Metody Obliczeniowe w Nauce i Technice

Metody całkowania równań różniczkowych zwyczajnych

Marian Bubak, Katarzyna Rycerz

Department of Computer Science
AGH University of Science and Technology
Krakow, Poland
kzajac@agh.edu.pl
dice.cyfronet.pl

Contributors
Paweł Urban
Jakub Ptak
Kamil Doległo

Outline

Model

$$\boxed{\begin{cases} \frac{du}{dt} + f(u,t) = 0 & \mathrm{u} = \mathrm{u}(\mathrm{t}) \\ \text{warunki początkowe:} & u(t^0) = u^0 \end{cases}}$$

- $\frac{df}{du} > 0$ - równanie typu "rozpadu"
- $\frac{df}{du} < 0$ - równanie typu "wzrostu"

- u zespolone - równanie typu "oscylacyjnego"
  gdy - u, f zespolone $\Rightarrow$ para równań
- schematy całkowania po czasie $\Rightarrow$ przejście do równań różnicowych

Przenosi się to na:

- układy równań różniczkowych zwyczajnych
- układy równań różniczkowych cząstkowych

Model c.d.

Równanie modelowe można scałkować na siatce czasowej:

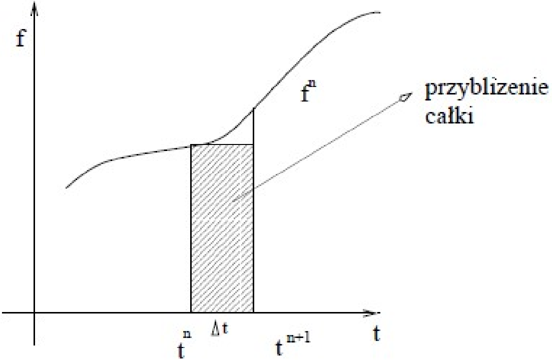
$$\Delta t = t^{n+1} - t^n$$

$$u^{n+1} = u^n - \underbrace{\int_{t^n}^{t^{n+1}} f(u,t)dt}_{}$$

różne aproksymacje $\Rightarrow$ metody całkowania

Metoda Eulera pierwszego rzędu

$\underline{f(u,t) = \,?}$ dla $t \in [t^n, t^{n+1}] \approx f(u^n, t^n)$



Cechy metody

Algorytm:

$$u^{n+1} = u^n - f(u^n, t^n) \cdot \Delta t$$

Cechy:

- jawna
- 1-go rzędu - bład zmienia liniowo się ze względu na Δt: $\varepsilon = 0(\Delta t)$
- prosta
- efektywna

Stabilność

w czasie t^n mamy u^n z błędem ε^n
w czasie t^{n+1} mamy u^{n+1} z błędem ε^{n+1}

$\boxed{\varepsilon^{n+1} = g \cdot \varepsilon^n \qquad \text{g - współczynnik wzmocnienia błędu}}$

$$u^{n+1} + \varepsilon^{n+1} = u^n + \varepsilon^n - f(u^n + \varepsilon^n, t^n) \cdot \Delta t \qquad (*)$$

<u>zał.:</u> ε^n - mały $\Rightarrow$ rozwijamy w szereg Taylora, pomijamy człony nieliniowe:

$$f(u^n + \varepsilon^n, t^n) = f(u^n, t^n) + \frac{\partial f}{\partial u}\bigg|_{u^n} \cdot \varepsilon^n$$

Po podst. do (*):

$$\boxed{u^{n+1}} + \varepsilon^{n+1} = \boxed{u^n} + \varepsilon^n - (\boxed{f(u^n, t^n)} + \frac{\partial f}{\partial u}\bigg|_{u^n} \cdot \varepsilon^n) \cdot \boxed{\Delta t}$$

$$\varepsilon^{n+1} = \varepsilon^n - \frac{\partial f}{\partial u}\Big|_n \cdot \Delta t \cdot \varepsilon^n$$

współczynnik wzmocnienia: $g = \frac{\epsilon^{n+1}}{\epsilon^n} = 1 - \frac{\partial f}{\partial u}\Big|_n \cdot \Delta t$
warunek stabilności: $|g| \leqslant 1$
warunek stabilności dla $\frac{\partial f}{\partial u} > 0$

$$\frac{\partial f}{\partial u}\Big|_n \cdot \Delta t \leqslant 2 \quad \rightarrow \quad \Delta t \leqslant \frac{2}{\frac{\partial f}{\partial u}\Big|_n} \Rightarrow krok$$

(gdy $\frac{\partial f}{\partial u}\Big|_n < 0$ - metoda niestabilna)
stabilność - fundamentalna własność metody różnicowej

Przykład

$$\frac{du}{dt} + \frac{u}{\tau} = 0, \qquad u(0) = 1$$

rozpad promieniotwórczy

- rozwiązanie otrzymane analitycznie: $u = e^{-\frac{t}{\tau}}$

krok czasowy gwarantujący stabilność:

$$\Delta t \leqslant \frac{2}{\frac{\partial f}{\partial u}\Big|_n}$$

$$\left.\frac{\partial f}{\partial u}\right|_n = \frac{1}{\tau} \qquad \Delta t \leqslant 2\tau$$

Przykład 2 - równanie typu oscylacyjnego

oscylator harmoniczny: $\frac{d^2x}{dt^2} + \omega^2 x = 0$

$\Rightarrow$ układ 2 równań (dla $v = \frac{dx}{dt}$):

$$\begin{cases} \frac{dv}{dt} & + & \omega^2 x & = & 0 \\ \frac{dx}{dt} & - & v & = & 0 & | \cdot(-i\omega) \end{cases}$$

po dodaniu mamy:

$$\frac{dv}{dt} - i\omega\frac{dx}{dt} + \omega^2 x + i\omega v = 0$$

$$\frac{dv}{dt} - i\omega\frac{dx}{dt} + -i^2\omega^2 x + i\omega v = 0$$

$$\frac{dv}{dt} - i\omega\frac{dx}{dt} + i\omega(v - i\omega x) = 0$$

Przykład 2 c.d.

wprowadzenie $\underline{u = v - i\omega x}$ $\Rightarrow$ pojedyncze równanie 1-go rzędu:

$$\frac{du}{dt} + i\omega u = 0$$

<u>współczynnik wzmocnienia:</u> $(f(u) = i\omega u)$

$$g = 1 - \left.\frac{\partial f}{\partial u}\right|_n \cdot \Delta t \quad \Rightarrow \quad g = 1 - i \cdot \omega \cdot \Delta t \qquad \mathit{zespolony!}$$

$$|g|^2 = g \cdot g^* = 1 + \omega^2 \Delta t^2 \quad \Rightarrow |g| > 1$$

Metoda Eulera jest niestabilna dla równań typu oscylacyjnego.

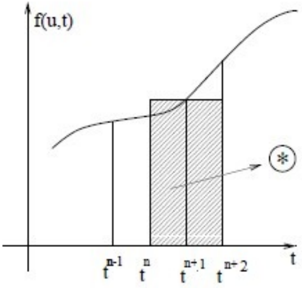
W zagadnieniach *nieliniowych* $\frac{\partial f}{\partial u}$ jest funkcją u $\Rightarrow$ należy na każdym etapie wybierać Δt spełniające warunki stabilności.

Metoda skokowa

Założenia

- całkujemy na kroku czasowym o podwójnej długości
- punkt czasowy w środku takiego kroku używamy dla obliczenia całki metodą prostokątów
- metoda wycentrowana w czasie $\longrightarrow$ dokładność drugiego rzędu $\varepsilon = 0(\Delta t^2)$

Algorytm



Algorytm:

$$
\begin{aligned}
u^{(n+1)}: &\quad = \quad u^{n-1} - f(u^n, t^n) \cdot 2 \cdot \Delta t \\
u^{(n+2)}: &\quad = \quad u^{n} \quad - \underbrace{f(u^{n+1}, t^{n+1}) \cdot 2 \cdot \Delta t}_{*}
\end{aligned}
$$

Trudności

1. znamy $u^0 = u(0)$, potrzebne $u^1 = u(\Delta t)$; od u^1 zależy całkowita dokładność
   - podprzedziały w pierwszym Δt, metoda Eulera,
   - rozwinięcie w szereg wyższych potęg (szereg Taylora),
2. zagadnienie nieliniowe $\Rightarrow$ zmienne Δt metoda przestaje być wycentrowana w czasie

Współczynnik wzmocnienia błędu

$$\varepsilon^{n+1} = \varepsilon^{n-1} - \left.\frac{\partial f}{\partial u}\right|_n \cdot 2 \cdot \Delta t \cdot \varepsilon^n \quad | : \varepsilon^{n-1}$$

$$\alpha = \left.\frac{\partial f}{\partial u}\right|_n \cdot \Delta t \qquad g = \frac{\varepsilon^{n+1}}{\varepsilon^n} \approx \frac{\varepsilon^n}{\varepsilon^{n-1}} \qquad g^2 = \frac{\varepsilon^{n+1}}{\varepsilon^n} \cdot \frac{\varepsilon^n}{\varepsilon^{n-1}}$$

$$g^2 = 1 - \alpha \cdot 2g \qquad \underline{g = -\alpha \pm \sqrt{\alpha^2 + 1}}$$

Mamy dwa pierwiastki.
jeśli α - rzeczywiste $\Rightarrow \quad |g| > 1 \quad \rightarrow$ niestabilna
Jeśli α urojone i równe $i\beta$, gdzie $\beta \leqslant 1$, rzeczywiste
$\Rightarrow g = -i\beta \pm \sqrt{1 - \beta^2} \qquad |g|^2 = g \cdot g^* = 1$
np. dla równania oscylatora: $\frac{du}{dt} + i\omega u = 0 \rightarrow \quad \beta = \omega \Delta t$
warunek $\beta \leqslant 1 \quad \rightarrow \quad \Delta t \leqslant \frac{1}{\omega}$

Ulepszona metoda Eulera - metoda jawna

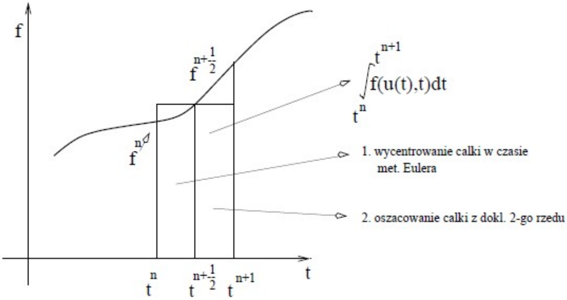
1. wyliczenie zmiennej u dla pośredniego $t^{n+\frac{1}{2}}$ metodą Eulera:

$$u^{n+\frac{1}{2}} = u^n - f(u^n, t^n)\frac{\Delta t}{2} \qquad \text{(wzór pomocniczy)}$$

2. Wyliczenie właściwej wartości u_{n+1} na podstawie $u_{n+\frac{1}{2}}$ z poprzedniego kroku:

$$u^{n+1} = u^n - f(u^{n+\frac{1}{2}}, t^{n+\frac{1}{2}})\Delta t \qquad \text{(wzór główny)}$$

Metoda jawna - tzn w każdym kroku znamy wszystkie wartości do wyznaczania kolejnego kroku wprost.



$u^{n+\frac{1}{2}}$- pomocnicze, nie zachowywane powyżej t^{n+1}

Stabilność

$$\varepsilon^{n+1} = \varepsilon^n - \left.\frac{\partial f}{\partial u}\right|_n \Delta t \left[1 - \left.\frac{\partial f}{\partial u}\right|_n \frac{\Delta t}{2}\right] \varepsilon^n$$

$$\alpha = \left.\frac{\partial f}{\partial u}\right|_n \Delta t$$

$$\varepsilon^{n+1} = \varepsilon^n - \alpha \left[1 - \frac{\alpha}{2}\right] \varepsilon^n$$

$$g = 1 - \alpha + \frac{1}{2}\alpha^2$$

- stabilna dla α rzeczywistego, gdy $\Delta t \leqslant \frac{2}{\left.\frac{\partial f}{\partial u}\right|_n}$
- może być niestabilna dla α urojonego

Metoda niejawna drugiego rzędu

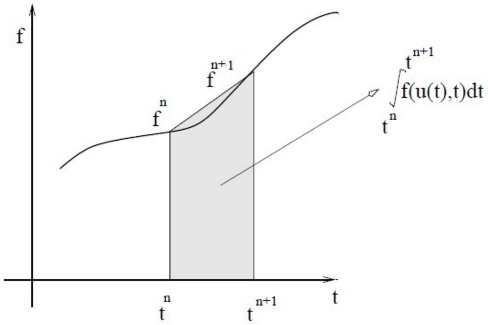
$$u^{n+1} = u^n - [f(u^n, t^n) + f(u^{n+1}, t^{n+1})] \cdot \frac{\Delta t}{2}$$

- u^{n+1} - uwikłane (jako nieznany argument funckcji $f(u^{n+1}, t^{n+1})$), której wartości potrzebujemy, aby go wyliczyć
- gdy f bardziej skomplikowana to w każdym kroku czasowym może wymagać rozwiązania równania
- metoda ma dokładność 2-go rzędu

Stabilność

$$g = 1 - \frac{\partial f}{\partial u}\bigg|_n \cdot \frac{\Delta t}{2} - \frac{\partial f}{\partial u}\bigg|_{n+1} \cdot \frac{\Delta t}{2} g$$

$$g = \frac{1 - \frac{\partial f}{\partial u}\big|_n \cdot \frac{\Delta t}{2}}{1 + \frac{\partial f}{\partial u}\big|_{n+1} \cdot \frac{\Delta t}{2}}$$



Stabilność c.d.

- równanie rozpadu: $\frac{\partial f}{\partial u} > 0, \quad |g| < 1$ (zawsze)
- równanie oscylacyjne $\frac{\partial f}{\partial u}$ - urojone, $|g| = 1$

W obu przypadkach metoda stabilna niezależnie od wyboru kroku

- bezwzględnie stabilna, ważne w zagadnieniach nieliniowych
- cena: konieczność rozwiązywania równania algebraicznego na u^{n+1} lub stosowania wzoru iteracyjnego.

Metoda Rungego-Kutty dla pojedynczego równania

$$\frac{du(t)}{dt} = f(t,u) \qquad t \in [a,b] \quad u_0 = u(t_0)$$

Rozwiązanie $u(t)$ wyznaczamy w punktach

$$t_i = t_0 + i \cdot h \qquad h = \frac{b-a}{n}, \quad i = 0, 1, \ldots, n$$

W metodach jednokrokowych jawnych rozwiązanie:

$$u_{i+1} = u_i + h \cdot F(t_i, u_i, h) \qquad i = 0, 1, \ldots, n-1 \qquad (*)$$

F - zadana funkcja

Metody R-K - szczególny przypadek (*)

$$u(t_{i+1}) = u(t_i) + h \cdot u'(t_1) + \frac{h^2}{2} \cdot u''(t_i) + \ldots + \frac{h^q}{q!} \cdot u^{(q)}(t_i) + R$$

dla $\mathbf{i = 0}$:

$$u_1 = u(t_0) + h \cdot \underbrace{u'}_{!!}(t_0) + \frac{h^2}{2} \cdot \underbrace{u''}_{!!}(t_0) + \ldots + \frac{h^q}{q!} \cdot \underbrace{u^{(q)}}_{!!}(t_0) \qquad (**)$$

!! - należy wyznaczyć

$u'(t_0) = f(t_0, u_0)$
$u''(t_0) = \frac{\partial f}{\partial t}\|_0 + \frac{\partial f}{\partial u}\|_0 \cdot u'_0 = \frac{\partial f}{\partial t}\|_0 + \frac{\partial f}{\partial u}\|_0 \cdot f(t_0, u_0)$ (pochodna zupełna)
itd...
$u'(t_0)$ - prosto, wyższe - uciążliwe obliczenia!

Równanie (**) można zapisać jako:

$$\begin{aligned} u_1 &= u_0 + h \cdot \left[u'(t_0) + \tfrac{h}{2} u''(t_0) + \ldots + \tfrac{h^{q-1}}{q!} u^{(q)}(t_0) \right] = \\ &= u_0 + h \cdot \underbrace{\left[f(t_0, u_0) + \frac{h}{2} \cdot f^{(1)}(t_0, u_0) + \ldots + \frac{h^{q-1}}{q!} f^{q-1}(t_0, u_0) \right]}_{F_T} \end{aligned}$$

gdzie (pochodna zupełna):

$$f^{(s)}(t_0, u_0) = \frac{d^s}{dt^s} f(t, u(t)) \bigg|_{t=t_0, u(t)=u_0} \qquad s = 1, 2, \ldots, q-1$$

Idea metod Rungego-Kutty

Ideą metod jest dobranie funkcji F (wzór *):

- bez wyznaczania pochodnych $f^{(s)}(t_0, u_0)$
- tak, by jak najlepiej przybliżyła funkcję F_T

R-punktowa metoda Rungego-Kutty

$$(***) \begin{cases} u_{i+1} = u_i + hF(t_i, u_i, h), \qquad i = 0, 1, 2, \ldots, n-1 \\ F(t, u, h) = c_1 k_1(t, u, h) + c_2 k_2(t, u, h) + \ldots + c_r k_r(t, u, h) \\ k_1(t, u, h) = f(t, u) \\ k_j(t, u, h) = f(t + ha_j, u + h\sum_{s=1}^{j-1} b_{js} k_s(t, u, h)), \quad j = 2, \ldots, r \\ a_j = \sum_{s=1}^{j-1} b_{js} \end{cases}$$

a_j, b_{js}, c_k - stałe rzeczywiste, dobrane tak, aby:

$$\phi(h) = F(t, u, h) - F_T(t, u, h)$$

zawierała jedynie potęgi h możliwie wysokiego rzędu.

Dla $r = 3$ wzory (***) mają postać:

$$
\begin{aligned}
F(t,u,h) &= c_1 \cdot k_1 + c_2 \cdot k_2 + c_3 \cdot k_3 \\
k_1 &= f(t,u) \\
k_2 &= f(t + h \cdot a_2, u + h \cdot b_{21} \cdot k_1 = \\
&= f(t + a_2 \cdot h, u + h \cdot a_2 \cdot k_1) \\
k_3 &= f(t + a_3 \cdot h, u + h \cdot (b_{31} \cdot k_1 + b_{32} \cdot k_2)) = \\
&= f(t + a_3 \cdot h, u + h \cdot (a_3 - b_{32}) \cdot k_1 + h \cdot b_{32} \cdot k_2)
\end{aligned}
$$

Potrzebna nam będzie pochodna zupełna

$$
f^{(1)}(t,u) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial u} \cdot f(t,u)
$$

Niech (dla uproszczenia):

$$
\begin{aligned}
\alpha &= \tfrac{\partial f}{\partial t} + f \tfrac{\partial f}{\partial u} \\
\beta &= \tfrac{\partial^2 f}{\partial t^2} + 2f \tfrac{\partial^2 f}{\partial t \partial u} + f^2 \tfrac{\partial^2 f}{\partial u^2} \qquad \mathrm{f} = \mathrm{f(t,u)}
\end{aligned}
$$

k_2, k_3 rozwijamy w szereg Taylora wokół (t,u):

$$k_2 =$$
$$f(t,u) + ha_2\left(\frac{\partial f}{\partial t} + k_1\frac{\partial f}{\partial u}\right) + \frac{1}{2}h^2 a_2^2\left(\frac{\partial^2 f}{\partial t^2} + 2k_1\frac{\partial^2 f}{\partial t \partial u} + k_1^2\frac{\partial^2 f}{\partial u^2}\right) + O(h^3)$$
$$k_2 = f(t,u) + ha_2\alpha + \frac{1}{2}h^2 a_2^2\beta + O(h^3)$$
$$k_3 = f(t,u) + ha_3\alpha + h^2(a_2 \cdot b_{32} \cdot \alpha\frac{\partial f}{\partial u} + \frac{1}{2}a_3^2\beta) + O(h^3)$$

stąd:

$$\varphi(h) = F(t,u,h) - F_T(t,u,h) =$$
$$= (c_1 + c_2 + c_3 - 1) \cdot f(t,u) + h \cdot \alpha \cdot (c_2 \cdot a_2 + c_3 \cdot a_3 - \frac{1}{2}) +$$
$$+ h^2 \cdot \left[c_3 a_3 b_{32}\alpha\frac{\partial f}{\partial u} + \frac{1}{2}(c_2 \cdot a_2^2 + c_3 a_3^2) - \frac{1}{6}(\alpha \cdot \frac{\partial f}{\partial u} + \beta)\right] + O(h^3)$$

Uwaga!

Należy dobrać $c_1, c_2, c_3, a_2, a_3, b_{32}$ w taki sposób, aby w $\varphi(h)$ były tylko potęgi możliwie wysokiego stopnia.

$\boxed{\mathbf{r = 1}}$

$$c_2 = c_3 = a_2 = a_3 = b_{32} = 0$$

Przyjmuje się $c_1 = 1 \qquad \Rightarrow \qquad \begin{array}{l} \varphi(h) = O(h) \\ F(t,u,h) = f(t,u) \end{array}$

$$u_{i+1} = u_i + h \cdot f(t_i, u_i), \qquad i = 0, 1, \ldots, n-1$$

Wniosek

Jest to metoda Eulera.

$\boxed{\mathbf{r=2}}$

$$c_3 = b_{32} = 0$$

$$\begin{cases} c_1 + c_2 - 1 & = & 0 \\ c_2 \cdot a_2 - \frac{1}{2} & = & 0 \end{cases} \quad \Rightarrow \quad \varphi(h) = O(h^2)$$

1. $c_1 = 0, c_2 = 1, a_2 = \frac{1}{2} \quad \Rightarrow \quad F(t,u,h) = f(t + \frac{1}{2}h, u + \frac{1}{2}h \cdot f(t,u)) \rightarrow$ zmodyfikowana metoda Eulera
2. $c_1 = \frac{1}{2}, c_2 = \frac{1}{2}, a_2 = 1 \quad \Rightarrow \quad F(t,u,h) = \frac{1}{2} \cdot [f(t,u) + f(t+h, u + h \cdot f(t,u))]$

$\boxed{\mathbf{r = 3}}$

Dobieramy $c_1, c_2, c_3, a_2, a_3, b_{32}$, tak, aby w $\varphi(h)$ współczynniki przy h^0, h, h^2 - były zerowe:

$$\begin{cases} c_1 + c_2 + c_3 = 1 \\ c_2 a_2 + c_3 a_3 = \frac{1}{2} \\ c_3 a_2^2 + c_3 a_3^2 = \frac{1}{3} \\ c_3 a_2 b_{32} = \frac{1}{6} \end{cases} \longrightarrow$$

4 równania, 6 niewiadomych

$\Rightarrow$

Trójpunktowa metoda Rungego-Kutty

$$c_1 = \frac{1}{6}, \quad c_2 = \frac{2}{3}, \quad c_3 = \frac{1}{6}$$
$$a_2 = \frac{1}{2}, \quad a_3 = 1, \quad b_{32} = 2$$

Praktyczne znaczenie - 4-punktowe wersje metody Rungego-Kutty

a)

$$
\begin{aligned}
u_{i+1} &= u_i + \tfrac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \qquad i = 0, 1, 2, \dots, n-1\\
k_1 &= f(t_i, u_i),\\
k_2 &= f(t_i + \tfrac{1}{2}h, u_i + \tfrac{1}{2}hk_1),\\
k_3 &= f(t_i + \tfrac{1}{2}h, u_i + \tfrac{1}{2}hk_2),\\
k_4 &= f(t_i + h, u_i + hk_3)
\end{aligned}
$$

b)

$$
\begin{aligned}
u_{i+1} &= u_i + \tfrac{h}{8}(k_1 + 3k_2 + 3k_3 + k_4), \qquad i = 0, 1, 2, \dots, n-1\\
k_1 &= f(t_i, u_i),\\
k_2 &= f(t_i + \tfrac{1}{3}h, u_i + \tfrac{1}{3}hk_1),\\
k_3 &= f(t_i + \tfrac{2}{3}h, u_i - \tfrac{1}{3}hk_1 + hk_2),\\
k_4 &= f(t_i + h, u_i + hk_1 - hk_2 + hk_3)
\end{aligned}
$$

Dla obu: $\varphi_4(h) = O(h^4)$

Wada metod R-4

Wada: np. dla $\mathbf{r} = \mathbf{4}$, na etapie 1 kroku obliczeń wartość f jest wyznaczana 4 razy i nie będą to wartości użyteczne na dalszych etapach.

Metody R-K dla układów ODE

$$\begin{cases} \vec{u}(t) = (u_1(t), u_2(t), \ldots, u_m(t)) \\ \vec{f}(t, \vec{u}) = (f_1(t, \vec{u}), f_2(t, \vec{u}), \ldots, f_m(t, \vec{u})) \\ \vec{u_0} \end{cases}$$

układ równań:

$$\begin{cases} \frac{d}{dt}\vec{u}(t) = \vec{f}(t, \vec{u}(t)), \qquad t \in [a, b] \\ \text{z warunkiem początkowym } \vec{u}(t_0) = \vec{u_0} \end{cases}$$

Postać r-punktowej metody R-K

$$\vec{u}_{i+1} = \vec{u}_i + h \cdot \vec{F}(t_i, \vec{u}_i, h)$$

gdzie:

$$\begin{aligned}
\vec{F}(t, \vec{u}, h) &= c_1 \cdot k_1(t, \vec{u}, h) + c_2 \cdot k_2(t_1, \vec{u}, h) + \ldots + c_r k_r(t, \vec{u}, h) \\
\vec{k}_1(t, \vec{u}, h) &= f(t, \vec{u}), \\
\vec{k}_j(t, \vec{u}, h) &= \vec{f}(t + h \cdot a_{j_1} \vec{u} + h \cdot \textstyle\sum_{s=1}^{j-1} b_{js} \vec{k}_s(t, \vec{u}, h)), \; j = 2, 3, \ldots, r
\end{aligned}$$

$$a_j = \sum_{s=1}^{j-1} b_{js}, \qquad j = 2, 3, \ldots, r$$

a_j, b_{js}, c_r - stałe rzeczywiste, wartości takie same jak dla pojedynczego równania.

Literatura

Zastosowanie metod do symulacji układów ciał niebieskich:
`http://www.artcompsci.org/`
`http://www.artcompsci.org/vol_1/v1_web/`