underline{F_3} = h \underline{f}\left(t_i + \frac{h}{2}, \underline{y}_i + \frac{1}{2}\underline{F_2}\right)$$
$$\underline{F_4} = h \underline{f}\left(t_i + h, \underline{y}_i + \underline{F_3}\right)$$
$$\underline{y}_{i+1} = \underline{y}_i + \frac{1}{6}\left(\underline{F_1} + 2\underline{F_2} + 2\underline{F_3} + \underline{F_4}\right)$$

Any differential eq of the form:

$$y^{(n)} = f(t, y, y', y^{(2)}, \dots, y^{(n-1)})$$

can be transformed into a first order system. First we introduce the var:

$$y_1 = y, \quad y_2 = y', \quad y_3 = y'', \quad \dots, \quad y_n = y^{(n-1)}$$

Then the new variables satisfy the system:

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ y_3' = y_4 \\ \vdots \\ y_n' = f(t, y_1, y_2, \dots, y_n) \end{cases}$$