

1


## Initial-value problem solvers

- Our error analysis showed that for each step of these techniques, the error was $\mathrm{O}\left(h^{N}\right)$ for a single step,
but for multiple steps, it was $\mathrm{O}\left(h^{N-1}\right)$
- This was based on a similar analysis we saw for integration
- Problem:

With integration, we could assume each value was exact

- With IVP solvers,
all but once, we estimate $y_{k+1}$ using a previous estimate

3

## Euler's method

- Fortunately, the error is never-the-less proportional to $h^{N-1}$, but, the coefficient may be slightly larger than expected
- We will do a slightly deeper dive on Euler's method
- First: the error depends on $y^{(2)}(\tau)$, but can we estimate this?
- Recall that

$$
y^{(1)}(t)=f(t, y(t))
$$

- Thus

$$
\begin{aligned}
y^{(2)}(t) & =\frac{\mathrm{d}}{\mathrm{~d} y} f(t, y(t)) \\
& =\frac{\partial}{\partial t} f(t, y(t))+\frac{\partial}{\partial y} f(t, y(t)) \frac{\mathrm{d}}{\mathrm{~d} y} y(t) \\
& =\frac{\partial}{\partial t} f(t, y(t))+\frac{\partial}{\partial y} f(t, y(t)) f(t, y(t))
\end{aligned}
$$

4


4

## Euler's method

- Let's look at our two examples
- Given $y^{(1)}(t)=-y(t)$

$$
\begin{aligned}
y^{(2)}(t) & =0+(-1)(-y(t)) \\
& =y(t)
\end{aligned}
$$

- Given $y^{(1)}(t)=-0.2 y(t)-\sin (t)-0.1$

$$
\begin{aligned}
y^{(2)}(t) & =-\cos (t)+(-0.2)(-0.2 y(t)-\sin (t)-0.1) \\
& =-\cos (t)+0.04 y(t)+0.2 \sin (t)+0.02
\end{aligned}
$$

5

## Lipschitz constant

- A function $f$ on a domain $\left[t_{0}, t_{f}\right] \times\left[y_{\mathrm{L}}, y_{\mathrm{U}}\right]$ is said to satisfy a Lipschitz condition if

$$
\left|f\left(t, y_{2}\right)-f\left(t, y_{1}\right)\right| \leq L\left|y_{2}-y_{1}\right|
$$

- For example, given $f(t, y)=-y$

$$
\left|\left(-y_{2}\right)-\left(-y_{1}\right)\right|=1 \cdot\left|y_{2}-y_{1}\right|
$$

- For example, given $f(t, y)=-0.2 y-\sin (t)-0.1$

$$
\begin{aligned}
\mid\left(-0.2 y_{2}-\sin (t)-0.1\right)-(-0.2 & \left.y_{1}-\sin (t)-0.1\right) \mid \\
& =\left|0.2\left(y_{1}-y_{2}\right)\right| \\
& =0.2\left|y_{2}-y_{1}\right|
\end{aligned}
$$

## Euler's method

- You can therefore show the error may be larger than expected:

$$
\begin{aligned}
\left|y\left(t_{f}\right)-y_{N}\right| & \leq h \frac{\mid y^{(2)}(\tau)}{2 L}\left(e^{L\left(t_{f}-t_{0}\right)}-1\right) \\
& =h \frac{\left|y^{(2)}(\tau)\right|\left(\left(t_{f}-t_{0}\right)+\frac{1}{2} L\left(t_{f}-t_{0}\right)^{2}+\frac{1}{6} L^{2}\left(t_{f}-t_{0}\right)^{3}+\cdots\right)}{2} \\
& =h \frac{\left|y^{(2)}(\tau)\right|}{2}\left(t_{f}-t_{0}\right)+h \frac{y^{(2)}(\tau)}{2}\left(\frac{1}{2} L\left(t_{f}-t_{0}\right)^{2}+\frac{1}{6} L^{2}\left(t_{f}-t_{0}\right)^{3}+\cdots\right)
\end{aligned}
$$



7

## A comment on error

- With previous algorithms,
we required successive approximations to be within $\varepsilon_{\text {abs }}$
- Previous problems have included:
- Convergence of the root finding algorithms
- Approximating a solution to an IVP is an open-ended problem
- We may not know at which point we wish to stop
- Thus, we will use the maximum absolute error per unit time
- Given an IVP and we want to approximate a solution at $t_{p}$ we will expect the error to be less than $\varepsilon_{\mathrm{abs}}\left(t_{f}-t_{0}\right)$


## Muler's method

## Is our approximation good enough?

- Question: With Euler's method, how large should $h$ be?
- Consider the following strategy:
- Choose an $n$ and find $y_{n}$ which approximates $y\left(t_{f}\right)$
- Repeat this process but now use $2 n$ intervals
- Let these approximations be denoted $z_{k}$ so that $z_{2 n}$ also approximates $y\left(t_{f}\right)$


9

Is our approximation good enough?

- Thus, we have two approximations of $y\left(t_{f}\right)$ :

$$
\begin{gathered}
y\left(t_{f}\right) \approx y_{n}+C h \\
y\left(t_{f}\right) \approx z_{2 n}+C \frac{h}{2} \\
0 \approx\left(y_{n}+C h\right)-\left(z_{2 n}+C \frac{h}{2}\right)
\end{gathered}
$$

- Thus, we have

$$
z_{2 n}-y_{n} \approx C \frac{h}{2}
$$

- Therefore, the error of $z_{2 n}$ is approximately

$$
y\left(t_{f}\right)-z_{2 n} \approx z_{2 n}-y_{n}
$$

## Euler's method

## Is our approximation good enough?

- Thus, by subtracting the worse approximation from the better approximation, we have an approximation of the error of $z_{2 n}$

$$
y\left(t_{f}\right)-z_{2 n} \approx z_{2 n}-y_{n}
$$

- Not only that, but we can do one better:

$$
y\left(t_{f}\right) \approx 2 z_{2 n}-y_{n}
$$

- Thus, having calculated $y_{n}$ and $z_{2 n}$, we can find an even better approximation of $y\left(t_{f}\right)$
- This weighted average is an $\mathrm{O}\left(h^{2}\right)$ approximation!

11

## Is our approximation good enough?

- Consequently, given our two approximations
- The difference between the last two is an approximation of the error of the last $z_{2 n}$
- Also, for each $y_{k}$ we have that $y\left(t_{k}\right), 2 z_{2 k}-y_{k}$ is a better approximation of $y\left(t_{k}\right)$



## Euler's method

## Is our approximation good enough?

- Thus, suppose we want our approximation to $y\left(t_{f}\right)$ to be no larger than $\varepsilon_{\mathrm{abs}}\left(t_{f}-t_{0}\right)$
- Thus, we will do the following:
- Choose an $n$ and calculate the approximations $y_{1}, \ldots, y_{n}$
- With $2 n$ intervals, calculate the approximations $z_{1}, \ldots, z_{2 n}$
- Recall that $\left|y_{n}-z_{2 n}\right|$ is only an estimate of the error
- To be safe, check if $2\left|y_{n}-z_{2 n}\right|<\varepsilon_{\text {abs }}\left(t_{f}-t_{0}\right)$
- If this is true, we are finished and approximate $y\left(t_{k}\right)$ with the $n$ values $y\left(t_{k}\right) \approx 2 z_{2 k}-y_{k}$
- Otherwise, repeat this process again but now with $4 n$ intervals
- Because these techniques require us to approximate solutions with $n$, then $2 n$ and possibly $4 n, 8 n$, etc. intervals, we will call these approaches iterative solvers


## Is our approximation good enough?

- We can do this with Heun and RK4

$$
\begin{array}{cc}
\text { Heun } & \text { RK4 } \\
y\left(t_{f}\right) \approx y_{n}+C h^{2} & y\left(t_{f}\right) \approx y_{n}+C h^{4} \\
y\left(t_{f}\right) \approx z_{2 n}+C \frac{h^{2}}{4} & y\left(t_{f}\right) \approx z_{2 n}+C \frac{h^{4}}{16} \\
0 \approx\left(y_{n}+C h^{2}\right)-\left(z_{2 n}+C \frac{h^{2}}{4}\right) & 0 \approx\left(y_{n}+C h^{2}\right)-\left(z_{2 n}+C \frac{h^{4}}{16}\right) \\
z_{2 n}-y_{n} \approx C \frac{3}{4} h^{2} & z_{2 n}-y_{n} \approx C \frac{15}{16} h^{4} \\
\frac{z_{2 n}-y_{n}}{3} \approx C \frac{h^{2}}{4} & \frac{z_{2 n}-y_{n}}{15} \approx C \frac{h^{4}}{16} \\
y\left(t_{f}\right)-z_{2 n} \approx \frac{z_{2 n}-y_{n}}{3} & y\left(t_{f}\right)-z_{2 n} \approx \frac{z_{2 n}-y_{n}}{15} \\
y\left(t_{f}\right) \approx \frac{4 z_{2 n}-y_{n}}{3} & y\left(t_{f}\right) \approx \frac{16 z_{2 n}-y_{n}}{15}
\end{array}
$$

## Issues

- This approach, however, does have its drawbacks:
- It's expensive: we are making a minimum of $3 n$ approximations
- Even if the approximation at $y\left(t_{f}\right)$ is sufficiently accurate,
this does not mean that each point is sufficiently accurate...
- Also, suppose the solution is smooth at some points, and variable at others
- Where the solution is smooth, we can get away with a larger step size
- Where the solution varies, we probably require a smaller step size
- This approach, however, requires us to use that smaller step size everywhere

15


- We will, instead, look at adaptive techniques that allow us to dynamically vary the step size
- Question: How can we estimate the error if we don't know what the solution actually is?


## Summary

- Following this topic, you now
- Understand how to get $y^{(2)}(t)$ given $y^{(1)}(t)=f(t, y(t))$
- Are aware that the error may be more significant
- You are not required to master this in this course
- You must simply be aware of the issue, many of you will never come across this issue
- Understand we will define the maximum absolute error in terms of an acceptable error $\varepsilon_{\text {abs }}$ per unit time
- Because $\left|y_{n}-z_{2 n}\right|$ only approximates the error, we'll double it
- Know how to iteratively estimate the error with Euler's, Heun's and the $4^{\text {th }}$-order Runge-Kutta method
- Know how to get an even better approximation using two lessoptimal approximations
- Understand that there are issues with this approach

17



19


## Disclaimer

These slides are provided for the ECE 204 Numerical methods course taught at the University of Waterloo. The material in it reflects the author's best judgment in light of the information available to them at the time of preparation. Any reliance on these course slides by any party for any other purpose are the responsibility of such parties. The authors accept no responsibility for damages, if any, suffered by any party as a result of decisions made or actions based on these course slides for any other purpose than that for which it was intended.

