KTH Computer Science
and Communication

Övning 4

Convergence of One-step Methods

We want to solve the ODE

$$\frac{dy}{dt} = f(t,y), \qquad y(0) = y_0 \tag{1}$$

with a one-step method in a fixed interval $0 \le t \le T$. The general form of a one-step method is

$$u_{n+1} = u_n + h\phi(h, t_n, u_n, u_{n+1}), \qquad u_0 = y_0, \tag{2}$$

where ϕ obviously depends on f. (Ex. for forward Euler, $\phi = f$.) If $\phi = \phi(h, t_n, u_n)$ the method is explicit, otherwise implicit. The approximate solution u_n actually also depends on h and when we need to be more precise we write $u_{n,h}$. For simplicity we only consider the case of a constant timestep, $t_n = nh$.

Consistency

Let $y(t)$ be the exact solution to the ODE. Then the *local truncation error*, $\tau_{n,h}$ is defined as the residual when $y(t_n)$ is entered into (2) instead of u_n, scaled by h,

$$y(t_{n+1}) = y(t_n) + h\phi(h, t_n, y(t_n), y(t_{n+1})) + h\tau_{n,h}.$$

The quantity $h\tau_{n,h}$ corresponds to the error made in one single step with the scheme, when starting from $y(t_n)$. If

$$\max_n |\tau_{n,h}| = \mathcal{O}(h^p), \qquad p \ge 1,$$

the method is said to be *consistent* with an *order of accuracy* p. Typically, hence,

$$\max_n |\tau_{n,h}| \le Ch^p,$$

where C depends on the size of the derivatives of the solution in the interval.

Example. For Forward Euler we have

$$y(t_{n+1}) = y(t_n) + hy'(t_n) + \frac{h^2}{2}y''(\xi) = y(t_n) + hf(y(t_n)) + \frac{h^2}{2}y''(\xi), \qquad \xi \in [t_n, t_{n+1}].$$

Hence, $\tau_{n,h} = hy''(\xi)/2$ and if $M = \max_{0\le t\le T} |y''(t)|$ we can bound

$$\max_{0\le n\le N_h} |\tau_{n,h}| \le \frac{1}{2}Mh,$$

where $T = hN_h$ defines N_h.

We also define the *global truncation error* as

$$e_{n,h} := y(t_n) - u_n, \tag{3}$$

hence the total error at $t = t_n$ when the timestep is h. We would expect that $e_{n,h}$ should be roughly the sum of the errors made in each step, $h\tau_n$, and since we take $\mathcal{O}(h^{-1})$ steps in the scheme, it should be of the order $\mathcal{O}(h^p)$. This is indeed true as long as h is small enough and the scheme remains stable, which will be the case for one-step methods. In that case consistency with an order of accuracy p is equivalent to $|e_{n,h}|$ being $\mathcal{O}(h^p)$, with $p \geq 1$.

Convergence

We say that the method in (2) is *convergent* if, for every ODE (1), with a Lipschitz function f, and for all $T > 0$,

$$\lim_{h\to 0} \max_{0\leq n\leq N_h} ||e_{n,h}|| = 0, \qquad T = hN_h.$$

We have the following convergence theorem.

Theorem 1 *Suppose ϕ is Lipschitz continuous in its last two arguments, uniformly in h and t, i.e. there are constants L and $\delta < 1$ such that for all $0 \leq t \leq T$, and $0 < h \leq \frac{1-\delta}{L}$,*

$$||\phi(h, t, u_n, u_{n+1}) - \phi(h, t, v_n, v_{n+1})|| \leq L\Big[||u_n - v_n|| + ||u_{n+1} - v_{n+1}||\Big]. \tag{4}$$

Then,

$$\max_{0\leq n\leq N_h} ||u_{n,h} - y(t_n)|| \leq C \max_{0\leq n\leq N_h} |\tau_{n,h}|, \qquad T = hN_h, \tag{5}$$

where C is a constant that depends on T, L and δ but not on h or $y(t)$. Therefore, if the method is consistent with an order of accuracy p it is convergent and

$$\max_{0\leq n\leq N_h} ||u_{n,h} - y(t_n)|| = \mathcal{O}(h^p), \qquad T = hN_h. \tag{6}$$

Proof: Let us drop the h subscript in the local and global truncation errors, writing τ_n and $e_n = y(t_n) - u_n$. Then

$$\begin{aligned} e_{n+1} &= y_{n+1} - u_{n+1} = y(t_n) + h\phi(h, t_n, y(t_n), y(t_{n+1})) + h\tau_n - u_n - h\phi(h, t_n, u_n, u_{n+1}), \\ &= e_n + h\Big[\phi(h, t_n, y(t_n), y(t_{n+1})) - h\phi(h, t_n, u_n, u_{n+1})\Big] + h\tau_n. \end{aligned}$$

By the Lipschitz condition (4),

$$\begin{aligned} ||e_{n+1}|| &\leq ||e_n|| + hL\Big[||y(t_n) - u_n|| + ||y(t_{n+1}) - u_{n+1}||\Big] + h \max_{0\leq n\leq N_h} |\tau_n| \\ &= (1 + hL)||e_n|| + hL||e_{n+1}|| + h|\tau|_\infty, \end{aligned}$$

where we have defined

$$|\tau|_\infty = \max_{0\leq n\leq N_h} |\tau_n|.$$

Since $h \leq \frac{1-\delta}{L} < \frac{1}{L}$, we can subtract $hL||e_{n+1}||$ from both sides and divide by $1 - hL$,

$$||e_{n+1}|| \leq \beta||e_n|| + \frac{h}{1 - hL}|\tau|_\infty, \qquad \beta = \frac{1 + hL}{1 - hL}. \tag{7}$$

We claim that (7) implies

$$||e_n|| \leq \frac{1}{2L}(\beta^n - 1)|\tau|_\infty. \tag{8}$$

This holds trivially for $n = 0$ since $||e_0|| = ||y_0 - u_0|| = 0$. Suppose that (8) holds for $n = 0, \ldots, p$. Then, by (7) and (8),

$$
\begin{aligned}
||e_{p+1}|| &\leq \{\text{use (7)}\} \leq \beta||e_p|| + \frac{h}{1-hL}|\tau|_\infty \leq \{\text{use (8)}\} \leq \frac{\beta}{2L}(\beta^p - 1)|\tau|_\infty + \frac{h}{1-hL}|\tau|_\infty \\
&= \frac{\beta^{p+1}}{2L}|\tau|_\infty + \frac{1}{2L}\left(\frac{2hL}{1-hL} - \beta\right)|\tau|_\infty = \frac{\beta^{p+1}}{2L}|\tau|_\infty + \frac{1}{2L}\left(\frac{2hL}{1-hL} - \frac{1+hL}{1-hL}\right)|\tau|_\infty \\
&= \frac{1}{2L}\left(\beta^{p+1} - 1\right)|\tau|_\infty.
\end{aligned}
$$

Hence, (8) then holds also for $n = p + 1$ and the claim follows by induction.

We note that (8) is of the same form as the inequality that we want to prove (5), provided that β^n is bounded by a constant when $n \leq N_h$, i.e. $hn \leq T$. Therefore, we try to estimate how fast β^n grows with n. We get

$$
\beta = \frac{1+hL}{1-hL} = 1 + \frac{2hL}{1-hL} \leq \left\{1 + x \leq e^x\right\} \leq \exp\left(\frac{2hL}{1-hL}\right),
$$

and, consequently,

$$
\beta^n \leq \exp\left(\frac{2hLn}{1-hL}\right) = \exp\left(\frac{2Lt_n}{1-hL}\right) \leq \left\{h \leq \frac{1-\delta}{L}\right\} \leq \exp\left(\frac{2Lt_n}{\delta}\right) \leq \exp\left(\frac{2LT}{\delta}\right).
$$

Inserting this estimate of β^n in (8) gives the result (5) with

$$
C = \frac{1}{2L}\left(\exp\left(\frac{2LT}{\delta}\right) - 1\right).
$$

If the method is consistent with an order of accuracy p, then $|\tau|_\infty = \mathcal{O}(h^p)$ and (6) follows from (5) since C is independent of h. $\square$

Remark. The Lipschitz condition (4) follows almost always directly from the fact that f itself is Lipschitz. For instance, for the trapezoidal rule, $\phi(h, t, u_n, u_{n+1}) = [f(u_n) + f(u_{n+1})]/2$ and

$$
\begin{aligned}
||\phi(h, t, u_n, u_{n+1}) - \phi(h, t, v_n, v_{n+1})|| &\leq \frac{1}{2}||f(u_n) - f(v_n)|| + \frac{1}{2}||f(u_{n+1}) - f(v_{n+1})|| \\
&\leq \frac{L_f}{2}\Big[||u_n - v_n|| + ||u_{n+1} - v_{n+1}||\Big],
\end{aligned}
$$

where L_f is the Lipschitz constant for f. In practice therefore "almost all" reasonable consistent one-step methods are convergent, in particular, forward/backward Euler, the trapezoidal rule and Runge-Kutta methods.