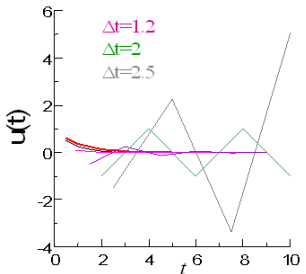


pozbyć się ograniczenia na krok czasowy ze strony bezwzględnej stabilności: niejawna metoda Eulera



$$\frac{du}{dt} = f(t, u)$$

$$u_n = u_{n-1} + f(t_{n-1}, u_{n-1})\Delta t$$

jawna metoda Eulera

jawna metoda Eulera
(funkcjonuje jak podstawienie)

„metoda odważna”

$$u_n = u_{n-1} + f(t_n, u_n)\Delta t$$

niejawna metoda Eulera

funkcjonuje jak równanie nieliniowe

„metoda ostrożna”
zmiana u zgodna z prawą stroną
w punkcie docelowym

niejawna metoda Eulera:

problem początkowy:

$$u' = -100u, \quad u(0) = 1$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

jawny Euler

t_n	u_n
0	1
0.05	-4
0.1	8
0.15	-16
0.2	256
0.25	-1024
0.3	4096

niejawny Euler

$$u_n = u_{n-1} - 100\Delta t u_n$$

$$u_n = \frac{u_{n-1}}{(1 - \Delta t \lambda)}$$

W naszym problemie $u_n = 1/6^n$

t_n	u_n	$e(t_n)$
0	1	0
0.05	.166(6)	-.15992
0.1	.027(7)	-.02773
0.15	.004(629)	-.00462
0.2	.0007716	
0.25	.0001286	
0.3	.00002143	

itd..

$\exp(-100 t_n)$ gaśnie
znacznie szybciej niż $1/6^n$

mało dokładne, ale zawsze to lepiej niż
eksplodująca oscylacja jawnego Eulera

niejawna metoda Eulera:
region bezwzględnej stabilności

$$\frac{du}{dt} = \lambda u$$

$$u_n = u_{n-1} + \Delta t \lambda u_n$$

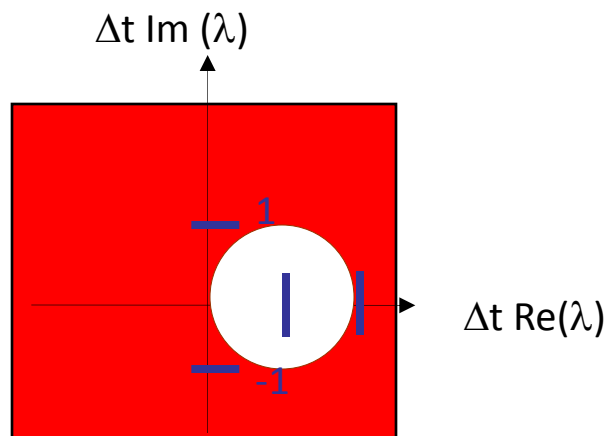
$$u_n = \frac{u_{n-1}}{(1 - \Delta t \lambda)}$$

$$u_n = \frac{u_0}{(1 - \Delta t \lambda)^n}$$

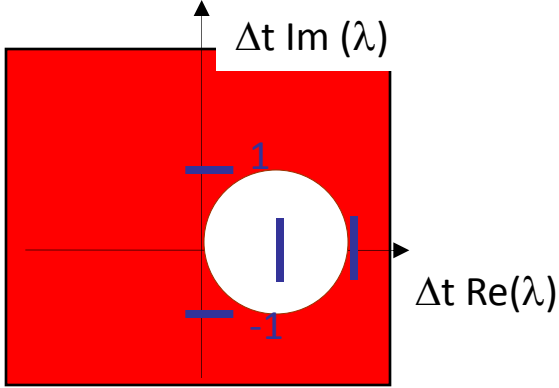
$$\longrightarrow \left| \frac{1}{(1 - \Delta t \lambda)^n} \right| \leq 1$$



$$|1 - z| \geq 1$$



region bezwzględnej stabilności:
dopełnienie pustego koła o środku w (1,0)
i promieniu 1

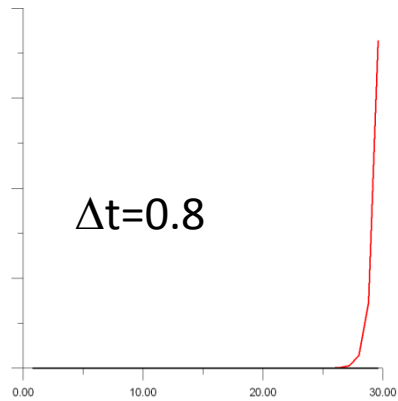
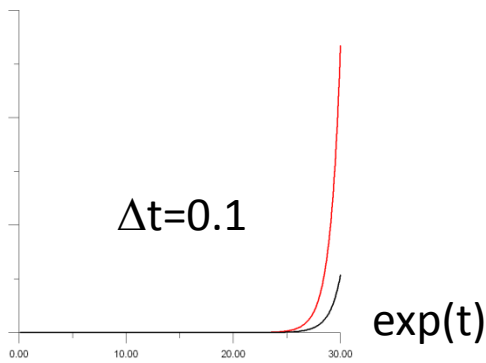


$$\frac{du}{dt} = \lambda u$$

$$u_n = \frac{u_{n-1}}{(1 - \Delta t \lambda)}$$

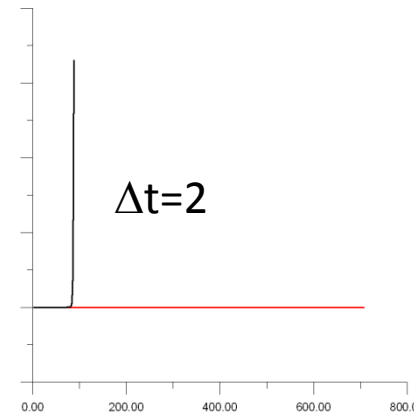
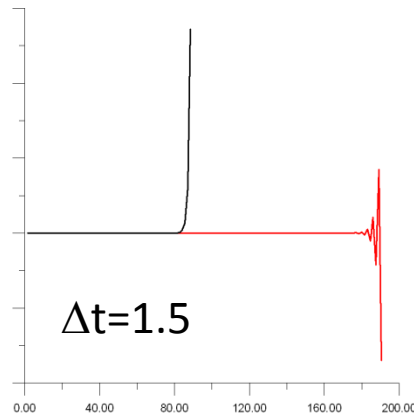
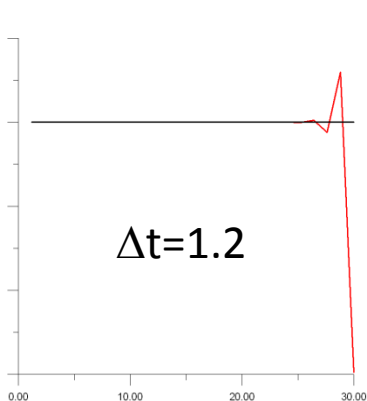
Niejawny schemat Eulera

$\lambda=1$ – zakres niestabilności $\Delta t \in (0,2)$



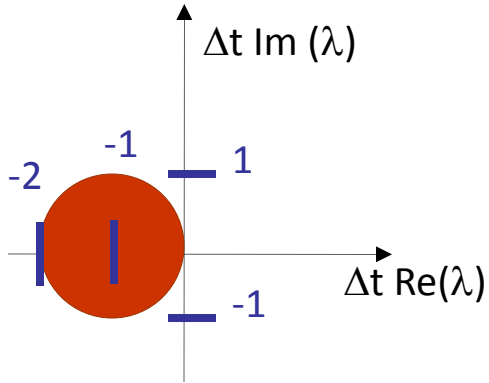
Zbliżamy się do $\Delta t=1$ –
wyniki schematu rosną coraz szybciej

Dla $\Delta t=1$ –
nieskończoność w pierwszym kroku

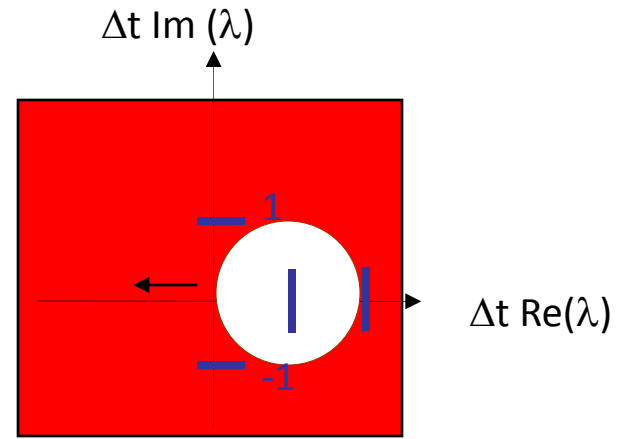


1,-1,1,-1 itd

regiony stabilności metod Eulera

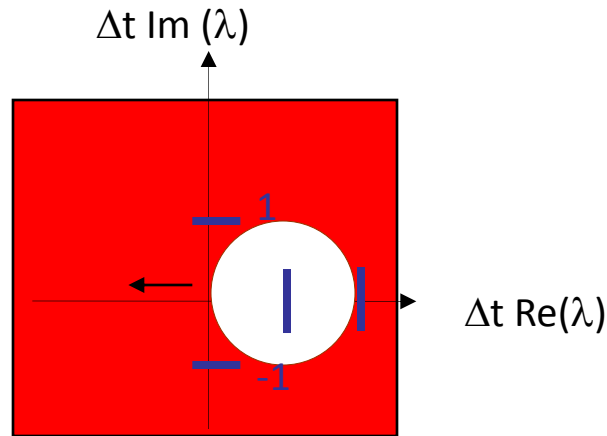


metoda Eulera jawna



niejawna metoda Eulera

niejawna metoda Eulera:
region bezwzględnej stabilności



$$\frac{du}{dt} = \lambda u$$

$\operatorname{Re}(\lambda) < 0$: niejawny Euler bezwzględnie stabilny
dla dowolnego kroku czasowego

take metody: tzw. A-stabilne

dla $\operatorname{Re}(\lambda) > 0$, poza kołem metoda Eulera
jest bezwzględnie stabilna
mimo, że rozwiązania równania różniczkowego
są niestabilne (patrz wyżej)
w tym obszarze metoda jest **nadstabilna**
daje skończone wartości, mimo że
rozwiązania dokładne dążą do nieskończoności

bezwzględna stabilność
nie oznacza dobrej dokładności.
W regionie nadstabilności
dla $\operatorname{Re}(\lambda) > 0$ błędy będą rosły w nieskończoność.

$$\frac{du}{dt} = f(t, u)$$

$$u_n = u_{n-1} + f(t_n, u_n) \Delta t$$

jak rozwiązać, gdy nie można rozwinąć równania (f nieliniowe względem u)
lub gdy f nieznane w formie wzoru (np. gdy piszemy program dla ogólnego f)

1) iteracja funkcjonalna

$$u_n^0 = u_{n-1}$$

$$u_n^1 = u_{n-1} + \Delta t f(t_n, u_n^0)$$

$$u_n^\mu \overset{\swarrow}{=} u_{n-1} + \Delta t f(t_n, u_n^{\mu-1})$$

iterować do zbieżności

jeśli się zbiegnie $u^\mu = u^{\mu-1}$ i mamy rozwiązanie równania nieliniowego

iteracja funkcjonalna przykład

problem początkowy:

$$u' = -100u, \quad u(0) = 1, \quad dt = 0.05$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

$$u_n = \frac{u_{n-1}}{(1 + \lambda \Delta t)} = \frac{u_{n-1}}{6}$$

t_n	u_n	$e(t_n)$
0	1	0
0.05	.166(6)	-.15992
0.1	.027(7)	-.02773
0.15	.004(629)	-.00462
0.2	.0007716	
0.25	.0001286	
0.3	.00002143	

$$u_n^\mu = u_{n-1} + \Delta t f(t_n, u_n^{\mu-1})$$

$$u_1^0 = 1$$

$$u_1^\mu = 1 - 5u_1^{\mu-1}$$

kolejne oszacowania:

1, -4, 21, -104, 521, -2604, ...

iteracja się nie zbiega ☹️

cały zysk z niejawności stracony
bo nie potrafimy wykonać kroku

iteracja funkcjonalna przykład

iteracja się nie zbiega ☹️. zmniejszymy krok dt ,
zaczynając iterację od u_{n-1} będziemy bliżej rozwiązania. Może się zbiegnie.

$$dt=0.01$$

$$u_1^0 = 1$$

$$u_1^\mu = 1 - u_1^{\mu-1}$$

(1,0,1,0,1,0)

$$dt=0.001$$

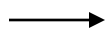
$$u_1^0 = 1$$

$$u_1^\mu = 1 - u_1^{\mu-1}/10$$

1, 0.9, 0.91, 0.909, 0.9091, 0.90909, 0.909091,
... 0.90909090909

iteracja funkcjonalna się zbiega gdy $\Delta t \max |f_u(t,u)| \leq 1$ (w otoczeniu u)

$$u' = -100u,$$



$\Delta t 100 < 1$ uwaga: w tej sytuacji
jawny Euler jest bezwzględnie stabilny dla 2-krotnie większego kroku!
[dla jawnego Eulera $\Delta t 100 < 2$]
Z iteracją funkcjonalną stosować wsteczny Euler nie ma sensu.

problem początkowy:

$$u' = -100u, \quad u(0) = 1, \quad dt = 0.05$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

1, -4, 21, -104, 521, -2604, ...

oscylująca rozbieżność
- stłumimy ją:

iteracja funkcjonalna – zapewniamy zbieżność modyfikując przepis iteracyjny

zamiast:

$$u_n^\mu = u_{n-1} + \Delta t f(t_n, u_n^{\mu-1})$$

„mieszając” nowe i stare rozwiązania z wagą w , $0 \leq w \leq 1$

$$u_n^\mu = (1 - w)u_n^{\mu-1} + w \left(u_{n-1} + \Delta t f(t_n, u_n^{\mu-1}) \right)$$

Zabieg podobny do “podrelaksacji”

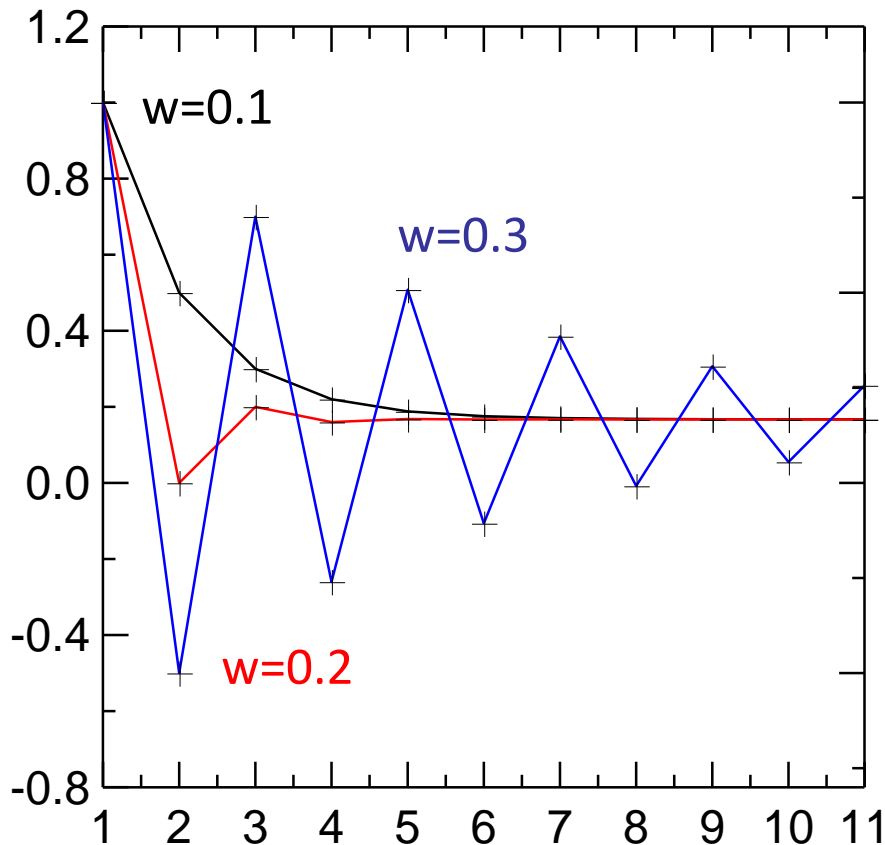
jeśli się zbiegnie –
to do rozwiązania schematu
niejawnego

problem początkowy:

$$u' = -100u, \quad u(0) = 1, \quad dt = 0.05$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

$$u_n^\mu = (1 - w)u_n^{\mu-1} + w \left(u_{n-1} + \Delta t f(t_n, u_n^{\mu-1}) \right)$$



iterujemy $u(dt)$

wybierając w odpowiedni sposób wagę w : potrafimy ustabilizować iterację i doprowadzić ją do zbieżności.

$$u_n^\mu = (1 - w)u_n^{\mu-1} + w (u_{n-1} + \Delta t f(t_n, u_n^{\mu-1}))$$

dt=0.01 w=0 (1,0,1,0,1,0)

w=.2 (optymalne dla dt=0.05)

0.8,0.68, 0.608, 0.5648, 0.53888, 0.5233, 0.51399,
0.50839, 0.50503, 0.503, 0.5018, 0.5010, 0.50065,
0.50039, ..., 1/2

tutaj optymalne byłoby w=1/2
(zbieżność w jednej iteracji)

dla w=.7
0.3,0.58,0.468,0.512,0.4948,0.5003,0.4998

dt=0.001

w=1

1, 0.9, 0.91, 0.909, 0.9091,
0.90909, 0.909091,
... 0.90909090909

w=0.2

```

1  0.9799999998752028
2  0.964399997863173
3  0.952231997235827
4  0.942740956797946
5  0.935337945296528
6  0.929563596356724
7  0.925059604208092
8  0.921546490351203
9  0.918806261557684
10 0.916668883110325
11 0.915001727930422
12 0.913701346897147
13 0.912687049696691
14 0.911895897884623
15 0.911278799474556
16 0.910797462717313
17 0.910422020048698
18 0.910129174768766
19 0.909900755451658
20 0.909722588385279
21 0.909583618074257
22 0.909475221232247
23 0.909390671695937

```

$$u_n^\mu = (1 - w)u_n^{\mu-1} + w (u_{n-1} + \Delta t f(t_n, u_n^{\mu-1}))$$

dt=0.01 w=0 (1,0,1,0,1,0)

w=.2 (optymalne dla dt=0.05)

0.8,0.68, 0.608, 0.5648, 0.53888, 0.5233, 0.51399,
0.50839, 0.50503, 0.503, 0.5018, 0.5010, 0.50065,
0.50039, ..., 1/2

dt=0.001

w=1

1, 0.9, 0.91, 0.909, 0.9091, 0.90909, 0.909091,
... 0.90909090909

w=0.2

1	0.979999998752028
2	0.964399997863173
3	0.952231997235827
4	0.942740956797946
5	0.935337945296528
6	0.929563596356724
7	0.925059604208092
8	0.921546490351203
9	0.918806261557684
10	0.916668883110325
11	0.915001727930422
12	0.913701346897147
13	0.912687049696691
14	0.911895897884623
15	0.911278799474556
16	0.910797462717313
17	0.910422020048698
18	0.910129174768766
19	0.909900755451658
20	0.909722588385279
21	0.909583618074257
22	0.909475221232247
23	0.909390671695937

tutaj optymalne byłoby w=.5
(zbieżność w jednej iteracji natychmiastowa)

Problem:

- 1) w trzeba odpowiednio dobrać (mniejszy krok czasowy, w bliższe 1)
- 2) dla źle dobranego w iteracja może być wolnozbieżna

Proces optymalizacji (np. dynamicznej) w może być kłopotliwy.

$$\frac{du}{dt} = f(t, u)$$

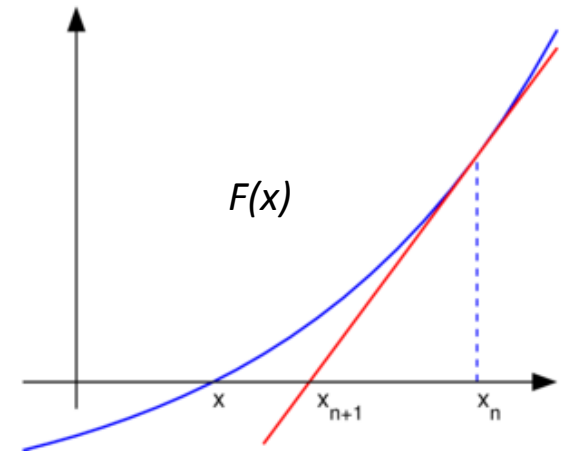
$$u_n = u_{n-1} + f(t_n, u_n) \Delta t$$

Na szczęście nie jesteśmy skazani na iterację funkcjonalną

2) metoda Newtona-Raphsona (stycznych)

szukamy zera równania nieliniowego

$$F(u_n) = u_n - u_{n-1} - f(t_n, u_n) \Delta t$$



$$F(x_n + \Delta x) = F(x_n) + \Delta x F'(x_n)$$

$$F(x_n + \Delta x) = 0$$

$$x_{n+1} = x_n - F(x_n) / F'(x_n)$$

2) metoda Newtona-Raphsona

szukamy zera równania nieliniowego

$$F(u_n) = u_n - u_{n-1} - f(t_n, u_n) \Delta t$$

$$F(x_n + \Delta x) = F(x_n) + \Delta x F'(x_n)$$

$$F(x_n + \Delta x) = 0$$

$$x_{n+1} = x_n - F(x_n) / F'(x_n)$$

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$

niejawny schemat Eulera z metodą Newtona-Raphsona, zastosowanie

problem początkowy:

$$u' = -100u, \quad u(0) = 1$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

$$F(u_n) = u_n - u_{n-1} - f(t_n, u_n) \Delta t$$

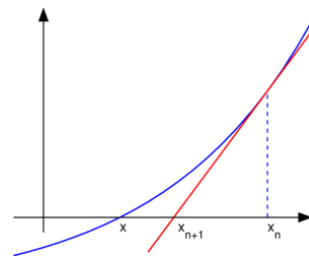
$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} + 100 \Delta t u_n^{\mu-1}}{1 + 100 \Delta t}$$

kolejne przybliżenia: $\Delta t = 0.05$ (jawny Euler stabilny bezwzględnie dla $\Delta t < 0.02$)
1, 0.1666677, 0.1666677 zbieżność w jednej iteracji - F jest liniowa w u

*Wniosek: dla liniowych f liniowe jest również F
wtedy iteracja Newtona zbiega się w jednej iteracji niezależnie od wielkości Δt*

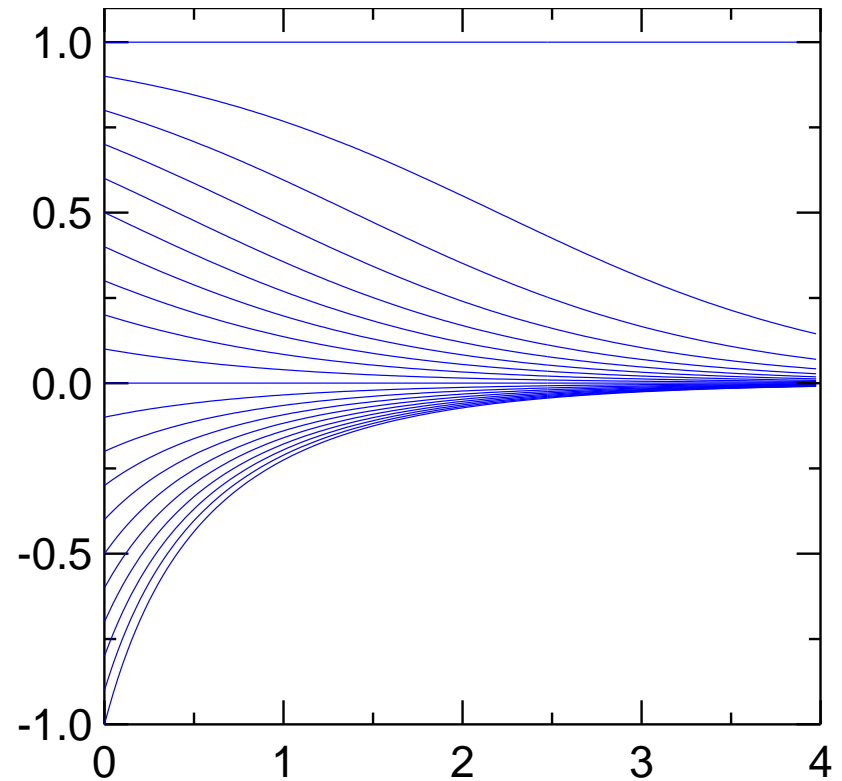
zakres zbieżności: w praktyce Δt znacznie większy
niż w iteracji funkcjonalnej
ale: niedostępne proste oszacowane przedziału zbieżności
w praktyce iteracja Newtona – szybsza i szerzej zbieżna niż iteracja funkcjonalna

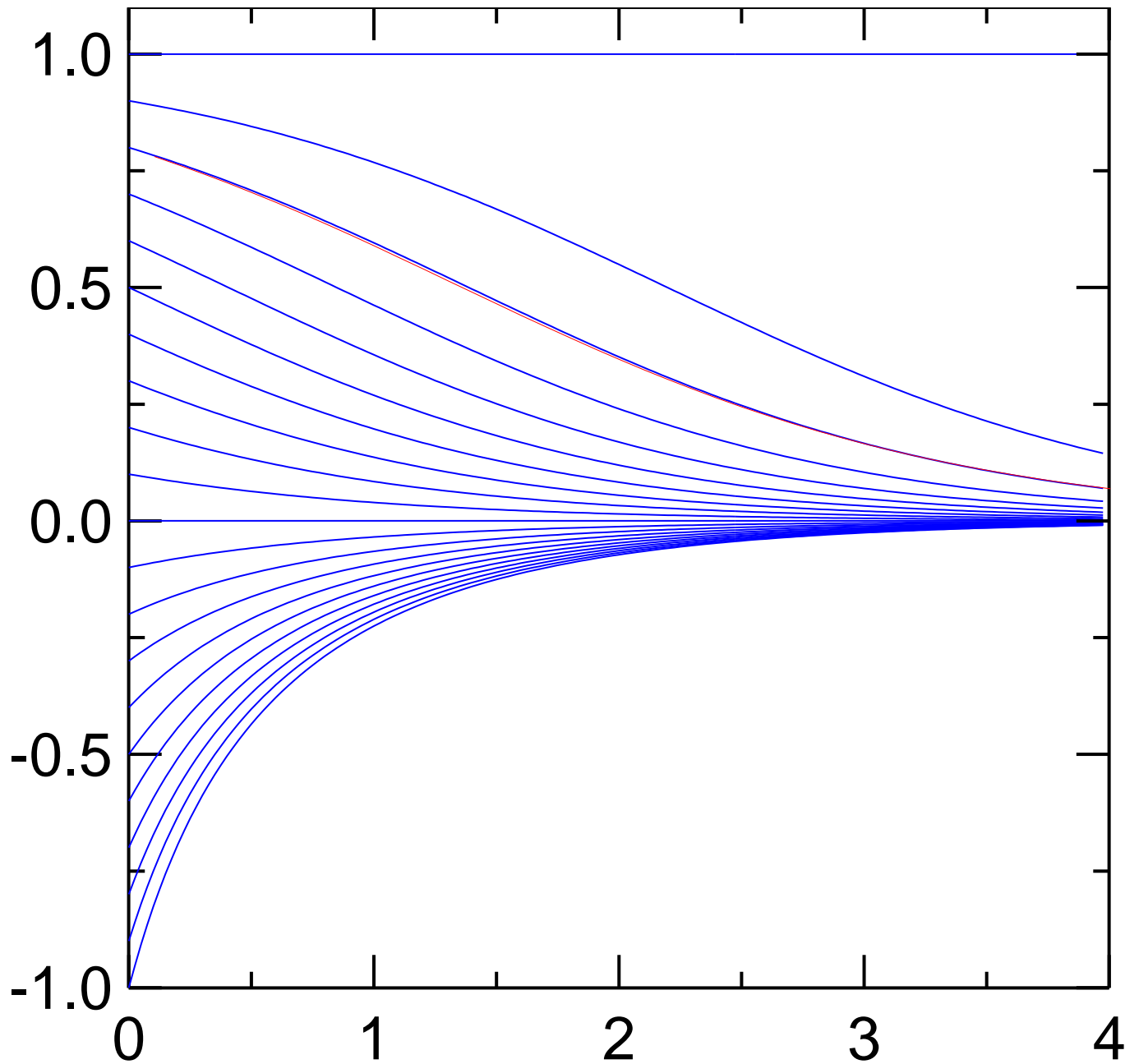


problem początkowy: dla równania:
 $u' = u(u-1)$

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t u_n^{\mu-1} (u_n^{\mu-1} - 1)}{1 - \Delta t (2u_n^{\mu-1} - 1)}$$






czerwone
niejawny
Euler z krokiem
 $\Delta t=0.1$ $u(0)=0.8$

iteracja dla
 $u(\Delta t)$ ze
startem w
 $u(0)$:

0.80000
0.78297
0.78300
0.78300

niejawny schemat Eulera z metodą Newtona-Raphsona

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$


gdy przepis funkcyjny nieznan (np. programujemy metodę dla dowolnego f)
można szacować z ilorazu różnicowego (poniżej centralny = dokładnie różniczuje parabolę)

$$f'_u(t_n, u_n) = \frac{f(t_n, u_n + du) - f(t_n, u_n - du)}{2du}$$

cena zastąpienia dokładnej pochodnej ilorazem różnicowym?
przy osiągniętej zbieżności - nie zmieni rozwiązania! może tylko spowolnić iterację!
dla naszego przykładu $u' = u(u-1)$ centralny iloraz różnicowy zadziała dokładnie dla dowolnego du

żeby przykład był ciekawszy: policzmy pochodną z wstecznego ilorazu różnicowego

$$f'_u(t_n, u_n) = \frac{f(t_n, u_n) - f(t_n, u_n - du)}{du} = 2u - 1 - du$$

$u(0)=0.8$, pierwszy krok $t=\Delta t$:

$$u'=f(u)=u(u-1)$$

metoda Newtona dla pochodnej f liczonej numerycznie w każdej iteracji:

dokładna pochodna	iloraz wsteczny $du=u/10$	iloraz wsteczny $du=u/2$	przybliżenie w liczeniu pochodnej nie zmienia wyniku do którego iteracja zbiega bo:
0.80000	0.80000	0.80000	
0.78297	0.78312	0.78367	
0.78300	0.78301	0.78303	
0.78300	0.78300	0.78301	
		0.78300	

$$x_{n+1}=x_n-F(x_n)/F'(x_n)$$

niewielko spowalnia iterację

numeryczne liczenie pochodnych w każdej iteracji może być kosztowne

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$

w praktyce
można np. wstawić tutaj u_{n-1}

można również używać oszacowania pochodnej w wielu kolejnych iteracjach
odnawiać pochodną gdy iteracja zwalnia

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_{n-1}, u_{n-1})}$$

zamiast

$$u_n^\mu = u_n^{\mu-1} - \frac{u_n^{\mu-1} - u_{n-1} - \Delta t f(t_n, u_n^{\mu-1})}{1 - \Delta t f'_u(t_n, u_n^{\mu-1})}$$

**iteracja Newtona z pochodną
liczoną w poprzednim kroku
(nieiterowaną)**

dla naszego przykładu: $f(u)=u(u-1)$ z $dt=0.1$

iterowana pochodna	pochodna brana z punktu t_{n-1}, u_{n-1}
$f'_u(t_n, u_n^{\mu-1})$	
0.80000	0.80000
0.78297	0.78297
0.78300	0.78300
0.78300	0.78300

bez różnicy!

dt=0.5		dt=1	
iterowana	z poprzedniego kroku	iterowana	stara
0.8	0.8	0.8	0.8
0.6857	0.6857	0.4	0.4
0.6937	0.6950	0.5333	0.8
0.6937	0.6935	0.5523	0.4
	0.6938	0.5527	0.8
	0.6937	0.5527	

wolniej

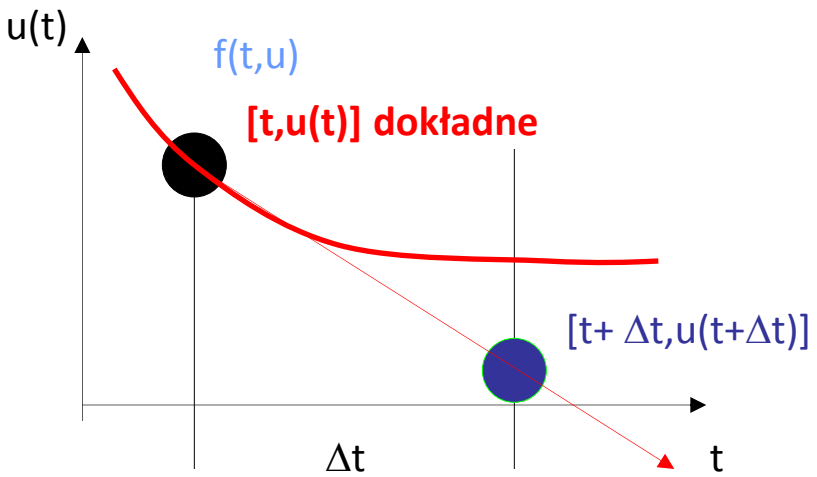
brak zbieżności ☹️

w mianowniku: $1-dt(2u-1)$
stara: 0.94 , doiterowana 0.9434

stara: 0.7
doiterowana: 0.806

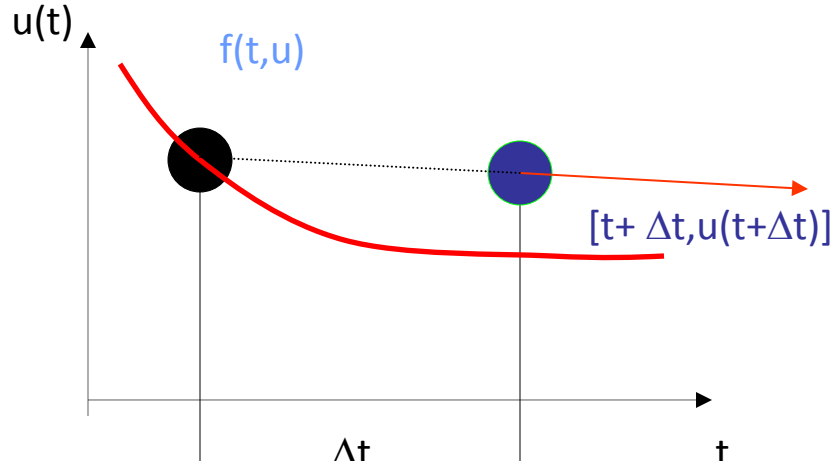
stara 0.4
doiterowana 0.89

jawny schemat Eulera [globalny błąd $O(\Delta t)$]

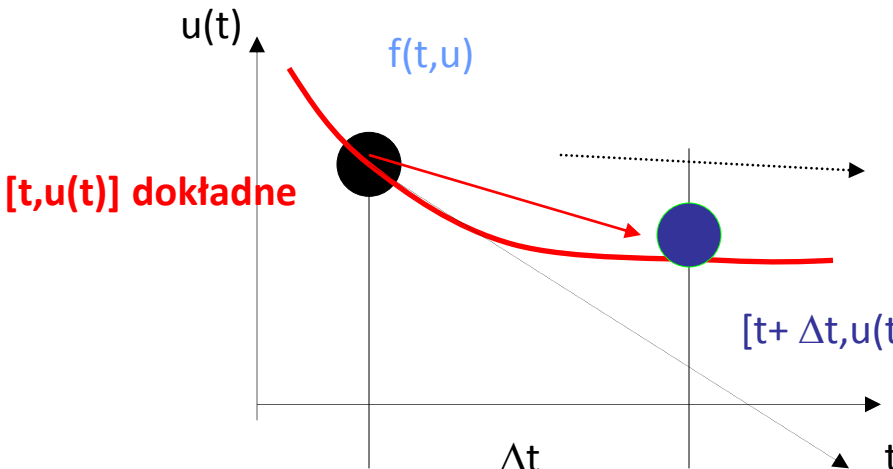


$$u_{n+1} = u_n + \Delta t f(t_n, u_n) + \frac{dt^2}{2} u''(\xi_n)$$

niejawny schemat Eulera [globalny błąd $O(\Delta t)$]



$$u_{n+1} = u_n + \Delta t f(t_{n+1}, u_{n+1}) - \frac{dt^2}{2} u''(\xi_n)$$



przesunięcie wyliczane na podstawie średniej arytmetycznej z chwil t i $t+\Delta t$

[wzór trapezów]

$$u_{n+1} = u_n + \frac{\Delta t}{2} [f(t_n, u_n) + f(t_{n+1}, u_{n+1})] + \dots$$

dokładność wzoru trapezów a jawnego schematu Eulera:

Równanie:

$$\frac{du}{dt} = f(t, u) = tu$$

Warunek początkowy: $u_1 = u(t_1=0) = 1$

Rozwiązanie: $u(t) = \exp(t^2/2)$

Punkt $t_2=0.5$ $u_2 = ?$ [dokładnie: 1.1331]

Euler jawny jeden krok: $u_2 = u_1 + \Delta t t_1 u_1 = u_1 = 1$

wzór trapezów

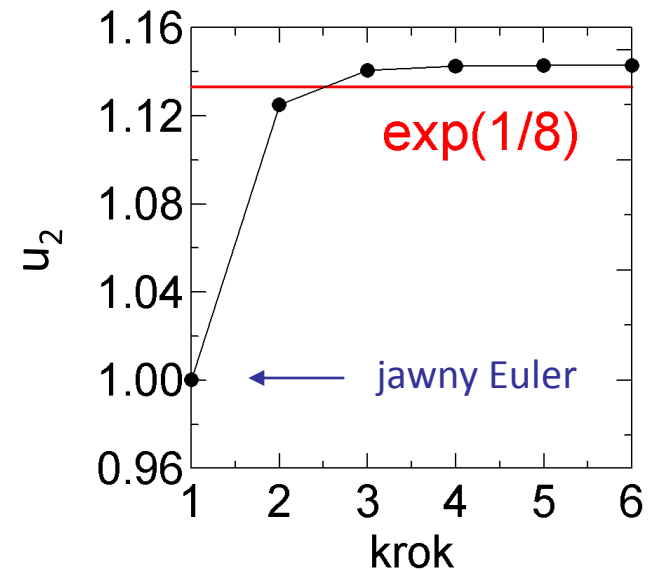
$$u_2 = u_1 + (t_1 u_1 + t_2 u_2) \Delta t / 2 = u_1 + t_2 u_2 \Delta t / 2$$

$$u_2 := u_1 + u_2 / 8$$

iteracja funkcjonalna

wynik 8/7

TR wygląda na bardziej dokładny od E:



Oszacować błąd lokalny wzoru trapezów

$$\frac{du}{dt} = f(t, u)$$

1. rozw. Taylora wstecz


$$u(t) = u(t + \Delta t) - \Delta t \frac{du(t + \Delta t)}{dt} + \frac{\Delta t^2}{2} \frac{d^2 u(t + \Delta t)}{dt^2} + O(\Delta t^3)$$

2. dla dowolnej funkcji ciągłej $f(t) = f(t + \Delta t) + O(\Delta t)$ 

$$\longrightarrow \frac{d^2 u(t + \Delta t)}{dt^2} = \frac{d^2 u(t)}{dt^2} + O(\Delta t) \quad (\text{wstawimy, rząd błędu pozostanie trzeci})$$

$$u(t) = u(t + \Delta t) - \Delta t \frac{du(t + \Delta t)}{dt} + \frac{\Delta t^2}{2} \frac{d^2 u(t)}{dt^2} + O(\Delta t^3)$$

3. Rozwiązać na $u(t + \Delta t)$

$$u(t + \Delta t) = u(t) + \Delta t \frac{du(t + \Delta t)}{dt} - \frac{\Delta t^2}{2} \frac{d^2 u(t)}{dt^2} + O(\Delta t^3)$$


Euler miał rząd obcięcia Δt^2 pozbyć się go.

3. Rozwiązać na $u(t+\Delta t)$ [przepisane]

$$u(t + \Delta t) = u(t) + \Delta t \frac{du(t + \Delta t)}{dt} - \frac{\Delta t^2}{2} \frac{d^2 u(t)}{dt^2} + O(\Delta t^3)$$

4. Uśrednić z rozwinięciem Taylora do przodu

$$u(t + \Delta t) = u(t) + \Delta t \frac{du(t)}{dt} + \frac{\Delta t^2}{2} \frac{d^2 u(t)}{dt^2} + O(\Delta t^3)$$

5. Wynik

$$u(t + \Delta t) = u(t) + \frac{\Delta t}{2} \left(\frac{du(t)}{dt} + \frac{du(t + \Delta t)}{dt} \right) + O(\Delta t^3)$$

6. Korzystamy z równania

$$\frac{du}{dt} = f(t, u)$$

$$u(t + \Delta t) = u(t) + \frac{\Delta t}{2} (f(t, u(t)) + f(t + \Delta t, u(t + \Delta t))) + O(\Delta t^3)$$

jawny i niejawny Euler
wzór trapezów

- lokalny błąd rzędu drugiego (rzęd dokładności 1)
- lokalny błąd rzędu trzeciego (rzęd dokładności 2)

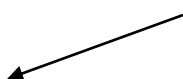
stabilność bezwzględną wzoru trapezów

problem modelowy:

$$\frac{du}{dt} = \lambda u \quad \text{WP: } u(t=0)=1.$$
$$\text{rozwiązanie } u = \exp(\lambda t)$$

$$u_n = u_{n-1} + \frac{\Delta t}{2} \lambda (u_{n-1} + u_n)$$

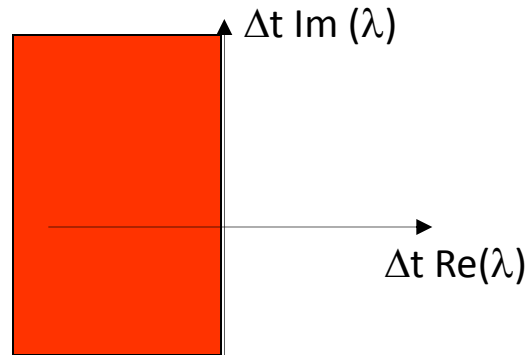
$$u_n = \frac{1 + \frac{\Delta t}{2} \lambda}{1 - \frac{\Delta t}{2} \lambda} u_{n-1}$$

$$u_n = \left(\frac{1 + \frac{\Delta t}{2} \lambda}{1 - \frac{\Delta t}{2} \lambda} \right)^n u_0$$
$$\left| \frac{1 + \lambda \frac{\Delta t}{2}}{1 - \lambda \frac{\Delta t}{2}} \right| \leq 1$$


$$|2 + z| \leq |2 - z|$$

zbiór punktów na p. Gaussa, które są nie dalej od $(-2,0)$ niż od $(2,0)$

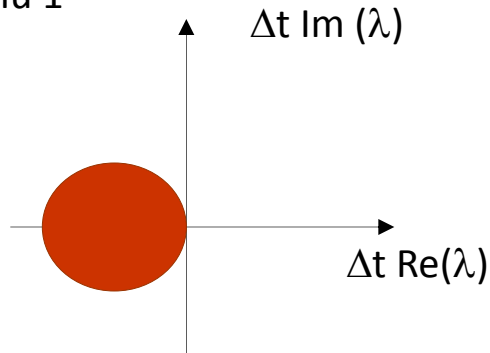
region bzw. stabilności wzoru trapezów



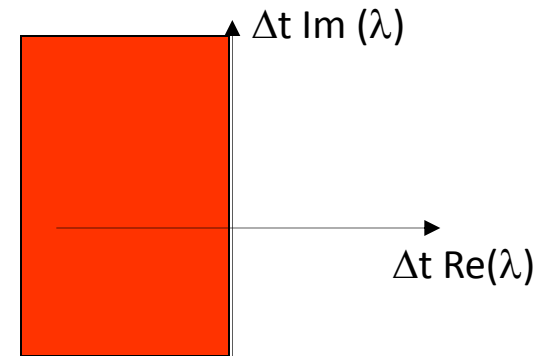
**Wniosek: dla $\lambda < 0$ wzór trapezów bezwzględnie stabilny dla dowolnego kroku czasowego !
A-stabilny**

druga bariera Dahlquist: maksymalny rząd dokładności metody A-stabilnej = 2
schemat trapezów
jest najdokładniejszą metodą A-stabilną spośród liniowych metod
wielokrokowych Implementowana np. w SPICE.

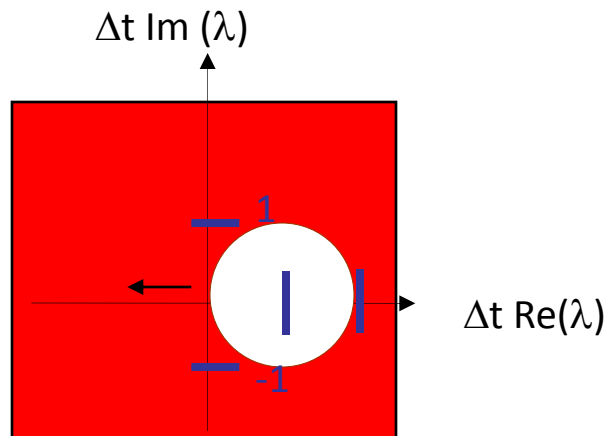
region bzw. stabilności Eulera:
koło o promieniu 1
i środku (-1,0)



region bzw. stabilności wzoru trapezów



niejawna metoda Eulera:
region bezwzględnej stabilności



między metodami można przechodzić w sposób ciągły

$$u_n = u_{n-1} + \Delta t [(1 - \theta)f(t_{n-1}, u_{n-1}) + \theta f(t_n, u_n)]$$

$\theta=0, 1, 1/2$ – Euler jawny, niejawny i wzór trapezów odpowiednio
w wykładzie na temat niejawnych formuł RK zobaczymy, że
dokładność rzędu 2 uzyskana tylko dla $\theta=1/2$
region stabilności ?

iteracja funkcjonalna a wzór trapezów

problem początkowy:

$$u' = -100u, \quad u(0) = 1$$

z rozwiązaniem dokładnym $u(t) = \exp(-100t)$

$$u_n^\mu = u_{n-1} + \frac{\Delta t}{2} \left(\underline{f(t_{n-1}, u_{n-1})} + f(t_n, u_n^{\mu-1}) \right)$$

$\Delta t = 0.01$ = graniczny dla zbieżności IF dla niejawnego Eulera

1,0, 0.5, 0.25, 0.375, 0.3125, 0.34375, 0.328125, 0.33593, 0.33203, 0.333984, ... ,0.333333

wzór trapezów = używa prawej strony z poprzedniego kroku czasowego z wagą 0.5 – co nieco stabilizuje iterację.

niestety iteracja funkcjonalna dla $\Delta t = 0.02$ już przestaje być zbieżna (+1,-1,+1,-1,itd..) wzór trapezów zwiększa zakres zbieżności iteracji dwukrotnie (wyróżnienie podkreślenie stabilizuje iterację) ale to wciąż mało

metoda Newtona-Raphsona pozostaje

$$\frac{du}{dt} = f$$

poznane metody:

$$1) u_{n+1} = u_n + \Delta t f(t_n, u_n) + O(\Delta t^2)$$

$$2) u_{n+1} = u_n + \Delta t f(t_{n+1}, u_{n+1}) + O(\Delta t^2)$$

$$3) u_{n+1} = u_n + \frac{\Delta t}{2} (f(t_{n+1}, u_{n+1}) + f(t_n, u_n)) + O(\Delta t^3)$$

Poznane metody: jednokrokowe (1-3), jawna (1) i niejawne (2-3), pierwszego (1-2) i drugiego (3) rzędu dokładności
Metody (2-3) A stabilne, metoda (2) nadstabilna

jawne metody różnicowe wysokiej dokładności ??

jawne metody jednokrokowe wyższego rzędu dokładności niż jawny Euler

$$u' = f(t, u), u(0) = u_0$$

rozwinięcie Taylora ponownie:

$$u(t_n) = u(t_{n-1}) + \Delta t u'(t_{n-1}) + \frac{\Delta t^2}{2} u''(t_{n-1}) + \dots + \frac{\Delta t^k}{k!} u^{(k)}(t_{n-1}) + O(\Delta t^{k+1})$$

liczymy pochodne:

$$u'(t_{n-1}) = f(t_{n-1}, u(t_{n-1})) \quad \text{z RR.}$$

RR różniczkujemy po czasie

$$u''(t_{n-1}) = [f'_t + f'_u u']_{(t_{n-1}, u_{n-1})}$$

czyli

$$u''(t_{n-1}) = [f'_t + f'_u f]_{(t_{n-1}, u_{n-1})}$$

podobnie

$$u'''(t_{n-1}) = [f''_{tt} + 2f''_{tu} f + f''_{uu} f^2 + (f'_u)^2 f + f'_u f'_t + f''_{tu} f]_{(t_{n-1}, u_{n-1})}$$

Zależnie od tego gdzie się zatrzymamy uzyskamy błąd lokalny
żądanego rzędu

$$u(t_n) = u(t_{n-1}) + \Delta t u'(t_{n-1}) + \frac{\Delta t^2}{2} u''(t_{n-1}) + \dots + \frac{\Delta t^k}{k!} u^{(k)}(t_{n-1}) + O(\Delta t^{k+1})$$

Zależnie od tego gdzie się zatrzymamy uzyskamy błąd lokalny zadanego rzędu np.

$$u(t_n) = u(t_{n-1}) + \Delta t f(t_{n-1}, u_{n-1}) + \frac{\Delta t^2}{2} [f'_t + f'_u f]_{(t_{n-1}, u_{n-1})} + \frac{\Delta t^3}{6} u'''(\xi_n)$$

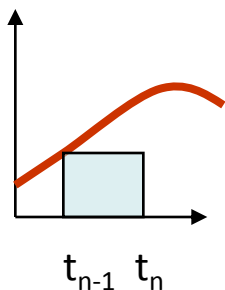
pomysł: mało przydatny w praktyce ze względu na konieczność analitycznego wyliczenia pochodnych cząstkowych f .

Dla metod ogólnych: nie powinniśmy liczyć, że f jest dane wzorem

podjęcie alternatywne: inspirowane całkowaniem

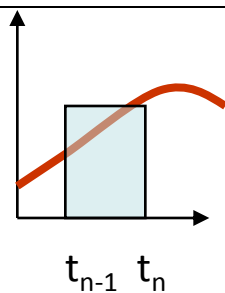
prawa strona = funkcja tylko t $u' = f(t), u(0) = u_0$

z rozwiązaniem: $u(t) = u_0 + \int_0^t f(\tau) d\tau$



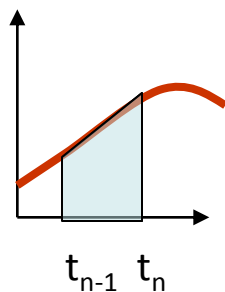
jeśli zastąpimy całkę kwadraturą prostokątów z wywołaniem funkcji w lewym końcu przedziału

$u(t_n) = u(t_{n-1}) + \Delta t f(t_{n-1}) + O(\Delta t^2)$ - rozpoznajemy jawny schemat Eulera



kwadratura prostokątów z wywołaniem funkcji w prawym końcu przedziału

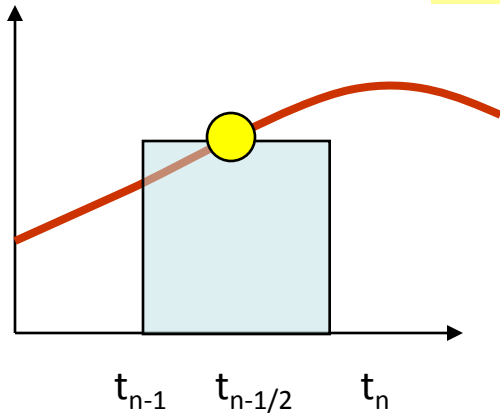
$u(t_n) = u(t_{n-1}) + \Delta t f(t_n) + O(\Delta t^2)$ - rozpoznajemy niejawny schemat Eulera



kwadratura trapezów

$u(t_n) = u(t_{n-1}) + \Delta t f(t_n)/2 + \Delta t f(t_{n-1})/2 + O(\Delta t^3)$
- rozpoznajemy niejawny schemat trapezów

reguła punktu środkowego



wzór prostokątów z wywołaniem
funkcji w środku przedziału
(dokładny dla funkcji liniowej, znoszenie błędów)

$$u_n = u_{n-1} + \Delta t f(t_{n-1} + \Delta t/2) + O(\Delta t^3)$$

uogólniony wzór na równanie różniczkowe $u' = f(t, u)$

$$u_n = u_{n-1} + \Delta t f\left(t_{n-\frac{1}{2}}, u_{n-\frac{1}{2}}\right)$$

ale - skąd rozwiązanie w środku przedziału?

np. ze schematu Eulera:

$$u_{n-\frac{1}{2}} = u_{n-1} + \frac{\Delta t}{2} f(t_{n-1}, u_{n-1})$$

błąd lokalny Eulera $O(\Delta t^2)$, czy **reguła punktu środkowego** zachowa trzeci rząd błędu lokalnego?

sprawdźmy to rozważając bardziej ogólny schemat:

$$u_n = u_{n-1} + \Delta t (b_1 k_1 + b_2 k_2)$$

$$k_1 = f(t_{n-1}, u_{n-1}) \quad \text{obliczone na początku kroku}$$

$$k_2 = f(t_{n-1} + c\Delta t, u_{n-1} + a\Delta t k_1)$$

obliczone gdzie
w środku przedziału
(t_{n-1}, t_n)

z odpowiednio oszacowanym
rozwiązaniem u dla tego t (wzór typu Eulera)

jest to jawny dwustopniowy schemat **Rungego-Kutty**. potencjalna wyższa dokładność od jawnego Eulera kosztem dwóch wywołań f (podobnie jak we wzorze trapezów, ale RK: jawny)

b_1, b_2, a, c – parametry metody – jakie muszą być aby RK2 (2 = rząd dokładności)

reguła punktu środkowego:

należy do tej klasy z
 $b_1=0, b_2=1, c=1/2, a=1/2$

$$u_n = u_{n-1} + \Delta t f\left(t_{n-\frac{1}{2}}, u_{n-\frac{1}{2}}\right)$$

$$u_{n-\frac{1}{2}} = u_{n-1} + \frac{\Delta t}{2} f(t_{n-1}, u_{n-1})$$

$$\begin{aligned}
 u_n &= u_{n-1} + \Delta t (b_1 k_1 + b_2 k_2) \\
 (*) \quad k_1 &= f(t_{n-1}, u_{n-1}) \\
 k_2 &= f(t_{n-1} + c\Delta t, u_{n-1} + a\Delta t k_1)
 \end{aligned}$$

$$u' = f(t, u)$$

jak dobrać b_1, b_2, c, a ? – metodą brutalnej siły -
 tak aby rozwinięcie Taylora metody zgadzało się
 z rozwinięciem Taylora dokładnego równania różniczkowego
 do wyrazów tak wysokiego rzędu jak to tylko możliwe

przypominamy: rozwinięcie Taylora dla funkcji dwóch zmiennych

$$\begin{aligned}
 g(x + \Delta x, y + \Delta y) &= g(x, y) \\
 &+ \left(\Delta x \frac{\partial g(x, y)}{\partial x} + \Delta y \frac{\partial g(x, y)}{\partial y} \right) \\
 &+ \frac{1}{2!} \left(\Delta x^2 \frac{\partial^2 g(x, y)}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 g(x, y)}{\partial x \partial y} + \Delta y^2 \frac{\partial^2 g(x, y)}{\partial y^2} \right) + \dots
 \end{aligned}$$

wstawiamy rozwiązanie dokładne $u(t_n)$, $u(t_{n-1})$ do (*) i rozwijamy względem t_{n-1} , u_{n-1}

$$u(t_n) = u(t_{n-1}) + \Delta t (b_1 k_1 + b_2 k_2)$$

$$k_1 = f(t_{n-1}, u(t_{n-1}))$$

$$k_2 = f(t_{n-1} + c\Delta t, u(t_{n-1}) + a\Delta t k_1) \leftarrow \text{to trzeba rozwinąć}$$

$$k_2 = f + c\Delta t f'_t + a\Delta t f f'_u + \frac{1}{2} (c^2 \Delta t^2 f''_{tt} + a^2 \Delta t^2 f^2 f''_{uu} + 2ac\Delta t^2 f f'_{tu}) + O(\Delta t^3)$$

wstawmy k_2 do rozwinięcia.

Zachowajmy człony do Δt^2 :

(wszystkie pochodne liczone
w t_{n-1}, u_{n-1})

$$u(t_n) = u(t_{n-1}) + \Delta t [b_1 f + b_2 (f + c\Delta t f'_t + a\Delta t f f'_u)] + O(\Delta t^3)$$

rozwinięcie Taylora rozwiązania dokładnego uzyskaliśmy kilka slajdów wcześniej

$$u(t_n) = u(t_{n-1}) + \Delta t f(t_{n-1}, u_{n-1}) + \frac{\Delta t^2}{2} [f'_t + f'_u f]_{(t_{n-1}, u_{n-1})} + \frac{\Delta t^3}{6} u'''(\xi_n)$$

czyli:

rzęd Δt : $b_1 + b_2 = 1$, rzęd Δt^2 : $b_2 c = b_2 a = 1/2$

czyli reguła punktu środkowego: $b_1 = 0, b_2 = 1, c = 1/2, a = 1/2$ ma błąd lokalny rzędu $O(\Delta t^3)$
mamy metodę równie dokładną co wzór trapezów – ale jawną (co ma swoje zalety i wady)

Wyższy rząd błędu do uzyskania tylko w metodach o większej niż 2 liczbie stopni

cztery parametry i trzy równania

$$b_1 + b_2 = 1$$

$$b_2 c = b_2 a = 1/2$$

- pozostaje swoboda w wyborze parametrów

reguła punktu środkowego RK2

$$b_1 = 0, b_2 = 1, c = 1/2, a = 1/2$$

dwa zastosowania jawnego schematu Eulera

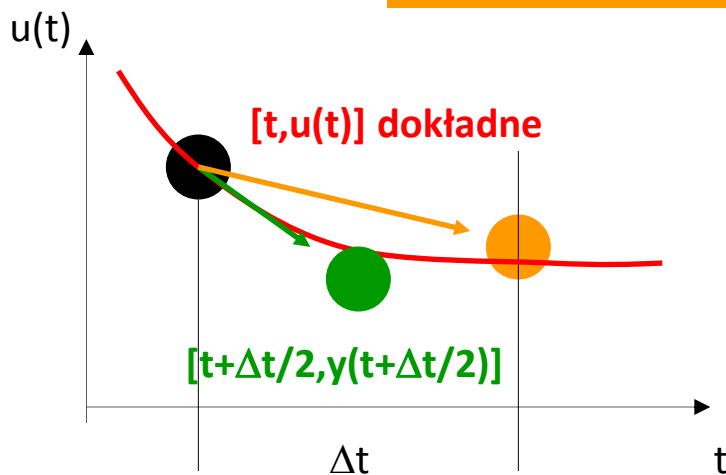
$$u_{n-\frac{1}{2}} = u_{n-1} + \frac{\Delta t}{2} f(t_{n-1}, u_{n-1})$$

oszacowanie wstępne w punkcie pośrednim
(błąd lokalny rzędu drugiego)

$$u_n = u_{n-1} + \Delta t f(t_{n-\frac{1}{2}}, u_{n-\frac{1}{2}})$$

oszacowanie docelowe
(błąd lokalny oszacowania: rzędu trzeciego)

$$u_n = u_{n-1} + \Delta t f\left(t_{n-1} + \frac{\Delta t}{2}, u_{n-1} + \frac{\Delta t}{2} f(t_{n-1}, u_{n-1})\right)$$



- 1) Szacujemy metodą Eulera punkt środkowy $[t + \Delta t/2, u(t + \Delta t/2)]$ korzystając z $f(t, u)$ w lewym końcu przedziału
- 2) Wykorzystujemy wartość f w tym punkcie do wyliczenia zmiany y na całym przedziale Δt

RK 2 : $b_1+b_2=1$,
 $b_2c=b_2a=1/2$

inny wybór niż p.ś.:
 $b_1=b_2=1/2$, wtedy musi $a=c=1$

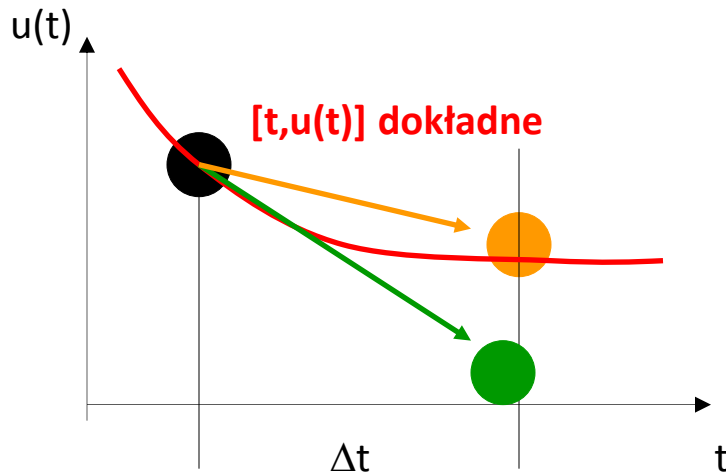
$$u_n = u_{n-1} + \Delta t (b_1 k_1 + b_2 k_2)$$

$$k_1 = f(t_{n-1}, u_{n-1})$$

$$k_2 = f(t_{n-1} + c\Delta t, u_{n-1} + a\Delta t k_1)$$

metoda RK podobna do wzoru trapezów
(ale jawna)

$$u_n = u_{n-1} + \frac{\Delta t}{2} f(t_{n-1}, u_{n-1}) + \frac{\Delta t}{2} f(t_{n-1} + \Delta t, u_{n-1} + \Delta t f(t_{n-1}, u_{n-1}))$$



metoda RK2 trapezów

- 1) Szacujemy metodą Eulera punkt końcowy $[t+\Delta t, u(t+\Delta t)]$ korzystając z $f(t, u)$ w lewym końcu przedziału
- 2) krok z t do $t+\Delta t$ wykonujemy biorąc średnią arytmetyczną z f na początku i końcu

dla błędu lokalnego $O(\Delta t^3)$ potrzeba aby,
rzęd Δt : $b_1 + b_2 = 1$, rzęd Δt^2 : $b_2 c = b_2 a = 1/2$

punkt środkowy $b_2 = 1$, $b_1 = 0$ [$b_1 + b_2$]=1

czy ma sens $b_1 = 1$, $b_2 = 0$?

$$u(t_n) = u(t_{n-1}) + \Delta t (b_1 k_1 + b_2 k_2)$$

$$k_1 = f(t_{n-1}, u(t_{n-1}))$$

są to metody jednokrokowe, o przepisie ogólnym:

$$u_n = u_{n-1} + \Delta t \phi(t_{n-1}, u_{n-1}, \Delta t)$$

metoda RK w s-odstępach (*stage*) (unikamy słowa „krok”)

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

z

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} k_j)$$

wzory przedstawiane w formie tabel Butchera

c **A**
b

c_1	a_{11}	a_{12}	\cdots	a_{1s}
c_2	a_{21}	a_{22}	\cdots	a_{2s}
\vdots	\vdots	\vdots	\ddots	\vdots
c_s	a_{s1}	a_{s2}	\cdots	a_{ss}
	b_1	b_2	\cdots	b_s

Metody Rungego-Kutty, forma ogólna

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} k_j)$$

czasem zapisywane w postaci:

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i f(t_{n-1} + c_i \Delta t, U_i)$$

$$U_i = u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} f(t_{n-1} + c_j \Delta t, U_j)$$

tutaj U_i – przybliżone rozwiązanie w chwili $t_{n-1} + c_i \Delta t$
zazwyczaj niższej dokładności niż rozwiązanie końcowe

jawne metody Rungego-Kutty

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} k_j)$$

jawne: $a_{ij}=0$ dla $j \geq i$

c_1	a_{11}	a_{12}	\cdots	a_{1s}
c_2	a_{21}	a_{22}	\cdots	a_{2s}
\vdots	\vdots	\vdots	\ddots	\vdots
c_s	a_{s1}	a_{s2}	\cdots	a_{ss}
	b_1	b_2	\cdots	b_s

obcięte sumowanie:

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^{i-1} a_{ij} k_j)$$

odstona i -ta wyliczana na podstawie tylko wcześniejszych odstępów

historycznie wszystkie RK były jawne, uogólnienie okazało się przydatne dla problemów sztywnych

Wyprowadzanie formuł RK (a,b,c)

- 1) Rozwijamy rozwiązanie dokładne w szereg Taylora względem t_{n-1}
- 2) Podstawiamy rozwiązanie dokładne do ogólnej formy RK i rozwijamy względem t_{n-1}
- 3) Wartości parametrów a,b,c uzyskujemy z porównania. zazwyczaj w sposób niejednoznaczny

najbardziej popularne: jawne formuły 4-etapowe RK4:
o 4-tym stopniu zbieżności (4-tym rzędzie dokładności)
i 5-tym rzędzie błędu lokalnego

ogólna tabela Butchera:
dla jawnych RK4

0	0	0	0	0
c_2	a_{21}	0	0	0
c_3	a_{31}	a_{32}	0	0
c_4	a_{41}	a_{42}	a_{43}	0
	b_1	b_2	b_3	b_4

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^{i-1} a_{ij} k_j)$$

$$c_1 = 0$$

(dla każdej jawnej RK, zaczynamy – k_1
od wyliczenia prawej strony w kroku początkowym)

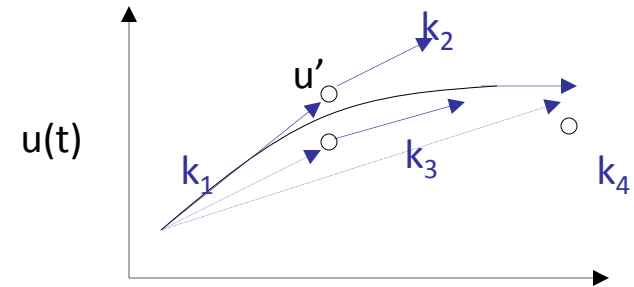
klasyczna formuła RK4:
$$u_n = u_{n-1} + \frac{\Delta t}{6} (k_1 + 2k_2 + 2k_3 + k_4)$$

$$k_1 = f(t_{n-1}, u_{n-1})$$

$$k_2 = f\left(t_{n-1} + \frac{\Delta t}{2}, u_{n-1} + \frac{\Delta t k_1}{2}\right)$$

$$k_3 = f\left(t_{n-1} + \frac{\Delta t}{2}, u_{n-1} + \frac{\Delta t k_2}{2}\right)$$

$$k_4 = f(t_{n-1} + \Delta t, u_{n-1} + \Delta t k_3)$$



4 wywołania f na krok,
błąd lokalny $O(\Delta t^5)$

gdy f tylko funkcja czasu RK4
redukuje się do formuły Simpsona :

$$u_n = u_{n-1} + \frac{\Delta t}{6} (f(t_{n-1}) + 4f(t_{n-1}/2) + f(t_n))$$

$$\frac{du^1}{dt} = f^1(t, u^1, u^2)$$

$$\frac{du^2}{dt} = f^2(t, u^1, u^2)$$

2 zmienne zależne u^1, u^2 ,
2 prawe strony f^1, f^2

2 równania, s-odstęp
($i=1,2,\dots,s$)

$$U_i^1 = u_{n-1}^1 + \Delta t \sum_{j=1}^{i-1} a_{ij} f^1(t_{n-1} + c_j \Delta t, U_j^1, U_j^2)$$

$$U_i^2 = u_{n-1}^2 + \Delta t \sum_{j=1}^{i-1} a_{ij} f^2(t_{n-1} + c_j \Delta t, U_j^1, U_j^2)$$

$$u_n^1 = u_{n-1}^1 + \Delta t \sum_{i=1}^s b_i f^1(t_{n-1} + c_i \Delta t, U_i^1, U_i^2)$$

$$u_n^2 = u_{n-1}^2 + \Delta t \sum_{i=1}^s b_i f^2(t_{n-1} + c_i \Delta t, U_i^1, U_i^2)$$

zapis wektorowy

$\mathbf{u}_{n-1}, \mathbf{u}_n, \mathbf{f}, \mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_N$ są wektorami o 2 składowych

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(t, \mathbf{u})$$

$$\mathbf{U}_i = \mathbf{u}_{n-1} + \Delta t \sum_{j=1}^{i-1} a_{ij} \mathbf{f}(t_{n-1} + c_j \Delta t, \mathbf{U}_j)$$

$$\mathbf{u}_n = \mathbf{u}_{n-1} + \Delta t \sum_{i=1}^s b_i \mathbf{f}(t_{n-1} + c_i \Delta t, \mathbf{U}_i)$$

Tabela Butchera dla klasycznej jawnej RK4

$$u_n = u_{n-1} + \frac{\Delta t}{6} (k_1 + 2k_2 + 2k_3 + k_4)$$

$$k_1 = f(t_{n-1}, u_{n-1})$$

$$k_2 = f\left(t_{n-1} + \frac{\Delta t}{2}, u_{n-1} + \frac{\Delta t k_1}{2}\right)$$

$$k_3 = f\left(t_{n-1} + \frac{\Delta t}{2}, u_{n-1} + \frac{\Delta t k_2}{2}\right)$$

$$k_4 = f(t_{n-1} + \Delta t, u_{n-1} + \Delta t k_3)$$

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

$$k_i = f\left(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^{i-1} a_{ij} k_j\right)$$

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

0	0	0	0	0
c_2	a_{21}	0	0	0
c_3	a_{31}	a_{32}	0	0
c_4	a_{41}	a_{42}	a_{43}	0
	b_1	b_2	b_3	b_4

Dlaczego RK4 najbardziej popularna:

Liczba kroków a rząd zbieżności jawnych metod RK:

rząd	1	2	3	4	5	6	7	8
minimalna liczba odstępów	1	2	3	4	6	7	9	11

RK4 – wyjątkowo opłacalna

RK1 – metoda RK w jednej odstępnie

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} k_j)$$

$b_1 + b_2 = 1$, przy $b_1 = 1$, $b_2 = 0$
dostaniemy jawnego Eulera
warunek $a_{12} b_2 = c_2 b_2 = 1/2$
nie będzie spełniony

jawny schemat Eulera to jawna metoda RK1

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^s b_i k_i$$

jawny Euler

$$k_i = f(t_{n-1} + c_i \Delta t, u_{n-1} + \Delta t \sum_{j=1}^s a_{ij} k_j)$$

0	0
	1

tabela Butchera

RK2 trapezów

$$\mathbf{b_1=b_2=1/2, a=c=1}$$

0	0	0
1	1	0
	$\frac{1}{2}$	$\frac{1}{2}$

$$u_n = u_{n-1} + \Delta t (b_1 k_1 + b_2 k_2)$$

$$k_1 = f(t_{n-1}, u_{n-1})$$

$$k_2 = f(t_{n-1} + c\Delta t, u_{n-1} + a\Delta t k_1)$$

RK2 punktu środkowego

$$u_{n+1} = u_n + \Delta t f\left(t_n + \frac{\Delta t}{2}, u_n + \frac{\Delta t}{2} f(t_n, u_n)\right)$$

0	0	0
$1/2$	$1/2$	0
	0	1