

## Advanced Topics from Scientific Computing

TU Berlin Winter 2023/24

### Notebook 09

 Jürgen Fuhrmann

GRBackend()

```
1 begin
2     using LinearAlgebra
3     using ForwardDiff, DiffResults, NLSolve, RecursiveArrayTools
4     using RungeKutta
5     using DifferentialEquations
6     using DiffEqDevTools
7     using DataStructures
8     using Colors, Plots
9     Plots.gr()
10 end
```

# ODE and DAE Solvers

---

Literature:

- E. Hairer, S.P. Nørsett, G. Wanner: Solving Ordinary Differential Equations I. Nonstiff Problems
- E. Hairer, G. Wanner: Solving Ordinary Differential Equations II. Stiff and Differential Algebraic Problems
- P. Kunkel, V. Mehrmann: Differential-Algebraic Equations. Analysis and Numerical Solution

Further sources used:

- [TU Chemnitz lecture notes by A. Naumann](#) (in German)
- HU Berlin course "Numerik gewöhnlicher Differentialgleichungen I" (implementation of Explicit Runge-Kutta)

While this course focuses on the solution of partial differential equations which describe coupled processes in space and time, we first talk about the solution of ordinary differential equations, providing methods for discretization in time.

Look for  $\mathbf{u} : [0, T] \rightarrow \mathbb{R}^m$  be a time dependent, differentiable function such that it fulfills the initial value problem for the ordinary differential equation (ODE) system:

$$\begin{cases} \dot{\mathbf{u}} &= \mathbf{f}(\mathbf{u}, \mathbf{p}, t) \\ \mathbf{u}|_{t=0} &= \mathbf{u}_0 \end{cases}$$

with  $\mathbf{f} : \mathbb{R}^m \times \mathbb{R}^k \times [0, T] \rightarrow \mathbb{R}^m$ .

Here,  $\mathbf{p}$  is a  $k$ -vector of parameters.

The system is called *autonomous* if  $\mathbf{f}$  does not depend on  $t$ .

Subdivide  $[0, T]$  into  $N - 1$  intervals of time step size  $\tau = \frac{T}{N-1}$ . Let  $t_n = (n - 1)\tau$  and let  $\mathbf{u}_n = \mathbf{u}(t_n)$ .

## Explicit Euler method

The simplest way of discretizing in time is to equate the finite difference in time to the right hand side value calculated in the last timestep.

$$\frac{\mathbf{u}_{n+1} - \mathbf{u}_n}{\tau} = \mathbf{f}(\mathbf{u}_n, \mathbf{p}, t_n)$$

We implement this *explicit Euler method* such that we describe the right hand side  $\mathbf{du}=\mathbf{f}(\mathbf{u}, \mathbf{p}, t)$  via a mutating function  $\mathbf{f}!(\mathbf{du}, \mathbf{u}, \mathbf{p}, t)$ . It returns a vector of solution vectors and the chosen time values.

```

1 function explicit_euler(f!::F, # ODE right hand side
2                       p; # Parameters
3                       u0 = [0.0], # vector of initial values
4                       dt = 0.1, # time step size
5                       tspan = [0.0, 1.0],) where {F}
6     u = [u0]
7     t = Float64[tspan[1]]
8     du = zeros(length(u0))
9     while t[end] < tspan[end]
10      uold = u[end]
11      told = t[end]
12      f!(du, uold, p, told)
13      unew = uold + dt * du
14      tnew = told + dt
15      push!(u, unew)
16      push!(t, tnew)
17     end
18     u, t
19 end;
```

# Test problem

Let us first look at the equation

$$\dot{u} = \lambda u$$

```
1 f_test!(du, u, λ, t) = du[1] = λ * u[1];
```

The exact solution is given by  $u = u_0 e^{\lambda t}$ .

```
1 exact_test(u0, p, t) = u0 * exp(p * t);
```

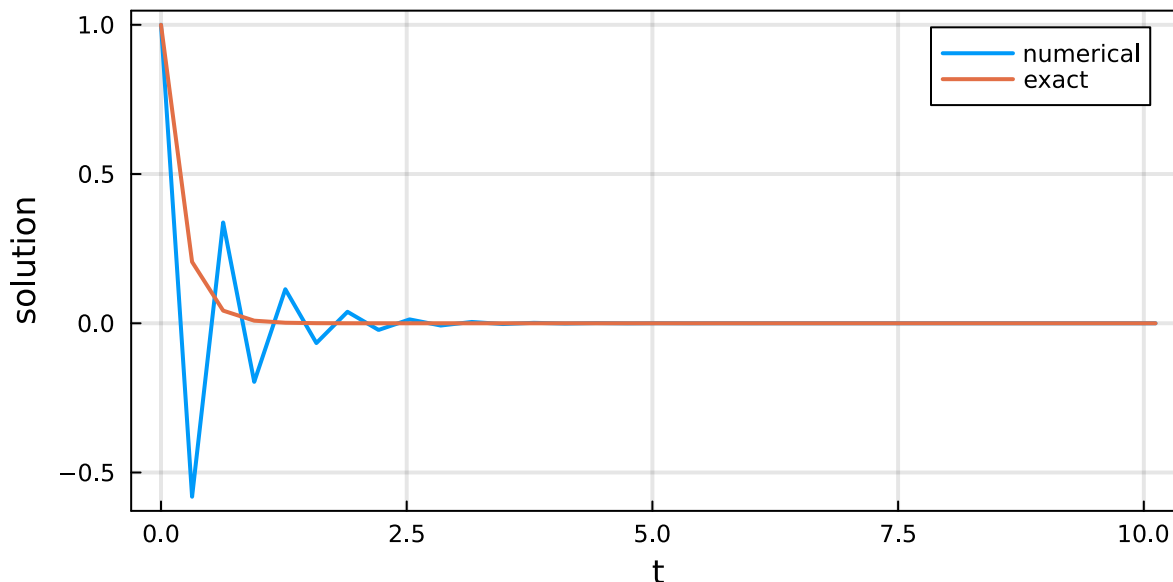
For  $\lambda = \frac{\log \frac{1}{2}}{t_{\frac{1}{2}}}$ , this equation e.g. describes radioactive decay with half value time  $t_{\frac{1}{2}}$

`p_test = -5`

```
1 p_test = -5
```

`log( $\tau$ ) =`  `-0.5`

```
1 u_test, t_test = explicit_euler(f_test!, p_test; u0 = [1.0], dt = 10logdt_test,  
tspan = (0, 10));
```



Obviously we may have a problem with the step size, and we may try to find other methods.

# $\theta$ - and Runge-Kutta methods

Main theorem of calculus gives

$$\int_{t_n}^{t_{n+1}} \dot{u} = \int_{t_n}^{t_{n+1}} f(u(t), p, t)$$

Approximate the integral using a quadrature rule with points  $c_i$  and weights  $b_i$ :

$$u_{n+1} - u_n = \tau \sum_{i=1}^s b_i f(u_{n,i}, p, t_n + c_i \tau)$$

Using some simple quadratures leads to

$$b = \{1\}, c = \{\theta\}, u_{n,1} = \theta u_{n+1} + (1 - \theta) u_n$$

- $\theta = 0$ : explicit Euler method
- $\theta = 1$ : implicit Euler method
- $\theta = \frac{1}{2}$ : implicit midpoint method

$$\frac{u_{n+1} - u_n}{\tau} = f(\theta u_{n+1} + (1 - \theta) u_n, p, t_n + \theta \tau)$$

Unless  $\theta=0$  (explicit Euler method) we need to solve an equation or a system of equations in each timestep. Here we implement this for general  $\theta$ .

## $\theta$ method: implementation

This implementation is slightly more general:

- if `linear=true`, instead of Newton's method to solve the implicit equations it performs just one Newton step (assuming the time step size is small enough). This results in a *linear implicit* method (linear implicit Euler or linear implicit midpoint)
- it implements the solution of a slightly more general problem: for a matrix  $M$  (unit matrix by default), solve

$$M\dot{u} = f(u, p, t).$$

```

1 function θmethod(f!::F, # right hand side function
2                 p; # Parameters
3                 u0 = [0.0], # Initial value
4                 dt = 0.1, # time step size
5                 tspan = (0, 1), # time interval
6                 θ = 0, # θ parameter
7                 linear = false, # switch between linear method and newton solver
8                 M = nothing, # Mass matrix
9                 tol = 1.0e-10,) where {F}
10 # Start arrays for the result
11 u = [u0]
12 t = Float64[tspan[1]]
13 uold = u[end]
14 told = t[end]
15
16 # If no mass matrix is given as parameter, let it be the unit matrix of
17 # appropriate size
18 if M == nothing
19     M = Diagonal(ones(length(u0)))
20 end
21
22 # Each timestep involves the solution of F(uneu)=0
23 # Here, uneu is the current iteration and F is the result
24 function fstep!(F, uneu)
25     f!(F, θ * uneu + (1 - θ) * uold, p, told + θ * dt)
26     F .= M * (uneu .- uold) ./ dt .- F
27 end
28
29 if linear
30     # Handle differentiation results
31     diffresult = DiffResults.JacobianResult(u0)
32     y = zero(u0)
33     cfg = ForwardDiff.JacobianConfig(fstep!, y, u0)
34 end
35
36 while t[end] < tspan[2]
37     # Loop over the timesteps
38     told = t[end]
39     uold = u[end]
40     if linear
41         # Perform one Newton step
42         ForwardDiff.jacobian!(diffresult, fstep!, y, uold, cfg)
43         uneu = uold - DiffResults.jacobian(diffresult) \
44         DiffResults.value(diffresult)
45     else
46         # Solve time step problem and check for convergence
47         result = nlsolve(fstep!, uold; autodiff = :forward, xtol = tol)
48         !converged(result) && throw("convergence error")
49         uneu = result.zero
50     end
51     # Extend solution arrays
52     push!(u, uneu)
53     push!(t, told + dt)
54 end

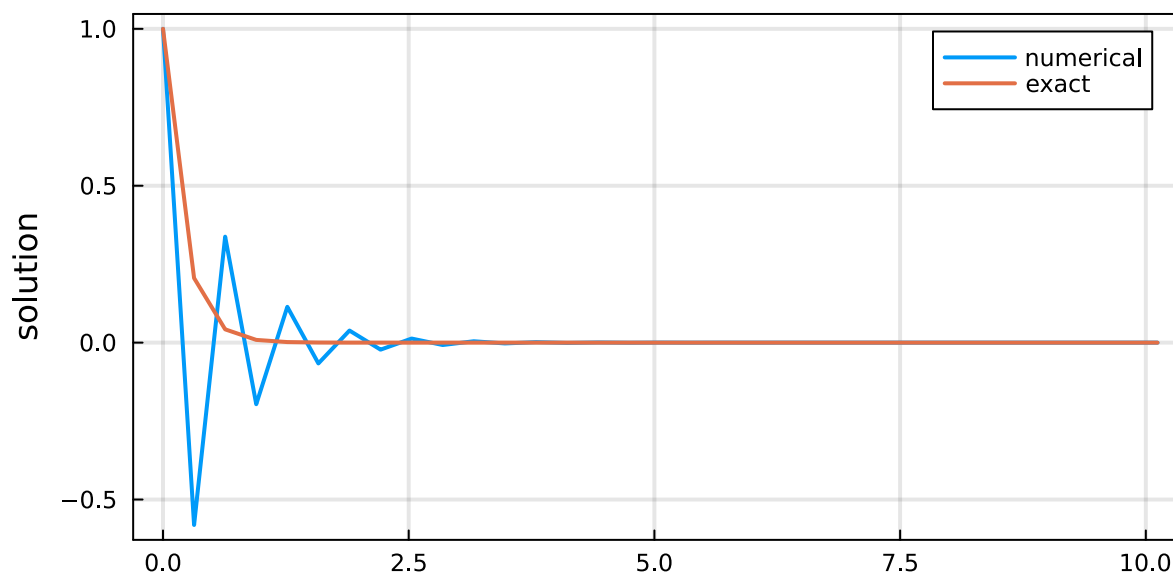
```

```
54 # Return the solution as DiffEqArray
55 # from RecursiveArrayTools.jl (easy to plot)
56 DiffEqArray(u, t)
57 end;
```

So let us use this for our test example:

```
1 uθ_test = @method(f_test!, p_test; u0 = [1.0], dt = 10^logdtx, tspan = (0, 10), θ = θx);
```

θ=  0.0  $\log(\tau)=$   -0.5



## Runge-Kutta methods

A generalization of the the choice of internal stages  $u_{n,i}$  in the quadrature based method starting with

$$\int_{t_n}^{t_{n+1}} \dot{u} = \int_{t_n}^{t_{n+1}} f(u(t), p, t)$$

$$u_{n+1} - u_n = \tau \sum_{i=1}^s b_i f(u_{n,i}, p, t_n + c_i \tau)$$

chooses another quadrature rule with the points  $c_1 \dots c_s$  and weights  $a_{ij}$ :

$$u_{n,i} - u_n = \tau \sum_{j=1}^s a_{ij} f(u_{n,j}, p, t_n + c_j \tau)$$

This Ansatz gives  $s$ -stage Runge-Kutta methods.

If  $A$  is strictly lower triangular, the method is explicit, i.e. it does only involve the evaluation of  $f$ , no solution of a system of equations involving  $f$ .

Otherwise, it is implicit, and it involves the solution of one or more, possibly coupled systems of equations of dimension  $m \times s$ . As a consequence, implicit Runge-Kutta methods are hard to use for large systems.

Runge-Kutta methods thus are described by the points  $c = \{c_1 \dots c_s\}$ , the weights  $b = \{b_1 \dots b_s\}$  and the  $s \times s$  matrix  $A = (a_{ij})$ . Usually these are arranged into *Butcher tableaux*:

$$\begin{array}{c|c} c & A \\ \hline & b \end{array}$$

These tableaux can be conveniently managed using the `RungeKutta.jl` package.

The  $\theta$  method is a Runge-Kutta method:

Runge-Kutta Tableau ExplicitEuler with 1 stage and order 1:

$$\begin{array}{c|c} 0//1 & 0//1 \\ \hline & 1//1 \end{array}$$

1 `TableauExplicitEuler(Rational)`

Runge-Kutta Tableau ImplicitEuler with 1 stage and order 1:

1//1	1//1
1//1	1//1

1 `TableauImplicitEuler(Rational)`

Runge-Kutta Tableau ImplicitMidpoint with 1 stage and order 2:

1//2	1//2
1//1	1//1

1 `TableauImplicitMidpoint(Rational)`

The classical Runge-Kutta method is "RK4":

Runge-Kutta Tableau RK416 with 4 stages and order 4:

0//1	0//1	0//1	0//1	0//1
1//2	1//2	0//1	0//1	0//1
1//2	0//1	1//2	0//1	0//1
1//1	0//1	0//1	1//1	0//1
1//6	1//3	1//3	1//3	1//6

1 `TableauRK4(Rational)`

With RungeKutta.jl, one can e.g. define the Dormand-Prince tableau:

Runge-Kutta Tableau DoPri5 with 7 stages and order 5:

0//1	0//1	0//1	0//1	0//1	0//1	0//1	...
1//5	1//5	0//1	0//1	0//1	0//1	0//1	...
3//10	3//40	9//40	0//1	0//1	0//1	0//1	...
4//5	44//45	-56//15	32//9	0//1	0//1	0//1	...
8//9	19372//6561	-25360//2187	64448//6561	-212//729	0//1	0//1	...
1//1	9017//3168	-355//33	46732//5247	49//176	-5103//18656	0//1	...
1//1	35//384	0//1	500//1113	125//192	-2187//6784	0//1	...
35//384	0//1	500//1113	125//192	-2187//6784	0//1	0//1	...

2 columns omitted

1 `TableauDP(Rational, 5)`

TableauDP (generic function with 4 methods)

The Radau IIA method is an implicit Runge-Kutta method:



Runge-Kutta Tableau RadauIIA with 3 stages and order 5:

0.155051	0.196815	-0.0655354	0.023771
0.644949	0.394424	0.292073	-0.0415488
1.0	0.376403	0.512486	0.111111
	0.376403	0.512486	0.111111

```
1 TableauRadauIIA(3)
```

Implementation of the explicit Runge-Kutta scheme:

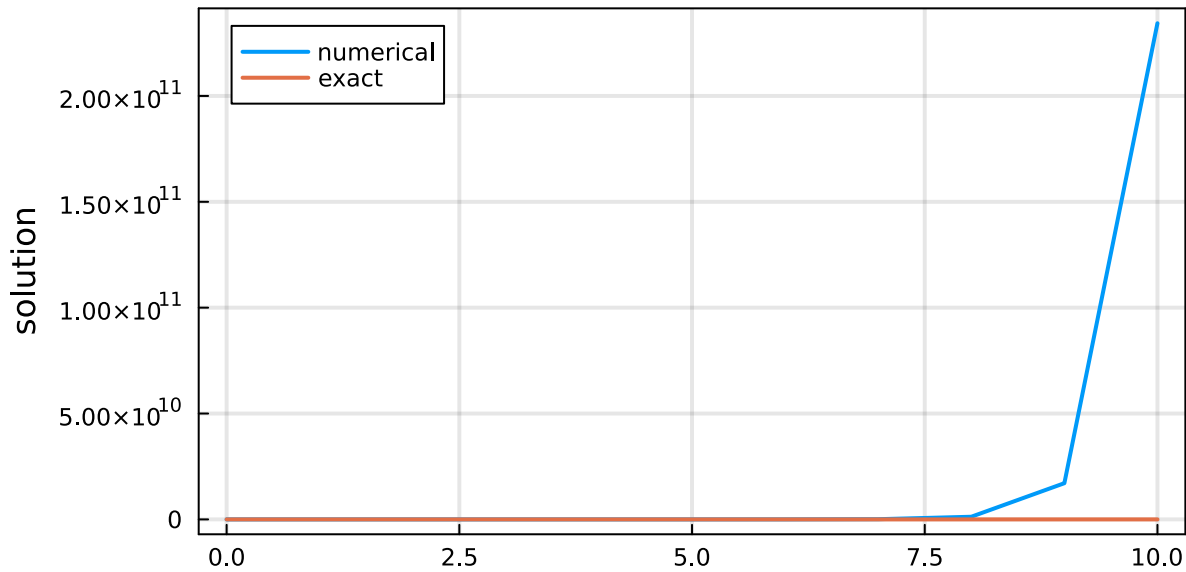
explicit\_rk (generic function with 1 method)

```
1 function explicit_rk(f!::F, # ODE right hand side
2     p; # Parameters
3     tableau = TableauRK4(),
4     u0 = [0.0], # vector of initial values
5     dt = 0.1, # time step size
6     tspan = (0, 1),) where {F}
7     u = [u0]
8     t = Float64[tspan[1]]
9     a = tableau.a
10    b = tableau.b
11    c = tableau.c
12    m = length(u0)
13    s = length(b)
14    du = zeros(m)
15    z = zeros(m, s)
16    while t[end] < tspan[end]
17        z .= 0
18        for j = 1:s
19            du .= 0
20            for k = 1:(j - 1)
21                @views du .+= z[:, k] * a[j, k]
22            end
23            @views f!(z[:, j], u[end] .+ dt * du, p, t[end] + c[j] * dt)
24        end
25        du .= 0
26        for j = 1:s
27            @views du .+= b[j] * z[:, j]
28        end
29        u_new = u[end] + dt * du
30        t_new = t[end] + dt
31        push!(u, u_new)
32        push!(t, t_new)
33    end
34    DiffEqArray(u, t)
35 end
```

```
1 tableau = TableauRK4();
```

```
1 urk_test = explicit_rk(f_test!, p_test; u0 = [1.0], tableau, dt = 10^logdt_rk,
  tspan = (0, 10));
```

$\log(\tau) =$   0.0



Once again we see that for a "decent" solution, the time step size should be not too large. We need to discuss two phenomena:

- What about the "wiggles" which destroy the overall quality of solutions  $\Rightarrow$  stability
- How good is the approximation  $\Rightarrow$  approximation order

## Stability

Let us discuss aspects of stability (resp. instability).

Accidentally, the test problem  $\dot{u} = \lambda u$  chosen here is also called *Dahlquist test equation*. It is used to judge the stability of ODE methods.

## A-Stability

**Definition:** An ODE method is called unconditionally A-stable if for  $\lambda \in \mathbb{C}$  with  $\operatorname{Re}\lambda < 0$ , the approximate solution of the Dahlquist test equation fulfills  $|u_{n+1}| \leq |u_n|$ .

The method is called conditionally A stable if  $|u_{n+1}| \leq |u_n|$  is fulfilled for  $h\lambda \in \mathcal{S} \subset \mathbb{C}$ , where  $\mathcal{S}$  is called stability region.

In order to investigate A-Stability, it is useful to define the stability function  $R$  such that

$$u_{n+1} = R(\tau\lambda)(u_n)$$

For the  $\theta$ -methods this gives:

$$1 \quad R_{\theta 0}(z) = 1 + z;$$

$$1 \quad R_{\theta 0.5}(z) = (1 + z / 2) / (1 - z / 2);$$

$$1 \quad R_{\theta 1}(z) = 1 / (1 - z);$$

For generic Runge-Kutta methods, we have:

```
1 function R_rk(z, tableau = TableauRK4())
2     A = tableau.a
3     b = tableau.b
4     Ib = ones(length(b))
5     1 + z * b' * inv(I - z * A) * Ib
6 end;
```

## L-Stability

An A-stable one-step method is called L-stable if for all  $\lambda \in \mathbb{C}$  with  $\operatorname{Re}\lambda < 0$ , the solution of the Dahlquist test problem fulfills

$$\lim_{n \rightarrow \infty} |u_n| = 0$$

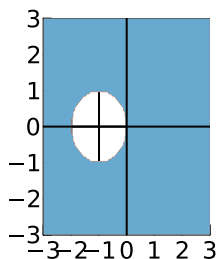
independent of the stepsize  $\tau$ .

L-Stability is fulfilled if and only if for all  $\lambda \in \mathbb{C}$  with  $\operatorname{Re}\lambda < 0$ ,

$$\lim_{\tau \rightarrow \infty} R(\tau\lambda) = 0$$

stabregion (generic function with 1 method)

Stability of the theta methods:

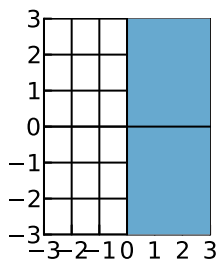


Explicit Euler

$$R(z) = 1 + z$$

Conditionally A-stable

not L-stable

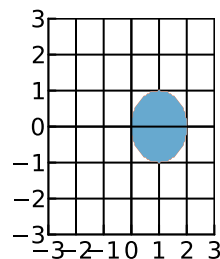


Implicit Midpoint

$$R(z) = \frac{2+z}{2-z}$$

A-stable

not L-stable



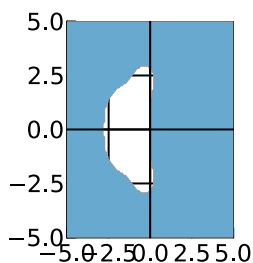
Implicit Euler

$$R(z) = \frac{1}{1-z}$$

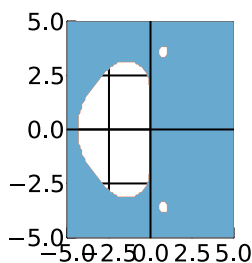
A-stable

L-stable

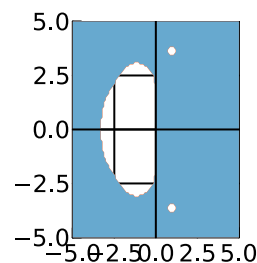
Explicit Runge-Kutta methods are conditionally A stable:



RK4

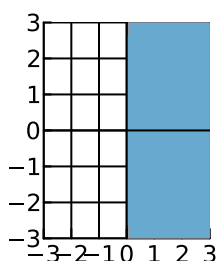


DP4



DP5

Implicit Runge-Kutta methods are A-stable, some of them are L-stable like RadauIIA(5):

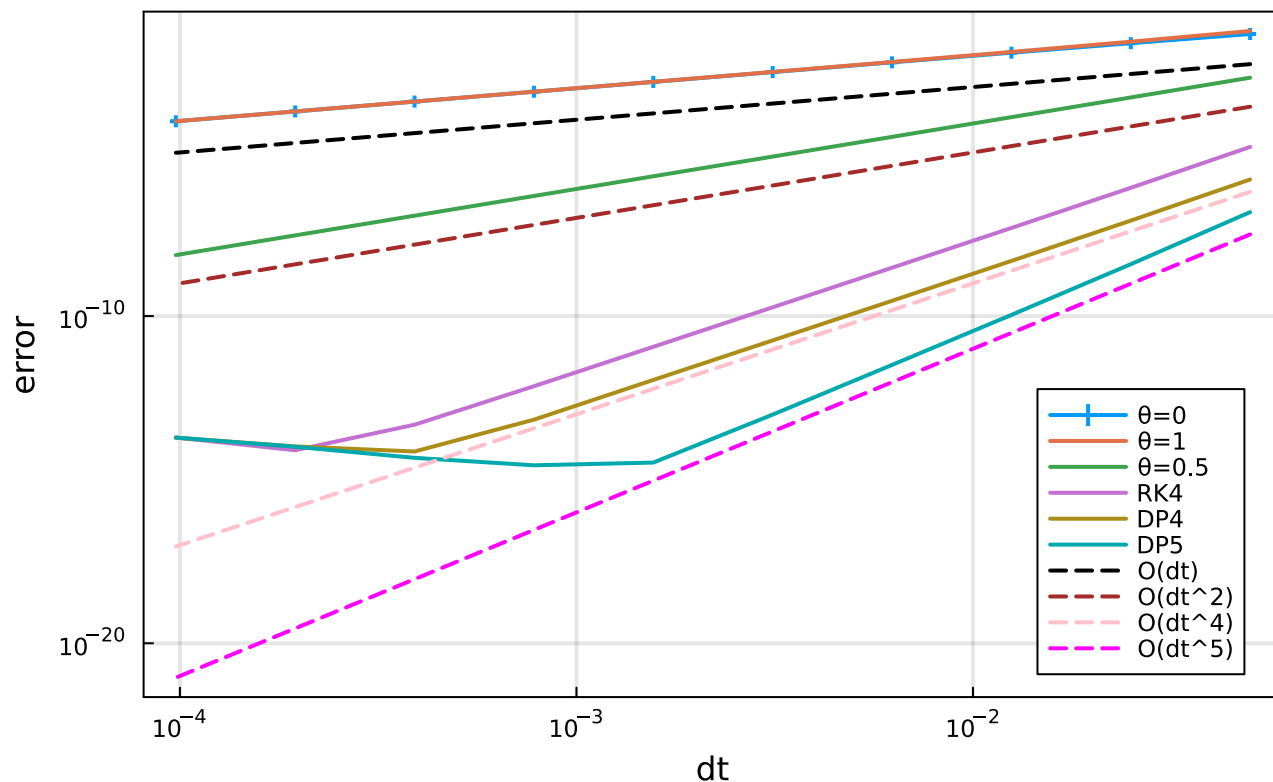


```
1 stabregion(z -> R_rk(z, TableauRadauIIA(3)))
```

## Approximation order

Approximation order can generally be analysed using Taylor expansion to estimate the consistency order. We only demonstrate the results here.

errortest (generic function with 3 methods)



- Implicit Euler and explicit Euler exhibit first order of approximation
- Implicit midpoint has second approximation order
- Runge-Kutta methods can have higher order

These methods are one-step methods, which just involve two subsequent time steps. Multistep methods (e.g. BDF - backward differencing formula) achieve higher order by taking into account more of the older time steps, however they have a problem to keep their order at the start.

## An ODE system

We demonstrate the Lotka-Volterra system describing predator-prey dynamics:

- species  $x$ : prey which growth rate  $a$  being eaten with rate  $by$  by the predators
- species  $y$ : predators growing with rate  $dx$  by eating prey and dying with rate  $c$

lotkavolterra! (generic function with 1 method)

```
1 function lotkavolterra!(du, u, p, t)
2   (a, b, c, d) = p
3   x, y = u[1], u[2]
4   du[1] = a * x - b * x * y
5   du[2] = -c * y + d * x * y
6 end
```

The function  $V$  is constant along trajectories – it is easy to show that  $\frac{dV}{dt} = 0$ . Therefore in the phase-space the trajectories are isolines of  $V(x, y)$ , and the solution must be periodic.

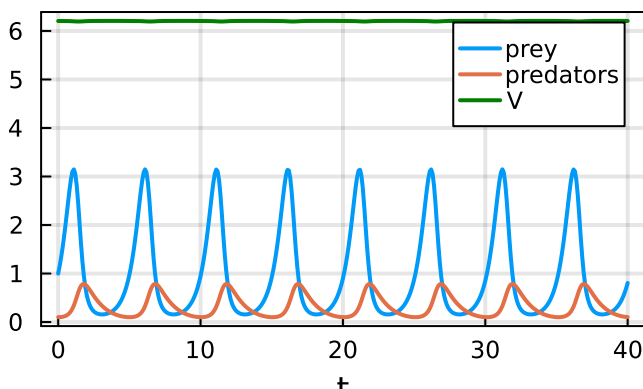
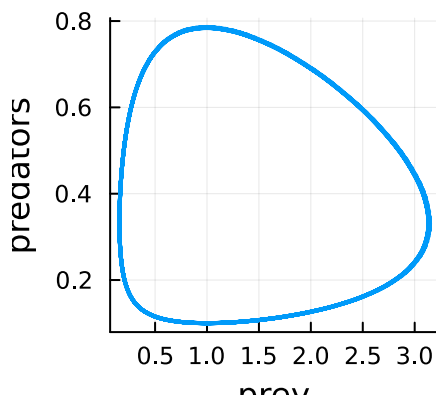
```
1 function V(x::Number, y::Number, param)
2   a, b, c, d = param
3   x = max(1.0e-15, x)
4   y = max(1.0e-15, y)
5   d * x - c * log(x) + b * y - a * log(y)
6 end;
```

param = (a = 2, b = 6, c = 1, d = 1)

```
1 param = (a = 2, b = 6, c = 1, d = 1)
```

```
1 ulv = let
2   rklv ?
3   explicit_rk(lotkavolterra!, param; u0 = [1.0, 0.1], dt = 10^logdtlv, tspan =
4   (0, 40)) :
5   θmethod(lotkavolterra!,
6   param;
7   u0 = [1.0, 0.1],
8   dt = 10^logdtlv,
9   tspan = (0, 40),
10  θ = θlv,
11  linear = linlv,)
11 end;
```

θ =  0.5 log(dt) =  -1.0 linear =  RK4 =



- The implicit midpoint rule catches periodicity - surprisingly also for rather coarse time steps. The reason is that it preserves more of the structure of the system. This is the simplest *geometrical integrator* - a class of methods which exactly preserves certain invariants of the solution
- The Runge-Kutta methods does a quite good job in catching periodicity
- Implicit and explicit Euler are unable to catch the dynamics in the right way.
- The linear implicit approach works well

## Stiff problems

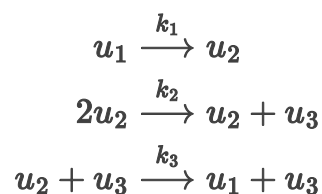
Problems with a high ratio between the real parts of the largest and the smallest eigenvalues of  $\frac{\partial f}{\partial u}$  (spectral condition number  $\kappa$  for spd problems) are called *stiff*.

This is e.g. the case with different time scales which may be due to fast and slow reactions, and for systems arising from the discretization of PDEs, where the condition number increases with  $O(h^{-2})$ , where  $h$  is the space discretization parameter.

Stability then depends on the stability constraint for the smallest timescale.

Unconditionally A-stable and L-stable methods are to be preferred if we want to guarantee stability for stiff problems.

Let us consider a chemical reaction ("Robertson reaction" example from matlab)



fstiff! (generic function with 1 method)

```
1 function fstiff!(du, u, p, t)
2     k1, k2, k3 = p
3     du[1] = -k1 * u[1] + k3 * u[2] * u[3]
4     du[2] = k1 * u[1] - k3 * u[2] * u[3] - k2 * u[2]^2
5     du[3] = k2 * u[2]^2
6 end
```

For this problem,  $d$  can be seen as "stiffness" parameter. The larger  $d$  the shorter the time scale of the second equation.

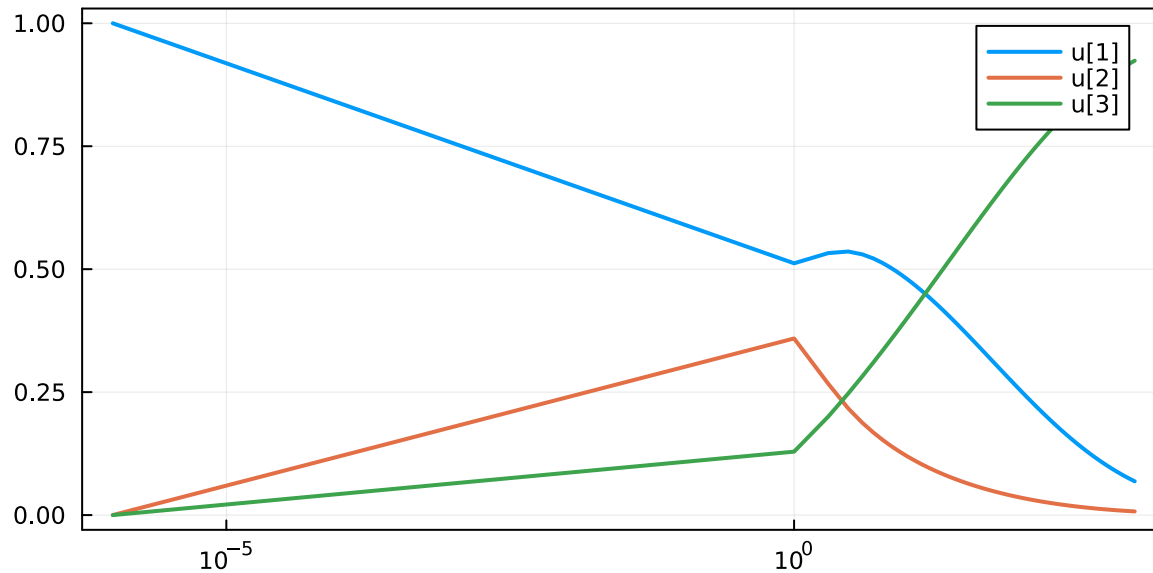
```
probertson = (k1 = 10, k2 = 1.0, k3 = 100)
```

```
1 probertson = (k1 = 10, k2 = 10loga, k3 = 100)
```

```
1 ustiff = @method(fstiff!,
2                 probertson;
3                 u0 = [1.0, 0.0, 0.0],
4                 dt = 10logdt,
5                 θ = θx2,
6                 linear = lin2,
7                 tspan = (1.0e-6, 1.0e3));
```

θ=  1.0 log(dt)=  0.0 linear=

Stiffness log(a)=  0.0



```
1 plot(ustiff;
2     label = ["u[1]" "u[2]" "u[3]"],
3     size = (600, 300),
4     linewidth = 2,
5     framestyle = :box,
6     xscale = :log10)
```

Due to the fact that the implicit midpoint rule is not L-stable, we can have problems with large stiffness.

## DAE problems



Differential-algebraic equation (DAE) system:

$$0 = f(\dot{u}, u, p, t)$$

Here, we will focus on mass matrix DAE systems (MMDAE):

$$M\dot{u} = f(u, p, t)$$

M is a diagonal matrix, some entries may be zero. The corresponding equations in the system turn from differential equations to algebraic equations.

We can see a DAE as the "large stiffness limit" case of an ODE system, so once again we want to use A-stable and L-stable methods.

Here, we re-define the Robertson example using the fact that the sum of all three reactants must be 1.



fdae! (generic function with 1 method)

```
1 function fdae!(du, u, p, t)
2     k1, k2, k3 = p
3     du[1] = -k1 * u[1] + k3 * u[2] * u[3]
4     du[2] = k1 * u[1] - k3 * u[2] * u[3] - k2 * u[2]^2
5     du[3] = u[1] + u[2] + u[3] - 1
6 end
```

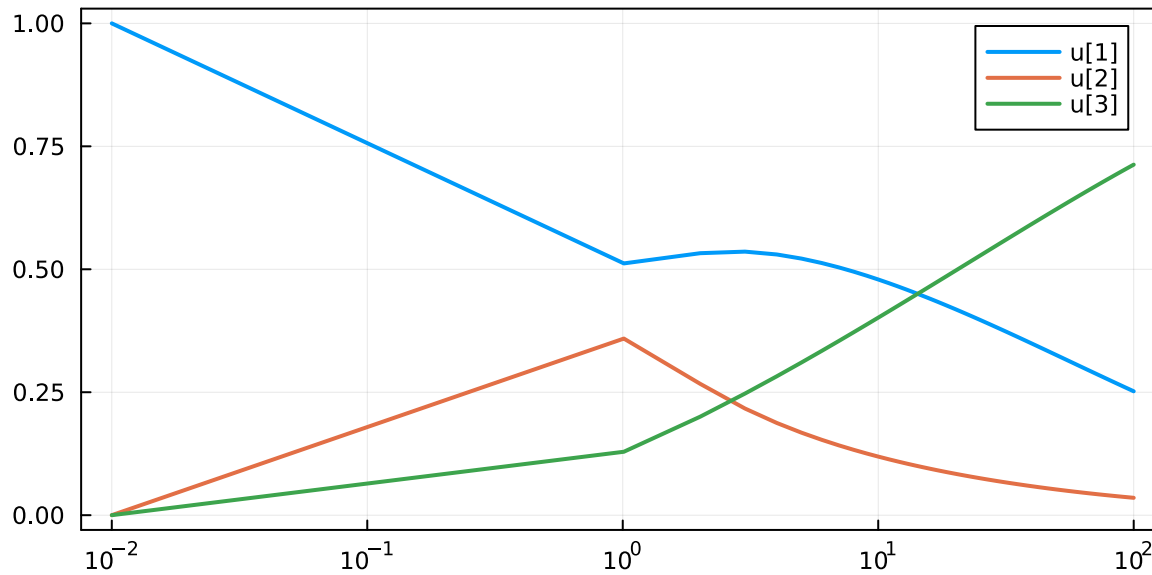
M\_DAE = 3×3 Diagonal{Float64, Vector{Float64}}:

```
 1.0  .  .
 .  1.0  .
 .  .  0.0
```

```
1 M_DAE = Diagonal([1.0, 1, 0])
```

θ=  1.0 log(dt)=  0.0 linear=

Stiffness log(a)=  0.0



```
1 plot(udae;  
2     label = ["u[1]" "u[2]" "u[3]"],  
3     size = (600, 300),  
4     linewidth = 2,  
5     framestyle = :box,  
6     xscale = :log10)
```

As for stiff problems, we want implicit methods to solve DAE problems.

## Multistep methods

Multistep methods attempt to take into account information from timesteps prior to  $t_n$ .

## BDF Methods

For a given vector  $X = x_1 \dots x_n$ , and function values  $Y = \{y_i = f(x_i)\}_{i=1 \dots n}$  a Lagrange polynomial  $L[X, Y]$  interpolates the function  $f$  in the interval  $[x_1, x_n]$ . It is given by

$$L[X, Y](\xi) = \sum_{j=1}^n l_j[X](\xi) y_j$$

with

$$l_j[X](\xi) = \prod_{i=1, i \neq j}^n \frac{x - x_i}{x_j - x_i}$$

From the very definition we have  $l_j[X](x_i) = \delta_{ij}$ . The interesting point is that we can hope that  $L'[X](x_n)$  is a good approximation of  $f'(x_n)$ . This idea gives another way to approximate the derivative on the left hand side of the ODE, taking into account results from older timesteps.

In particular, for the case of three interpolation points  $X = \{a, b, c\}$  we get

$$\begin{aligned} l_a[X](\xi) &= \frac{(\xi - b)(\xi - c)}{(a - b)(a - c)} & l'_a[X](\xi) &= \frac{2\xi - b - c}{(a - b)(a - c)} & l'_a[X](c) &= \frac{c - b}{(a - b)(a - c)} \\ l_b[X](\xi) &= \frac{(\xi - a)(\xi - c)}{(b - a)(b - c)} & l'_b[X](\xi) &= \frac{2\xi - a - c}{(b - a)(b - c)} & l'_b[X](c) &= \frac{c - a}{(b - a)(b - c)} \\ l_c[X](\xi) &= \frac{(\xi - a)(\xi - b)}{(c - a)(c - b)} & l'_c[X](\xi) &= \frac{2\xi - a - b}{(c - a)(c - b)} & l'_c[X](c) &= \frac{2c - a - b}{(c - a)(c - b)} \end{aligned}$$

Let  $a = t^{n-1}, b = t^n, c = t^{n+1}$

Setting  $Y = \{y_a, y_b, y_c\}$  gives

$$L'[X, Y](c) = \frac{1}{2\tau} (y_a - 4y_b + 3y_c)$$

For  $y_c = u^{n+1}, y_b = u^n, y_a = u^{n-1}$  gives rise to the BDF2 (Backward Differencing Formula) scheme:

$$\frac{1}{2} (3u^{n+1} - 4u^n + u^{n-1}) = \tau f(u^{n+1}, p, t^{n+1})$$

BDF schemes are implicit. Unlike the higher order pendants, BDF1 and BDF2 are A-stable, and thus useful for stiff problems. BDF2 has consistency order 2.

It is easy to see that this procedure can be adapted to varying step sizes. Further refined variants ensure that under stepsize variations the coefficient before  $u_{n+1}$  stays unchanged (Fixed leading coefficient BDF in SUNDIALS; FBDF in Julia). More information is available via Scholarpedia.

A higher order correction of the BDF methods leads to the QNDF method (quasi constant timestep numerical differentiation formula) which is equivalent to the `ode15s` method implemented in the matlab ODE suite. QNDF is the preferred default solver for stiff systems in Julia

## Adams methods

A similar interpolation approach can be applied to the right hand side of the ODE system, giving rise e.g. to the explicit Adams-Bashforth methods.

# DifferentialEquations.jl

---

So far, we did our own implementations of some methods discussed. These implementations miss a number of points:

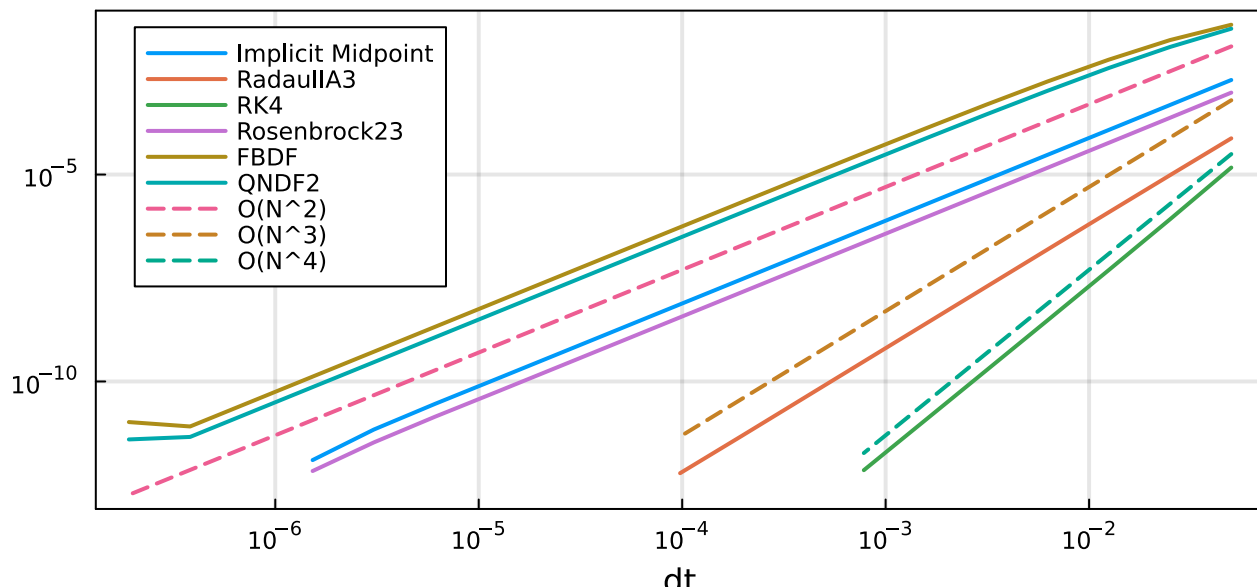
- Time step adaptivity. A general approach to time step adaptivity considers the comparison between two solutions of different consistency order to obtain an error estimate.
- Multistage linear implicit methods (Rosenbrock methods). The simple methods of this class are the linear implicit methods which replace the nonlinear solution by just one Newton step. Higher order variants combine a fixed number of intermediate linear steps.

Julia provides a state of the art toolbox for the solution of systems of differential equations: DifferentialEquations.jl. The overview on implemented methods is given in the documentation:



We choose a subset of the methods implemented there and return to the example problems we discussed:

```
diffeqmethods =  
OrderedDict("Implicit Euler" => ImplicitEuler, "Implicit Midpoint" => ImplicitMidpoint,
```



## Lotka-Volterra revisited

analyze (generic function with 1 method)

```
param2 = (a = 2, b = 6, c = 1, d = 1)
```

```
1 param2 = (a = 2, b = 6, c = 1, d = 1)
```

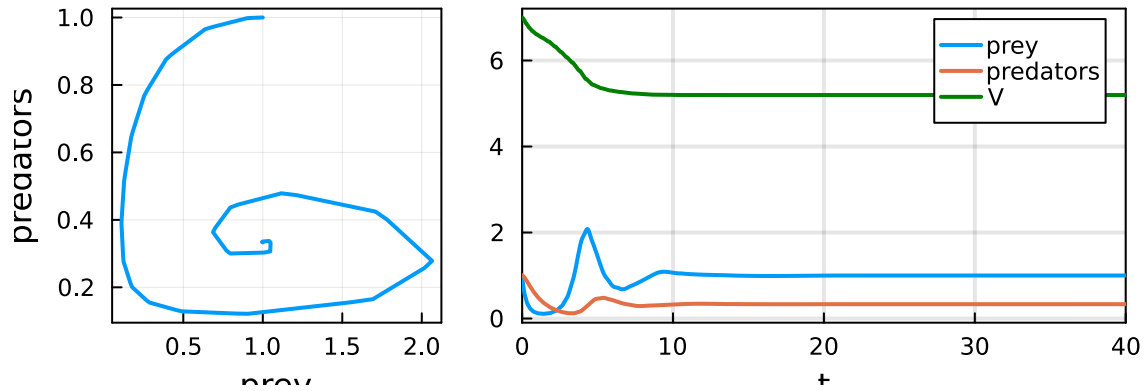
```
1 ulv_deq = let
2   prob = ODEProblem(lotkavolterra!, [1.0, 1.0], (0, 40), param2)
3   method = diffeqmethods[lvmethod]
4   if method == ImplicitMidpoint
5     dt = 10logreltol
6     @info "implicit midpoint: dt=$(dt)"
7     solve(prob, method(); dt)
8   else
9     reltol = 10logreltol
10    @info "reltol=$(reltol)"
11    solve(prob, method(); reltol)
12  end
13 end;
```

reltol=0.1

method:

(timesteps = 53, smallest = 1.715e-5, largest = 16.6289)

```
1 analyze(ulv_deq)
```



## Stiffness + DAE revisited

```

1  ustiff_deq = let
2      prob = ODEProblem(fstiff!, [1.0, 0, 0], (1.0e-2, 10), (k1 = 10, k2 = 10logst,
3          k3 = 100))
4      solve(prob, diffeqmethods[stmethod](); reltol = 10logstreltol)
5  end;

```

(timesteps = 43, smallest = 3.4641e-11, largest = 4.02916)

```

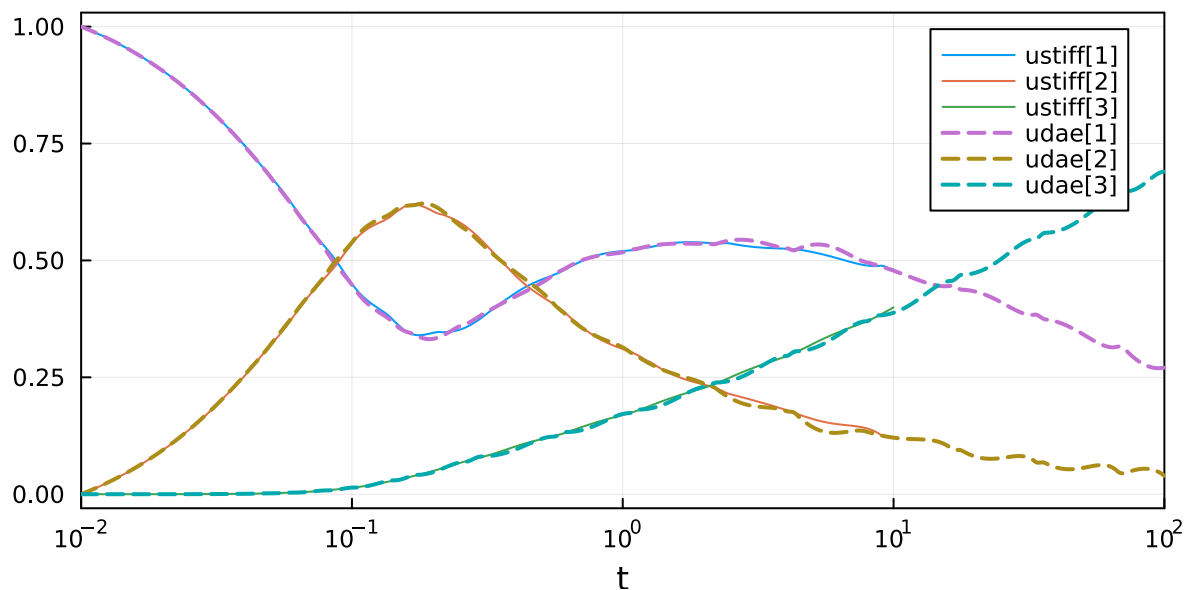
1  udae_deq = let
2      fdae = ODEFunction(fdae!; mass_matrix = Diagonal([1, 1, 0]))
3      prob = ODEProblem(fdae, [1.0, 0, 0], (1.0e-2, 100), (k1 = 10, k2 = 10logst, k3
4          = 100))
5      solve(prob, diffeqmethods[stmethod](); reltol = 10logstreltol)
6  end;

```

(timesteps = 42, smallest = 2.0e-7, largest = 30.6452)

Method:  ▾

Stiffness  $\log(d)$  =  0.0  $\log(\text{reltol})$ :  -1.0



## Conclusions

---

- There is a whole world of ODE and DAE solvers. `DifferentialEquations.jl` covers a significant part of this world
- ODE systems created from PDE discretizations are *large* and they are *stiff*, i.e. they contain strongly differing timescales
- Stability of explicit methods is limited by that of the fastest timescale, they are not feasible for stiff problems
- Explicit methods do not work for DAEs by construction
- A- and L- stable methods are required for stiff problems

The general advise when solving a particular problem is to use the capabilities of `DifferentialEquations.jl` and to find out which methods work best

---

### Helper cells



# Table of Contents

## ODE and DAE Solvers

- Explicit Euler method

  - Test problem

- $\theta$ - and Runge-Kutta methods

  - $\theta$  method: implementation

  - Runge-Kutta methods

- Stability

  - A-Stability

  - L-Stability

- Approximation order

- An ODE system

- Stiff problems

- DAE problems

- Multistep methods

  - BDF Methods

  - Adams methods

- DifferentialEquations.jl

  - Lotka-Volterra revisited

  - Stiffness + DAE revisited

- Conclusions

@statement\_str (macro with 1 method)