

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

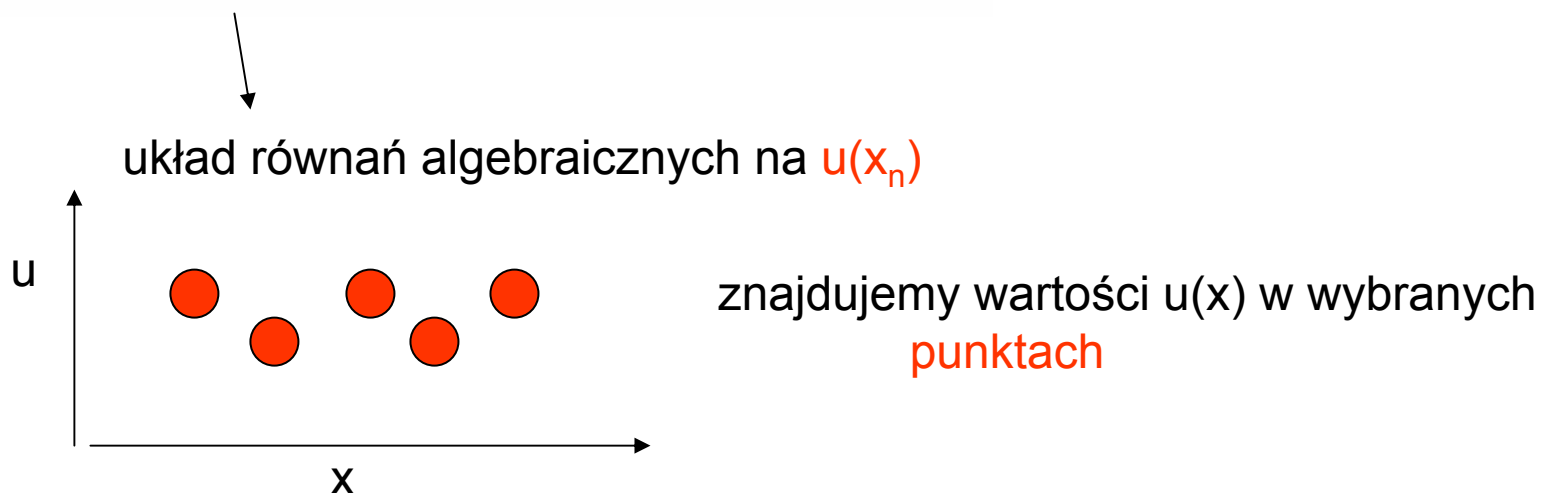
| do metody elementów
skończonych

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

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

metoda różnic skończonych:

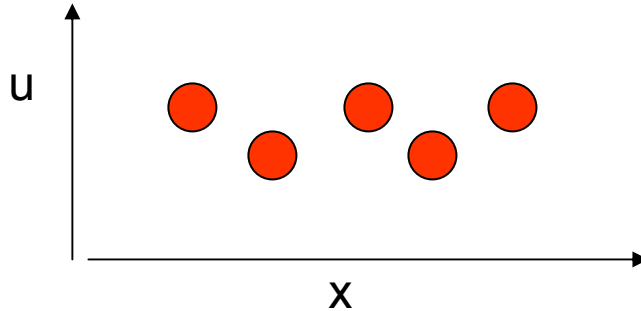
$$\frac{d^2 u}{dx^2} \simeq \frac{u(x+dx) + u(x-dx) - 2u(x)}{dx^2}$$



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

metoda różnic skończonych

układ równań algebraicznych na $u(x_n)$



znajdujemy wartości $u(x)$ w wybranych punktach

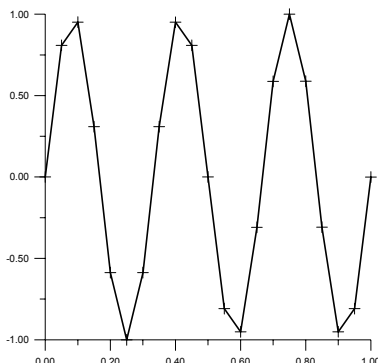
Główna (jedyna) zaleta MRS: prosta dyskretyzacja równań.

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

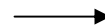
: niełatwy opis objętości o konturze odbiegającym od prostokątnego

: 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 ☹

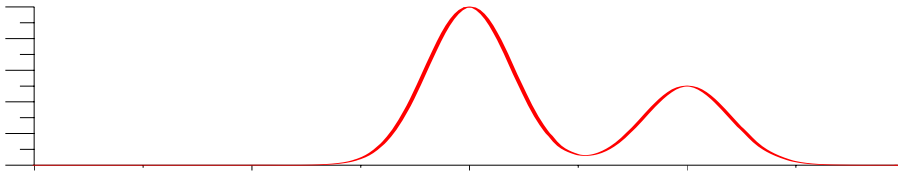


użyć bazy funkcyjnej i
 $\sin(6\pi x)$ włączyć do bazy funkcyjnej
w której poszukujemy rozwiązania...

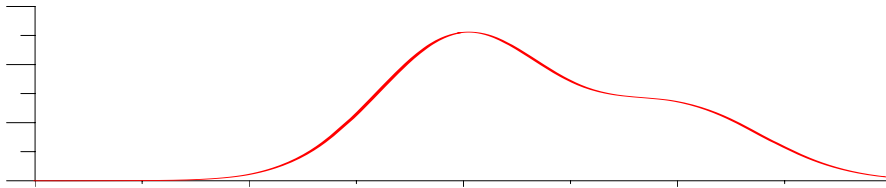
motywacja do pracy z bazą funkcyjną cd.



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 $\sin(nx)$

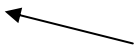
- rozwiązanie w takiej bazie da nam automatycznie dyskretną transformatę Fouriera rozwiązania

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

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

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 rozpin

optymalne rozwiązanie znaczy optymalne współczynniki c_i

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

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 rozpiną

o $v(x)$ – funkcja próbna
 (test function, trial function)

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

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

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 $v(x)$:

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

$$u(-1)=0$$

$$u(1)=0$$

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: zawęźa przestrzeń poszukiwań optymalnego rozwiązania

optymalne rozwiązanie znaczy optymalne współczynniki c_i

błąd rozwiązania przybliżonego $v(x)$:

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

błąd, reszta, residuum

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

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

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

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 $[i=1,2,3]$

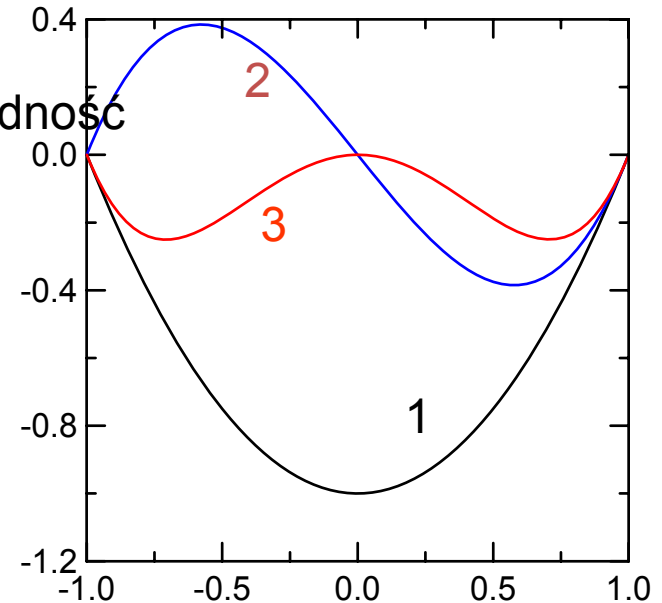
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

(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

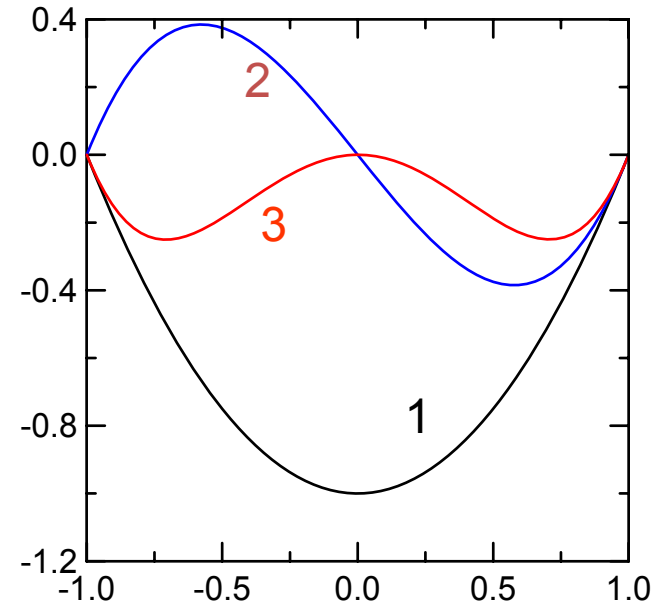
wektor c dany przez warunek $E(x_i)=0$

$E(x_j)=0$ – 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

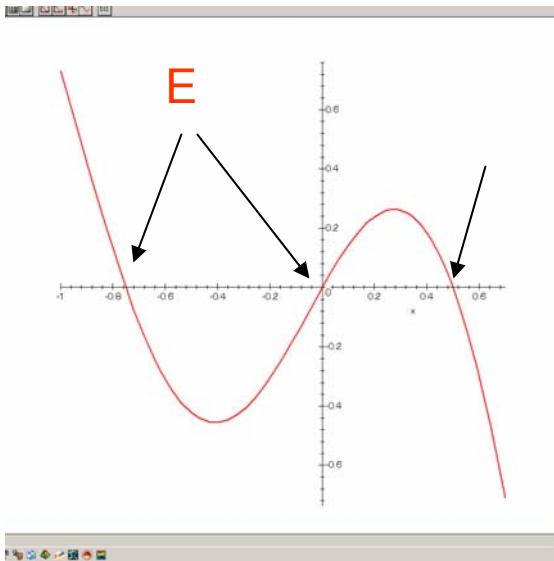
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

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

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

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

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



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

metoda kolokacji

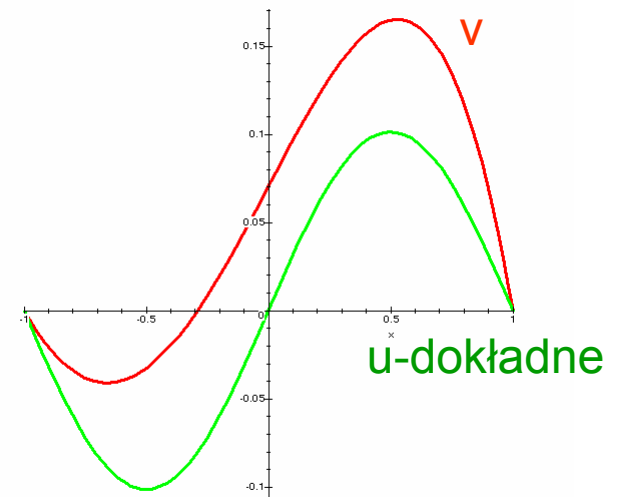
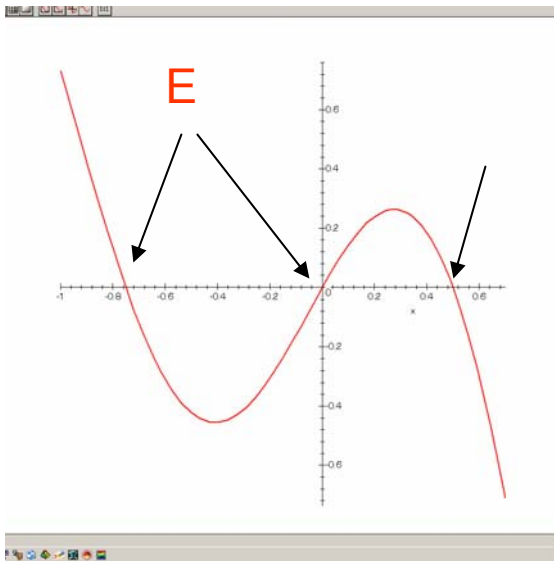
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

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

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

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

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



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

metoda kolokacji

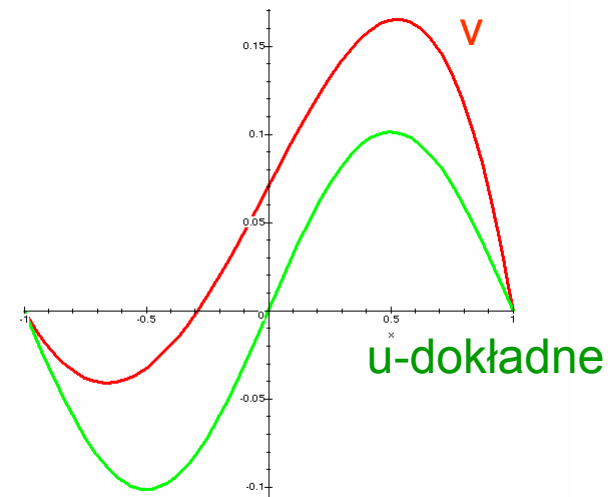
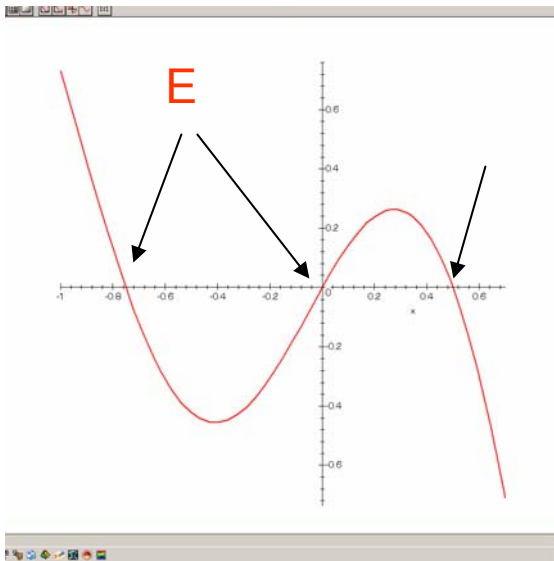
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

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

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

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

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



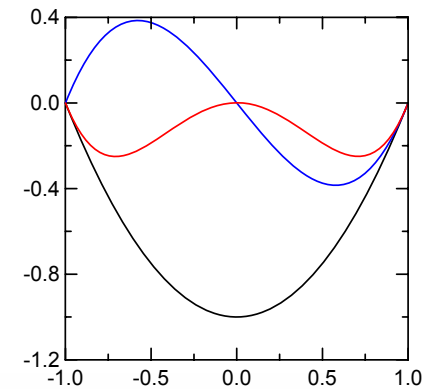
Uwaga: $E(x_a)=0$ NIE znaczy $v(x_a)=u(x_a)$ bo błąd to nie jest odchylenie od wartości dokładnej. W naszym równaniu $E(x_a)=0$ znaczy: $v''(x_a)=u''(x_a)$

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

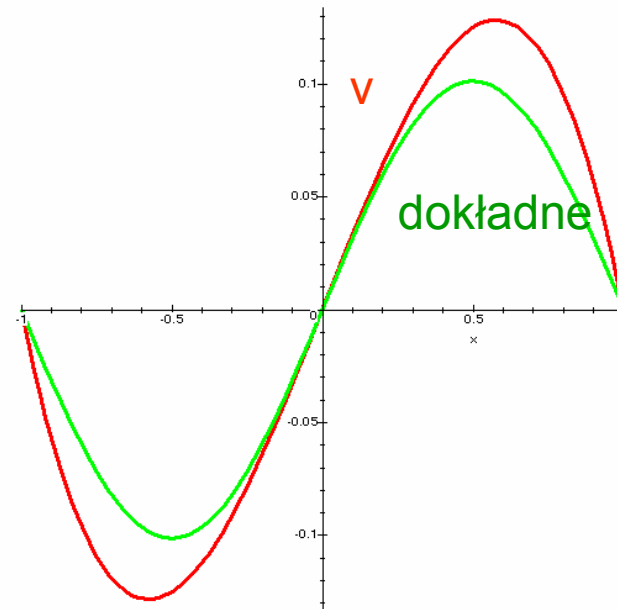
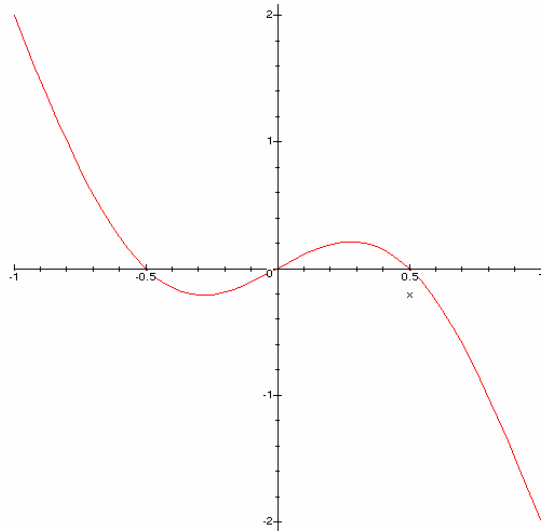
metoda kolokacji

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]

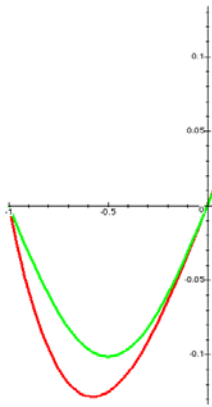


lepiej niż poprzednio, mimo że
tylko jedna funkcja bazowa pracuje

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

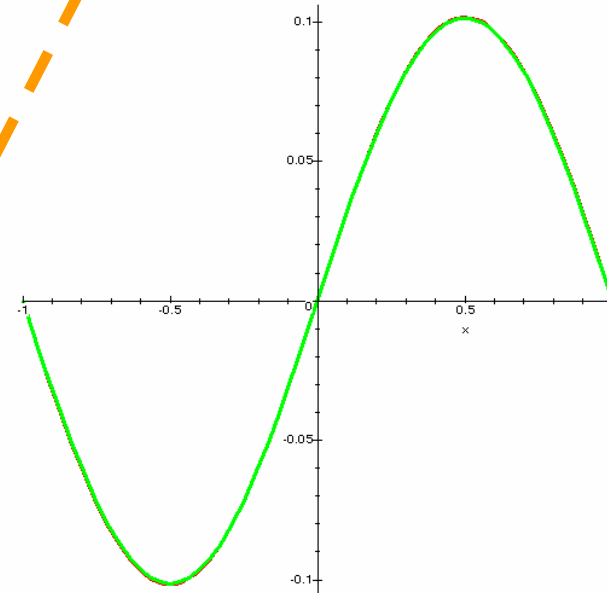
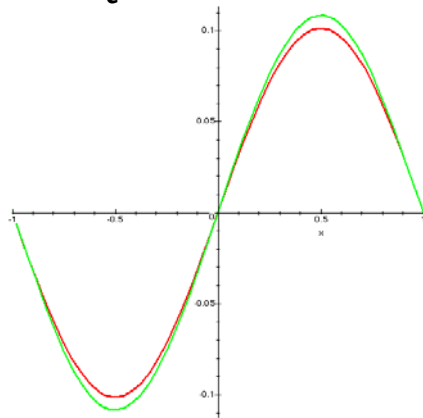
$$E(x) = 2c_1 + 6c_2x + 12c_3x^2 - 2c_3 + \sin(\pi x)$$

metoda kolokacji



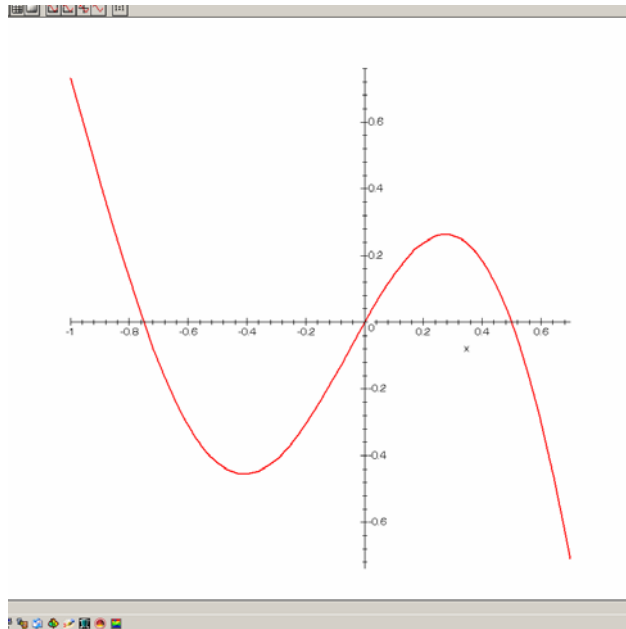
poprzedni wybór:
jedna funkcja bazowa
 f_2 oraz punkt kolokacji $\frac{1}{2}$
[symetria gwarantuje
również znikanie błędu E
w 0 i $-\frac{1}{2}$]

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



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

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

dostaniemy znowu układ równań liniowych

- 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]

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 $[i=1,2,3]$

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

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

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

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}) ;

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

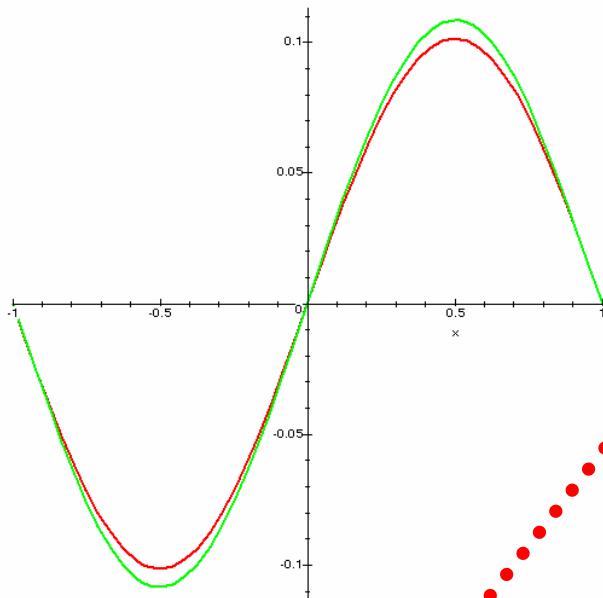
odrzucone funkcje bazowe o złej symetrii

dwie funkcje bazowe: f_2 oraz f_4

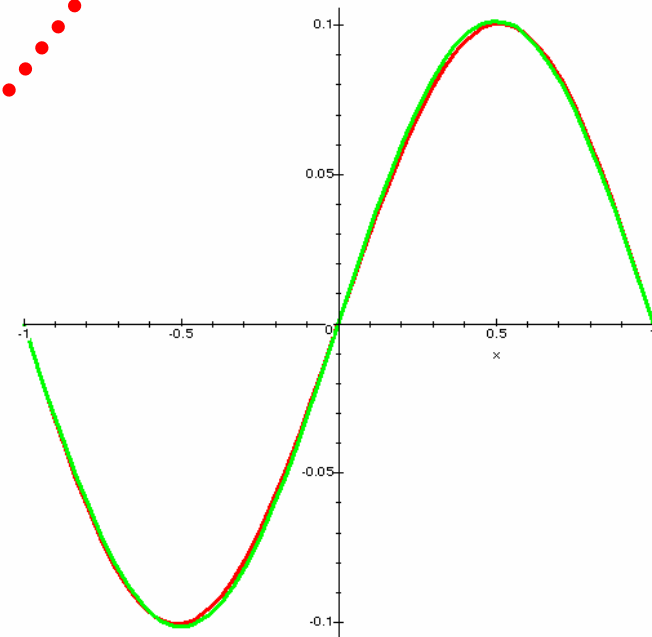
punkty kolokacji: $\frac{1}{2}$ oraz $\frac{3}{4}$

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

min ☐



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 Galerkina: $w_j = f_j$ (wagi tożsame z funkcjami bazowymi)

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

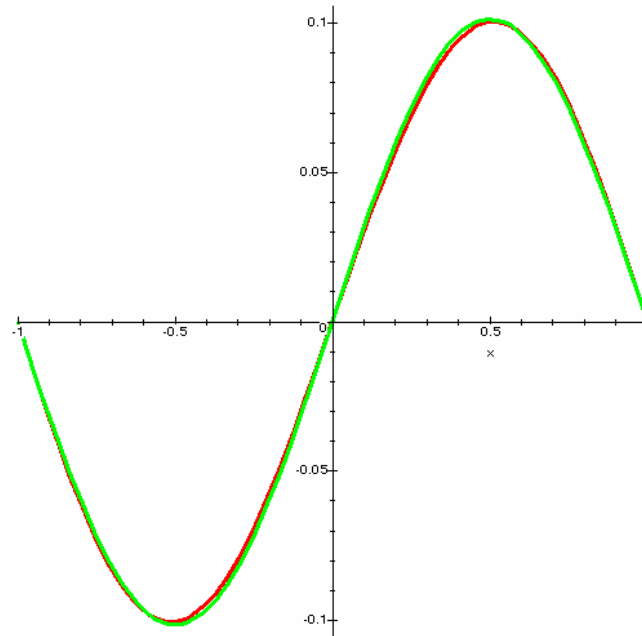
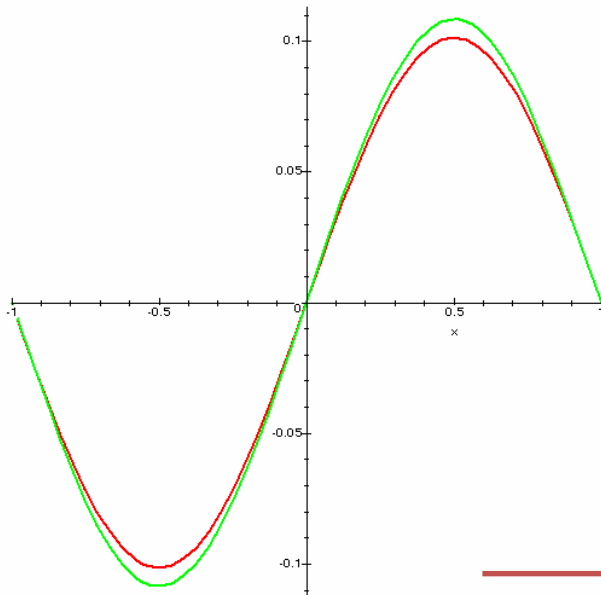
$$f_i(x) = (x+1)(x-1)x^{i-1}$$

$R_j =$ układ równań na c_j

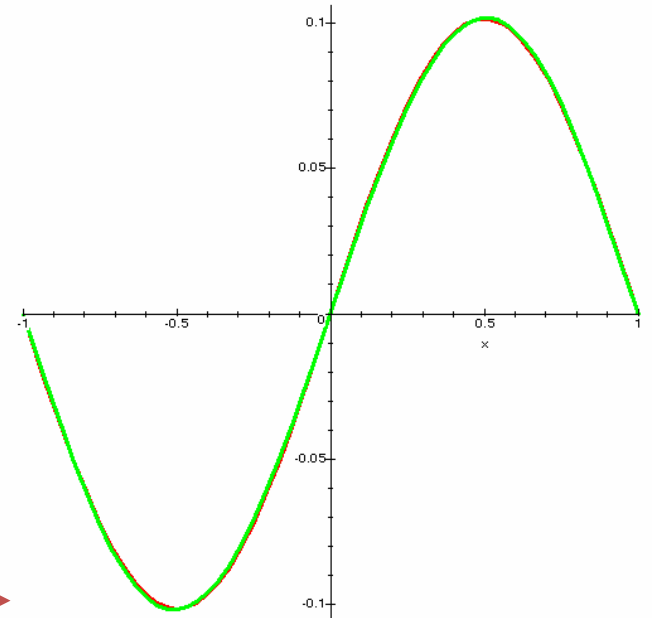
baza 2 funkcji f_2, f_4
porównanie metod

min ☐

Kolokacja



Galerkin



jakość rozwiązania



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 \mathbf{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 bazowy

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

słaba forma równania

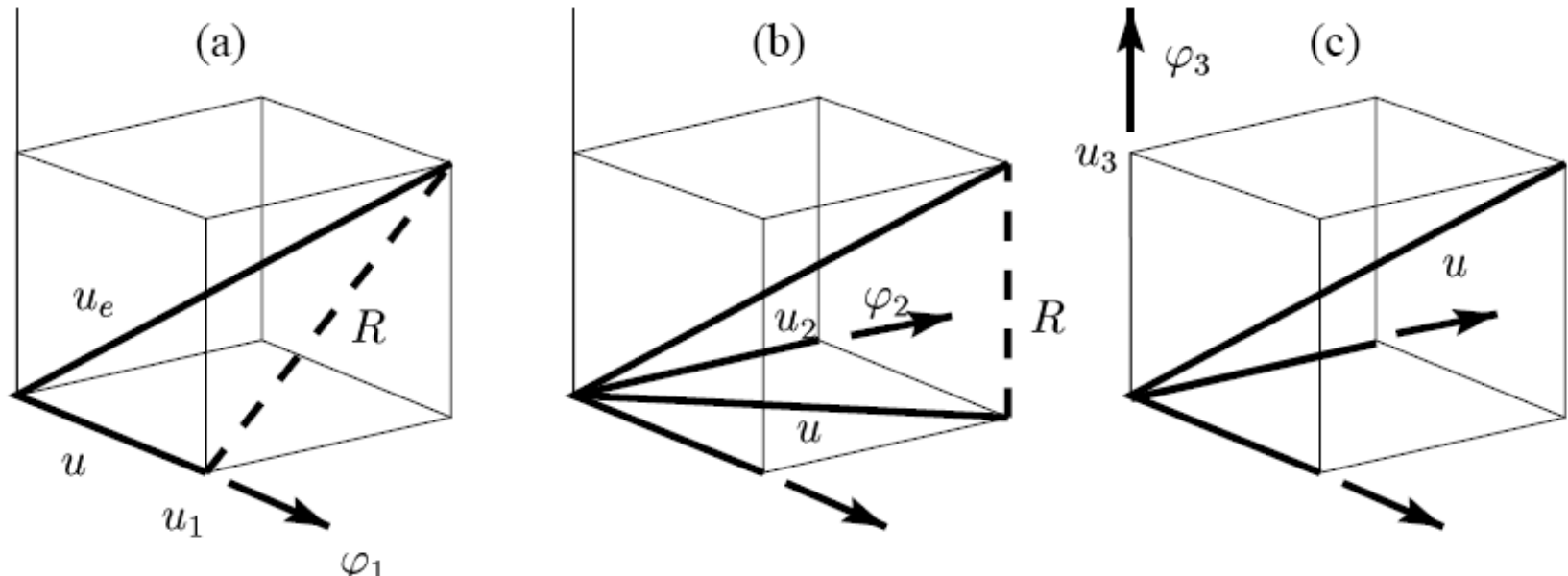
błąd $E=Au-f$: ortogonalny do każdego wektora bazowego

$$(E, v_j) = 0$$

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 ϕ_1, ϕ_2, ϕ_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

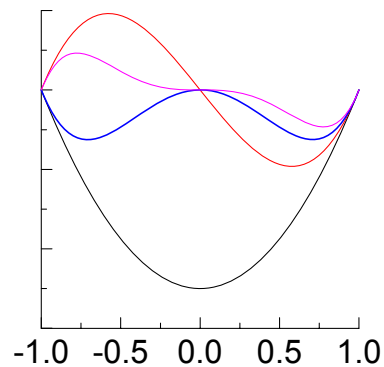
Przykład: z laboratorium

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

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

baza spełniająca Dirichleta

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



Galerkin

baza spełniająca warunki Dirichleta

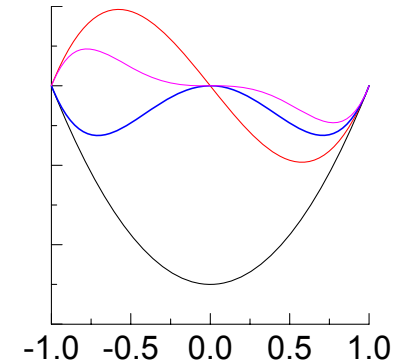
$$\frac{d^2 u}{dx^2} = -\sin(\pi x)$$

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

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

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

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



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

całkowanie przez części

z warunków
brzegowych

dla i oraz j tej samej parzystości

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

prawa strona:

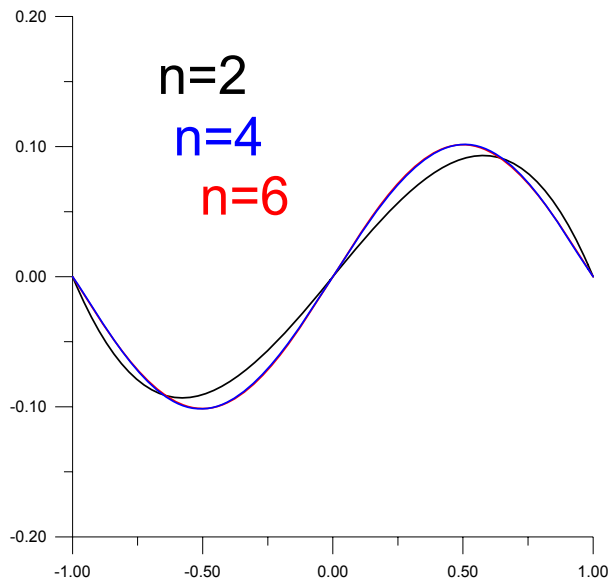
$$I(k) = - \int_{-1}^1 dx \sin(\pi x) x^k$$
$$I(k) = -\frac{2}{\pi} - \frac{k(k-1)}{\pi^2} I(k-2)$$

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

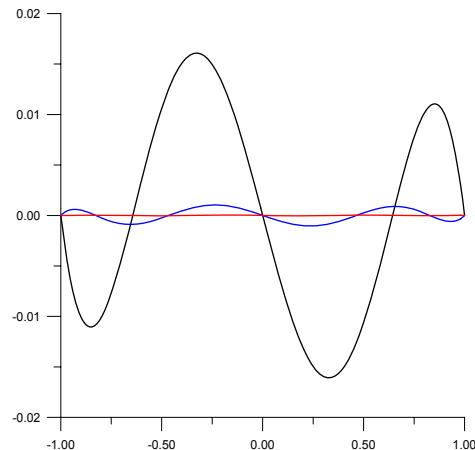
$$F(j) = I(j+1) - I(j-1)$$

dla j nieparzystych

rozwiązanie



błąd ε (**nie residuum** tylko różnica dokładne – Galerkina):

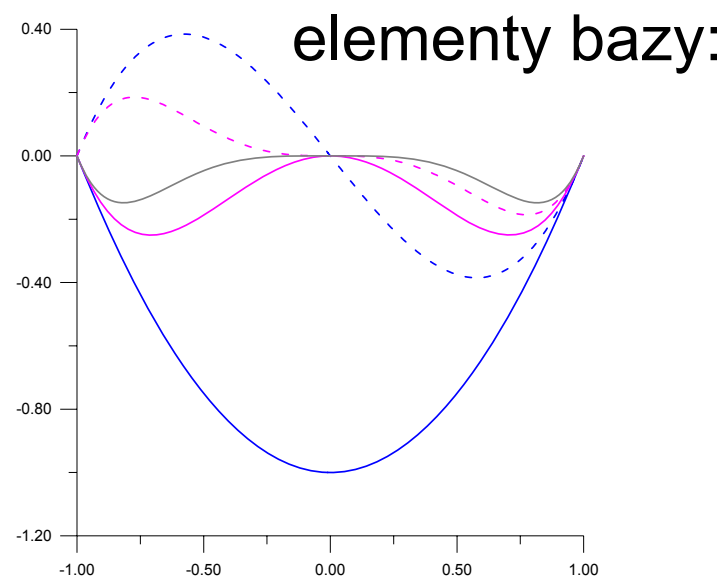
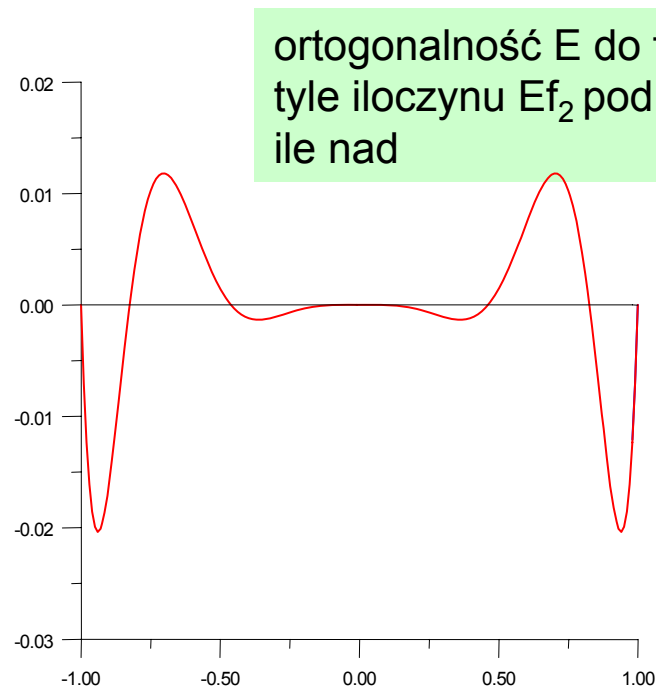


ortogonalność residuum do wektorów bazowych

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

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

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 Galerkina - 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 Ω , z jednorodnym warunkiem brzegowym $u=0$ na brzegu Γ ,
 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

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

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

$$Au=f$$

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

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

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

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

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

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

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

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

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

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

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

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):

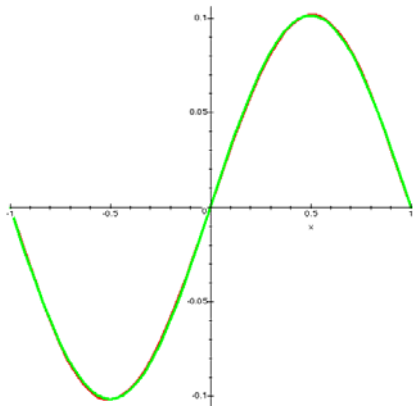
$$\frac{d^2 u}{dx^2} = -\sin(\pi x) \quad \begin{array}{l} u(-1)=0 \\ u(1)=0 \end{array}$$

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

\mathbf{F}

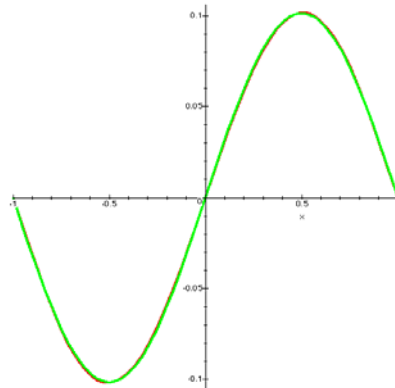
wyberzmy bazę $\Phi_i(x) = (x+1)(x-1)x^{i-1}$

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



$$c_1 = c_3 = c_5 = 0$$

wynik dla bazy funkcji $F2, \Phi4$



$$c2 := \frac{105}{8} \frac{2\pi^2 - 27}{\pi^5}$$

$$c4 := -\frac{315}{8} \frac{2\pi^2 - 21}{\pi^5}$$

wynik dokładnie ten sam co w metodzie Galerkina!

URL z zasady wariacyjnej:

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

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 \mathbf{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) Φ_j
- 2) Wybór funkcji wagowych w_j
- 3) ... które często wybierane są jako maksymalnie rozłączne przestrzennie
wtedy podział przestrzeni jest kolejnym problemem

w stronę metody elementów skończonych

metoda ważonych reszt - ogólnie

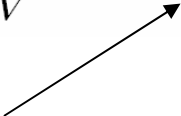
$Lu=f$ (na Ω) \longrightarrow Rozwiązanie dokładne (silnej postaci równania)
 $Bu=g$ (na $d\Omega$) jest „trudne”.
szukamy rozwiązania przybliżonego w bazie funkcji

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

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:

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

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

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

```
graph TD; A["u = sum_{i=1}^n y_i v_i"] --> B["(Lu, v_j) = (f, v_j)"]; A --> C["sum_{i=1}^n (Lv_i, v_j) y_i = (f, v_j)"]; B --> C; C --> D["S_{ij} = (Lv_i, v_j)"]; C --> E["F_j = (f, v_j)"]; D --> F["SY = F"]; E --> F; F --> G["stiffness matrix<br/>macierz sztywności"]; F --> H["load vector<br/>wektor obciążeń"];
```

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

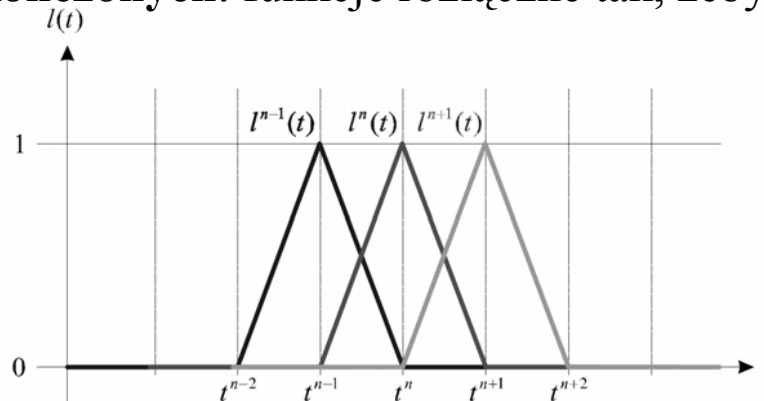
Wysokie potęgi wielomianów niewygodne w użyciu: całkowanie, efekt Rungego,
powód najważniejszy:

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 \mathbf{S} = rzadka



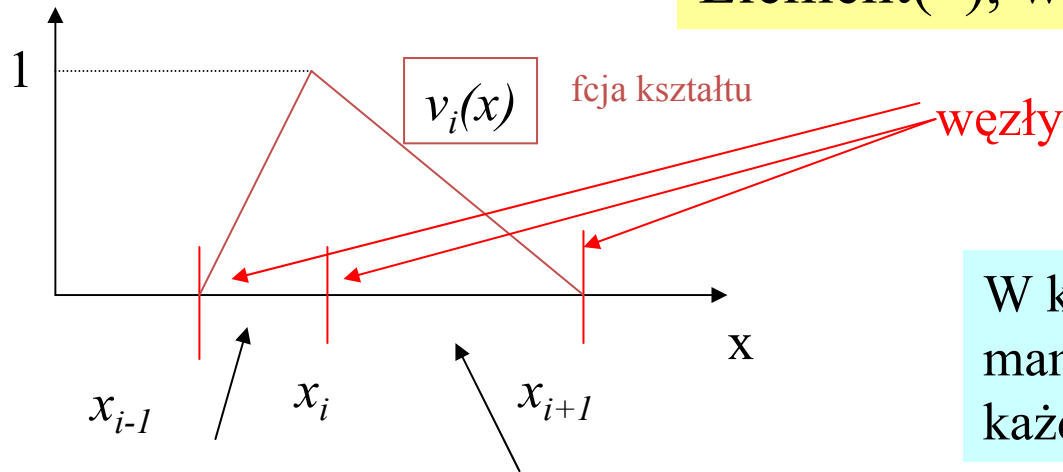
najprostszy wybór *funkcji kształtu*(*): baza funkcji odcinkami liniowymi
zbieżność dostaniemy w przestrzeni funkcji odcinkami liniowymi

(*) trzecie pojęcie z żargonu MES

Zobaczmy 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



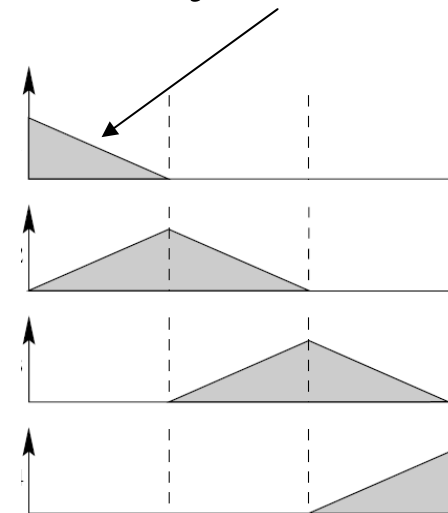
W każdym elemencie:
mamy 2 funkcje,
każda z innym węzłem związana

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



$$u = \sum_{i=1}^n y_i v_i \longrightarrow \text{Dla (jednorodnych) warunków Dirichleta mamy } y_{\text{pierwsze}} = y_{\text{ostatnie}} = 0$$

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

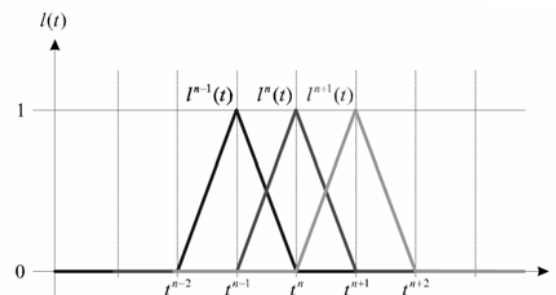
$$\frac{d^2 u}{dx^2} = -\sin(\pi x)$$

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

$$S_{ij} = (Lv_i, v_j) \quad \begin{array}{l} \text{niezerowe tylko dla} \\ i=j, i=j-1 \text{ oraz } i=j+1 \text{ [bez przekrywania} \\ \text{całka znika]} \end{array}$$

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

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

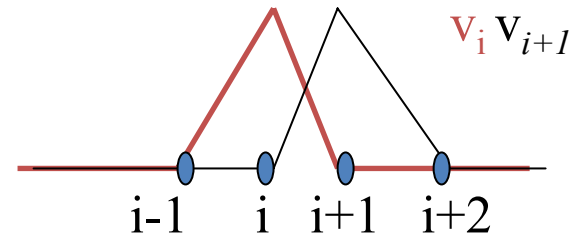


$$S_{ij} = (Lv_i, v_j) \quad \text{niech } j = i+1 \quad v'_i(x) = \begin{cases} \frac{1}{x_i - x_{i-1}} & x \in K_i \\ -\frac{1}{x_{i+1} - x_i} & x \in K_{i+1} \\ 0 & x \notin K_i \cup K_{i+1} \end{cases}$$

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

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

gdy jedna pochodna
dodatnia druga ujemna

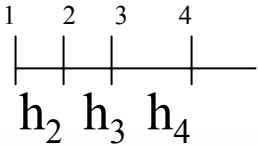


$$S_{i,i+1} = - \left(-\frac{1}{x_{i+1} - x_i} \times \frac{1}{x_{i+1} - x_i} \times (x_{i+1} - x_i) \right) = \frac{1}{h_{i+1}}$$

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

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

Macierz sztywności dla n węzłów
+ warunek $y_1=y_n=0$



wiersz $n-l \rightarrow$

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

$$F_i=(v_i f)$$

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

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

po elemencie K_i

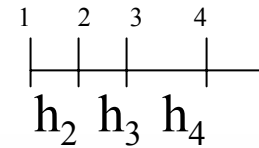
$$+ \int_{x_i}^{x_{i+1}} \frac{-x + x_{i+1}}{x_{i+1} - x_i} f(x) dx$$

po K_{i+1}

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

Macierz sztywności dla n węzłów

+ warunek $y_1=y_n=0$



wiersz $n-1 \rightarrow$

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

$$F_i = (v_i f)$$

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

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

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

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

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

Układ równań z macierzą trójkątniową – 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

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

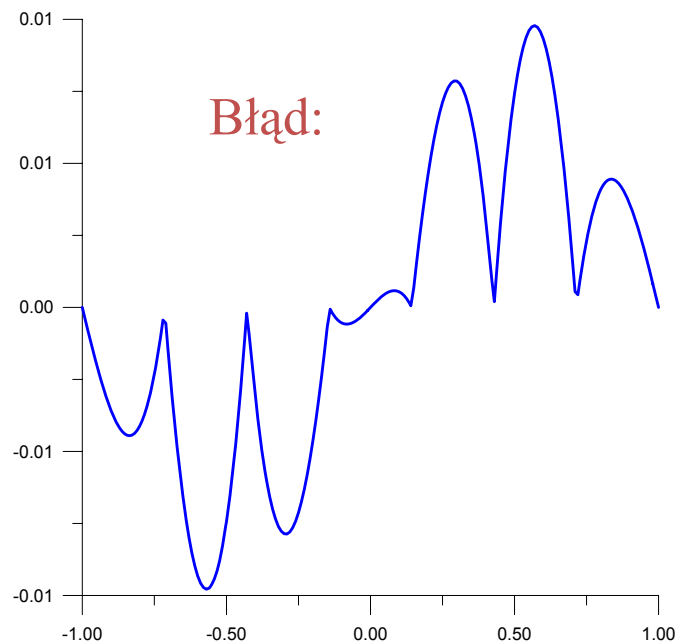
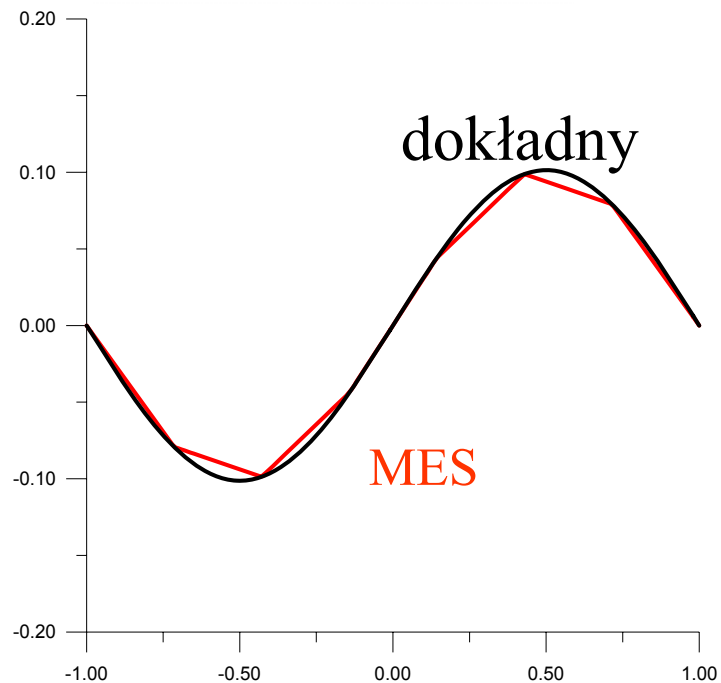
$$U = \begin{pmatrix} \alpha_1 & c_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & c_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \alpha_3 & c_3 & 0 & \dots & 0 \\ & & & \dots & & & \\ 0 & 0 & 0 & \dots & 0 & \alpha_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & \alpha_n \end{pmatrix}$$

bez zmian

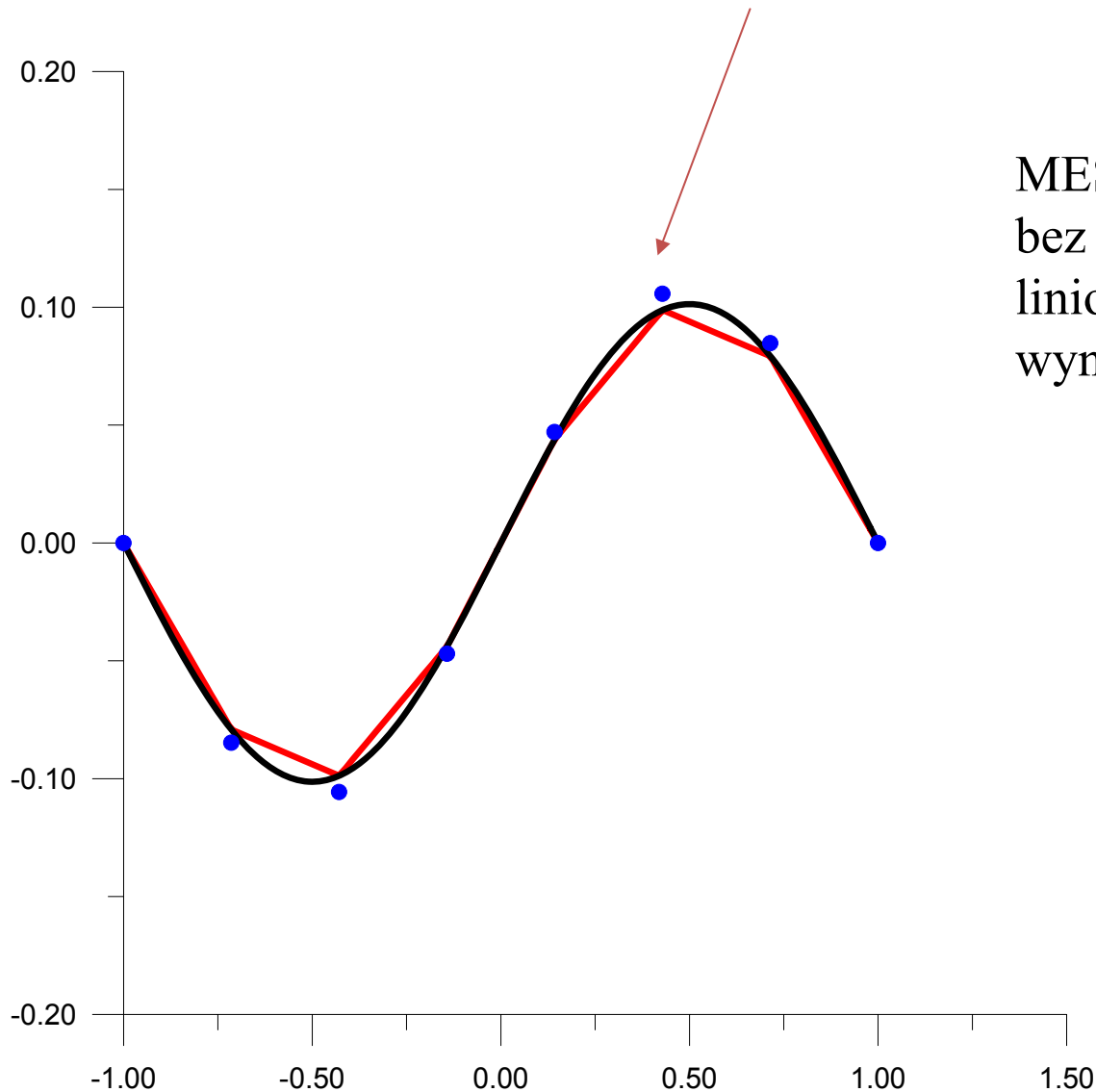
$$\alpha_1 = a_1 \quad \left| \quad \begin{array}{l} \beta_i = \frac{b_i}{\alpha_{i-1}} \\ \alpha_i = a_i - \beta_i c_{i-1} \end{array} \right. \quad \text{dla } i > 1$$

Wynik: równoodległe węzły

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

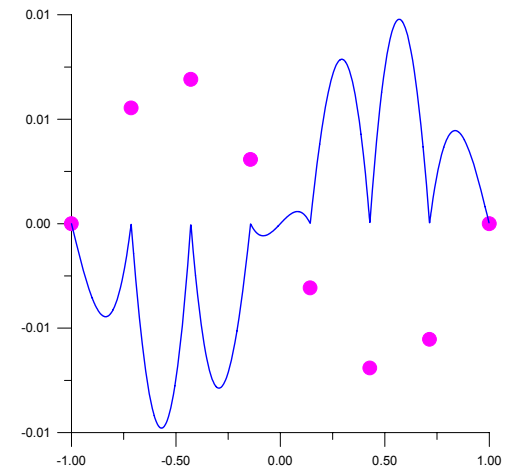


MES (równoodległe węzły) a **MRS** (węzły w tych samych punktach):

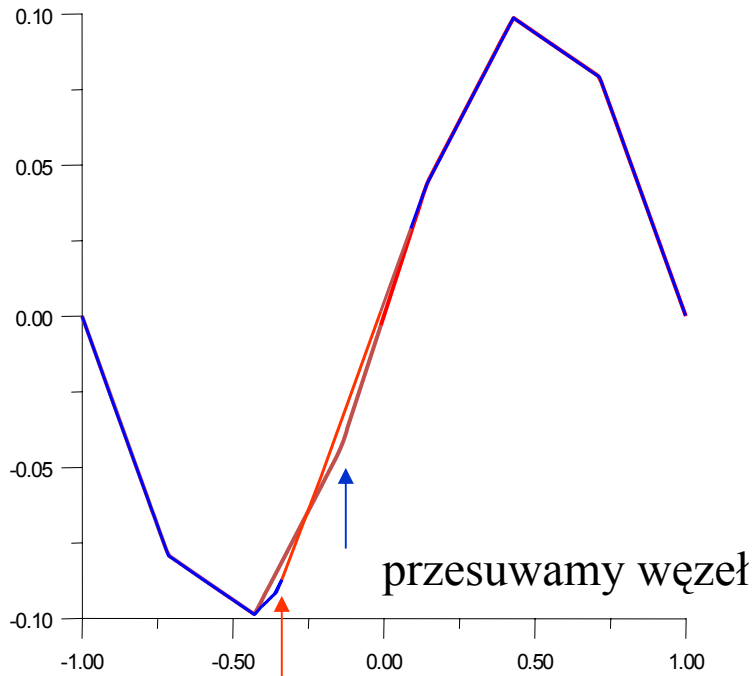


MES dla laplasjanu
bez pochodnej z funkcjami
liniowymi: w węzłach
wynik dokładny !!!

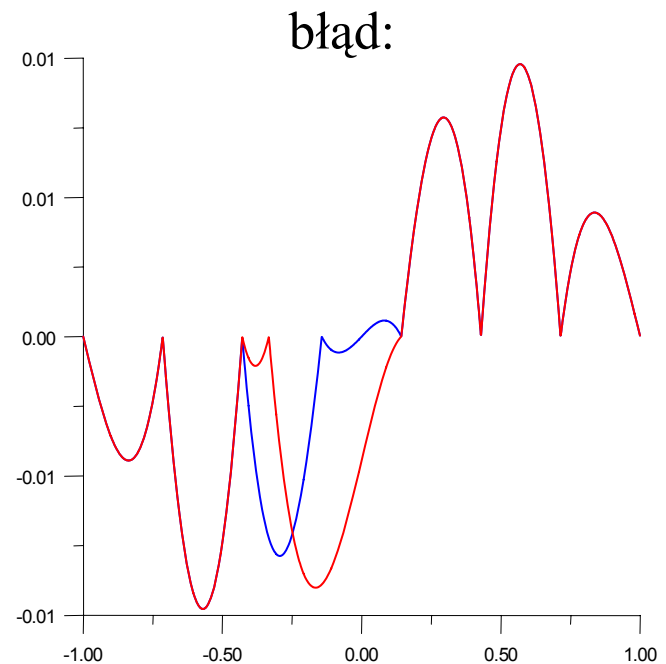
błąd **MRS** i **MES**

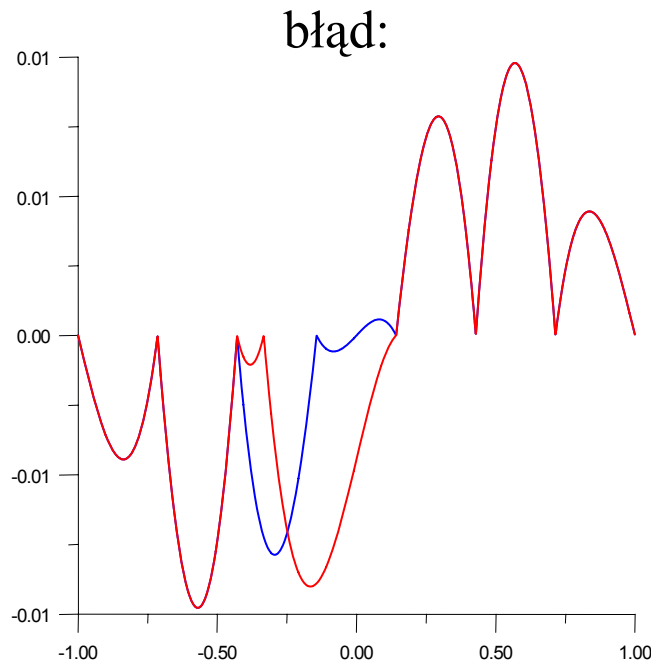


znikanie błędu MES (1D, liniowe f.kształtu) w węzłach
zachodzi również dla nierównomiernego rozkładu węzłów:



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