

## Równania różniczkowe:

- rozwiązania w bazie funkcyjnej  
(porzucamy metodę różnic skończonych)

## Plan:

metoda kolokacji

metoda najmniejszych kwadratów

metoda Galerkina

formalizm reszt ważonych

metoda wariacyjna Reyleigha-Ritza

| do metody elementów  
skończonych

Przykład:  $\frac{d^2u}{dx^2} = -\sin(\pi x)$   $u(-1)=0$   
 $u(1)=0$

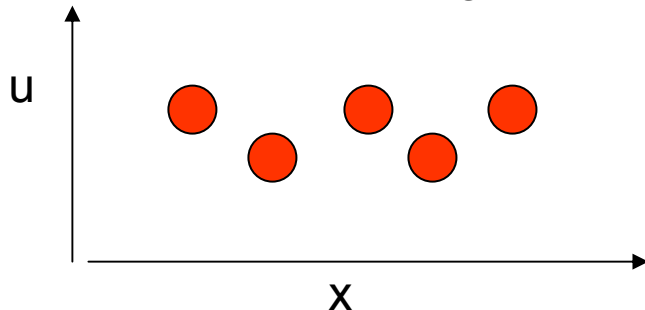
analityczne:  $u(x) = \frac{1}{\pi^2} \sin(\pi x)$  .

---

metoda różnic skończonych:

$$\frac{d^2u}{dx^2} \simeq \frac{u(x+dx) + u(x-dx) - 2u(x)}{dx^2}$$

↓  
układ równań algebraicznych na  $u(x_n)$

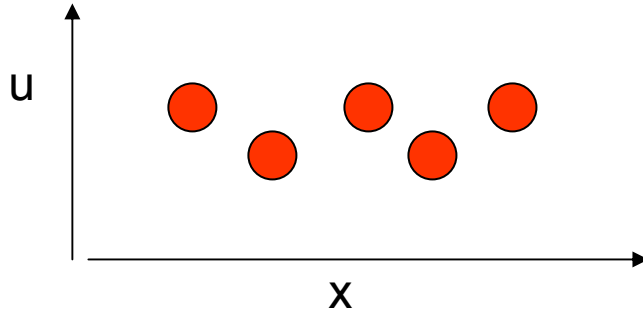


znajdujemy wartości  $u(x)$  w wybranych  
**punktach**

$$\frac{d^2u}{dx^2} = -\sin(\pi x) \quad \frac{d^2u}{dx^2} \simeq \frac{u(x+dx) + u(x-dx) - 2u(x)}{dx^2}$$

metoda różnic skończonych

układ równań algebraicznych na  $u(x_n)$



znajdujemy wartości  $u(x)$  w wybranych punktach

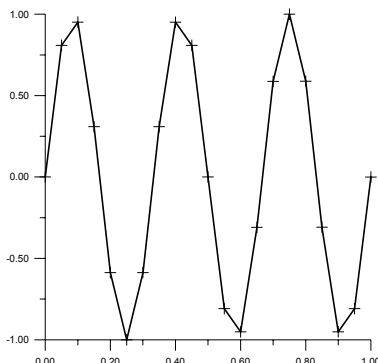
Główna (jedyna) zaleta MRS: prosta dyskretyzacja równań.

Wady: niełatwe lokalne zagęszczanie siatki (drobne, lecz ważne) szczegóły

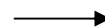
: niełatwy opis objętości o konturze odbiegającym od prostokątnego

: duże zużycie pamięci (istotne ograniczenie dokładności w trzech (i więcej) wymiarach)

wyobraźmy sobie, że rozwiązanie jest szybkozmiennie  
powiedzmy, że bliskie  $\sin(6\pi x)$



$\sin(6\pi x)$  opisany  
na 20 punktach  
dokładność ☹

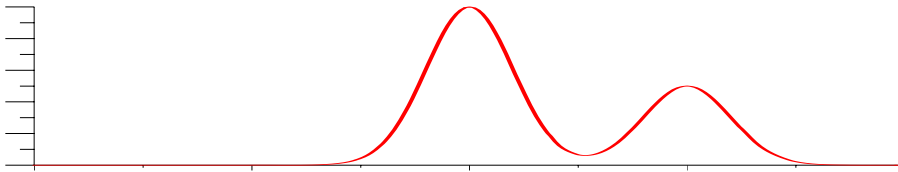


użyć bazy funkcyjnej i  
 $\sin(6\pi x)$  włączyć do bazy funkcyjnej  
w której poszukujemy rozwiązania...

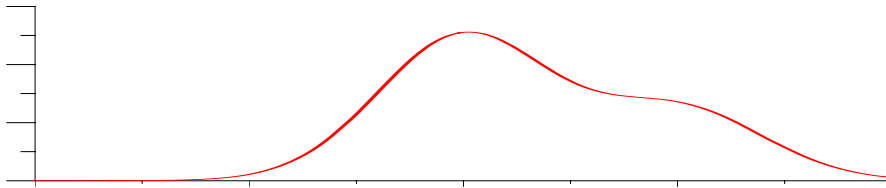
## motywacja do pracy z bazą funkcyjną cd.



wyobraźmy sobie, że mamy siatkę złożoną z dwóch punktów w metodzie różnic skończonych znamy tylko wartości rozwiązania w węzłach ...



baza złożona z dwóch funkcji gaussowskich opisuje rozwiązanie również między węzłami siatki



... a parametrami bazy (funkcji gaussowskich) można dodatkowo manipulować  
... znacznie więcej informacji zawartej w bazie  
... wyniki rachunku zbiegają do dokładnych szybciej w funkcji liczby elementów bazowych niż w funkcji oczek siatki (szczególnie >1D)

---

wyobraźmy sobie, że jako funkcji bazowych użyjemy funkcji  $\sin(nx)$

- rozwiązanie w takiej bazie da nam automatycznie dyskretną transformatę Fouriera rozwiązania

podobnie – informacje użyteczne uzyskamy, jeśli funkcje bazowe mają określoną interpretację

---

Przykład cd.:  $\frac{d^2u}{dx^2} = -\sin(\pi x)$   $u(-1)=0$   
 $u(1)=0$

---

poszukujemy rozwiązania w bazie funkcyjnej

**!**  
 $u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$  funkcje bazowe [trafny wybór: dobre przybliżenie przy minimalnym N]

wybór bazy: zawęża przestrzeń poszukiwań  
optymalnego rozwiązania do wektorowej przestrzeni liniowej,  
którą baza rozpina

optymalne rozwiązanie znaczy optymalne współczynniki  $c_i$

Przykład cd.:  $\frac{d^2u}{dx^2} = -\sin(\pi x)$   $u(-1)=0$   
 $u(1)=0$

---

poszukujemy rozwiązania w bazie funkcyjnej

**!**  
 $u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$  funkcje bazowe [trafny wybór: dobre przybliżenie przy minimalnym N]

wyбір bazy: zaważa przestrzeń poszukiwań

optymalnego rozwiązania do wektorowej przestrzeni liniowej,  
która może rozpiąć

**$v(x)$  – funkcja próbna**

**(test function, trial function)**

optymalne rozwiązanie z zadanych funkcji próbnych

$$\frac{d^2u}{dx^2} = -\sin(\pi x)$$

$$u(-1)=0$$
$$u(1)=0$$

poszukujemy rozwiązania w bazie analitycznie zadanych funkcji

!

$$u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$$

funkcje bazowe [trafny wybór: dobre przybliżenie przy minimalnym  $N$ ]

wybór bazy: zawęża przestrzeń poszukiwań optymalnego rozwiązania

optymalne rozwiązanie znaczy optymalne współczynniki  $c_i$

błąd rozwiązania przybliżonego  $v(x)$ :

$$\frac{d^2u}{dx^2} = -\sin(\pi x)$$

$$u(-1)=0$$
$$u(1)=0$$

poszukujemy rozwiązania w bazie analitycznie zadanych funkcji

!

$$u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$$

funkcje bazowe [trafny wybór: dobre przybliżenie przy minimalnym N]

wybór bazy: zawęża przestrzeń poszukiwań optymalnego rozwiązania

optymalne rozwiązanie znaczy optymalne współczynniki  $c_i$

błąd rozwiązania przybliżonego  $v(x)$ :

$$E(x) = \frac{d^2v}{dx^2} + \sin(\pi x)$$

jeśli  $u=v$ ,  $E=0$   
tak dobieramy  $c_i$  aby  $E$  był „mały”  
Wybór kryterium małości generuje wiele metod.  
Na laboratorium ćwiczymy 3 :  
kolokacji, najmniejszych kwadratów, Galerkina

błąd, reszta, residuum



$$\frac{d^2 u}{dx^2} = -\sin(\pi x)$$

$$\begin{aligned} u(-1) &= 0 \\ u(1) &= 0 \end{aligned}$$

poszukujemy rozwiązania w bazie analitycznie zadanych funkcji

!

$$u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$$

funkcje bazowe [trafny wybór: dobre przybliżenie przy minimalnym N]

wybór bazy: zawęź przestrzeń poszukiwań optymalnego rozwiązania

optymalne rozwiązanie znaczy optymalne współczynniki  $c_i$

błąd rozwiązania przybliżonego  $v(x)$ :

$$E(x) = \frac{d^2 v}{dx^2} + \sin(\pi x)$$

jeśli  $u=v$ ,  $E=0$

tak dobieramy  $c_i$  aby  $E$  był „mały”

Wybór kryterium generuje wiele metod.

Na laboratorium ćwiczymy 3 metody:

kolokacji, najmniejszych kwadratów, Galerkina

dlaczego nie wprowadzić metod w oparciu o bardziej naturalny wybór  $E = u - v$  ?  
... bo  $u$  w praktycznych zastosowaniach jest nieznane  
problem minimalizacji  $\|u-v\|$  gdy  $u$  znane, to problem aproksymacji

wyberzmy bazę  $f_i(x)=(x+1)(x-1)x^{i-1}$

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

Każda z funkcji bazowych spełnia warunki brzegowe.

niech w bazie będą  $N = 3$  funkcje  $[i=1,2,3]$

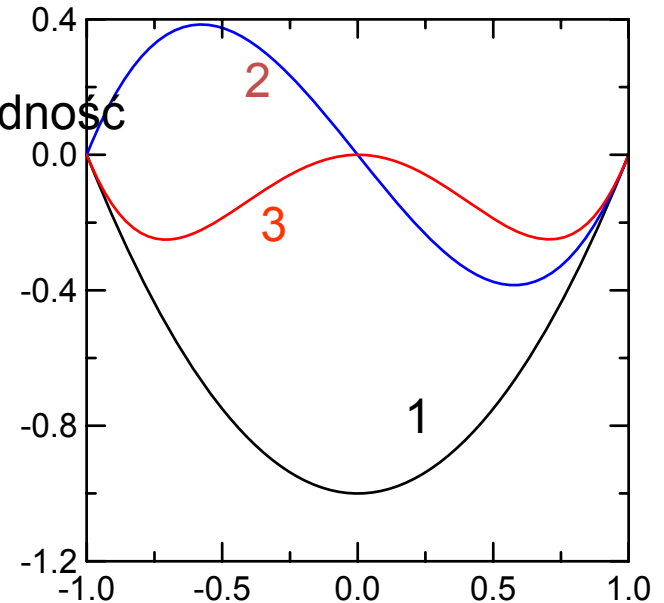
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

(nawet jeśli rozwiązanie analityczne nie istnieje –  
Znamy  $E$  w formie analitycznej jeśli tylko niejednorodność  
równania dana jest wzorem)

$$\frac{d^2u}{dx^2} = -\sin(\pi x)$$

$$u(-1)=0$$

$$u(1)=0$$



wyberzmy bazę  $f_i(x)=(x+1)(x-1)x^{i-1}$

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

Każda z funkcji bazowych spełnia warunki brzegowe.

niech w bazie będą  $N=3$  funkcje  $[i=1,2,3]$

$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

**metoda kolokacji:** niech błąd  $E$  znika w  $N$  punktach przestrzeni (niech funkcja  $v$  spełnia dokładnie równanie różniczkowe w  $N$  wybranych punktach)

$N$  punktów  $x_i$

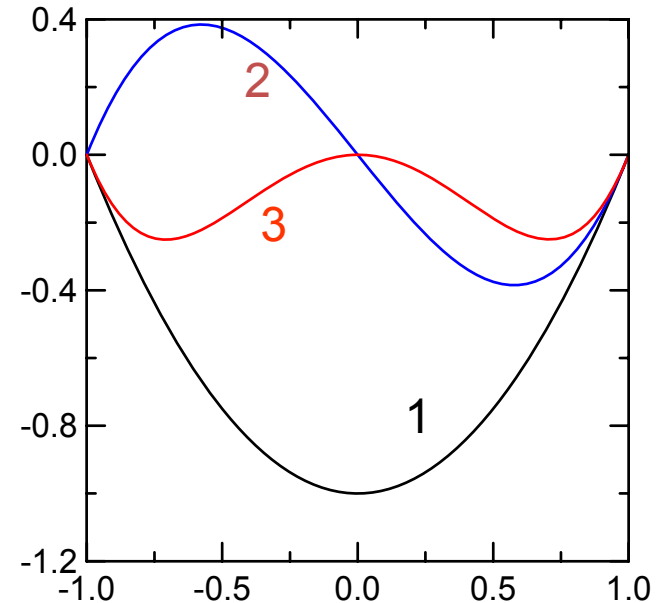
wektor  $c$  dany przez warunek  $E(x_i)=0$

$E(x_j)=0$  – układ  $N$  równań na  $N$  niewiadomych

$$\frac{d^2u}{dx^2} = -\sin(\pi x)$$

$$u(-1)=0$$

$$u(1)=0$$



$E(x_j)=0$  – układ N równań na N niewiadomych

metoda kolokacji

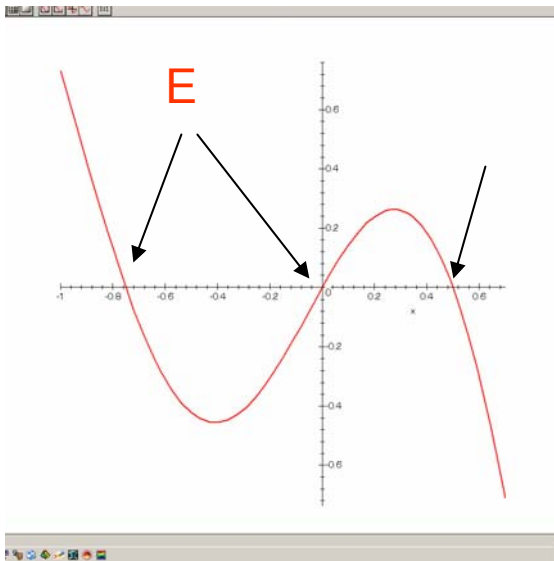
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

$$x_1=0 : 2c_1 - 2c_3 = 0$$

$$x_2=1/2 : 2c_1 + 3c_2 + c_3 + 1 = 0$$

$$x_3=-3/4 : 2c_1 - 4.5c_2 + 4.75c_3 - \sqrt{2}/2 = 0$$

$$c_1 = -\frac{2}{15} + \frac{2}{45}\sqrt{2} \quad c_2 = -\frac{1}{5} - \frac{2}{45}\sqrt{2} \quad c_3 = -\frac{2}{15} + \frac{2}{45}\sqrt{2}$$



$E(x_j)=0$  – układ  $N$  równań na  $N$  niewiadomych

metoda kolokacji

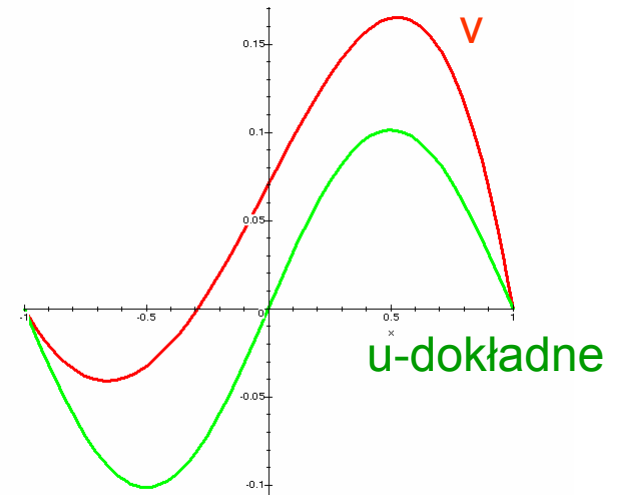
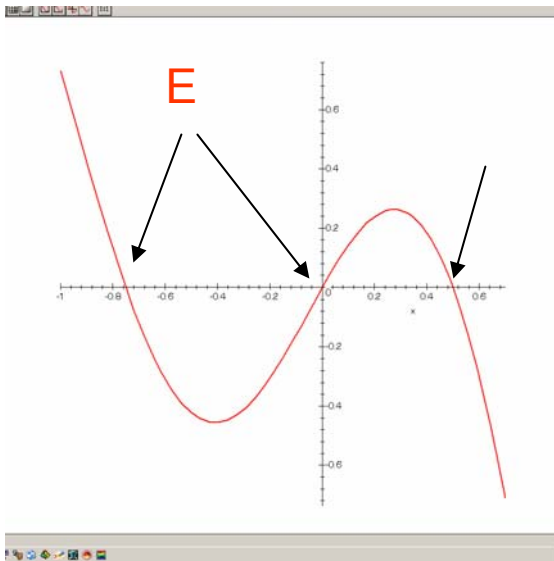
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

$$x_1=0 : 2c_1 - 2c_3 = 0$$

$$x_2=1/2 : 2c_1 + 3c_2 + c_3 + 1 = 0$$

$$x_3=-3/4 : 2c_1 - 4.5c_2 + 4.75c_3 - \sqrt{2}/2 = 0$$

$$c_1 = -\frac{2}{15} + \frac{2}{45}\sqrt{2} \quad c_2 = -\frac{1}{5} - \frac{2}{45}\sqrt{2}, \quad c_3 = -\frac{2}{15} + \frac{2}{45}\sqrt{2}.$$



$E(x_j)=0$  – układ N równań na N niewiadomych

metoda kolokacji

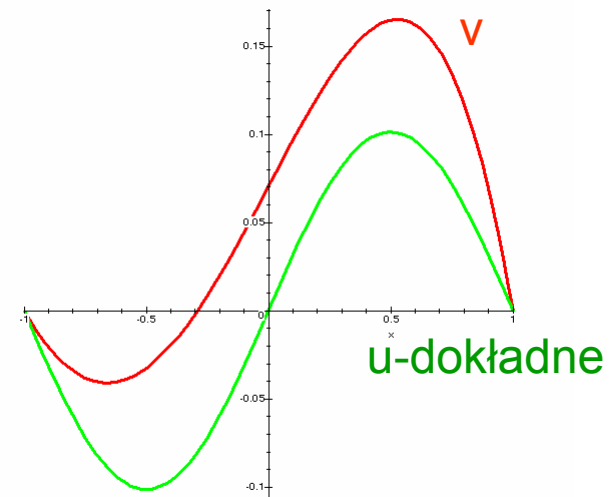
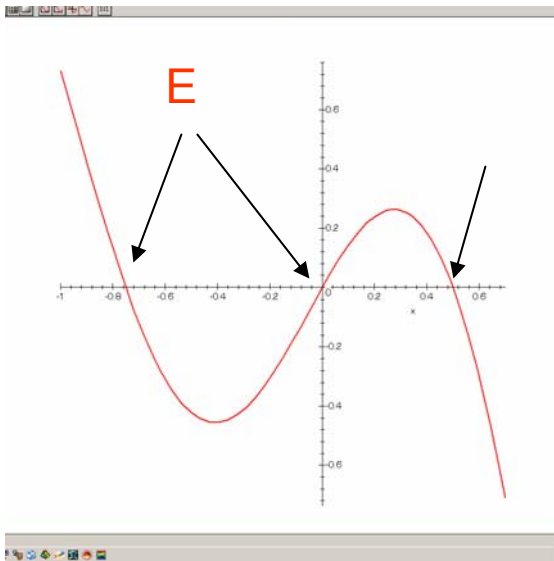
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

$$x_1=0 : 2c_1 - 2c_3 = 0$$

$$x_2=1/2 : 2c_1 + 3c_2 + c_3 + 1 = 0$$

$$x_3=-3/4 : 2c_1 - 4.5c_2 + 4.75c_3 - \sqrt{2}/2 = 0$$

$$c_1 = -\frac{2}{15} + \frac{2}{45}\sqrt{2} \quad c_2 = -\frac{1}{5} - \frac{2}{45}\sqrt{2}, \quad c_3 = -\frac{2}{15} + \frac{2}{45}\sqrt{2}.$$



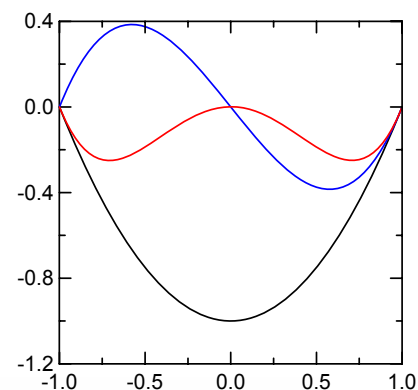
Uwaga:  $E(x_a)=0$  NIE znaczy  $v(x_a)=u(x_a)$  bo błąd to nie jest odchylenie od wartości dokładnej. W naszym równaniu  $E(x_a)=0$  znaczy:  $v''(x_a)=u''(x_a)$

$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

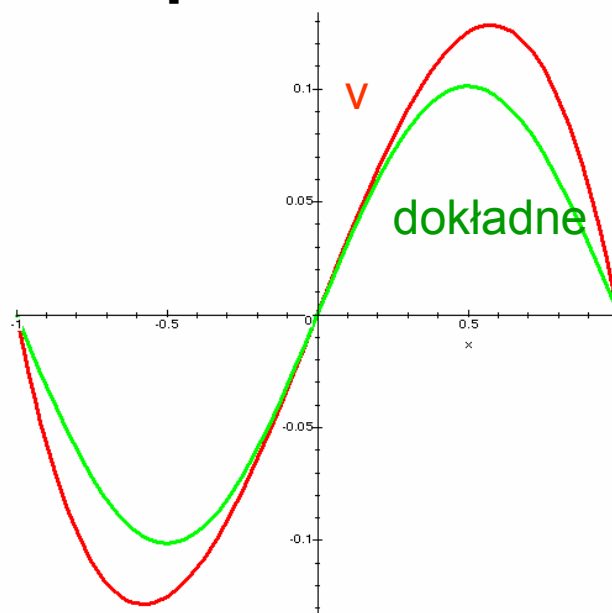
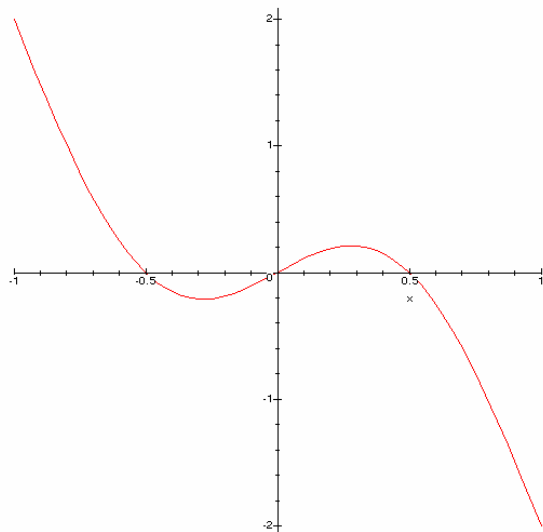
metoda kolokacji

Jakość rozwiązania :  
zależy od wyboru punktów kolokacji

$$\begin{aligned} x_1=0 & : 2c_1 - 2c_3 = 0 \\ x_2=1/2 & : 2c_1 + 3c_2 + c_3 + 1 = 0 \\ x_3=-1/2 & : 2c_1 - 3c_2 + c_3 - 1 = 0 \end{aligned} \rightarrow \begin{aligned} c_1 &= c_3 = 0 \\ c_2 &= -1/3 \end{aligned}$$



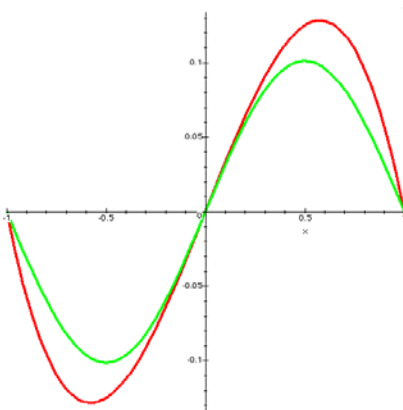
[symetria odrzuca parzyste elementy bazowe]



lepiej niż poprzednio, mimo że tylko jedna funkcja bazowa pracuje

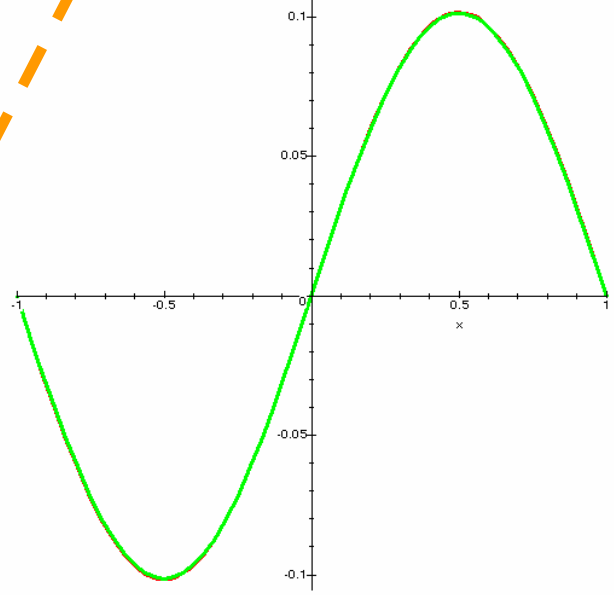
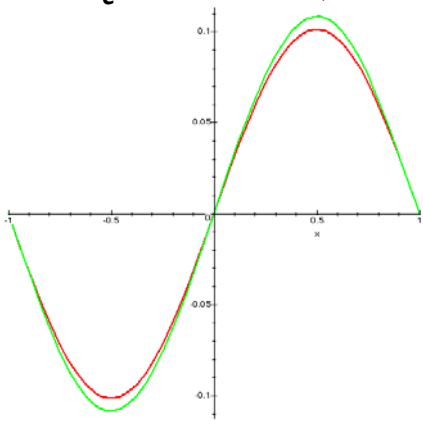
wyberzmy bazę  $f_i(x) = (x+1)(x-1)x^{i-1}$

$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

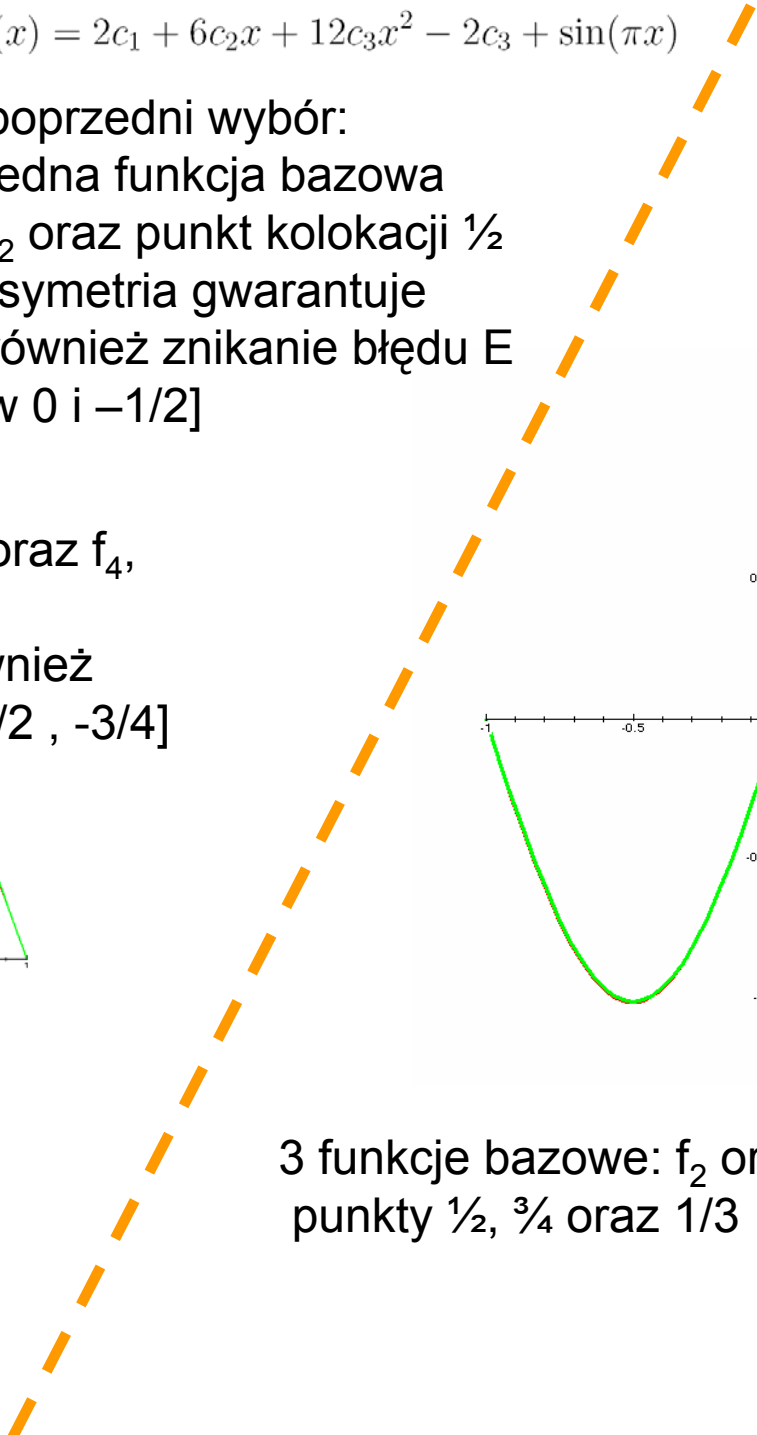


poprzedni wybór:  
jedna funkcja bazowa  
 $f_2$  oraz punkt kolokacji  $\frac{1}{2}$   
[symetria gwarantuje  
również znikanie błędu  $E$   
w  $0$  i  $-1/2$ ]

dwie funkcje bazowe:  $f_2$  oraz  $f_4$ ,  
punkty:  $\frac{1}{2}$  oraz  $\frac{3}{4}$   
[symetria gwarantuje również  
znikanie błędu  $E$  w:  $0$ ,  $-1/2$ ,  $-3/4$ ]

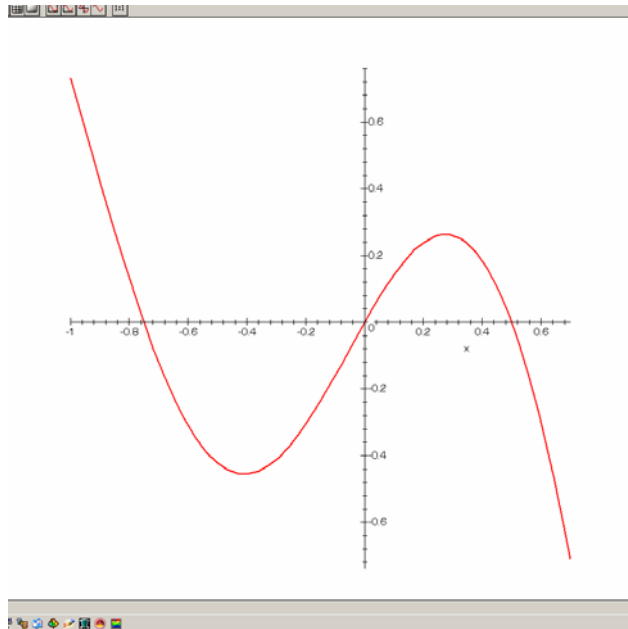


3 funkcje bazowe:  $f_2$  oraz  $f_4$ ,  $f_6$   
punkty  $\frac{1}{2}$ ,  $\frac{3}{4}$  oraz  $\frac{1}{3}$





problem z metodą kolokacji: jeśli nawet  $E$  znika w wybranych punktach  $E(x)$  może znacznie od zera odbiegać w pozostałych



$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

pomysł: uznamy za optymalną funkcję próbną  $v$  dla której przeciętne  $E^2$  jest minimalne

$$F = \int_{-1}^1 E^2(x) dx$$

$$\min F(c_1, c_2, \dots, c_N)$$



metoda najmniejszych kwadratów

- 1) odpada wybór punktów kolokacji
- 2) pojawia się konieczność całkowania  
[ kolokacja jest jedyną metodą, w której całkować nie trzeba, co okazuje się zaletą gdy problem jest wielowymiarowy i gdy funkcje bazowe i niejednorodność są w całkowaniu trudne  
[np. – baza trygonometryczna]

$$\frac{\partial F}{\partial c_j} = 0$$

dostaniemy znowu układ równań liniowych

wyberzmy bazę  $f_i(x)=(x+1)(x-1)x^{i-1}$

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

każda z funkcji bazowych spełnia warunki brzegowe.

niech w bazie będą N=3 funkcje [i=1,2,3]

$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

$$F = \int_{-1}^1 E^2(x) dx$$

$$F := \frac{1}{5} \frac{40 c_1^2 \pi + 80 c_1 c_3 \pi + 120 c_2 + 168 c_3^2 \pi + 120 c_2^2 \pi + 5 \pi}{\pi}$$

metoda najmniejszych  
kwadratów

$$F := \frac{1}{5} \frac{40 c1^2 \pi + 80 c1 c3 \pi + 120 c2 + 168 c3^2 \pi + 120 c2^2 \pi + 5 \pi}{\pi}$$

**r1 := diff (F, c1) ;**

$$r1 := \frac{1}{5} \frac{80 c1 \pi + 80 c3 \pi}{\pi}$$

**r2 := diff (F, c2) ;**

$$r2 := \frac{1}{5} \frac{120 + 240 c2 \pi}{\pi}$$

**r3 := diff (F, c3) ;**

$$r3 := \frac{1}{5} \frac{336 c3 \pi + 80 c1 \pi}{\pi}$$

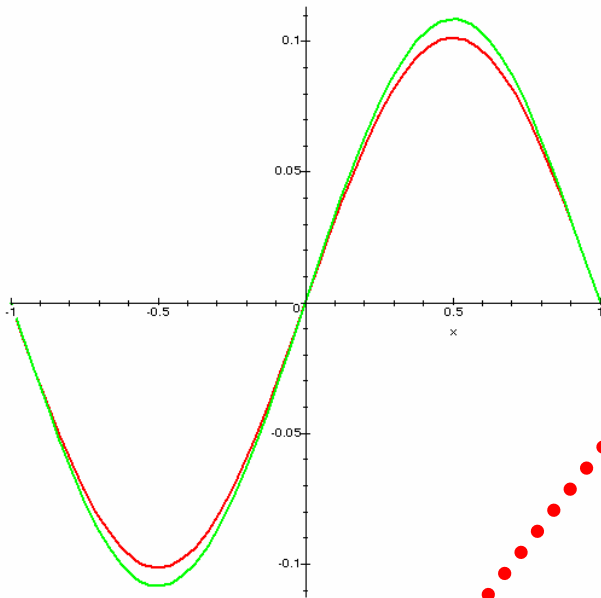
**solve ({r1, r2, r3}, {c1, c2, c3}) ;**

$$\{c3 = 0, c2 = -\frac{1}{2} \frac{1}{\pi}, c1 = 0\}$$

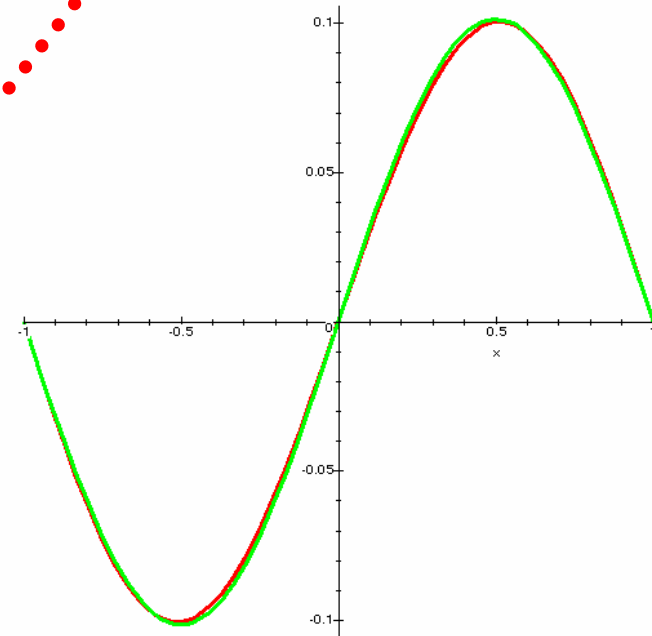
odrzucone funkcje bazowe o złej symetrii

dwie funkcje bazowe:  $f_2$  oraz  $f_4$

punkty kolokacji:  $\frac{1}{2}$  oraz  $\frac{3}{4}$   
[symetria gwarantuje również 0,  $-\frac{1}{2}$ ,  $-\frac{3}{4}$ ]



kolokacja



metoda najmniejszych kwadratów przy tej samej bazie

min □ okazują się lepsze  
od kolokacji w sensie przeciętnej wartości  $|u-v|$

## Metoda reszt ważonych

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

$$E(x) = \frac{d^2 v}{dx^2} + \sin(\pi x)$$

$$R_j = \int_{-1}^1 E(x) w_j(x) dx$$

aby wyznaczyć  $N$  wartości  $c$ , wybieramy  $N$  liniowo niezależnych funkcji wagowych  $w_j$ , i żądamy znikania całki błędu z funkcjami wagowymi  $w_j$

Jeden z możliwych wyborów funkcji wagowych: daje **metodę Galerkiną**:  $w_j = f_j$  (wagi tożsame z funkcjami bazowymi)

$$E(x) = 2c_1 + 6c_2 x + 12c_3 x^2 - 2c_3 + \sin(\pi x)$$

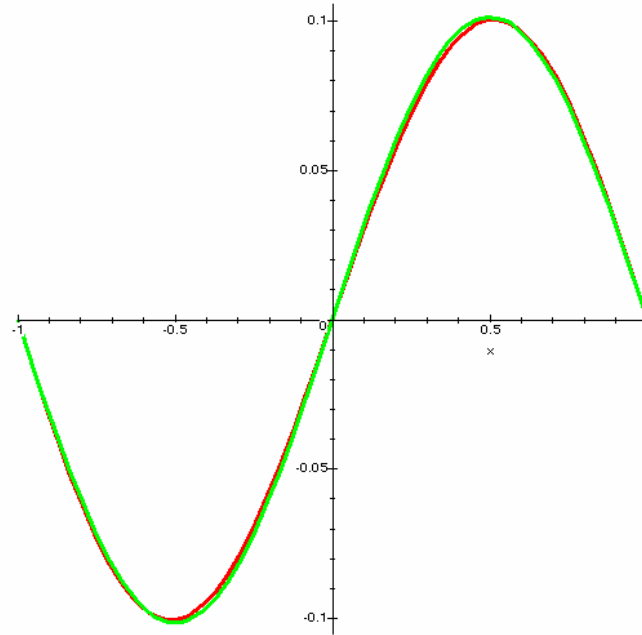
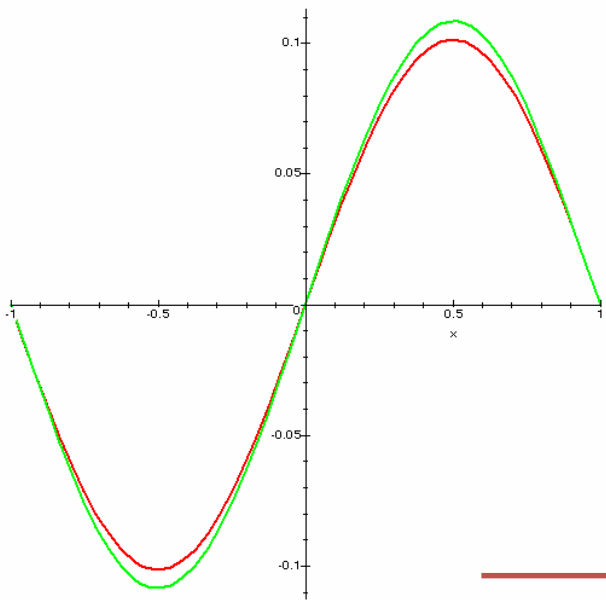
$$f_i(x) = (x+1)(x-1)x^{i-1}$$

$R_j =$  układ równań na  $c_j$

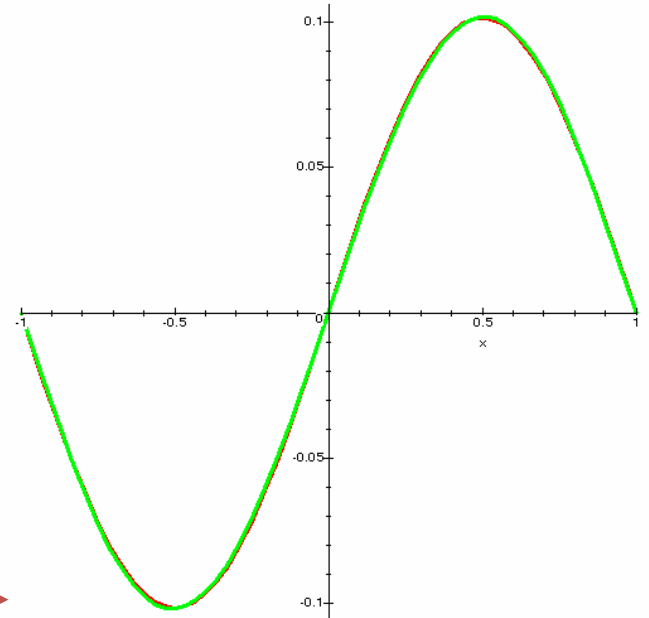
baza 2 funkcji  $f_2, f_4$   
porównanie metod

min □

Kolokacja



Galerkin



jakość rozwiązania



# problem istnienia jednoznacznego rozwiązania dla kolokacji

metoda kolokacji: wybieramy  $N$  punktów  $x_k$ , układ równań na  $c_i$ :  $E(x_k)=0$

$$Lv(x_k)=g(x_k)$$

$$u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$$

załóżmy, że  $L$  – liniowy

$$\sum_{i=1}^N Lf_i(x_k)c_i = g(x_k)$$

$$\mathbf{Ac}=\mathbf{g}$$

$$\mathbf{A}_{ki} = L f_i(x_k)$$

$$\begin{bmatrix} Lf_1(x_1) & Lf_2(x_1) & Lf_3(x_1) \\ Lf_1(x_2) & Lf_2(x_2) & Lf_3(x_2) \\ Lf_1(x_3) & Lf_2(x_3) & Lf_3(x_3) \end{bmatrix}$$

Aby istniało jednoznaczne rozwiązanie URL, potrzeba aby?

# problem istnienia jednoznacznego rozwiązania dla kolokacji

metoda kolokacji: wybieramy  $N$  punktów  $x_k$ , układ równań na  $c_i$ :  $E(x_k)=0$

$$Lv(x_k)=g(x_k)$$

$$u(x) \simeq v(x) = \sum_{i=1}^N c_i f_i(x)$$

załóżmy, że  $L$  – liniowy

$$\sum_{i=1}^N Lf_i(x_k)c_i = g(x_k)$$

$$\mathbf{Ac}=\mathbf{g}$$

$$\mathbf{A}_{ki} = L f_i(x_k)$$

$$\begin{bmatrix} Lf_1(x_1) & Lf_2(x_1) & Lf_3(x_1) \\ Lf_1(x_2) & Lf_2(x_2) & Lf_3(x_2) \\ Lf_1(x_3) & Lf_2(x_3) & Lf_3(x_3) \end{bmatrix}$$

Aby istniało jednoznaczne rozwiązanie URL, wartości funkcji w kolejnych punktach kolokacji (wiersze)  $Lf_i(x_k)$  muszą być liniowo niezależne

funkcje  $f_i$  są liniowo niezależne [tak wybieramy bazę]  
czy mamy gwarancję, że również funkcje  $Lf_i$  - są niezależne liniowo?



$$f_i = x^i \quad [i=0, 1, 2, \dots]$$

$u''(x) = -\rho(x)$  [ $L$  = druga pochodna], wtedy  $Lf_0 = Lf_1 = 0$  (z tak wybraną bazą **kolokacja zawiedzie** niezależnie od wyboru punktów **bo zbiór funkcji  $Lf$  nie jest bazą (mimo, że  $f$  – jest).**

Czy jest to problem?

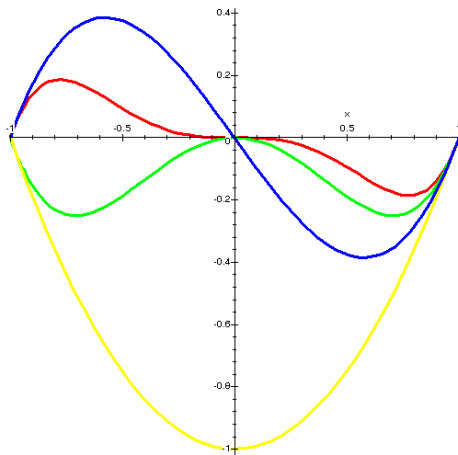
- 1)  $(ax+b)$  nie ma wpływu na  $v''$ .
- 2)  $(ax+b)$  może przydać się przy określeniu warunku brzegowego
- 3)  $(ax+b)$  nie jest potrzebne w bazie, może być wprowadzone poza kolokacją (dodane do rozwiązania kolokacji)

przerabiany przykład:

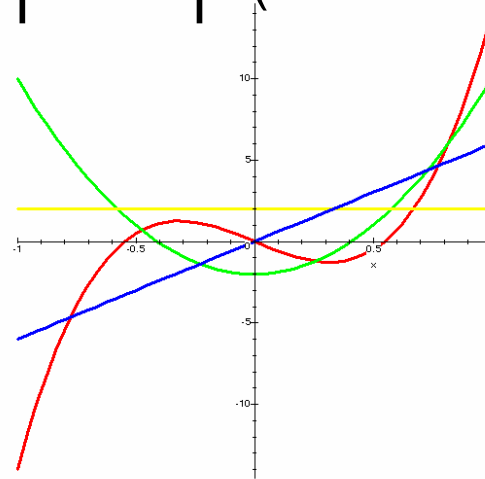
baza  $f_i(x)=(x+1)(x-1)x^{i-1}$

$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \begin{array}{l} u(-1)=0 \\ u(1)=0 \end{array}$$

$f_i$



$h_i = Lf_i$  (wielomiany różnych stopni)



Lf jest bazą

Jeśli  $Lf_i$  – układ funkcji liniowo niezależnych:

**na pewno** istnieje taki wybór punktów kolokacji, że problem (układ równań na  $c$ ) ma jednoznaczne rozwiązanie

nie jest jednak tak, że dla dowolnego wyboru punktów próbkowania problem kolokacji będzie miał jednoznaczne rozwiązanie

Np. baza funkcji parzystych

oraz symetrycznie względem zera wybrane punkty kolokacji

ogólnie baza funkcji, które przyjmują tę samą wartość w dwóch różnych punktach [słaba baza]

[jeśli punkty kolokacji wybrane zostały mało szczęśliwie – dowiemy się o tym na podstawie wyznacznika macierzy URL – będzie bliski zera]

# metoda najmniejszych kwadratów, problem istnienia i jednoznaczności rozwiązania

$$Lu=g$$

$$E(x)=Lv(x)-g(x)$$

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

$$F = \int_{-1}^1 E^2(x) dx$$

minimalne

zał: L – liniowy

$$Lv(x) = \sum_{i=1}^N c_i Lf_i(x)$$

$$F = \int_{-1}^1 dx \left( \sum_i c_i Lf_i(x) - g(x) \right)^2$$

# metoda najmniejszych kwadratów, problem istnienia i jednoznaczności rozwiązania

$$Lu=g$$

$$E(x)=Lv(x)-g(x)$$

$$v(x) = \sum_{i=1}^N c_i f_i(x)$$

$$F = \int_{-1}^1 E^2(x) dx \quad \text{minimalne}$$

zał: L – liniowy

$$Lv(x) = \sum_{i=1}^N c_i Lf_i(x)$$

$$F = \int_{-1}^1 dx \left( \sum_i c_i Lf_i(x) - g(x) \right)^2$$

$$F = \int_{-1}^1 \left( \sum_i \sum_j c_i c_j Lf_i(x) Lf_j(x) + g(x)^2 - 2g(x) \sum_i c_i Lf_i(x) \right) dx$$

oznaczenie  $(a, b) = \int_{-1}^1 dx a(x)b(x)$  iloczyn skalarny w przestrzeni rzeczywistych funkcji całkowalnych z kwadratem

$$F = \sum_i \sum_j c_i c_j (Lf_i, Lf_j) + (g, g) - 2 \sum_i c_i (g, Lf_i)$$

## iloczyn skalarny

Iloczyn skalarny  $(u,v)$  : parze wektorów (funkcji) przyporządkowuje liczbę zespoloną, taką że

- 1)  $(u,v)=(v,u)^*$  [przemienny ze sprzężeniem]
- 2)  $(u,bv)=b(u,v)$  [liniowy względem mnożenia przez skalar]
- 3)  $(u,v+w)=(u,v)+(u,w)$  [liniowy względem dodawania wektorów]
- 4)  $(u,u)\geq 0$  dodatnio określony [równość tylko gdy  $u=0$ ]

np. dla funkcji całkowalnych z kwadratem w  $V$

$$(u, v) = \int_V u^*(x)v(x)dx$$

dwie funkcje  $u$  oraz  $v$  są ortogonalne o ile  $(u,v)=0$

$$F = \sum_i \sum_j c_i c_j (Lf_i, Lf_j) + (g, g) - 2 \sum_i c_i (g, Lf_i)$$

$$\frac{\partial F}{\partial c_k} = 0 \longrightarrow \sum_j c_j (Lf_k, Lf_j) = (g, Lf_k)$$

$$h_k := Lf_k$$

problem posiada jednoznaczne rozwiązanie jeśli macierz  $H_{kl}$  utworzona z iloczynów skalarnych  $(h_k, h_l)$  jest macierzą nieosobliwą

powinniśmy wybrać  $f_k$  tak, aby zbiór  $h_k$  tworzył bazę,

czy niezależność liniowa funkcji  $h_k$

wystarcza aby problem posiadał jednoznaczne rozwiązanie?

[wątpliwość stąd,

że iloczyn skalarny dwóch różnych par funkcji może być identyczny]

Założmy, że mamy dużo szczęścia i

$h_k := Lf_k$  tworzą bazę ortogonalną [np.  $L=d^2/dx^2$ ,  $f_k=\sin(kx)$ ]  
tzn.  $(h_k, h_l) = N_k \delta_{kl}$

wtedy:

$$\sum_j c_j (Lf_k, Lf_j) = (g, Lf_k)$$

Redukuje się do:  $c_k = (g, h_k) / (h_k, h_k)$

macierz  $(h_k, h_l)$  – diagonalna i z konieczności nieosobliwa  
[osobliwa byłaby tylko w sytuacji, gdy jedna z funkcji  $h_k$  była  
tożsamościowo równa zero, lecz wtedy zbiór  $h_k$  nie tworzyłby bazy]



zazwyczaj baza  $h_k$  nie jest ortogonalna,  
bazę można jednak zortogonalizować  
(stworzyć nowy zbiór funkcji ortogonalnych  $u_k$ )

h (baza zwykła) ->  
u (b. ortonormalna)

(ortonormalizacja Grama-Schmidta):

$$u_1 = h_1 / (h_1, h_1)^{1/2}$$

$$u_2 = h_2 - (h_2, u_1)u_1$$

$u_2$  ma być ortogonalne do  $u_1$  i ... jest:

$$(u_2, u_1) = (h_2, u_1) - (h_2, u_1)(u_1, u_1)$$

$$u_n = h_n - \sum_{i=1}^{n-1} (h_n, u_i)u_i$$

$$u_n := u_n / (u_n, u_n)^{1/2}$$

jako kolejny element bazy ortonormalnej  $u$   
bierzemy element bazy oryginalnej  $h$ , liczymy  
jego przekrywanie z wcześniej przyjętymi  
elementami bazy ortonormalnej  $u$  i  
odcinamy odpowiednie przyczynki od  
orzyjmowanego do bazy  $u$  elementu  $h$

# Ortogonalizacja Grama-Schmidta

Przedział  $[-1, 1]$ .

Mamy zbiór niezależnych liniowo funkcji  $h_0=1$ ,  $h_1=x$ ,  $h_2=x^2$ ,  $h_3=x^3$ , ...  
które nie są ortogonalne [iloczyn skalarny określony z funkcją wagową  $w(x)$ ].

Chcemy skonstruować bazę wielomianów ortogonalnych.

Dostaniemy wielomiany Legendre'a.

$$u_0 = 1$$

$$u_1 = a + x$$

Jakie  $a$  aby  $(u_0, u_1) = 0$  ? : odp.:  $a = 0$

$$u_1 = x$$

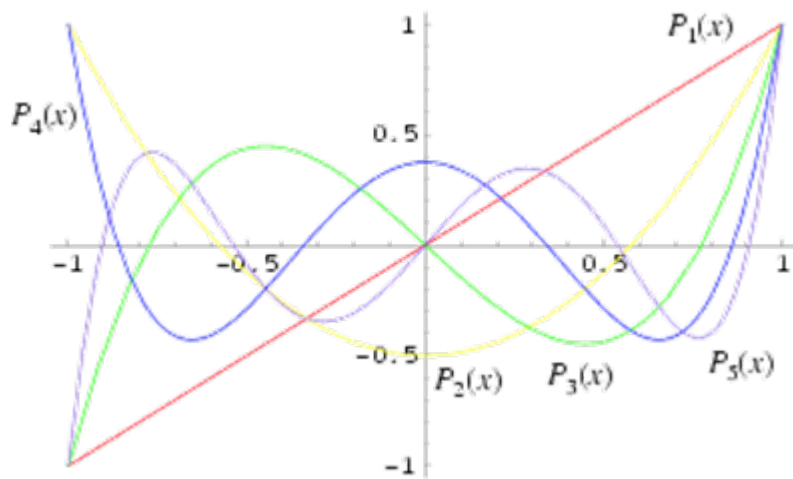
$$u_2 = x^2 + bx + c$$

$$(u_2, u_0) = 2/3 + 2c = 0$$

$$(u_2, u_1) = 0 \rightarrow b = 0$$

$$u_2 = (x^2 - 1/3)$$

W literaturze wielomiany Legendre'a normalizowane tak  
aby  $P_k(1) = 1$  :  $1, x, 3/2 (x^2 - 1/3)$



ltd.

ortonormalizacja GS:

z jednej bazy przechodzimy do drugiej (ortonormalnej)

przestrzeń rozpięta przez obydwie bazy jest identyczna

[baza  $1, x, x^2$  generuje tą samą przestrzeń wielomianów 2 stopnia

jak baza  $L_0 = 1, L_1 = x, L_2 = x^2 - 1/3$

baza nieortonormalna jest mniej wygodna, ale równie elastyczna]

ortonormalizacja GS:

z jednej bazy przechodzimy do drugiej (ortonormalnej)

przestrzeń rozpięta przez obydwie bazy jest identyczna

[przy pomocy  $1, x, x^2$  można wygenerować przestrzeń wielomianów 2 stopnia

tak samo dobrze jak przy pomocy  $L_0, L_1, L_2$

baza nieortonormalna jest mniej wygodna, ale równie elastyczna]

$$F = \int_{-1}^1 E^2(x) dx \quad E(x) = Lv(x) - g(x)$$

problem znalezienia takiego  $v$  aby  $F$  – minimalny

ma to samo rozwiązanie dla bazy przed i po ortonormalizacji

ortonormalizacja GS:

z jednej bazy przechodzimy do drugiej (ortonormalnej)

przestrzeń rozpięta przez obydwie bazy jest identyczna

[przy pomocy  $1, x, x^2$  można wygenerować przestrzeń wielomianów 2 stopnia

tak samo dobrze jak przy pomocy  $L_0, L_1, L_2$

baza nieortonormalna jest mniej wygodna, ale równie elastyczna]

$$F = \int_{-1}^1 E^2(x) dx \quad E(x) = Lv(x) - g(x)$$

problem znalezienia takiego  $v$  aby  $F$  – minimalny

ma to samo rozwiązanie dla bazy przed i po ortonormalizacji

w bazie ortonormalnej problem ma niewątpliwie jednoznaczne rozwiązanie ...

...

ma więc je również w każdej innej bazie skonstruowanej przez kombinacje liniowe elementów tej bazy.

wniosek: Niezależność liniowa zbioru  $Lf_k$  wystarczy do istnienia jednoznacznego rozwiązania optymalnego w sensie najmniejszej całki z kwadratu błędu.

## Metoda Galerkina

problem różniczkowy:  $Au=f$

(silna forma równania,  
równość funkcji)

1)  $E=Au-f$

2)  $u = \sum_{j=1}^N c_j \Phi_j$

3)  $\int dx \Phi_k(x) \times$

$$\sum_j c_j (\Phi_k, A\Phi_j) - (\Phi_k, f) = 0$$

(forma słaba równania.  
równość N liczb  
zamiast równości funkcji)

$$\mathbf{A} \mathbf{c} = \mathbf{F}$$

układ równań na  $c$

# metoda Galerkina: residuum a przestrzeń wektorowa rozpięta przez wektory wybranej bazy

$$Au=f$$

zamiast wprowadzać błąd  $E$ , można po prostu wstawić funkcję próbną do oryginalnego równania

$$u(x) = \sum_{i=1}^n c_i v_i(x)$$

a potem wyrzutować lewą i prawą stronę na  $j$ -ty element bazy

chcemy znaleźć taki element przestrzeni aby:  $(Au, v_j) = (f, v_j)$

**słaba forma** równania

błąd  $E=Au-f$ : **ortogonalny do każdego wektora bazowego**

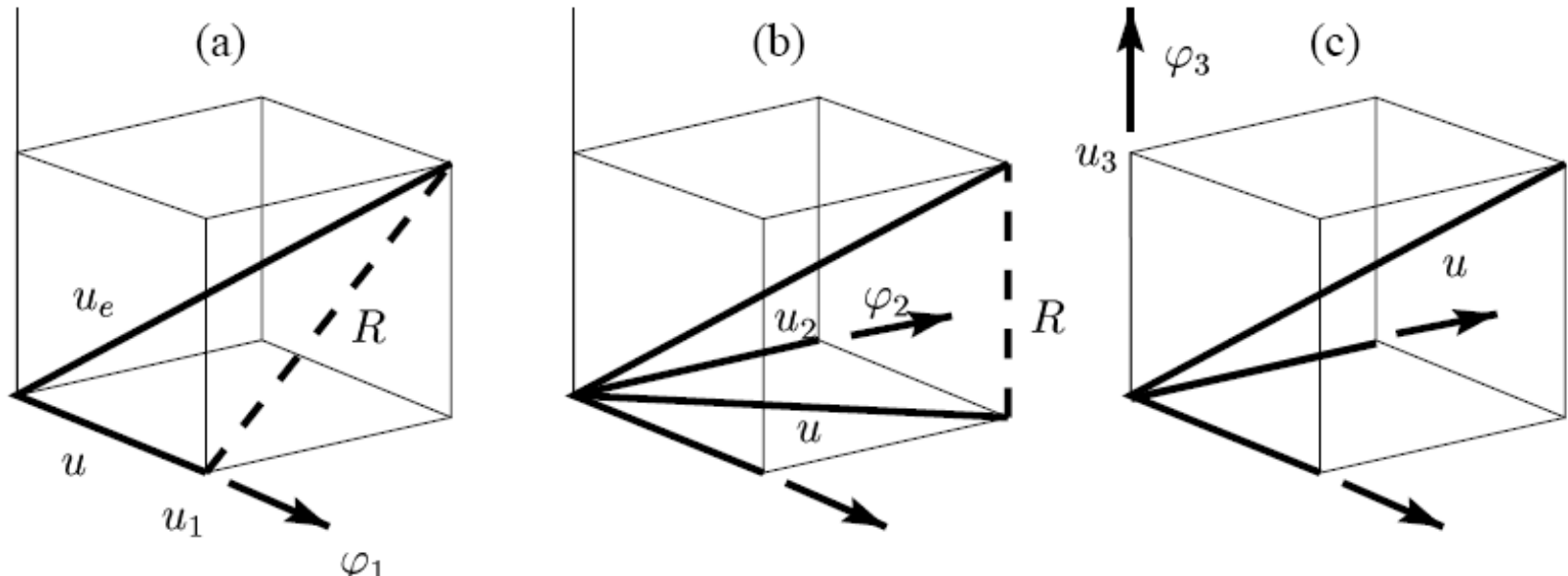
$$(E, v_j) = 0$$

błąd (residuum) znajduje się poza przestrzenią generowaną przez wybraną bazę

ilustracja:  $u_e$  to rozwiązanie dokładne (przekątna sześcianu),  
 $u$  to rozwiązanie przybliżone

$R$  tutaj to  $u_e - u$

od (a) do (c) dodajemy elementy bazowe  $\phi_1, \phi_2, \phi_3$ .



metoda Galerkina rzutuje rozwiązanie dokładne na wektory wybranej bazy

błąd metody: residuum – jest ortogonalne do podprzestrzeni wyznaczonej przez wektory bazy

metoda Galerkina jest zbieżna: gdy baza obejmuje całą przestrzeń – nie ma miejsca na residuum



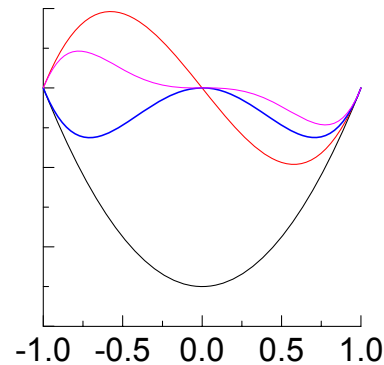
Przykład: z laboratorium

$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \begin{array}{l} u(-1)=0 \\ u(1)=0 \end{array} \quad \text{Dirichleta}$$

analityczne:  $u(x) = \frac{1}{\pi^2} \sin(\pi x)$

baza spełniająca Dirichleta

$$v_i = (x + 1)(x - 1)x^{i-1}$$



# Galerkin

$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \mathbf{SY}=\mathbf{F}$$

$$L = \frac{d^2}{dx^2}$$

$$S_{ij} = (Lv_i, v_j) = \int_{-1}^1 dx \frac{d^2 v_i}{dx^2} v_j$$

$$S_{ij} = \cancel{\frac{dv_j}{dx} v_j} \Big|_{-1}^1 - \int_{-1}^1 dx \frac{dv_i}{dx} \frac{dv_j}{dx}$$

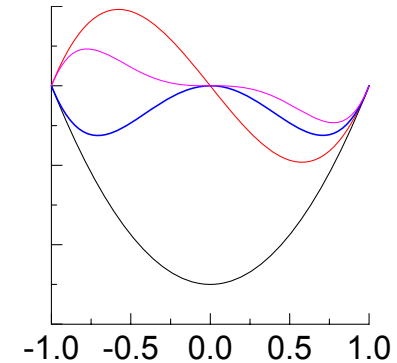
z warunków  
brzegowych

dla  $i$  oraz  $j$  tej samej parzystości

$$S_{ij} = -2 \left[ \frac{(i+1)(j+1)}{i+j+1} + \frac{(i-1)(j-1)}{i+j-3} - \frac{(i+1)(j-1) + (i-1)(j+1)}{i+j-1} \right]$$

baza spełniająca warunki Dirichleta

$$v_i = (x+1)(x-1)x^{i-1}$$



całkowanie przez części

prawa strona:

$$I(k) = - \int_{-1}^1 dx \sin(\pi x) x^k$$

$$I(1) = -\frac{2}{\pi}$$

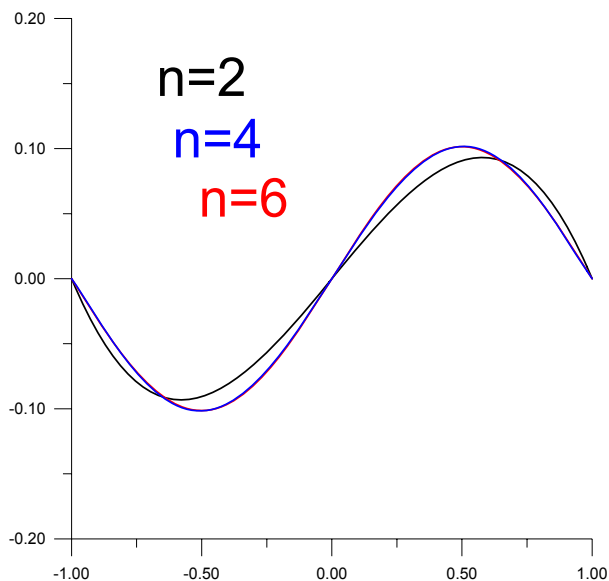
$$I(k) = -\frac{2}{\pi} - \frac{k(k-1)}{\pi^2} I(k-2)$$

$$F(j) = I(j+1) - I(j-1)$$

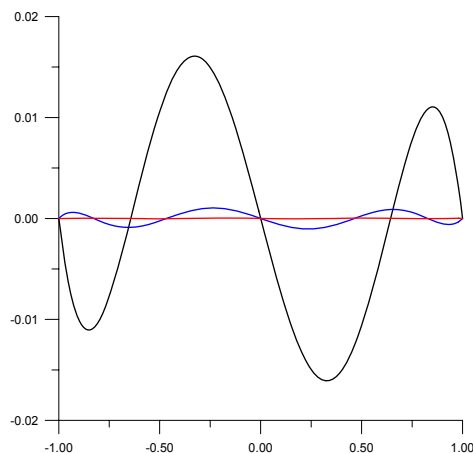
dla  $j$  nieparzystych

---

rozwiązanie



błąd  $\varepsilon$  (nie residuum tylko różnica dokładne – Galerkina):

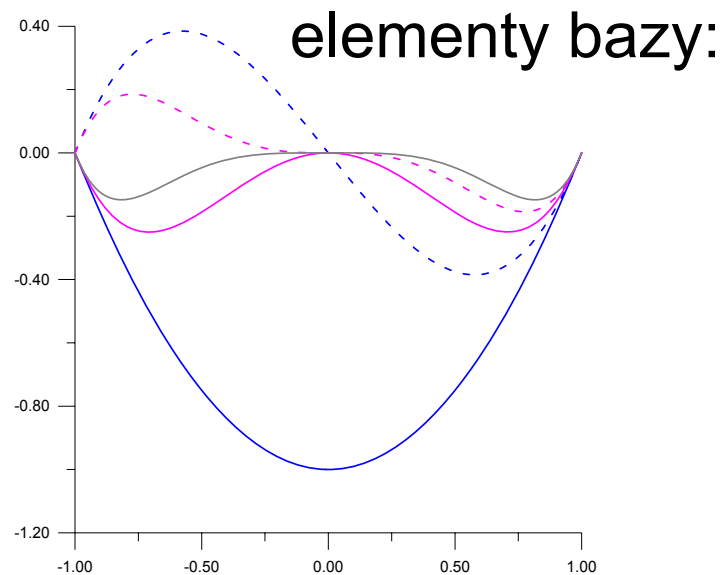
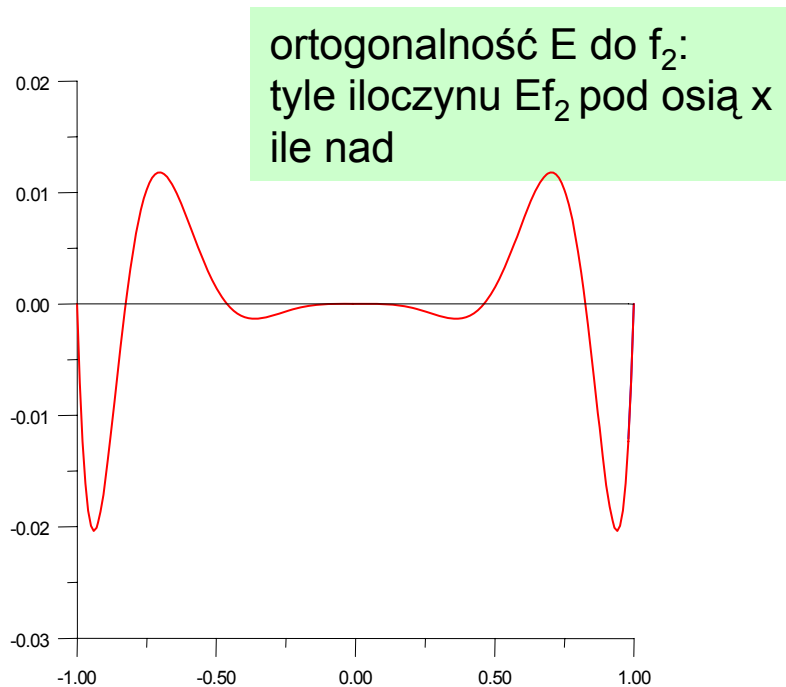


## ortogonalność residuum do wektorów bazowych

$$E(x) = Lv - f$$

$$E(x) = \frac{d^2v}{dx^2} + \sin(\pi x)$$

zgodnie z naszą wiedzą: ma być  $(E, v_i) = 0$   
E ortogonalne do elementów bazy  
z  $i=1, 3$  oraz  $5$  – bo te są funkcjami parzystymi



# Metoda Galerkinia - równoważna metodzie wariacyjnej, (gdy ta stosowalna)

## metoda wariacyjna (Reyleigha-Ritza)

Na jednym z poprzednich wykładów pokazaliśmy, że

$$\nabla^2 \phi = -\rho \longleftrightarrow S = \min$$

$$S = \int \left( \frac{1}{2} (\nabla \phi)^2 - \rho \phi \right) dv$$

S – używaliśmy jako parametr zbieżności  
metod iteracyjnego rozwiązywania równania Poissona

Warunek minimum funkcjonału + baza funkcyjna = metoda wariacyjna RR

## Wariacyjne sformułowanie równania różniczkowego

r. różniczkowe na **rzeczywistą** funkcję  $u$ :

$Au=f$  w  $\Omega$ , z jednorodnym warunkiem brzegowym  $u=0$  na brzegu  $\Gamma$ ,  
 $A$  liniowy, dodatnio określony, samosprężony operator różniczkowy:

liniowy  $A(f_1+f_2)=Af_1+Af_2$

dodatnio określony 
$$\int_{\Omega} uAu \geq 0$$

samosprężony 
$$\int_{\Omega} uAv = \int_{\Omega} vAu$$
 (zakładamy, że funkcje rzeczywiste)

wtedy rozwiązanie równania różniczkowego  $Au=f$  jest takie, że

$$S(u) = \int_{\Omega} u \left( \frac{1}{2} Au - f \right) \quad \text{minimalne}$$

Przykład:  $A = -d^2/dx^2$  jest operatorem liniowym i dodatnio określonym w przestrzeni funkcji całkowalnych z kwadratem i znikających na granicy pudła obliczeniowego [  $u(1)=u(0)=0$  ]

całkowanie przez części:  $(fg)' = f'g + fg'' \rightarrow -fg'' = -(fg)'+ f'g'$

$$\int_0^1 u \left( -\frac{d^2 u}{dx^2} \right) dx = - \cancel{u(x) \frac{du}{dx} \Big|_0^1} + \int_0^1 \left( \frac{du}{dx} \right)^2 dx$$

Przykład:  $A = -d^2/dx^2$  jest operatorem samosprężonym

$$- \int_0^1 u \frac{d^2 v}{dx^2} dx = - \cancel{u \frac{dv}{dx} \Big|_0^1} + \int_0^1 \frac{du}{dx} \frac{dv}{dx} dx$$

$$- \int_0^1 v \frac{d^2 u}{dx^2} dx = - \cancel{v \frac{du}{dx} \Big|_0^1} + \int_0^1 \frac{du}{dx} \frac{dv}{dx} dx$$

# metoda wariacyjna Reyleigha-Ritza

$$u(0)=u(1)=0$$

$$-\frac{d^2u}{dx^2} = \rho(x)$$

$$Au=f$$

$$S(u) = \int_0^1 u \left( -\frac{1}{2} \frac{d^2u}{dx^2} - \rho(x) \right) dx$$

$$S(u) = \int_{\Omega} u \left( \frac{1}{2} Au - f \right)$$

---

iloczyn skalarny w przestrzeni rzeczywistych funkcji całkowalnych z kwadratem

$$(a, b) = \int_{\Omega} a(x)b(x)dx$$

$$S(u) = \frac{1}{2}(u, Au) - (u, f)$$



# metoda wariacyjna Reyleigha-Ritza

Baza:  $\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_N$  funkcji spełniających jednorodny warunek brzegowy

$$u = \sum_{i=1}^N c_i \Phi_i \longrightarrow \text{poszukujemy } c_i \text{ dla których } S(u) \text{ minimalny w wybranej bazie}$$

$$S(u) = \frac{1}{2}(u, Au) - (u, f)$$

# metoda wariacyjna Reyleigha-Rit

$$S(u) = \frac{1}{2}(u, Au) - (u, f)$$

$$u = \sum_{i=1}^N c_i \Phi_i \text{ baza funkcji } \underline{\text{rzeczywistych}}$$

liniowość A

$$S(u) = \frac{1}{2} \left( \sum_i c_i \Phi_i, \sum_j c_j A \Phi_j \right) - \left( \sum_i c_i \Phi_i, f \right)$$

$$S(u) = \frac{1}{2} \sum_i \sum_j c_i c_j (\Phi_i, A \Phi_j) - \sum_i c_i (\Phi_i, f)$$

liniowość iloczynu skalarnego

$$\frac{dS(u)}{dc_k} = 0$$

$$S(u) = \frac{1}{2}(u, Au) - (u, f)$$

$$u = \sum_{i=1}^N c_i \Phi_i \text{ baza funkcji rzeczywistych}$$

liniowość A

$$S(u) = \frac{1}{2} \left( \sum_i c_i \Phi_i, \sum_j c_j A \Phi_j \right) - \left( \sum_i c_i \Phi_i, f \right)$$

$$S(u) = \frac{1}{2} \sum_i \sum_j c_i c_j (\Phi_i, A \Phi_j) - \sum_i c_i (\Phi_i, f)$$

liniowość iloczynu skalarnego

$$\frac{dS(u)}{dc_k} = 0$$

$$\frac{1}{2} \sum_i \sum_j \delta_{ik} c_j (\Phi_i, A \Phi_j) + \frac{1}{2} \sum_i \sum_j \delta_{jk} c_i (\Phi_i, A \Phi_j) - \sum_i \delta_{ik} c_i (\Phi_i, f) = 0$$

sumowanie po deltach

$$\frac{1}{2} \sum_j c_j (\Phi_k, A \Phi_j) + \frac{1}{2} \sum_i c_i (\Phi_i, A \Phi_k) - (\Phi_k, f) = 0$$

przepisane:

$$\frac{1}{2} \sum_j c_j (\Phi_k, A\Phi_j) + \frac{1}{2} \sum_i c_i (\Phi_i, A\Phi_k) - (\Phi_k, f) = 0$$

zmiana indeksu i / j

$$\frac{1}{2} \sum_j c_j (\Phi_k, A\Phi_j) + \frac{1}{2} \sum_j c_j (\Phi_j, A\Phi_k) - (\Phi_k, f) = 0$$

przemienność iloczynu skalarnego + **samosprężoność A**

$$\sum_j c_j (\Phi_k, A\Phi_j) - (\Phi_k, f) = 0$$

$$\mathbf{A} \mathbf{c} = \mathbf{F}$$

układ równań na c

zastosowanie metody wariacyjnej (wracamy do przerobionego problemu):

$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \begin{array}{l} u(-1)=0 \\ u(1)=0 \end{array}$$

$$\mathbf{A} \mathbf{c} = \mathbf{F}$$

wyberzmy bazę  $\Phi_i(x) = (x+1)(x-1)x^{i-1}$

$$\sum_j c_j (\Phi_k, A\Phi_j) - (\Phi_k, f) = 0$$

$$\begin{pmatrix} -\frac{8}{3} & 0 & -\frac{8}{15} & 0 & -\frac{8}{35} \\ 0 & -\frac{8}{5} & 0 & -\frac{24}{35} & 0 \\ -\frac{8}{15} & 0 & -\frac{88}{105} & 0 & -\frac{8}{15} \\ 0 & -\frac{24}{35} & 0 & -\frac{184}{315} & 0 \\ -\frac{8}{35} & 0 & -\frac{8}{15} & 0 & -\frac{104}{231} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{12}{\pi^3} \\ 0 \\ 4 \frac{7\pi^2 - 60}{\pi^5} \\ 0 \end{pmatrix}$$

macierz operatora samosprężonego  
- symetryczna

zera tam, gdzie symetria  
się nie zgadza

$$c_1 = c_3 = c_5 = 0$$

zastosowanie metody wariacyjnej (wracamy do przerobionego problemu):

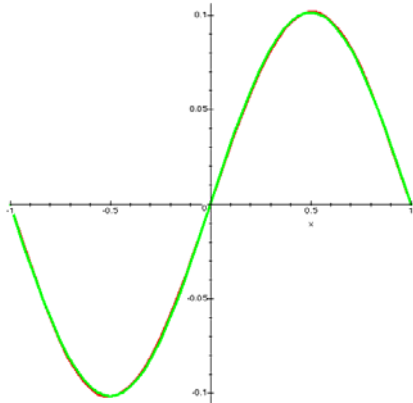
$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \begin{array}{l} u(-1)=0 \\ u(1)=0 \end{array}$$

$$\mathbf{A} \mathbf{c} =$$

$\mathbf{F}$

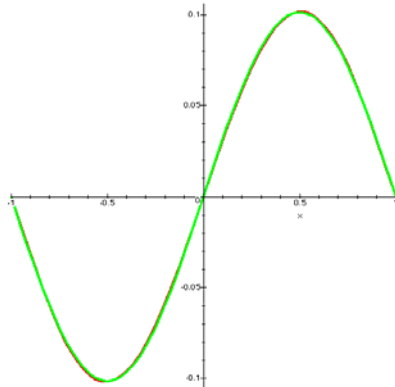
wyberzmy bazę  $\Phi_i(x) = (x+1)(x-1)x^{i-1}$

$$\sum_j c_j (\Phi_k, A\Phi_j) - (\Phi_k, f) = 0$$



$$c_1 = c_3 = c_5 = 0$$

wynik dla bazy funkcji  $F_2, \Phi_4$



$$c_2 := \frac{105}{8} \frac{2\pi^2 - 27}{\pi^5}$$

$$c_4 := -\frac{315}{8} \frac{2\pi^2 - 21}{\pi^5}$$

wynik dokładnie ten sam co w metodzie Galerkina!

URL z zasady wariacyjnej:

$$\sum_j c_j (\Phi_k, A\Phi_j) - (\Phi_k, f) = 0$$

$j$  URL identyczny z układem produkowanym przez metodę Galerkina

1)  $E=Au-f$

2)  $u = \sum_{j=1}^N c_j \Phi_j$

3)  $\int dx \Phi_k(x) \times$

zapis równania na  $c$  w metodzie wariacyjnej Reyleigha-Ritza - identyczny jak w metodzie Galerkina.  
Gdy podejście funkcjonalne obowiązuje: metody RR i G są tożsame.

metoda Galerkina - bardziej ogólna  
- działa również dla operatorów, które nie są samosprężone / liniowe / dodatnio określone  
to jest -  
dla operatorów, dla których funkcjonał osiągający minimum dla rozwiązania równania nie jest znany



## M. Galerkina, a podejście wariacyjne cd.:

Funkcja próbna: z naszego przykładu numerycznego zawiera tylko *liniowe parametry wariacyjne*  $c$ :

$$v(x) = \sum_i c_i f_i(x)$$

$c$  wyznaczone przez URL  
Jeśli tylko równanie różniczkowe  
Jest liniowe

Bardziej elastyczny: rachunek z funkcjami bazowymi zależnymi od *nieliniowych parametrów wariacyjnych*

$$v(x) = \sum_i c_i f_i(x, \alpha_i)$$

Sposób postępowania: dla ustalonej bazy – optymalne  $c$  szukamy jak wyżej. Optymalną bazę (optymalne nieliniowe parametry wariacyjne) znajdujemy minimalizując funkcjonał (zadanie – nieliniowe).

W MES: bazę (podział przestrzeni na elementy) będziemy w ten sposób optymalizować.

# Metoda Galerkina to szczególny przypadek metody reszt ważonych

problem różniczkowy:  $Au=f$  (silna forma równania)

1)  $E=Au-f$

2) 
$$u = \sum_{j=1}^N c_j \Phi_j$$

3)  $\int dx w_k(x) \times$  jeśli jako wag  
użyjemy funkcji  
bazowych  $w_k = \Phi_k$   
mamy m. Galerkina

$$\sum_j c_j (w_k, A\Phi_j) - (w_k, f) = 0 \quad (\text{forma słaba})$$

$$\mathbf{A} \mathbf{c} = \mathbf{F}$$

układ równań na  $c$

$$\sum_j c_j (w_k, A\Phi_j) - (w_k, f) = 0$$

Metoda reszt ważonych: główne punkty  
(i różnice między różnymi wariantami metody):

- 1) Wybór podprzestrzeni wektorowej (bazy)  $\Phi_j$
- 2) Wybór funkcji wagowych  $w_j$
- 3) ... które często wybierane są jako maksymalnie rozłączne przestrzennie  
wtedy podział przestrzeni jest kolejnym problemem

## metoda różnic skończonych dla problemu początkowego w formalizmie reszt ważonych

$$\frac{dy}{dt} = f(t, y) \quad y(t=0)=y_0$$

rozwiązać na  $t$  z przedziału  $[0, T]$

Zadanie: znaleźć przybliżone rozwiązanie  $\tilde{y}$  w  $(N+1)$  chwilach czasowych  $t_n = n\Delta t$ ,  $n=0, 1, \dots, N$   
krok czasowy  $\Delta t = T/N$ .

$$y_n = \tilde{y}(t_n)$$

Pochodna szacowana ilorazem centralnym

$$\frac{dy}{dt} \simeq \frac{y_{n+1} - y_{n-1}}{2\Delta t}$$

$$y_{n+1} = y_{n-1} + 2\Delta t f(t_n, y_n)$$

reguła punktu pośredniego [żabiego skoku]

# Wyprowadzenie metody różnic skończonych w formalizmie reszt ważonych

A)  $y_n$  określone na równoodległych punktach.

Zakładamy, że między punktami  $t_n$  rozwiązanie zmienia się liniowo z  $t$ .



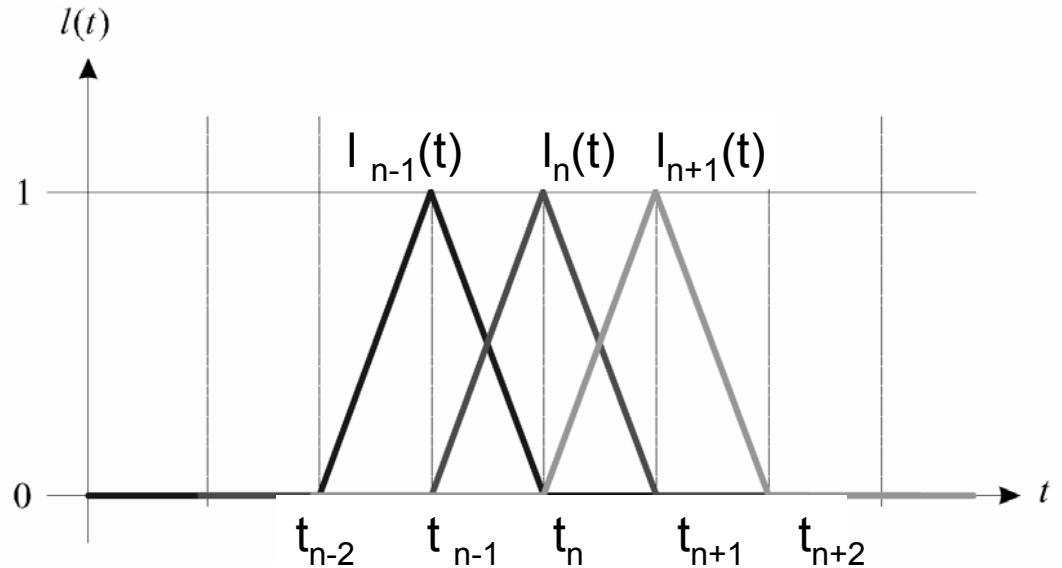
B)  $\Omega=[0,T]$ , i dyskretyzacja na przedziały czasowe  $\tau_m=[t_{m-1},t_m]$

$$C) \quad \tilde{y}(t) = \sum_{i=0}^N c_i l_i(t)$$

baza ma zapewniać odcinkowo liniową zmienność przybliżonego rozwiązania

$$\tilde{y}(t) = \sum_{i=0}^N c_i l_i(t)$$

funkcje bazowe wybieramy tak, aby każda rozwinięta w nich funkcja była ciągła i odcinkami liniowa.



$$l_n(t) = \begin{cases} 0 & t \leq t_{n-1} \\ \frac{t-t_{n-1}}{\Delta t} & t_{n-1} \leq t \leq t_n \\ \frac{t_{n+1}-t}{\Delta t} & t_n \leq t \leq t_{n+1} \\ 0 & t_{n+1} \leq t \end{cases}$$

każda funkcja bazowa określona na dwóch fragmentach  $\omega_m$

Widzimy, że:  $l_m(t_n) = \delta_{nm}$

Wyliczyć współczynniki rozwinięcia:

$$\tilde{y}(t_n) = \sum_{i=0}^N c_i l_i(t_n) = \sum_{i=0}^N c_i \delta_{in} = c_n = y_n$$

Współczynniki rozwinięcia  $c_i$  równe wartościom rozwiązania w węzłach.

Tą samą bazę stosujemy do prawej strony równania  $\frac{dy}{dt} = f(t, y)$

$$f(t, y) = \sum_{i=0}^N f_i l_i(t)$$

---

przepis na  $y_n$ :

$$\longrightarrow \int_0^T \left[ \sum_{i=0}^N c_i l_i'(t) - \sum_{i=0}^N f_i l_i(t) \right] w_j(t) dt = 0$$

---

Potrzebne dookreślenie funkcji wagowych  $w_j$ .

D)

Wagi: rozwiązanie chcemy znać tylko w chwilach  $t_n$  - wagi powinny je wyłuskać

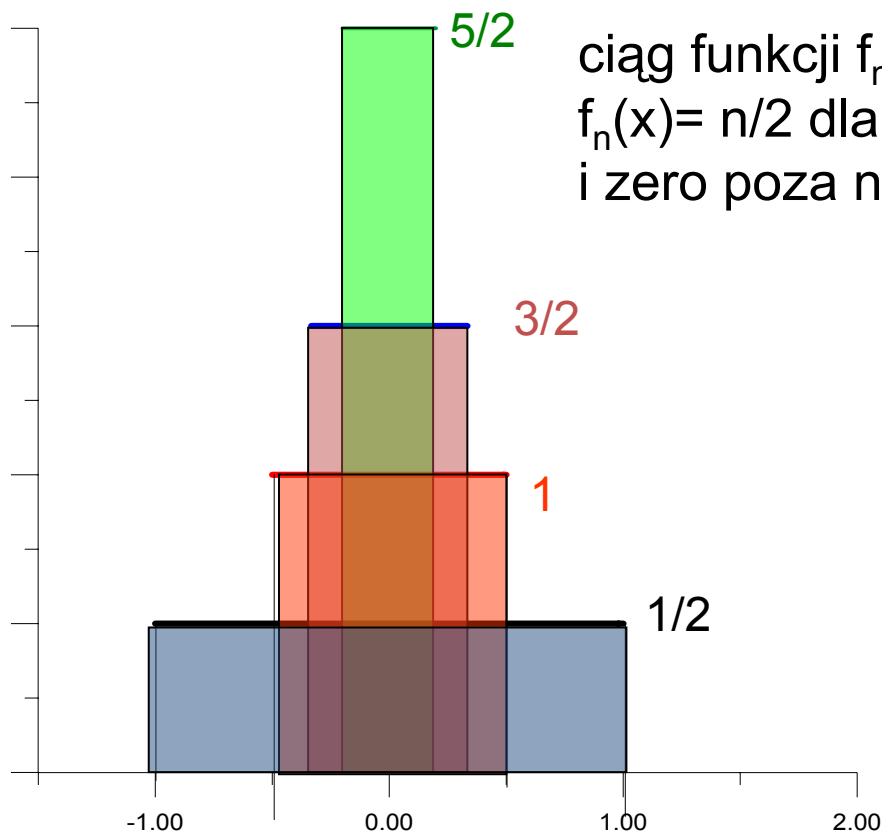
$$\int_0^T \left[ \sum_{i=0}^N c_i l'_i(t) - \sum_{i=0}^N f_i l_i(t) \right] w_j(t) dt = 0$$

--                                  --

$$w_j(t) = \delta(t - t_j) \longleftarrow \text{delta Diraca}$$



# dystrybucja delta Diraca



ciąg funkcji  $f_n$  :

$$f_n(x) = \begin{cases} n/2 & \text{dla } x \in [-1/n, 1/n] \\ 0 & \text{i zero poza nim.} \end{cases} \longrightarrow \int_{-\infty}^{\infty} f_n(x) dx = 1$$

$$\lim_{n \rightarrow \infty} f_n(x) = \delta(x)$$

granica tego ciągu „funkcja”  
(dystrybucja) delta Diraca:

własności:

‘jednostkowy impuls’

$$\delta(x) = \begin{cases} 0 & \text{dla } x \neq 0 \\ \infty & \text{dla } x = 0 \end{cases} \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

konsekwencja, dla ciągłej funkcji  $\int_{-\infty}^{\infty} \delta(x) g(x) dx = g(0)$

$$\lim_{n \rightarrow \infty} f_n(x) = \delta(x)$$

uzasadnienie:

$$\int_{-\infty}^{\infty} \delta(x)g(x)dx = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x)g(x)dx$$

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x)g(x)dx = \lim_{n \rightarrow \infty} \frac{n}{2} \int_{-\frac{1}{n}}^{\frac{1}{n}} g(x)dx$$

$$\lim_{n \rightarrow \infty} \frac{n}{2} \int_{-\frac{1}{n}}^{\frac{1}{n}} g(x)dx = \lim_{n \rightarrow \infty} \frac{n}{2} \frac{2}{n} g(\xi)$$

tw. o wartości  
średniej,  $\xi$   
z przedziału całkowania  
[-1/n, 1/n]

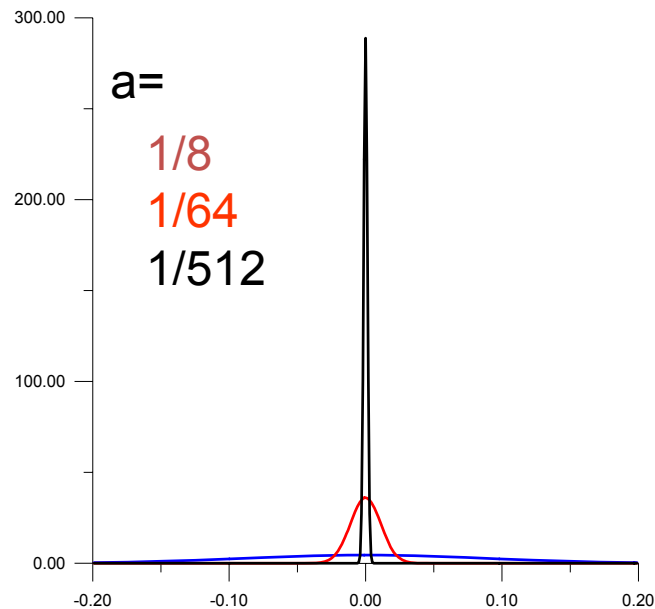
$$\lim_{n \rightarrow \infty} g(\xi) = g(0)$$

---

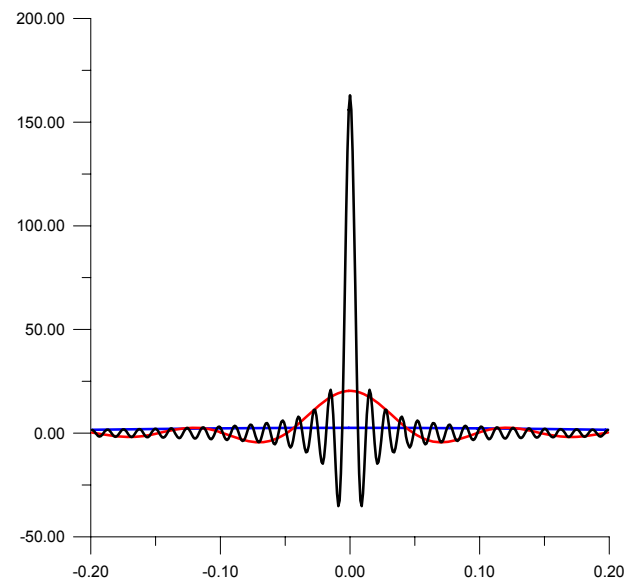
$$\int_{-\infty}^{\infty} \delta(x - a)g(x)dx = g(a)$$

## Inne funkcje dążące do delty Diraca

$$\delta(x) = \lim_{a \rightarrow \infty} \frac{1}{a\sqrt{\pi}} \exp(-x^2/a^2)$$



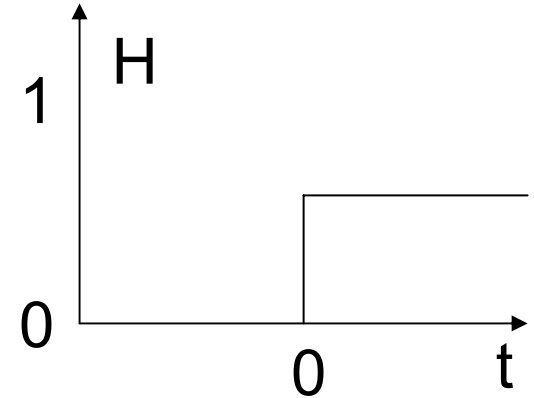
$$\delta(x) = \lim_{a \rightarrow \infty} \frac{1}{\pi x} \sin(x/a)$$



delta Diraca = nachylenie funkcji Heavyside'a

$$\delta(x - a) = H'(x - a)$$

$$H(t) = \begin{cases} 1; & \text{dla } t > 0 \\ 0; & \text{dla innych } t \end{cases}$$



---

dwa i więcej wymiary

$$\delta(x, y) = \delta(x)\delta(y)$$

D)

Wagi: rozwiązanie chcemy znać tylko w chwilach  $t_n$  - wagi powinny je wyłuskać

$$w_j(t) = \delta(t - t_j) \quad \longleftarrow \text{delta Diraca}$$

$$\int_0^T \left[ \sum_{i=0}^N c_i l'_i(t) - \sum_{i=0}^N f_i l_i(t) \right] w_j(t) dt = 0$$

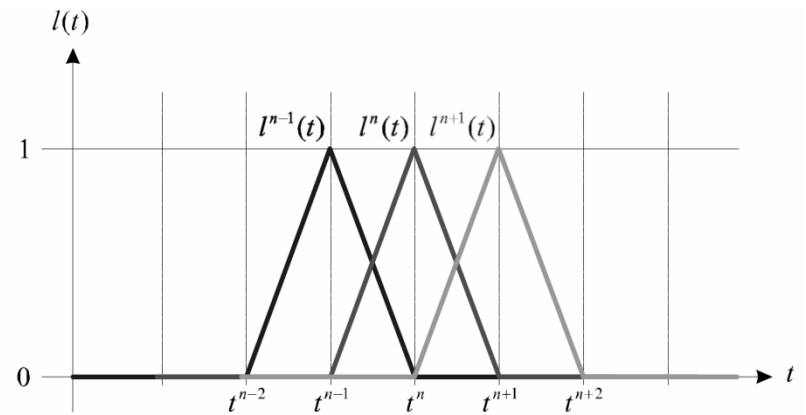
$$\sum_{i=0}^N (c_i l'_i(t_j) - f_i l_i(t_j)) = 0$$

$$\sum_{i=0}^N c_i l'_i(t_j) = f_j$$

$$l'_i(t_j) = ?$$

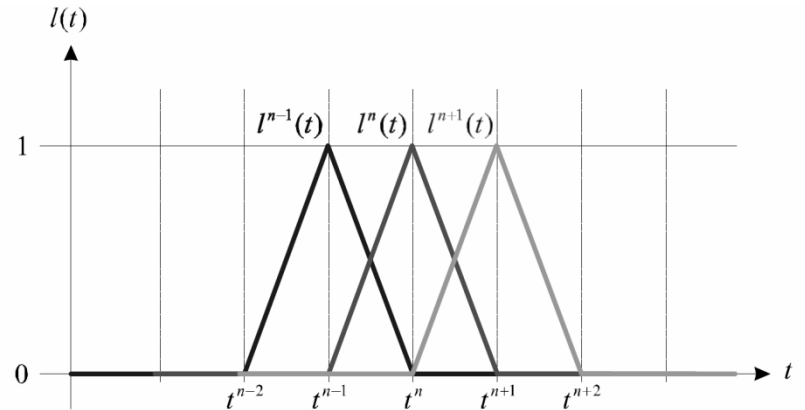
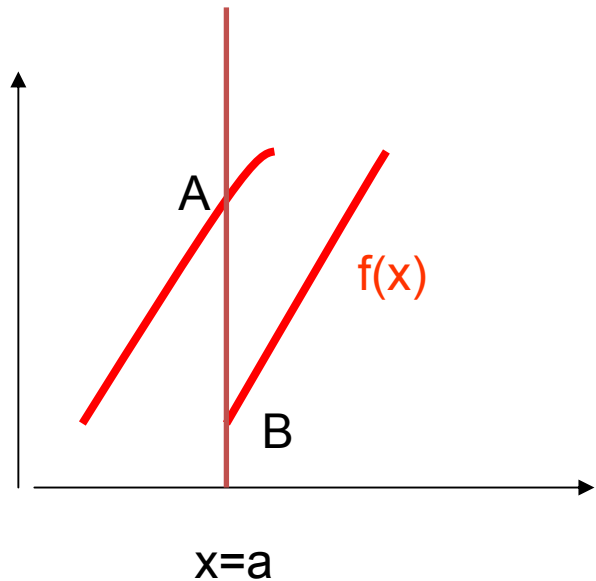
$$l'_i(t_j) = ?$$

$l(t)$  ... nie jest różniczkowalna w punktach węzłowych...



$$l'_i(t_j) = ?$$

$l(t)$  nie jest różniczkowalna w punktach węzłowych...  
 $l'$  ma nieciągłą pochodną ...



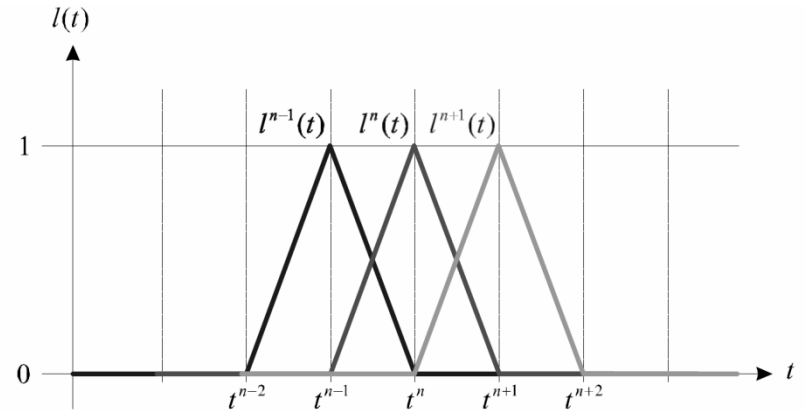
$$\int_{-\infty}^{\infty} dx f(x) \delta(x - a) = \frac{A + B}{2}$$

$$l'_i(t_j) = \left( \lim_{t \rightarrow t_j^+} l'_i(t) + \lim_{t \rightarrow t_j^-} l'_i(t) \right) / 2 \quad (\text{tzw. główna wartość pochodnej})$$

$$l'_n(t_j) = ?$$

$$l'_i(t_j) = \left( \lim_{t \rightarrow t_j^+} l'_i(t) + \lim_{t \rightarrow t_j^-} l'_i(t) \right) / 2$$

$$l_n(t) = \begin{cases} 0 & t \leq t_{n-1} \\ \frac{t-t_{n-1}}{\Delta t} & t_{n-1} \leq t \leq t_n \\ \frac{t_{n+1}-t}{\Delta t} & t_n \leq t \leq t_{n+1} \\ 0 & t_{n+1} \leq t \end{cases}$$



główna wartość pochodnej (średnia z lewo i prawostronnej pochodnej)

$$l'_n(t_j) = \begin{cases} -\frac{1}{2\Delta t} & n = j - 1 \\ 0 & n \neq j \pm 1 \\ \frac{1}{2\Delta t} & n = j + 1 \end{cases}$$



$$l'_n(t_j) = \begin{cases} -\frac{1}{2\Delta t} & n = j - 1 \\ 0 & n \neq j \pm 1 \\ \frac{1}{2\Delta t} & n = j + 1 \end{cases} \longrightarrow l'_n(t_j) = \frac{1}{2\Delta t} (\delta_{n,j+1} - \delta_{n,j-1})$$

$$\sum_{i=0}^N c_i l'_i(t_j) = f_j$$

$$\frac{1}{2\Delta t} \sum_{i=0}^N c_i (\delta_{i,j+1} - \delta_{i,j-1}) = f_j$$

$$\frac{1}{2\Delta t} (c_{j+1} - c_{j-1}) = f_j \xrightarrow{c_j = y_j} y_{j+1} = y_{j-1} + 2\Delta t f_j$$

Metoda różnic skończonych jest przypadkiem szczególnym: metody reszt ważonych dla odcinkowo liniowej bazy i funkcji wagowych typu delta Diraca

# w stronę metody elementów skończonych

metoda ważonych reszt - ogólnie

$Lu=f$  (na  $\Omega$ )  $\longrightarrow$  Rozwiązanie dokładne (silnej postaci równania)  
jest „trudne”.

$Bu=g$  (na  $d\Omega$ ) szukamy rozwiązania przybliżonego w bazie funkcji

$$\tilde{u} = \sum_{i=1}^N c_i v_i(x) \quad \text{(rozwiązanie w podprzestrzeni wektorowej rozpiętej przez wektory bazy)}$$

Działając operatorami  $L$  i  $B$  na rozwiązanie przybliżone dostajemy funkcje reszkowe (rezydualne) zamiast zera:

$$\begin{aligned} L\tilde{u} - f &= r \\ B\tilde{u} - g &= s \end{aligned} \longrightarrow \text{zależy nam, aby reszty } r \text{ i } s \text{ były jak najmniejsze}$$

$c$  wyznaczamy z ważenia reszty:

$$\int_V r(x) w_j(x) dx = 0$$

dla metody Galerkina bierzemy funkcje bazowe jako wagi:  $w_j = v_j$

### **Silna forma równania:**

$$Lu=f \quad (\text{równość funkcji w każdym punkcie} \quad \tilde{u} = \sum_{i=1}^N c_i v_i(x) \\ \text{obszaru całkowania})$$

ważone reszty:

$$(L\tilde{u} - f, w_j) = (r, w_j) = 0$$

$$(L\tilde{u}, w_j) = (f, w_j) \quad \text{\textbf{słaba forma równania,}} \\ \text{(równość } N \text{ liczb)}$$

# żargon MES: macierz sztywności i wektor obciążeń

$$u = \sum_{i=1}^n y_i v_i$$
$$(Lu, v_j) = (f, v_j)$$
$$\sum_{i=1}^n (Lv_i, v_j) y_i = (f, v_j)$$

$$\mathbf{S}_{ij} = (Lv_i, v_j)$$

$$\mathbf{F}_j = (f, v_j)$$

$$\mathbf{SY} = \mathbf{F}$$

*stiffness matrix*  
macierz sztywności

*load vector*  
wektor obciążeń

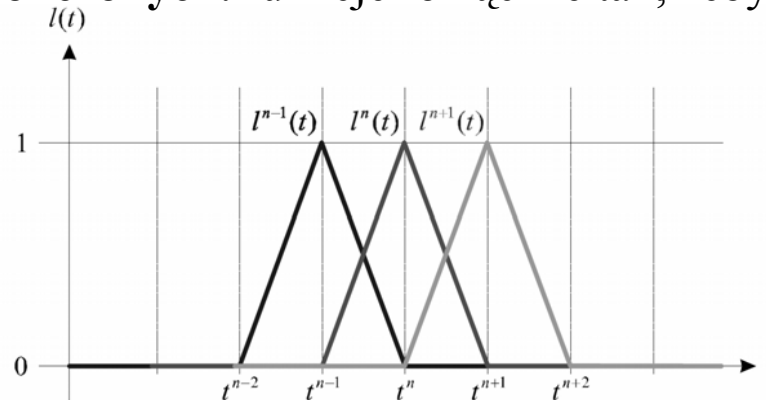
powyższy przykład: baza wielomianów określonych na całym pudle obliczeniowym.  
Z wielu powodów jest to zły pomysł.

Wysokie potęgi wielomianów niewygodne w użyciu: całkowanie, efekt Rungego,  
powód najważniejszy:  
macierz **S** byłaby gęsta, problem nie do rozwiązania przy dużym **N**.

$$\mathbf{SY}=\mathbf{F}$$

Galerkin z bazą odcinkami wielomianowych funkcji zdefiniowanych w sposób rozłączny przestrzennie → metoda elementów skończonych

**Metoda elementów skończonych:** funkcje rozłączne tak, żeby **S** = rzadka



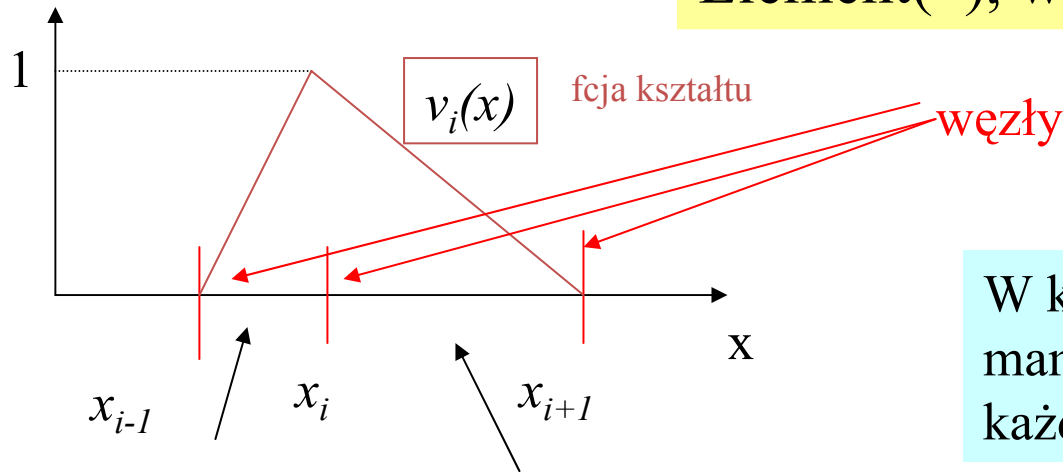
najprostszy wybór *funkcji kształtu*(\*): baza funkcji odcinkami liniowych  
zbieżność dostaniemy w przestrzeni funkcji odcinkami liniowych

(\*) trzecie pojęcie z żargonu MES

Zobaczymy w działaniu metodę elementów skończonych,  
ale na razie: bez jej charakterystycznych narzędzi:  
bez lokalnych macierzy sztywności związanych z każdym elementem  
bez ich składania do macierzy globalnej  
bez mapowania przestrzeni fizycznej do przestrzeni referencyjnej

będziemy mówili o metodzie z punktu widzenia węzłów:  
tak najłatwiej wprowadzić metodę, ale dla 2D i 3D takie podejście okazuje się niepraktyczne  
podejście związane z elementami zobaczymy później

# Element(\*), węzły(\*), funkcje kształtu



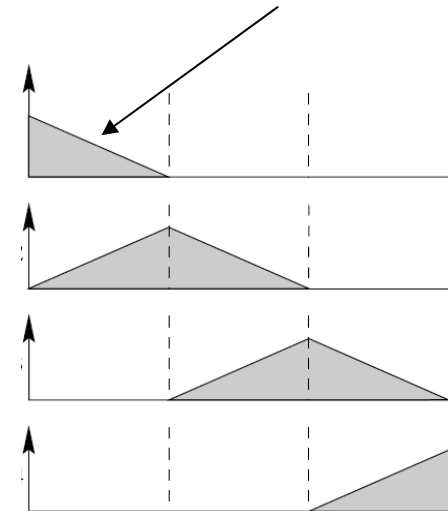
W każdym elemencie:  
mamy 2 funkcje,  
każda z innym węzłem związana

**element**  $K_i$  długości  
 $h_i = x_i - x_{i-1}$

**element**  $K_{i+1}$  długości  
 $h_{i+1} = x_{i+1} - x_i$

$$v_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}} & x \in K_i \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & x \in K_{i+1} \\ 0 & x \notin K_i \cup K_{i+1} \end{cases}$$

funkcje bazowe i brzeg



$$u = \sum_{i=1}^n y_i v_i$$

Dla (jednorodnych) warunków Dirichleta  
mamy  
 $y_{\text{pierwsze}} = y_{\text{ostatnie}} = 0$

$$v_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}} & x \in K_i \\ \frac{x_{i+1}-x}{x_{i+1}-x_i} & x \in K_{i+1} \\ 0 & x \notin K_i \cup K_{i+1} \end{cases}$$

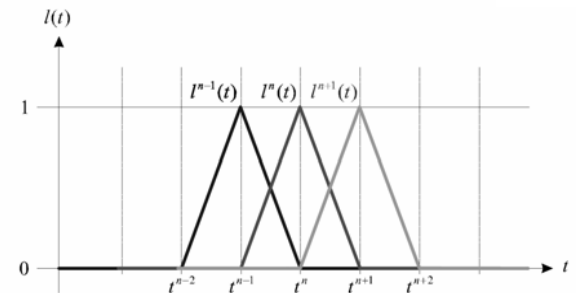
$$\frac{d^2 u}{dx^2} = -\sin(\pi x)$$

$$v'_i(x) = \begin{cases} \frac{1}{x_i-x_{i-1}} & x \in K_i \\ -\frac{1}{x_{i+1}-x_i} & x \in K_{i+1} \\ 0 & x \notin K_i \cup K_{i+1} \end{cases}$$

$S_{ij} = (Lv_i, v_j)$  niezerowe tylko dla  $i=j, i=j-1$  oraz  $i=j+1$  [bez przekrywania całka znika]

$$S_{ii} = \cancel{v'_i(x)v_i(x)} \Big|_{x_{i-1}}^{x_{i+1}} - \int_{x_{i-1}}^{x_{i+1}} dx v'_i(x)v'_i(x) dx$$

$$S_{ii} = -\left(\frac{1}{h_i} + \frac{1}{h_{i+1}}\right)$$

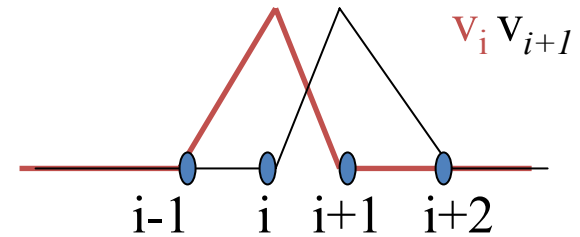




$$S_{ij} = (Lv_i, v_j) \quad \text{niech } j = i+1 \quad v'_i(x) = \begin{cases} \frac{1}{x_i - x_{i-1}} & x \in K_i \\ -\frac{1}{x_{i+1} - x_i} & x \in K_{i+1} \\ 0 & x \notin K_i \cup K_{i+1} \end{cases}$$

$$S_{i,i+1} = + v'_i(x) v_{i+1}(x) \Big|_{x_{i-1}}^{x_{i+2}} - \int_{x_{i-1}}^{x_{i+2}} dx v'_i(x) v'_{i+1}(x) dx$$

$$S_{i,i+1} = - \int_{x_i}^{x_{i+1}} dx v'_i(x) v'_{i+1}(x) dx$$



gdy jedna pochodna  
dodatnia druga ujemna

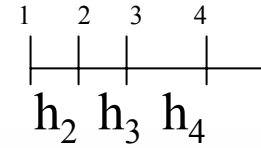
$$S_{i,i+1} = - \left( -\frac{1}{x_{i+1} - x_i} \times \frac{1}{x_{i+1} - x_i} \times (x_{i+1} - x_i) \right) = \frac{1}{h_{i+1}}$$

$$S_{i,i-1} = \frac{1}{h_i} \quad \leftarrow \text{długość elementu o numerze większym z dwóch indeksów } S$$

# SY=F

Macierz sztywności dla  $n$  węzłów

+ warunek  $y_1=y_n=0$



$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{h_2} & -(\frac{1}{h_2} + \frac{1}{h_3}) & \frac{1}{h_3} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{h_3} & -(\frac{1}{h_2} + \frac{1}{h_3}) & \frac{1}{h_4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \frac{1}{h_{n-1}} & -(\frac{1}{h_{n-1}} + \frac{1}{h_n}) & \frac{1}{h_n} \\ 0 & \dots & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

wiersz  $n-1 \rightarrow$

$$F_i = (v_i f)$$

$$F_i = \int_{x_{i-1}}^{x_{i+1}} v_i(x) f(x) dx$$

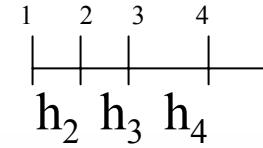
$$F_i = \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} f(x) dx + \int_{x_i}^{x_{i+1}} \frac{-x + x_{i+1}}{x_{i+1} - x_i} f(x) dx$$

po elemencie  $K_i$                       po  $K_{i+1}$

# SY=F

Macierz sztywności dla  $n$  węzłów

+ warunek  $y_1=y_n=0$



wiersz  $n-1 \rightarrow$

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{h_2} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \frac{1}{h_3} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{h_3} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \frac{1}{h_4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \frac{1}{h_{n-1}} & -\left(\frac{1}{h_{n-1}} + \frac{1}{h_n}\right) & \frac{1}{h_n} \\ 0 & \dots & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$F_i = (v_i f)$$

$$F_i = \int_{x_{i-1}}^{x_{i+1}} v_i(x) f(x) dx$$

$$F_i = \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} f(x) dx + \int_{x_i}^{x_{i+1}} \frac{-x + x_{i+1}}{x_{i+1} - x_i} f(x) dx$$

po elemencie  $K_i$

po  $K_{i+1}$

dla równoodległych węzłów  $S$  jak macierz metody RS (razy  $h=dx$ ),  
ale wektor obciążeń  $F$  – nie! w MRS mielibyśmy  $F_i=f(x_i) dx$

$$F_i = \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} f(x) dx + \int_{x_i}^{x_{i+1}} \frac{-x + x_{i+1}}{x_{i+1} - x_i} f(x) dx$$

dla  $f(x) = -\sin(\pi x)$

$$F_i = -\frac{\sin(x_i \pi)}{\pi^2(x_i - x_{i-1})} + \frac{\sin(x_{i-1} \pi)}{\pi^2(x_i - x_{i-1})} - \frac{\sin(x_{i+1} \pi)}{\pi^2(x_i - x_{i+1})} + \frac{\sin(x_i \pi)}{\pi^2(x_i - x_{i+1})}$$

warunki brzegowe (jednorodne Dirichleta): forma  $S$  oraz  $F_1 = F_n = 0$

ten URL wygląda prawie jak dla MRS...  
zobaczmy wyniki

$$\mathbf{SY}=\mathbf{F}$$

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{h_2} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \frac{1}{h_3} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{h_3} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \frac{1}{h_4} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \frac{1}{h_{n-1}} & -\left(\frac{1}{h_{n-1}} + \frac{1}{h_n}\right) & \frac{1}{h_n} \\ 0 & \dots & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Układ równań z macierzą trójprzekątniową – przypomnienie.  
 Jak rozwiązać?

# Dekompozycja LU macierzy trójkątnej

$$SY=F$$

$$S=LU \quad (LU - \text{trójkątne})$$

$$(LU)Y=F$$

$$UY=x$$

$Lx=F$  - najpierw rozwiązujemy ten układ

$$S = \begin{pmatrix} a_1 & c_1 & 0 & 0 & 0 & \dots & 0 \\ b_2 & a_2 & c_2 & 0 & 0 & \dots & 0 \\ 0 & b_3 & a_3 & c_3 & 0 & \dots & 0 \\ & & & & & \dots & \\ 0 & 0 & \dots & 0 & 0 & b_n & a_n \end{pmatrix}$$

dwuprzekątne

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ \beta_2 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \beta_3 & 1 & 0 & 0 & \dots & 0 \\ & & & \dots & & & \\ 0 & 0 & 0 & \dots & 0 & \beta_n & 1 \end{pmatrix}$$

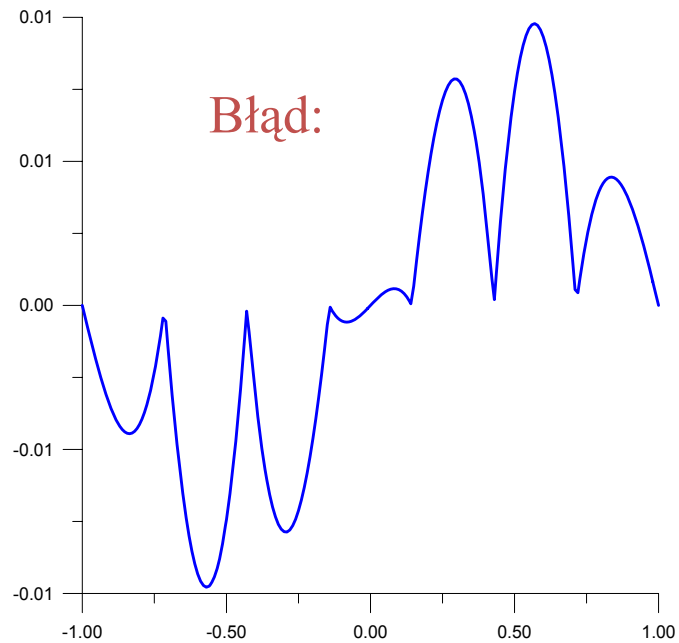
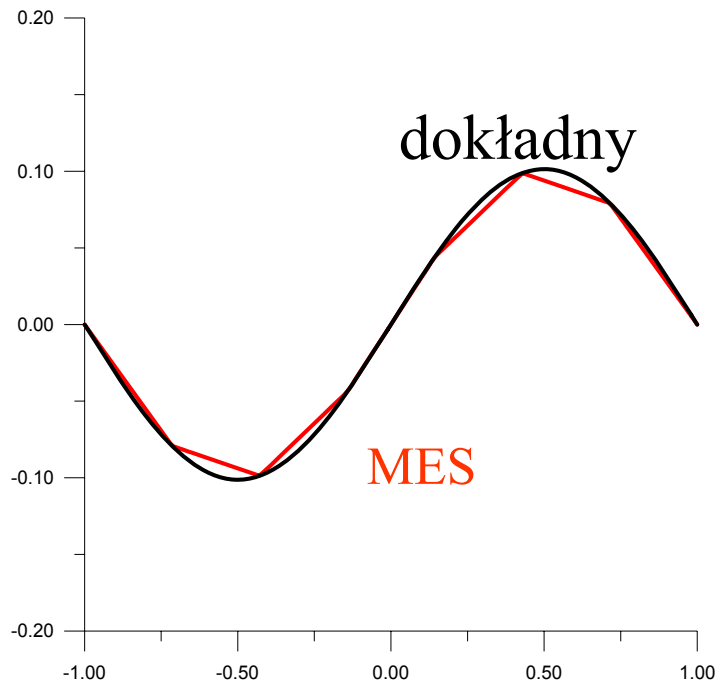
$$U = \begin{pmatrix} \alpha_1 & c_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & c_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \alpha_3 & c_3 & 0 & \dots & 0 \\ & & & \dots & & & \\ 0 & 0 & 0 & \dots & 0 & \alpha_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & \alpha_n \end{pmatrix}$$

bez zmian

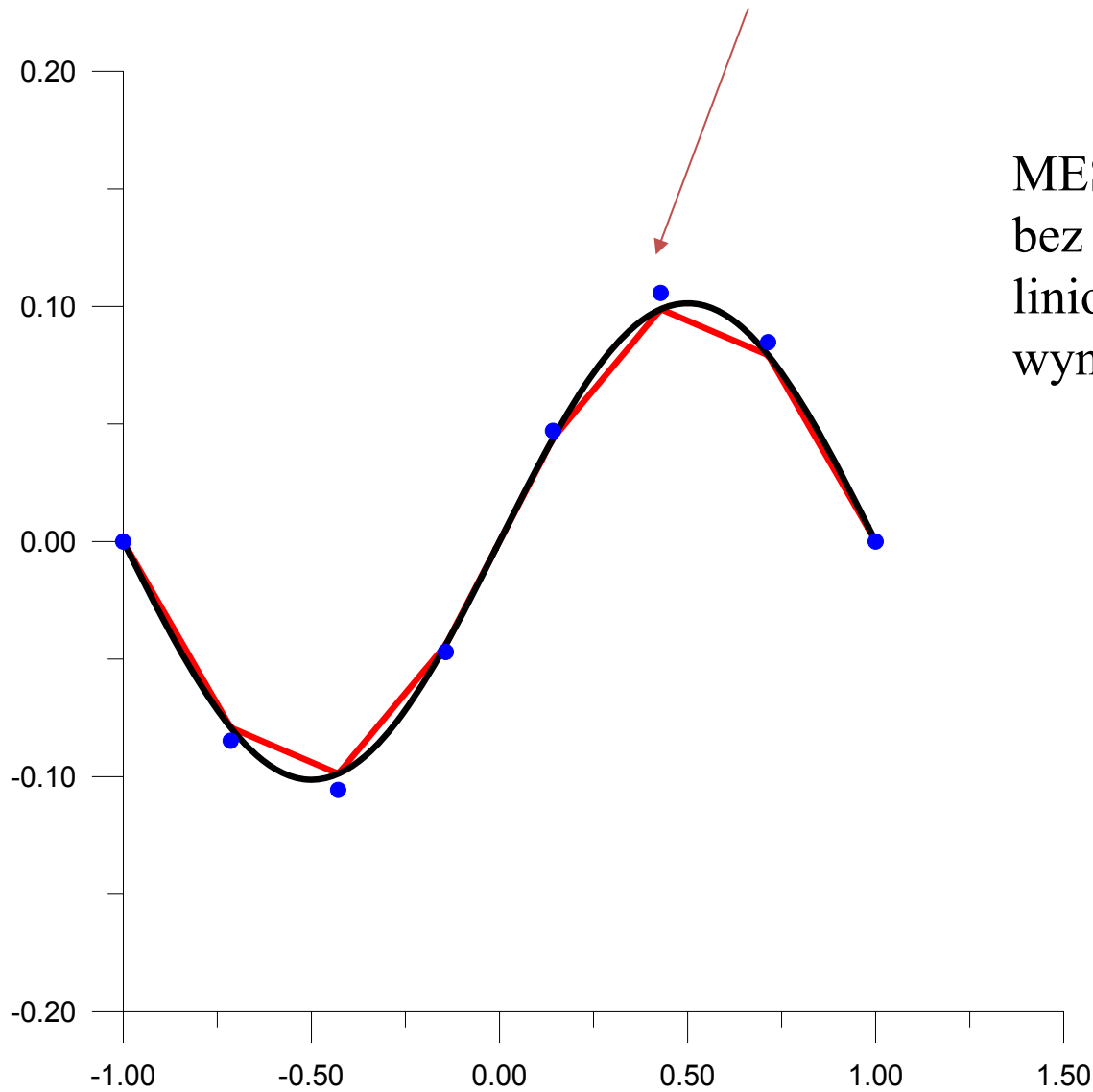
$$\alpha_1 = a_1 \quad \left| \quad \begin{array}{l} \beta_i = \frac{b_i}{\alpha_{i-1}} \\ \alpha_i = a_i - \beta_i c_{i-1} \end{array} \right. \quad \text{dla } i > 1$$

Wynik: równoodległe węzły

$$\frac{d^2u}{dx^2} = -\sin(\pi x)$$

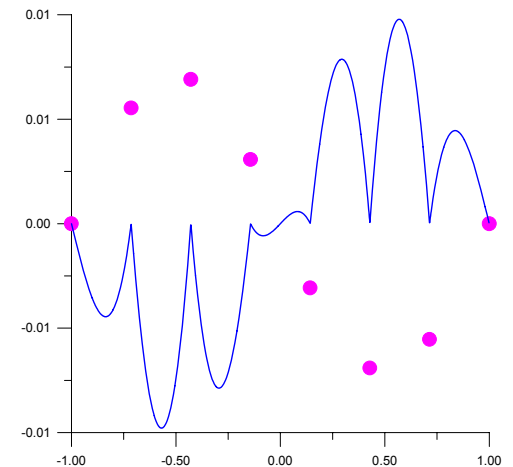


MES (równoodległe węzły) a **MRS** (węzły w tych samych punktach):



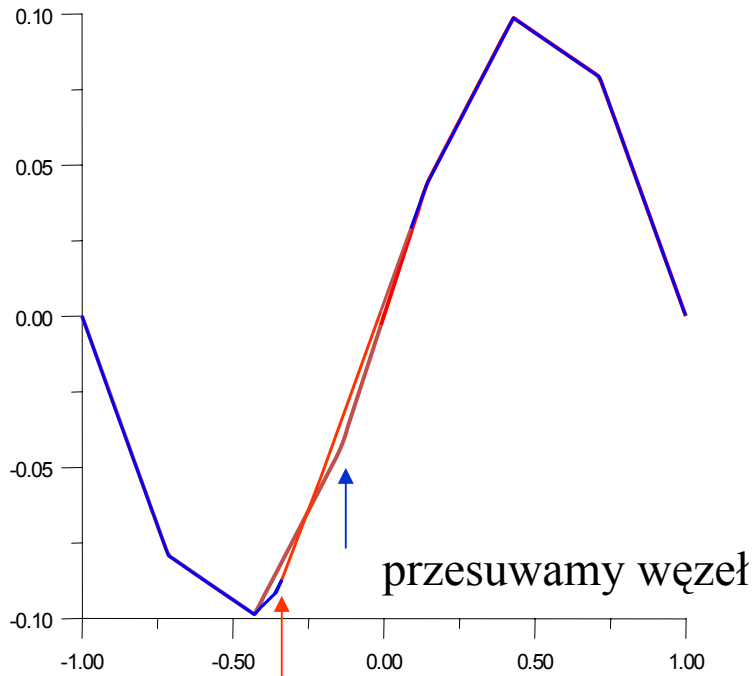
MES dla laplasjanu  
bez pochodnej z funkcjami  
liniowymi: w węzłach  
wynik dokładny !!!

błąd **MRS** i **MES**

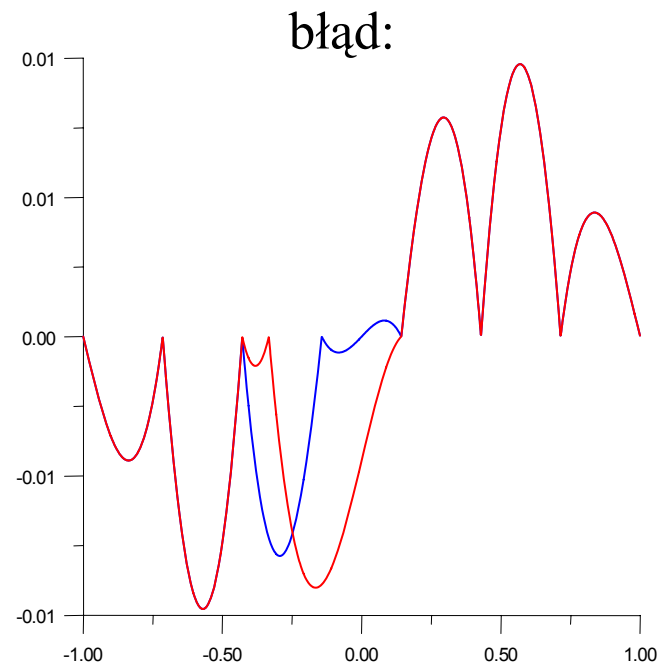


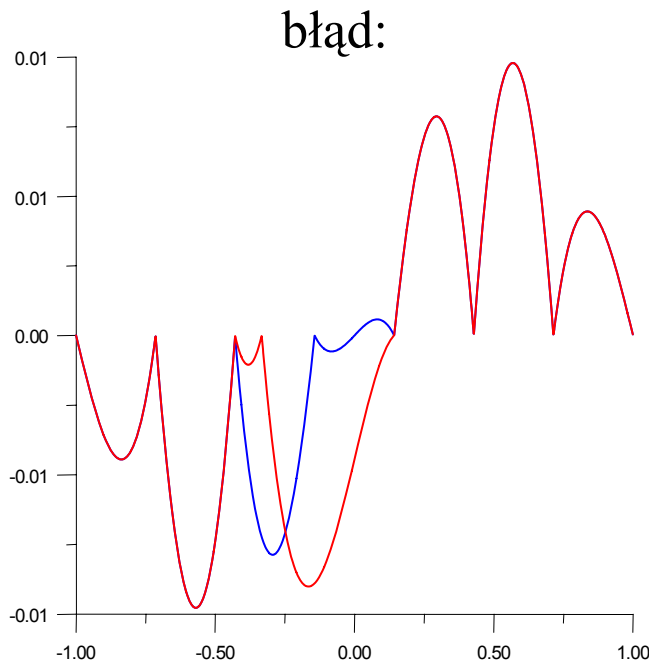


znikanie błędu MES (1D, liniowe f.kształtu) w węzłach  
zachodzi również dla nierównomiernego rozkładu węzłów:



Dla MRS: dla nierównomiernej siatki  
musielibyśmy używać  
niesymetrycznych ilorazów o  
[jak widzieliśmy] niższej dokładności





Równanie Poissona,  
funkcje kształtu liniowe  
wynik MES **dokładny** w węzłach

MES: produkuje oszacowanie  
wyniku również między węzłami

MRS: tylko w węzłach

MRS: wartości w węzłach,  
są dokładne **TYLKO**  
w granicy  $\Delta x \rightarrow 0$

dowód dokładności MES w tej wersji - za parę folii