Initial-Value Problems for Ordinary Differential Equations

Simon Fraser University – Surrey Campus

MACM 316 – Spring 2005

Instructor: Ha Le

Overview

- Problem Description
- Existence and Uniqueness
- Euler's Method
- Higher-Order Taylor Methods
- Runge-Kutta Methods
- Error Control and the Runge-Kutta-Fehlberg Method
- Multistep Methods
- Stability

Initial-Value Problems for ODE's

Problem. Approximating the solution $y(t)$ to the ODE

$$y'(t) = f(t, y), \quad a \leq t \leq b,$$

subject to an initial condition

$$y(a) = \alpha.$$

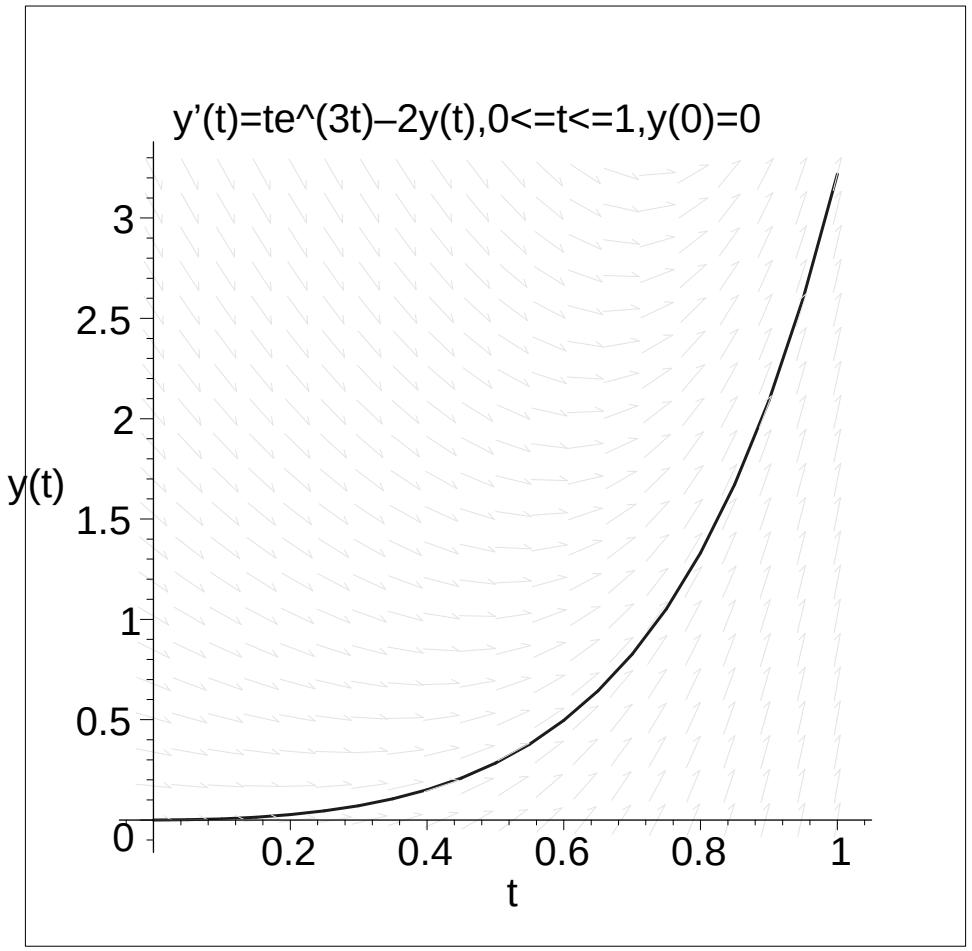
Two components of approximating methods:

- **time step**: find a suitable discrete set of t points

  $$t_0 < t_1 < \cdots < t_N$$

  for evaluation of the solution;

- **solution step**: compute a set of values $w_0, w_1, \ldots, w_N$ such that w_i approximates the exact value of the solution at t_i, that is, approximate $y(t_i)$.


y'(t)=te^(3t)–2y(t),0<=t<=1,y(0)=0
y(t)
t
0
0.5
1
1.5
2
2.5
3
0.2
0.4
0.6
0.8
1

Solution Step

$$y'(t) = f(t, y), \quad a \le t \le b, \;\; y(a) = \alpha. \tag{1}$$

General technique. differential equation $\Longrightarrow$ difference equation.

Single-step methods and multi-step methods.

Example 1 *(single-step method)*

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hf(t_i, w_i), \;\; i = 0, 1, \ldots, N-1. \end{aligned}$$

Example 2 *(multi-step method)*

$$w_0 = \alpha,\ w_1 = \alpha_1,\ w_2 = \alpha_2,\ w_3 = \alpha_3,$$
$$w_{i+1} = w_i + \tfrac{h}{24}(55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) +$$
$$37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})),\ i = 3, 4, \ldots, N-1.$$

Explicit methods and implicit methods.

Example 3 *(implicit methods)*

$$w_{i+1} = w_i + \frac{h}{2}(F(t_i, w_i) + F(t_{i+1}, w_{i+1})).$$

Existence and Uniqueness

Theorem 1

$$\left.\begin{array}{l} D = \{(t,y) \,|\, a \le t \le b,\ -\infty < y < \infty\} \\ f(t,y) \text{ continuous on } D \\ f \text{ satisfies } \textit{Lipschitz condition} \text{ on } D \text{ in } y \end{array}\right\} \begin{array}{l} \text{the IVP (1)} \\ \text{has a unique solution} \\ y(t) \text{ for } a \le t \le b. \end{array}$$

Definition 1 *A function $f(t,y)$ satisfies a* **Lipschitz condition** *in y on $D \in \mathbb{R}^2$ if a constant $L > 0$ exists with*

$$|f(t,y_1) - f(t,y_2)| \le L\,|y_1 - y_2|\,,$$

whenever (t,y_1), $(t,y_2) \in D$.

The constant L is called a **Lipschitz constant** *for f.*

Example 4 Let

$$D = \{(t, y) \,|\, 1 \le t \le 2,\ -3 \le y \le 4\}, \text{ and } f(t, y) = t|y|.$$

$$|f(t, y_1) - f(t, y_2)| = |t|y_1| - t|y_2|| = |t||y_1| - |y_2|| \le \underbrace{2}_{L} |y_1 - y_2|.$$

Hence, f satisfies a Lipschitz condition on D in the variable y.

Example 5 Consider the IVP

$$y'(t) = 1 + t\sin(ty), \ \ 0 \le t \le 2, \ \ y(0) = 0.$$

When $y_1 < y_2$, it follows from the Mean Value Theorem that

$$\exists\, \xi \in (y_1, y_2) \text{ s.t. } \frac{f(t, y_2) - f(t, y_1)}{y_2 - y_1} = \frac{\partial}{\partial y} f(t, \xi) = t^2 \cos(\xi t).$$

Hence, $|f(t, y_2) - f(t, y_1)| = |y_2 - y_1||t^2 \cos(\xi t)| \le 4|y_2 - y_1|$. Since $f(t, y)$ is continuous when $0 \le t \le 2$, and $-\infty < y < \infty$, Theorem 1 implies that a unique solution exists to this IVP.

The following theorem provides *sufficient conditions* for a Lipschitz condition to hold.

Theorem 2 *Suppose $f(t, y)$ is defined on a convex set $D \in \mathbb{R}^2$. If a constant $L > 0$ exists with*

$$\left|\frac{\partial f}{\partial y}(t, y)\right| \leq L, \quad \forall\, (t, y) \in D, \tag{2}$$

then f satisfies a Lipschitz condition on D in the variable y with Lipschitz constant L.

Definition 2 *A set $D \in \mathbb{R}^2$ is said to be* convex *if whenever (t_1, y_1) and (t_2, y_2) belong to D and λ is in $[0, 1]$, the point*

$$((1 - \lambda)t_1 + \lambda t_2,\ (1 - \lambda)y_1 + \lambda y_2)$$

also belongs to D.

Exercise 1 For any constant a and b, the set

$$D = \{(t, y) \,|\, a \leq t \leq b,\ -\infty < y < \infty\} \text{ is convex.}$$

Definition 3 *The IVP*

$$y'(t) = f(t, y), \quad a \leq t \leq b, \;\; y(a) = \alpha,$$

is a well-posed problem *if:*

1. *A unique solution,* $y(t)$*, to the problem exists;*
2. $\forall\, \varepsilon > 0, \exists\, k(\varepsilon) > 0$ *s.t. if* $|\varepsilon_0| < \varepsilon$ *and* $\delta(t)$ *is continuous with* $|\delta(t)| < \varepsilon$ *on* $[a, b]$*, a unique solution,* $z(t)$*, to the* perturbed problem

$$z'(t) = f(t, z) + \delta(t), \quad a \leq t \leq b, \;\; z(a) = \alpha + \varepsilon_0,$$

exists with

$$|z(t) - y(t)| < k(\varepsilon)\varepsilon, \;\; \forall\, a \leq t \leq b.$$

Theorem 3 *Suppose* $D = \{(t, y) \mid a \leq t \leq b \text{ and } -\infty < y < \infty\}$. *If* f *is continuous and satisfies a Lipschitz condition in the variable* y *on the set* D*, then the IVP (1) is well-posed.*

Example 6 Let $D = \{(t, y) \mid 0 \leq t \leq 1,\ -\infty < y < \infty\}$. Consider the IVP

$$y'(t) = y - t^2 + 1, \ \ 0 \leq t \leq 2, \ \ y(0) = 0.5.$$

Since

$$\left|\frac{\partial(y - t^2 + 1)}{\partial y}\right| = |1| = 1,$$

Since D is convex (Exercise 1), Theorem 2 implies that $f(t, y)$ satisfies a Lipschitz condition. Since f is continuous on D, Theorem 3 implies that the problem is well-posed.

Euler's Method

Time step. $t_i = a + ih, \quad \text{for each} \;\; i = 0, 1, 2, \ldots, N.$

Step size: $h = (b - a)/N.$

Solution step.

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + (t_{i+1} - t_i) y'(t_i) + \frac{(t_{i+1} - t_i)^2}{2} y''(\xi_i), \;\; \xi_i \in (t_i, t_{i+1}) \\ &= y(t_i) + h y'(t_i) + \frac{h^2}{2} y''(\xi_i) \\ &= y(t_i) + h f(t_i, y(t_i)) + \frac{h^2}{2} y''(\xi_i). \end{aligned} \tag{3}$$

Euler's method.

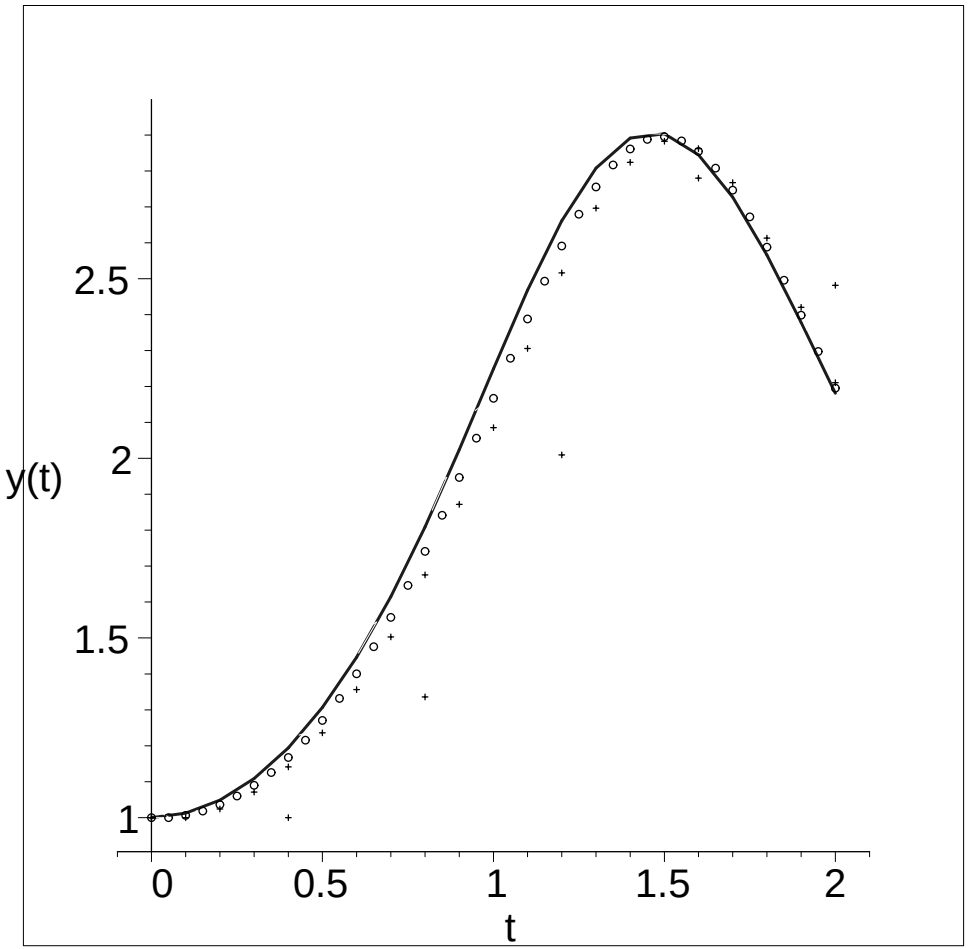
$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + h f(t_i, w_i), \;\; i = 0, 1, \ldots, N - 1. \end{aligned}$$

Example 7

$$y'(t) = \underbrace{y(t)(2.5t - t^2\sqrt{y(t)}}_{f(t,y(t))}, \;\; y(0) = 1, \; 0 \leq t \leq 2, \; h = 0.4.$$

$$\begin{aligned} w_0 &= 1.0, \\ w_{i+1} &= w_i + hf(t_i, w_i), \;\; i = 0, 1, 2, 3, 4. \end{aligned}$$

i	0	1	2	3	4	5
t_i	0	0.4	0.8	1.2	1.6	2.0
w_i	1.0	1.0	1.34	2.01	2.78	2.48


y(t)
2.5
2
1.5
1
0
0.5
1
1.5
2
t

Euler's Method – Error Bound

Lemma 1

$$\forall\, x \geq -1, \forall\, m \in \mathbb{N} \setminus \{0\},\ 0 \leq (1+x)^m \leq e^{mx}.$$

Proof.

$$e^x = 1 + x + \frac{1}{2}x^2 e^{\xi}, \quad \xi \text{ between } x \text{ and } 0$$

$$\Longrightarrow \quad 0 \leq 1 + x \leq 1 + x + \frac{1}{2}x^2 e^{\xi} = e^x$$

$$\Longrightarrow \quad (1+x)^m \leq (e^x)^m = e^{mx}.$$

■

Lemma 2 *If*

1. $s,\ t \in \mathbb{Z}^+$,
2. *the sequence* $\{a_i\}_{i=0}^k$ *satisfies* $a_0 \geq -t/s$, *and*
3. $a_{i+1} \leq (1+s)a_i + t$, *for each* $i = 0, 1, 2, \ldots, k$, *then*

$$a_{i+1} \leq e^{(i+1)s}\left(a_0 + \frac{t}{s}\right) - \frac{t}{s}. \tag{4}$$

Proof.

$$\begin{aligned} a_{i+1} &\leq (1+s)a_i + t \leq (1+s)((1+s)a_{i-1} + t) + t \leq \cdots \\ &\leq (1+s)^{i+1}a_0 + \left(1 + (1+s) + (1+s)^2 + \cdots + (1+s)^i\right) t \\ &= (1+s)^{i+1}\left(a_0 + \frac{t}{s}\right) - \frac{t}{s}. \end{aligned}$$

Applying Lemma 1 with $x = s$ gives (4). ■

Theorem 4 *Suppose f is continuous and satisfies a Lipschitz condition with constant L on*

$$D = \{(t, y) \,|a \leq t \leq b,\ -\infty < y < \infty\},$$

and that a constant M exists with

$$|y''(t)| \leq M, \quad \forall\, t \in [a, b].$$

Let $y(t)$ denote the unique solution to the IVP

$$y'(t) = f(t, y),\ \ a \leq t \leq b,\ \ y(a) = \alpha,$$

and $w_0, w_1, \ldots, w_N$ be the approximations generated by Euler's method for some positive integer N. Then, for each $i = 0, 1, 2, \ldots, N$,

$$|y(t_i) - w_i| \leq \frac{hM}{2L}\left(e^{L(t_i - a)} - 1\right).$$

Proof.

- $i = 0$, $y(t_0) = w_0 = \alpha$, $t_0 - a = 0$, $\implies 0 \leq 0$.
- For $i = 0, 1, \ldots, N-1$, writting $y_i = y(t_i)$, $y_{i+1} = y(t_{i+1})$:

$$\begin{aligned} y_{i+1} &\overset{(3)}{=} y_i + hf(t_i, y_i) + \frac{h^2}{2} y''(\xi_i), \\ w_{i+1} &= w_i + hf(t_i, w_i). \end{aligned}$$

Hence,

$$\begin{aligned} y_{i+1} - w_{i+1} &= y_i - w_i + h(f(t_i, y_i) - f(t_i, w_i)) + \frac{h^2}{2} y''(\xi_i) \\ |y_{i+1} - w_{i+1}| &\leq |y_i - w_i| + h|f(t_i, y_i) - f(t_i, w_i)| + \frac{h^2}{2} |y''(\xi_i)| . \end{aligned}$$

Since f satisfies a Lipschitz condition in the second variable with constant L, and since $|y''(t)| \leq M$,

$$|y_{i+1} - w_{i+1}| \leq (1 + hL)|y_i - w_i| + \frac{h^2 M}{2}.$$

Let

$$s = hL, \;\; t = \frac{h^2 M}{2}, \;\; a_i = |y_i - w_i|.$$

By Lemma 2,

$$\underbrace{|y_{i+1} - w_{i+1}|}_{a_{i+1}} \leq e^{(i+1)\underbrace{hL}_{s}} \left(\underbrace{|y_0 - w_0|}_{a_0} + \underbrace{\frac{h^2 M}{2hL}}_{t/s} \right) - \frac{h^2 M}{2hL}.$$

Since $|y_0 - w_0| = 0$, and $(i+1)h = t_{i+1} - t_0 = t_{i+1} - a$, we have

$$|y_{i+1} - w_{i+1}| \leq \frac{hM}{2L} \left(e^{(t_{i+1} - a)L} - 1 \right),$$

for each $i = 0, 1, \ldots, N-1$. ■

Example 8 For the IVP

$$y' = y - t^2 + 1, \;\; 0 \le t \le 2, \;\; y(0) = 0.5,$$

the exact solution is

$$y(t) = (t+1)^2 - \frac{1}{2}e^t.$$

Hence,

$$y''(t) = 2 - 0.5e^t, \;\; |y''(t)| \le 0.5e^2 - 2, \; t \in [0, 2].$$

With $h = 0.2$, $L = 1$ (see Example 6), and $M = 0.5e^2 - 2$,

$$\begin{aligned} |y_i - w_i| &\le \frac{hM}{2L}\left(e^{(t_i - a)L} - 1\right) \\ &= 0.1(0.5e^2 - 2)(e^{t_i} - 1). \end{aligned}$$

Remark 1 If $\partial f/\partial t$ and $\partial f/\partial y$ both exist, the chain rule for partial differentiation implies that

$$y''(t) = \frac{dy'}{dt}(t) = \frac{df}{dt}(t, y(t)) = \frac{\partial f}{\partial t}(t, y(t)) + \frac{\partial f}{\partial y}(t, y(t)) \cdot f(t, y(t)).$$

This provides a way to obtain an error bound for $y''(t)$ without explicitly knowing $y(t)$.

Local Truncation Error

Definition 4 *The difference method*

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + h\phi(t_i, w_i), \quad i = 0, 1, \ldots, N-1 \end{aligned}$$

has local truncation error

$$\tau_{i+1}(h) = \frac{y_{i+1} - (y_i + h\phi(t_i, y_i))}{h} = \frac{y_{i+1} - y_i}{h} - \phi(t_i, y_i),$$

for each $i = 0, 1, \ldots, N-1$.

Example 9 For Euler's method, $\phi(t_i, w_i) = f(t_i, w_i)$, and

$$\begin{aligned} \tau_{i+1}(h) &= \frac{y_{i+1} - y_i}{h} - f(t_i, y_i), \quad i = 0, 1, \ldots, N-1 \\ &\overset{(3)}{=} \frac{h}{2} y''(\xi_i), \quad \xi_i \in (t_i, t_{i+1}). \end{aligned}$$

On $[a, b]$, if $|y''(t)| \leq M$, then

$$|\tau_{i+1}(h)| \leq \frac{h}{2} M,$$

and the local truncation error in Euler's method is $O(h)$.

Higher-Order Taylor Methods

Time step. $t_i = a + ih, \quad \text{for each} \;\; i = 0, 1, 2, \ldots, N.$

Step size: $h = (b - a)/N.$

Solution step.

$$\begin{aligned} y(t_{i+1}) &= y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + \cdots + \frac{h^n}{n!}y^{(n)}(t_i) \\ &\quad + \frac{h^{n+1}}{(n+1)!}y^{(n+1)}(\xi_i), \;\; \xi_i \in (t_i, t_{i+1}). \end{aligned}$$

Since

$$y'(t) = f(t, y(t)), \;\; y''(t) = f'(t, y(t)), \;\; \ldots, \;\; y^{(k)}(t) = f^{(k-1)}(t, y(t)),$$

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2} f'(t_i, y(t_i)) + \cdots + \frac{h^n}{n!} f^{(n-1)}(t_i, y(t_i)) + \frac{h^{n+1}}{(n+1)!} f^{(n)}(\xi_i, y(\xi_i)). \quad (5)$$

Taylor method of order n.

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hT^{(n)}(t_i, w_i), \quad i = 0, 1, \ldots, N-1, \end{aligned}$$

where

$$T^{(n)}(t_i, w_i) = f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i) + \cdots + \frac{h^{n-1}}{n!} f^{(n-1)}(t_i, w_i).$$

Example 10 Apply Taylor's method of order two to the IVP

$$y' = y - t^2 + 1, \quad 0 \le t \le 2, \quad y(0) = 0.5.$$

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hT^{(2)}(t_i, w_i), \quad i = 0, 1, \ldots, N-1, \end{aligned} \tag{6}$$

where

$$T^{(2)}(t_i, w_i) = f(t_i, w_i) + \frac{h}{2} f'(t_i, w_i).$$

Since $f'(t, y(t)) = \dfrac{d}{dt}(y - t^2 + 1) = y' - 2t = y - t^2 - 2t + 1,$

$$\begin{aligned} T^{(2)}(t_i, w_i) &= w_i - t_i^2 + 1 + \frac{h}{2}(w_i - t_i^2 - 2t_i + 1) \\ &= \left(1 + \frac{h}{2}\right)(w_i - t_i^2 + 1) - ht_i. \end{aligned}$$

$$w_0 \xrightarrow{(6)} w_1 \xrightarrow{(6)} w_2 \xrightarrow{(6)} w_3 \xrightarrow{(6)} w_4$$

Theorem 5 *If Taylor's method of order n is used to approximate the solution to*

$$y'(t) = f(t, y(t)), \;\; a \leq t \leq b, \;\; y(a) = \alpha,$$

with step size h and if $y \in C^{n+1}[a, b]$, then the local truncation error is $O(h^n)$.

Proof.

$$\tau_{i+1}(h) = \frac{y_{i+1} - y_i}{h} - T^{(n)}(t_i, y_i) \overset{(5)}{=} \frac{h^n}{(n+1)!} f^{(n)}(\xi_i, y(\xi_i)),$$

for $i = 0, 1, \ldots, N-1$.

Since $y \in C^{n+1}[a, b]$, $y^{(n+1)}(t) = f^{(n)}(t, y(t))$ is bounded on $[a, b]$. ■

Runge-Kutta Methods

- Taylor methods are seldom used in practice because they require the computation and evaluation of the derivatives of $f(t, y)$. These evaluations can be complicated and time-consuming.
- Runge-Kutta methods have the high-order local truncation error of the Taylor methods but do not need to compute and evaluate derivatives of $f(t, y)$.
- To give some idea of how Runge-Kutta methods are developed, we will show the derivation of a simple second order method.

Taylor's Theorem in Two Variables

Theorem 6 *Suppose that $f(t,y)$ and all its partial derivatives of order less than or equal to $n+1$ are continuous on $D = \{(t,y) \,|\, a \le t \le b,\ c \le y \le d\}$, and let $(t_0, y_0) \in D$. For every $(t,y) \in D$, there exists ξ between t and t_0, and μ between y and y_0 with*

$$f(t,y) = P_n(t,y) + R_n(t,y), \quad P_n \in \mathbb{R}[t,y],$$

where

$$R_n(t,y) = \frac{1}{(n+1)!} \sum_{j=0}^{n+1} \binom{n+1}{j} (t-t_0)^{n+1-j} (y-y_0)^j \frac{\partial^{n+1} f}{\partial t^{n+1-j} \partial y^j}(\xi, \mu),$$

$$
\begin{aligned}
P_n(t,y) \quad = \quad & f(t_0,y_0) + \left((t-t_0)\frac{\partial f}{\partial t}(t_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(t_0,y_0) \right) + \\
& \left(\frac{(t-t_0)^2}{2}\frac{\partial^2 f}{\partial t^2}(t_0,y_0) + (t-t_0)(y-y_0)\frac{\partial^2 f}{\partial t \partial y}(t_0,y_0) + \right. \\
& \left. \frac{(y-y_0)^2}{2}\frac{\partial^2 f}{\partial y^2}(t_0,y_0) \right) + \cdots + \\
& \left(\frac{1}{n!}\sum_{j=0}^{n} \binom{n}{j} (t-t_0)^{n-j}(y-y_0)^j \frac{\partial^n f}{\partial t^{n-j}\partial y^j}(t_0,y_0) \right) .
\end{aligned}
$$

Example 11 For

$$
f(t,y) = e^{-1/4\,(t-2)^2 - 1/4\,(y-3)^2} \cos\left(2\,t + y - 7\right),
$$

$$
P_2(t,y) = 1 - 9/4\,(t-2)^2 - 2\,(t-2)\,(y-3) - 3/4\,(y-3)^2 .
$$

A Second-Order Runge-Kutta Method

- Second-order Taylor's method.

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hT^{(2)}(t_i, w_i), \quad i = 0, 1, \ldots, N-1, \\ T^{(2)}(t_i, w_i) &= f(t_i, w_i) + \tfrac{h}{2} f'(t_i, w_i). \end{aligned}$$

Idea. Find a_1, α_1, β_1 with the property that $a_1 f(t + \alpha_1, y + \beta_1)$ approximates $T^{(2)}(t, y)$.

- Since

$$f'(t, y) = \frac{df}{dt}(t, y) = \frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) \cdot y'(t), \quad y'(t) = f(t, y).$$

Hence,

$$T^{(2)}(t, y) = f(t, y) + \frac{h}{2}\frac{\partial f}{\partial t}(t, y) + \frac{h}{2}\frac{\partial f}{\partial y}(t, y) \cdot f(t, y). \qquad (7)$$

• Expanding $f(t+\alpha_1, y+\beta_1)$ in its Taylor polynomial of degree one about (t, y) gives

$$\begin{aligned} a_1 f(t+\alpha_1, y+\beta_1) &= a_1 f(t,y) + a_1\alpha_1 \frac{\partial f}{\partial t}(t,y) + \\ & \qquad a_1\beta_1 \frac{\partial f}{\partial y}(t,y) + a_1 R_1(t+\alpha_1, y+\beta_1), \quad (8) \end{aligned}$$

where

$$\begin{aligned} R_1(t+\alpha_1, y+\beta_1) &= \frac{\alpha_1^2}{2}\frac{\partial^2 f}{\partial t^2}(\xi,\mu) + \\ & \qquad \alpha_1\beta_1 \frac{\partial^2 f}{\partial t \partial y}(\xi,\mu) + \frac{\beta_1^2}{2}\frac{\partial^2 f}{\partial y^2}(\xi,\mu), \quad (9) \end{aligned}$$

• Matching the coefficients of f and its derivative in (7) and (8):

$$a_1 = 1, \quad \alpha_1 = \frac{h}{2}, \quad \beta_1 = \frac{h}{2} f(t,y).$$

Hence,

$$\begin{aligned} T^{(2)}(t,y) &= f\left(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)\right) - \\ &\quad R_1\left(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)\right), \\ R_1\left(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)\right) &= \frac{h^2}{8}\frac{\partial^2 f}{\partial t^2}(\xi,\mu) + \frac{h^2}{4}f(t,y)\frac{\partial^2 f}{\partial t \partial y}(\xi,\mu) + \\ &\quad \frac{h^2}{8}(f(t,y))^2\frac{\partial^2 f}{\partial y^2}(\xi,\mu). \end{aligned}$$

If all the second-order partial derivatives of f are bounded, then

$$R_1\left(t+\frac{h}{2}, y+\frac{h}{2}f(t,y)\right)$$

is $O(h^2)$.

Midpoint Method.

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + hf(t_i + \tfrac{h}{2}, w_i + \tfrac{h}{2} f(t_i, w_i)), \quad i = 0, 1, \ldots, N-1. \end{aligned}$$

Other $O(h^2)$ Methods: modified Euler method, Heune's method.

Modified Euler Method.

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + \tfrac{h}{2}(f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))), \end{aligned}$$

for each $i = 0, 1, \ldots, N-1$.

Heune's Method.

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + \tfrac{h}{4}\left(f(t_i, w_i) + 3f(t_i + \tfrac{2}{3}h, w_i + \tfrac{2}{3}hf(t_i, w_i))\right), \end{aligned}$$

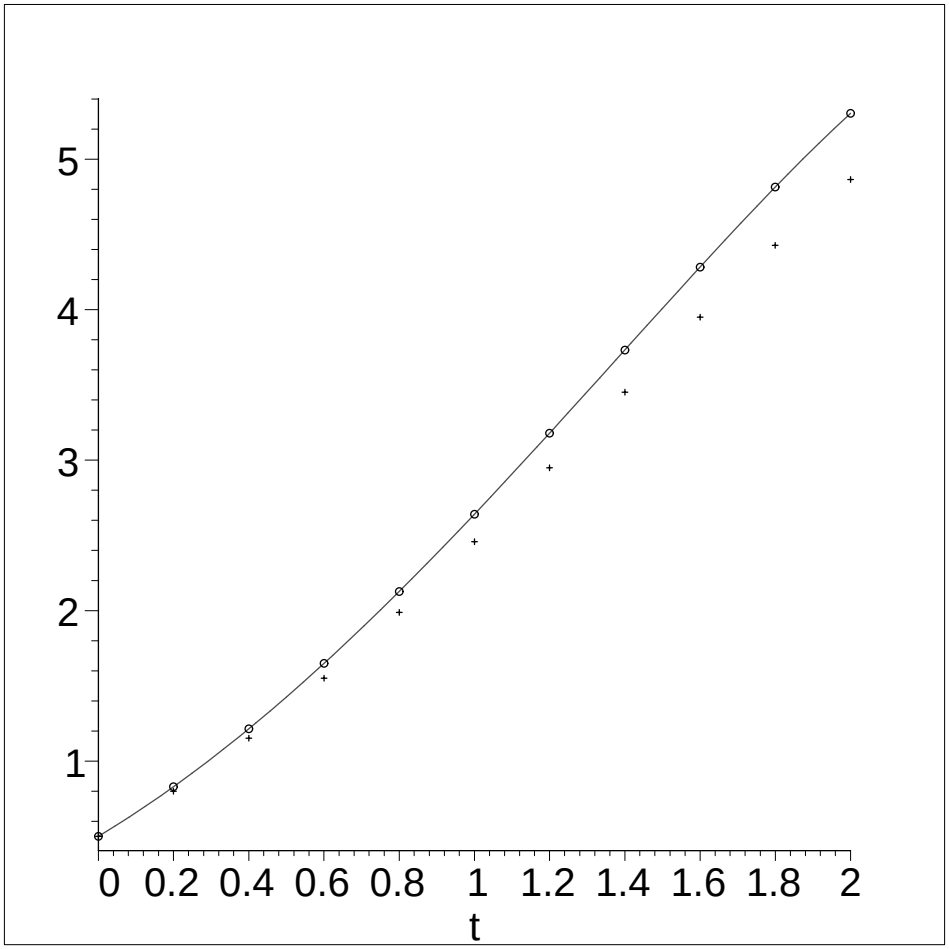
for each $i = 0, 1, \ldots, N-1$.

Runge-Kutta Order Four.

$$\begin{aligned} w_0 &= \alpha, \\ k_1 &= hf(t_i, w_i), \\ k_2 &= hf\left(t_i + \tfrac{h}{2}, w_i + \tfrac{1}{2}k_1\right), \\ k_3 &= hf\left(t_i + \tfrac{h}{2}, w_i + \tfrac{1}{2}k_2\right), \\ k_4 &= hf(t_{i+1}, w_i + k_3), \\ w_{i+1} &= w_i + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k4), \end{aligned}$$

for each $i = 0, 1, \ldots, N-1$.

IVP: $y' = y - t^2 + 1, \;\; 0 \leq t \leq 2, \;\; y(0) = 0.5.$



Exact solution: $y\,(t) = (t+1)^2 - 1/2\, e^t$

Error Control and Runge-Kutta-Fehlberg Method

- Up to this stage, we have discussed using *constant time steps* for computing the solutions. In practice this is a bad idea since the size of a time step should depend on the shape of the particular function.
- In order to take advantage of the changing shapes of a function we need to detect when we need to increase or decrease the timestep in response to how rapidly the solution $y(t)$ is changing.
- The general approach is to use *methods of differing orders* to predict the local truncation error, and consequently, choose a step size that keeps the local and global errors in check.

nth-order Taylor method	$(n+1)$st-order Taylor method
$y(t_{i+1}) = y(t_i) + h\phi(t_i, y(t_i), h) + O(h^{n+1})$	$y(t_{i+1}) = y(t_i) + h\tilde{\phi}(t_i, y(t_i), h) + O(h^{n+2})$
$w_0 = \alpha$	$\tilde{w}_0 = \alpha$
$w_{i+1} = w_i + h\phi(t_i, w_i, h)$	$\tilde{w}_{i+1} = \tilde{w}_i + h\tilde{\phi}(t_i, \tilde{w}_i, h)$
$\tau_{i+1}(h) = O(h^n)$	$\tilde{\tau}_{i+1}(h) = O(h^{n+1})$.

ASSUME (1) $w_i \approx y(t_i) \approx \tilde{w}_i$, and (2) a fixed step size h is used to generate w_{i+1} and $\tilde{w}_{i+1}$ to $y(t_{i+1})$.

$$\begin{aligned}
\tau_{i+1}(h) &= \frac{y(t_{i+1}) - y(t_i)}{h} - \phi(t_i, y(t_i), h) \\
&= \frac{y(t_{i+1}) - w_i}{h} - \phi(t_i, w_i, h) \\
&= \frac{y(t_{i+1}) - (w_i + h\phi(t_i, w_i, h))}{h} \\
&= \frac{1}{h}(y(t_{i+1}) - w_{i+1}).
\end{aligned}$$

Similarly,

$$\tilde{\tau}_{i+1}(h) = \frac{1}{h}(y(t_{i+1}) - \tilde{w}_{i+1}).$$

Consequently,

$$\begin{aligned}
\tau_{i+1}(h) &= \frac{1}{h}(y(t_{i+1}) - w_{i+1}) \\
&= \frac{1}{h}((y(t_{i+1}) - \tilde{w}_{i+1}) + (\tilde{w}_{i+1} - w_{i+1})) \\
&= \tilde{\tau}_{i+1}(h) + \frac{1}{h}(\tilde{w}_{i+1} - w_{i+1}).
\end{aligned}$$

Since $\tau_{i+1}(h)$ is $O(h^n)$, and $\tilde{\tau}_{i+1}(h)$ is $O(h^{n+1})$,

$$\begin{aligned}
\tau_{i+1}(h) &\approx Kh^n, \quad K \in \mathbb{R}, \\
\tau_{i+1}(h) &\approx \frac{1}{h}(\tilde{w}_{i+1} - w_{i+1}).
\end{aligned}$$

To bound $\tau_{i+1}(qh)$ by a tolerance ε, we choose q so that

$$\begin{aligned} |\tau_{i+1}(qh)| &\approx |K(qh)^n| = |q^n(Kh^n)| \approx |q^n \tau_{i+1}(h)| \\ &\approx \left| \frac{q^n}{h}(\tilde{w}_{i+1} - w_{i+1}) \right| \leq \varepsilon. \end{aligned}$$

Hence,

$$q \leq \left(\frac{\varepsilon h}{|\tilde{w}_{i+1} - w_{i+1}|} \right)^{1/n}.$$

The value of q determined at the ith step is used for two purposes:

1. to reject, if necessary, the initial choice of h at the ith step, and repeat the calculations using qh; and
2. to predict an appropriate initial choice of h for the $(i+1)$st step.

One popular technique is the Runge-Kutta-Fehlberg method. This technique uses a Runge-Kutta method with local truncation error of order five to estimate the local error in a Runge-Kutta method of order four. See textbook (p284-p285) for the formulas.

One-step Methods: Stability and Convergence

As we make the step size smaller and smaller, we want the numerical approximation to become closer and closer to the exact result. More precisely, we want a *convergent method*:

Definition 5 *A one-step difference equation method is convergent with respect to the differential equation it approximates if*

$$\lim_{h \to 0} \max_{1 \le i \le N} |w_i - y(t_i)| = 0.$$

Example 12 The error bound formula for Euler's method is

$$|w_i - y_i| \le \frac{Mh}{2L} \left| e^{(t_i - a)L} - 1 \right|.$$

Hence,

$$\max_{1 \le i \le N} |w_i - y(t_i)| \le \frac{Mh}{2L} \left| e^{L(b-a)} - 1 \right|,$$

and Euler's method is convergent.

Instead of deriving global error bound formulae to show convergence (which is difficult), we prefer to use local properties. One local property which will be particularly useful is *consistency*:

Definition 6 *A one-step difference equation method with a local truncation error $\tau_i(h)$ at the i-th step is* consistent *with the differential equation it approximates if*

$$\lim_{h \to 0} \max_{1 \leq i \leq N} |\tau_i(h)| = 0.$$

A one-step method is *consistent precisely* when the local truncation error approaches zero as the step size approaches zero.

We also want *stable methods*, i.e., methods that have the property that small changes or perturbations in the initial conditions produce correspondingly small changes in the subsequent approximations.

Stability, consistency, and convergence are all tied together in the following theorem:

Theorem 7 *Suppose the initial-value problem*

$$y' = f(t, y), \quad a \leq t \leq b, \quad y(a) = \alpha,$$

is approximated by a one-step difference method in the form

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + h\phi(t_i, w_i). \end{aligned}$$

Suppose also that a number $h_0 > 0$ exists and that $\phi(t, w, h)$ is continuous and satisfies a Lipschitz condition in the variable w with Lipschitz constant L on the set

$$D = \{(t, w, h) \mid a \le t \le b,\ -\infty < w < \infty,\ 0 \le h \le h_0\}.$$

Then

(i) *The method is stable;*

(ii) *The difference method is convergent if and only if it is consistent, which is equivalent to*

$$\phi(t, y, 0) = f(t, y), \quad \textit{for all} \quad a \le t \le b;$$

(iii) *If a function τ exists and, for each $i = 1, 2, \ldots, N$, the local truncation error $\tau_i(h)$ satisfies $|\tau_i(h)| \le \tau(h)$ whenever $0 \le h \le h_0$, then*

$$|y(t_i) - w_i| \le \frac{\tau(h)}{L}\, e^{L(t_i - a)}.$$

Example 13 Consider the Modified Euler method:

$$\begin{aligned} w_0 &= \alpha, \\ w_{i+1} &= w_i + \tfrac{h}{2}(f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))), \end{aligned}$$

for each $i = 0, 1, \ldots, N-1$. For

$$\phi(t, w, h) = \frac{1}{2}(f(t, w) + f(t + h, w + hf(t, w)),$$

we verify that this method satisfies the hypothesis of Theorem 7.

- ϕ is continuous. If f is continuous on

$$\{(t, w) \mid a \le t \le b,\ -\infty < w < \infty\},$$

then ϕ is continuous on

$$\{(t, w, h) \mid a \le t \le b,\ -\infty < w < \infty,\ 0 \le h \le h_0\}.$$

- ϕ satisfies a Lipschitz condition in w. That is,

$$\exists\, L > 0 \text{ s.t. } |\phi(t, w, h) - \phi(t, \bar{w}, h)| \leq L|w - \bar{w}|.$$

$$\begin{aligned} \phi(t, w, h) - \phi(t, \bar{w}, h) &= \frac{1}{2} f(t, w) + \frac{1}{2} f(t + h, w + hf(t, w)) \\ &\quad - \frac{1}{2} f(t, \bar{w}) - \frac{1}{2} f(t + h, \bar{w} + hf(t, \bar{w})). \end{aligned}$$

If f satisfies a Lipschitz condition on

$$\{(t, w) \mid a \leq t \leq b,\ -\infty < w < \infty\}$$

in the variable w with constant L_f, then

$$
\begin{aligned}
|\phi(t,w,h) - \phi(t,\bar{w},h)| &\leq \frac{1}{2}L_f|w-\bar{w}| + \\
&\quad \frac{1}{2}L_f|w + hf(t,w) - \bar{w} - hf(t,\bar{w})| \\
&\leq L_f|w-\bar{w}| + \frac{1}{2}L_f|hf(t,w) - hf(t,\bar{w})| \\
&\leq L_f|w-\bar{w}| + \frac{1}{2}hL_f^2|w-\bar{w}| \\
&= \left(\underbrace{L_f + \frac{1}{2}hL_f^2}_{L}\right)|w-\bar{w}|.
\end{aligned}
$$

Hence, ϕ satisfies a Lipschitz condition in w on the set

$$\{(t,w,h) \mid a \leq t \leq b,\ -\infty < w < \infty,\ 0 \leq h \leq h_0\},$$

for any $h_0 > 0$ with constant L.

Consequently, Theorem 7 implies that the Modified Euler method is stable. Letting $h = 0$, we have

$$\phi(t, w, 0) = f(t, w),$$

so the consistency condition expressed in Theorem 7, part (ii) holds. Thus, the method is convergent. Moreover, we have seen that for this method the local truncation error is $O(h^2)$, so the convergence of the Modified Euler method is also $O(h^2)$.

Multistep Methods

- Runge-Kutta and Taylor methods are examples of one-step methods because the approximation at $t = t_{i+1}$ only involves previous mesh point t_i.
- Since the approximate solution is available at the mesh points $t_0, t_1, \ldots, t_i$, an alternate approach is to use these earlier approximations to develop an approximation for $y(t_{i+1})$.

Definition 7 *An m-step multistep method for solving the IVP*

$$y' = f(t, y), \;\; a \leq t \leq b, \;\; y(a) = \alpha,$$

has a difference equation for finding the approximation w_{i+1} at the mesh point t_{i+1} represented by the following equation, where m is an integer greater than 1:

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + \cdots + a_0 w_{i+1-m} +$$

$$h(b_m f(t_{i+1}, w_{i+1}) + b_{m-1} f(t_i, w_i) + \cdots + b_0 f(t_{i+1-m}, w_{i+1-m})),$$

for $i = m-1, m, \ldots, N-1$, where $h = (b-a)/N$, the a_i's, b_j's are constants, and the starting values

$$w_0 = \alpha, \;\; w_1 = \alpha_1, \;\; w_2 = \alpha_2, \;\; \ldots, \;\; w_{m-1} = \alpha_{m-1}$$

are specified.

General technique.

$$\int_{t_i}^{t_{i+1}} y'(t)\,dt = \int_{t_i}^{t_{i+1}} f(t,y(t))\,dt$$

$$y(t_{i+1}) \approx w_i + \int_{t_i}^{t_{i+1}} f(t,y(t))\,dt$$

Let $f(t,y(t)) = P_{m-1}(t) + R_{m-1}(t)$ where $P_{m_1}(t)$ is the interpolating polynomial which interpolates the m points

$$(t_i,f(t_i,y(t_i))),(t_{i-1},f(t_{i-1},y(t_{i-1}))),\ldots,(t_{i+1-m},f(t_{i+1-m},y(t_{i+1-m}))),$$

and the remainder term

$$R_{m-1}(t) = \frac{f^{(m)}(\xi_i,y(\xi_i))}{m!}(t-t_i)(t-t_{i-1})\cdots(t-t_{i+1-m}), \xi_i \in (t_{i+1-m},t_i).$$

Hence,

$$y(t_{i+1}) \approx w_i + \int_{t_i}^{t_{i+1}} P_{m-1}(t)\,dt + \int_{t_i}^{t_{i+1}} R_{m-1}(t)\,dt.$$

Example 14 .

Fourth-order Adams-Bashforth technique.

$$w_0 = \alpha,\ w_1 = \alpha_1,\ w_2 = \alpha_2,\ w_3 = \alpha_3,$$
$$w_{i+1} = w_i + \frac{h}{24}(55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) +$$
$$37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3})),\ i = 3, 4, \ldots, N-1.$$

Fourth-order Adams-Moulton technique.

$$w_0 = \alpha,\ w_1 = \alpha_1,\ w_2 = \alpha_2,$$
$$w_{i+1} = w_i + \frac{h}{24}(9f(t_{i+1}, w_{i+1}) + 19f(t_i, w_i) -$$
$$5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2})),\ i = 2, 3, \ldots, N-1.$$

Remark 2 .

- The fourth-order Adams-Moulton method is an *implicit* method since w_{i+1} occurs on both sides and is specified *implicitly*.
- To apply an implicit method, one must solve the implicit equation for w_{i+1}. It is not clear that this can be done in general or that a unique solution exists.

Remark 3 .

- The textbook (p.294,p.295) gives 2, 3, 4, and 5 step Adams-Bashforth (A-B) schemes. They are *explicit* methods.
- The textbook (p.295) also gives 2, 3, and 4 step Adams-Moulton (A-M) schemes. They are *implicit* methods.

m-step A-B explicit methods vs $(m-1)$ step A-M implicit methods.

	Local Truncation Error		Local Truncation Error
3 step A-B	$\frac{3}{8}y^{(4)}(\mu_i)h^3$	2 step A-M	$-\frac{1}{24}y^{(4)}(\mu_i)h^3$
4 step A-B	$\frac{251}{720}y^{(5)}(\mu_i)h^4$	3 step A-M	$-\frac{19}{720}y^{(5)}(\mu_i)h^4$
5 step A-B	$\frac{95}{288}y^{(6)}(\mu_i)h^5$	4 step A-M	$-\frac{3}{160}y^{(6)}(\mu_i)h^5$

- Both have the terms $y^{(m+1)}(\mu_i)h^m$ in their local truncation errors.
- The coefficients are smaller for the A-M implicit methods. This contributes to a significantly smaller roundoff error in the implicit methods.

Predictor-corrector method. A combination of explicit and implicit methods. The explicit method *predicts* an approximation, and the implicit method *corrects* the prediction.

Example 15 Consider a predictor-corrector method based on the 4-step A-B explicit method, and the 3-step A-M implicit method.

- $w_0 \longrightarrow w_1 \longrightarrow w_2 \longrightarrow w_3$ (1-step method, e.g., Runge-Kutta of order 4).
- 4-step explicit A-B $\longrightarrow$ $w_4^{(0)}$:

  $$w_4^{(0)} = w_3 + \frac{h}{24}(55f(t_3, w_3) - 59f(t_2, w_2) + 37f(t_1, w_1) - 9f(t_0, w_0)).$$

- 3-step implicit A-M $\longrightarrow$ $w_4^{(1)} \approx y(t_4)$:

  $$w_4^{(1)} = w_3 + \frac{h}{24}(9f(t_4, w_4^{(0)}) + 19f(t_3, w_3) - 5f(t_2, w_2) + f(t_1, w_1)).$$

- $w_5^{(0)}$, $w_5^{(1)} \approx y(t_5)$, etc.