

Komputerowa analiza zagadnień różniczkowych

10. Równania różniczkowe z niezmiennikami

Równania algebraiczno-różniczkowe — zarys
problematyki

P. F. Góra

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

2013

I. Równania różniczkowe z niezmiennikami

Przypuśćmy, iż dany jest problem Cauchy'ego ($y, f, y_0 \in \mathbb{R}^n$)

$$\begin{cases} \frac{dy}{dx} = f(x, y) \\ y(x_0) = y_0 \end{cases} \quad (1)$$

oraz pewna funkcja $g : \mathbb{R}^n \rightarrow \mathbb{R}^m, m < n$, o tej własności, że

$$g(y(x)) = \text{const}, \quad (2)$$

gdzie $y(x)$ jest rozwiązaniem problemu (1). Funkcję g nazywamy wówczas *całką* lub *niezmiennikiem ruchu*, jako że jej wartość jest zachowana „wzdłuż” *rozwiązań równania (1)*. Ściślej rzecz biorąc, każda składowa g jest całką ruchu — w tym sensie (2) definiuje m całek ruchu. (Różne warunki początkowe mogą prowadzić do różnych wartości liczbowych całek ruchu.)

W fizyce całki ruchu typowo związane są z zasadami zachowania: energii, momentu pędu itd.

Jak pokazuje się w ramach standardowego kursu analizy matematycznej, każda nietrywialna *i ciągła* całka ruchu redukuje wymiarowość problemu: jedno (skalarne) równanie postaci (2) pozwala wyrazić pewną składową y przez pozostałe składowe *oraz* przez wartości całek ruchu. Ostatecznie pozostaje więc do rozwiązania problem $n - m$ wymiarowy.

Przykład

Rozważmy układ dwu identycznych, sprzężonych cząstek, opisanych lagranżianem

$$\mathcal{L} = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) - \left(V(x_1) + V(x_2) + \frac{1}{2}k(x_1 - x_2)^2 \right). \quad (3)$$

Prowadzi on do następujących równań ruchu:

$$\dot{x}_1 = v_1, \quad (4a)$$

$$\dot{x}_2 = v_2, \quad (4b)$$

$$\dot{v}_1 = - \left. \frac{dV}{dx} \right|_{x=x_1} - k(x_1 - x_2), \quad (4c)$$

$$\dot{v}_2 = - \left. \frac{dV}{dx} \right|_{x=x_2} - k(x_2 - x_1). \quad (4d)$$

Układ (3) posiada całkę ruchu — energię — co pozwala na wyrażenie jednej zmiennej — powiedzmy x_2 — przez pozostałe zmienne i wartość całki ruchu:

$$x_2 = \mathcal{G}(x_1, v_1, v_2, E). \quad (5)$$

Ostatecznie dostajemy

$$\dot{x}_1 = v_1, \quad (6a)$$

$$\dot{v}_1 = - \left. \frac{dV}{dx} \right|_{x=x_1} - k(x_1 - \mathcal{G}(x_1, v_1, v_2, E)), \quad (6b)$$

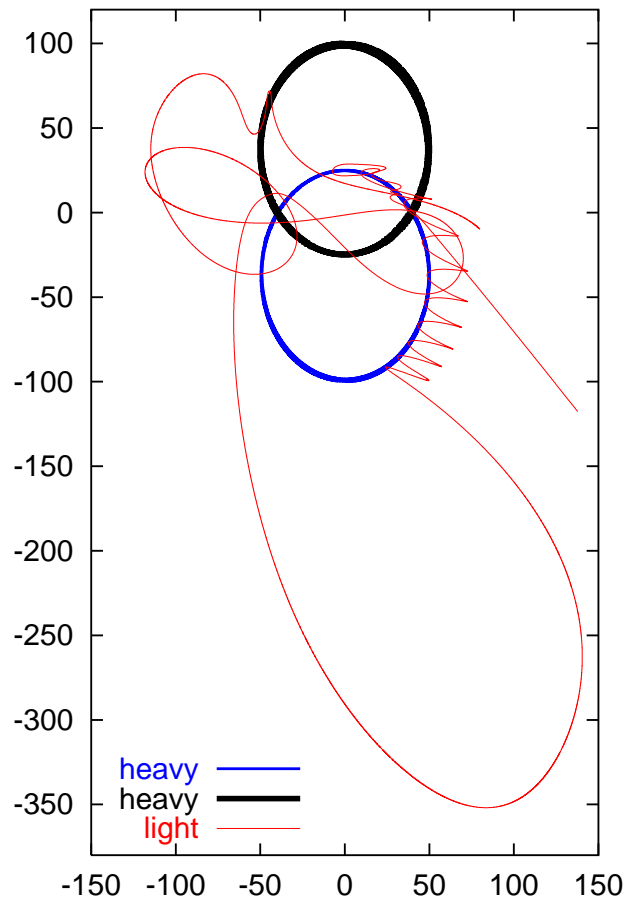
$$\dot{v}_2 = - \left. \frac{dV}{dx} \right|_{x=\mathcal{G}(x_1, v_1, v_2, E)} - k(\mathcal{G}(x_1, v_1, v_2, E) - x_1). \quad (6c)$$

Układ równań różniczkowych (6) jest tylko *pozornie* prostszy od układu równań (4). Zauważmy, że

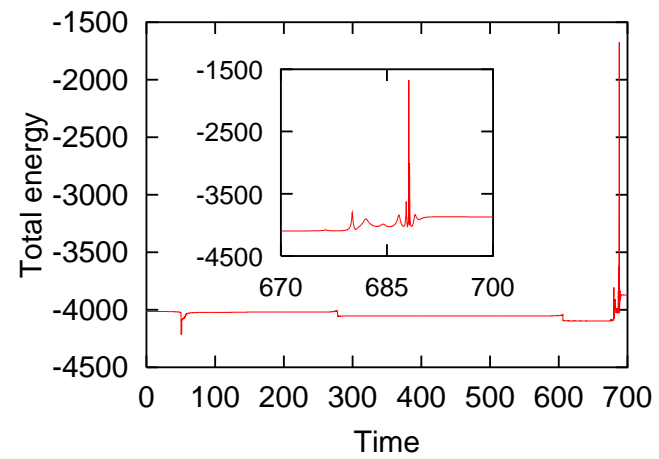
- ☹ Zniszczyliśmy symetrię pomiędzy x_1 a x_2 . Na skutek tego obie zmienne będą traktowane *numerycznie* inaczej, co może być istotne choćby ze względu na błąd zaokrąglenia.
- ☹ Jeżeli obliczanie (5) zawiera niejednoznaczne operacje algebraiczne (na przykład pierwiastkowanie), nie mamy gwarancji otrzymania prawidłowego wyniku.
- ☹ Równań do rozwiązania jest mniej, ale sumaryczny *czas obliczania* prawej strony (6) może być dłuższy (wyobraźmy sobie, że nie mamy jednej cząstki, ale $N \gg 1$; całka ruchu nadal pozwala na wyeliminowanie tylko jednego równania).
- ☹ Kod realizujący (6) będzie bardziej skomplikowany, a zatem trudniejszy do konserwacji i potencjalnie ułatwiający popełnienie błędu, niż kod realizujący (4).

Dlatego też w praktyce numerycznej na ogół (poza niewielką ilością prostych przypadków, na przykład gdy całka ruchu jest liniową funkcją zmiennych dynamicznych) nie dokonuje się eliminacji zmiennych korzystając z całek ruchu.

Całki ruchu są jednak bardzo ważne, zaś standardowe algorytmy *nie gwarantują*, że wartości tych całek faktycznie będą zachowane.



A three-body problem



Co zatem robić?

- Zastanowić się, czy przypadkiem samo całkowanie ODE nie wystarczy — warunek (2) nie będzie wówczas spełniony *ściśle*, ale *być może* odchylenia od równości są akceptowalne, to znaczy mniejsze od zadanej tolerancji. **Warunek ten nie jest spełniony w szczególności wtedy, gdy użyta metoda numeryczna powoduje, że $g(y)$ systematycznie oddala się od swojej analitycznej wartości.**
- Zamienić układ (1)–(2) na pewien system DAE (patrz niżej (8)) i rozwiązać metodami właściwymi dla DAE.
- Zastosować algorytm inteligentnie rzutujący numeryczne rozwiązanie ODE na hiperpowierzchnię zadaną przez niezmiennik.
- Jeżeli niezmienniki, o których mowa, pojawiają się jako rezultat zasad zachowania mechaniki klasycznej, należy użyć *algorytmów symplektycznych*, mimo iż, technicznie rzecz biorąc, nie gwarantują one zachowania niezmienników.

Zamiana ODE z niezmiennikiem na DAE

Zapiszmy (1)–(2) w postaci (zakładam, że prawa strona (1) nie zależy jawnie od zmiennej niezależnej)

$$\frac{dy}{dx} = f(y), \quad (7a)$$

$$g(y) = 0, \quad (7b)$$

co teraz można zamienić na DAE w postaci Hessenberga o indeksie 2:

$$\frac{dy}{dx} = f(y) - D(y)z, \quad (8a)$$

$$g(y) = 0, \quad (8b)$$

gdzie D jest pewną macierzą. **Ścisłe rozwiązanie wyjściowego ODE z niezmiennikiem (7) odpowiada rozwiązaniu DAE (8) spełniającemu $z = 0$.**

Konstrukcja macierzy D

Jak dobrać macierz D ? Sensownie ☺. Niech

$$\mathbf{G} = \nabla_{\mathbf{y}} \mathbf{g}, \quad (9)$$

to znaczy $G_{ij} = \partial g_i / \partial y_j$, $\mathbf{G} \in \mathbb{R}^{m \times n}$. Najczęściej przyjmuje się

$$\mathbf{D}(\mathbf{y}) = \mathbf{G}^T \in \mathbb{R}^{n \times m}. \quad (10)$$

Układ algebraiczno-różniczkowy (8) na ogół *można* teraz rozwiązać, ale bywa to uważane za “*overkill*”, to znaczy, że używa się zbyt wymagającej (i kosztownej!) metody dla osiągnięcia zadanego celu.

Algorytmy rzutu

Niech $\mathcal{M} = \{\mathbf{y} : \mathbf{g}(\mathbf{y}) = 0\}$. Równanie

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(\mathbf{y}) - \gamma \mathbf{F}(\mathbf{y}) \mathbf{g}(\mathbf{y}), \quad (11)$$

gdzie \mathbf{F} jest pewną *inną* macierzą, ma to samo rozwiązanie na \mathcal{M} , co (7a). Co więcej, jeżeli macierz \mathbf{GF} jest dodatnio określona, γ zaś jest dostatecznie duże, rozwiązanie będzie “ściągane” na \mathcal{M} .

Aby to zobaczyć, po pierwsze zauważmy, że $\frac{1}{2}g^T g \geq 0$, przy czym równość oznacza, że warunek (7b) jest spełniony. $g^T g$ jest liczbową miarą odstępstwa od tego warunku; jeżeli $g^T g$ *maleje*, oznacza to, że układ zdąża w stronę \mathcal{M} .

Po drugie, załóżmy, że $\mathbf{G}\mathbf{F}$ jest dodatnio określona, ma zatem najmniejszą dodatnią wartość własną. Oznaczmy ją λ_0 .

Po trzecie, załóżmy, że dla \mathbf{y} leżących dostatecznie blisko \mathcal{M}

$$\exists \gamma_0 \geq 0 : \|\mathbf{G}\mathbf{f}(\mathbf{y})\| \leq \gamma_0 \|\mathbf{g}(\mathbf{y})\|. \quad (12)$$

Jeżeli spełnione jest równanie (11), zachodzi

$$\begin{aligned} \frac{1}{2} \frac{d}{dx} (\mathbf{g}^T \mathbf{g}) &= \mathbf{g}^T \frac{d\mathbf{g}(\mathbf{y})}{dx} = \mathbf{g}^T \mathbf{G} \frac{d\mathbf{y}}{dx} = \mathbf{g}^T \mathbf{G} (\mathbf{f} - \gamma \mathbf{F} \mathbf{g}) \\ &= \mathbf{g}^T \mathbf{G} \mathbf{f} - \gamma \mathbf{g}^T \mathbf{G} \mathbf{F} \mathbf{g} \leq (\gamma_0 - \gamma \lambda_0) \mathbf{g}^T \mathbf{g}. \end{aligned} \quad (13)$$

Jeśli więc $\gamma > \gamma_0/\lambda_0$, $d(\mathbf{g}^T \mathbf{g})/dx \leq 0$, a więc rozwiązanie równania (11) jest ściągane na \mathcal{M} .

Jak dobrać \mathbf{F} ? Sensownie ☺. Na przykład

$$\mathbf{F}(\mathbf{y}) = \mathbf{D} (\mathbf{G} \mathbf{D})^{-1}, \quad (14)$$

gdzie \mathbf{D} jest dane przez (10). Wówczas $\mathbf{G} \mathbf{F} = \mathbb{I}$ i $\lambda_0 = 1$.

Przykład

Założmy, że badamy układ, który ma jedną całkę ruchu, energię (lub “energię”) $E(\mathbf{y})$. Wówczas

- $\mathbf{G} = \nabla_{\mathbf{y}} E(\mathbf{y}) = [\partial E / \partial y_1, \partial E / \partial y_2, \dots, \partial E / \partial y_n] \in \mathbb{R}^{1 \times n}$.
- $\mathbf{D} = \mathbf{G}^T = [\partial E / \partial y_1, \partial E / \partial y_2, \dots, \partial E / \partial y_n]^T \in \mathbb{R}^n$ — to jest wektor!
- $\mathbf{GD} = \sum_{i=1}^n (\partial E / \partial y_i)^2$ — to jest liczba.
- Ostatecznie numerycznie rozwiązujemy

$$\frac{d\mathbf{y}}{dx} = \mathbf{f}(\mathbf{y}) - \gamma \frac{E(\mathbf{y}) - E(\mathbf{y}_0)}{\sum_{i=1}^n (\partial E(\mathbf{y}) / \partial y_i)^2} \mathbf{D}, \quad (15)$$

gdzie $\gamma > \gamma_0$ (γ_0 zdefiniowane jest przez (12)) oraz $\mathbf{y}(0) = \mathbf{y}_0$.

II. Układy algebraiczno–różniczkowe

Spróbujmy rozwiązać równanie płaskiego wahadła matematycznego we współrzędnych kartezjańskich. Lagranżian wydaje się mieć postać

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy, \quad (16)$$

z czego wynikają **zupełnie złe** równania ruchu:

$$\ddot{x} = 0, \quad (17a)$$

$$\ddot{y} = -g. \quad (17b)$$

Równania (17) są złe, gdyż nie uwzględniliśmy warunku $x^2 + y^2 = r^2$. Musimy wprowadzić mnożnik Lagrange'a:

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy - \lambda(x^2 + y^2 - r^2), \quad (18)$$

z czego wynika ($z = 2\lambda/m$)

$$\ddot{x} = -zx, \quad (19a)$$

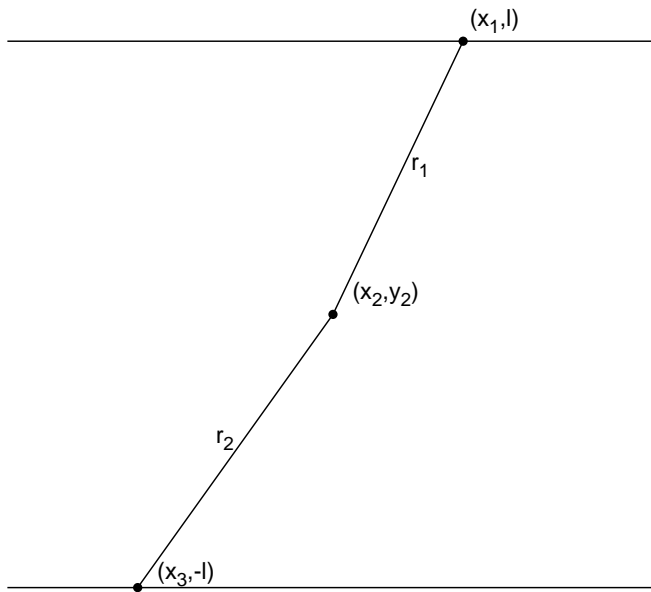
$$\ddot{y} = -g - zy, \quad (19b)$$

$$x^2 + y^2 = r^2. \quad (19c)$$

Człony proporcjonalne do z opisują *sily reakcji więzów*. Równania (19) **nie** stanowią układu ODE: Po zamienieniu na równania pierwszego rzędu, (19) stanowi układ z pięcioma zmiennymi $(x, \dot{x}, y, \dot{y}, z)$, ale tylko cztery równania są różniczkowe — jedno jest równaniem algebraicznym. Układ tego typu nazywa się układem *algebraiczno-różniczkowym*.

Przez podstawienie $x = r \cos \theta$, $y = r \sin \theta$ układ (19) można łatwo sprowadzić do znanego ODE, ale w ogólnym przypadku nie jest to ani łatwe, ani wygodne.

Przykład: układ mechaniczny z więzami



Rozpatrzmy układ składający się z trzech punktów. Dwa z nich mogą poruszać się bez tarcia po poziomych liniach odległych o $2l$. Punkty te połączone są z punktem środkowym za pomocą sztywnych, nieważkich prętów o długościach r_1 , r_2 . Połączenia są przegubowe i pręty mogą się w połączeniach swobodnie obracać.

Lagranżian (wraz z czynnikami Lagrange'a) ma postać

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m_1\dot{x}_1^2 + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2) + \frac{1}{2}m_3\dot{x}_3^2 \\ &\quad - z_1 \left((x_1 - x_2)^2 + (y_2 - l)^2 - r_1^2 \right) \\ &\quad - z_2 \left((x_2 - x_3)^2 + (y_2 + l)^2 - r_2^2 \right)\end{aligned}\tag{20}$$

i prowadzi do równań

$$m_1\ddot{x}_1 + 2z_1(x_1 - x_2) = F_1,\tag{21a}$$

$$m_2\ddot{x}_2 + 2z_1(x_2 - x_1) + 2z_2(x_2 - x_3) = 0,\tag{21b}$$

$$m_2\ddot{y}_2 + 2z_1(y_2 - l) + 2z_2(y_2 + l) = 0,\tag{21c}$$

$$m_3\ddot{x}_3 + 2z_2(x_3 - x_2) = F_3,\tag{21d}$$

$$(x_1 - x_2)^2 + (y_2 - l)^2 = r_1^2,\tag{21e}$$

$$(x_3 - x_2)^2 + (y_2 + l)^2 = r_2^2,\tag{21f}$$

gdzie uwzględniliśmy też poziome **siły zewnętrzne** działające na układ.

Układu (21) nie da się *łatwo*, przez proste podstawienie, zredukować do układu ODE.

Skomplikowane układy mechaniczne z więzami stanowią typowe źródło problemów modelowanych przez DAE. Innym klasycznym źródłem DAE są obwody elektryczne — rolę więzów algebraicznych spełniają prawa Kirchhoffa.

Jeszcze innym źródłem DAE może być dyskretyzacja niektórych równań różniczkowych cząstkowych — na przykład równań Naviera-Stokesa dla cieczy *nieściśliwej*.

DAE — formalna definicja

Rozważmy równanie

$$\mathbf{F}(t, \mathbf{y}, \mathbf{y}') = 0. \quad (22)$$

gdzie $\mathbf{F}, \mathbf{y} \in \mathbb{R}^n$. Jeżeli $\forall t : \partial \mathbf{F} / \partial \mathbf{y}' \neq 0$, równanie (22) jest uwikłaną formą układu ODE. Jeżeli $\partial \mathbf{F} / \partial \mathbf{y}' = 0$ (choćby lokalnie), z (22) nie da się wyznaczyć tylu równań różniczkowych, ile jest zmiennych: (22) jest wówczas DAE. Postać ta zwana jest *w pełni uwikłanym DAE*.

Postać szczególna

Rozważmy równanie

$$\mathbf{M}\mathbf{y}' = \mathbf{f}(t, \mathbf{y}), \quad (23)$$

gdzie $\mathbf{y}, \mathbf{f} \in \mathbb{R}^n$ oraz $\mathbf{M} \in \mathbb{R}^{n \times n}$ jest macierzą osobliwą, $\det \mathbf{M} = 0$. Postać (23) jest szczególną postacią (22); można też uważać, że (23) powstaje z (22) w trakcie próby rozwikłania ze względu na \mathbf{y}' (\mathbf{M} jest wówczas jacobianem \mathbf{F} po ostatnim argumencie).

Poza przypadkiem trywialnym, istnieją nieosobliwe macierze \mathbf{P}, \mathbf{Q} , takie, że

$$\mathbf{P}\mathbf{M}\mathbf{Q} = \begin{bmatrix} \mathbb{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{N} \end{bmatrix} \quad (24)$$

gdzie \mathbb{I}_r jest macierzą jednostkową $r \times r$, $r < n$, przy czym istnieje liczba naturalna μ taka, że $\mathbf{N}^\mu = \mathbf{0}$. Równanie (23) przybiera wówczas postać

$$\tilde{\mathbf{y}}' = \mathbf{f}_r(t, \tilde{\mathbf{y}}, \mathbf{z}), \quad (25a)$$

$$\mathbf{Nz}' = \mathbf{f}_a(t, \tilde{\mathbf{y}}, \mathbf{z}), \quad (25b)$$

gdzie

$$\begin{bmatrix} \mathbf{f}_r \\ \mathbf{f}_a \end{bmatrix} = \mathbf{P}\mathbf{f}, \quad \begin{bmatrix} \tilde{\mathbf{y}} \\ \mathbf{z} \end{bmatrix} = \mathbf{Q}^{-1}\mathbf{y}. \quad (26)$$

$\tilde{\mathbf{y}}, \mathbf{f}_r \in \mathbb{R}^{r \times r}$, $\mathbf{z}, \mathbf{f}_a \in \mathbb{R}^{n-r \times n-r}$.

Ile wynosi $\det \mathbf{N}$? $0 = \det \mathbf{N}^\mu = (\det \mathbf{N})^\mu$, a zatem $\det \mathbf{N} = 0$ oraz $\det \mathbf{M} = \det \mathbf{P}^{-1} \cdot 1 \cdot \det \mathbf{N} \cdot \det \mathbf{Q}^{-1} = 0$. Liczbę μ nazywamy **indeksem układu DAE (23)**. Indeks określa ilość niezależnych różniczkowań potrzebnych do zamienienia DAE na ODE (w postaci $\mathbf{x}' = \mathbf{f}(t, \mathbf{x})$). W ogólności indeks może zależeć od rozwiązania DAE, od warunków początkowych, a nawet może zmieniać się wzdłuż rozwiązania. Na szczęście w wielu istotnych z praktycznego punktu widzenia przypadkach indeks jest stały i zależy tylko od formy DAE.

DAE — postaci Hessenberga

W postaciach (22) i (23) równania różniczkowe i algebraiczne mieszają się. W praktyce często występują układy, w których można wskazać *zmiennne różniczkowe* i *zmiennne algebraiczne*.

Postać Hessenberga o indeksie 1:

$$y' = f(t, y, z), \quad (27a)$$

$$0 = g(t, y, z) \quad (27b)$$

gdzie $\partial g / \partial z \neq 0$.

Postać Hessenberga o indeksie 2:

$$y' = f(t, y, z), \quad (28a)$$

$$0 = g(t, y) \quad (28b)$$

gdzie $(\partial g / \partial y) \cdot (\partial f / \partial z) \neq 0$.

Postać Hessenberga o indeksie 3:

$$\mathbf{x}' = \mathbf{f}(t, \mathbf{x}, \mathbf{y}, \mathbf{z}), \quad (29a)$$

$$\mathbf{y}' = \mathbf{g}(t, \mathbf{x}, \mathbf{y}), \quad (29b)$$

$$0 = \mathbf{h}(t, \mathbf{y}) \quad (29c)$$

gdzie $(\partial \mathbf{h} / \partial \mathbf{y}) \cdot (\partial \mathbf{g} / \partial \mathbf{x}) \cdot (\partial \mathbf{f} / \partial \mathbf{z}) \neq 0$.

Ta właśnie postać pojawia się w wyniku modelowania układów mechanicznych z więzami.

Każdą postać Hessenberga można oczywiście sprowadzić do postaci w pełni uwikłanej w odpowiednio większej ilości wymiarów.

DAE — problem warunków początkowych

Warunki początkowe dla DAE muszą być *konsyistentne* — niektóre warunki mogą prowadzić do nieistnienia rozwiązania. Dzieje się tak dlatego, że oprócz warunków podawanych jawnie, muszą być także spełnione pewne warunki ukryte.

Przykład: Rozważmy układ

$$x' = z, \quad (30a)$$

$$0 = x - t. \quad (30b)$$

Rozwiązaniem jest $x = t$, $z = 1$ i żadne warunki początkowe nie są potrzebne. Gdybyśmy chcieli zadać warunek $x(0) = c$, będzie on *konsyistentny* z problemem (30) wtedy i tylko wtedy, gdy $c = 0$.

DAE — związek z układami sztywnymi

Przepiszmy (27) w postaci

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}, \mathbf{z}), \quad (31a)$$

$$\varepsilon \mathbf{z}' = \mathbf{g}(t, \mathbf{y}, \mathbf{z}) \quad (31b)$$

przy czym rozpatrujemy granicę $\varepsilon \rightarrow 0$. Sugeruje to zastąpienie problemu DAE przez problem (31) z $|\varepsilon| \ll 1$, rozwiązanie go jak ODE dla szeregu takich ε i ekstrapolowanie wyników do $\varepsilon \rightarrow 0$.

DAE i metody niejawne — motywacja

Podejście opisane na poprzedniej stronie jest niepraktyczne, między innymi dlatego, że prowadzi do rozwiązywania **skrajnie** sztywnych ODE, sugeruje jednak zastosowanie metod właściwych dla układów sztywnych, a więc metod niejawnych.

Inna motywacja za metodami niejawnymi bierze się z następującego rozumowania: Gdybyśmy używali metod jawnych, nie wiemy co zrobić z ograniczeniami algebraicznymi. Jeśli problem zadany jest w postaci Hessenberga, moglibyśmy myśleć o zastosowaniu metody jawnej do części różniczkowej, a następnie rozwiązaniu równania algebraicznego dla części algebraicznej, nie ma jednak wówczas żadnej gwarancji, że otrzymane rozwiązanie będzie poprawne. Cały problem trzeba sprowadzić do rozwiązania układu równań algebraicznych **jednocześnie** dla wszystkich zmiennych.

Przykład: Niejawna metoda Eulera

Rozważmy problem (27). Zastosowanie niejawnej metody Eulera prowadzi do

$$y_{n+1} = y_n + hf(t_{n+1}, y_{n+1}, z_{n+1}), \quad (32a)$$

$$0 = g(t_{n+1}, y_{n+1}, z_{n+1}). \quad (32b)$$

W ogólności, dla zapewnienia stabilności, DAE o indeksie μ należy rozwiązywać przy pomocy metod BDF o rzędzie $k > \mu - 1$.

DAE i niejawne metody RK

Do DAE w postaci w pełni uwikłanej

$$0 = \mathbf{f}(t, \mathbf{y}, \mathbf{y}') \quad (33)$$

możemy stosować niejawne metody RK:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h \sum_{i=1}^s w_i \mathbf{K}_i, \quad (34)$$

przy czym w każdym punkcie pośrednim rozwiązujemy

$$T_i = t_n + \alpha_i h, \quad (35a)$$

$$0 = \mathbf{f}(T_i, \mathbf{Y}_i, \mathbf{K}_i), \quad (35b)$$

$$\mathbf{Y}_i = \mathbf{y}_n + h \sum_{j=1}^s \beta_{ij} \mathbf{K}_j. \quad (35c)$$

Podobnie postępujemy dla DAE w postaci Hessenberga — ta postać pozwala nam na zmniejszenie rozmiaru układu *jednocześnie* rozwiązywanych równań algebraicznych.

Ostrzeżenie: Aby metoda RK nadawała się do DAE, musi spełniać pewne dodatkowe warunki.