

Wprowadzenie do MES

Krzysztof Banaś

24 marca 2020

MES¹ (Metoda Elementów Skończonych²) jest jednym z podstawowych narzędzi komputerowego wspomaganie badań naukowych i analiz inżynierskich, o bardzo szerokim zakresie zastosowań i dużej popularności. Znajomość MES obejmuje wiedzę z różnych dziedzin zastosowań (działy fizyki, mechanika konstrukcji, itp.), wiedzę matematyczną dotyczącą podstaw oraz wiedzę informatyczną o realizacji na sprzęcie komputerowym. Znajdujące się poniżej omówienie obejmuje głównie aspekty matematyczne i informatyczne, stara się zrównoważyć prostotę i ścisłość. Składa się z dwóch części, pierwsza stanowi wprowadzenie ogólne (bez wzorów!), druga przedstawia przykład konkretny zastosowania MES.

1. MES w mniej niż 2000 słów

Metoda Elementów Skończonych jest metodą aproksymacji (czyli otrzymywania rozwiązań przybliżonych) równań różniczkowych cząstkowych (w jednym wymiarze przestrzennym także równań różniczkowych zwyczajnych). **Równania różniczkowe (oznaczane dalej skrótowo jako RR) stanowią model matematyczny**, najczęściej jakiegoś procesu lub stanu układu fizycznego. Proces lub stan opisywane są za pomocą parametrów będących funkcjami położenia w przestrzeni i ewentualnie czasu. RR opisują zależności między tymi funkcjami oraz ich pochodnymi. Znalezienie rozwiązania RR to znalezienie tych funkcji, stanowiących funkcje niewiadome dla konkretnego zagadnienia.

Zastosowanie MES do rozwiązania konkretnego zadania naukowego lub inżynierskiego składa się z dwóch odrębnych procesów:

- stworzenia modelu obliczeniowego
- rozwiązania konkretnego zadania za pomocą uzyskanego modelu

¹Popularnie używane na uczelniach rozwinięcie nazwy MES jako Metoda Eliminacji Studentów nie oddaje merytorycznej istoty MES, jednak stało się jedną z podstawowych motywacji do napisania tekstu.

²Angielski termin *the finite element method* jest w polskiej literaturze tłumaczony dwojako: w analizie numerycznej używa się formy „metoda elementu skończonego” (wiernie oddającej treść angielską), w naukach inżynierskich przyjęło się stosować nazwę „metoda elementów skończonych”. Ta druga forma, nawiązująca do tytułu polskiego tłumaczenia klasycznego podręcznika prof. Olgierda Zienkiewicza, jest konsekwentnie stosowana w niniejszym skrypcie.

1.1. Stworzenie modelu obliczeniowego MES

Stworzenie modelu dokonywane jest najczęściej przez fizyków, inżynierów, matematyków. Polega na wyborze lub stworzeniu modelu matematycznego w postaci RR i przekształceniu do tzw. sformułowania MES. **Sformułowanie MES dla pewnego typu problemów składa się z dwóch elementów:**

- **równania całkowitego związanego z RR**
- **definicji z jakich funkcji konstruowane będzie rozwiązanie przybliżone**

Każde zadanie MES określane jest dla pewnego fizycznego obiektu lub grupy obiektów zajmujących miejsce w przestrzeni. To zajmowane miejsce, czyli **obszar w którym zdefiniowane są RR nazywane jest obszarem obliczeniowym**. Każdy taki obszar jest skończony, a więc posiada brzeg. **Elementem modelu matematycznego zjawiska, oprócz RR, które określają zachowanie funkcji niewiadomych wewnątrz obszaru obliczeniowego, są także dodatkowe równania określające zachowanie funkcji niewiadomych na brzegu (warunki brzegowe)**. Także te równania ujmowane są w sformułowaniu MES.

Istotą metody elementów skończonych jest sposób aproksymacji RR polegający na podziale obszaru obliczeniowego na małe podobszary o prostych kształtach, zwane elementami skończonymi oraz specjalny sposób konstruowania funkcji aproksymujących opierający się na funkcjach zdefiniowanych w elementach skończonych. MES jako jedna z niewielu metod potrafi modelować zjawiska w skomplikowanych obszarach obliczeniowych, co stanowi jedną z jej podstawowych zalet w zastosowaniach praktycznych.

W elementach skończonych definiuje się proste funkcje, najczęściej funkcje liniowe lub wielomiany niskiego stopnia, zwane funkcjami kształtu. Elementem sformułowania MES jest określenie w jaki sposób z funkcji kształtu konstruuje się funkcje aproksymujące rozwiązanie. Konstrukcja przebiega w dwóch etapach:

- **z funkcji kształtu określonych w pojedynczych elementach konstruuje się funkcje określone w całym obszarze obliczeniowym (będącym sumą elementów). Proces ten można określić jako sklejanie wielu funkcji w małych podobszarach w jedną funkcję w całym podobszarze – te funkcje określane w całym obszarze nazywane są funkcjami bazowymi**
- **definiuje się w jaki sposób ostateczne rozwiązanie ma być otrzymywane z funkcji bazowych. Zasada jest tutaj prosta, rozwiązanie przybliżone MES jest kombinacją liniową funkcji bazowych (czyli sumą ze współczynnikami różnymi dla każdej funkcji).** Zbiór wszystkich możliwych kombinacji liniowych funkcji bazowych stanowi zbiór wszystkich możliwych rozwiązań przybliżonych danego problemu, czyli stanowi przestrzeń funkcji, w której poszukiwane jest rozwiązanie konkretnego zadania.

Współczynniki występujące w kombinacji liniowej dla aproksymowanej funkcji stanowią zbiór liczb. Znając ten zbiór liczb oraz definicje funkcji bazowych można uzyskać rozwiązanie przybliżone danego problemu w dowolnym punkcie obszaru obliczeniowego. Stąd często mówi się, że **MES jest metodą dyskretyzacji**, czyli uzyskiwania rozwiązania w postaci dyskretnej (a więc w postaci skończonego zbioru liczb), na podstawie którego

można uzyskać rozwiązanie w dowolnym miejscu obszaru. Często, dzięki odpowiednim definicjom funkcji kształtu i funkcji bazowych, dyskretny zbiór liczb określających rozwiązanie MES to zbiór wartości w wybranych punktach obszaru obliczeniowego, zwanych węzłami. Wtedy można powiedzieć, że MES jest metodą uzyskiwania rozwiązania w wybranych punktach obszaru obliczeniowego, a następnie interpolowania rozwiązania w pozostałych punktach obszaru za pomocą funkcji bazowych (lub patrząc z punktu widzenia pojedynczego elementu funkcji kształtu).

Wykorzystanie podanych powyżej elementów składowych sformułowania MES

- równania całkowego
- definicji sposobu tworzenia funkcji aproksymujących

proceeds to the transformation of the integral equation to the form of a linear system of equations. Co dzieje się z całkami z równania? Teraz każdy współczynnik macierzy układu równań liniowych, nazywanej w MES tradycyjnie macierzą sztywności³, jest sumą odpowiednich całek. Całki odpowiadają początkowemu układowi RR, są zdefiniowane dla całego obszaru obliczeniowego, ale oblicza się je jako sumę całek po elementach⁴. Odpowiednie definicje funkcji bazowych powodują, że w otrzymanym układzie równań liniowych zdecydowana większość wartości to zera (niekiedy udział zer może przekraczać 99,99%). To czy dany element macierzy układu jest zerem czy nie, wiadomo na podstawie definicji elementów i funkcji kształtu, dlatego znając te definicje, z góry wiadomo, które wyrazy macierzy są zerami, a które nie. Wyrazów zerowych nie oblicza się. **Rozwiązanie zadania za pomocą MES sprowadza się więc do rozwiązania najczęściej wielkiego, rzadkiego układu równań liniowych. To właśnie ten fakt, że macierz układu równań liniowych w MES jest tak rzadka powoduje, że MES jest atrakcyjna z obliczeniowego punktu widzenia.**

1.2. Rozwiązanie konkretnego zadania MES

Stworzenie modelu obliczeniowego MES jest zadaniem realizowanym, jak to było już powiedziane, przez matematyków, inżynierów, fizyków dla konkretnych grup zadań. Model w postaci równania całkowego i definicji przestrzeni funkcji aproksymujących jest równoważny modelowi w postaci sposobu konstrukcji układu równań liniowych MES. Ta druga postać lepiej nadaje się do tworzenia algorytmów numerycznych i jest podstawową postacią wykorzystywaną przy tworzeniu programów MES. Często definicje elementów i funkcji kształtu są tak proste, że udaje się wyliczyć całki tworzące wyrazy macierzy sztywności, w zależności od pewnych parametrów elementów, i podać gotowe wzory na wyrazy macierzy sztywności. **Przyjmować będziemy, że końcowym etapem tworzenia modelu MES jest podanie algorytmicznego przepisu na tworzenie układu równań liniowych odpowiadających sformułowaniu MES.**

Otrzymany model MES zazwyczaj ma zastosowanie do pewnej grupy problemów. Na jego podstawie tworzone są programy komputerowe, które realizują model w konkretnych obliczeniach. Przeciętny użytkownik MES nie tworzy modelu MES, ale korzysta z gotowych modeli dostarczanych wraz z odpowiednimi programami.

³W nawiązaniu do pierwszych zastosowań MES w dynamice konstrukcji

⁴Całka po całym obszarze całkowania jest sumą całek po podobszarach, w tym wypadku elementach

Rozwiązanie konkretnego zadania MES polega na skorzystaniu z programu MES zawierającego odpowiedni model MES. Proces ten **rozpoczyna się od realizacji następujących wstępnych etapów:**

- **wybór modelu** (każdy program MES oferuje zazwyczaj kilka modeli MES, odpowiadających konkretnym RR i warunkom brzegowym)
- **definicja obszaru obliczeniowego i podział obszaru na elementy** (w programach MES mogą pojawiać się rozmaite elementy jedno, dwu i trójwymiarowe)
- **definicja przestrzeni, w której poszukiwane są funkcje niewiadome czyli wybór aproksymacji** (programy MES mogą oferować rozmaite zestawy funkcji kształtu dla elementów i funkcji bazowych w całym obszarze obliczeniowym)
- **określenie parametrów modelowanego procesu** (najczęściej parametry te związane są ze współczynnikami RR i warunków brzegowych, w praktyce odpowiadają np. parametrom materiałów w modelowanych obiektach, zasadom interakcji obiektów ze światem zewnętrznym, itp.)

Bardzo istotnym elementem w praktyce jest zdefiniowanie obszaru obliczeniowego oraz jego podziału na elementy. Etap ten, zwany **pre-processingiem**, w przypadku skomplikowanych obszarów obliczeniowych może zająć wiele czasu i wysiłku, może wymagać współpracy z urządzeniami pomiarowymi, oprogramowaniem CAD, odrębnymi programami generowania siatek MES. Raz zdefiniowana siatka MES może być używana przy rozwiązaniu szeregu zadań dla tego samego obiektu.

Po pełnym zdefiniowaniu rozwiązywanego zadania uruchamiane są procedury obliczeniowe. Ten etap nosi nazwę **processingu** i składa się z dwóch podstawowych faz:

- **utworzenie układu równań liniowych**
- **rozwiązanie układu równań liniowych**

Tworzenie układu równań liniowych odbywa się na podstawie przepisu stanowiącego model obliczeniowy MES. Rozwiązanie układu równań liniowych następuje poprzez zastosowanie odpowiedniego algorytmu, czasem są to algorytmy ogólnego przeznaczenia, czasem metody specjalnie opracowane dla MES.

Efektem rozwiązania układu równań liniowych jest zbiór liczb, tworzących wektor niewiadomych (często, jak to było już wspomniane, jest to zbiór wartości w konkretnych punktach obszaru, np. wierzchołkach elementów). Znając ten zbiór wartości oraz definicje funkcji kształtu w elementach można odtworzyć rozwiązanie przybliżone w całym obszarze obliczeniowym. Na podstawie rozwiązania przybliżonego można także obliczać inne parametry zjawiska. Takie obliczenia, a także wizualizacja rozwiązań i obliczonych na ich podstawie wielkości, są realizowane w etapie zwanym **post-processingiem**. Także tutaj istnieją odrębne programy realizujące te funkcje.

Otrzymanie rozwiązania za pomocą programu MES nie powinno nigdy być końcem procedury rozwiązywania problemu. Trzeba mieć świadomość, że uzyskany wynik prawie zawsze obciążony jest błędem. Istnieje wiele możliwych źródeł błędu rozwiązania. Kilka najważniejszych to:

- błąd modelowania (zastosowany model matematyczny nie odzwierciedla dokładnie rzeczywistości)
- błąd wartości współczynników (przyjęte wartości współczynników RR i warunków brzegowych, czyli np. dane materiałowe, dane o interakcji obiektu ze światem zewnętrznym obciążone są błędem)
- błąd odwzorowania obszaru (obszar obliczeniowy nie odpowiada dokładnie rzeczywistemu obszarowi zajmowanemu przez analizowany obiekt)
- błąd numeryczny (błąd dyskretyzacji, zastosowana metoda aproksymacji wprowadza błąd w stosunku do rozwiązania dokładnego problemu wyjściowego w postaci RR)
- błąd zaokrąglenia (ze względu na zastosowanie ograniczonej dokładności reprezentacji liczb w komputerze, rozwiązanie uzyskane programem komputerowym nie odpowiada rozwiązaniu przybliżonemu, które zostałyby otrzymane przy dokładnej reprezentacji liczb)

Po uzyskaniu rozwiązania wyniki należy poddać weryfikacji. W przypadku błędu modelowania mówimy o walidacji modelu. Model matematyczny jest opracowywany przez inżynierów, fizyków, matematyków – przeciętny użytkownik programów MES powinien sprawdzić jak dobrze zastosowany przez niego model matematyczny odwzorowuje rzeczywistość, np. jak wiele osób dotychczas stosowało ten model, jakie uzyskały wyniki itp.

Z kolei błędy wartości współczynników i błąd odwzorowania obszaru należą do fazy przygotowania danych do rozwiązywanego problemu. Matematyczna analiza sformułowania problemu może przynieść odpowiedź na pytanie, jak wrażliwy jest model na zmiany powyższych parametrów, w jaki sposób zmiany parametrów wpływają na zmianę rozwiązania, czy wiedząc, że informacje o danych i obszarze obciążone są pewnym błędem nadal możemy zakładać, że rozwiązanie MES wystarczająco dokładnie opisuje badane zjawisko.

Błąd odwzorowania obszaru może wynikać nie tylko z błędów danych wejściowych przy definicji problemu, może zostać wprowadzony w fazie dyskretyzacji obszaru, czyli generowania siatki MES. Tutaj także analiza matematyczna zagadnienia może prowadzić do prób oszacowania jak duży jest błąd i w jaki sposób można go zmniejszyć.

Kolejnym typem błędów jest błąd numeryczny. MES jako metoda aproksymacji, w zdecydowanej większości zastosowań (poza niezwykle prostymi zadaniami) prowadzi do błędów dyskretyzacji⁵. Błąd dyskretyzacji możemy określić jako różnicę rozwiązania dokładnego RR i przybliżonego rozwiązania MES. W teorii MES bada się jaka jest zależność błędów numerycznych od sformułowania MES i parametrów rozwiązania, takich jak np. maksymalna wielkość elementów w siatce MES lub stopień wielomianów przyjętych jako funkcje kształtu.

Teoria dostarcza także informacji jak dla konkretnego zadania poprawić rozwiązanie. Mówimy wtedy o adaptacji zadania, polegającej najczęściej na modyfikacji siatki lub

⁵Błąd dyskretyzacji związany jest z zamianą rozwiązania dokładnego z przestrzeni nieskończenie wymiarowej, na rozwiązanie z przestrzeni skończenie wymiarowej, czyli rozwiązanie, które daje się przedstawić w postaci skończonej liczby wartości, np. wartości rozwiązania w węzłach siatki MES

doboru funkcji kształtu. Zdecydowana większość współczesnych programów MES zawiera mechanizmy adaptacji. Ich zastosowanie polega najczęściej na wstępnym rozwiązaniu zadania, oszacowaniu popełnionego błędu numerycznego, a następnie modyfikacji zadania i ponownym rozwiązaniu. **Informacje o procedurach szacowania błędu oraz procedurach modyfikacji zadania (siatki i aproksymacji) powinny znajdować się w dokumentacji programu MES. Ich znajomość jest często warunkiem koniecznym uzyskiwania wiarygodnych i dokładnych wyników za pomocą MES.**

Ostatni typ błędu, błąd zaokrągleń jest specyficzny dla komputerowej realizacji algorytmów MES. Użytkownik powinien mieć świadomość, w których momentach obliczeń mogą pojawić się błędy zaokrągleń, jak bardzo są one istotne dla dokładności wyników i czy istnieją alternatywne algorytmy unikające tych błędów. Informacje takie powinny także znaleźć się w podręczniku użytkownika programu komputerowego MES.

2. Prosty przykład zastosowania MES dla prostego problemu w przestrzeni jednowymiarowej.

Zastosujemy MES do rozwiązania następującego RR:

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

określonego w jednowymiarowym obszarze obliczeniowym będącym odcinkiem $(0, 1)$. Warunkami brzegowymi są:

- warunek Dirichleta (określający wartość funkcji) dla $x = 0$:

$$u(0) = 0$$

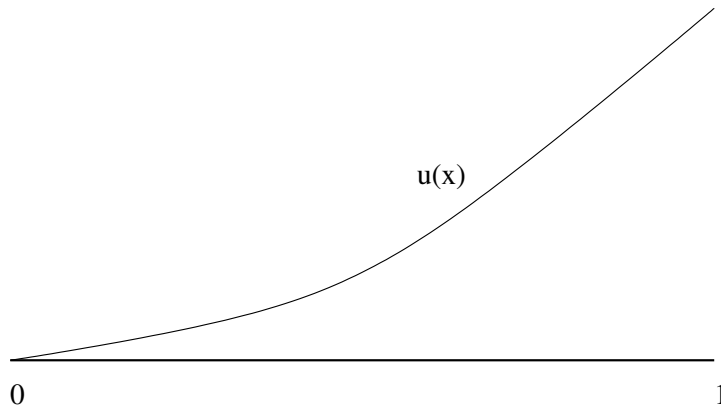
- warunek brzegowy Neumanna (określający pochodną funkcji) dla $x = 1$:

$$\frac{du}{dx}(1) = 2$$

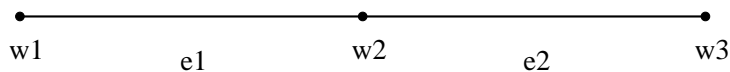
Rozwiązaniem dokładnym zadania jest funkcja kwadratowa $u = x^2$ przedstawiona na rys. 1. Pierwszym krokiem na drodze do rozwiązania zadania za pomocą MES jest dyskretyzacja obszaru, czyli podział na elementy skończone. W naszym przypadku zakładamy, że obszar dzielimy na dwa elementy, e_1 i e_2 , wierzchołki elementów oznaczamy jako w_1, w_2, w_3 . Sytuację ilustruje rys. 2

W każdym elemencie definiujemy dwie liniowe funkcje kształtu, ϕ_1 i ϕ_2 , pokazane na rys.3.

Z funkcji kształtu określonych w elementach konstruujemy (sklejamy) funkcje bazowe określone w całym obszarze obliczeniowym. Zakładamy, że funkcje bazowe związane są z wierzchołkami elementów, mamy więc trzy funkcje bazowe, ψ_1, ψ_2 i ψ_3 , pokazane na rys. 4. Ich istotną cechą jest to, że funkcja bazowa związana z wierzchołkiem w_i ma wartość 1 w tym wierzchołku i wartość 0 w pozostałych wierzchołkach.



Rysunek 1: Rozwiązanie dokładne przykładowego zagadnienia brzegowego



Rysunek 2: Podział obszaru obliczeniowego na elementy skończone

Definiujemy teraz zbiór funkcji, w którym będziemy poszukiwać przybliżonego rozwiązania zagadnienia brzegowego. Zbiór ten, czyli przestrzeń funkcji oznaczaną przez V^h , określamy jako zbiór wszystkich funkcji będących kombinacjami liniowymi funkcji bazowych ψ_i , czyli funkcji mających postać:

$$u^h(x) = U_1^h \psi_1(x) + U_2^h \psi_2(x) + U_3^h \psi_3(x)$$

Wartości U_i^h , będące współczynnikami kombinacji liniowej, stanowią niewiadome naszego problemu, tzw. stopnie swobody. Nasz problem MES ma więc trzy niewiadome, trzy stopnie swobody. Znając wartości niewiadomych i funkcje bazowe ψ_i możemy odtworzyć funkcję $u^h(x)$ w dowolnym punkcie obszaru obliczeniowego. Na przykład dla wartości wektora $U^h = \{0.5, 0.3, 1.0\}$ funkcja $u^h(x)$ ma postać:

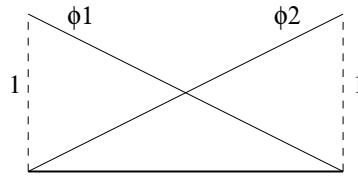
$$u^h(x) = 0.5\psi_1(x) + 0.3\psi_2(x) + \psi_3(x)$$

zilustrowaną na rys. 5. Jak widać, w przestrzeni V^h , dzięki specjalnemu doborowi funkcji bazowych, wartości stopni swobody są wartościami rozwiązania w wierzchołkach elementów.

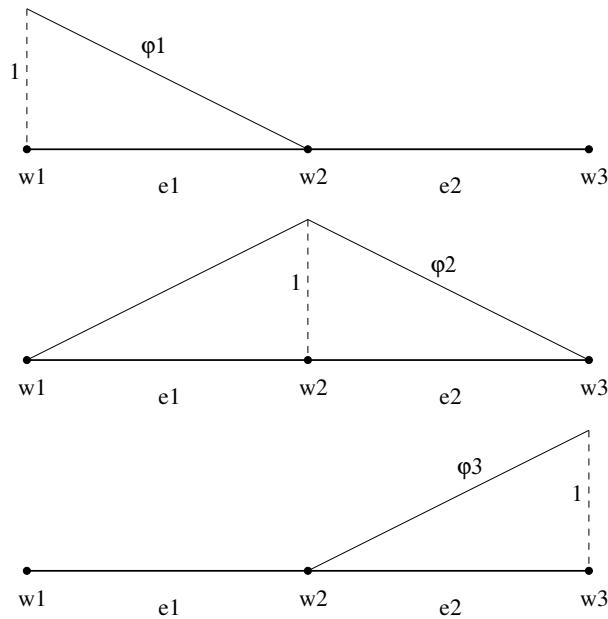
Kolejnym etapem procedury jest uzyskanie sformułowania MES. Sformułowanie MES jest równaniem całkowym odpowiadającym równaniu różniczkowemu (funkcje spełniające sformułowanie MES są aproksymacjami zagadnienia brzegowego). Sformułowanie MES odpowiadające przykładowemu zagadnieniu brzegowemu ma postać:

Znajdź niewiadomą funkcję $u^h(x)$ należącą do przestrzeni V^h i spełniającą warunek Dirichleta $u^h(0) = 0$, taką że dla dowolnej funkcji testującej $w(x)$ należącej do V^h i spełniającej warunek $w(0) = 0$, prawdziwe jest równanie całkowite:

$$-\int_0^1 \frac{du^h}{dx} \frac{dw}{dx} dx = \int_0^1 2w(x) dx - 2w(1)$$



Rysunek 3: Liniowe funkcje kształtu w elemencie jednowymiarowym



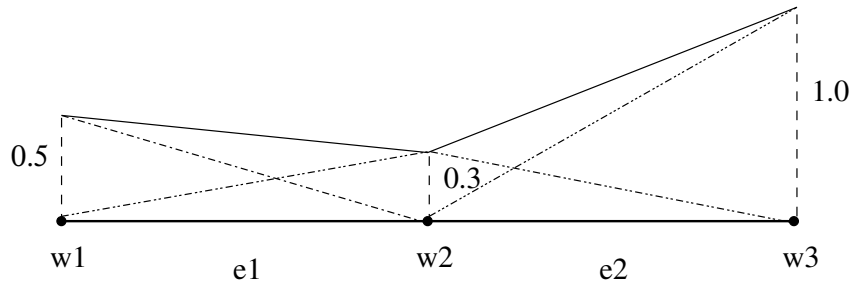
Rysunek 4: Funkcje bazowe w obszarze obliczeniowym

Pomijając szczegóły wyprowadzenia powyższego wzoru przyjmijmy, że sformułowanie MES zostało zadane, wraz z dowodem, że rozwiązanie $u^h(x)$ rzeczywiście przybliży rozwiązanie dokładne $u(x)$. Sformułowanie MES zawiera w sobie przetransformowane równanie różniczkowe (całki po obu stronach równania) oraz wyraz po prawej stronie pochodzący z uwzględnienia warunku brzegowego Neumanna. Warunek brzegowy Dirichleta uwzględniony jest jawnie poprzez założenie, że $u^h(0) = 0$.

Sformułowanie MES zapisane w powyższej postaci, z wykorzystaniem funkcji testujących $w(x)$, jest trudne do intuicyjnego zrozumienia przez osoby nieobite z aparatem matematycznym teorii MES. W postaci takiej używane jest często przez matematyków lub specjalistów od analizy numerycznej do przeprowadzania dowodów istnienia i jednoznaczności rozwiązania przybliżonego oraz jego dokładności, czyli różnicy pomiędzy rozwiązaniem dokładnym $u(x)$ a rozwiązaniem przybliżonym $u^h(x)$.

Sformułowanie MES nie stanowi jeszcze podstawy tworzenia algorytmów numerycznych. Dla inżynierów i naukowców stosujących MES oraz dla informatyków zaangażowanych w implementację MES w programach komputerowych bardziej od postaci całkowej sformułowania MES przydatna jest postać uzyskana przez transformację do układu równań liniowych.

Podstawą transformacji do postaci układu równań liniowych jest wykorzystanie założenia,



Rysunek 5: Przykładowa funkcja z przestrzeni V^h

że funkcja niewiadoma należy do przestrzeni V^h , tzn. jest kombinacją liniową funkcji bazowych:

$$u^h(x) = \sum_{K=1}^3 U_K^h \psi_K(x)$$

Podobnie zakładamy o funkcjach testujących $w(x)$:

$$w(x) = \sum_{L=1}^3 W_L \psi_L(x)$$

Podstawienie powyższych wzorów do sformułowania MES i kilka prostych transformacji prowadzi do układu równań liniowych MES, mającego w przypadku naszego zadania przykładowego postać:

$$\sum_{K=1}^3 A_{LK} U_K^h = F_L + B_L \quad L = 1, 2, 3$$

Jest to układ trzech równań z trzema niewiadomymi: U_1^h, U_2^h, U_3^h . Współczynniki macierzy układu A (macierzy sztywności) dane są wzorami:

$$A_{LK} = - \int_0^1 \frac{d\psi_K}{dx} \frac{d\psi_L}{dx} dx \quad K, L = 1, 2, 3$$

Składowe wektora prawej strony F_L oblicza się jako:

$$F_L = \int_0^1 2\psi_L(x) dx \quad L = 1, 2, 3$$

Dodatkowo w układzie występują wyrazy odpowiadające warunkom brzegowym, w naszym przypadku jest to wektor $B_L = [0, 0, -2]$.

Jak widać postać układu równań liniowych MES bezpośrednio nawiązuje do sformułowania MES. Wzór na współczynniki A_{LK} odpowiada całce po lewej stronie sformułowaniu MES, w której w miejsce funkcji niewiadomej i testującej wstawione zostały funkcje bazowe. Składowe F_L odpowiadają całce po prawej stronie sformułowania słabego, a wektor B wyrazowi związanemu z warunkiem brzegowym. W naszym przypadku mamy do czynienia z małym układem tylko trzech równań z trzema niewiadomymi, ale konstrukcja układów równań dla

złożonych zagadnień z milionami stopni swobody przebiega w sposób identyczny. Ostateczna postać układu równań dla zadania przykładowego może zostać zapisana jako:

$$\begin{aligned} A_{11}U_1^h + A_{12}U_2^h + A_{13}U_3^h &= F_1 + B_1 \\ A_{21}U_1^h + A_{22}U_2^h + A_{23}U_3^h &= F_2 + B_2 \\ A_{31}U_1^h + A_{32}U_2^h + A_{33}U_3^h &= F_3 + B_3 \end{aligned}$$

lub w postaci macierzowej

$$\mathbf{A} \cdot \mathbf{U}^h = \mathbf{F} + \mathbf{B}$$

czyli

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \cdot \begin{pmatrix} U_1^h \\ U_2^h \\ U_3^h \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix}$$

Pierwsze równanie otrzymanego układu odpowiada pierwszej funkcji bazowej ψ_1 , drugie równanie drugiej funkcji ψ_2 , a trzecie funkcji bazowej ψ_3 . W zapisie macierzowym układu równań oznacza to że każdy wiersz odpowiada jednej, kolejnej funkcji bazowej, przy czym funkcje bazowe pochodzą z reprezentacji w postaci sumy dla funkcji testującej $w(x)$.

Z kolei kolumny macierzy układu \mathbf{A} odpowiadają funkcjom bazowym pochodzącym z rozkładu funkcji niewiadomej na postać kombinacji liniowej. Pierwsza kolumna odpowiada pierwszej funkcji bazowej, druga drugiej, trzecia trzeciej. Tak więc każdy wyraz macierzy sztywności A_{LK} odpowiada parze funkcji bazowych ψ_K, ψ_L . Podobnie każda składowa F_L i B_L odpowiadają pojedynczej funkcji bazowej ψ_L .

W przyjętej przez nas postaci aproksymacji, za pomocą liniowych funkcji bazowych, pojedynczej funkcji bazowej odpowiada pojedynczy wierzchołek elementów w obszarze obliczeniowym. Taki wierzchołek wraz z odpowiadającą mu funkcją bazową nazywany jest węzłem siatki MES. Jeśli gdziekolwiek używany jest zwrot funkcja bazowa, można kojarzyć to z węzłem siatki MES i na odwrót.

Postać układu równań jest już postacią nadającą się do implementacji w programach komputerowych. Pierwszą naiwną implementacją mogłoby być dokonanie podwójnej pętli po wszystkich funkcjach bazowych (węzłach siatki MES) w obszarze obliczeniowym, dla każdej pary ψ_K, ψ_L obliczenie A_{LK} , dla każdej funkcji ψ_L obliczenie F_L , a następnie sprawdzenie czy dla ψ_L nie zachodzi potrzeba uwzględnienia warunku brzegowego w postaci odpowiedniego wyrazu B_L . Dla tak utworzonego układu równań liniowych rozwiązanie w postaci wektora \mathbf{U}^h (wartości $u^h(x)$ w węzłach siatki MES) stanowiłoby podstawę do stworzenia przybliżonego rozwiązania $u^h(x)$ w całym obszarze obliczeniowym.

W praktyce jednak najczęściej stosuje się inny algorytm wykorzystujący pętlę po elementach, własności całkowania oraz funkcje kształtu w miejsce funkcji bazowych. Wyrazy A_{LK} zdefiniowane są jako całki funkcji bazowych po całym obszarze obliczeniowym. Korzystamy z faktu, że całka po obszarze złożonym z podobszarów jest równa sumie całek po podobszarach. Tak więc

$$A_{LK} = - \int_0^1 \frac{d\psi_K}{dx} \frac{d\psi_L}{dx} dx = - \int_{e_1} \frac{d\psi_K}{dx} \frac{d\psi_L}{dx} dx - \int_{e_2} \frac{d\psi_K}{dx} \frac{d\psi_L}{dx} dx = A_{LK}^{e_1} + A_{LK}^{e_2}$$

Cała macierz sztywności w naszym przykładzie ma postać

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} A_{11}^{e_1} + A_{11}^{e_2} & A_{12}^{e_1} + A_{12}^{e_2} & A_{13}^{e_1} + A_{13}^{e_2} \\ A_{21}^{e_1} + A_{21}^{e_2} & A_{22}^{e_1} + A_{22}^{e_2} & A_{23}^{e_1} + A_{23}^{e_2} \\ A_{31}^{e_1} + A_{31}^{e_2} & A_{32}^{e_1} + A_{32}^{e_2} & A_{33}^{e_1} + A_{33}^{e_2} \end{pmatrix}$$

Wyrazy macierzy \mathbf{A} można obliczyć w pętli po elementach. W każdym elemencie należałoby zrobić podwójną pętlę po wszystkich funkcjach bazowych (węzłach siatki) i obliczyć odpowiednie wyrazy. Procedura taka byłaby jeszcze mniej efektywna niż algorytm naiwny, na szczęście można w tym momencie skorzystać z ważnej cechy układu równań MES co znacznie skraca czas obliczeń.

Zadajemy pytanie czy musimy robić pętlę po wszystkich funkcjach bazowych, czy dla każdej pary funkcji bazowych musimy liczyć wyrazy A_{LK} ? Odpowiedź brzmi: nie, musimy liczyć A_{LK} tylko wtedy kiedy wiemy, że wyraz ten jest niezerowy. Czy mamy taką wiedzę przed przystąpieniem do obliczeń? Odpowiedź daje analiza sposobu aproksymacji w MES.

Wyraz A_{LK} jako całka, w której występują odpowiednie funkcje bazowe, jest niezerowy wtedy, kiedy istnieje choć jeden element e_i , dla którego $A_{LK}^{e_i}$ jest różne od zera. Oznacza to, że w elemencie e_i obie funkcje bazowe ψ_K i ψ_L są różne od zera.

Rzut oka na rys. 4 wystarczy, żeby zorientować się, że nie istnieje żaden element w obszarze obliczeniowym, w którym jednocześnie ψ_1 i ψ_3 są różne od zera. Tak więc A_{13} i A_{31} są równe zero. Macierz \mathbf{A} sprowadza się do postaci:

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{pmatrix}$$

Prosta analiza obszaru obliczeniowego prowadzi do wniosku, że gdybyśmy mieli więcej elementów w obszarze obliczeniowym to liczba zer w macierzy sztywności byłaby znacznie większa. Zwiększając liczbę elementów w obszarze, zwiększalibyśmy liczbę węzłów siatki MES. Funkcja bazowa ψ_1 jest niezerowa tylko w pierwszym elemencie. Funkcja ψ_2 w pierwszym i drugim, ψ_3 w drugim i trzecim, każda następna byłaby zerowa i w pierwszym i w drugim elemencie. To oznacza, że żadna z następnych funkcji bazowych nie byłaby jednocześnie niezerowa z ψ_1 i ψ_2 w żadnym elemencie. W efekcie w pierwszym wierszu macierzy sztywności pozostałyby dwa elementy niezerowe, a w drugim trzy tak jak to jest w naszym przykładzie. Podobnie w każdym następnym wierszu byłyby tylko trzy wyrazy niezerowe, za wyjątkiem wiersza ostatniego, gdzie byłyby dwa.

Jeśli nasza siatka MES miałaby milion węzłów, czyli macierz sztywności byłaby macierzą 1000000×1000000 to w każdym wierszu (z wyjątkiem pierwszego i ostatniego) mielibyśmy 3 wyrazy niezerowe i 999997 zer czyli procent zer w macierzy wynosiłby znacznie ponad 99,99%. Sytuacja taka jest typowa dla zadań MES z wielką liczbą stopni swobody.

Wróćmy do algorytmu, w którym próbujemy obliczyć wyrazy macierzy sztywności w pętli po elementach. Powiedzieliśmy, że pomysł, żeby wykonać w elemencie podwójną pętlę po wszystkich funkcjach bazowych jest niezwykle kosztowny obliczeniowo i niepotrzebny, gdyż wiele wyrazów macierzy sztywności jest zerowych. W jaki sposób możemy będąc w konkretnym elemencie zorientować się, które wyrazy w tym elemencie są zerowe, a które nie? Musimy rozważyć funkcje bazowe niezerowe w tym elemencie. Które funkcje bazowe są niezerowe w

danym elemencie? Te, które powstały przez sklejenie funkcji kształtu danego elementu. Jak uwzględnić wszystkie funkcje bazowe niezerowe w elemencie? Zamiast wykonywać pętle po funkcjach bazowych i sprawdzać czy dana funkcja jest niezerowa w elemencie, możemy zrobić pętle po funkcjach kształtu w elemencie i sprawdzić jakie funkcje bazowe tworzone są z funkcji kształtu. Tym sposobem możemy uwzględnić wszystkie funkcje bazowe niezerowe w danym elemencie. Musimy tylko w każdym elemencie przechowywać informację, które funkcje bazowe powstały przez sklejenie lokalnych funkcji kształtu elementu.

W praktyce powyższe zależności wyraża się korzystając z pojęcia węzła siatki. Numerację węzłów w całym obszarze obliczeniowym (będącą jednocześnie numeracją funkcji bazowych) określa się mianem numeracji globalnej (patrz rys. 2). Numeracja globalna jest numeracją odpowiadającą pozycjom w macierzy A , nazywanej przez to często globalną macierzą sztywności.

Niezależnie od numeracji globalnej, w każdym elemencie wprowadza się numerację lokalną. W przypadku naszego obszaru obliczeniowego i liniowych funkcji kształtu, lewy węzeł i związaną z nim funkcję kształtu opatruje się numerem 1, a prawy węzeł i jego funkcje kształtu oznacza numerem 2.

Następnie wprowadza się, dla każdego elementu, odwzorowanie numerów lokalnych w globalne. Dla elementu e_1 będzie to $[1, 2] \Rightarrow [1, 2]$, a dla drugiego elementu $[1, 2] \Rightarrow [2, 3]$.

W konsekwencji sposób postępowania przy obliczaniu niezerowych elementów macierzy sztywności A może więc być następujący. W pętli po elementach, dla każdego elementu e_i oblicza się lokalną elementową macierz sztywności a . Wyrazami macierzy są elementowe przyczynki do globalnej macierzy sztywności A . Korzystając z odwzorowania lokalnej numeracji węzłów (funkcji kształtu) w globalną numerację węzłów (funkcji bazowych) przeprowadza się dodanie wyrazów lokalnej macierzy sztywności do odpowiednich wyrazów macierzy globalnej. Ta procedura nosi zwyczajową nazwę agregacji lokalnej elementowej macierzy sztywności. Należy zwrócić uwagę, że w praktycznych zastosowaniach MES lokalne elementowe macierze sztywności są macierzami małymi i gęstymi, natomiast globalna macierz sztywności jest wielka i rzadka.

W przypadku naszej siatki MES agregacja oparta jest na następującym odwzorowaniu wyrazów macierzy elementowej z numeracją lokalną (po lewej) do odpowiednich wyrazów macierzy globalnej (po prawej):

$$\begin{pmatrix} a_{11}^{e_1} & a_{12}^{e_1} \\ a_{21}^{e_1} & a_{22}^{e_1} \end{pmatrix} \Rightarrow \begin{pmatrix} A_{11}^{e_1} & A_{12}^{e_1} \\ A_{21}^{e_1} & A_{22}^{e_1} \end{pmatrix}$$

a dla drugiego elementu ma postać:

$$\begin{pmatrix} a_{11}^{e_2} & a_{12}^{e_2} \\ a_{21}^{e_2} & a_{22}^{e_2} \end{pmatrix} \Rightarrow \begin{pmatrix} A_{22}^{e_2} & A_{23}^{e_2} \\ A_{32}^{e_2} & A_{33}^{e_2} \end{pmatrix}$$

W ostateczności globalna macierz sztywności A złożona będzie z następujących wyrazów lokalnych:

$$A = \begin{pmatrix} a_{11}^{e_1} & a_{12}^{e_1} & 0 \\ a_{21}^{e_1} & a_{22}^{e_1} + a_{11}^{e_2} & a_{12}^{e_2} \\ 0 & a_{21}^{e_2} & a_{22}^{e_2} \end{pmatrix}$$

W jaki sposób obliczane są elementowe macierze sztywności? Dla każdego elementu e_i wykonywana jest podwójna pętla po węzłach elementu (funkcjach kształtu elementu). Dla

pojedynczej pary funkcji kształtu ϕ_k i ϕ_l obliczany jest wyraz $a_{kl}^{e_i}$ elementowej macierzy sztywności. Jeśli posiadamy algebraiczny wzór na $a_{kl}^{e_i}$ algorytm jest niezwykle prosty. Jeśli nie, to musimy obliczyć $a_{kl}^{e_i}$ jako odpowiednią całkę związaną ze sformułowaniem MES. Dla bardziej złożonych problemów (np. nieliniowych) lub dla bardziej złożonych elementów (np. krzywoliniowych) obliczenia wyrazów lokalnych macierzy sztywności dokonuje się przez całkowanie numeryczne, co dodatkowo komplikuje całą procedurę.

W naszym prostym przykładzie obliczeniowym nie musimy korzystać z całkowania numerycznego. Pojedynczy wyraz elementowej macierzy sztywności dany jest wzorem:

$$a_{kl}^{e_i} = - \int_{e_i} \frac{d\phi_k}{dx} \frac{d\phi_l}{dx} dx$$

Zakładając postać funkcji kształtu zilustrowaną na rys. 3, widać że $\frac{d\phi_1}{dx} = -1/h_i$, a $\frac{d\phi_2}{dx} = 1/h_i$, gdzie h_i oznacza długość elementu. Dokonując całkowania otrzymujemy wzór na elementową macierz sztywności (taką samą dla e_1 i e_2 , jako że w obu przypadkach występują te same funkcje kształtu i ta sama długość elementu $h_i = 0.5$):

$$\mathbf{a} = \begin{pmatrix} -1/h & 1/h \\ 1/h & -1/h \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}$$

Agregacja obu lokalnych macierzy sztywności prowadzi do globalnej macierzy sztywności postaci

$$\mathbf{A} = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -4 & 2 \\ 0 & 2 & -2 \end{pmatrix}$$

W sposób w pełni analogiczny do tworzenia globalnej macierzy sztywności przebiega tworzenie globalnego wektora prawej strony. W pętli po wszystkich elementach tworzone są lokalne wektory prawej strony i następnie agregowane do globalnego wektora \mathbf{F} . Ponownie wyrazy lokalnego wektora prawej strony, tworzone w pętli po wszystkich funkcjach kształtu (węzłach elementu), mogą być obliczone analitycznie lub numerycznie (przez całkowanie).

W przypadku naszego problemu przykładowego lokalny wektor prawej strony obliczony na podstawie sformułowania słabego i postaci funkcji kształtu (taki sam dla e_1 i e_2) ma postać:

$$\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

co prowadzi do zagregowanej postaci wektora globalnego:

$$\mathbf{F} = \begin{pmatrix} 0.5 \\ 1.0 \\ 0.5 \end{pmatrix}$$

Końcowym krokiem konstruowania układu równań liniowych jest uwzględnienie warunków brzegowych. W przypadku warunków brzegowych Neumanna, włączonych do sformułowania MES, dokonuje się tego w sposób zbliżony do obliczania pozostałych wyrazów wektora prawej strony. W pętli po wszystkich elementach sprawdza się czy wierzchołki danego elementu stanowią brzeg obszaru obliczeniowego i jeśli tak jest, to modyfikuje się odpowiadające im wyrazy

lokalnego wektora prawej strony, a następnie lokalny wektor agreguje do globalnego wektora prawej strony.

W przypadku zadania przykładowego prowadzi to do modyfikacji globalnego wektora prawej strony do postaci:

$$\begin{pmatrix} 0.5 \\ 1.0 \\ -1.5 \end{pmatrix}$$

W przypadku warunków brzegowych Dirichleta, związanych z jawnym określeniem wartości funkcji aproksymującej u^h , jedną z możliwości jest bezpośrednio uwzględnienie warunku przez modyfikację funkcji aproksymującej (a dokładniej przestrzeni funkcji aproksymujących), co prowadzi do modyfikacji układu równań przez usunięcie pewnych równań i przyjęcie pewnych stopni swobody jako zadanych⁶.

W przypadku naszego zagadnienia brzegowego warunek Dirichleta ma postać $u^h(0) = 0$. Jako że w przyjętej postaci aproksymacji wartości stopni swobody są wartościami funkcji aproksymującej w punkcie, spełnienie warunku brzegowego prowadzi do zerowania składowej U_1^h wektora niewiadomych. Ze względu na matematyczną poprawność sformułowania MES wymóg zerowania w miejscu określenia warunku brzegowego Dirichleta dotyczy także funkcji testującej $w(x)$. W naszym przypadku oznacza to, że W_1 jest zawsze równe zero (w każdej funkcji testującej). W konsekwencji z układu równań liniowych znika pierwsze równanie odpowiadające W_1 . Jednocześnie na skutek zerowania U_1^h znika pierwsza kolumna układu. Ostateczna postać układu równań jest następująca:

$$\begin{aligned} -4U_2^h + 2U_3^h &= 1 \\ 2U_2^h - 2U_3^h &= -1.5 \end{aligned}$$

Rozwiązaniem układu jest wektor $\mathbf{U}^h = [0.25, 1.0]$, który wraz z założeniem $U_1^h = 0$ daje w wyniku funkcję aproksymującą

$$u^h(x) = 0.25\psi_2(x) + \psi_3(x)$$

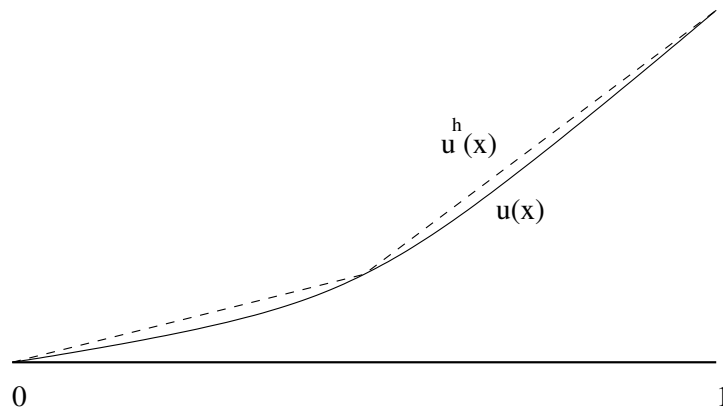
zilustrowaną na rys. 6.

Na rysunku widać jak rozwiązanie przybliżone $u^h(x)$ odbiega od rozwiązania dokładnego $u(x)$. Błąd numeryczny aproksymacji można mierzyć korzystając z normy błędu $\|u(x) - u^h(x)\|$, określającej precyzyjnie jak bardzo funkcja przybliżona odbiega od rozwiązania dokładnego. Dla wielu problemów udaje się ustalić oszacowania normy błędu, które często przyjmują postać zbliżoną do wzoru:

$$\|u(x) - u^h(x)\| < C \cdot h^p \cdot \left\| \frac{d^2 u}{dx^2} \right\|$$

(C jest pewną stałą zależną od sformułowania MES, h maksymalnym rozmiarem elementu w siatce, p stopniem aproksymacji, a rodzaj użytych norm ściśle zależy od sformułowania MES). Istotą powyższego wzoru są następujące fakty:

⁶Istnieją też inne sposoby uwzględnienia warunków brzegowych Dirichleta, bez usuwania równań z układu, jak np. tzw. metoda funkcji kary



Rysunek 6: Rozwiązanie dokładne i rozwiązanie przybliżone MES przykładowego zagadnienia brzegowego

- błąd numeryczny zależy od h i p – zmniejszając h (czyli zwiększając liczbę elementów w obszarze) lub zwiększając p można zmniejszać błąd numeryczny
- rozwiązanie MES zbiega się do rozwiązania dokładnego (błąd maleje do zera) dla h zmierzającego do zera.

Zbieżność rozwiązań przybliżonych MES jest jednym z podstawowych kryteriów poprawności modelu MES.

W powyższym wzorze na oszacowanie błędu numerycznego występuje niestety po prawej stronie nieznanne rozwiązanie dokładne $u(x)$. Oznacza to, że nie jesteśmy w stanie dokładnie zmierzyć błędu. Istnieje jednak szereg metod, które pozwalają w przybliżeniu określić błąd aproksymacji, i to nie tylko globalnie dla całego obszaru, ale lokalnie dla poszczególnych elementów. Metody te są podstawą tzw. adaptacyjnej MES, w której na podstawie lokalnego oszacowania błędu dokonuje się lokalnej modyfikacji siatki (np. zmniejszenia h lub zwiększenia p) i kolejno na coraz lepszych siatkach uzyskuje coraz dokładniejsze rozwiązania.