

WYKŁAD 5

WPROWADZENIE DO METOD NUMERYCZNYCH DLA RÓWNAŃ RÓŻNICZKOWYCH ZWYCZAJNYCH (ZAGADNIENIA POCZĄTKOWE)

Celem tego wykładu jest elementarne wprowadzenie do problematyki analizy numerycznej **zagadnień początkowych formułowanych dla równań różniczkowych zwyczajnych** i ich układów. Są to zagadnienia o wielkim znaczeniu dla praktyki inżynierskiej, a to z następujących powodów:

1. Równania różniczkowe zwyczajne są narzędziem stosowanym powszechnie do modelowania dynamiki rozmaitych zjawisk i układów fizycznych i technicznych (mechanicznych, elektrycznych, elektronicznych, chemicznych, biologicznych, itd.), w których fakt rozciągłości przestrzennej może być pominięty lub nie prowadzi do konieczności użycia równań różniczkowych cząstkowych (np. dynamika ciała sztywnego). O układach takich mówimy, że mają „parametry skupione”, tj. stan takich układów może być opisane za pomocą zbioru funkcji zależnych wyłącznie od czasu.
2. Znakomita większość istotnych dla zastosowań technicznych zagadnień formułowanych w terminach równań różniczkowych nie może być rozwiązana metodami analitycznymi, a jeśli nawet to jest to trudne i wymaga rozległej wiedzy teoretycznej.

Rozważmy zagadnienie początkowe dla równania różniczkowego 1-szego rzędu w postaci standardowej, tj.

$$\begin{cases} y'(t) = F[t, y(t)] \\ y(t_0) = y_0 \end{cases}$$

Funkcja F i występujące w warunku początkowym wartości t_0 i y_0 są zadane. Jeśli funkcja prawej strony F nie zależy jawnie od czasu, tj. $F = F(y)$, powyższe równanie nazywamy autonomicznym.

Pominiemy dalej prezentację ogólnej teorii związanej z istnieniem i jednoznacznością rozwiązań. Zamiast tego, rozważymy dwa pouczające przykłady.

Pokażemy po pierwsze, że rozwiązanie nie musi być jednoznaczne. Oto „standardowy” przykład takiego zagadnienia

$$\begin{cases} y'(t) = 3y^{2/3}(t) \\ y(0) = 0 \end{cases}$$

Łatwo sprawdzić, że posiada ono dwa rozwiązania, a mianowicie

$$Y_1(t) \equiv 0 \quad , \quad Y_2(t) = t^3$$

Powodem takiej sytuacji jest pogwałcenie przez funkcję prawej strony **warunku Lipschitza** w otoczeniu zera. Istotnie, pochodna $F'(y) = 2y^{-1/3}$ jest nieograniczona, gdy $y \rightarrow 0$.

Następny przykład pokazuje, że rozwiązanie (choć jedyne) może nie istnieć dla dowolnie długiego czasu. Rozważmy bowiem zagadnienie

$$\begin{cases} y'(t) = y^2(t) \\ y(0) = 1 \end{cases}$$

Rozwiązanie ma postać $Y(t) = \frac{1}{1-t}$ i jest oczywiste, że $\lim_{t \rightarrow 1^-} Y(t) = +\infty$, tj. rozwiązanie to „eksploduje” w skończonym czasie $t = 1$.

Metoda numeryczna – główna idea

Rozwiązując numerycznie zagadnienie początkowe staramy się określić przybliżoną wartość rozwiązania w skończonym (i zwykle bardzo liczny) zbiorze dyskretnych chwil czasu w pewnym przedziale $[t_0, t_0 + T]$. Dystans pomiędzy dwiema sąsiednimi chwilami nazywamy krokiem całkowania, tj. $h_k = t_{k+1} - t_k$.

W najprostszym podejściu, krok całkowania jest stały, ale często stosuje się (o zastosowana metoda na to pozwala) strategię zmiany kroku opartą na oszacowaniu dokładności przybliżonego rozwiązania (**adaptacja kroku całkowania**).

Rozwiązanie przybliżone wyznaczamy „z kroku na krok” posługując się pewną receptą (zwaną metodą lub schematem całkowania), która mówi jak obliczyć przybliżone rozwiązanie w kroku $k+1$ -szym (w czasie $t_{k+1} = t_k + h_k$) znając je w kroku k -tym (i – ewentualnie – poprzednich).

Ogólna postać schematu całkowania naszego zagadnienia może być zapisana następująco:

$$y_{k+1} = y_k + G[t_{k-(n-1)}, y_{k-(n-1)}; t_{k-(n-2)}, y_{k-(n-2)}; \dots; t_k, y_k; t_{k+1}, y_{k+1}]$$

gdzie y_j , $j = k - (n - 1), \dots, k + 1$ oznacza przybliżoną wartość rozwiązania w chwili $t = t_j$.

Funkcja G nie musi być zadana wprost wzorem – powyższy zapis symbolizuje pewną procedurę obliczeniową (być może bardzo złożoną), której celem jest wyliczenie przyrostu rozwiązania w przedziale czasu $[t_k, t_{k+1}]$,

W ogólności, powyższa „recepta” może być **wielokrokowa**, tj. może wykorzystywać wartości rozwiązania i/lub wartości funkcji prawej strony z kilku poprzednich kroków (czyli liczna n w powyższym wzorze jest większa od 1). Jest oczywistym, że tylko metoda **jednokrokowa** ($n = 1$) jest **samostartująca**, albowiem w zagadnieniu warunek początkowy jest postawiony tylko dla jednej (początkowej) chwili czasu.

Mówimy, że **schemat całkowania jest jawny** (otwarty, ang. explicit) jeśli procedura całkowania (symbolizowana przez funkcje G) nie zawiera w sposób uwikłany odwołania do wielkości wyznaczanej, czyli y_{k+1} . Innymi słowy, na liście argumentów funkcji G nie występuje y_{k+1} ani t_{k+1} .

W przeciwnym wypadku schemat całkowania nazywamy **niejawnym** (zamkniętym, ang. implicit). Jeśli schemat jest niejawny to formuła metody całkowania jest uwikłana i w każdym kroku czasowym **trzeba rozwiązać pewien problem algebraiczny** (zwykle nieliniowy), aby wyznaczyć y_{k+1} . Wynika stąd, że metody niejawne jest bardziej skomplikowana w użyciu, posiadają jednak pewną istotną przewagę nad metodami jawnymi, o czym później.

Metoda Eulera. Zgodność, stabilność i zbieżność.

Najprostsza metoda całkowania to **Metoda Eulera**. Formułę tej metody wyprowadzamy stosując rozwinięcie Taylora. Załóżmy, że $y = Y(t)$ jest rozwiązaniem ścisłym zagadnienia początkowego, przy czym funkcja Y posiada ciągłe pochodne do rzędu równego (przynajmniej) 2. Możemy wówczas napisać równość (na podstawie tw. Taylora)

$$Y(t+h) = Y(t) + h Y'(t) + \frac{1}{2}h^2 Y''(t + \theta \cdot h), \quad \theta \in (0,1) \\ = F(t, Y(t))$$

Dla dostatecznie małego kroku całkowania h człon z h^2 może być zaniedbany, co prowadzi do następującej formuły

$$y_{k+1} = y_k + h \underbrace{F[t_k, y_k]}_{F_k} = y_k + hF_k$$

gdzie y_k, y_{k+1} to przybliżone wartości, odpowiednio, $Y(t_k)$ i $Y(t_k + h)$.

Jaki błąd popełniamy stosując powyższą formułę? Dokładniej: jaka jest różnica pomiędzy rozwiązaniem ścisłym i przybliżonym w chwili $t = t_{k+1}$, jeśli założyć, że $y_k \equiv Y(t_k)$?

Różnicę tę nazywamy **lokalnym błędem całkowania** (LBC) i oznaczamy $e(h)$. Jest to podstawowa charakterystyka każdej metody całkowania przybliżonego równań różniczkowych.

Mówimy, że **metoda całkowania jest zgodna** wtedy i tylko wtedy, gdy $\lim_{h \rightarrow 0} e(h) = 0$.

Każda sensowna metoda musi być zgodna.

Zgodność metody Eulera jest łatwa do wykazania. Z wyprowadzenia formuły tej metody wynika natychmiast, że

$$e(h) \equiv Y(t_k + h) - \hat{y}_{k+1} = \frac{1}{2}h^2 Y''(t_k + \theta \cdot h)$$

gdzie oznaczyliśmy $\hat{y}_{k+1} = Y(t_k) + hF[t_k, Y(t_k)]$. Jasnym jest, że $\lim_{h \rightarrow 0} e(h) = 0$.

Ważną charakterystyką każdej metody całkowania jest jej **rząd dokładności**. Mówimy, że metoda całkowania ma **rząd dokładności równy r** (lub po prostu – jest rzędu r) wtedy i tylko wtedy, gdy

$$e(h) = Ch^{r+1} + o(h^{r+1}), \quad C \neq 0, \quad \lim_{h \rightarrow 0} \frac{o(h^{r+1})}{h^{r+1}} = 0,$$

Wnioskujemy natychmiast, że **Metoda Eulera jest metodą 1-szego rzędu**.

Innym ważnym zagadnieniem jest **problem stabilności** metody całkowania. Jest to zagadnienie szerokie i zbyt zaawansowane aby omawiać je szczegółowo, toteż ograniczymy się do wyjaśnienia idei stabilności na przykładzie.

Rozważmy następujący problem „testowy”

$$y'(t) = \lambda y(t) \quad , \quad y(0) = Y_0$$

gdzie $\lambda = \lambda_R + i\lambda_I$ jest zadaną liczbą zespoloną. Ścisłe rozwiązanie tego problemu ma postać:

$$Y(t) = Y_0 \exp(\lambda t) = Y_0 e^{\lambda_R t} (\cos \lambda_I t + i \sin \lambda_I t)$$

Zauważmy, że jeśli $\lambda_R < 0$, to $\lim_{t \rightarrow \infty} |Y(t)| = 0$.

Zastosujmy do zagadnienia testowego metodę Eulera. Dla rozważanego przypadku metoda ta generuje geometryczny ciąg liczb (zespolonych), opisany następującą **regułą rekurencyjną** (ciąg geometryczny)

$$y_{k+1} = y_k + hF_k = y_k + h\lambda y_k = y_k (1 + h\lambda)$$

z której wynika, że kolejne wartości przybliżonego rozwiązania zadane są wzorem

$$y_k = (1 + h\lambda)^k Y_0$$

Powstaje pytanie: kiedy otrzymany ciąg liczb adekwatnie odwzorowuje zachowanie rozwiązana dokładnego dla dużych czasów?

Jeśli $\lambda_R < 0$ to ciąg ten powinien być ograniczony i jego elementy powinny dążyć do zera gdy $k \rightarrow \infty$. Tak będzie jedynie wtedy, gdy $|1 + h\lambda| < 1$, tj. gdy liczba zespolona $h\lambda$ należy do wnętrza okręgu (na płaszczyźnie zespolonej) o równaniu $|z - (-1, 0)| = 1$. Zauważmy, że jeśli krok czasowy jest zbyt wielki i liczba $h\lambda$ leży poza tym okręgiem, to ciąg liczbowy $\{y_k, k = 0, 1, 2, \dots\}$ jest nieograniczony – mówimy, że metoda jest wówczas **niestabilna**.

Jeżeli $\lambda \in R$ i $\lambda = -\mu < 0$ to **warunek stabilności** przyjmuje postać

$$-1 < 1 - h\mu < 1 \Leftrightarrow h < \frac{2}{\mu}$$

Mówimy, że metoda **Eulera jest warunkowo stabilna, tj. stabilność ma miejsce pod warunkiem, że krok całkowania jest dostatecznie mały.**

Zauważmy, że dla ujemnych i rzeczywistych wartości parametru λ rozwiązanie dokładne maleje monotonicznie do zera. Ciąg wartości rozwiązania przybliżonego będzie poprawnie odtwarzał taki charakter rozwiązania jedynie wtedy jeśli spełniony będzie jeszcze mocniejszy warunek, a mianowicie $0 < 1 - h\mu < 1$, tj. gdy $h < 1/\mu$.

Rozważmy zagadnienie wyznaczenia przybliżonego rozwiązania w pewnym przedziale czasu o długości T , tj. $t \in (t_0, t_0 + T)$. Załóżmy, że liczba kroków całkowania wynosi n , a wszystkie kroki są tej samej długości $h = T / n$.

Globalnym błędem całkowania (GBC) danej metody numerycznej nazywać będziemy wielkość

$$E(h) = \max_{0 < k \leq n} |Y(t_0 + kh) - y_k|$$

Zakładamy oczywiście, że $y_0 \equiv Y(t_0)$.

Mówimy, że **metoda całkowania jest zbieżna** wtedy i tylko wtedy, gdy $\lim_{\substack{h \rightarrow 0 \\ nh=T}} E(h) = 0$.

Zbieżność metody to własność, która gwarantuje, że największe odchylenie rozwiązania przybliżonego od rozwiązania ścisłego może być uczynione dowolnie małym poprzez odpowiedni dobór kroku całkowania. Zauważmy, że im mniejszy krok tym więcej ich trzeba wykonać! W **metodzie zbieżnej nie następuje zatem katastrofalna kumulacja błędów** popełnianych w każdym kroku całkowania.

Pokażemy, że **metoda Eulera jest zbieżna**. Jeśli rozważymy problem testowy, to wniosek jest natychmiastowy, mamy bowiem

$$y_n = (1 + h\lambda)^n Y_0 = (1 + \frac{\lambda T}{n})^n Y_0 \xrightarrow{n \rightarrow \infty} Y_0 e^{\lambda T} = Y(T)$$

W przypadku ogólnym dowód opiera się na założeniu, że funkcja prawej strony F spełnia warunek Lipschitza względem drugiego argumentu, tzn.

$$\exists L > 0, \forall t \in I, \forall y_1, y_2 \in R: |F(t, y_1) - F(t, y_2)| \leq L \cdot |y_1 - y_2|$$

Zgodnie z tw. Taylora rozwiązanie dokładne spełnia równość

$$Y(t_{k+1}) = Y(t_k) + hF[t_k, Y(t_k)] + B_k, \quad B_k = \frac{1}{2} h^2 Y''(\tau_k), \quad \tau_k \in (t_k, t_{k+1})$$

Rozwiązanie przybliżone w chwili $t = t_{k+1}$ obliczane jest natomiast ze wzoru

$$y_{k+1} = y_k + hF(t_k, y_k)$$

Odejmując stronami otrzymujemy związek pomiędzy $E_k = Y(t_k) - y_k$ i $E_{k+1} = Y(t_{k+1}) - y_{k+1}$:

$$E_{k+1} = E_k + h[F(t_k, Y_k) - F(t_k, y_k)] + B_k$$

Z warunku Lipschitza i oczywistej nierówności $|a + b| \leq |a| + |b|$ wynika, że:

$$|E_{k+1}| \leq (1 + hL)|E_k| + |B_k| \leq (1 + hL)|E_k| + B$$

gdzie $B = \frac{1}{2}h^2M_2$, $M_2 = \max_{t \in [t_0, t_0+T]} |Y''(t)|$. Otrzymana nierówność ma formę

$$|E_{k+1}| \leq a|E_k| + b \quad , \quad a = 1 + hL \quad , \quad b = B$$

Stosując metodę indukcji można wykazać prawdziwość następującej nierówności (ćwiczenie!)

$$|E_k| \leq b \frac{a^k - 1}{a - 1} + a^k |E_0|$$

czyli

$$|E_k| \leq B \frac{(1 + hL)^k - 1}{hL} + (1 + hL)^k |E_0|$$

Dalej, korzystając z oczywistej nierówności $e^x > 1 + x$ mamy dla $k = n$

$$|E_n| \leq B \frac{(1 + hL)^n - 1}{hL} + (1 + hL)^n |E_0| \leq B \frac{e^{Lhn} - 1}{hL} + e^{Lhn} |E_0| = \frac{B}{hL} (e^{LT} - 1) + e^{LT} |E_0|$$

Kładąc $E_0 = 0$ (warunek początkowy zadany jest dokładnie) i podstawiając $B = \frac{1}{2}h^2 M_2$ otrzymujemy oszacowanie błędu na końcu przedziału całkowania o długości T

$$|E_n| \leq \frac{M_2}{2L} (e^{LT} - 1)h$$

Jak widać, dla dowolnego T mamy $\lim_{h \rightarrow 0} |E_n| = 0$, co dowodzi zbieżności. Należy zauważyć, że GBC maleje proporcjonalnie do h .

Mogłoby się wydawać, że zgodność metody całkowania jest warunkiem wystarczającym jej zbieżności. Następujący kontrprzykład pokazuje, że wcale nie musi tak być!

Rozważmy tzw. schemat skokowy zadany wzorem

$$y_{k+1} = y_{k-1} + 2h F(t_k, y_k)$$

Schemat ten jest zgodny i jego rząd wynosi 2. Dowodzi tego następujący rachunek

$$Y(t_k + h) = Y(t_k) + hY'(t_k) + \frac{1}{2}h^2Y''(t_k) + \frac{1}{6}h^3Y'''(\xi_+)$$

$$Y(t_k - h) = Y(t_k) - hY'(t_k) + \frac{1}{2}h^2Y''(t_k) - \frac{1}{6}h^3Y'''(\xi_-)$$

□

$$Y(t_k + h) - Y(t_k - h) = 2hY'(t_k) + O(h^3)$$

⇓

$$Y(t_k + h) = Y(t_k - h) + 2h \underbrace{Y'(t_k)}_{F[t_k, Y(t_k)]} + \underbrace{O(h^3)}_{e(h)}$$

Zastosujmy schemat skokowy do następującego zagadnienia początkowego

$$\begin{cases} y'(t) = -2y(t) + 1 \\ y(0) = 1 \end{cases}$$

Rozwiązanie dokładne (sprawdzić!) ma postać

$$Y(t) = \frac{1}{2}(e^{-2t} + 1)$$

Otrzymujemy wzór rekurencyjny drugiego rzędu postaci

$$y_{k+1} = y_{k-1} + 2h(-2y_k + 1) = -4hy_k + y_{k-1} + 2h$$

Wyprowadzimy formułę ogólną ciągu $\{y_k, k = 0, 1, 2, \dots\}$. W tym celu musimy rozwiązać **równanie różnicowe** (nie różniczkowe!) 2-ego rzędu postaci

$$y_{k+1} + 4hy_k - y_{k-1} = 2h$$

Podobnie jak równanie różniczkowe, równanie różnicowe 2-ego rzędu potrzebuje dwóch warunków początkowych – trzeba określić wartość y_0 i y_1 . Wykorzystamy w tym celu znane rozwiązanie analityczne i przyjmiemy, że

$$y_0 = 1, y_1 = \frac{1}{2}(e^{-2h} + 1)$$

Analogicznie jak w przypadku liniowego równania różniczkowego, rozwiązanie zagadnienia różnicowego jest sumą rozwiązania ogólnego równania jednorodnego i rozwiązania szczególnego równania pełnego. Równanie jednorodne ma postać

$$z_{k+1} + 4hz_k - z_{k-1} = 0, \quad k = 1, 2, \dots$$

Przewidujemy rozwiązanie równania jednorodnego w postaci potęgowej

$$z_k = s^k$$

gdzie liczba rzeczywista s podlega wyznaczeniu. Po podstawieniu do równania różnicowego i podzieleniu otrzymanej równości przez s^{k-1} (jest to wykonalne bo $k > 1$) otrzymujemy tzw. równanie charakterystyczne

$$s^2 + 4hs - 1 = 0$$

które ma dwa pierwiastki (rzeczywiste i różne)

$$s_1 = -2h + \sqrt{1 + 4h^2} \quad , \quad s_2 = -2h - \sqrt{1 + 4h^2}$$

Rozwiązaniem ogólnym równania jednorodnego jest ciąg liczbowy postaci

$$z_k = C_1 s_1^k + C_2 s_2^k$$

przy czym stałe C_1 i C_2 są (na razie) dowolne.

Rozwiązanie szczególnym równania niejednorodnego (czyli pełnego) jest poszukiwane w formie ciągu stałego $q_k \equiv M$. Po podstawieniu otrzymujemy natychmiast $M = \frac{1}{2}$.

Ostatecznie zatem rozwiązanie ogólne równania niejednorodnego ma postać

$$y_k = C_1 s_1^k + C_2 s_2^k + \frac{1}{2}$$

Stałe C_1 i C_2 dobieramy tak, aby otrzymać założone wartości y_0 i y_1 . Nie musimy wyznaczać ich explicite, dla naszych potrzeb wystarczy stała C_2 nie jest (na ogół) równa zeru.

Aby zrozumieć w czym tkwi problem wystarczy zauważyć, że przy dowolnym (i oczywiście dodatnim) kroku całkowania h pierwiastki s_1 and s_2 spełniają nierówności $|s_1| < 1$, $|s_2| > 1$.

Jeśli zatem $C_2 \neq 0$ to drugi składnik rozwiązania rośnie nieograniczenie wraz ze wzrostem wykładnika k (czyli „w czasie”) i w konsekwencji $\lim_{k \rightarrow \infty} |y_k| = \infty$.

Takie zachowanie przybliżonego rozwiązania jest kompletnie błędne – rozwiązanie dokładne jest dla dużych czasów ograniczone i dąży do wartości $\frac{1}{2}$. Innymi słowy, **schemat skokowy** - chociaż zgodny – **jest bezwarunkowo** (tj. dla dowolnie małego kroku całkowania) **niestabilny** i tym samym bezużyteczny.

Podsumowując: każdy użyteczna metoda całkowania powinna być nie tylko zgodna, ale również stateczna, przynajmniej pod warunkiem, że krok całkowania jest dostatecznie mały. Tylko metoda zgodna i stabilna zarazem może być zbieżna. Dla szerokiej klasy metod prawdziwe jest również stwierdzenie odwrotne. **Krótko:**

Zgodność + Stabilność = Zbieżność

Istnieją metody całkowania, które są stabilne bez względu na wielkość kroku całkowania (nazywamy je **bezw warunkowo stabilnymi**). Przykładem takiej metody jest **niejawna metoda Eulera**

$$y_{k+1} = y_k + h \underbrace{F[t_{k+1}, y_{k+1}]}_{F_{k+1}} = y_{k+1} + hF_{k+1}$$

Podobnie jak zwykła (tj. jawna) metoda Eulera, metoda niejawna jest 1-ego rzędu.

Ćwiczenie: pokazać, że niejawna metoda Eulera zastosowana do zagadnienia testowego $y'(t) = \lambda y(t)$, $y(0) = Y_0$ generuje ciąg wartości $\{y_0, y_1, y_2, \dots\}$ taki, że ciąg $\{|y_0|, |y_1|, \dots, |y_k|, \dots\}$ zbiega monotonicznie do zera, bez względu na wielkość kroku całkowania h .

Bezwarunkowo zbieżna jest również **metoda Cranka-Nicholsona** (znana też jako **metoda Adamsa-Moultona**). Formuła tej metody przedstawia się następująco

$$y_{k+1} = y_k + \frac{1}{2}h[F_k + F_{k+1}]$$

Metoda C-N jest 2-ego rzędu. Zauważmy, że jej formuła może być otrzymana jako „średnia arytmetyczna” jawnej i niejawnej metody Eulera.

Ćwiczenie:

- Dowieść, że metoda C-N jest 2-ego rzędu.
- Wykazać (posługując się zagadnieniem testowym), że metoda C-N jest bezwarunkowo stabilna, ale jeśli krok całkowania jest zbyt wielki może generować sztuczne (tj. nie występujące w ścisłym rozwiązaniu) oscylacje (należy rozważyć przypadek rzeczywistej i ujemnej wartości parametru λ).

Konstrukcja metod całkowania wyższych rzędów

Omówimy krótko jedną z metod konstruowania jawnych metod całkowania wyższych rzędów. Naturalne podejście polega na wykorzystaniu rozwinięcia w szereg Taylora. Pokażemy jak działa ta metoda na przykładzie schematu 2-ego rzędu.

Rozważmy formułę Taylora zakładając, że rozwiązanie zagadnienia początkowego $Y = Y(t)$ ma dostateczną regularność (przynajmniej trzy pierwsze pochodne ciągłe)

$$Y(t+h) = Y(t) + hY'(t) + \frac{1}{2}h^2Y''(t) + O(h^3)$$

Występujące w powyższej formule pochodne można wyrazić za pomocą funkcji prawej strony i jej pochodnych, a mianowicie

$$Y'(t) = F[t, Y(t)]$$

$$Y''(t) = \partial_t F[t, Y(t)] + \partial_y F[t, Y(t)] \cdot Y'(t) = \partial_t F[t, Y(t)] + \partial_y F[t, Y(t)] \cdot F[t, Y(t)]$$

Po podstawieniu i odrzuceniu członów $O(h^3)$ otrzymujemy **metodę całkowania 2-ego rzędu**

$$y(t_k + h) = y(t_k) + hF(t_k, y_k) + \frac{1}{2}h^2(\partial_t F + \partial_y F \cdot F)(t_k, y_k)$$

Opisane podejście można uogólnić dla metod dowolnie wysokiego rzędu. Jednak, takie podejście do konstruowania metod całkowania nie jest praktyczne, bowiem wymaga jawnego obliczenia pochodnych cząstkowych wysokiego rzędu dla funkcji F , co prowadzi do bardzo złożonych i żmudnych w wyznaczaniu wzorów. Może też być i tak, że jawna postać funkcji F jest tak skomplikowana, że analityczne obliczenie formy jej pochodnych cząstkowych jest praktycznie niewykonalne.

Jak zatem skonstruować metodę wyższego rzędu dokładności bez obliczania pochodnych funkcji prawej strony F ?

Jednym z możliwych podejść prowadzi do ważnej rodziny metod znanych jako **metody Rungego-Kutty**. Pokażemy jak przebiega wyprowadzenie metody R-K 2-ego rzędu.

Poszukujemy jawnej metody całkowania o następującej formie

$$\begin{cases} y(t_k + h) = y(t_k) + w_1 K_1 + w_2 K_2 \\ K_1 = hF(t_k, y_k) \\ K_2 = hF[t_k + \alpha h, y_k + \beta K_1] \end{cases}$$

Musimy wyznaczyć wartości parametrów w_1 , w_2 , α i β w taki sposób, aby uzyskać 2-gi rząd dokładności. W tym celu przekształcimy wyjściową postać metody R-K w taki sposób, aby można ją było porównać z generyczną metodą 2-ego rzędu wyprowadzoną wcześniej z rozwinięcia w szereg Taylora.

Rachunek przebiega następująco

$$\begin{aligned} Y(t_k + h) &= Y(t_k) + w_1 h F[t_k, Y(t_k)] + w_2 h F\{t_k + \alpha h, Y(t_k) + \beta h F[t_k, Y(t_k)]\} = \\ &= Y(t_k) + w_1 h F[t_k, Y(t_k)] + w_2 h \{F[t_k, Y(t_k)] + \partial_t F[t_k, Y(t_k)] \cdot \alpha h + \\ &+ \partial_y F[t_k, Y(t_k)] \cdot \beta h F[t_k, Y(t_k)] + O(h^2)\} \end{aligned}$$

Po prostych przekształceniach mamy

$$Y(t_k + h) = Y(t_k) + (w_1 + w_2) h F[t_k, Y(t_k)] + h^2 (w_2 \alpha \partial_t F + w_2 \beta \partial_y F \cdot F)[t_k, Y(t_k)] + O(h^3)$$

Otrzymana formuła jest równoważna uzyskanej wcześniej jeśli tylko parametry metody R-K spełniają warunki

$$w_1 + w_2 = 1, \quad \alpha = \beta, \quad \alpha w_2 = \frac{1}{2}$$

Mamy 3 warunki (nieliniowe) dla 4-ech parametrów – rozwiązanie zapewne nie jest jedyne. W istocie – teoretycznie mamy nieskończenie wiele metod R-K 2-ego rzędu. W praktyce wykorzystywane są (bywają) dwa warianty szczególne:

- metoda Heuna (zwana też po prostu RK2): $w_1 = w_2 = \frac{1}{2}$, $\alpha = \beta = 1$

$$\begin{cases} y(t_k + h) = y(t_k) + \frac{1}{2}(K_1 + K_2) \\ K_1 = hF(t_k, y_k) \\ K_2 = hF[t_k + h, y_k + K_1] \end{cases}$$

- metoda Cauchy’ego-Eulera: $w_1 = 0$, $w_2 = 1$, $\alpha = \beta = \frac{1}{2}$

$$\begin{cases} y(t_k + h) = y(t_k) + K_2 \\ K_1 = hF(t_k, y_k) \\ K_2 = hF[t_k + \frac{1}{2}h, y_k + \frac{1}{2}K_1] \end{cases}$$

Metody wyższego rzędu otrzymywane są w analogiczny sposób. Potrzebne jest wprowadzenie większej liczby wielkości oznaczanych wielką literą K i większej liczby parametrów. Wyznaczanie ich wartości robi się skomplikowane (warunki mają formę nieliniowych równań algebraicznych), ale jest wykonalne.

Wyznaczenie postaci metody 4-ego rzędu wymaga wprowadzenia czterech wielkości typu K i 13-stu swobodnych parametrów. Liczba warunków do spełnienia wynosi „tylko” 11, zatem ponownie metod 4-ego rzędu jest wiele.

Najbardziej popularny wariant oznaczany po prostu jako **metoda RK4** wygląda następująco

$$y(t_k + h) = y(t_k) + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

$$K_1 = hF(t_k, y_k)$$

$$K_2 = hF[t_k + \frac{1}{2}h, y_k + \frac{1}{2}K_1]$$

$$K_3 = hF[t_k + \frac{1}{2}h, y_k + \frac{1}{2}K_2]$$

$$K_4 = hF[t_k + h, y_k + K_3]$$

W każdym kroku czasowym metody RK4 funkcja prawej strony F jest obliczana 4 razy (w RK2 – dwa razy, w metodzie Eulera (de facto RK1) – 1 raz). Należy podkreślić, że w metodach RK rzędy wyższego niż 4 liczba odwołań do funkcji prawej strony jest większa niż rząd. Np. metody 5-ego rzędu „wywołują” funkcję F co najmniej 6-cio krotnie, a rzędu 6-ego – co najmniej 8-krotnie w każdym kroku czasowym. Z tego powodu metody RK rzędy wyższego niż 5 nie są w praktyce używane.

Uogólnienie na układy równań różniczkowych zwyczajnych

Omówione wcześniej schematy całkowania można uogólnić na układy równań różniczkowych zwyczajnych, o ile mają one postać standardową, a mianowicie

$$\frac{d}{dt} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f_1(t, y_1(t), y_2(t), \dots, y_n(t)) \\ f_2(t, y_1(t), y_2(t), \dots, y_n(t)) \\ \vdots \\ f_n(t, y_1(t), y_2(t), \dots, y_n(t)) \end{bmatrix}, \quad \begin{bmatrix} y_1(t_0) \\ y_2(t_0) \\ \vdots \\ y_n(t_0) \end{bmatrix} = \begin{bmatrix} y_{10} \\ y_{20} \\ \vdots \\ y_{n0} \end{bmatrix} \text{ - warunek początkowy}$$

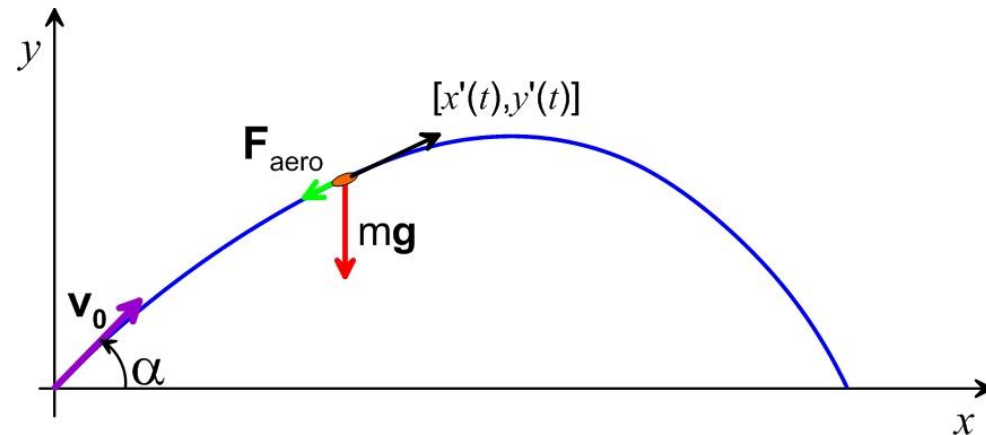
W zwartym zapisie macierzowo-wektorowym mamy

$$\mathbf{y}'(t) = \mathbf{F}(t, \mathbf{y}(t)) \quad , \quad \mathbf{y}(t_0) = \mathbf{y}_0$$

Układy RRZ otrzymywane z zastosowaniami naukowych i technicznych na ogół nie mają postaci standardowej. Przed zastosowaniem tej czy innej metody numerycznego całkowania należy wówczas sprowadzić dany układ do (równoważnej) postaci układu równań 1-ego rzędu.

Procedurę sprowadzenia do takiej postaci pokażemy na przykładzie. Rozważmy układ równań różniczkowych opisujących lot pocisku z uwzględnieniem oporu powietrza. Zgodnie z 2-gą zasadą dynamiki mamy

$$\begin{cases} mx'' = -cx'\sqrt{x'^2 + y'^2} \\ my'' = -mg - cy'\sqrt{x'^2 + y'^2} \end{cases}$$



W powyższym układzie symbolem m oznaczyliśmy masę pocisku (stałą), c to współczynnik oporu aerodynamicznego, g jest przyspieszeniem grawitacyjnym. Położenie pocisku określone jest za pomocą dwóch współrzędnych: poziomej x i pionowej y . warunki początkowe dla lotu balistycznego mogą mieć postać

$$x(0) = 0, \quad y(0) = 0$$

$$x'(0) = v_0 \cos(\alpha), \quad y'(0) = v_0 \sin(\alpha)$$

gdzie v_0 jest prędkością początkową, zaś α oznacza kąt nachylenia trajektorii pocisku względem kierunku poziomego w chwili wystrzału.

Układ równań opisujących lot pocisku nie jest w postaci standardowej – równania są 2-ego rzędu, a nie 1-ego. W celu sprowadzenia go do postaci standardowej wprowadzamy następujące funkcje

$$z_1(t) = x(t), z_2(t) = x'(t), z_3(t) = y(t), z_4(t) = y'(t)$$

Następujący układ równań różniczkowych 1-ego rzędu jest równoważny oryginalnemu i jest jednocześnie w formie standardowej

$$\begin{cases} z_1' = z_2 \\ z_2' = -k z_2 \sqrt{z_2^2 + z_4^2} \\ z_3' = z_4 \\ z_4' = -g - k z_4 \sqrt{z_2^2 + z_4^2} \end{cases}$$

gdzie oznaczyliśmy $k = c / m$. W postaci macierzowo-wektorowej mamy

$$\mathbf{z}'(t) = \mathbf{F}[t, \mathbf{z}(t)]$$

Wektorowa funkcja prawych stron ma w tym przypadku następującą postać

$$\mathbf{F}[t, \mathbf{z}(t)] \equiv \begin{cases} f_1(t, z_1, \dots, z_4) = z_2 \\ f_2(t, z_1, \dots, z_4) = -k z_2 \sqrt{z_2^2 + z_4^2} \\ f_3(t, z_1, \dots, z_4) = z_4 \\ f_4(t, z_1, \dots, z_4) = -k z_4 \sqrt{z_2^2 + z_4^2} \end{cases}$$

Wreszcie, warunki początkowe „przetłumaczone” na nowe niewiadome mają formę

$$z_1(0) = 0 \quad , \quad z_2(0) = v_0 \cos(\alpha) \quad , \quad z_3(0) = 0 \quad , \quad z_4(0) = v_0 \sin(\alpha)$$

Poznane wcześniej metody mogą być zapisane w formie macierzowo-wektorowej, tj. dostosowanej do układu równań różniczkowych w postaci standardowej. Mamy w szczególności:

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{F}[t_k, \mathbf{y}_k]$$

Metoda Euler (rz = 1)

$$\begin{cases} \mathbf{y}_{k+1} = \mathbf{y}_k + \frac{1}{2}(\mathbf{k}_1 + \mathbf{k}_2) \\ \mathbf{k}_1 = h\mathbf{F}[t_k, \mathbf{y}_k] \\ \mathbf{k}_2 = h\mathbf{F}[t_k + h, \mathbf{y}_k + \mathbf{k}_1] \end{cases}$$

Metoda RK2 (Heuna, rz = 2)

$$\begin{cases} \mathbf{y}_{k+1} = \mathbf{y}_k + \frac{1}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4) \\ \mathbf{k}_1 = h\mathbf{F}[t_k, \mathbf{y}_k] \\ \mathbf{k}_2 = h\mathbf{F}[t_k + \frac{1}{2}h, \mathbf{y}_k + \frac{1}{2}\mathbf{k}_1] \\ \mathbf{k}_3 = h\mathbf{F}[t_k + \frac{1}{2}h, \mathbf{y}_k + \frac{1}{2}\mathbf{k}_2] \\ \mathbf{k}_4 = h\mathbf{F}[t_k + h, \mathbf{y}_k + \mathbf{k}_3] \end{cases}$$

Metoda RK4 (rz = 4)

Uwagi końcowe

- W praktyce metody R-K w formie „surowej” są rzadko stosowane. Współcześnie używa się ich z reguły w połączeniu z pewną strategią automatycznego doboru kroku całkowania wykorzystującą szacowanie błędu. Prosta strategia bazuje na porównaniu rozwiązania przybliżonego otrzymanego w wyniku wykonania pojedynczego kroku o długości h z rozwiązaniem otrzymanym w wyniku wykonania dwóch kroków o długości $h/2$. Bardziej wyrafinowane strategie polegają na porównaniu rozwiązań otrzymanym dwiema różnymi, ale specjalnie dobranymi metodami R-K (różniącymi się rzędem dokładności o 1). Przykładem takiej metody jest popularna metoda Rungego-Kutty-Fehlberga, oznaczana często jako RKF45 (wykorzystywana jest w niej kombinacja metod 4-ego i 5-ego rzędu). Innym przykładem odpowiednio skonstruowanej pary metod RK rzędu 4 i 5 jest para Dormanda-Prince’a. Metoda ta jest zaimplementowana m.in. w pakiecie MATLAB (jako funkcja *ode45*).
- Alternatywnym sposobem konstrukcji metod całkowania wysokiego rzędu jest wykorzystanie historii rozwiązania tj. informacji o nim z kilku poprzednich kroków. Otrzymywane w ten sposób schematy całkowania nazywane są metodami wielokrokowe. Popularną rodziną takich metod są **metody Adamsa** (jawne i niejawne), a także wywodzące się z nich **metody typu predyktor-korektor**. Zaletą metod wielokrokowych jest to, że w każdym kroku całkowania funkcja prawych stron (czyli pochodne) są obliczane tylko raz. Z drugiej strony – **metody wielokrokowe nie są samostartujące**, problematyczne jest również wykorzystanie w nich dynamicznej adaptacji kroku całkowania.