

Zaawansowane metody numeryczne
Komputerowa analiza zagadnień różniczkowych

6. Metody Rungego-Kutty I

P. F. Góra

<http://th-www.if.uj.edu.pl/zfs/gora/>

semestr letni 2007/08

Definicja metody

Poszukujemy rozwiązania problemu zmienności pochodnej (prawej strony standardowego problemu Cauchy'ego) w trakcie wykonywania kroku całkowania. **Postulujemy, że należy używać pewnej *średniej pochodnej* — średniej ważonej wyliczonej z pochodnych w pewnych punktach pośrednich:**

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{i=1}^s w_i \mathbf{k}_i + O(h^{p+1}), \quad (1a)$$

$$\mathbf{k}_i = \mathbf{f}\left(t_n + h\alpha_i, \mathbf{y}_n + h \sum_{j=1}^s \beta_{ij} \mathbf{k}_j\right). \quad (1b)$$

Równania (1) definiują *s-krokową metodę Rungego-Kutty rzędu p*. Liczby w_i , α_i , β_{ij} są parametrami metody. $\mathbf{y}_l, \mathbf{k}_i \in \mathbb{R}^N$, $\mathbf{f} : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$.

Zgodność metod RK

Najważniejszym kryterium, jakie musi spełniać każda metoda numerycznego całkowania równań różniczkowych, jest zgodność metody — wymaganie, aby metoda odtwarzała wyjściowy problem Cauchy'ego w granicy nieskończenie małych kroków. Dla metod RK z (1a) mamy

$$\lim_{h \rightarrow 0^+} \frac{y_{n+1} - y_n}{h} = \sum_{i=1}^s w_i \lim_{h \rightarrow 0^+} \mathbf{k}_i. \quad (2)$$

Granica lewej strony (2) jest oczywiście $\left. \frac{dy}{dt} \right|_n = \mathbf{f}(t_n, \mathbf{y}_n)$. Z (1b) wynika, że $\lim_{h \rightarrow 0^+} \mathbf{k}_i = \mathbf{f}(t_n, \mathbf{y}_n)$, a zatem **warunkiem zgodności metod RK** jest

$$\sum_{i=1}^s w_i = 1. \quad (3)$$

Zapis metod RK w postaci tabeli

Współczynniki metody RK zapisuje się na ogół w postaci tabeli

$$\begin{array}{c|cccc} \alpha_1 & \beta_{11} & \beta_{12} & \dots & \beta_{1s} \\ \alpha_2 & \beta_{21} & \beta_{22} & \dots & \beta_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_s & \beta_{s1} & \beta_{s2} & \dots & \beta_{ss} \\ \hline & w_1 & w_2 & \dots & w_s \end{array} = \frac{\alpha}{\mathbf{w}^T} \cdot \mathbf{B} \quad (4)$$

Jawne i niejawne metody RK

Jeżeli macierz \mathbf{B} ma niezerowe elementy na i powyżej diagonalu, metoda jest *niejawna*: do obliczenia pochodnej w punkcie pośrednim trzeba znać tęże pochodną (diagonalu) i ewentualnie pochodne w kolejnych punktach pośrednich (wyraży ponad diagonalu), tak więc proces obliczania wektorów k_i staje się rozwiązywaniem skomplikowanego (na ogół nieliniowego) układu równań na te wielkości. Wymiar tego układu wynosi $s \times N$.

Jeżeli macierz \mathbf{B} ma niezerowe elementy tylko pod diagonalu, metoda jest *jawna*: do obliczenia kolejnych k_i potrzebna jest wyłącznie znajomość $k_{j < i}$.

Zależność pomiędzy ilością kroków (s) a rzędem metody (p)

Na pierwszy rzut oka może się wydawać, że każde dodatkowe obliczenie prawej strony równania, czyli każde obliczenie pochodnej w punkcie pośrednim, zwiększa rząd metody o jeden. W rzeczywistości tak nie jest. Dla metod *jawnych* zawsze zachodzi $p \leq s$; minimalną ilość kroków potrzebnych do zrealizowania metody jawnej danego rzędu podaje poniższa tabela:

Rząd metody	1	2	3	4	5	6	7	8
Minimalna ilość kroków	1	2	3	4	6	7	9	11

Przykłady metod RK

Metody Eulera, jawna i niejawna, są, formalnie rzecz biorąc, jednokrokowymi metodami RK rzędu 1. Ich tabelki mają postać

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \\ \text{(jawna)} & \end{array} \quad \begin{array}{c|c} 1 & 1 \\ \hline & 1 \\ \text{(niejawna)} & \end{array} \quad (5)$$

Jawna metoda punktu środkowego jest dwukrokową metodą rzędu drugiego

$$\mathbf{k}_1 = \mathbf{f}(x_n, y_n), \quad (6a)$$

$$\mathbf{k}_2 = \mathbf{f}\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}\mathbf{k}_1\right), \quad (6b)$$

$$y_{n+1} = y_n + h \mathbf{k}_2 + O(h^3) \quad (6c)$$

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}$$

Niejawna metoda punktu środkowego jest jednokrokową metodą rzędu drugiego

$$\mathbf{k}_1 = \mathbf{f} \left(t_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_1 \right), \quad (7a)$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\mathbf{k}_1 + O(h^3). \quad (7b)$$

Tabela dla tej metody ma postać

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$$

Klasyczna metoda czterokrokowa

Najczęściej stosowaną metodą RK jest *klasyczna metoda czterokrokowa* — większość ludzi uważa, że termin “metoda Rungego-Kutty” określa tylko tę metodę 😊

$$\mathbf{k}_1 = \mathbf{f}(t_n, \mathbf{y}_n), \quad (8a)$$

$$\mathbf{k}_2 = \mathbf{f}\left(t_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_1\right), \quad (8b)$$

$$\mathbf{k}_3 = \mathbf{f}\left(t_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_2\right), \quad (8c)$$

$$\mathbf{k}_4 = \mathbf{f}(t_n + h, \mathbf{y}_n + h\mathbf{k}_3), \quad (8d)$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\left(\frac{1}{6}\mathbf{k}_1 + \frac{1}{3}\mathbf{k}_2 + \frac{1}{3}\mathbf{k}_3 + \frac{1}{6}\mathbf{k}_4\right) + O(h^5). \quad (8e)$$

Tabela dla klasycznej metody czterokrokowej

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Zasady wyprowadzania współczynników metod RK

Współczynniki metod RK rzędu p otrzymuje się rozwijając obie strony wyrażenia (dla uproszczenia stosujemy tu notację skalarną — nie ma to jednak wpływu na samo wyprowadzenie)

$$y_{n+1} - y_n = h \sum_i w_i k_i \quad (9)$$

w szereg potęgowy względem h z dokładnością do wyrazów rzędu $O(h^{p+1})$ i żądając równości współczynników rozwinięcia. Ponieważ wektory (pochodne w punktach pośrednich) k_i są wyrażone przez funkcję f (prawą stronę równania), żądamy, aby funkcja f była klasy co najmniej C^{p-1} .

Lewa strona*

Rozwijając lewą stronę wyrażenia (9) otrzymujemy

$$y_{n+1} - y_n = \sum_{l=1}^p \frac{1}{l!} \left. \frac{d^l y}{dt^l} \right|_{t_n} h^l + O(h^{p+1}). \quad (10)$$

Mamy teraz

$$\frac{dy}{dt} = f(t, y) \quad (11a)$$

$$\begin{aligned} \frac{d^2 y}{dt^2} &= \frac{d}{dt} f(t, y) = \frac{\partial f}{\partial t} + \frac{dy}{dt} \frac{\partial f}{\partial y} = \frac{\partial f}{\partial t} + f \frac{\partial f}{\partial y} \\ &= \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) f(t, y) \end{aligned} \quad (11b)$$

$$\begin{aligned}
\frac{d^3 y}{dt^3} &= \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) f(t, y) \\
&= \frac{\partial^2 f}{\partial t^2} + 2f \frac{\partial^2 f}{\partial t \partial y} + f^2 \frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial t} \frac{\partial f}{\partial y} + f \left(\frac{\partial f}{\partial y} \right)^2
\end{aligned} \tag{11c}$$

$$\begin{aligned}
\frac{d^4 y}{dt^4} &= \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) \left(\frac{\partial}{\partial t} + f \frac{\partial}{\partial y} \right) f(t, y) \\
&= \frac{\partial^3 f}{\partial t^3} + 3f \frac{\partial^3 f}{\partial t^2 \partial y} + 3f^2 \frac{\partial^3 f}{\partial t \partial y^2} + f^3 \frac{\partial^3 f}{\partial y^3} \\
&\quad + \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial t^2} + 3 \frac{\partial f}{\partial t} \frac{\partial^2 f}{\partial t \partial y} + 5f \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial t \partial y} + 3f \frac{\partial f}{\partial t} \frac{\partial^2 f}{\partial y^2} + 4f^2 \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial y^2} \\
&\quad + \frac{\partial f}{\partial t} \left(\frac{\partial f}{\partial y} \right)^2 + f \left(\frac{\partial f}{\partial y} \right)^3 .
\end{aligned} \tag{11d}$$

Prawa strona*

$$\begin{aligned} h \sum_i w_i k_i &= h \sum_i w_i \left(\sum_{l=0}^{p-1} \frac{1}{l!} \frac{d^l k_i}{dh^l} \Big|_0 h^l + O(h^p) \right) \\ &= \sum_{l=0}^{p-1} \frac{1}{l!} \left(\sum_i w_i \frac{d^l k_i}{dh^l} \Big|_0 \right) h^{l+1} + O(h^{p+1}), \end{aligned} \quad (12)$$

a zatem musimy obliczyć pochodne k_i tylko do rzędu $p - 1$, nie zaś p . Różnicz-

kując (1b) i korzystając z twierdzenia o funkcjach uwikłanych otrzymujemy

$$\frac{dk_i}{dh} = \alpha_i \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right). \quad (13a)$$

(13a) nie jest jawnym wyrażeniem na pochodne dk_i/dh , ale układem równań, z którego pochodne te można wyliczyć. W wyrażeniu tym wszystkie k_j oraz pochodne f określone są w punkcie $\left(t_n + \alpha_l h, y_n + h \sum_m \beta_{lm} k_m \right)$, gdzie indeks l jest równy, odpowiednio j, i . Postępując analogicznie otrzymujemy

$$\begin{aligned}
\frac{d^2 k_i}{dh^2} &= \alpha_i \frac{d}{dh} \left(\frac{\partial f}{\partial t} \right) + \left[\frac{d}{dh} \left(\frac{\partial f}{\partial y} \right) \right] \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&+ \frac{\partial f}{\partial y} \frac{d}{dh} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&= \alpha_i^2 \frac{\partial^2 f}{\partial t^2} + 2\alpha_i \frac{\partial^2 f}{\partial t \partial y} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&+ \frac{\partial^2 f}{\partial y^2} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^2 \\
&+ \frac{\partial f}{\partial y} \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right)
\end{aligned} \tag{13b}$$

$$\begin{aligned}
\frac{d^3 k_i}{dh^3} &= \alpha_i^2 \frac{d}{dh} \left(\frac{\partial^2 f}{\partial t^2} \right) + 2\alpha_i \left[\frac{d}{dh} \left(\frac{\partial^2 f}{\partial t \partial y} \right) \right] \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&+ 2\alpha_i \frac{\partial^2 f}{\partial t \partial y} \frac{d}{dh} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&+ \left[\frac{d}{dh} \left(\frac{\partial^2 f}{\partial y^2} \right) \right] \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^2 + \frac{\partial^2 f}{\partial y^2} \frac{d}{dh} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^2 \\
&+ \left[\frac{d}{dh} \left(\frac{\partial f}{\partial y} \right) \right] \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right) \\
&+ \frac{\partial f}{\partial y} \frac{d}{dh} \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right)
\end{aligned}$$

$$\begin{aligned}
&= \alpha_i^3 \frac{\partial^3 f}{\partial t^3} + 3\alpha_i^2 \frac{\partial^3 f}{\partial t^2 \partial y} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
&+ 3\alpha_i \frac{\partial^3 f}{\partial t \partial y^2} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^2 + \frac{\partial^3 f}{\partial y^3} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^3 \\
&+ 3\alpha_i \frac{\partial^2 f}{\partial t \partial y} \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right) \\
&+ 3 \frac{\partial^2 f}{\partial y^2} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right) \\
&+ \frac{\partial f}{\partial y} \left(3 \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} + h \sum_j \beta_{ij} \frac{d^3 k_j}{dh^3} \right)
\end{aligned} \tag{13c}$$

Jak już powiedziano, powyższe wzory nie dają jawnych wyrażeń na pochodne k_i , a tylko układ równań, z których pochodne te możnaby wyliczyć. Jednak dla $h = 0$ sytuacja znacznie się upraszcza: położenie $h = 0$ usuwa z prawej strony człony najwyższego rzędu, wobec czego otrzymujemy jawne wyrażenia na pochodne w zerze. I tak

$$k_i(0) = f, \quad (14a)$$

$$\left. \frac{dk_i}{dh} \right|_0 = \alpha_i \frac{\partial f}{\partial t} + \sum_j \beta_{ij} f \frac{\partial f}{\partial y}, \quad (14b)$$

$$\begin{aligned} \left. \frac{d^2 k_i}{dh^2} \right|_0 &= \alpha_i^2 \frac{\partial^2 f}{\partial t^2} + 2\alpha_i \sum_j \beta_{ij} f \frac{\partial^2 f}{\partial t \partial y} + \left(\sum_j \beta_{ij} \right)^2 f^2 \frac{\partial^2 f}{\partial y^2} \\ &+ 2 \sum_j \beta_{ij} \alpha_j \frac{\partial f}{\partial t} \frac{\partial f}{\partial y} + 2 \sum_j \sum_q \beta_{ij} \beta_{jq} f \left(\frac{\partial f}{\partial y} \right)^2, \end{aligned} \quad (14c)$$

$$\begin{aligned}
\left. \frac{d^3 k_i}{dh^3} \right|_0 &= \alpha_i^3 \frac{\partial^3 f}{\partial t^3} + 3\alpha_i^2 \sum_j \beta_{ij} f \frac{\partial^3 f}{\partial t^2 \partial y} + 3\alpha_i \left(\sum_j \beta_{ij} \right)^2 f^2 \frac{\partial^3 f}{\partial t \partial y^2} + \left(\sum_j \beta_{ij} \right)^3 f^3 \frac{\partial^3 f}{\partial y^3} \\
&+ 3 \sum_j \beta_{ij} \alpha_j^2 \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial t^2} + 6\alpha_i \sum_j \beta_{ij} \alpha_j \frac{\partial f}{\partial t} \frac{\partial^2 f}{\partial t \partial y} \\
&+ 6 \sum_{j,q} \beta_{ij} (\alpha_i + \alpha_j) \beta_{jq} f \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial t \partial y} + 6 \sum_j \beta_{ij} \sum_q \beta_{iq} \alpha_q f \frac{\partial f}{\partial t} \frac{\partial^2 f}{\partial y^2} \\
&+ 3 \left(\sum_j \beta_{ij} \left(\sum_q \beta_{jq} \right)^2 + 2 \sum_j \beta_{ij} \sum_{q,r} \beta_{iq} \beta_{qr} \right) f^2 \frac{\partial f}{\partial y} \frac{\partial^2 f}{\partial y^2} \\
&+ 6 \sum_{j,q} \beta_{ij} \beta_{jq} \alpha_q \frac{\partial f}{\partial t} \left(\frac{\partial f}{\partial y} \right)^2 + 6 \sum_{j,q,r} \beta_{ij} \beta_{jq} \beta_{qr} f \left(\frac{\partial f}{\partial y} \right)^3. \tag{14d}
\end{aligned}$$

Funkcję f i jej pochodne należy teraz obliczać w punkcie (t_n, y_n) .

Warunki

Pozostaje nam teraz uzgodnić współczynniki rozwinięcia (9) w kolejnych potęgach h . Zauważmy, że ponieważ metoda ma działać dla dowolnych f spełniających podane założenia, współczynniki występujące po obu stronach (9) przy pewnej kombinacji f i jej pochodnych muszą być sobie równe.

$p = 1$

W najniższym rzędzie uzgadniamy tylko współczynniki przy wyrazach rzędu h^1 .
Otrzymujemy

$$\frac{1}{1!}f = \frac{1}{0!} \sum_{i=1}^s w_i f ,$$

czyli

$$\sum_{i=1}^s w_i = 1 , \tag{15}$$

co, jak już wiemy, jest warunkiem zgodności metody RK. Widać zatem, że każda zgodna metoda RK, bez względu na to, czy zachodzą jakieś dodatkowe związki pomiędzy jej współczynnikami, czy nie, jest metodą rzędu co najmniej pierwszego.

$p = 2$

Aby metoda była rzędu drugiego, żądamy spełnienia (15) oraz następujących warunków dodatkowych, wynikających z postulatu zgodności współczynników przy wyrazach rzędu h^2 :

$$\sum_{i=1}^s w_i \alpha_i = \frac{1}{2}, \quad (16a)$$

$$\sum_{i=1}^s w_i \sum_{j=1}^s \beta_{ij} = \frac{1}{2}. \quad (16b)$$

Punkt ten wymaga pewnej dyskusji. Odejmując stronami (16a), (16b) otrzymujemy

$$\sum_{i=1}^s \left(\alpha_i - \sum_{j=1}^s \beta_{ij} \right) w_i = 0. \quad (17)$$

Gdyby w_i były liniowo niezależne, powyższy warunek oznaczałby, że koniecznie $\alpha_i = \sum_{j=1}^s \beta_{ij}$. Ale w_i *nie* są liniowo niezależne, jako że wiąże je warunek zgodności (15)! W tej sytuacji (17) narzuca pewne dodatkowe więzy na współczynniki metody: można go potraktować jako dodatkowy więz na wagi w_i (tylko $s - 2$ jest wówczas liniowo niezależnych), jednak znacznie wygodniej jest potraktować to jako warunek znikania współczynników kombinacji liniowej stojącej po

prawej stronie (17) i tak też uczynimy. Można też na to spojrzeć i od innej strony: Zgodność wyrazów rozwinięcia ma zachodzić w każdym rzędzie *niezależnie*, co jest możliwe tylko, gdy $\alpha_i = \sum_{j=1}^s \beta_{ij}$.

Wszystkie metody RK rzędu drugiego opisane są zatem układem równań

$$\sum_{i=1}^s w_i = 1, \quad (18a)$$

$$\sum_{i=1}^s w_i \alpha_i = \frac{1}{2}, \quad (18b)$$

$$\sum_{j=1}^s \beta_{ij} = \alpha_i. \quad (18c)$$

Zauważmy, iż z (18c) wynika, że we wszystkich metodach jawnych $\alpha_1 = 0$.

$p = 3$

W trzecim rzędzie dostajemy

$$\sum_{i=1}^s w_i \alpha_i^2 = \frac{1}{3}, \quad (19a)$$

$$\sum_{i=1}^s w_i \sum_{j=1}^s \beta_{ij} \alpha_j = \frac{1}{6}, \quad (19b)$$

gdzie skorzystaliśmy z (18c) w celu zmniejszenia ilości niezależnych równań. Równania te, wraz z równaniami (18a)–(18c), opisują wszystkie metody RK rzędu trzeciego. Widać, że o ile równania (18a)–(18c) narzucały związki liniowe na współczynniki α , β , równania (19a)–(19b) narzucają związki kwadratowe.

$p = 4$

W czwartym rzędzie, po skorzystaniu z (18c), jako nietrywialne równania dostajemy

$$\sum_{i=1}^s w_i \alpha_i^3 = \frac{1}{4}, \quad (20a)$$

$$\sum_{i=1}^s w_i \alpha_i \sum_{j=1}^s \beta_{ij} \alpha_j = \frac{1}{8}, \quad (20b)$$

$$\sum_{i=1}^s w_i \sum_{j=1}^s \beta_{ij} \alpha_j^2 = \frac{1}{12}, \quad (20c)$$

$$\sum_{i=1}^s w_i \sum_{j=1}^s \beta_{ij} \sum_{l=1}^s \beta_{jl} \alpha_l = \frac{1}{24}. \quad (20d)$$

Warto zaznaczyć, iż po zastosowaniu (18c) niektóre z równań na współczynniki są spełnione tożsamościowo, co stanowi niejako test wewnętrznej spójności teorii. Równania (20a)–(20d), łącznie z równaniami (19a)–(19b) oraz (18a)–(18c), opisują wszystkie metody RK rzędu czwartego.

Równania określające metody czwartego rzędu narzucają sześciennie związki pomiędzy współczynnikami α , β .

Wyrażenia, które muszą spełniać współczynniki metod wyższego rzędu, wyprowadza się na tej samej zasadzie co powyższe, są one jednak bardziej skomplikowane. W dowolnym rzędzie p równania te zawierają człon

$$\sum_{i=1}^s w_i \alpha_i^{p-1} = \frac{1}{p} \quad (21)$$

i narzucają związki rzędu $p - 1$.

Stabilność metod RK

W celu zbadania stabilności metody RK, zaburzymy y_n w (1a) oraz (1b) przez pewne ε_n , przy czym $\|\varepsilon_n\| \ll 1$. W miejsce (1a) otrzymamy

$$\begin{aligned} y_{n+1} + \varepsilon_{n+1} &= y_n + \varepsilon_n + h \sum_{i=1}^s w_i \mathbf{k}_i(y_n + \varepsilon_n) \\ &\simeq y_n + \varepsilon_n + h \sum_{i=1}^s w_i \mathbf{k}_i(y_n) + h \sum_{i=1}^s w_i \mathbf{L}_i \varepsilon_n \\ &= y_{n+1} + \varepsilon_n + h \mathbf{w}^T \vec{\mathbf{L}} \varepsilon_n = y_{n+1} + \mathbf{G} \varepsilon_n, \end{aligned} \quad (22)$$

$\mathbf{L}_i, i = 1, \dots, s$, są macierzami o składowych

$$[\mathbf{L}_i]_{\mu\nu} = \frac{\partial (\mathbf{k}_i)_\mu}{\partial (\boldsymbol{\varepsilon}_n)_\nu}, \quad (23)$$

$\vec{\mathbf{L}}$ oznacza wektor złożony z tych macierzy, $\mathbf{w}^T = [w_1, \dots, w_s]^T$, zaś $\mathbf{G} = \mathbb{I} + h\vec{\mathbf{w}}^T \vec{\mathbf{L}}$ jest macierzą wzmocnienia. Metoda jest stabilna, gdy wszystkie wartości własne macierzy wzmocnienia \mathbf{G} spełniają zależność $|g| < 1$ (wartości własne g mogą być zespolone). Kluczowe staje się obliczenie macierzy \mathbf{L}_i . Mamy

$$\begin{aligned}
\frac{\partial (\mathbf{k}_i)_\mu}{\partial (\boldsymbol{\varepsilon}_n)_\nu} &= \frac{\partial \left(\mathbf{f} \left(x_n + \alpha_i h, \mathbf{y}_n + \boldsymbol{\varepsilon}_n + h \sum_j^s \beta_{ij} \mathbf{k}_j (\mathbf{y}_n + \boldsymbol{\varepsilon}_n) \right) \right)_\mu}{\partial (\boldsymbol{\varepsilon}_n)_\nu} \\
&= \sum_{\sigma=1}^N \mathbf{J}_{\mu\sigma} \left(\delta_{\sigma\nu} + h \sum_{j=1}^s \beta_{ij} \frac{\partial (\mathbf{k}_j)_\sigma}{\partial (\boldsymbol{\varepsilon}_n)_\nu} \right). \tag{24}
\end{aligned}$$

\mathbf{J} jest macierzą Jacobiego funkcji \mathbf{f} po drugim argumencie. Formalnie rzecz biorąc, jej elementy należałoby obliczać w każdym punkcie pośrednim z osobna. Przyjmijmy jednak przybliżenie polegające na tym, iż obliczamy tę macierz w lewym krańcu przedziału — prowadzi to do poprawek wyższego rzędu w h .

Reprezentacja diagonalna

Przejdźmy do reprezentacji, w której macierz ta jest diagonalna i oznaczmy przez λ pewną wartość własną macierzy \mathbf{J} . Formalnie rzecz biorąc, poniższą procedurę należy powtórzyć dla wszystkich wartości własnych \mathbf{J} , okaże się jednak, że nie jest to konieczne. Równanie (24) możemy przepisać teraz w postaci

$$\mathbf{L}_i = \lambda \left(1 + h \sum_{j=1}^s \beta_{ij} \mathbf{L}_j \right), \quad i = 1, \dots, s. \quad (25)$$

(25) jest układem równań, z którego można wyliczyć poszukiwane wielkości \mathbf{L}_i .

Współczynnik wzmocnienia

Układ ten można zapisać w postaci

$$[\mathbb{I} - \lambda h \mathbf{B}] \vec{\mathbf{L}} = \lambda \hat{\mathbf{e}}, \quad (26)$$

gdzie $\hat{\mathbf{e}} = [1, 1, \dots, 1]^T$, zaś $[\mathbf{B}]_{ij} = \beta_{ij}$. Oznaczając $z = \lambda h$, ostatecznie otrzymujemy następujące wyrażenie na współczynnik wzmocnienia:

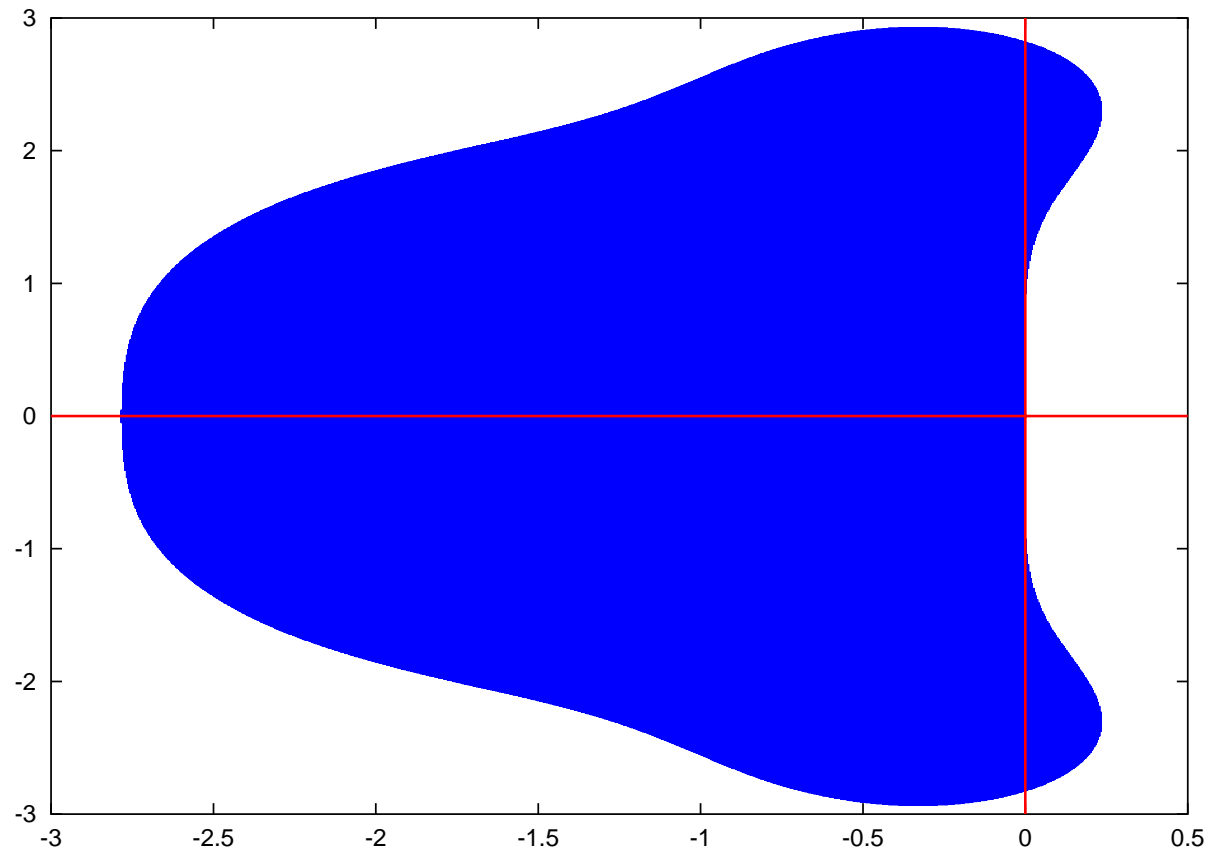
$$g = 1 + z \vec{\mathbf{w}}^T [\mathbb{I} - z \mathbf{B}]^{-1} \hat{\mathbf{e}}. \quad (27)$$

g jest liczbą, nie macierzą; “macierzowość” macierzy wzmocnienia \mathbf{G} brała się z “macierzowości” jacobianu, a skoro jacobian zastąpiliśmy wartością własną, \mathbf{G} zastąpiliśmy jej wartościami własnymi.

Obszar stabilności

Badanie stabilności metod RK sprowadza się teraz do wyznaczenia *obszaru stabilności* metody, to jest zbioru takich $z \in \mathbb{C}$, że współczynniki g wyliczone dla wszystkie wartości własnych macierzy Jacobiego spełniają $|g| < 1$. Metoda RK jest stabilna, jeśli dla *wszystkich* wartości własnych macierzy Jacobiego we *wszystkich* punktach pośrednich, $z = h \times$ (wartość własna) należy do obszaru stabilności. Zauważmy, że w ten sposób “uwolniliśmy się” niejako od szczegółów równania, które rozwiązujemy: własności równania (funkcji f) tkwią w wartościach własnych macierzy Jacobiego, czyli w liczbach z . **Kształt obszaru stabilności zależy tylko od wybranej metody RK.**

Obszar stabilności klasycznej czterokrokowej metody Rungego-Kutty



Obszary stabilności metod jawnych i niejawnych

Współczynnik wzmocnienia g może być albo funkcją wymierną, albo wielomianem w z . W tym drugim przypadku obszar stabilności metody jest z całą pewnością skończony, jako że każdy wielomian rośnie nieograniczenie (na moduł) przy $z \rightarrow \infty$. Jeśli natomiast współczynnik g jest funkcją wymierną w z , to jest możliwe, że obszar stabilności metody jest nieograniczony. Jeśli w szczególności cała lewa półpłaszczyzna płaszczyzny zespolonej należy do obszaru stabilności, mówimy, że metoda jest A–stabilna. Oznacza to, że dowolnie wielki krok h nie prowadzi do utraty stabilności (oczywiście nie znaczy to, że rozwiązanie numeryczne uzyskane z krokiem $h \gg 0$ będzie miało wiele wspólnego z rozwiązaniem analitycznym — znaczy to tylko, że błąd z poprzednich kroków nie przenosi się do kroków następnych).

W jaki sposób g może stać się funkcją wymierną w z ? Może się tak stać tylko przy obliczaniu $[\mathbb{I} - z\mathbf{B}]^{-1}$, a jeszcze ściślej, przy dzieleniu przez wyznacznik $\det [\mathbb{I} - z\mathbf{B}]$. Zauważmy, że dla metody jawnej macierz \mathbf{B} ma tylko elementy poddiagonalne, a zatem $\det [\mathbb{I} - z\mathbf{B}] = 1$, a skoro tak, to g są wielomianami w z . Widzimy zatem, że **jawne metody RK mają ograniczony obszar stabilności** — *tylko metody niejawne mogą mieć nieograniczony obszar stabilności.*