

Waldemar Rachowicz  
**PRACA HABILITACYJNA**  
ADAPTACYJNA METODA ELEMENTÓW  
SKOŃCZONYCH DO ROZWIĄZYWANIA  
RÓWNAŃ NAVIERA-STOKESA  
DLA PRZEPŁYWÓW ŚCIŚLIWYCH

7/1997

W A R S Z A W A 1 9 9 7

<http://rcin.org.pl>

Praca wpłynęła do Redakcji dnia 14 maja 1997r.

Recenzent – dr Jacek Rokicki



56552



**Praca habilitacyjna**

---

Instytut Podstawowych Problemów Techniki PAN  
Nakład 100 egz. Ark. wyd. 8,00 Ark. druk. 10,00  
Oddano do drukarni w lipcu 1997r.

---

**ATOS** Poligrafia-Reklama, Warszawa, Stawki 14

<http://rcin.org.pl>

# Rozdział 1

## Wstęp

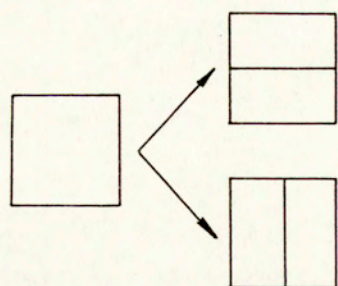
### 1.1 Tematyka pracy

Problematyka przepływów naddźwiękowych z małą lepkością jest jednym z trudniejszych wyzwań dla metod numerycznych w dynamice płynów. Najistotniejszą trudnością w zadaniach z tej dziedziny jest dokładna aproksymacja przepływu w bardzo cienkiej warstwie przyściennej formującej się przy powierzchni opływającego obiektu stałego. Mimo niewielkiej objętości zajmowanej przez warstwy przyścienne, znajomość przepływu właśnie w tych małych obszarach determinuje możliwość określenia obciążeń aerodynamicznych wywieranych na obiekt opływany przez gaz. Metodą, którą często na podstawie przesłanek intuicyjnych sugeruje się jako optymalną dla warstw przyściennych, jest zastosowanie dyskretyzacji z gęstością węzłów wzdłuż warstwy istotnie mniejszą niż w kierunku prostopadłym. W niniejszej pracy przedstawiamy praktyczne zastosowanie tej idei w postaci metody elementów skończonych z elementami silnie wydłużonymi. Metodę tę określamy jako „anizotropową adaptacyjną metodę elementów skończonych typu  $h$ ”. Technika ta pozwala na automatyczną konstrukcję siatek skończenie elementowych, w których rozmiar i wydłużenie elementów dopasowuje się do rozwiązania tak, aby osiągnąć wymaganą jego dokładność możliwie najmniejszym kosztem. Wspomniane dopasowanie – adaptacja – odbywa się poprzez podziały elementów, w tym podziały kierunkowe (anizotropowe) prowadzące do istotnego wydłużenia elementów (rys. 1.1). Podstawą do dokonywania tych podziałów jest oszacowanie błędu rozwiązania i analiza jego charakterystyk. Motywacją do opracowania metody adaptacyjnej był zamysł, aby w zakresie jej możliwości znalazła się klasa zadań interakcji fali uderzeniowej z warstwą przyścienną, gdzie napotykamy na swoistą koncentrację nieregularności rozwiązań.

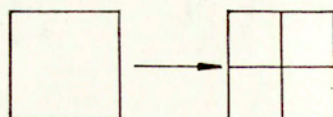
Praktyczne wykorzystanie anizotropowej metody adaptacyjnej stało się możliwe dzięki zastosowaniu następujących nowych wyników teoretycznych i procedur

1. Oszacowań *a priori* błędu interpolacji dla elementów o nieograniczonym wydłużeniu.
2. Oszacowań dokładności strumieni lepkich na ścianie opływającego obiektu.
3. Strategii anizotropowej adaptacji siatki skończenie elementowej.
4. Algorytmów iteracyjnego rozwiązywania równań możliwych do zastosowania dla siatek adaptacyjnych i niewrażliwych na obecność elementów silnie wydłużonych.

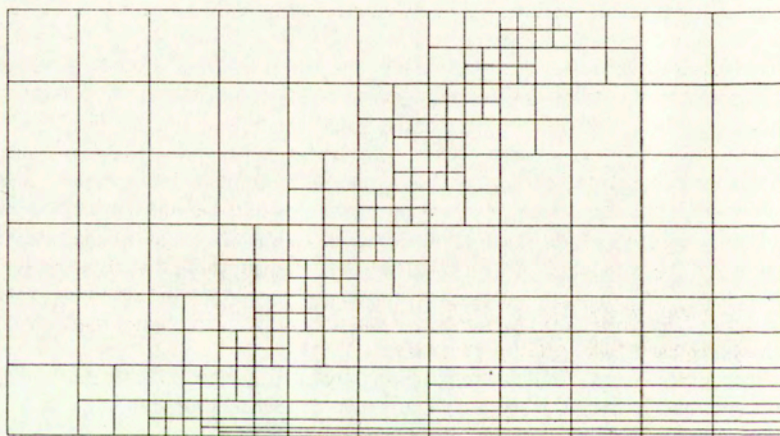
Wymienione oszacowania służą do skonstruowania strategii anizotropowej adaptacji. Pozwała ona na znaczną oszczędność liczby elementów, zwłaszcza w obszarach warstw przy-



podziały anizotropowe



podział izotropowy



Rysunek 1.1: Anizotropowe i izotropowe podziały elementów oraz przykładowa siatka adaptacyjna.

ściennych. Algorytm iteracyjny umożliwia efektywne rozwiązywanie układów równań wynikających z aproksymacji na złożonych siatkach adaptacyjnych. Omówienie tych zagadnień zajmuje najistotniejszą część pracy i stanowi jej twórczy wkład. Plan całej pracy jest zaś następujący:

Pozostała część rozdziału wstępnego zajmuje omówienie dotychczasowych wyników z zastosowaniem adaptacji do przepływów ściśliwych oraz przedstawienie motywacji do zastosowania metod adaptacyjnych. Rozdział 2. definiuje problem fizyczny będący przedmiotem rozważań, równania Naviera–Stokesa dla przepływów ściśliwych. W rozdziale 3. przedstawiono metodę SUPG jako technikę rozwiązywania problemów przepływów. Następna część pracy omawia najistotniejsze składniki instrumentarium adaptacyjnej metody elementów skończonych: technikę podziału elementów, oszacowania *a priori* i *a posteriori* będące podstawą kształtowania optymalnych siatek. Rozdział 5. jest poświęcony strategii adaptacji siatek dla przepływów z falami uderzeniowymi i warstwami przyściennymi. Szczególnie dużo miejsca poświęcamy w następnej części pracy algorytmom iteracyjnego rozwiązywania równań wynikających z dyskretyzacji skończonej elementowej. Doświadczenie obliczeniowe z zadaniami o większej skali wskazuje, iż jest to kluczowy algorytm w procedurze rozwiązywania problemu. Z drugiej zaś strony metodom rozwiązywania równań odpowiadających aproksymacji przepływów na siatkach adaptacyjnych nie poświęcono dotąd zbyt wiele uwagi. Algorytm proponowany przez nas to połączenie znanych procedur gradientów sprzężonych i GMRES z uwarunkowaniem wstępnym w postaci klasycznych metod Jacobiego i Gaussa–Seidela. Istotnym dokonaniem w tym zakresie jest zaprojektowanie wspomnianych metod w wersjach pozwalających na ich zastosowania dla siatek adaptacyjnych z elementami wydłużonymi oraz analiza teoretyczna zbieżności procedur iteracyjnych. Analiza ta jest nietypowa i „uciążliwa” ze względu na fakt, że dotyczy zadania, które nie ma cech regularnego problemu eliptycznego, cech, które „standardowo” w podobnych sytuacjach się zakłada.

W końcu ostatni rozdział pracy przedstawia zastosowanie anizotropowej metody adaptacyjnej do rozwiązania kilku testowych zadań. Są wśród nich problemy z „gładkim opływem” a także zadania z interakcją fal uderzeniowych z warstwą przyścienną. Zakres liczby Reynoldsa sięga  $2 \cdot 10^6$  w pierwszym przypadku i o rząd mniej w drugim. Dla analizowanych problemów uzyskano dobrą zgodność wyników obliczeń z pomiarami eksperymentalnymi lub z rozwiązaniami analitycznymi.

## 1.2 Dotychczasowe doświadczenia z zastosowaniem adaptacji do przepływów ściśliwych

Rzeczony adaptacyjnej metody elementów skończonych został zapoczątkowany przez I. Babuškę i W.C. Rheinboldta [10]–[13] w końcu lat 70. Od tamtego czasu I. Babuška koncentruje wokół siebie grupę współpracowników i studentów, z którymi rozwija zwłaszcza teoretyczne podstawy metod adaptacyjnych. Niemal równolegle zainicjowane zostały prace nad praktycznymi zastosowaniami metod adaptacyjnych w różnych dziedzinach mechaniki. Szczególnie aktywnymi promotorami tego kierunku badań są O.C. Zienkiewicz i J.T. Oden. Wraz ze swymi współpracownikami stworzyli podstawy zastosowań adaptacji w szerokiej klasie problemów od mechaniki ciała stałego, poprzez dynamikę płynów aż po zjawiska falowe i elektromagnetyczne.

Zastosowanie adaptacji w problemach przepływów zapoczątkowali w 1983 roku R.

Löhner, K. Morgan i O.C. Zienkiewicz pracą [67]. W następnych latach ukazały się dalsze prace tych autorów [69],[65] i [66], które przedstawiały zastosowanie techniki adaptacyjnej polegającej na generowaniu nowej siatki z zgęszczeniem elementów w obszarach dużego błędu (*remeshing*). Prace dotyczyły przepływów nielepkich. W większości podobną tematykę z różnymi wersjami algorytmu całkującego równania przepływu podejmują artykuły [48], [86],[68],[115],[74], w których do grona autorów dołączyli O. Hassan, J. Peraire, J. Reiro i M. Vahdata. W [68] Löhner zastosował w adaptacji możliwość wydłużenia elementów uzyskując istotną poprawę aproksymacji fal uderzeniowych. Algorytmy symulujące przepływy lepkie przedstawiono w pracach [47] i [103], z których szczególnie ostatnia prezentuje zakończone sukcesem symulacje przepływów z interakcją fali uderzeniowej i warstwy przyściennej. Jako technikę adaptacji zastosowano regenerację siatki zaś obszar warstwy przyściennej pokryto gęstą siatką strukturalną.

W pierwszej połowie lat 80. rozpoczął także prace nad adaptacją w mechanice płynów J.T. Oden z grupą współpracowników. Artykuł [31] L. Demkowicza, J.T. Odena, T. Strouboulisa i P. Devloo prezentuje ideę adaptacji, w której zagęszczenie elementów otrzymuje się przez ich podział na elementy mniejsze. Następne artykuły tych autorów [77] i [81] zawierały rozwinięcie idei adaptacji z dzieleniem elementów zaś w [99] zastosowano ideę przemieszczania węzłów dla zmiany rozkładu ich gęstości. Wspomniane prace a także nieco późniejsze [78],[80] i [36] podejmowały problem przepływów nielepkich i ograniczały się do adaptacji typu  $h$ , tj. z podziałem elementów. W artykule [35] z 1988 roku Devloo, Oden i Pattani przedstawili metodę adaptacyjną, w której możliwość podziałów uzupełniono podwyższeniem stopnia aproksymacji w warstwie przyściennej dla zadania lepkiego. Pozwoliło to na uzyskanie wiarygodnych wartości strumieni lepkich na opływanej powierzchni. Ideę zastosowania wyższych stopni aproksymacji  $p$  wraz z podziałem elementów, tzw. adaptacji typu  $hp$ , do przepływów lepkich przedstawili też L. Demkowicz, J.T. Oden i W. Rachowicz w [29]. W metodzie tej zastosowano także wielokrokowy algorytm całkowania zadania w czasie w oparciu o ideę G. Stranga pozwalającą na rozkład operatora różniczkowego problemu na składniki dla przepływów nielepkich i składniki dyfuzyjne: tarcie lepkie i przepływ ciepła. Artykuły [79],[30] i [82], których współautorami byli też O. Hardy, T. Liszka i T. Westermann, podejmują tematykę adaptacji typu  $hp$  i szacowania błędu w problemach przepływu. Adaptacyjne rozwiązywanie przepływów lepkich z dodatkową optymalizacją schematu całkowania problemu i uwzględnieniem interakcji płynu i opływającego ciała prezentują W. Tworzydło, J.T. Oden, E.A. Thornton i S.K. Youn w [107] i [104], przepływy z turbulencją C.Y. Huang, W. Tworzydło i J.T. Oden w [108]. W pewnym stopniu w sekwencję tych prac wpisuje się także artykuł [89] autora, z którego wiele idei zostało przedstawionych w niniejszej pracy.

Z upływem czasu zastosowanie różnych wersji adaptacyjnej metody elementów skończonych do symulacji przepływów naddźwiękowych zyskiwało na popularności ze względu na spektakularną poprawę dokładności rozwiązań przy względnie niewielkim koszcie obliczeń. Szczególnie szybki postęp w rozszerzaniu skali obliczeń następował dla przepływów nielepkich. Wśród artykułów z przełomu lat 80. i 90. można tu wymienić prace takich autorów jak G.S. Iannelli i A.J. Backer [58], R. Ramakrishnan, K.S. Bey, E.A. Thornton [93], T.E. Tezduyar i J. Liou [102], E.A. Thornton, P. Dechaumphai, G. Venaganti [104] czy P. Hansbo i C. Johnson [46] i J. Argyris, I.S. Doltsinis i H.Friz [8] a także C.E. Baumann, M.A. Stori, S.R. Idelson [16] i J. Szmelter, M.J. Marchant, A. Evans, N.P. Weatherill [100] prezentujących skuteczność różnych odmian adaptacji.

Wśród prac nieco późniejszych spektakularną efektywnością adaptacji odznaczają się

wyniki prezentowane przez M.J. Castro-Díaz i jego współpracowników [24],[25] oraz przez zespół M. Fortina [43],[44]. Proponowane przez nich procedury stosują zarówno regenerację siatek jak i lokalne ich modyfikacje oparte na przesuwaniu i tworzeniu nowych węzłów dla zmniejszenia błędu interpolacji rozwiązania, prowadząc do siatek o bardzo zróżnicowanych rozmiarach i proporcjach elementów. Efektywną generację siatek o gwałtownie zmieniającej się gęstości elementów prezentują też D.L. Marcum i N.P. Weatherill [72] oraz R.C. Almeida i A.C. Galeao [5].

Adaptację na siatkach strukturalnych dla zadań o większej skali prezentują R. Hentschel i E.H. Hirschel [49] dla dwu- i trójwymiarowych przepływów nielepkich. J.C. Jouhand i M. Barrel w [62] zaproponowali natomiast rodzaj adaptacji na kawałkami równomiernych siatkach o strukturze hierarchicznej, która nie jest szczególnie efektywna w redukcji liczby zmiennych, lecz pozwala za to na zastosowanie algorytmów wielosiatkowych. Bardzo szybkie i niezawodne algorytmy regeneracji siatek adaptacyjnych zaprezentowali R. Tilch, K. Morgan i N.P. Weatherill w [106], stosując je do symulacji zadań niestacjonarnych z siatkami zmieniającymi się wraz z ewolucją rozwiązania. Adaptację dla tej klasy problemów na siatkach strukturalnych stosowali K. Banaś i L. Demkowicz [14].

Szczególnie obiecujące dla rozszerzenia skali problemów, których rozwiązanie jest osiągalne przy danej mocy obliczeniowej, jest połączenie adaptacji z obliczeniami równoległymi (na wielu procesorach). Zastosowanie adaptacji w kontekście równoległości podejmują m.in. I.St. Doltsinis i J. Urban [37], R. Richter i P. Leyland [94], Z. Johan, T.J.R. Hughes, K.K. Mathur i S.L. Johnsson [59],[60] oraz K. Banaś i J. Płażek [15].

Istotnie wolniejszy okazał się postęp w obliczeniach przepływów lepkich. Oprócz wspomnianych już artykułów [62] i [103] uwieńczone powodzeniem symulacje przepływów z interakcją fal uderzeniowych i warstw przyściennych z liczbą Reynoldsa rzędu  $10^5$  udokumentowano w pracach G. Vemaganti i A. Wietinga [111], P. Dechaumphai [28] oraz G.H. Klopfera i H.C. Yee [63].

### 1.3 Motywacja

Zastosowanie metody elementów skończonych do symulacji przepływów ma kilkunastoletnią już historię. Mimo rozwoju wielu efektywnych algorytmów numerycznej mechaniki płynów, istnieje powszechna świadomość ograniczoności klasy zagadnień, które można rozwiązać na dostępnym obecnie sprzęcie komputerowym. Otrzymanie wiarygodnych symulacji przepływów wokół rzeczywistych obiektów inżynierskich wymaga niezwykle małych komórek obliczeniowych – elementów, co przy ich równomiernej dystrybucji prowadzi do ogromnej liczby zmiennych modelu dyskretnego. Przykładowo, dla obiektu lotniczego o jednostkowym wymiarze grubości warstwy przyściennej można oszacować zgrubnie jako  $Re^{-1/2}$ , gdzie  $Re$  jest liczbą Reynoldsa. Zakładając, że 10 elementów wystarcza do dyskretyzacji takiego przedziału z realistyczną dokładnością, liczbę zmiennych dyskretnych problemu możemy oszacować jako rzędu  $(10 \cdot Re^{-1/2})^2$  w dwu i  $(10 \cdot Re^{-1/2})^3$  w trzech wymiarach. Dla typowych zadań przemysłu lotniczego rząd liczby Reynoldsa „zaczyna się” od  $10^6$ , co oznacza odpowiednio  $10^8$  i  $10^{12}$  punktów obliczeniowych. Słowem kluczowym w tym oszacowaniu jest „równomierna dystrybucja”: dyskretyzacja o przytoczonej gęstości jest konieczna jedynie w pewnych podobszarach o dość ograniczonej objętości, w warstwach przyściennych czy, w mniejszym stopniu, wzdłuż fal uderzeniowych. W pozostałej części obszaru dyskretyzacja może być o wiele rzadsza, rozwiązanie zadania przepływu jest tam bardzo gładkie, często niemal stałe. Co więcej, rozwiązanie

w warstwach przyściennych zmienia się bardzo gwałtownie tylko wzdłuż grubości warstwy, podczas gdy wzdłuż opływającego obiektu jest względnie gładkie, przez co zastosowanie elementów bardzo silnie wydłużonych w tych obszarach daje dodatkowe zmniejszenie liczby zmiennych. Kombinacja dwu wymienionych technik prowadzi do radykalnej redukcji liczby stopni swobody koniecznych dla wiarygodnej symulacji przepływu. Pozwala zmniejszyć rozmiar zadań dyskretnych do poziomu będącego w zasięgu mocy obliczeniowych współczesnych komputerów.

Zastosowanie siatek z nierównomiernym rozkładem gęstości elementów dopasowanym do nieregularności rozwiązania (np. warstw przyściennych i fal uderzeniowych) wymaga informacji o ich rozmiarach i lokalizacji. Nie jest ona *a priori* znana. Procedura, którą można w praktyce zastosować, polega na otrzymaniu rozwiązania zgrubnego na pewnej siatce początkowej oraz znalezieniu oszacowania *a posteriori* błędu tej aproksymacji. W następnym kroku dokonujemy podziału elementów ze wskaźnikiem błędu przekraczającym pewien próg otrzymując nową siatkę, *adaptowaną* do rozwiązania. Procedurę tę powtarzamy kilkakrotnie, dopóki otrzymane rozwiązanie nie spełni wymaganych kryteriów dokładności. Przedstawiona technika to adaptacyjna metoda elementów skończonych.

W niniejszej pracy przedstawiamy ogół zagadnień związanych z zastosowaniem adaptacyjnej metody elementów skończonych w mechanice płynów. Wiele z aspektów tej metodologii jest bardzo podobnych lub wręcz identycznych z procedurami dla zadań eliptrycznych. Nowym zagadnieniem, wokół którego koncentruje się większość tej pracy, jest zastosowanie elementów silnie wydłużonych do aproksymacji warstw przyściennych. Technika ta wydaje się jak dotąd jedyną możliwą drogą rozwiązywania problemów z dużą liczbą Reynoldsa. Model fizyczny, do którego się ograniczono, to naddźwiękowy przepływ gazu doskonałego z małą lepkością w warunkach stanu ustalonego. Klasą zagadnień testowych, dla których proponowana technika powinna być skuteczna, są tzw. zadania interakcji fali uderzeniowej z warstwą przyścienną. W zadaniach tego rodzaju występuje dodatkowe znaczne pocienienie warstwy przyściennej i związany z nim lokalny silny wzrost ciśnienia, strumienia ciepła i tarcia powierzchniowego. Z tego względu zadania interakcji uchodzą za szczególnie trudne do symulacji. Mają one jednocześnie istotne znaczenie praktyczne w inżynierii lotniczej. Wspomniana koncentracja obciążeń aerotermicznych może być bowiem przyczyną przepalenia lub stopienia konstrukcji lotniczej.

Symulacja problemów z dużą liczbą Reynoldsa jest możliwa także w dość ograniczonym zakresie bez podejścia adaptacyjnego. Jest ona oparta na możliwości zgrubnego oszacowania grubości warstw przyściennych i na zastosowaniu siatek strukturalnych (tj. odpowiednio zniekształconych siatek równomiernych) z oczkiem o właściwym rozmiarze w obszarze przewidywanej warstwy przyściennej. Równomierność siatki – w sensie topologicznym – pozwala na zastosowanie bardzo efektywnych algorytmów przystosowanych do obliczeń równoległych, dzięki czemu czas obliczeń mimo wielkiej liczby węzłów ulega znacznej redukcji. Tym niemniej podejście to trudno jest zastosować do ogólnych problemów interakcji, z możliwą złożoną geometrią, w których grubość warstwy przyściennej może zmieniać się wzdłuż brzegu w trudny do przewidzenia sposób o jeden czy nawet dwa rzędy wielkości. Poza tym obliczenia na siatkach adaptacyjnych także mogą być prowadzone równoległe, choć z nieco mniejszą efektywnością niż dla siatek równomiernych.



## 1.4 Efektywność adaptacji

Idea adaptacji w metodzie elementów skończonych wyraża bardzo intuicyjny pomysł wzbogacania dyskretyzacji selektywnie w obszarach, gdzie rozwiązanie przybliżone jest istotnie mniej dokładne niż w pozostałej części obszaru obliczeniowego. Intencją takiego działania jest ulepszenie rozwiązania przy możliwie najmniejszym koszcie obliczeń. Największą efektywność opisaney procedury obserwuje się w przypadku rozwiązań bardzo nieregularnych.

Ilościową ocenę korzyści wynikających z adaptacji w metodzie elementów skończonych można otrzymać rozpatrując zagadnienie optymalnej interpolacji danej funkcji. Pomysł konstruowania optymalnych siatek, z którego korzystamy poniżej, pochodzi z pracy [10] I. Babuški i W.C. Rheinboldta z 1978 roku.

Poszukujemy siatki skończenie elementowej, która byłaby optymalna w sensie minimalizacji błędu interpolacji przy ustalonej liczbie elementów. Ograniczymy się dla uproszczenia do siatek prostokątnych. Rozpatrzmy oddzielnie adaptację z podziałem elementów na 4 „elementy-synów”, izotropową, i adaptację z podziałem na 2 elementy, anizotropową.

Dla elementów w kształcie kwadratu z biliniowymi funkcjami kształtu błąd interpolacji w pólnormie Sobolewa  $H^1$  szacuje się jako

$$\|u - u_I\|_{1,K}^2 \leq Ch_K^2 \int_K (u_{,xx}^2 + u_{,yy}^2) dx \quad (1.4.1)$$

gdzie,  $u_I$  jest interpolantem funkcji  $u$ ,  $h_K$  jest rozmiarem elementu  $K$ , [26]. Będziemy minimalizować sumę wyrażeń z prawej strony (1.4.1) przy założeniu, że liczba elementów wynosi  $N_i$ . Oznacza to, że zamiast minimalizować właściwy błąd, zminimalizujemy raczej analityczne wyrażenie ograniczające błąd, nazywając je jednak, dla uproszczenia, ciągle „błędem”. Zakładając, że  $h(\mathbf{x})$  jest poszukiwaną (kawalkami stałą) funkcją formułujemy następujący problem minimalizacji:

Znaleźć  $h(\mathbf{x})$  takie że:

$$e^2 := \int_{\Omega} h(\mathbf{x})^2 (u_{,xx}^2 + u_{,yy}^2) dx = \min \int_{\Omega} \frac{dx}{h^2(\mathbf{x})} = N_i \quad (1.4.2)$$

Druga z całek w (1.4.2) wyraża liczbę elementów,  $e$  oznacza błąd. Do rozwiązania problemu stosujemy standardową metodę mnożników Lagrange'a:

$$\int_{\Omega} 2h(u_{,xx}^2 + u_{,yy}^2) \delta h dx - \lambda \int_{\Omega} 2 \frac{dx}{h^3} \delta h = 0, \forall \delta h \quad (1.4.3)$$

gdzie  $\delta h(\mathbf{x})$  jest wariacją funkcji  $h(\mathbf{x})$ . Powyższy warunek pociąga za sobą kryterium optymalności:

$$h^4 (u_{,xx}^2 + u_{,yy}^2) = \lambda \quad (1.4.4)$$

Ponieważ  $h(\mathbf{x})$  musi być kawalkami stałe, powyższy warunek można spełnić jedynie w przybliżeniu. Jeśli dodatkowo weźmiemy pod uwagę, że dla odpowiednio małych elementów pochodne  $u(\mathbf{x})$  zmieniają się w nich nieznacznie, będziemy mogli zinterpretować (1.4.4) następująco:

$$\frac{h^4}{h^2} (u_{,xx}^2 + u_{,yy}^2) = \lambda \quad (1.4.5)$$

Kryterium ((1.4.5)) jest określane jako *zasada ekwidystribucji błędów* dla siatek optymalnych. Jest ona uzasadnieniem dla intuicyjnie akceptowanej strategii adaptacji siatek: „dziel elementy z największymi błędami”, gdyż taka procedura prowadzi do przybliżonej ekwidystribucji błędów.

Równania (1.4.2)<sub>II</sub> i (1.4.5) wskazują, że  $\lambda = e^2/N_u$ . Wobec tego, wyrażając  $h(x)$  z (1.4.4),  $h(x) = [e^2/N_i(u_{,xx}^2 + u_{,yy}^2)^{-1}]^{1/4}$  i wstawiając do ((1.4.2))<sub>I</sub> otrzymujemy następującą zależność między  $e$  i  $N_i$

$$N_i = 1/e^2 \left[ \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2)^{1/2} dx \right]^2 \quad (1.4.6)$$

Analogiczny związek między liczbą elementów  $N_u$  i błędem dla siatek równomiernych otrzymamy zauważając, że

$$e^2 = h^2 \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2) dx \quad (1.4.7)$$

i  $h^2 = S/N_u$ , gdzie  $S = \text{meas}(\Omega)$ . Tak więc mamy

$$N_u = 1/e^2 \cdot S \cdot \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2) dx \quad (1.4.8)$$

Stosując nierówność Cauchy'ego-Schwarza

$$\left[ \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2)^{1/2} \cdot 1 dx \right]^2 \leq \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2) dx \cdot \int_{\Omega} 1 dx \quad (1.4.9)$$

znajdujemy, że, wobec (1.4.6) i (1.4.9),

$$N_i \leq N_u \quad (1.4.10)$$

co dowodzi oczekiwanej przewagi siatek adaptacyjnych nad siatkami równomiernymi.

W przypadku minimalizacji błędu w normie  $L^2$  oszacowanie ((1.4.1)) zastępujemy przez

$$\|u - u_I\|_{0,K}^2 \leq Ch_K^4 \int_K (u_{,xx}^2 + u_{,yy}^2) dx \quad (1.4.11)$$

zaś rozumowanie analogiczne do przedstawionego powyżej prowadzi do związku liczby elementów i błędu dla siatek optymalnych i równomiernych jako

$$N_i = 1/e \left[ \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2)^{1/3} dx \right]^{3/2} \quad (1.4.12)$$

i

$$N_u = 1/e \cdot S \cdot \left[ \int_{\Omega} (u_{,xx}^2 + u_{,yy}^2) dx \right]^{1/2} \quad (1.4.13)$$

Stosując nierówność Höldera

$$\int_{\Omega} |wv| dx \leq \left( \int_{\Omega} |v|^p dx \right)^{1/p} \left( \int_{\Omega} |w|^q dx \right)^{1/q}, \text{ dla } 1 \leq p, q, \frac{1}{p} + \frac{1}{q} = 1 \quad (1.4.14)$$

przy  $v \equiv (u_{,xx}^2 + u_{,yy}^2)^{1/3}$ ,  $w \equiv 1$ ,  $p = 3$ ,  $q = 3/2$ , podobnie jak poprzednio znajdujemy, że

$$\text{http://rcin.org.pl} \quad (1.4.15)$$

Przejdźmy obecnie do podobnych oszacowań dla anizotropowej metody adaptacji. W tym przypadku ze względu na dopuszczalne (a nawet intencjonalne) silne wydłużenie elementów, geometrię elementu prostokątnego charakteryzujemy jego dwoma wymiarami  $h_x$  i  $h_y$ . Oszacowanie *a priori* błędu interpolacji dla elementów wydłużonych, które dyskutujemy w szczególności w rozdziale 4.3, ma następującą postać

$$\|u - u_I\|_{1,K}^2 \leq Ch_x^2 \int_K u_{,xx}^2 dx + Ch_y^2 \int_K u_{,yy}^2 dx \quad (1.4.16)$$

Zadanie minimalizacji błędu analogiczne do (1.4.2) możemy przedstawić w postaci:

Znaleźć  $h_x(\mathbf{x})$ ,  $h_y(\mathbf{x})$ , takie że

$$e^2 := \int_{\Omega} [h_x^2(\mathbf{x})u_{,xx}^2 + h_y^2(\mathbf{x})u_{,yy}^2] dx = \min \int_{\Omega} \frac{dx}{h_x(\mathbf{x})h_y(\mathbf{x})} = N_a \quad (1.4.17)$$

Tak jak i poprzednio stosujemy metodę mnożników Lagrange'a do optymalizacji z więzami

$$\int_{\Omega} (2h_x u_{,xx}^2 \delta h_x + 2h_y u_{,yy}^2 \delta h_y) dx - \lambda \int_{\Omega} \left( \frac{\delta h_x}{h_x^2 h_y} + \frac{\delta h_y}{h_x h_y^2} \right) dx = 0, \quad \forall \delta h_x, \delta h_y \quad (1.4.18)$$

gdzie  $\delta h_x, \delta h_y$  są wariacjami funkcji  $h_x(\mathbf{x})$  i  $h_y(\mathbf{x})$ . Otrzymujemy stąd, że

$$h_x^3 h_y u_{,xx}^2 = \lambda/2 \quad \text{i} \quad h_y^3 h_x u_{,yy}^2 = \lambda/2, \quad (1.4.19)$$

co z kolei, dla odpowiednio małych elementów, da się zapisać jako

$$h_{x,K}^2 \int_K u_{,xx}^2 dx = \lambda/2, \quad h_{y,K}^2 \int_K u_{,yy}^2 dx = \lambda/2 \quad (1.4.20)$$

Warunki te są odpowiednikiem zasady ekwidystribucji błędów dla adaptacji anizotropowej. Z (1.4.20) i (1.4.17)<sub>1</sub> wynika, że  $\lambda = e^2/N_a$ . Z (1.4.19) możemy więc wyrazić  $h_x$  i  $h_y$

$$h_x = \left( \frac{e^2 u_{,yy}}{N_a u_{,xx}} \right)^{1/4}, \quad h_y = \left( \frac{e^2 u_{,xx}}{N_a u_{,yy}} \right)^{1/4} \quad (1.4.21)$$

co z kolei po wstawieniu do (1.4.17)<sub>1</sub> prowadzi do następującej zależności błędu i liczby elementów:

$$N_a = 1/e^2 \left( \int_{\Omega} \sqrt{2|u_{,xx}u_{,yy}|} dx \right)^2 \quad (1.4.22)$$

Zauważmy, że, ponieważ  $2|u_{,xx}u_{,yy}| \leq u_{,xx}^2 + u_{,yy}^2$ , (1.4.8) i (1.4.22) wskazują, że

$$N_a \leq N_i \leq N_u \quad (1.4.23)$$

co jest oczekiwanym wynikiem, gdyż siatki izotropowe są szczególnym przypadkiem siatek anizotropowych, tj. optymalizacja (1.4.17) odbywa się po większym zbiorze niż (1.4.2).

Błąd interpolacji w normie  $L^2$  dla elementów wydłużonych szacuje się następująco (zob. rozdział 4.3)

$$\|u - u_I\|_{0,K}^2 \leq \frac{1}{6} \left( \int_K u_{,xx}^2 dx + \int_K u_{,yy}^2 dx \right) \quad (1.4.24)$$

Norma	a	$N_u/N_i$	$N_u/N_a$	Norma	a	$N_u/N_i$	$N_u/N_a$
$H^1$	5	1.65	2.66	$L^2$	5	1.47	4.53
	10	2.67	5.33		10	2.16	9.36
	20	5.11	12.74		20	3.79	21.98
	40	10.16	29.85		40	7.17	53.49

Tablica 1.1: Zależność  $N_u/N_i$  i  $N_u/N_a$  od parametru  $a$  dla funkcji  $u_1$  (1.4.27) i błędów w półnormie  $H^1$  i w normie  $L^2$ .

Norma	a	$N_u/N_i$	$N_u/N_a$	Norma	a	$N_u/N_i$	$N_u/N_a$
$H^1$	5	1.69	11.27	$L^2$	5	1.47	6.17
	10	2.88	72.23		10	2.35	20.03
	20	5.35	607.27		20	4.23	76.28
	40	10.34	5058.46		40	8.06	298.94

Tablica 1.2: Zależność  $N_u/N_i$  i  $N_u/N_a$  od parametru  $a$  dla funkcji  $u_2$  (1.4.28) i błędów w półnormie  $H^1$  i w normie  $L^2$ .

Rozumowanie podobne do przedstawionego powyżej pozwala znaleźć zależność błędu w normie  $L^2$  i liczby elementów siatki optymalnej w postaci

$$N_a = 1/e \left( \int_{\Omega} |2u_{,xx}u_{,yy}|^{1/3} dx \right)^{2/3} \quad (1.4.25)$$

co, ze względu na (1.4.12) i  $2|u_{,xx}u_{,yy}| \leq u_{,xx}^2 + u_{,yy}^2$  pokazuje, że

$$N_a \leq N_i \leq N_u \quad (1.4.26)$$

Przedstawione wyżej oszacowania wykazują przewagę siatek optymalnych anizotropowych nad optymalnymi izotropowymi a tych z kolei nad równomiernymi (nierówności (1.4.26) i (1.4.23)). Faktyczna proporcja między liczbami elementów  $N_a$ ,  $N_i$  i  $N_u$  dla ustalonego błędu  $\epsilon$  zależy oczywiście od interpolowanej funkcji  $u(\mathbf{x})$ . Dla zilustrowania tych relacji rozpatrzmy dwie modelowe funkcje

$$u_1 = \arctan a(r - r_0), \quad r = \sqrt{x^2 + y^2}, \quad r_0 = 1/2 \quad (1.4.27)$$

oraz

$$u_2 = (1 - e^{a(x-x_0)})(1 - e^{a(y-y_0)}), \quad x_0 = y_0 = 1 \quad (1.4.28)$$

w obszarze  $(x, y) \in [0, 1]^2$ . Funkcja  $u_1(\mathbf{x})$  jest modelem rozwiązania z rozmytą nieciągłością wzdłuż okręgu  $x^2 + y^2 = r_0^2$ , zaś  $u_2(\mathbf{x})$  charakteryzuje się warstwą przysięnną wzdłuż boków  $x = 1$  i  $y = 1$ . Parametr  $a$  służy do kontroli nieregularności obydwu funkcji: ze wzrostem  $a$  rośnie stromość  $u_1$  wzdłuż okręgu  $x^2 + y^2 = r_0^2$  oraz stromość  $u_2$  w warstwie przysięnniej. Tabele 1.1 i 1.2 przedstawiają proporcje  $N_u/N_a$  i  $N_u/N_i$  dla błędów w półnormie  $H^1$  i w normie  $L^2$  odpowiednio dla funkcji  $u_1$  i  $u_2$  przy różnych wartościach parametru  $a$ . Zauważmy, że redukcja liczby elementów wynikająca z adaptacji rośnie bardzo znacząco wraz ze wzrostem  $a$  osiągając spektakularne wartości. Przytoczone przykłady wskazują, że dla pewnej klasy silnie nieregularnych rozwiązań zastosowanie adaptacji jest absolutną koniecznością. Jest kwestia możliwości bądź braku możliwości rozwiązania zadania.

## Uwagi krytyczne

Przytoczone powyżej oszacowanie efektywności metod adaptacyjnych może wzbudzać entuzjazm, jednakże nie zawiera ono „całej prawdy”. Na oczekiwania dotyczące korzyści z adaptacji powinniśmy spojrzeć także z pewnym krytycyzmem. Poniżej przedstawiamy kilka jego przyczyn, zarówno formalnych jak i natury praktycznej.

1. W przeprowadzonych rachunkach zaniedbywaliśmy wartości stałych występujących w oszacowaniach *a priori*. Nie zmienia to relacji między  $N_u$ ,  $N_i$  i  $N_a$  w przypadku błędów w normie  $L^2$ , dla których stałe w oszacowaniu izotropowych i anizotropowym są jednakowe (por. rozdział 4.3). Dla oszacowania w półnormie  $H^1$  nie znamy jednak proporcji stałych w oszacowaniu izo- i anizotropowym. Należy się spodziewać, że jej uwzględnienie mogłoby nieco zmodyfikować  $N_i$  w tym przypadku.

(Oczywiście stałe w oszacowaniach *a priori* można zawsze zrównać przez odpowiednie zwiększenie mniejszej z nich – psuje to jednak optymalność oszacowania, co omawiamy poniżej.)

2. Wielkością, którą minimalizowaliśmy, był nie tyle faktyczny błąd interpolacji, co wartość analitycznego wyrażenia z oszacowania *a priori*. Wynik minimalizacji mógłby okazać się zupełnie bezwartościowy, gdyby oszacowanie to było nieoptymalne, tj. gdyby wyniki oszacowania znacznie przekraczały faktyczne błędy. Przez optymalność oszacowania takiego jak np. (1.4.1) rozumiemy możliwość obustronnego ograniczenia błędu przez dane wyrażenie, tj. sytuację, gdy istnieją takie możliwie bliskie stałe  $C > c > 0$ , że

$$ch_K^2 \int_K (u_{,xx}^2 + u_{,yy}^2) dx \leq |u - u_I|_{1,K}^2 \leq Ch_K^2 \int_K (u_{,xx}^2 + u_{,yy}^2) dx \quad (1.4.29)$$

dla pewnej klasy funkcji. Relacja taka pozwala na ocenę proporcji faktycznych błędów interpolacji przy równości ograniczających wyrażen. Pierwszego z oszacowań w (1.4.29), które wydaje się być bardzo naturalnym i pożądanym uzupełnieniem drugiej, standardowej części (1.4.29), nie udało nam się znaleźć w popularnych podręcznikach metody elementów skończonych. W dodatku A pokazujemy, że jest ono prawdziwe asymptotycznie dla funkcji klasy  $C^3(\bar{\Omega})$ .

3. W praktycznych zadaniach interesuje nas błąd *aproksymacji* problemu nie zaś błąd hipotetycznej interpolacji rozwiązania. Jest korzystne, jeśli istnieje relacja między tymi dwoma rodzajami błędów, np. dla liniowych zadań eliptycznych drugiego rzędu błąd interpolacji ogranicza błąd aproksymacji (lemat Cea). Jeśli nie dysponujemy związkiem tego typu, możemy jedynie mieć intuicyjne przekonanie, że dla poprawy dokładności rozwiązania korzystnie jest stosować siatkę, która pozwala to rozwiązanie dokładnie interpolować.

4. Wiele faktycznych algorytmów dla złożonych zagadnień mechaniki może nie dopuszczać stosowania skomplikowanych siatek adaptacyjnych, bądź też reagować na nie pogarszaniem się stabilności, zbieżności czy innych istotnych charakterystyk. Powoduje to konieczność różnych ograniczeń w adaptacji a zatem i zmniejszenie spodziewanych korzyści.

## 1.5 Uwagi o terminologii i oznaczeniach

Przez tytułowe „równania Naviera–Stokesa dla przepływów ściśliwych” rozumiemy równania dynamiki gazu doskonałego z uwzględnieniem ściśliwości i lepkości zapisane w postaci praw zachowania. Odpowiedni termin angielski to *compressible Navier–Stokes equations*.

Dla pojęć, których polskie nazewnictwo nie jest utrwalone, podano odpowiednie nazwy angielskie przy pierwszym pojawieniu się terminu w tekście.

Stosujemy standardowe oznaczenia przestrzeni funkcyjnych Sobolewa dla  $m$  całkowitych nieujemnych i dla  $1 \leq p \leq \infty$

$$\begin{aligned} W^{m,p}(\Omega) &= \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \text{ dla } 0 \leq |\alpha| \leq m\} \\ H^m(\Omega) &= W^{m,2}(\Omega) \end{aligned} \quad (1.5.30)$$

gdzie dla  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $\alpha_i \geq 0$ , całkowitych,  $D^\alpha u$  oznacza dystrybucyjną pochodną cząstkową funkcji  $u$  rzędu  $\alpha_1$  względem  $x_1, \dots$ , rzędu  $\alpha_n$  względem  $x_n$ ,  $|\alpha| = \alpha_1 + \dots + \alpha_n$ ;  $\Omega \subset \mathbb{R}^n$  oraz

$$\begin{aligned} L^p(\Omega) &= \{u : \Omega \rightarrow \mathbb{R} \mid \int_{\Omega} |u|^p dx < \infty\}, \quad 1 \leq p < \infty \\ L^\infty(\Omega) &= \{u : \Omega \rightarrow \mathbb{R} \mid \operatorname{ess\,sup}_{x \in \Omega} |u| < \infty\}. \end{aligned} \quad (1.5.31)$$

W przestrzeniach  $L^p(\Omega)$ ,  $W^{m,p}(\Omega)$  i  $H^m(\Omega)$  definiujemy normy

$$\begin{aligned} \|u\|_{L^p(\Omega)} &= \left( \int_{\Omega} |u|^p dx \right)^{1/p}, \quad 1 \leq p < \infty, \quad \|u\|_{L^\infty(\Omega)} = \operatorname{ess\,sup}_{x \in \Omega} |u| \\ \|u\|_{W^{m,p}(\Omega)} &\equiv \|u\|_{m,p,\Omega} = \left( \sum_{|\alpha| \leq m} \int_{\Omega} |D^\alpha u|^p dx \right)^{1/p}, \quad 1 \leq p < \infty \\ \|u\|_{W^{m,\infty}(\Omega)} &\equiv \|u\|_{m,\infty,\Omega} = \max_{|\alpha| \leq m} \operatorname{ess\,sup}_{x \in \Omega} |D^\alpha u| \\ \|u\|_{H^m(\Omega)} &\equiv \|u\|_{m,\Omega} = \|u\|_{m,2,\Omega} \end{aligned} \quad (1.5.32)$$

Wektory z przestrzeni  $\mathbb{R}^n$  i macierze oznaczane są tłustą czcionką, np.  $\mathbf{Ax} = \mathbf{b}$ ,  $c = \mathbf{x}^T \mathbf{y}$ .

Stosujemy umowę sumacyjną Einsteina oraz zapis z przecinkiem pochodnych względem czasu i zmiennych przestrzennych,  $\partial a / \partial t = a_{,t}$ ,  $\partial a / \partial x_i = a_{,i}$ .

W wielu sekwencjach oszacowań  $C$  i  $c$  (z możliwymi wskaźnikami) oznaczają stałe o niekoniernie tej samej wartości, niezależne od parametrów dyskretyzacji  $h$  i funkcji występujących w oszacowaniach, o ile zależność ta nie jest jawnie zasygnalizowana.

## Rozdział 2

# Równania Naviera–Stokesa dla przepływów ściśliwych

Równania Naviera–Stokesa dla przepływów ściśliwych reprezentują podstawowe prawa zachowania mechaniki, prawo zachowania masy, pędu i całkowitej energii wraz ze związkami konstytutywnymi dla gazu doskonałego. Wymienione prawa zachowania często zapisujemy w postaci całkowitej z użyciem pojęcia pochodnej materialnej całki:

$$\frac{D}{Dt} \int_{\Omega} a \, dx \equiv \int_{\Omega} \left[ \frac{\partial a}{\partial t} + \frac{\partial}{\partial x_i} (u_i a) \right] dx \quad (2.0.1)$$

przedstawiającej szybkość zmian wielkości  $\int_{\Omega} a \, dx$  w czasie, jeśli  $\Omega$  jest obszarem związanym z cząstkami materialnymi zawartymi w  $\Omega$  w chwili początkowej. Jeśli stan gazu opiszemy za pomocą następujących zmiennych (konserwatywnych, *conservative variables*):

$\rho$  - gęstości płynu,

$\mathbf{m} = (m_1, m_2)$  - wektora pędu na jednostkę objętości ( $\mathbf{m} = (m_1, m_2, m_3)$  w trzech wymiarach),

$e$  - gęstości całkowitej energii, tj. energii wewnętrznej i kinetycznej płynu na jednostkę objętości,

oraz jeśli oznaczymy przez:

$\boldsymbol{\tau} = \{\tau_{ij}\}_{i,j=1,2}$  - tensor naprężeń wywołanych lepkością,

$p$  - ciśnienie termodynamiczne,

$\mathbf{q} = \{q_i\}_{i=1,2}$  - strumień ciepła przewodzonego przez płyn,

to zasady zachowania przyjmą następującą postać:

zasada zachowania masy:

$$\frac{D}{Dt} \int_{\Omega} \rho \, dx = 0 \quad (2.0.2)$$

zasada zachowania pędu:

$$\frac{D}{Dt} \int_{\Omega} \mathbf{m} \, dx = \int_{\partial\Omega} (\boldsymbol{\tau} - p\mathbf{I})\mathbf{n} \, dS \quad (2.0.3)$$

zasada zachowania energii:

$$\frac{D}{Dt} \int_{\Omega} e \, dx = \int_{\partial\Omega} [\mathbf{u} \cdot (\boldsymbol{\tau} - p\mathbf{I})\mathbf{n} - \mathbf{q} \cdot \mathbf{n}] \, dS \quad (2.0.4)$$

gdzie dodatkowo  $\mathbf{n}$  oznacza wektor normalny do brzegu  $\partial\Omega$ ,  $\mathbf{u} = \mathbf{m}/\rho$  jest wektorem prędkości płynu,  $\mathbf{I}$  oznacza macierz jednostkową. Stosując związek (2.0.1) oraz

korzystając z twierdzenia Stokesa o zamianie całki objętościowej na powierzchniową możemy wyrazić równości (2.0.2), (2.0.3) i (2.0.4) przez wyłącznie całki objętościowe, których (postulowane) znikanie dla dowolnie wybranych obszarów  $\Omega$  pociąga za sobą lokalną postać praw zachowania

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j} [m_j] = 0 \quad (2.0.5)$$

$$\frac{\partial m_i}{\partial t} + \frac{\partial}{\partial x_j} \left[ \frac{m_i m_j}{\rho} + \delta_{ij} p \right] = \frac{\partial}{\partial x_j} [\tau_{ij}] \quad (2.0.6)$$

$$\frac{\partial e}{\partial t} + \frac{\partial}{\partial x_j} \left[ \frac{m_j}{\rho} (e + p) \right] = \frac{\partial}{\partial x_j} \left[ \tau_{ij} \frac{m_i}{\rho} - q_i \right] \quad (2.0.7)$$

Równania konstytutywne dla gazu doskonałego mają następującą postać:

- Równanie stanu:

$$\frac{p}{\rho \theta} = R \quad (2.0.8)$$

gdzie  $\theta$  jest temperaturą,  $R$  stałą gazową. Energia wewnętrzna na jednostkę objętości,  $\iota = e - (m_1^2 + m_2^2)/(2\rho)$ , ze względu na stałość ciepła właściwego przy stałej objętości  $C_v$  jest proporcjonalna do temperatury i gęstości:

$$\iota = C_v \rho \theta \quad (2.0.9)$$

Biorąc pod uwagę związek ciepła właściwych przy stałej objętości  $C_v$  i stałym ciśnieniu  $C_p$ ,  $C_p - C_v = R$ , oraz przyjmując oznaczenie  $\gamma = C_p/C_v$ , na podstawie (2.0.8) i (2.0.9) znajdujemy, że

$$p = (\gamma - 1)\iota \quad (2.0.10)$$

Dla powietrza  $\gamma = 1.4$ .

- Zależność naprężeń lepkich od prędkości odkształceń:

$$\tau_{ij} = 2\mu \varepsilon_{ij} + \lambda \varepsilon_{kk}, \quad (2.0.11)$$

gdzie  $\varepsilon_{ij} = 1/2(u_{i,j} + u_{j,i})$  jest tensorem prędkości odkształceń,  $\mu$  oznacza dynamiczny współczynnik lepkości,  $\lambda = -2/3\mu$ . Lepkość gazu jest funkcją temperatury, określa ją prawo Sutherlanda

$$\mu = \mu_0 \left( \frac{\theta}{\theta_0} \right)^{1.5} \frac{c_1 + \theta_0}{c_1 + \theta}, \quad (2.0.12)$$

gdzie  $\mu_0$  jest współczynnikiem lepkości w temperaturze odniesienia  $\theta_0$ ,  $c_1 = 110^0 K$ .

- Prawo przewodnictwa cieplnego (prawo Fouriera):

$$q_i = -\kappa \theta_{,i}, \quad (2.0.13)$$

gdzie  $\kappa$  jest współczynnikiem przewodnictwa cieplnego zależnym od temperatury podobnie jak  $\mu$ , tak iż

$$\kappa = \frac{C_p}{Pr} \mu, \quad (2.0.14)$$

gdzie  $Pr$  jest liczbą Prandtla, dla powietrza  $Pr = 0.72$ .

Układ równań (2.0.5), (2.0.6) i (2.0.7) wraz z powyższymi związkami fizycznymi stanowi równania Naviera-Stokesa dla przepływów sekwencyjnych.



## 2.1 Bezwymiarowa postać równań Naviera–Stokesa

Posługiwanie się bezwymiarową postacią sformułowania danego problemu fizycznego ma na celu znalezienie parametrów podobieństwa zjawiska, tj. parametrów, od których faktycznie zależy rozwiązanie, jeśli pominąć stałe czynniki skalujące. Dodatkową korzystną cechą sformułowań bezwymiarowych jest łatwość porównywania wyników obliczeń czy pomiarów doświadczalnych. Poniżej cytujemy za [7] najczęściej stosowaną wersję ubezwymiarowania równań Naviera–Stokesa.

Wprowadzamy następujące wielkości odniesienia: dla długości -  $L$ , prędkości -  $V_\infty$ , gęstości -  $\rho_\infty$ , temperatury -  $\theta_\infty$ , lepkości -  $\mu_\infty$ . Wielkości te są w zasadzie dowolne, jednak zwyczajowo najczęściej przyjmuje się  $L$  jako charakterystyczny wymiar geometrii problemu,  $V_\infty$ ,  $\rho_\infty$ ,  $\theta_\infty$  i  $\mu_\infty$  jako prędkość, gęstość, temperaturę i lepkość niezaburzonego przepływu (tj. charakterystyki przepływu z dala od opływanych obiektów).

Jako nowe bezwymiarowe zmienne niezależne przyjmujemy

$$x_i^* := \frac{x_i}{L} \quad \text{i} \quad t^* := \frac{V_\infty}{L} t \quad (2.1.15)$$

Nowe bezwymiarowe zmienne zależne są zaś następujące

$$\varrho^* := \frac{\varrho}{\rho_\infty}, \quad m_i^* := \frac{m_i}{\rho_\infty V_\infty}, \quad i = 1, 2, \quad e^* := \frac{e}{\rho_\infty V_\infty^2}. \quad (2.1.16)$$

Dodatkowo definiujemy bezwymiarowe odpowiedniki wielkości będących funkcjami rozwiązania

$$\theta^* := \frac{\theta}{\theta_\infty}, \quad \iota^* := \frac{\iota}{\rho_\infty V_\infty^2}, \quad u_i^* := \frac{u_i}{V_\infty}, \quad i = 1, 2, \quad p^* := \frac{p}{\rho_\infty V_\infty^2} \quad (2.1.17)$$

oraz

$$\tau_{ij}^* := \frac{\tau_{ij}}{\rho_\infty V_\infty^2}, \quad i, j = 1, 2, \quad q_i^* := \frac{q_i}{\rho_\infty V_\infty^3}, \quad i = 1, 2, \quad \mu^* := \frac{\mu}{\mu_\infty}, \quad \lambda^* := \frac{\lambda}{\mu_\infty} \quad (2.1.18)$$

Zastosowanie podstawień (2.1.15-2.1.18) w równaniach różniczkowych (2.0.5)-(2.0.7) prowadzi do identycznej postaci tych równań z formalną zamianą  $t \rightarrow t^*$ ,  $x_i \rightarrow x_i^*$ ,  $\varrho \rightarrow \varrho^*$ ,  $m_i \rightarrow m_i^*$ ,  $p \rightarrow p^*$ ,  $\tau_{ij} \rightarrow \tau_{ij}^*$  i  $q_i \rightarrow q_i^*$ . Dlatego też zwyczajowo pomija się tę zmianę oznaczeń zakładając domniemane użycie zmiennych bezwymiarowych. Nieco inaczej przedstawia się sprawa związków fizycznych (2.0.9)-(2.0.14). Zamiana zmiennych (2.1.15)-(2.1.18) sprawia, że związki te zachowują swą postać przy odpowiednim przeskalowaniu występujących w nich stałych fizycznych. Skalowanie to wygodnie jest scharakteryzować za pomocą dwu bezwymiarowych parametrów zależnych od  $L$ ,  $V_\infty$ ,  $\rho_\infty$ ,  $\theta_\infty$  i  $\mu_\infty$ :

liczby Macha

$$M_\infty := \frac{V_\infty}{c_\infty} = \frac{V_\infty}{\sqrt{\gamma(\gamma-1)C_v\theta_\infty}} \quad (2.1.19)$$

(gdzie  $c_\infty = \sqrt{\gamma p_\infty / \rho_\infty} = \sqrt{\gamma(\gamma-1)C_v\theta_\infty}$  jest prędkością dźwięku w temperaturze  $\theta_\infty$ )  
oraz

liczby Reynoldsa

$$\text{Re} := \frac{\rho_\infty V_\infty L}{\mu_\infty} \quad (2.1.20)$$

Łatwy rachunek pokazuje, że równania (2.0.9)-(2.0.13) w nowych zmiennych przyjmują postać:

$$\iota^* = \varrho^* \bar{C}_v \theta^* \quad (2.1.21)$$

$$p^* = (\gamma - 1) \iota^* \quad (2.1.22)$$

$$\tau_{ij}^* = 2\bar{\mu}\varepsilon_{ij}^* + \bar{\lambda}\delta_{ij}\varepsilon_{kk}^* \quad (\text{przy } \varepsilon_{ij}^* := \frac{1}{2} \left( \frac{\partial u_i^*}{\partial x_j^*} + \frac{\partial u_j^*}{\partial x_i^*} \right)) \quad (2.1.23)$$

$$q_i^* = -\bar{\kappa}^* \frac{\partial \theta^*}{\partial x_i^*} \quad (2.1.24)$$

jeśli przyjmiemy następujące definicje stałych fizycznych

$$\bar{C}_v := \frac{1}{\gamma(\gamma-1)M_\infty^2}, \quad \bar{\mu} := \frac{\mu^*}{\text{Re}}, \quad \bar{\kappa} := \frac{\gamma \bar{C}_v \bar{\mu}}{\text{Pr}} \quad (2.1.25)$$

Podsumowując, bezwymiarowa postać równań Naviera-Stokesa jest formalnie identyczna z postacią używającą standardowych jednostek, o ile stałe fizyczne  $\bar{C}_v$ ,  $\bar{\mu}$  i  $\bar{\kappa}$  są zdefiniowane jak w (2.1.25). Dlatego też tradycyjnie pomija się przedstawioną zmianę oznaczeń przyjmując po prostu definicje (2.1.25).

Formalizm równań Naviera-Stokesa zyskuje na przejrzystości po zastosowaniu zapisu wektorowo-macierzowego. Prawa zachowania (2.0.5)-(2.0.7) można przedstawić w tym ujęciu następująco:

$$U_{,t} + F_{i,i} = F_i^v, \quad (2.1.26)$$

gdzie

$$U = \{\varrho, m_1, m_2, e\}^T \quad (2.1.27)$$

jest wektorem zmiennych konserwatywnych opisujących stan gazu,

$$F_i = \{m_i, m_1 m_i / \varrho + \delta_{1i} p, m_2 m_i / \varrho + \delta_{2i} p, (e + p) m_i / \varrho\}^T, \quad i = 1, 2 \quad (2.1.28)$$

są strumieniami eulerowskimi,

$$F_i^v = \{0, \tau_{1i}, \tau_{2i}, \tau_{ij} u_j - q_i\}^T, \quad i = 1, 2 \quad (2.1.29)$$

są strumieniami lepkości. W (2.1.26) a także w dalszej części pracy  $(\cdot)_{,i}$  oznacza różniczkowanie względem współrzędnych kartezjańskich,  $(\cdot)_{,i} = \partial/\partial x_i$ , zaś  $(\cdot)_{,t}$  różniczkowanie względem czasu,  $(\cdot)_{,t} = \partial/\partial t$ . Zastosowanie ma także umowa sumacyjna Einsteina.

Inną często używaną formą zapisu równań (2.0.5)-(2.0.7) jest postać quasi-liniowa

$$U_{,t} + A_i U_{,i} = (K_{ij}(U) U_{,j})_{,i} \quad (2.1.30)$$

gdzie  $A_i := \partial F_i / \partial U$  oznaczają jacobiany eulerowskie, zaś prawa strona (2.1.30) wyraża fakt, iż  $F_i^v$  jest liniową funkcją  $\nabla U$ ,  $F_i^v = K_{ij}(U) U_{,j}$ . Postać macierzy  $A_i$  i  $K_{ij}$  podano w dodatku A.2.

Dynamikę gazu nielepkiego opisuje szczególnie przypadek powyższych równań, równania Eulera

$$U_{,t} + A_i U_{,i} = 0. \quad (2.1.31)$$

## 2.2 Symetryzacja równań Naviera–Stokesa

Macierze  $\mathbf{A}_i$  oraz  $\mathbf{K}_{ij}$ ,  $i, j = 1, 2$  można sprowadzić do postaci symetrycznej za pomocą macierzy  $\mathbf{A}_0$  zdefiniowanej jako

$$\mathbf{A}_0 = \left( \frac{\partial^2 H}{\partial U^2} \right)^{-1} \quad (2.2.32)$$

gdzie  $H$  jest matematyczną funkcją entropii,  $H = -\rho \ln(p\rho^{-\gamma})$ , pokrywającą się co do znaku z termodynamiczną entropią gazu doskonałego na jednostkę objętości (jawną postać  $\mathbf{A}_0$  podano w dodatku A.2). Macierz  $\mathbf{A}_0$  jest symetryzatorem równań Navier–Stokesa w następującym znaczeniu

1.  $\mathbf{A}_0^T = \mathbf{A}_0 > 0$
2.  $(\mathbf{A}_i \mathbf{A}_0)^T = \mathbf{A}_i \mathbf{A}_0$ ,  $i = 1, 2$
3.  $(\mathbf{K}_{ij} \mathbf{A}_0)^T = \mathbf{K}_{ji} \mathbf{A}_0$ ,  $i, j = 1, 2$
4.  $U_{,i} \mathbf{K}_{ij} \mathbf{A}_0 U_{,j} \geq 0$ ,  $\forall U_{,i}$

Powyższe własności  $\mathbf{A}_0$  pozwalają otrzymać symetryczne sformułowanie równań Naviera–Stokesa za pomocą nieliniowej zamiany zmiennych

$$\mathbf{V}^T = \frac{\partial H(\mathbf{U})}{\partial \mathbf{U}} \quad (2.2.34)$$

w postaci

$$\mathbf{A}_0 \mathbf{V}_{,t} + \tilde{\mathbf{A}}_i \mathbf{V}_{,i} = (\tilde{\mathbf{K}}_{ij} \mathbf{V}_{,j})_{,i} \quad (2.2.35)$$

gdzie

$$\tilde{\mathbf{A}}_i = \mathbf{A}_i \mathbf{A}_0 \quad \text{i} \quad \tilde{\mathbf{K}}_{ij} = \mathbf{K}_{ij} \mathbf{A}_0 \quad (2.2.36)$$

są operatorami symetrycznymi zgodnie z (2.2.33)<sub>2,3</sub>. Zamiana zmiennych (2.2.34) implikuje bowiem, że

$$\frac{\partial \mathbf{U}}{\partial \mathbf{V}} = \left( \frac{\partial^2 H}{\partial U^2} \right)^{-1} = \mathbf{A}_0 \quad (2.2.37)$$

co pozwala wyrazić odpowiednie pochodne funkcji  $\mathbf{U}$  w (2.1.30) przez pochodne  $\mathbf{V}$  jak w (2.2.35). Parametry  $\mathbf{V}$  nazywają się zmiennymi entropowymi (*entropy variables*).

## 2.3 Fale uderzeniowe

Równania Eulera (2.1.31) opisujące ruch gazu nielepkiego dopuszczają istnienie rozwiązań nieciągłych. Sugeruje to nieobecność w nich drugich pochodnych, zaś formalnie możliwość nieciągłości zauważyć można przy przejściu od postaci całkowej praw zachowania do postaci lokalnej. Najprostszym rozwiązaniem nieciągłym jest prostopadła nieruchoma fala uderzeniowa. Jest to przepływ jednowymiarowy, dla którego parametry gazu zmieniają się skokowo przy przejściu przez płaszczyznę prostopadłą do kierunku przepływu, zaś w obu półprzestrzeniach oddzielonych tą płaszczyzną rozwiązania są stałe. Z analizy praw zachowania w obszarze płaszczyzny nieciągłości można otrzymać związek między parametrami gazu po obu stronach. Jest to <http://rcin.org.pl>  $\rho_0 u_0 = \rho_1 u_1$ ,  $\rho_0 u_0^2 + p_0 = \rho_1 u_1^2 + p_1$ ,  $\rho_0 u_0^3 + p_0 u_0 = \rho_1 u_1^3 + p_1 u_1$ , gdzie  $\rho$  to gęstość,  $u$  to prędkość,  $p$  to ciśnienie i

temperaturę po stronie napływającego gazu, zaś  $\varrho_2, u_2, p_2, \theta_2$  analogiczne parametry po stronie przeciwnej, to

$$\frac{\varrho_2}{\varrho_1} = \frac{u_1}{u_2} = \frac{(\gamma + 1)M_1^2}{(\gamma - 1)M_1^2 + 2} \quad \frac{p_1}{p_2} = \frac{2\gamma M_1^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1} \quad (2.3.38)$$

$$\frac{\theta_2}{\theta_1} = \frac{2(\gamma - 1)(1 + \frac{\gamma-1}{2}M_1^2)(\frac{2\gamma}{\gamma-1}M_1^2 - 1)}{(\gamma + 1)^2 M_1^2} \quad (2.3.39)$$

gdzie  $M_1 = u_1/\sqrt{\gamma p_1/\varrho_1}$  jest liczbą Macha po stronie napływającego gazu, przy czym spełnienie II zasady termodynamiki – zapewnienie dodatniej produkcji entropii – wymaga dodatkowo, aby  $M_1 > 1$ . Wówczas po stronie przeciwnej

$$M_2^2 = \frac{M_1^2 + \frac{2}{\gamma-1}}{\frac{2\gamma}{\gamma-1}M_1^2 - 1} < 1 \quad (2.3.40)$$

Przedstawione zależności oznaczają następujące relacje parametrów przed i za falą:  $M_1 > 1 > M_2$ ,  $p_1 < p_2$ ,  $\theta_1 < \theta_2$ ,  $\varrho_1 < \varrho_2$ .

Stacjonarne ukośne fale uderzeniowe można rozpatrywać w niewielkich obszarach (lokalnie) jako złożenie prostopadłej nieruchomej fali uderzeniowej z ruchem obserwatora równoległym do płaszczyzny fali.

Fale uderzeniowe dla przepływów rzeczywistych z lepkością nie są nieciągłe, lecz charakteryzują się gwałtowną zmianą parametrów przepływu, która zachodzi na grubości rzędu kilku średnich dróg swobodnych cząstek gazu. Przykładowo dla powietrza w temperaturze  $0^\circ\text{C}$  pod ciśnieniem 1 atm średnia droga swobodna wynosi ok.  $2 \cdot 10^{-5}$  cm. Grubość fal uderzeniowych jest więc na ogół nieporównywalnie mała w stosunku do skali opływanych obiektów.

## 2.4 Warstwy przyścienne

W przypadku modelu gazu nielepkiego, opisywanego równaniami Eulera, nie występują naprężenia styczne, tensor naprężeń jest tensorem kulistym. Nie występuje też przewodnictwo ciepła. Sprawia to, iż z fizycznego punktu widzenia nie ma możliwości zapobieżenia poślizgowi gazu na powierzchni opływającego obiektu ani „ogrzania” gazu do temperatury tej powierzchni, czyli narzucenia takich właśnie warunków brzegowych. Przepływ nielepki w pobliżu opływanej przeszkody spełnia jedynie warunek zerowania się prędkości prostopadłej do ciała, co nie powoduje jednak szczególnej nieregularności rozwiązania bezpośrednio przy przeszkodzie. Lepkość i przewodnictwo ciepłe gazu rzeczywistego sprawiają, że zarówno przejście cząstek gazu od przylegania do ciała stałego do ruchu we wnętrzu obszaru jak i zmiana temperatury od panującej na powierzchni do tej we wnętrzu musi następować na pewnej grubości tym mniejszej im mniejsze są stałe lepkości i przewodnictwa.

Przedstawione argumenty fizyczne znajdują odbicie w formie rzeczywistych przepływów z małą lepkością (tj. z dużą liczbą Reynoldsa). Charakteryzują się one występowaniem cienkich warstw przyściennych wzdłuż powierzchni opływającego obiektu, w których prędkość i temperatura zmieniają się bardzo gwałtownie do wartości charakteryzujących przepływ z dala od przeszkody. Dla wielu rodzajów przepływów z małą lepkością sprawdza się hipoteza, iż lepkość wpływa wyłącznie na stan gazu w obszarze warstw przyściennych

natomiast przepływ w pozostałej części obszaru jest praktycznie nielepki. Przyjmując tę hipotezę można otrzymać istotne uproszczenie równań gazodynamiki dla przepływów z małą lepkością, tzw. równania warstwy przyściennej. Ich praktyczne zastosowanie jest wprawdzie ograniczone, przytaczamy je jednak dla pełni prezentacji za monografią J. D. Andersona Jr [6]. Rozważane są rozwiązania stacjonarne, zmiennymi niezależnymi są położenie wzdłuż brzegu ciała  $x$  i odległość od brzegu  $y$ . Stan gazu opisują gęstość  $\rho$ , prędkości  $u$  i  $v$  (w kierunkach  $x$  i  $y$ ) oraz temperatura  $\theta$  i ciśnienie  $p = (\gamma - 1)C_v \rho \theta$ . Równania mają postać:

$$\begin{aligned} \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} &= 0 \\ \rho u \frac{\partial u}{\partial x} + \rho v \frac{\partial v}{\partial y} &= -\frac{\partial p_e}{\partial x} + \frac{\partial}{\partial y} \left( \mu \frac{\partial u}{\partial y} \right) \\ \frac{\partial p}{\partial y} &= 0 \\ C_p \rho u \frac{\partial \theta}{\partial x} + C_p \rho v \frac{\partial \theta}{\partial y} &= \frac{\partial}{\partial y} \left( \kappa \frac{\partial \theta}{\partial y} \right) + u \frac{\partial p_e}{\partial x} + \mu \left( \frac{\partial u}{\partial y} \right)^2 \end{aligned} \tag{2.4.41}$$

gdzie  $p_e = p_e(x)$  jest z założenia znanym ciśnieniem na zewnętrznej krawędzi warstwy przyściennej. Podstawowe wnioski wynikające z (2.4.41) to stałość ciśnienia wzdłuż grubości warstwy przyściennej oraz związana z tym zależność temperatury i gęstości  $\rho \theta = p/[(\gamma - 1)C_v] = c$ . Korzystając z równań warstwy przyściennej otrzymano szereg analitycznych rozwiązań, z których najbardziej znane to przepływ nad płaską płytą i opływ ciała zatępionego w pobliżu punktu zastoju. Jednym z wniosków z analizy rzędów wielkości dokonywanej przy wyprowadzeniu równań warstwy przyściennej jest oszacowanie jej grubości  $\delta$  w postaci  $\delta \sim 1/\sqrt{\text{Re}}$ , gdzie  $\text{Re}$  jest liczbą Reynoldsa.

## Rozdział 3

# Dyskretyzacja skończenie elementowa równań przepływów metodą SUPG

### 3.1 Metoda SUPG

Zastosowanie standardowej metody Galerkin do konstrukcji sformułowania słabego równań Naviera–Stokesa prowadzi do niepowodzenia. Metoda dyskretna z elementami skończonymi użytymi do aproksymacji przestrzennej w takim sformułowaniu słabym charakteryzuje się niedostatecznym tłumieniem numerycznym, co prowadzi do rozwiązań oscylacyjnych, szczególnie w pobliżu nieciągłości rozwiązania ścisłego. Większość algorytmów aproksymacji skończenie elementowej dla przepływów stosuje sztuczną lepkość (dyssypację), której zadaniem jest tłumienie niefizycznych oscylacji w rozwiązaniu numerycznym. Dodanie sztucznej lepkości modyfikuje rozwiązywany problem. Dlatego też jest szczególnie istotne, aby sztuczna lepkość była wystarczająca do tłumienia pasożytniczych oscylacji, lecz jednocześnie aby była na tyle mała, aby nie zaburzyć istotnie rozwiązania. Tę bardzo delikatną równowagę dostrojenia mechanizmu sztucznej dyssypacji zapewnia z dużym powodzeniem metoda SUPG zaproponowana i propagowana przez T.J.R. Hughesa [51]–[57]. SUPG jest skrótem od *Streamline Upwind Petrov/Galerkin* i w wolnym tłumaczeniu oznacza metodę elementów skończonych z funkcjami testowymi zniekształconymi w kierunku linii prądu. Podstawą tej metody jest bowiem koncepcja zastosowania niesymetrycznych funkcji testowych o większych wartościach w kierunku, z którego transportowana jest informacja: z kierunku „pod prąd”. Oryginalna wersja metody SUPG powstała w kontekście równań Naviera–Stokesa ze zmiennymi entropowymi opisującymi stan gazu,  $\mathbf{V} = \{V_1, \dots, V_4\}^T$ , związanymi ze zmiennymi konserwatywnymi  $\mathbf{U} = \{\rho, m_1, m_2, e\}^T$  wzorem (2.2.34). Jednak w ostatnim czasie powstała także wersja SUPG dla zmiennych konserwatywnych [17],[4]. Jest ona często preferowana ze względu na natychmiastową interpretację fizyczną zmiennych  $\rho, m_i, e$  w odróżnieniu od wielkości  $V_i$ , mniejszy koszt obliczeń oraz ze względu na podobieństwa metody do wielu innych algorytmów.

Podstawą konstrukcji metody jest zdefiniowanie przestrzeni skończenie elementowej. Zakładamy, że obszar obliczeniowy  $\Omega$  jest pokryty siatką elementów czworokątnych  $K$  z biliniowymi funkcjami kształtu. Tworzą one standardową przestrzeń skończenie elementową  $V^h$ :

$$V^h := \{v : \Omega \rightarrow R^4 \mid v \in C(\bar{\Omega}) \text{ i } v|_K \in (Q_1(K))^4, \forall K\} \quad (3.1.1)$$

gdzie  $Q_1(K)$  oznacza biliniowe funkcje kształtu na elemencie  $K$ . O funkcjach  $v \in V^h$  zakładamy dodatkowo, że spełniają jednorodne warunki brzegowe typu Dirichleta na części brzegu  $\partial\Omega_u$ . Metodę SUPG definiuje następujące sformułowanie słabe:

Znaleźć  $U \in V^h + \mathbf{u}_0$ , takie że

$$\int_{\Omega} \mathbf{W}^T (U_{,t} + \mathbf{A}_i U_{,i}) dx + \int_{\Omega} \mathbf{W}_{,i}^T \mathbf{K}_{ij} U_{,j} dx + \sum_K \int_K \nu_d \mathbf{W}_{,i}^T U_{,i} dx + \sum_K \int_K \tau_a \mathbf{W}_{,k}^T \mathbf{A}_{,k} [U_{,t} + \mathbf{A}_i U_{,i} - (\mathbf{K}_{ij} U_{,j})_{,i}] dx = \int_{\partial\Omega_t} \mathbf{W}^T \mathbf{H} dS, \quad \forall \mathbf{W} \in V^h \quad (3.1.2)$$

gdzie  $\mathbf{u}_0$  i  $\mathbf{H}$  są danyimi odpowiadającymi niejednorodnym warunkom brzegowym typu Dirichleta i typu Neumanna, z których ostatnie związane są z częścią brzegu  $\partial\Omega_t$ . W (3.1.2)  $\mathbf{A}_i$ ,  $i = 1, 2$  są jacobianami eulerowskimi,  $\mathbf{A}_i = \partial \mathbf{F}_i / \partial \mathbf{U}$ ,  $\mathbf{K}_{ij}$  są macierzami lepkości. W sformułowaniu (3.1.2) pierwsze dwa składniki po lewej stronie oraz wyraz z prawej strony odpowiadają standardowemu sformułowaniu słabemu Galerkin dla równań Naviera-Stokesa (2.1.26). Pozostałe dwa składniki to dodatkowe wyrazy stabilizujące metody SUPG. Pierwszy z nich,  $\sum_K \int_K \nu_d \mathbf{W}_{,i}^T U_{,i} dx$  nosi w oryginale nazwę *shock capturing operator* i ma na celu stabilizowanie rozwiązania wzdłuż nieciągłości fali uderzeniowej. Drugi z wyrazów można interpretować jako residuum równań Naviera-Stokesa (wyrażenie w nawiasie kwadratowym) testowane zmodyfikowaną funkcją testową  $\tau_a \mathbf{A}_{,k}^T \mathbf{W}_{,k}$ . Współczynniki  $\nu_d$  i  $\tau_a$  regulują wielkość wyrazów stabilizujących. Zależą one od aktualnego rozwiązania  $U$  oraz od siatki skończenie elementowej. Można je przedstawić następująco:

$$\tau_a = \frac{\max_{i=1,2} (h_i |\beta_i|)}{2(c + |\mathbf{u} \cdot \boldsymbol{\beta}|)}, \quad \nu_d = \left( \frac{(\mathbf{A}_i U_{,i})^T \mathbf{A}_0^{-1} \mathbf{A}_j U_{,j}}{(\boldsymbol{\xi}_t, U_{,i})^T \mathbf{A}_0^{-1} \boldsymbol{\xi}_{t,j} U_{,j}} \right)^{1/2} \quad (3.1.3)$$

gdzie  $\mathbf{A}_0$  oznacza macierz symetryzującą równań Naviera-Stokesa zdefiniowaną w (2.2.32).  $\boldsymbol{\beta}$  w (3.1.3) jest znormalizowanym wektorem równoległym do

$$\nabla \|U\|_{\mathbf{A}_0^{-1}} \quad \text{przy} \quad \|U\|_{\mathbf{A}_0^{-1}} := U^T \mathbf{A}_0^{-1} U,$$

$c$  jest prędkością dźwięku,  $c = \sqrt{\gamma p / \rho}$ , parametry  $h_i$ ,  $i = 1, 2$  są wymiarami elementu,  $h_i = (\sum_{j=1}^2 (\partial x_i / \partial \xi_j)^2)^{1/2}$ . Funkcje  $x_i(\xi_j)$ ,  $i = 1, 2$  reprezentują odwzorowanie elementu wzorcowego  $\hat{K} = [-1, 1]^2$  w element rzeczywisty  $K$ ,  $\xi_i(x_j)$  odpowiadają odwzorowaniu odwrotnemu.

Następnym krokiem jest zdefiniowanie dyskretyzacji w czasie. W tym celu uprościmy najpierw drugi z wyrazów stabilizujących w (3.1.2) poprzez pominięcie  $U_{,t}$ . Modyfikacja ta nie zmienia rozwiązań stacjonarnych, które są naszym celem, ułatwia natomiast iteracyjne rozwiązywanie dyskretnego układu równań dzięki uniknięciu istotnej niesymetrii, jaką powoduje uwzględnienie tego składnika (wpływ niesymetrii na iteracyjne rozwiązywanie równań omówiono w rozdziale 6.11). Otrzymane sformułowanie dyskretyzujemy w czasie przez zastosowanie wersji schematu Eulera typu „semi-implicit”, w której  $U_{,t} = (U_{n-1} - U_n) / \Delta t$ , zaś podział pozostałych wyrazów sformułowania na jawne i niejawne jest taki, że prowadzi do następującego liniowego problemu brzegowego:

Znaleźć takie  $U^{n+1} \in V^h + \mathbf{u}_0$ , że

$$\int_{\Omega} \mathbf{W}^T U^{n-1} dx + \Delta t \int_{\Omega} \mathbf{W}_{,i}^T (\mathbf{K}_{ij} + \tau_a \mathbf{A}_{,i} \mathbf{A}_{,j} + \nu_d \delta_{ij}) U_{,i}^{n-1} dx = \quad (3.1.4)$$

$$\int_{\Omega} \mathbf{W}^T \mathbf{U}^n dx - \Delta t \int_{\Omega} \mathbf{W}^T \mathbf{A}_i \mathbf{U}_{,i}^n dx + \Delta t \sum_K \int_K \tau_a \mathbf{W}_{,k}^T \mathbf{A}_k (\mathbf{K}_{ij} \mathbf{U}_{,j}^n)_{,i} dx + \Delta t \int_{\partial\Omega_i} \mathbf{W}^T \mathbf{H} dS \quad \forall \mathbf{W} \in V^h$$

gdzie współczynniki  $\tau_a$  i  $\nu_d$  i macierze  $\mathbf{K}_{ij}$  i  $\mathbf{A}_i$ ,  $i, j = 1, 2$ , są funkcjami  $\mathbf{U}^n$ . Przedstawiona konstrukcja sformułowania jest motywowana tym, że schemat z eliptyczną konytrybucją w postaci drugiej całki po lewej stronie (3.1.4) i z jawnym traktowaniem wyrazu  $\Delta t \int_{\Omega} \mathbf{W}^T \mathbf{A}_i \mathbf{U}_{,i}^n dx$  jest warunkowo stabilny w sensie stabilności liniowej, niezależnie od aproksymacji przestrzennej (co omawiamy dalej).

Na koniec zauważmy, że problem (3.1.4) można zapisać w standardowej postaci abstrakcyjnego sformułowania słabego:

Znaleźć takie  $\mathbf{U}^{n-1} \in V^h + \mathbf{u}_0$ , że

$$B(\mathbf{U}^{n-1}, \mathbf{W}) = L(\mathbf{W}) \quad \forall \mathbf{W} \in V^h \quad (3.1.5)$$

gdzie

$$B(\mathbf{U}, \mathbf{W}) := \int_{\Omega} \mathbf{W}^T \mathbf{U} dx + \Delta t \int_{\Omega} \mathbf{W}_{,i}^T (\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}) \mathbf{U}_{,i} dx \quad (3.1.6)$$

$$L(\mathbf{W}) := \int_{\Omega} \mathbf{W}^T \mathbf{U}^n dx - \Delta t \int_{\Omega} \mathbf{W}^T \mathbf{A}_i \mathbf{U}_{,i}^n dx + \Delta t \sum_K \int_K \tau_a \mathbf{W}_{,k}^T \mathbf{A}_k (\mathbf{K}_{ij} \mathbf{U}_{,j}^n)_{,i} dx + \Delta t \int_{\partial\Omega_i} \mathbf{W}^T \mathbf{H} dS$$

są formą dwuliniową i funkcjonalem liniowym.

## 3.2 Zbieżność i stabilność metody SUPG

Metoda SUPG ze zmiennymi konserwatywnymi została zaproponowana w artykułach [17],[4] i [46]. Autorzy nie podają w nich jednak teoretycznych aspektów metody (choć wspominają o ich opracowaniu w materiałach nie opublikowanych). Dlatego też prezentację tytułowych zagadnień oprzemy na innych źródłach, które opisują je dla pokrewnych metod. Przedstawimy wynik teoretyczny opisujący zbieżność symetrycznej wersji SUPG wg. pracy [61] C. Johnsona, A. Szepessy i P. Hansbo. Stabilność schematu sprawdzimy zaś opierając się na wynikach pracy [30] poświęconej metodzie Taylora–Galerkina, którą można uważać za pewną modyfikację SUPG w zmiennych konserwatywnych.

**Zbieżność** (wg [61])

Rozważmy zadanie początkowe z układem praw zachowania (w szczególności z równaniami Eulera) w całej przestrzeni w postaci

$$\begin{aligned} \mathbf{U}_{,t} + \mathbf{F}_i(\mathbf{U})_{,i} &= 0, \quad t > 0, \quad \mathbf{x} \in \mathbb{R}^d, \quad d \geq 2 \\ \mathbf{U}(0, \mathbf{x}) &= \mathbf{U}_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d \end{aligned} \quad (3.2.7)$$

gdzie  $\mathbf{U} \in \mathbb{R}^m$ ,  $m \geq 2$ . Zakładamy, że dla tego problemu istnieje ściśle wypukła funkcja entropii  $\eta(\mathbf{U})$  (dla równań Eulera jest nią z przeciwnym znakiem entropia termodynamiczna na jednostkę objętości, tj.  $H(\mathbf{U})$ ) oraz stowarzyszone z nią strumienie entropii  $q_i$ ,  $i = 1, \dots, d$ , spełniające warunek zgodności

$$\frac{\partial \eta}{\partial \mathbf{U}} \frac{\partial \mathbf{F}_i}{\partial \mathbf{U}} = \frac{\partial q_i}{\partial \mathbf{U}} \quad \forall i = 1, \dots, d \quad (3.2.8)$$



(dla równań Eulera  $q_i = H(\mathbf{U})u_i$ , gdzie  $u_i$  są składowymi prędkościami). Wprowadzamy rozważaną już wcześniej (w kontekście symetryzacji równań Naviera–Stokesa) zamianę zmiennych

$$\mathbf{V}^T = \frac{\partial \eta}{\partial \mathbf{U}}, \quad (3.2.9)$$

prowadzącą do sformułowania

$$\begin{aligned} \mathbf{A}_0 \mathbf{V}_{,t} + \tilde{\mathbf{A}}_i \mathbf{V}_{,i} &= 0, \quad t > 0, \quad \mathbf{x} \in R^d \\ \mathbf{V}(0, \mathbf{x}) &= \mathbf{V}_0(\mathbf{x}) \end{aligned} \quad (3.2.10)$$

gdzie  $\mathbf{A}_0 = \partial \mathbf{U} / \partial \mathbf{V} = (\partial^2 \eta / \partial \mathbf{U}^2)^{-1}$ ,  $\tilde{\mathbf{A}}_i = \mathbf{A}_i \mathbf{A}_0$ . Warunek (3.2.8) implikuje symetrię  $\tilde{\mathbf{A}}_i$ . Dla problemu początkowego (3.2.10) stosujemy następującą aproksymację skończenie elementową:

1. Półprostą czasu,  $t > 0$ , dzielimy na odcinki  $I_n$  między chwilami  $0 = t_0 < t_1 < t_2 < \dots, I_n = (t_n, t_{n+1})$ .
2. Rozważamy „plastry” czasoprzestrzeni  $\Omega = [0, t] \times R^d$ :  $S_n := I_n \times R^d$ . Każdy  $S_n$  pokrywamy siatką quasi-równomierną  $T_h^n$  (por. rozdział 4.) sympleksów  $d + 1$ -wymiarowych (czworościanów dla  $d = 2$ ) o wymiarze  $h$ . Definiujemy przestrzeń skończenie elementową  $V_h^n$  dla  $S_n$  i dla czasoprzestrzeni  $\Omega$ :  $V_h^n$ :

$$\begin{aligned} V_h^n &= \{v \in [H^1(S_n)]^m : v|_K \in P_k(K), \quad K \in T_h^n\} \\ V_h &= \prod_{n \geq 0} V_h^n \end{aligned} \quad (3.2.11)$$

(gdzie  $P_k(K)$  są wielomianami stopnia  $k$ ). Metodę SUPG definiuje następujące sformułowanie słabe:

Znaleźć  $\mathbf{V} \in V_h$ , takie że

$$\begin{aligned} &\int_{S_n} [\mathbf{W} + \delta(\mathbf{A}_0(\mathbf{V})\mathbf{W}_{,t} + \tilde{\mathbf{A}}_i(\mathbf{V})\mathbf{W}_{,i})]^T (\mathbf{A}_0(\mathbf{V})\mathbf{V}_{,t} + \tilde{\mathbf{A}}_j(\mathbf{V})\mathbf{V}_{,j}) dt \, dx + \\ &\delta_1 \int_{S_n} (\nabla \mathbf{W})^T \frac{|\mathbf{A}_0(\mathbf{V})\mathbf{V}_{,t} + \tilde{\mathbf{A}}_i(\mathbf{V})\mathbf{V}_{,i}|}{\varepsilon + |\nabla \mathbf{V}|} \nabla \mathbf{V} dt \, dx + \\ &\delta_2 \int_{S_n} |\tilde{\mathbf{U}}| (\nabla_x \mathbf{W})^T \nabla_x \mathbf{V} dt \, dx + \int_{R^d} (\mathbf{W}_+^n)^T (\mathbf{U}_+^n - \mathbf{U}_-^n) dx = 0 \quad \forall \mathbf{W} \in V_h^n \end{aligned} \quad (3.2.12)$$

gdzie

$$\nabla_x \mathbf{V} = \{\partial \mathbf{V} / \partial x_i\}, \quad \nabla \mathbf{V} = \{\mathbf{V}_{,t}, \nabla_x \mathbf{V}\} \quad (3.2.13)$$

$$\tilde{\mathbf{U}}|_K = \begin{cases} (\mathbf{U}_+^n - \mathbf{U}_-^n)|_{K \cap (R^d \times \{t_n\})}, & \text{jeśli } \text{meas}[K \cap (R^d \times \{t_n\})] > 0 \\ 0 & \text{w przeciwnym wypadku} \end{cases} \quad (3.2.14)$$

$$\mathbf{V}_\pm^n(\mathbf{x}) = \lim_{s \rightarrow 0^\pm} \mathbf{V}(t_n + s, \mathbf{x}) \quad (3.2.15)$$

$$\delta = h \mathbf{A}_0^{-1/2} (\mathbf{I} + \sum_{i=1}^d \tilde{\mathbf{A}}_i^2)^{-1/2} \mathbf{A}_0^{-1/2} \quad \text{przy} \quad \tilde{\mathbf{A}}_i = \mathbf{A}_0^{-1/2} \tilde{\mathbf{A}}_i \mathbf{A}_0^{-1/2} \quad (3.2.16)$$

Zauważmy, że w sformułowaniu (3.2.12) aproksymacja w czasie i w przestrzeni są ze sobą sprzężone. Dla rozwiązań ustalonych (3.2.12) sprowadza się do następującej postaci

$$\int_{R^d} [\mathbf{W} + \delta \tilde{\mathbf{A}}_i(\mathbf{V})\mathbf{E}_{,i}]^T \tilde{\mathbf{A}}_j(\mathbf{V})\mathbf{V}_{,j} dt + \delta_1 \int_{R^d} (\nabla_x \mathbf{W})^T \frac{|\tilde{\mathbf{A}}_i(\mathbf{V})\mathbf{V}_{,i}|}{\varepsilon + |\nabla_x \mathbf{V}|} \nabla_x \mathbf{V} dx = 0 \quad (3.2.17)$$

w którym pierwszy składnik odpowiada stacjonarnej wersji oryginalnego równania (3.2.10), zaś dwa pozostałe są regularyzującymi wyrazami SUPG. Analogiczne równanie jest spełniane przez stacjonarne rozwiązanie niesymetrycznego schematu SUPG (3.1.4) (przy braku lepkości).

Aby przedstawić wynik opisujący zbieżność metody definiujemy pojęcie rozwiązania entropowego (*entropy solution*) układu (3.2.7):

**Definicja 3.2.1** Funkcja  $U \in [L^\infty(\Omega)]^m$ ,  $\Omega = (0, \infty) \times R^d$  jest rozwiązaniem entropowym problemu (3.2.7), jeśli dla wszystkich  $W \in [C_0^\infty(\Omega)]^m$ ,  $\bar{\Omega} = [0, \infty) \times R^d$  jest spełnione

$$\int_{\Omega} (W_t^T U + W_{,i}^T F_i(U)) dt dx + \int_{R^d} W(0, \cdot)^T U_0 dx = 0 \quad (3.2.18)$$

oraz dla wszystkich  $w \in C_0^\infty(\Omega)$ , takich że  $w \geq 0$

$$\int_{\Omega} (\eta w_t + q_i w_{,i}) dt dx \geq 0 \quad (3.2.19)$$

Zakładając wypukłość funkcji entropii, tj. warunek

$\exists \sigma > 0$  i zwarty zbiór rozwiązań  $D \subset R^m$ , taki że  $\forall V, W \in R^m$

$$\eta(V) - \eta(W) - \frac{\partial \eta}{\partial U}(W) \cdot (V - W) \geq \sigma |V - W|^2 \quad (3.2.20)$$

oraz następujące własności macierzy  $\delta$

$$\exists \alpha_1, \alpha_2 > 0: \alpha_1 h \leq \mathbf{x} \cdot \delta \mathbf{x}, \quad |\mathbf{x} \cdot \delta \mathbf{y}| \leq \alpha_2 h, \quad \forall \mathbf{x}, \mathbf{y} \in R^m, \quad |\mathbf{x}| = |\mathbf{y}| = 1 \quad (3.2.21)$$

mamy następujące

**Twierdzenie 3.2.1** Załóżmy, że ciąg rozwiązań  $V_h$  sformułowania (3.2.12), takich że  $Ry(V_h) \subset \eta_U(D)$ , zbiega prawie wszędzie w  $\Omega$  do funkcji  $V$  przy  $h \rightarrow 0$  oraz że mamy  $\|V_h\|_{L^\infty(\Omega)} \leq C$ . Wówczas  $U := (\eta_U)^{-1}(V)$  jest rozwiązaniem entropowym układu (3.2.7).

Najistotniejszy praktyczny wniosek z powyższego wyniku to fakt, że ciąg rozwiązań sformułowania ze sztuczną lepkością, jeśli jest zbieżny, ma granicę spełniającą (3.2.18), co implikuje spełnianie oryginalnych równań bez lepkości.

**Stabilność** (wg [30])

Sprawdzenie stabilności danego schematu całkowania dla nieliniowego problemu ewolucyjnego jest często trudne do osiągnięcia. Z konieczności ograniczamy się wówczas do zbadania stabilności liniowej metody, przez co rozumiemy jej stabilność dla problemu w postaci zlinearyzowanej przez ustalenie jego współczynników. W pracy [30] przedstawiono analizę stabilności liniowej dla metody Taylora–Galerkina, która ma wiele cech wspólnych z metodą SUPG. Poniżej przedstawiamy szkic tej analizy wraz z jej uogólnieniem dla schematu SUPG.

Rozważamy układ równań

$$U_t + A_1 U_{,1} = 0 \quad (3.2.22)$$

w obszarze  $\Omega = [0, 1]^2$  z periodycznymi warunkami brzegowymi

$$U(0, y) = U(1, y), \quad y \in [0, 1], \quad U(x, 0) = U(x, 1), \quad x \in [0, 1] \quad (3.2.23)$$

Zakładamy, że macierze  $\mathbf{A}_i$ ,  $i = 1, 2$  są stałe w  $\Omega$ . Wówczas metoda Taylora–Galerkina sprowadza się do wyznaczania rozwiązania  $\mathbf{U}^{n+1}$  z następującego sformułowania

$$\int_{\Omega} \mathbf{W}^T \mathbf{U}^{n+1} dx + \frac{\Delta t^2}{2} \int_{\Omega} \mathbf{W}_{,i}^T \mathbf{A}_i \mathbf{A}_j \mathbf{U}_{,j}^{n+1} dx = \int_{\Omega} \mathbf{W}^T \mathbf{U}^n dx + \Delta t \int_{\Omega} \mathbf{W}_{,i}^T \mathbf{A}_i \mathbf{U}^n dx \quad (3.2.24)$$

dla wszystkich  $\mathbf{W}$  spełniających periodyczne warunki brzegowe (3.2.23). Powyższe sformułowanie definiuje operator liniowy  $\mathbf{T} : V^h \rightarrow V^h$ ,  $\mathbf{T}\mathbf{U}^n = \mathbf{U}^{n+1}$ . Operator ten nie jest samosprężny, posiada więc na ogół zespolone wartości i wektory własne. Aby móc formalnie je rozpatrywać rozważamy zespolone uogólnienie  $\mathbf{T}$  otrzymane przez wprowadzenie przestrzeni zespolonych funkcji kształtu  $V^h + iV^h$ . Zastępując dodatkowo funkcję testową  $\mathbf{W}$  przez  $\mathbf{A}_0^{-1}\mathbf{W}$ , gdzie  $\mathbf{A}_0$  jest macierzą symetryzującą (2.2.32), stałą w  $\Omega$  podobnie jak i  $\mathbf{A}_i$ , otrzymujemy następującą postać sformułowania (3.2.24):

Znaleźć  $\mathbf{U} \in V^h + iV^h$ , takie że

$$a(\mathbf{U}^{n+1}, \mathbf{W}) + \frac{\Delta t^2}{2} b(\mathbf{U}^{n+1}, \mathbf{W}) = a(\mathbf{U}^n, \mathbf{W}) + \Delta t c(\mathbf{U}^n, \mathbf{W}), \quad (3.2.25)$$

gdzie

$$\begin{aligned} a(\mathbf{U}, \mathbf{W}) &= \int_{\Omega} \mathbf{W}^T \mathbf{A}_0^{-1} \mathbf{U} dx, & b(\mathbf{U}, \mathbf{W}) &= \int_{\Omega} \mathbf{W}_{,i}^T \mathbf{A}_i^T \mathbf{A}_0^{-1} \mathbf{A}_j \mathbf{U}_{,j} dx, \\ c(\mathbf{U}, \mathbf{W}) &= \int_{\Omega} \mathbf{W}_{,i}^T \mathbf{A}_0^{-1} \mathbf{A}_i \mathbf{U} dx \end{aligned} \quad (3.2.26)$$

Formy  $a$  i  $b$  są symetryczne,  $a$  jest dodatnio określona,  $b$  zaś dodatnio półokreślona dla argumentów rzeczywistych (wynika to z własności symetryzatora  $\mathbf{A}_0$ ).

Zakładając, że  $\mathbf{U}$  jest wektorem własnym,  $\lambda$  wartością własną operatora  $\mathbf{T}$ , podstawmy w (3.2.25)

$$\mathbf{U}^n = \mathbf{U}, \quad \mathbf{U}^{n+1} = \lambda \mathbf{U}, \quad \mathbf{W} = \bar{\mathbf{U}} \quad (3.2.27)$$

otrzymując zależność

$$\lambda \left( a + \frac{\Delta t^2}{2} b \right) = a + \Delta t c(\mathbf{U}, \bar{\mathbf{U}}), \quad (3.2.28)$$

gdzie  $a := a(\mathbf{U}, \bar{\mathbf{U}})$ ,  $b := b(\mathbf{U}, \bar{\mathbf{U}})$ . Wielkości  $a$  i  $b$  są rzeczywiste oraz  $a > 0$ ,  $b \geq 0$ . Przyjmując bowiem  $\mathbf{U} = \mathbf{u} + i\mathbf{v}$ ,  $\mathbf{u}, \mathbf{v} \in R^4$  mamy:

$$a(\mathbf{U}, \bar{\mathbf{U}}) = \int_{\Omega} (\mathbf{u}^T - i\mathbf{v}^T) \mathbf{A}_0^{-1} (\mathbf{u} + i\mathbf{v}) dx = \int_{\Omega} (\mathbf{u}^T \mathbf{A}_0^{-1} \mathbf{u} + \mathbf{v}^T \mathbf{A}_0^{-1} \mathbf{v}) dx > 0 \quad (3.2.29)$$

$$b(\mathbf{U}, \bar{\mathbf{U}}) = \int_{\Omega} (\mathbf{u}_{,k}^T - i\mathbf{v}_{,k}^T) \mathbf{A}_k^T \mathbf{A}_0^{-1} \mathbf{A}_j (\mathbf{u}_{,j} + i\mathbf{v}_{,j}) dx = \quad (3.2.30)$$

$$\int_{\Omega} (\mathbf{u}_{,k}^T \mathbf{A}_k^T \mathbf{A}_0^{-1} \mathbf{A}_j \mathbf{u}_{,j} + \mathbf{v}_{,k}^T \mathbf{A}_k^T \mathbf{A}_0^{-1} \mathbf{A}_j \mathbf{v}_{,j}) dx \geq 0$$

Podobnie podstawiając w (3.2.25)

$$\mathbf{U}^n = \bar{\mathbf{U}}, \quad \mathbf{U}^{n+1} = \bar{\lambda} \bar{\mathbf{U}}, \quad \mathbf{W} = \mathbf{U} \quad (3.2.31)$$

otrzymujemy

$$\bar{\lambda} \left( a + \frac{\Delta t^2}{2} b \right) = \bar{\lambda} c(\mathbf{U}, \bar{\mathbf{U}}) \quad (3.2.32)$$

Równania (3.2.28) i (3.2.32) pozwalają wyliczyć część rzeczywistą i urojoną  $\lambda$

$$\begin{aligned} \left(a + \frac{\Delta t^2}{2}b\right) \operatorname{Re} \lambda &= \Delta t \operatorname{Re}(c(\mathbf{U}, \bar{\mathbf{U}})) + a \\ \left(a + \frac{\Delta t^2}{2}b\right) \operatorname{Im} \lambda &= \Delta t \operatorname{Im}(c(\mathbf{U}, \bar{\mathbf{U}})) \end{aligned} \quad (3.2.33)$$

(wykorzystaliśmy fakt, że  $c(\bar{\mathbf{U}}, \mathbf{U}) = \overline{c(\mathbf{U}, \bar{\mathbf{U}})}$ ).

Wielkość  $c(\mathbf{U}, \bar{\mathbf{U}})$  jest urojona, gdyż

$$\begin{aligned} c(\mathbf{U}, \bar{\mathbf{U}}) + c(\bar{\mathbf{U}}, \mathbf{U}) &= \int_{\Omega} \bar{\mathbf{U}}_{,i}^T \mathbf{A}_0^{-1} \mathbf{A}_i \mathbf{U} dx + \int_{\Omega} \mathbf{U}_{,i}^T \mathbf{A}_0^{-1} \mathbf{A}_i \bar{\mathbf{U}} dx = \\ &= \int_{\Omega} (\bar{\mathbf{U}}^T \mathbf{A}_0^{-1} \mathbf{A}_i \mathbf{U})_{,i} dx = \int_{\partial\Omega} \bar{\mathbf{U}}^T \mathbf{A}_0^{-1} \mathbf{A}_i \mathbf{U} n_i ds = 0 \end{aligned} \quad (3.2.34)$$

ze względu na periodyczne warunki brzegowe (3.2.23). Wprowadzając oznaczenie  $c = \operatorname{Im} c(\mathbf{U}, \bar{\mathbf{U}})$  możemy więc wyliczyć wielkość  $|\lambda|$  następująco

$$|\lambda|^2 = \frac{a^2 + \Delta t^2 c^2}{(a + \Delta t^2/2 b)^2} \quad (3.2.35)$$

Parametr  $c^2$  da się oszacować przez  $ab$ , mamy bowiem

$$\begin{aligned} \operatorname{Im} c(\mathbf{U}, \bar{\mathbf{U}}) &= \frac{1}{i} \int_{\Omega} \bar{\mathbf{U}}_{,k}^T \mathbf{A}_0^{-1} \mathbf{A}_k \mathbf{U} dx = \frac{1}{i} \int_{\Omega} (\mathbf{u}_{,k}^T - i \mathbf{v}_{,k}^T) \mathbf{A}_0^{-1} \mathbf{A}_k (\mathbf{u} + i \mathbf{v}) dx = \\ &= - \int_{\Omega} (\mathbf{A}_k \mathbf{v}_{,k})^T \mathbf{A}_0^{-1} \mathbf{u} dx + \int_{\Omega} (\mathbf{A}_k \mathbf{u}_{,k})^T \mathbf{A}_0^{-1} \mathbf{v} dx \end{aligned} \quad (3.2.36)$$

Stosując nierówność Cauchy'ego-Schwarza dla iloczynu  $\int_{\Omega} (\cdot) \mathbf{A}_0^{-1} (\cdot) dx$  otrzymujemy

$$\begin{aligned} |c| &\leq \left( \int_{\Omega} (\mathbf{A}_k \mathbf{v}_{,k})^T \mathbf{A}_0^{-1} (\mathbf{A}_j \mathbf{v}_{,j}) dx \right)^{1/2} \left( \int_{\Omega} \mathbf{u}^T \mathbf{A}_0^{-1} \mathbf{u} dx \right)^{1/2} + \\ &\quad \left( \int_{\Omega} (\mathbf{A}_k \mathbf{u}_{,k})^T \mathbf{A}_0^{-1} (\mathbf{A}_j \mathbf{u}_{,j}) dx \right)^{1/2} \left( \int_{\Omega} \mathbf{v}^T \mathbf{A}_0^{-1} \mathbf{v} dx \right)^{1/2} \leq \sqrt{ab} \end{aligned} \quad (3.2.37)$$

przy czym ostatni krok wynika z teje nierówności dla iloczynu skalarnego w  $R^2$  oraz z (3.2.29) i (3.2.30).

Stosując powyższą nierówność w (3.2.35) otrzymujemy następujące oszacowanie wartości własnej  $\lambda$

$$|\lambda|^2 \leq \frac{a^2 + \Delta t^2 ab}{a^2 + \Delta t^2 ab + \Delta t^4/4 b^2} \leq 1 \quad (3.2.38)$$

Otrzymany wynik dowodzi bezwarunkowej (tj. dla dowolnego  $\Delta t$ ) stabilności liniowej schematu Taylora-Galerkina. Należy podkreślić, że rezultat ten jest niezależny od zastosowanej aproksymacji skończenie elementowej, tj. od przestrzeni  $V^h$ , stosuje się więc do złożonych nierównomiernych siatek adaptacyjnych.

Uogólnienie powyższego rozumowania dla schematu SUPG (3.1.4) z pominięciem lepkości i „shock capturing term”  $\nu_d$  otrzymujemy stosując w (3.1.4) funkcję testową  $\mathbf{A}_0^{-1} \mathbf{W}$  zamiast  $\mathbf{W}$ . Sformułowanie słabe przybiera postać zbliżoną do (3.2.25)

$$a(\mathbf{U}^{n-1}, \mathbf{W}) + \Delta t b(\mathbf{U}^{n-1}, \mathbf{W}) + \Delta t c(\mathbf{U}^n, \mathbf{W}) + \Delta t d(\mathbf{U}^n, \mathbf{W}) \quad (3.2.39)$$

gdzie

$$b^*(U, W) = \int_{\Omega} \tau_a W_{,i}^T A_i^T A_0^{-1} A_j U_{,j} dx \quad (3.2.40)$$

oraz gdzie wykorzystaliśmy fakt, że  $\int_{\Omega} W^T A_0^{-1} A_i U_{,i} = - \int_{\Omega} W_{,i}^T A_0^{-1} A_i U$ , co wynika z periodycznych warunków brzegowych. Oszacowanie wartości własnej  $\lambda$  operatora zdefiniowanego przez (3.2.39) otrzymujemy analogicznie jak (3.2.35) w postaci

$$|\lambda|^2 = \frac{a^2 + \Delta t^2 c^2}{(a + \Delta t b^*)^2} \quad (3.2.41)$$

gdzie  $a$  i  $c$  są jak poprzednio a  $b^* = b^*(U, \bar{U})$ . W zastosowanej poprzednio nierówności  $c^2 \leq ab$  musimy zastąpić  $b$  przez  $b^*$ . W tym celu zauważmy, że

$$b \leq \max_{x \in \Omega} 1/\tau_a b^* \quad (3.2.42)$$

co wraz z (3.2.41) i (3.2.37) pozwala oszacować  $|\lambda|^2$  następująco

$$|\lambda|^2 \leq \frac{a^2 + (\Delta t \max 1/\tau_a) \Delta t ab^*}{a^2 + 2\Delta t ab^* + \Delta t^2 b^{*2}} \quad (3.2.43)$$

Widoczne jest, że  $|\lambda|^2 \leq 1$ , o ile  $\Delta t \max 1/\tau_a \leq 2$ , czyli gdy

$$\Delta t \leq \min_K \frac{\max_i (h_i |\beta_i|)}{c + |\mathbf{u} \cdot \boldsymbol{\beta}|} \quad (3.2.44)$$

z oznaczeniami zastosowanymi w (3.1.3) oraz ze stałymi wektorami  $\mathbf{u}$  i  $\boldsymbol{\beta}$ . Wynik ten oznacza warunkową stabilność schematu.

Porównanie powyższych oszacowań stabilności metod Taylora–Galerkina i SUPG sugeruje korzystniejsze własności pierwszej z nich. Jest to jednakże przewaga pozorna. Łatwo zauważyć, że dla rozwiązań ustalonych parametr  $\Delta t$  nie znika ze sformułowania Taylora–Galerkina ((3.2.24) z  $U^n = U^{n-1}$ ), jest wielkością kontrolującą wyraz regularyzujący  $\int W_{,i}^T A_i A_j U_{,j}$ . Aby wpływ tej regularyzacji asymptotycznie zniknął,  $\Delta t$  musi być rzędu  $\min_K h_K$ , co jest ograniczeniem zbliżonym do (3.2.44). Z drugiej strony ustalenie  $\Delta t \sim \min_K h_K$  sprawia, że regularyzacja ta jest zbyt mała w obszarach grubszej siatki – wady tej nie ma SUPG, gdzie  $\tau_a$  zależy lokalnie od  $h_K$ . Poza tym należy dodać, że ograniczenie (3.2.44) nie jest warunkiem koniecznym stabilności, gdyż obecność wyrazu  $\Delta t^2 b^{*2}$  w mianowniku (3.2.43) oraz przyjęte „z zapasem” oszacowania (3.2.37) i (3.2.42) sprawiają, że możemy mieć  $|\lambda| \leq 1$  także dla  $\Delta t$  istotnie przekraczających ograniczenie (3.2.44). W końcu obecność lepkości naturalnej i sztucznej  $\nu_d \mathbf{I} \delta_{ij}$  dodatkowo zwiększa stabilność schematu.

### 3.3 Technika rozwiązywania problemów przepływów

Fizyczne zagadnienie przepływu otwartego rozpatrujemy teoretycznie w obszarze nieskończonym, w całej przestrzeni  $R^2$  czy  $R^3$ . Zagadnienie dyskretne z konieczności musimy sformułować w obszarze ograniczonym, co więcej, chcielibyśmy, aby był on możliwie niewielki, jeśli tylko pozwalałoby to na wystarczająco dokładne rozwiązanie przepływu. W praktyce procedura sprowadza się do zdefiniowania obszaru obliczeniowego otaczającego opływany obiekt ze sztucznymi warunkami brzegowymi na granicy tego obszaru. Noszą

one nazwę warunków na brzegu otwartym (*open boundary conditions*). Procedura taka zaburza w pewien sposób rozwiązywanie rzeczywistego problemu, jednakże modyfikacja ta może być akceptowalnie niewielka, jeśli obszar obliczeniowy jest wystarczająco duży.

Konstrukcja warunków na brzegu otwartym dla problemu przepływów ściśliwych oparta jest na rozróżnieniu, czy przepływ jest skierowany do wnętrza (wplyw) czy na zewnątrz obszaru i czy jest nad- czy poddźwiękowy. Szczegółową dyskusję tego problemu można znaleźć np. w pracach [30],[107], zwięźle omówienie zagadnienia podano w dodatku A.3, gdzie opisano też pozostałe rodzaje warunków brzegowych. W tym miejscu ograniczymy się jedynie do stwierdzenia, że warunki brzegowe zmieniają się od warunków typu Dirichleta dla wszystkich składowych rozwiązania dla naddźwiękowego wpływu aż po warunki typu Neumanna dla wszystkich składowych przy naddźwiękowym wypływie.

Istotnym elementem procedury rozwiązywania zadania przepływu jest przyjęcie kroku czasowego  $\Delta t$ . Algorytm (3.1.4) z formalnego punktu widzenia całkuje w czasie zagadnienie ewolucyjne zdefiniowane przez równania Naviera–Stokesa. Wykorzystanie go w takim właśnie celu wymagałoby dodatkowo zapewnienia kontroli dokładności aproksymacji przestrzennej dla całego rozpatrywanego przedziału czasowego. Jest to poza zakresem tej pracy. Ograniczając się do poszukiwania rozwiązań ustalonych wykorzystujemy procedurę całkowania w czasie jedynie jako szczególny rodzaj iteracji prostej prowadzącej do takich rozwiązań. Zgodnie z wcześniejszą dyskusją procedura (3.1.4) jest warunkowo stabilna w sensie stabilności liniowej. W praktyce możliwa do przyjęcia wielkość kroku czasowego  $\Delta t$  zależy od rozwiązywanego problemu. Typowym przedziałem czasowym, do którego odnosi się  $\Delta t$  jest minimalny czas przepływu informacji przez element siatki.  $\Delta t_0 = \min_K (h_K / (|u| + c))$ , gdzie  $h_K$  jest rozmiarem elementu,  $u$  prędkością przepływu zaś  $c$  prędkością dźwięku. Stosunek  $CFL = \Delta t / \Delta t_0$ , liczbę Couranta–Fridrichsa–Leviego, dla której algorytm zachowuje się stabilnie przyjmuje się w praktyce od 0.5 do kilkudziesięciu, „metodą prób i błędów”. Ogólnie można powiedzieć, że przyjmowanie dużej wartości  $CFL$  (a więc i dużych kroków czasowych  $\Delta t$ ) nawet jeśli nie powoduje utraty stabilności, to spowalnia zbieżność algorytmu iteracyjnego rozwiązywania układu równań liniowych. Dlatego też przy zwiększaniu  $CFL$  jesteśmy na ogół bardzo konserwatywni.

# Rozdział 4

## Adaptacyjna metoda elementów skończonych

### 4.1 Metoda siatek 1-nieregularnych

Znane są trzy techniki modyfikowania dystrybucji węzłów siatki w metodzie elementów skończonych: metoda przemieszczania węzłów, generowanie zmodyfikowanej nowej siatki (*remeshing*) oraz metoda podziału elementów na mniejsze „elementy-synów”. Zaletą ostatniej z technik jest precyzyjna lokalizacja modyfikacji siatki, możliwość tworzenia siatek z szybko zmieniającym się rozmiarem elementów  $h$  a także możliwość łatwego odwracania operacji podziału, tj. łączenia elementów na powrót. Co więcej, operacje podziałów i łączenia elementów nie są kosztowne numerycznie i nie wymagają projekcji rozwiązania z jednej siatki na drugą, jak to jest w przypadku poprzednich metod. Przy podziale elementu rozwiązanie na nowej siatce interpoluje w sposób ścisły rozwiązanie ze starej siatki, co zapobiega nagłej jego zmianie, gdy jest elementem zbieżnego ciągu rozwiązań iteracyjnych.

W niniejszej pracy stosowane są elementy czworokątne, które mogą być dzielone na „elementy-synów”, podejście to określamy jako metodę adaptacyjną typu  $h$ . Elementy mogą być dzielone na pół wzdłuż jednego z dwu możliwych kierunków (mówimy wtedy o anizotropowej adaptacji typu  $h$ ) lub na 4 elementy (metoda izotropowa). Pierwsza technika jest ogólniejsza, motywacją do jej zastosowania jest optymalna aproksymacja rozwiązań bliskich jednowymiarowym w warstwach przyściennych: metoda ta dopuszcza tworzenie elementów silnie wydłużonych.

Zakładamy, że siatka początkowa składa się z elementów czworokątnych i że jest ona quasi-równomierna, tj. że istnieją takie stałe  $C_1$  i  $C_2$ , że

$$H_K \geq C_1 h, \quad \frac{R_K}{H_K} \geq C_2, \quad (4.1.1)$$

dla wszystkich elementów początkowych  $K$ , gdzie  $H_K = \text{dia}(K)$ ,  $H = \max_K H_K$ ,  $R_K$  jest promieniem największego koła wpisanego w  $K$ . Elementy  $K$  rozważamy jako obrazy elementu wzorcowego (odniesienia)  $\hat{K} = [0, 1]^2$  w odwzorowaniu  $F$ . Jeśli ograniczamy się do elementów równoległobocznych, to odwzorowanie  $F$  jest afiniczne. W tym też przypadku rozmiary elementów wynikłych z  $l$  podziałów izotropowych spełniają warunki podobne do (4.1.1) a mianowicie

$$h_K \geq 2^{-l} C_1 H \quad \text{ i } \quad \frac{\rho_K}{h_K} \geq C_2 \quad (4.1.2)$$

gdzie znów  $h_K = \text{dia}(K)$  i  $\varrho_K$  jest promieniem największego koła wpisanego w  $K$ . Elementy wynikłe z takich podziałów także można rozważać jako obrazy elementu wzorcowego  $\hat{K}$  w odwzorowaniu afinicznym  $\phi$ , które jest odpowiednio przeskalowanym odwzorowaniem  $F$ . Odwzorowanie  $\phi$  spełnia następujące warunki

$$\|\nabla\phi\|_{\infty, \hat{K}} \leq \frac{h_K}{\rho_K}, \quad \|\nabla\phi^{-1}\|_{\infty, K} \leq \frac{h_{\hat{K}}}{\rho_K} \leq \frac{h_{\hat{K}}/C_2}{h_K}, \quad (4.1.3)$$

gdzie  $h_{\hat{K}} = \sqrt{2}$ ,  $\varrho_{\hat{K}} = 1/2$  są parametrami elementu  $\hat{K}$ . Warunki (4.1.3) przedstawiają znane własności odwzorowań afinicznych [84] oraz zależność (4.1.2)<sub>2</sub>.

W przypadku podziałów anizotropowych warunek (4.1.2)<sub>2</sub> na ogół przestaje być spełniony: parametr  $C_2$  musi być zmniejszony o czynnik  $2^{|\Delta l|}$ , gdzie  $\Delta l$  jest różnicą liczby podziałów w obu kierunkach. Dlatego też warunek (4.1.3)<sub>2</sub> również przestaje być spełniony, o ile nie zmienimy  $C_2$ .

W przypadku ogólnym możemy rozważać siatki krzywoliniowe z nieliniowym odwzorowaniem  $F$ . Wówczas regularność  $F$  i  $\phi$  nie może być scharakteryzowana wyłącznie przez parametry  $h_K$  i  $\varrho_K$ . Dlatego też dla odwzorowań nieliniowych postulujemy spełnianie *a priori* warunków typu (4.1.3)

$$\|\nabla F\|_{\infty, \hat{K}} \leq C_1 H, \quad \|\nabla F^{-1}\|_{\infty, K} \leq C_2/H, \quad \forall K, \quad (4.1.4)$$

co implikuje, że dla elementów otrzymanych z  $l$  podziałów izotropowych mamy

$$\|\nabla\phi\|_{\infty, \hat{K}} \leq C_1 h_K, \quad \|\nabla\phi^{-1}\|_{\infty, K} \leq C_2/h_K \quad (4.1.5)$$

z wielkością  $h_K = 2^{-l}H$  przyjętą z definicji jako rozmiar  $K$  i  $\phi: \hat{K} \rightarrow K$ . Warunki te formalnie odpowiadają (4.1.3).

Oczywiście każda siatka składająca się ze skończonej liczby elementów z  $h_K > \varrho_K > 0$  może być nazwana quasi-równomierną przy odpowiednio małych parametrach  $C_1$  i  $C_2$ . Ścisłej mówiąc pojęcie quasi-równomierności dotyczy raczej nieskończonej rodziny siatek i jest rozumiane jako możliwość scharakteryzowania wszystkich siatek tej rodziny wspólnymi parametrami  $C_1$  i  $C_2$ . Ponieważ w naszych rozważaniach obecna jest zwykle siatka pierwotna, wygodnie jest mówić o zachowaniu quasi-równomierności w stosunku do siatki pierwotnej. W tym znaczeniu siatki wynikłe z podziałów anizotropowych nie są quasi-równomierne.

Podziały elementów prowadzą do sytuacji, w których dany element może mieć dwa lub więcej mniejszych „elementów-sąsiadów” wzdłuż jednego boku. O siatkach takich mówimy, że są  $n$ -nieregularne, gdzie  $n$  jest maksymalną liczbą wierzchołków mniejszych sąsiadów leżących we wnętrzu boku większego elementu. W celu uniknięcia znacznego skomplikowania algorytmów ograniczamy się do siatek 1-nieregularnych. Praktycznie oznacza to, że jedynie dwa mniejsze elementy-sąsiedzi mogą pojawić się wzdłuż boku danego elementu. Jeśli zaś chcielibyśmy jeden z nich dodatkowo podzielić, wprawd należałoby podzielić większego sąsiada. Zachowanie 1-nieregularności określa się też czasem zasadą „dwa do jeden”.

Ciągłość aproksymacji dla siatek 1-nieregularnych zapewnia się przez odpowiednie związanie stopni swobody wzdłuż boku mniejszego sąsiada  $u_i^{(1)}$  ze stopniami swobody wzdłuż wspólnego boku większego sąsiada  $U_i^{(2)}$ :

$$u_i^{(1)} = \sum_{j=1}^{p+1} U_j^{(2)} / R_{ij} \quad i=1, \dots, p+1, \quad (4.1.6)$$



z nieosobliwą macierzą  $R_{ij}$  (zależną od wyboru elementowych funkcji kształtu) i  $p$  będącym stopniem aproksymacji wzdłuż rozważanego boku [33]. Wygodnie jest zdefiniować *faktyczne* stopnie swobody (*actual degrees-of-freedom*) mniejszego elementu  $U_i^{(1)}$  za pomocą relacji

$$u_i^{(1)} = \sum_{j=1}^{p+1} U_j^{(1)} R_{ji}, \quad i = 1, \dots, p+1 \quad (4.1.7)$$

co pozwala zapisać warunek ciągłości aproksymacji w standardowy sposób, jak dla siatek regularnych:  $U_i^{(1)} = U_i^{(2)}$ ,  $i = 1, \dots, p+1$ . Uzasadnia to także użycie słowa „faktyczne” w nazwie  $U_i^{(1)}$  – zarówno  $U_i^{(2)}$  jak i równe im  $U_i^{(2)}$  są utożsamiane z globalnymi stopniami swobody występującymi faktycznie w równaniach metody elementów skończonych. Dla wizualizacji faktycznych stopni swobody możemy wyobrazić sobie rozszerzenie oryginalnego obszaru elementu o boki przyległych większych sąsiadów (o ile takie występują) i przyporządkować faktyczne stopnie swobody  $U_i^{(1)}$  węzłom rozłożonym równomiernie wzdłuż tak wydłużonych boków (*faktycznym* węzłom). Takie podejście umożliwia spojrzenie na procedurę składania z elementów siatki zupełnie tak jak na składanie siatki regularnej: odpowiednie węzły sąsiednich elementów są łączone i związane z nimi stopnie swobody są przyrównywane, rys. 4.1. Związki typu (4.1.6) zebrane dla wszystkich boków ze związanymi stopniami swobody mogą być rozumiane jako zmiana bazy w przestrzeni funkcjonalów liniowych na elementowych funkcjach kształtu  $V^h(K)$ . Jeśli oznaczymy przez  $\bar{R}$  macierz odpowiadającą związkom (4.1.6) i dodatkowo jeśli uwzględnimy, że stopnie swobody wzdłuż pozostałych boków pozostają nie zmienione (tj.  $U_i^{(1)} = u_i^{(1)}$ , co daje blok macierzy identyczności w  $\bar{R}$ ), wówczas elementowe bazowe funkcje kształtu  $\psi_i^c$  dualne do  $U_i^{(1)}$  można wyrazić przez standardowe lagranżowskie funkcje kształtu  $\psi_i$  (dualne do  $u_i^{(1)}$ ) jako

$$\psi_i^c = \sum_{j=1}^{(p+1)^2} \bar{R}_{ij} \psi_j, \quad i = 1, \dots, (p+1)^2 \quad (4.1.8)$$

zgodnie z regułą transformacji baz dualnych. Funkcje  $\psi_i^c$  nazywamy *związanymi* bazowymi funkcjami kształtu (*constrained base shape functions*). Rys. 4.2 przedstawia takie funkcje dla elementu biliniowego i bikwadratowego z więzami. Związane funkcje kształtu spełniają (z definicji) warunek dualności

$$\psi_i^c(\mathbf{x}_j^a) = \delta_{ij}, \quad i, j = 1, \dots, (p+1)^2, \quad (4.1.9)$$

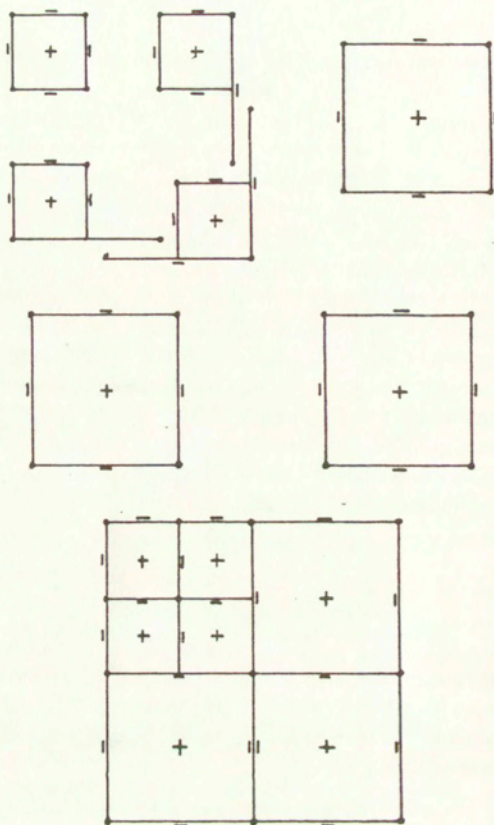
gdzie  $\mathbf{x}_j^a$  są faktycznymi węzłami.

Faktyczne stopnie swobody i związane funkcje kształtu pozwalają na zdefiniowanie uogólnionego operatora interpolacji  $I_h^c$  dla elementu z więzami

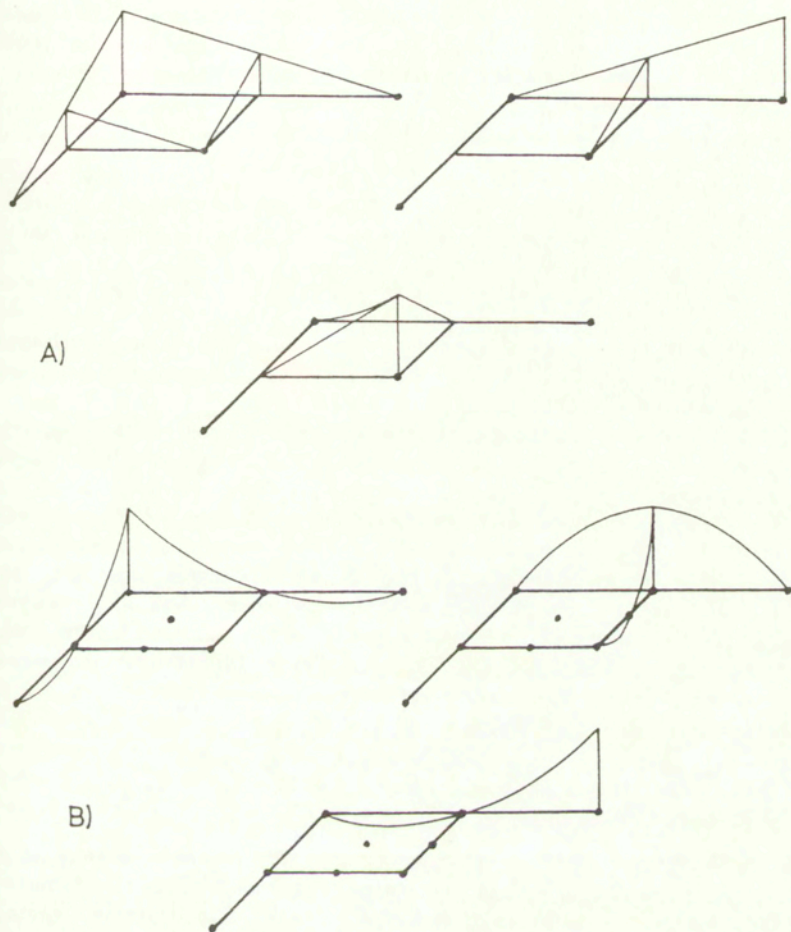
$$v_I = I_h^c v = \sum_{j=1}^{(p+1)^2} v(\mathbf{x}_j^a) \psi_j^c(\mathbf{x}) \in V^h(K). \quad (4.1.10)$$

Operator  $I_h^c$  gwarantuje, że przez sklepanie interpolantów  $v_I$  na elementach siatki 1-nieregularnej otrzymamy funkcję ciągłą z przestrzeni skończenie elementowej  $V^h$  (dla dowolnych sąsiednich elementów interpolanty są wielomianami stopnia  $p$  wzdłuż wspólnego brzegu i zgadzają się w  $p+1$  faktycznych węzłach  $\mathbf{x}_i^a$ ).

Wszystkie informacje oraz modyfikacje siatek adaptacyjnych są zawiadywane przez grupę algorytmów określaną jako struktura danych programu adaptacyjnego. Więcej informacji dotyczących obliczeń z wykorzystaniem siatek 1-nieregularnych można znaleźć np. w [33].



Rysunek 4.1: Składanie siatki 1-nieregularnej z elementów z więzami (· węzły narożne, - węzły na bokach elementów, + węzły centralne).



Rysunek 4.2: Bilineowe i bikwadratowe bazowe funkcje kształtu dla elementu z więzami.

## 4.2 Residualne oszacowania błędu a posteriori

Oszacowanie błędu istniejącej aproksymacji skończenie elementowej – oszacowanie *a posteriori* – jest istotnym czynnikiem efektywnej adaptacji siatek. Wiarygodny rozkład wskaźników błędu pozwala bowiem na trafne zagęszczenie aproksymacji tam, gdzie rzeczywisty błąd jest duży. Prócz oszacowania błędu rozwiązania, w procesie adaptacji można posłużyć się też wielkością residuum tego rozwiązania. Dla szerokiej klasy problemów (w tym i nieliniowych niestacjonarnych) dających się zapisać w ogólnej postaci:

Znaleźć  $u_0 \in X$ , takie że

$$F(u_0) = 0 \quad \text{w } Y' \quad (4.2.11)$$

gdzie  $X, Y$  są przestrzeniami Banacha,  $Y'$  zaś przestrzenią ciągłych liniowych funkcjonałów na  $Y$ , można pokazać, że jeśli pochodna  $DF(u)$  jest izomorfizmem, ciągłym w pewnym otoczeniu  $u_0$  i spełniającym warunek Lipschitza,  $u_h$  zaś przybliżeniem  $u_0$ , to:

$$C_1 \|F(u_h)\|_{Y'} \leq \|u_0 - u_h\|_X \leq C_2 \|F(u_h)\|_{Y'} \quad (4.2.12)$$

co wiąże wielkość błędu z wielkością residuum. Stałe  $C_1, C_2$  zależą od własności  $DF(u)$  w otoczeniu  $u_0$  [112]. (4.2.12) wskazuje, iż redukując residuum redukujemy jednocześnie błąd i na odwrót.

W przypadku równań Eulera, tj. równań przepływu nie uwzględniających lepkości, residuum rozwiązania skończenie elementowego jest funkcją kawałkami ciągłą, gdyż w równaniach występują pochodne co najwyżej pierwszego rzędu zaś aproksymacja jest klasy  $C^0$ . Funkcję taką można mierzyć używając wielu norm, jak choćby  $L^1, L^2$  czy  $L^\infty$ , które są lokalne a więc pozwalają wiązać wielkość residuum z poszczególnymi elementami. Inaczej przedstawia się sytuacja z równaniami Naviera–Stokesa, tj. równaniami przepływu uwzględniającymi lepkość. Strumienie lepkie zależą od pierwszych pochodnych rozwiązania, przez co residuum odpowiadające aproksymacji o ciągłości  $C^0$  musi być rozumiane jako funkcja uogólniona – dystrybucja. Dla tak rozumianego residuum  $R_h$  odpowiednią miarą jest norma w przestrzeni funkcjonałów liniowych np. w przestrzeni Sobolewa  $H^{-1}(\Omega) := (H_0^1(\Omega))'$ :

$$\|R_h\|_{H^{-1}(\Omega)} := \sup_{W \in H_0^1(\Omega)} \int_{\Omega} W R_h dx / \|W\|_{H^1(\Omega)} \quad (4.2.13)$$

Zaletą powyższej normy dualnej jest jej matematyczna poprawność, wadą zaś jej globalny charakter i złożoność obliczania wymagającego optymalizacji na zbiorze nieskończonym. Taką samą trudność napotykaemy w przypadku najlepiej teoretycznie rozpoznanych problemów – liniowych zadań eliptycznych. Dla tej klasy zadań opracowano szereg metod szacowania normy residuum bez potrzeby rozwiązywania optymalizacji (4.2.13). Ponieważ w przypadku zadań liniowych przy odpowiednim doborze normy w przestrzeni dualnej norma residuum okazuje się być identyczna z normą błędu rozwiązania, wspomniane techniki określa się najczęściej jako residualne oszacowania błędu. Użyteczność oszacowań residualnych dla zadań przepływów wiąże się z tym, że ich rozwiązania są całkowane w czasie poprzez rozwiązywanie ciągu liniowych problemów brzegowych, których charakter jest bliższy problemom eliptycznym.

Wśród technik residualnych rozważymy trzy metody, które okazują się przydatne w praktycznych zastosowaniach mechaniki płynów: metodę residualną na podobszarach, autorstwa Babuški i Rheinboldta [10], metodę residualną na osobnych elementach [82] oraz

metodę samozrównoważonych residuów Ainswortha-Odena [2],[3]. Elementem wspólnym tych metod jest rozwiązywanie lokalnych problemów brzegowych ograniczonych do niewielkich podobzszarów  $\Omega_i$ . Mają one następującą postać:

Znaleźć takie  $\Phi_i \in H \subset H^1(\Omega_i)$ , że

$$a(\Phi_i, \mathbf{W}) = L(\mathbf{W}) - B(\mathbf{U}_h, \mathbf{W}) + \int_{\partial\Omega_i} \mathbf{W}^T \mathbf{Q} dS \quad \forall \mathbf{W} \in H \quad (4.2.14)$$

gdzie  $B(\cdot, \cdot)$  i  $L(\cdot)$  są formą dwuliniową i liniową oryginalnego sformułowania słabego danego problemu,  $B(\mathbf{U}, \mathbf{W}) = L(\mathbf{W})$ , ograniczonymi do  $\Omega_i$ ,  $\mathbf{U}_h$  jest przybliżeniem skończenie elementowym. Forma  $a(\cdot, \cdot)$  jest pomocniczą symetryczną formą dwuliniową o tej własności, że  $B$  jest koercywna względem normy indukowanej przez  $a(\cdot, \cdot)$ ,  $\|\cdot\|_a := a(\cdot, \cdot)^{1/2}$ . Funkcjonal liniowy po prawej stronie (4.2.14) odpowiada pewnej formie residuum dla rozwiązania przybliżonego  $\mathbf{U}_h$ , składa się z części pochodzącej z wnętrza obszaru  $\Omega_i$ , tj.  $L(\mathbf{W}) - B(\mathbf{U}_h, \mathbf{W})$  oraz z kontrybucji brzegowej w postaci ostatniej całki w (4.2.14). Rozwiązania  $\Phi_i$  nazywają się funkcjami wskaźnika błędu. Globalne oszacowanie błędu przedstawia się za ich pomocą następująco:

$$\|\mathbf{U} - \mathbf{U}_h\|_a^2 \leq C \sum_i \|\Phi_i\|_a^2 \quad (4.2.15)$$

W dalszej części omawiamy szczegóły zastosowania metod residualnych do problemów przepływów.

#### Metoda residualna na podobzszarach.

Metoda Babuški i Rheinboldta powstała jako pierwsza z technik residualnych, przedstawiono ją w pracy [10] w 1978 roku. Dla zdefiniowania metody konieczne jest przyjęcie pewnego rozkładu jedności na obszarze obliczeniowym  $\Omega$ , tj. zbioru funkcji  $\{\phi_1, \dots, \phi_n\}$  takiego, że  $\phi_i \in H^1(\Omega)$  i  $\sum_{i=1}^n \phi_i(\mathbf{x}) = 1, \quad \forall \mathbf{x} \in \Omega$ . Podobzszary  $\Omega_i$  będące dziedziną lokalnych zadań (4.2.14) przyjmowane są jako nośniki funkcji  $\phi_i$ :  $\Omega_i = \text{supp}(\phi_i)$ . Elementem w pewnym stopniu dowolnym w metodzie jest symetryczna forma dwuliniowa  $a(\cdot, \cdot)$ . Swoboda jej wyboru jest ograniczona wymaganiem, aby oryginalna forma  $B(\cdot, \cdot)$  była koercywna względem  $\|\cdot\|_a$ . W przypadku formy dwuliniowej  $B(\cdot, \cdot)$  związanej ze sformulowaniem zadania przepływu (3.1.4) odpowiednim wyborem  $a$  jest zsymetryzowana postać  $B$  za pomocą macierzy symetryzatora równań Naviera-Stokesa:

$$a(\mathbf{U}, \mathbf{W}) = \int_{\Omega} \mathbf{W}^T \mathbf{A}_0^{-1} \mathbf{U} dx + \Delta t \int_{\Omega} \mathbf{W}_{,i}^T (\mathbf{A}_0^{-1} \mathbf{K}_{ij} + \tau_a \mathbf{A}_i^T \mathbf{A}_0^{-1} \mathbf{A}_j + \mathbf{A}_0^{-1} \delta_{ij}) \mathbf{U}_j dx \quad (4.2.16)$$

(W dodatku A.4 pokazano koercywność formy  $B$  względem normy  $\|\cdot\|_a$ .) Rozkład jedności obszaru  $\Omega$  można wprowadzić na wiele sposobów. W przypadku złożonych siatek adaptacyjnych najbardziej praktyczny wydaje się wybór  $\phi_i$  w postaci globalnych liniowych (lub biliniowych) bazowych funkcji kształtu odpowiadających danej siatce. Funkcje te oczywiście definiują rozkład jedności, gdyż funkcja stała jest przez nie interpolowana w sposób ścisły.

Lokalne problemy brzegowe (4.2.14) formułowane są z zerowym warunkiem brzegowym typu Dirichleta, w związku z czym nie uwzględnia się brzegowej kontrybucji obciążień w (4.2.14). Mają one więc postać:

Znaleźć takie  $\Phi_i$ , że  $\Phi_i = 0$  na  $\partial\Omega_i$  oraz

$$a(\Phi_i, \mathbf{W}) = L(\mathbf{W}) - B(\mathbf{U}_h, \mathbf{W}) + \int_{\partial\Omega_i} \mathbf{W}^T \mathbf{Q} dS, \quad \mathbf{W} = 0 \text{ na } \partial\Omega_i \quad (4.2.17)$$

Ostatnim krokiem w zdefiniowaniu metody jest przedstawienie błędu globalnego (4.2.15) w postaci kontrybucji  $e_K$  poszczególnych elementów  $K$ :

$$\|U - U_h\|_a^2 \leq C \sum_{i=1}^n \|\Phi_i\|_a^2 = \sum_K \left( \sum_{\Omega_j: \Omega_j \cap K \neq \emptyset} \|\Phi_{j|K}\|_a^2 \right) =: \sum_K e_K^2 \quad (4.2.18)$$

Przyjmujemy, że wielkości  $e_K$  są lokalnymi (elementowymi) wskaźnikami błędu. Rozwiązania ściśle lokalnych problemów (4.2.17) są w praktyce trudno osiągalne. Dlatego też zastępuje się je ich aproksymacjami skończenie elementowymi na siatce oryginalnej wzbogaconej przez podniesienie stopnia aproksymacji o 1.

### Metoda residualna na pojedynczych elementach.

Wadą procedury z użyciem podziału jednostki jest jej względnie duży koszt: podobszary  $\Omega_i$  składają się z kilku elementów (w przypadku regularnych siatek dwuwymiarowych liczba elementów składowych wynosi 4). Bardziej ekonomiczna jest technika residualna z lokalnymi problemami brzegowymi zdefiniowanymi na poszczególnych elementach. Ponieważ podobszary te nie zachodzą na siebie, konieczne jest w tym wypadku uwzględnienie składników residuum pochodzących z brzegów elementów. Lokalny problem brzegowy jest z założenia rozwiązywany w przestrzeni skończenie wymiarowej  $V_{h,2}^0(K)$  składającej się z funkcji bikwadratowych znikających w węzłach narożnych i ma następującą postać:

Znaleźć takie  $\Phi_K \in V_{h,2}^0(K)$ , że

$$B_K(\Phi_K, W) = \quad (4.2.19)$$

$$L_K(W) - B_K(U_h, W) + \int_{\partial K \setminus \partial \Omega} \left[ \sum_{k,l=1}^2 W^T \mathbf{a}_{kl} U_{h,l} \mathbf{n}_k \right] ds \quad \forall W \in V_{h,2}^0(K)$$

W sformułowaniu tym  $B_K(\cdot, \cdot)$  i  $L_K(\cdot)$  są formami z oryginalnego problemu brzegowego ograniczonymi do funkcji zdefiniowanych na  $K$ ,  $\mathbf{a}_{kl} = \Delta t(\mathbf{K}_{kl} + \tau_a \mathbf{A}_k \mathbf{A}_l + \nu_d \delta_{kl} \mathbf{I})$ ,  $[[\mathbf{v}]]$  oznacza skok funkcji  $\mathbf{v}$  na brzegu  $\partial K$ . Elementowe wskaźniki błędu wyrażają się przez  $\Phi_K$  następująco:

$$e_K^2 = B_K(\Phi_K, \mathbf{A}_0^{-1} \Phi_K), \quad (4.2.20)$$

gdzie  $\mathbf{A}_0$  jest macierzą symetryzującą równań Naviera–Stokesa.

Powyższe sformułowanie metody pozornie odbiega od ogólnego schematu: forma dwuliniowa lokalnych zadań brzegowych jest oryginalną (a więc niesymetryczną) formą  $B$ . Jednak zaprezentowana wersja metody jest uproszczeniem podejścia, w którym problemy lokalne sformułowane są z użyciem  $a(\cdot, \cdot)$ . Formę  $a$  można bowiem rozważać jako zsymetryzowaną wersję  $B$  w następującym rozumieniu:

$$a(\mathbf{A}_0^{-1} U, V) = B(U, V) \quad \forall U, V \quad (4.2.21)$$

o ile  $\mathbf{A}_0 = \text{const}$ , lub gdy pochodne  $\mathbf{A}_0$  są zaniedbywalne. Związek (4.2.21) pozwala zastąpić lewą stronę (4.2.19) przez  $a(\Psi_K, W)$ , gdzie  $\Psi_K = \mathbf{A}_0^{-1} \Phi_K$ , zaś wskaźniki błędu przedstawić jako  $e_K^2 = a(\Psi_K, \Psi_K)$ . Mimo że zaniedbanie pochodnych  $\mathbf{A}_0$  jest raczej dyskusyjne, wersja metody przedstawiona równaniem (4.2.19) okazuje się być efektywna w praktyce.

### Metoda samozrównoważonych residuów.

Oszacowaniem residualnym o szczególnej atrakcyjności jest metoda zaproponowana przez Ainswortha i Odena. Jej cechą jest asymptotyczna dokładność tj. asymptotyczne pokrywanie się rzeczywistego globalnego błędu  $e$  i globalnego estymowanego błędu  $e_{est}$ :

$$\frac{e}{e_{est}} \rightarrow 1, \text{ jeśli } h \rightarrow 0 \quad (4.2.22)$$

Trudnością metody jest konieczność rozdzielenia dystrybucyjnej części residuum związanej ze skokiem pochodnych rozwiązania  $U_h$  na sąsiednie elementy w taki sposób, aby dla wszystkich elementów  $K$  spełnić równość:

$$L_K(\mathbf{1}^{(k)}) - B_K(U_h, \mathbf{1}^{(k)}) + \int_{\partial K} \mathbf{1}^{(k)} \cdot \langle \mathbf{n}_K \cdot \mathbf{Q} \rangle dS = 0, \quad (4.2.23)$$

dla  $k = 1, \dots, 4$ , i gdzie

$$\{\mathbf{Q}\}_i = \Delta t(\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}) U_{h,j}, \quad i = 1, 2 \quad (4.2.24)$$

jest uogólnionym strumieniem,  $\mathbf{n}_K$  wektorem normalnym do  $\partial K$ , symbol  $\langle \mathbf{v} \rangle$  oznacza uśrednienie wielkości  $\mathbf{v}$  zmieniającej się skokowo na granicy elementów oraz

$$\mathbf{1}^{(1)} = \{1, 0, 0, 0\}^T, \dots, \mathbf{1}^{(4)} = \{0, 0, 0, 1\}^T.$$

Iloczyn skalarny w (4.2.23) jest rozumiany jako  $\mathbf{n} \cdot \mathbf{Q} := \sum_{i=1}^2 n_i \{\mathbf{Q}\}_i$ , tj. w wyniku otrzymujemy strumień przez powierzchnię prostopadłą do  $\mathbf{n}$ . (Analogiczne wielkości  $\mathbf{Q}$  w przypadku teorii sprężystości są naprężenia na  $\partial K$  zaś warunek (4.2.23) można zinterpretować jako zrównoważenie sił residualnych, skąd wywodzi się nazwa metody.)

Operację uśredniania naprężeń można przeprowadzić za pomocą obliczeń lokalnych (tj. bez konieczności rozwiązywania globalnego układu równań). Poniżej przedstawiamy wersję tej procedury zaproponowaną w [1].

Uśredniony strumień przedstawiamy jako średnią ważoną strumieni w sąsiednich elementach  $K$  i  $L$ :

$$\langle \mathbf{Q}^{(k)} \cdot \mathbf{n}_K \rangle = [\alpha_{KL}^{(k)}(s) \mathbf{Q}_K^{(k)}(s) + \alpha_{LK}^{(k)}(s) \mathbf{Q}_L^{(k)}(s)] \cdot \mathbf{n}_K \quad (4.2.25)$$

gdzie  $\alpha_{KL}^{(k)}(s), \alpha_{LK}^{(k)}(s)$  oznaczają funkcje wagowe,  $\alpha_{KL}^{(k)} + \alpha_{LK}^{(k)} = 1, k = 1, \dots, 4$  numeruje składowe strumienia. Okazuje się, iż wystarczy założyć, że są one liniowe wzdłuż boków elementów, aby możliwe było spełnienie (4.2.23). Funkcje wagowe możemy więc ogólnie przedstawić następująco:

$$\alpha_{KL}^{(k)}(s) = \sum_A \alpha_{KL,A}^{(k)} \psi_A(s) \quad (4.2.26)$$

gdzie  $\psi_A$  oznacza globalne liniowe (lub biliniowe) bazowe funkcje kształtu związane z narożnymi węzłami  $A$  siatki,  $\alpha_{KL,A}^{(k)}$  są stałymi współczynnikami. W celu ich wyznaczenia dla każdego nośnika funkcji kształtu  $\psi_A$ ,  $\text{supp}(\psi_A)$ , obliczamy następujące charakterystyki:

i) macierz  $\{T_{ij}\}_{i,j=1,n}$ , (gdzie  $n$  jest liczbą elementów w  $\text{supp}(\psi_A)$ ), taką że:

$$\{\mathbf{T}_A\}_{LM} = \begin{cases} C_L, & \text{jeśli } L = M \\ -1 & \text{jeśli elementy } L \text{ i } M \text{ mają wspólną krawędź} \\ 0, & \text{w pozostałych wypadkach} \end{cases} \quad (4.2.27)$$

gdzie  $C_L$  jest liczbą elementów-sasiadów elementu  $L$  w danym bloku elementów z  $\text{supp}(\psi_A)$ .

ii) wektor  $b_{K,A}^{(k)}$  i macierz  $\varrho_{KL,A}^{(k)}$  zdefiniowane za pomocą residuum na elemencie:

$$\begin{aligned} b_{K,A}^{(k)} &= L_K(\psi_A^{(k)}) - B_K(\mathbf{U}^{n+1}, \psi_A^{(k)}) + \int_{\partial K \setminus \partial \Omega} (\psi_A^{(k)})^T \langle \mathbf{n}_K \cdot \mathbf{Q} \rangle_{1/2} dS \\ \varrho_{KL,A}^{(k)} &= - \int_{\Gamma_{KJ}} (\psi_A^{(k)})^T [[\mathbf{n} \cdot \mathbf{Q}]] dS, \quad k = 1, \dots, 4 \end{aligned} \quad (4.2.28)$$

W powyższym  $\psi_A^{(1)} = (\psi_A, 0, 0, 0), \dots, \psi_A^{(4)} = (0, 0, 0, \psi_A)$  są wektorowymi bazowymi funkcjami kształtu odpowiadającymi węzłowi  $A$ ,  $\langle \mathbf{n}_K \cdot \mathbf{Q} \rangle_{1/2}$  oznacza strumień na brzegu elementu  $K$  wzięty jako średnia arytmetyczna strumieni na brzegu  $K$  i na brzegu elementu sąsiedniego, symbol  $[[v]]$  oznacza skok funkcji  $v$  przy przejściu wspólnej krawędzi sąsiednich elementów  $K, L$ ,  $\Gamma_{KL}$ .

Powyższe wielkości służą do sformułowania układów równań liniowych

$$\mathbf{T}_A \lambda_A^{(k)} = b_A^{(k)}, \quad k = 1, \dots, 4 \quad (4.2.29)$$

ze względu na parametry  $\lambda_A^{(k)} = \{\lambda_{1,A}^{(k)}, \dots, \lambda_{n,A}^{(k)}\}$  (przy  $b_A^{(k)} = \{b_{1,A}^{(k)}, \dots, b_{n,A}^{(k)}\}$ ,  $k = 1, \dots, 4$ ), przez które wyrażamy poszukiwane współczynniki  $\alpha_{KL,A}^{(k)}$ :

$$\alpha_{KL,A}^{(k)} = \frac{1}{2} + (\lambda_{K,A}^{(k)} - \lambda_{L,A}^{(k)}) / \varrho_{KL,A}^{(k)}. \quad (4.2.30)$$

(W przypadku, gdy  $|\varrho_{KL}^{(k)}|$  jest bardzo małe, tj. gdy skok strumieni  $\mathbf{Q}$  jest mały, procedurę uśredniania można pominąć).

Po wyznaczeniu samozrównoważonych residuów dla każdego elementu rozwiązujemy problem brzegowy:

Znaleźć takie  $\Phi_K \in H^1(K)$ , że

$$a(\Phi_K, \mathbf{W}) = L_K(\mathbf{U}_h) - B_K(\mathbf{U}_h, \mathbf{W}) + \int_{\partial K \setminus \partial \Omega} \mathbf{W}^T \langle \mathbf{n}_K \cdot \mathbf{Q} \rangle dS \quad \forall \mathbf{W} \in H^1(K) \quad (4.2.31)$$

a następnie obliczamy elementowe wskaźniki błędów:

$$e_K^2 = a(\Phi_K, \Phi_K) \quad (4.2.32)$$

W praktyce problem brzegowy (4.2.31) rozwiązuje się we wzbogaconej przestrzeni funkcji skończenie elementowych.

### 4.3 Anizotropowe oszacowania błędu interpolacji w normach $L^2$ i $H^1$

W przeciwieństwie do klasycznych problemów eliptycznych drugiego rzędu, dla których najodpowiedniejszą miarą błędu jest jego norma  $H^1$ , nie jest oczywiście, w jakiej normie powinniśmy mierzyć dokładność rozwiązań równań Naviera–Stokesa. Dyskusję tego zagadnienia przedstawimy nieco później, obecnie natomiast przedstawimy oszacowania interpolacyjne dla elementów wydłużonych w normie  $L^2$  i w półnormie  $H^1$ .

Anizotropowe siatki adaptacyjne powstają w wyniku ciągu podziałów elementów czworokątnych składających się na siatkę początkową. Zakładamy, że siatka początkowa jest



quasi-równomierna, tzn. że spełnia warunki (4.1.4). Jak wspomniano wcześniej, elementy powstałe z izotropowych podziałów także spełniają warunek ograniczonego wydłużenia (4.1.5)<sub>2</sub> w przeciwieństwie do elementów wygenerowanych za pomocą podziałów anizotropowych. Naszym celem jest znalezienie oszacowań błędu interpolacji dla elementów nie spełniających (4.1.5)<sub>2</sub> w odróżnieniu od dobrze znanych wyników dla elementów zachowujących to ograniczenie (*shape regular elements*).

Rozważmy na początek sytuację, gdy elementy siatki początkowej są obrazami elementu wzorcowego (odniesienia)  $\hat{K} = [0, 1]^2$  w odwzorowaniu afinicznym  $F$  (tj. elementy równoległoboczne). Wówczas każdy element  $K$  siatki adaptacyjnej jest także obrazem  $\hat{K}$  w pewnym innym odwzorowaniu afinicznym  $\phi$  (odpowiednio skalowanym  $F$ ). Załóżmy, że elementy są typu Lagrange'a stopnia  $p$ . Oznaczając przez  $u_h$  lagranżowski interpolant funkcji  $u$  możemy napisać:

$$\|u - u_h\|_{0,K}^2 = \quad (4.3.33)$$

$$\int_{\hat{K}} (\hat{u} - \hat{u}_h)^2 \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} d\xi \leq C \int_{\hat{K}} \left[ \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi_1^{p+1}} \right)^2 + \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi_2^{p+1}} \right)^2 \right] \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} d\xi$$

gdzie  $\hat{u} = u \circ \phi$ ,  $\hat{u}_h = u_h \circ \phi$ ,  $\partial(x_1, x_2)/\partial(\xi_1, \xi_2)$  oznacza jacobian odwzorowania  $\phi$ . W ostatnim kroku zastosowaliśmy klasyczny wynik interpolacyjny dla elementu Lagrange'a w kształcie kwadratu [26]. Oznaczając przez  $k$  i  $l$  wersory boków elementu  $K$  oraz przez  $h_k, h_l$  długości tych boków możemy zapisać (4.3.33) następująco:

$$\|u - u_h\|_{0,K}^2 \leq C \left[ h_k^{2(p+1)} \int_K \left( \frac{\partial^{p+1} u}{\partial k^{p+1}} \right)^2 dx + h_l^{2(p+1)} \int_K \left( \frac{\partial^{p+1} u}{\partial l^{p+1}} \right)^2 dx \right] \quad (4.3.34)$$

Uogólnienie powyższego wyniku na elementy zakrzywione jest łatwe i właściwie polega na wprowadzeniu odpowiedniej notacji. Rzeczywistym problemem jest podanie stosownej definicji rozmiarów  $h_k$  i  $h_l$  takich elementów. Pamiętając jednak, że w istocie kontrolujemy raczej liczby podziałów niż zmieniające się w sposób ciągly rozmiary elementów, możemy wprowadzić czysto formalne definicje  $h_k, h_l$  wynikające z liczby podziałów zastosowanych do generacji danego elementu. Załóżmy że odwzorowanie generujące element siatki pierwotnej  $K$ ,  $K = F(\hat{K})$ , jest homeomorfizmem spełniającym warunki (4.1.4). Niech  $L$  będzie elementem wynikłym z  $n_1$  poziomych i  $n_2$  pionowych podziałów elementu  $K$ . Wówczas  $L$  jest obrazem prostokąta  $\hat{L} \subset \hat{K}$ ,  $L = F(\hat{L})$ , o rozmiarach odpowiednio  $2^{-n_1}$  i  $2^{-n_2}$ . Oszacowanie może być zapisane następująco:

$$\|u - u_h\|_{0,L}^2 = \int_{\hat{L}} (\hat{u} - \hat{u}_h)^2 \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} d\xi \leq CH^2 \int_{\hat{L}} (\hat{u} - \hat{u}_h)^2 d\xi \quad (4.3.35)$$

gdzie zastosowaliśmy (4.1.4)<sub>1</sub>. Ostatnie wyrażenie możemy oszacować stosując (4.3.34) z  $K \equiv \hat{L}$ ,  $h_k = 2^{-n_1}$ ,  $h_l = 2^{-n_2}$  otrzymując:

$$\begin{aligned} & \|u - u_h\|_{0,L}^2 \leq \quad (4.3.36) \\ & CH^2 \left[ 2^{-2(p+1)n_1} \int_{\hat{L}} \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi_1^{p+1}} \right)^2 d\xi + 2^{-2(p+1)n_2} \int_{\hat{L}} \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi_2^{p+1}} \right)^2 d\xi \right] \leq CH^2 \times \\ & \times \left[ 2^{-2(p-1)n_1} \int_{\hat{L}} \left( \frac{\partial^{p-1} \hat{u}}{\partial \xi_1^{p-1}} \right)^2 \frac{\partial(\xi_1, \xi_2)}{\partial(x_1, x_2)} dx + 2^{-2(p-1)n_2} \int_{\hat{L}} \left( \frac{\partial^{p-1} \hat{u}}{\partial \xi_2^{p-1}} \right)^2 \frac{\partial(\xi_1, \xi_2)}{\partial(x_1, x_2)} dx \right] \end{aligned}$$

Stosując (4.1.4)<sub>2</sub> oraz wprowadzając formalne oznaczenia  $h_k = 2^{-n_1}H$ ,  $h_l = 2^{-n_2}H$  możemy zapisać oszacowanie następująco:

$$\|u - u_h\|_{0,L}^2 \leq C \left[ h_k^{2(p-1)} \int_L \left( \frac{1}{H^{p-1}} \frac{\partial^{p+1} \hat{u}}{\partial \xi_1^{p+1}} \right)^2 dx + h_l^{2(p+1)} \int_L \left( \frac{1}{H^{p+1}} \frac{\partial^{p+1} \hat{u}}{\partial \xi_2^{p+1}} \right)^2 dx \right] \quad (4.3.37)$$

gdzie pochodne są obliczane względem  $\xi_1, \xi_2$  odpowiednio wzdłuż krzywych  $F(\cdot, \xi_2)$  i  $F(\xi_1, \cdot)$ .

W przypadku półnormy  $H^1$  otrzymanie oszacowania anizotropowego jest bardziej złożone. Poniżej prezentujemy wyprowadzenie takiego oszacowania zaproponowane w [87].

W pierwszym kroku rozważamy błąd interpolacji na elemencie prostokątnym. Chcemy oszacować wielkość:

$$|e|_{1,K} := \int_K \left[ \left( \frac{\partial e}{\partial x} \right)^2 + \left( \frac{\partial e}{\partial y} \right)^2 \right] dx \quad (4.3.38)$$

gdzie  $e = u - u_h$  jest różnicą między daną funkcją  $u$  i jej biliniowym interpolantem  $u_h$  na prostokątnym elemencie  $K$  o bokach  $h_\xi$  i  $h_\eta$ . Element  $K$  jest obrazem elementu wzorcowego  $\hat{K} = [0, 1]^2$  w odwzorowaniu:

$$x = h_\xi \xi, \quad y = h_\eta \eta \quad \text{dla } (\xi, \eta) \in K \quad (4.3.39)$$

Poszukiwany błąd możemy wyrazić na elemencie wzorcowym następująco:

$$|e|_{1,K}^2 = \int_{\hat{K}} \left[ \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 \frac{1}{h_\xi^2} + \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 \frac{1}{h_\eta^2} \right] h_\xi h_\eta d\xi d\eta = \int_{\hat{K}} \left[ \frac{h_\eta}{h_\xi} \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 + \frac{h_\xi}{h_\eta} \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 \right] d\xi d\eta, \quad (4.3.40)$$

gdzie  $\hat{e}$  oznacza  $e$  jako funkcję  $\xi$  i  $\eta$ . Ograniczmy się do pierwszego składnika w powyższej całości:

$$\int_{\hat{K}} \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 d\xi d\eta = \int_0^1 d\eta \int_0^1 \left\{ \frac{\partial}{\partial \xi} [(\hat{u} - \hat{u}_h^*) + (\hat{u}_h^* - \hat{u}_h)] \right\}^2 d\xi \leq 2 \int_0^1 d\eta \int_0^1 \left[ \frac{\partial}{\partial \xi} (\hat{u} - \hat{u}_h^*) \right]^2 d\xi + 2 \int_0^1 d\eta \int_0^1 \left[ \frac{\partial}{\partial \xi} (\hat{u}_h^* - \hat{u}_h) \right]^2 d\xi \quad (4.3.41)$$

gdzie  $\hat{u}_h^*$  jest funkcją argumentów  $\xi$  i  $\eta$ , taką że dla ustalonego  $\eta$  jest ona jednowymiarowym interpolantem stopnia  $p$  funkcji  $\hat{u}$

$$\hat{u}_h^*(\xi, \eta) = \sum_{i=0}^p w_i(\xi) \hat{u}(\xi_i, \eta) \quad (4.3.42)$$

W powyższym  $\xi_i = i/p$ ,  $i = 0, \dots, p$ , zaś  $w_i(\xi)$  są wielomianami Lagrange'a stopnia  $p$ ,  $w_i(\xi_j) = \delta_{ij}$ ,  $i, j = 0, \dots, p$ , jeśli  $\delta_{ij}$  oznacza deltę Kroneckera. Stosując standardowe jednowymiarowe oszacowanie interpolacyjne różnicę  $\hat{u} - \hat{u}_h^*$  możemy oszacować następująco:

$$\int_0^1 d\eta \int_0^1 \left[ \frac{\partial}{\partial \xi} (\hat{u} - \hat{u}_h^*) \right]^2 d\xi \leq C \int_0^1 d\eta \int_0^1 \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi^{p+1}} \right)^2 d\xi \quad (4.3.43)$$

Rozważmy z kolei drugą z całek w (4.3.41) z  $\hat{u}_h^*$  i  $\hat{u}_h$ . Ponieważ  $\hat{u}_{h,\xi}$  jest wielomianem stopnia  $p$  w  $\eta$  oraz pokrywa się z  $\hat{u}_{h,\xi}^*$  w  $\eta = i/p$ ,  $i = 0, \dots, p$  (obydwie funkcje  $\hat{u}_h|_{\eta=i/p}$  i  $\hat{u}_h^*|_{\eta=i/p}$  są lagranżowskimi interpolantami rzędu  $p$  funkcji  $\hat{u}|_{\eta=i/p}$ ,  $i = 0, \dots, p$ , muszą się więc pokrywać), znów możemy zastosować klasyczne jednowymiarowe oszacowanie tym razem w normie  $L^2$ :

$$\int_0^1 [(\hat{u}_h^* - \hat{u}_h)_{,\xi}]^2 d\eta \leq C_{2i} \int_0^1 \left[ \frac{\partial^i}{\partial \eta^i} \hat{u}_{h,\xi}^* \right]^2 d\eta, \quad i = p, p+1 \quad (4.3.44)$$

W powyższym wzięliśmy pod uwagę możliwość, że pochodne mieszane funkcji  $u$  rzędu  $p+2$  mogłyby nie być całkowne w kwadracie i wówczas jedynie oszacowanie z  $i = p$  mogłoby mieć zastosowanie.

Na tym etapie chcielibyśmy zastąpić pochodne  $\hat{u}_h^*$  w (4.3.44) przez pochodne oryginalnej funkcji  $\hat{u}$ . W tym celu rozważmy dwie funkcje pomocnicze:

$$f(\xi) := \frac{\partial^i \hat{u}}{\partial \eta^i} |_{\eta=\eta_0} \quad \text{oraz} \quad g(\xi) := \frac{\partial^i \hat{u}_h^*}{\partial \eta^i} |_{\eta=\eta_0} \quad (4.3.45)$$

dla pewnej ustalonej wartości  $\eta_0$ , oraz  $i = p, p+1$ . Spełniają one następujące warunki:

- $f = g$  dla  $\xi = i/p$ ,  $i = 0, \dots, p$  ponieważ  $\hat{u}_h^* = \hat{u}$  dla  $\xi = i/p$ ,  $\eta \in [0, 1]$ ,
- $g$  jest wielomianem stopnia  $p$  zmiennej  $\xi$ , ponieważ  $\hat{u}_h^*$  jest takim wielomianem.

Rozważmy najpierw przypadek  $p = 1$ . Ponieważ wśród funkcji o ustalonych wartościach w 0 i 1 funkcje liniowe minimalizują całkę z kwadratu pierwszej pochodnej, możemy napisać, że

$$\int_0^1 \left( \frac{dg}{d\xi} \right)^2 d\xi \leq \int_0^1 \left( \frac{df}{d\xi} \right)^2 d\xi \quad (4.3.46)$$

to znaczy, że

$$\int_0^1 \left[ \frac{\partial}{\partial \xi} \left( \frac{\partial^i \hat{u}_h^*}{\partial \eta^i} \right) \right]^2 d\xi \leq \int_0^1 \left[ \frac{\partial}{\partial \xi} \left( \frac{\partial^i \hat{u}}{\partial \eta^i} \right) \right]^2 d\xi \quad (4.3.47)$$

z czego wynika poszukiwane oszacowanie pochodnych  $\hat{u}_h^*$  przez pochodne  $\hat{u}$ . Podobne rozumowanie możemy zastosować dla przypadku  $p > 1$ . W tym celu musimy się jednak oprzeć na następującym lemacie:

**Lemat 4.3.1** *Istnieje stała  $C > 0$  zależna tylko od  $p$ , taka że dla każdego wielomianu stopnia  $p$ ,  $g: [0, 1] \rightarrow \mathbb{R}$ , i dla kawałkami liniowej funkcji  $g_i$ , takiej że  $g(\xi_i) = g_i(\xi_i)$ ,  $\xi_i = i/p$ ,  $i = 0, \dots, p$  zachodzi:*

$$\int_0^1 \left( \frac{dg}{d\xi} \right)^2 d\xi \leq C \int_0^1 \left( \frac{dg_i}{d\xi} \right)^2 d\xi \quad (4.3.48)$$

Dowód. Funkcja  $h(a_0, \dots, a_p) := \int_0^1 (g')^2 d\xi / \int_0^1 (g_i')^2 d\xi$ , gdzie  $a_i := g(i/p)$ , jest jednorodna stopnia 0. Wystarczy ją więc rozpatrywać wyłącznie na sferze  $\sum_{i=0}^p a_i^2 = 1$ , gdzie jest ona ciągła a więc osiąga swe maksimum (sfera jest zwarta). Dowodzi to istnienia stałej  $C$  spełniającej (4.3.48).  $\square$

Stosując nierówność (4.3.48) oraz zauważając, że  $g_i$  jest liniowym interpolantem funkcji  $f$  w każdym z przedziałów  $[i/p, (i+1)/p]$ ,  $i = 0, \dots, p$ , a więc że (4.3.46) stosuje się do każdego z nich, możemy napisać, że

$$\int_{\eta_{i-1}}^{\eta_i} \left( \frac{dg}{d\xi} \right)^2 d\xi \leq C \int_{\eta_{i-1}}^{\eta_i} \left( \frac{dg_i}{d\xi} \right)^2 d\xi \leq C \int_0^1 \left( \frac{df}{d\xi} \right)^2 d\xi \quad (4.3.49)$$

czyli, wracając do definicji  $f$  i  $g$ , że

$$\int_0^1 \left[ \frac{\partial}{\partial \xi} \left( \frac{\partial^i \hat{u}_h}{\partial \eta^i} \right) \right]^2 d\xi \leq C \int_0^1 \left[ \frac{\partial}{\partial \xi} \left( \frac{\partial^i \hat{u}_h}{\partial \eta^i} \right) \right]^2 d\xi, \quad i = p, p+1 \quad (4.3.50)$$

Ostatecznie, wykorzystując (4.3.50), oszacowanie (4.3.44) możemy zapisać jako

$$\int_0^1 d\eta \int_0^1 \left[ \frac{\partial}{\partial \xi} (\hat{u}_h^* - \hat{u}_h) \right]^2 d\xi \leq CC_{2i} \int_K \left( \frac{\partial^{i+1} \hat{u}}{\partial \eta^i \partial \xi} \right)^2 d\xi d\eta \quad (4.3.51)$$

zaś oszacowanie dla  $\partial \hat{e} / \partial \xi$  zgodnie z (4.3.41) jest sumą (4.3.43) i powyższego

$$\int_K \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 d\xi d\eta \leq 2C_1 \int_K \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi^{p+1}} \right)^2 d\xi d\eta + 2CC_{2i} \int_K \left( \frac{\partial^{i+1} \hat{u}}{\partial \eta^i \partial \xi} \right)^2 d\xi d\eta. \quad (4.3.52)$$

W sposób zupełnie analogiczny możemy otrzymać oszacowanie dla  $\partial \hat{e} / \partial \eta$ . Łącząc te dwa składniki wzoru (4.3.40) znajdujemy, że

$$\begin{aligned} |e|_{1,K}^2 &= \int_K \left[ \frac{h_\eta}{h_\xi} \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 + \frac{h_\xi}{h_\eta} \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 \right] d\xi d\eta \leq \\ &2C_1 \left[ h_\xi^{2p} \int_K \left( \frac{\partial^{p+1} u}{\partial x^{p+1}} \right)^2 dx dy + h_\eta^{2p} \int_K \left( \frac{\partial^{p+1} u}{\partial y^{p+1}} \right)^2 dx dy \right] + \\ &2CC_{2i} \left[ h_\eta^{2i} \int_K \left( \frac{\partial^{i+1} u}{\partial y^i \partial x} \right)^2 dx dy + h_\xi^{2i} \int_K \left( \frac{\partial^{i+1} u}{\partial x^i \partial y} \right)^2 dx dy \right], \quad i = p, p+1 \end{aligned} \quad (4.3.53)$$

### Elementy równoległoboczne.

Argumenty, których użyto do uogólnienia oszacowania anizotropowego w normie  $L^2$  na elementy krzywoliniowe nie dadzą się niestety zastosować w przypadku półnormy  $H^1$ . Przyczyną jest różny sposób transformowania się wielkości  $\partial e / \partial x_1$  i  $\partial e / \partial x_2$  przy przejściu na element wzorowy. Tym niemniej możliwe jest otrzymanie oszacowania dla elementów równoległobocznych.

Rozważmy element  $K$  jako obraz elementu  $\hat{K} = [-h_\xi/2, h_\xi/2] \times [-h_\eta/2, h_\eta/2]$  w odwzorowaniu:

$$x = \xi + a\eta, \quad y = \eta \quad (4.3.54)$$

Błąd interpolacji na  $K$  może być wyrażony następująco:

$$\begin{aligned} |e|_{1,K}^2 &= \int_K \left[ \left( \frac{\partial e}{\partial x} \right)^2 + \left( \frac{\partial e}{\partial y} \right)^2 \right] dx dy = \int_{\hat{K}} \left[ \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 + \left( -a \frac{\partial \hat{e}}{\partial \xi} + \frac{\partial \hat{e}}{\partial \eta} \right)^2 \right] d\xi d\eta \\ &\leq \int_{\hat{K}} \left[ (1+a^2) \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 + \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 \right] d\xi d\eta + \\ &2|a| \left( \int_{\hat{K}} \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 d\xi d\eta \right)^{1/2} \left( \int_{\hat{K}} \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 d\xi d\eta \right)^{1/2} \end{aligned} \quad (4.3.55)$$

gdzie w ostatnim kroku zastosowaliśmy nierówność Cauchy'ego-Schwarza. Aby oszacować  $\partial \hat{e} / \partial \xi$  oraz  $\partial \hat{e} / \partial \eta$  zastosujemy wcześniej otrzymany wynik dla elementu prostokątnego:

$$\int_{\hat{K}} \left( \frac{\partial \hat{e}}{\partial \xi} \right)^2 d\xi d\eta \leq 2C_1 h_\xi^{2p} \int_K \left( \frac{\partial^{p+1} \hat{u}}{\partial \xi^{p+1}} \right)^2 d\xi d\eta + 2CC_{2i} h_\eta^{2i} \int_K \left( \frac{\partial^{i+1} \hat{u}}{\partial \eta^i \partial \xi} \right)^2 d\xi d\eta \quad (4.3.56)$$

$$\int_{\hat{K}} \left( \frac{\partial \hat{e}}{\partial \eta} \right)^2 d\xi d\eta \leq 2C_1 h_\eta^{2p} \int_K \left( \frac{\partial^{p+1} \hat{u}}{\partial \eta^{p+1}} \right)^2 d\xi d\eta + 2CC_{2i} h_\xi^{2i} \int_K \left( \frac{\partial^{i+1} \hat{u}}{\partial \xi^i \partial \eta} \right)^2 d\xi d\eta, \quad (4.3.57)$$

Następnie zamieniamy obszar całkowania z  $\hat{K}$  na  $K$  biorąc pod uwagę, że

$$\frac{\partial \hat{u}}{\partial \xi} = \frac{\partial u}{\partial x}, \quad \frac{\partial \hat{u}}{\partial \eta} = a \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = \sqrt{1+a^2} \frac{\partial u}{\partial l} \quad (4.3.58)$$

gdzie  $l = (0, 1)/\sqrt{a^2+1}$  jest wektorem równoległym do nachylonego boku  $K$ . Oznaczmy też przez  $h_l$  długość tego boku,  $h_l = \sqrt{1+a^2} h_\eta$  oraz, dla zgodności oznaczeń,  $h_k = h_\xi$ ,  $k = (1, 0)$ , długość i wektor pozostałego boku. Przy tych oznaczeniach i wzięwszy pod uwagę, że  $a = \cot \alpha$  możemy na podstawie (4.3.55)-(4.3.57) napisać poszukiwane oszacowanie w następującej postaci:

$$|e|_{1,K}^2 \leq \frac{1}{\sin^2 \alpha} (A + B + 2|\cot \alpha| \sqrt{AB}) \quad (4.3.59)$$

gdzie  $A$  i  $B$  oznaczają odpowiednio prawe strony (4.3.56) i (4.3.57). Ostatni wzór może być dodatkowo uproszczony przez zastosowanie nierówności  $2\sqrt{AB} \leq A + B$  do postaci

$$|e|_{1,K}^2 \leq \frac{1 + |\cot \alpha|}{\sin^2 \alpha} (A + B) \quad (4.3.60)$$

choć oznacza to pewną utratę ostrości oszacowania poza przypadkiem  $A = B$ .

# Rozdział 5

## Strategia anizotropowej adaptacji typu $h$

### 5.1 Ogólna idea adaptacji w problemach przepływów

Podanie ogólnego przepisu na konstruowanie siatek adaptacyjnych, które pozwalałyby otrzymać możliwie najlepsze rozwiązanie danego problemu przy ustalonym koszcie obliczeń jest zadaniem niezwykle złożonym. Przyczyną tego jest zarówno duża liczba czynników, jakie należałoby wziąć pod uwagę, jak i trudność przeprowadzenia takiej optymalizacji, nie wspominając o jej koszcie, który mógłby przewyższać koszt właściwych obliczeń. Dlatego też problemu tego nie formułuje się na ogół w kategoriach ściśle matematycznych, akceptując raczej pewne kryteria adaptacji siatek oparte na bardzo uproszczonych argumentach. O słuszności ich wyboru decyduje doświadczenie numeryczne.

Najbardziej popularna strategia adaptacji polega na redukcji rozmiaru elementów w obszarach, w których pewien wskaźnik błędu jest największy. Technika ta opiera się na intuicyjnie akceptowalnym warunku, że aby otrzymać dokładne rozwiązanie, należy eliminować elementy z dużym błędem. Ilościowe uzasadnienie tej strategii adaptacji opiera się na założeniu, że błąd związany z danym elementem zależy od jego rozmiaru jak  $f(x)h^a$ , gdzie  $f(x)$  jest pewną funkcją zależną od rozwiązania,  $h$  rozmiarem elementu,  $a = \text{const}$  i że chcemy zminimalizować sumę takich elementowych kontrybucji przy założeniu stałej liczby elementów [34]. Mimo iż wspomniany lokalny związek błędu i rozmiaru elementu jest słuszny praktycznie jedynie dla problemu interpolacji, strategia dzielenia elementów z dużym błędem a więc prowadząca do ekwidystribucji błędów jest powszechnie stosowana.

Przy konstruowaniu strategii anizotropowej adaptacji typu  $h$  dla problemu naddźwiękowych przepływów lepkich chcielibyśmy uwzględnić następujące założenia:

1. Strategia ma prowadzić do redukcji residuum rozwiązania odpowiadającego zregularyzowanym równaniom Naviera–Stokesa a jednocześnie do redukcji wpływu wyrazów stabilizujących na rozwiązanie.
2. Powinniśmy mieć możliwość decydowania nie tylko, które elementy zostaną podzielone lecz także, czy będzie to podział kierunkowy czy izotropowy.
3. Metoda adaptacyjna powinna dostarczać nam możliwości kontrolowania dokładności obciążeń aerodynamicznych (strumienia ciepła i tarcia powierzchniowego) w warstwach przyściennych.

#### 4. Procedura adaptacyjna powinna umożliwiać badanie zbieżności rozwiązań.

Powyższe zasady mają służyć otrzymaniu wiarygodnego rozwiązania możliwie małym kosztem. Konstrukcja strategii spełniającej warunki 1-4 napotyka jednak kilka pesymistycznych faktów wskazujących na ograniczenia, które musimy brać pod uwagę. Po pierwsze, redukcja residuum poniżej dowolnego progu jest praktycznie nierealistyczna w obszarze fal uderzeniowych, których rzeczywista grubość jest rzędu  $10^{-5}$  cm. Konieczność zaakceptowania tego matematycznie nieakceptowalnego ograniczenia („o jakiej zbieżności rozwiązania może być mowa, jeśli residuum nie zbiega do zera!”), ułatwia doświadczalnie potwierdzony fakt, iż najbardziej interesujące z inżynierskiego punktu widzenia charakterystyki rozwiązania otrzymuje się z bardzo dobrą dokładnością, mimo że aproksymowane fale uderzeniowe są o rzędy wielkości grubsze od rzeczywistych. Po drugie, znane residualne wskaźniki błędów (a takich jako najbardziej wiarygodnych chcielibyśmy użyć) nie dostarczają informacji dotyczącej zależności błędów od kierunkowych rozmiarów elementu, nie dają tym samym podstaw do adaptacji anizotropowej. W końcu ostatni argument „na nie”, oszacowania dokładności wielkości punktowych, jak strumienie lepkie na brzegu, nie są znane dla większości problemów, włączając sformułowanie skończenie elementowe dynamiki gazów (wyjątkiem są oszacowania przez matematyczny post-processing dla niektórych zagadnień eliptycznych, zob. [9]). Jednak mimo wymienionych wyżej trudności teoretycznych należy stwierdzić, że doświadczenie obliczeniowe wskazuje, iż zastosowanie siatek adaptacyjnych prowadzi do bardzo istotnego ulepszenia rozwiązań względnie małym kosztem, nawet jeśli użyta strategia adaptacji nie ma rygorystycznych podstaw matematycznych. Dajemy tu tym samym wyraz przekonaniu, iż lepiej jest stworzyć możliwość rozwiązania istotnie trudnego problemu stosując adaptację choćby częściowo *ad hoc*, niż nie podejmować tego zadania w ogóle.

Ogólny zarys strategii adaptacji dla zagadnień przepływów, którą chcemy zaproponować, jest następujący:

#### 1. Proces adaptacyjnego rozwiązywania problemu rozbijamy na dwa etapy:

- i) Adaptację siatki w całym obszarze obliczeniowym. Celem tej części procesu jest osiągnięcie akceptowalnej dokładności wartości rozwiązania w całym obszarze.
- ii) Adaptację ograniczoną wyłącznie do warstw przyściennych. Jej celem jest otrzymanie możliwie dokładnych wartości strumieni lepkich.

Powyższy podział wynika stąd, że stosowanie adaptacji dla całej siatki do momentu osiągnięcia dokładnych strumieni w warstwach przyściennych wiązałoby się z mimowolnym zagęszczaniem elementów wzdłuż fal uderzeniowych – algorytm dążyłby do dokładnej ich aproksymacji. Byłoby to podejście niezwykle kosztowne zważywszy na małą grubość rzeczywistych fal uderzeniowych.

Powrót do kroku i) po wykonaniu etapu ii) jest na ogół niemożliwy, ponieważ procedura ii) wprowadza niezwykle małe elementy, których obecność powoduje konieczność redukcji kroku czasowego  $\Delta t$  a to z kolei sprawia, że jakkolwiek dalsza ewolucja rozwiązania w całym obszarze staje się nie do zaakceptowania kosztowna.

2. Jedną z możliwych dróg adaptacji we wnętrzu obszaru  $\Omega$  jest oparcie jej o kryterium minimalizacji błędów interpolacji przy ustalonej liczbie elementów, tj. adaptowanie siatki do rozwiązania z punktu widzenia możliwości optymalnej jego interpolacji. Zaletą tego podejścia jest to, że wskazuje ono na optymalny kierunek podziału elementu (por. rozdziały 1.4 i 5.2). Istotną wadą jest zaś fakt, że decyzje dotyczące

podziałów są podejmowane na podstawie oszacowania drugich pochodnych rozwiązania, które jest znacznie mniej wiarygodną wielkością niż bez wątpienia rzeczywisty wskaźnik błędu: residuum. Przedstawione argumenty są podstawą następującej filozofii adaptacji siatki: decyzja, które elementy dzielić, jest podejmowana na podstawie *residualnych* wskaźników błędu – dzielone są elementy z największymi wskaźnikami; kierunek podziału ustalany jest zaś na podstawie optymalizacji siatki pod względem interpolacji. Podstawą dokonywania adaptacji jest więc redukcja residuum rozwiązania.

- Ocenę dokładności strumieni lepkich na ścianie opływającego obiektu oprzemy na punktowym oszacowaniu błędu interpolacji dla pochodnych rozwiązania. Proces adaptacji w warstwie przysciennej polegać będzie na podziale elementów przyległych do brzegu i uaktualnianiu rozwiązania na otrzymanych siatkach dopóki wartość wskaźnika błędu nie zostanie zredukowana poniżej danego progu. Procedurę tę możemy uzupełnić o dodatkowe podziały pozostałych elementów warstwy przysciennej dla zredukowania sztucznej dyssypacji i wskaźników błędu w całej warstwie.
- Zbieżność rozwiązań można badać poprzez obliczanie normy różnicy rozwiązań na kolejnych siatkach.

Poniżej przedstawiamy szczegóły realizacji wymienionych składników strategii adaptacji.

## 5.2 Kryteria optymalności dla anizotropowych siatek adaptacyjnych

W rozdziale 1.4 rozważaliśmy warunki optymalności dla siatek adaptacyjnych ograniczając się do siatek prostokątnych. Obecnie pokażemy, że otrzymane tam kryteria optymalności można uogólnić dla elementów równoległobocznych i krzywoliniowych, znajdziemy też optymalną wartość wydłużenia elementów.

Zagadnienie minimalizacji błędu interpolacji z ograniczeniem na liczbę elementów można dla wszystkich wymienionych wyżej przypadków sformułować następująco:

Znaleźć  $h_k(\mathbf{x})$ ,  $h_l(\mathbf{x})$  takie że:

$$e^2 := \int_{\Omega} [h_k^{\sigma} f_1(\mathbf{x}) + h_l^{\sigma} f_2(\mathbf{x})] dx = \min$$

$$\int_{\Omega} \frac{g(\mathbf{x}) dx}{h_k h_l} = N \quad (5.2.1)$$

Znaczenie  $\sigma$ ,  $f_1$ ,  $f_2$ ,  $g$  dla poszczególnych przypadków jest następujące:

i) Elementy równoległoboczne, błąd w półnormie  $H^1$ :

$$\sigma = 2p, \quad f_1 = \left( \frac{\partial^{p+1} u}{\partial k^{p+1}} \right)^2 \frac{1 + \cot \alpha}{\sin^2 \alpha}, \quad f_2 = \left( \frac{\partial^{p+1} u}{\partial l^{p+1}} \right)^2 \frac{1 + \cot \alpha}{\sin^2 \alpha}, \quad g = \frac{1}{\sin \alpha} \quad (5.2.2)$$

(gdzie zaniedbano kontrybucję wyrazów wyższego rzędu w oszacowaniu (4.3.60) z  $i = p + 1$ ).

ii) Elementy równoległoboczne, błąd w normie  $L^2$  (por. (4.3.34)):

$$\sigma = 2(p + 1), \quad f_1 = \left( \frac{\partial^{p+1} u}{\partial k^{p+1}} \right)^2 \frac{1}{\sin \alpha}, \quad f_2 = \left( \frac{\partial^{p+1} u}{\partial l^{p+1}} \right)^2 \frac{1}{\sin \alpha}, \quad g = \frac{1}{\sin \alpha} \quad (5.2.3)$$



iii) Elementy zakrzywione, błąd w normie  $L^2$  (por. (4.3.37)):

$$\sigma = 2(p+1), \quad f_1 = \left( \frac{1}{H^2} \frac{\partial^{p+1} \bar{u}}{\partial \xi^{p+1}} \right)^2, \quad f_2 = \left( \frac{1}{H^2} \frac{\partial^{p+1} \bar{u}}{\partial \eta^{p+1}} \right)^2, \quad g = H^2 \frac{\partial(\xi, \eta)}{\partial(x, y)} \quad (5.2.4)$$

Funkcje  $g(\mathbf{x})$  zostały przyjęte w taki sposób, aby  $\int_K g/(h_k h_l) dx = 1$ . (Postać  $g = 1/\sin \alpha$  jest oczywista dla równoległoboków. W przypadku elementów krzywoliniowych mamy:

$$H^2 \int_K \frac{\partial(\xi, \eta)}{\partial(x, y)} \frac{1}{h_k h_l} dx = H^2 \int_{F^{-1}(K)} \frac{\partial(\xi, \eta)}{\partial(x, y)} \frac{1}{H^2 2^{-n_1 - n_2}} \frac{\partial(x, y)}{\partial(\xi, \eta)} d\xi = 1 \quad (5.2.5)$$

gdzie wzięliśmy pod uwagę, że  $F^{-1}(K)$  jest prostokątem o wymiarach  $2^{-n_1}$  i  $2^{-n_2}$ .)

Rozwiązanie zadania optymalizacji (5.2.1) metodą mnożników Lagrange'a sprowadza się do przyrównania do zera wariacji funkcjonału:

$$L = \int_{\Omega} [h_k^\sigma f_1(\mathbf{x}) + h_l^\sigma f_2(\mathbf{x})] dx + \lambda \int_{\Omega} \frac{g(\mathbf{x}) dx}{h_k h_l} \quad (5.2.6)$$

to znaczy

$$\int_{\Omega} \left[ \sigma h_k^{\sigma-1} f_1(\mathbf{x}) \delta h_k + \sigma h_l^{\sigma-1} f_2(\mathbf{x}) \delta h_l - \lambda \frac{g(\mathbf{x})}{h_k^2 h_l} \delta h_k - \lambda \frac{g(\mathbf{x})}{h_k h_l^2} \delta h_l \right] dx = 0 \quad (5.2.7)$$

czyli

$$h_k^\sigma f_1 \frac{h_k h_l}{g} = \frac{\lambda}{\sigma} \quad \text{i} \quad h_l^\sigma f_2 \frac{h_k h_l}{g} = \frac{\lambda}{\sigma} \quad (5.2.8)$$

Dla dostatecznie małych elementów (asymptotycznie) można przyjąć, że  $h_k h_l / g \sim \text{meas}(K)$  i zinterpretować warunki (5.2.8) jako

$$\eta_{k,K}^2 := h_k^\sigma \int_K f_1(\mathbf{x}) dx = \lambda / \sigma, \quad \eta_{l,K}^2 := h_l^\sigma \int_K f_2(\mathbf{x}) dx = \lambda / \sigma \quad (5.2.9)$$

co odpowiada zasadzie ekwidystribucji kierunkowych wskaźników błędu  $\eta_{k,K}, \eta_{l,K}$ . Warunki (5.2.9) implikują strategię adaptacji, która polega na obliczeniu  $\eta_{k,K}, \eta_{l,K}$  dla wszystkich elementów siatki a następnie na dzieleniu tych spośród nich, dla których wskaźniki błędu przekraczają pewien próg, powiedzmy  $\beta \max_K(\eta_{k,K}, \eta_{l,K})$ ,  $0 < \beta < 1$ , gdyż taki proces prowadzi do przybliżonej ekwidystribucji.

Otrzymane kryterium optymalności pociąga za sobą w sposób oczywisty optymalną wartość stosunku  $\eta_{k,K}/\eta_{l,K}$  jako 1. W praktyce osiągnięcie takiej wartości na ogół nie jest możliwe, gdyż wymiarów elementów nie możemy zmieniać w sposób ciągły: mogą one przyjmować  $2^{-n}$ -tą część rozmiaru elementu początkowego,  $n = 0, 1, 2, \dots$ . Intuicja sugeruje, że najlepsza proporcja  $h_k/h_l$  powinna zostać wybrana tak, aby  $\eta_{k,K}/\eta_{l,K}$  było możliwie najbliższe 1. Optymalna wartość  $\eta_{k,K}/\eta_{l,K}$  powinna więc spełniać warunek:

$$\frac{1}{\sqrt{b}} \leq \eta_{k,K}/\eta_{l,K} \leq \sqrt{b} \quad (5.2.10)$$

gdzie  $b = 2^{p+1}$  dla błędów w normie  $L^2$ ,  $b = 2^p$  dla półnormy  $H^1$ , gdyż pojedynczy kierunkowy podział elementu zmienia  $\eta_{k,K}/\eta_{l,K}$  o czynnik  $b^{\pm 1}$ . Powyższy wniosek można też uzasadnić bardziej ścisłym rozumowaniem. Założymy, że w pewnym wystarczająco małym podobszarze pokrytym siatką  $\mathcal{T}$  z jednokierunkowymi elementami funkcje podcałkowe w

(5.2.9) są niemal stałe. Niech  $\eta_k > \eta_l$ . Załóżmy, że aby osiągnąć optymalny stosunek  $\eta_k/\eta_l$  musimy dokonać  $m$  podziałów redukujących  $h_k$   $2^m$  razy. Następnie zaś optymalne stają się podziały izotropowe, nie zmieniające  $\eta_k/\eta_l$ . Ponieważ redukcja kierunkowego wskaźnika błędu wynikająca z kierunkowych złamań wynosi  $a = b^2$ , wnioskujemy, że logarytmy błędów odpowiadającego  $m_a$  anizotropowym i  $m_i$  izotropowym podziałom są następującymi funkcjami:

$$y_a = \ln(e_1^2 a^{m_a} + e_2^2) \quad , \quad y_i = \ln(a^{m_i} \cdot c) \quad (5.2.11)$$

gdzie  $e_1^2 = s\eta_k^2$ ,  $e_2^2 = s\eta_l^2$ ,  $c$  jest stałą. Poziomy podziałów  $m_a, m_i$  odpowiadają następującym liczbom elementów:

$$N_a = c_1 2^{m_a}, \quad N_i = c_2 2^{2m_i}, \quad c_1, c_2 - \text{stałe} \quad (5.2.12)$$

Wobec tego logarytmy błędów  $y_a, y_i$  możemy wyrazić poprzez logarytmy z wielkości  $N_a$  i  $N_i$ :  $n_a = \log_2 c_1 + m_a$ ,  $n_i = \log_2 c_2 + 2m_i$  następująco:

$$y_a(n_a) = \ln(e_1^2 a^{\log_2 c_1 + n_a} + e_2^2), \quad y_i(n_i) = \ln c + \left(\frac{n_i}{2} - \frac{\log_2 c_2}{2}\right) \ln a \quad (5.2.13)$$

Liczba podziałów kierunkowych  $m$  jest optymalna, jeśli krzywa  $y_a(n)$  jest bardziej nachylona niż  $y_i(n)$  dla  $n \leq m$  i jeśli  $y_a(n)$  jest mniej stroma niż  $y_i(n)$  dla  $n \geq m$ , czyli gdy:

$$y_a(m) - y_a(m-1) \leq \frac{1}{2} \ln a \leq y_a(m+1) - y_a(m) \quad (5.2.14)$$

Oznaczając przez  $\beta = e_1^2 a^{\log_2 c_1 - m} / e_2^2$  optymalny stosunek  $\eta_k^2/\eta_l^2$ , możemy przepisać (5.2.14) następująco:

$$\frac{\beta + 1}{\beta/a + 1} \leq a^{1/2} \leq \frac{\beta a + 1}{\beta + 1} \quad (5.2.15)$$

co jest równoważne (5.2.10).

### 5.3 Strategia adaptacji

Zgodnie z przedstawioną wcześniej argumentacją w pierwszym etapie adaptacyjnego rozwiązywania zadania przepływu naszym celem jest poprawa dokładności wartości rozwiązania w całym obszarze obliczeniowym. Procedura ta polega na kilkukrotnych adaptacjach siatki skończenie elementowej z następującymi po nich rozwiązaniami zadania na aktualnej siatce. Wybór elementów do podziału jest dokonywany na podstawie wartości residualnych wskaźników błędu  $e_K$ , podczas gdy kierunek podziału elementu jest podyktowany optymalizacją błędu interpolacji.

W tym miejscu powinniśmy zdecydować jaka jest najodpowiedniejsza norma dla pomiaru błędu interpolacji w zadaniach przepływu. Często rozważanym podejściem w podobnych przypadkach jest odwołanie się do oszacowania *a priori* dla danego zagadnienia, tj. oszacowania błędu aproksymacji rozwiązania przez błąd interpolacji w pewnych normach. Przykładowo, dla liniowych zadań eliptycznych drugiego rzędu oszacowania *a priori* otrzymujemy z lematu Cea:

$$\|u - u_h\|_{1,\Omega} \leq C \|u - a_I\|_{1,\Omega} \quad (5.3.16)$$

gdzie  $u_h$  jest rozwiązaniem skończenie elementowym,  $u_I$  interpolantem rozwiązania ścisłego  $u$ . Wychodząc z oszacowania takiego rodzaju możemy stwierdzić, że redukując błąd interpolacji w normie wskazywanej przez oszacowanie *a priori* redukujemy także błąd aproksymacji (choć niekoniecznie w sposób optymalny), tj. spełniamy warunek konieczny efektywnej adaptacji.

Niestety, teoria matematyczna problemów przepływów ściśliwych a także ich aproksymacji jest niepełna i oszacowaniami *a priori* nie dysponujemy. Dlatego też przy wyborze normy, dla której siatka powinna optymalnie interpolować rozwiązanie, oprzemy się na przesłankach mniej precyzyjnych, po części intuicyjnych, wynikających z pewnych cech rozwiązań.

Dla zaskakującej klasy problemów przepływów naddźwiękowych z małą lepkością efekty z nią związane są obserwowane w pobliżu opływanych obiektów zaś w pozostałej części obszaru przepływ jest z dużą dokładnością nielepki („suchy”), można opisać go równaniami Eulera. Dlatego też racjonalne a przynajmniej możliwe wydaje się zastosowanie normy  $L^2$  jako poprawnej miary błędów rozwiązań równań Eulera: są to równania pierwszego rzędu – istnieje więc norma  $L^2$  ich residuum (dla aproksymacji z ciągłością  $C^0$ ) – przestrzeń  $L^2(\Omega)$  utożsamiamy z przestrzenią do niej dualną,  $L^2(\Omega) \sim (L^2(\Omega))'$ , która, zgodnie z (5.3.17) jest odpowiednia do pomiaru błędów. Z drugiej strony w obszarach warstw przyściennych, gdzie dominują efekty lepkie (decyduje o tym proporcja sił bezwładności do sił lepkości), które są związane z eliptycznym składnikiem równań Naviera–Stokesa, przez analogię do problemów eliptycznych odpowiednią normą dla miary błędów jest półnorma  $H^1$ .

Za zastosowaniem normy  $L^2$  we wnętrzu obszaru obliczeniowego przemawia dodatkowo następujący argument praktyczny związany z aproksymacją fal uderzeniowych. Rozważmy optymalizację siatki w pobliżu fali uderzeniowej. Jeśli kierunek tej fali jest ustalony oraz jeśli dana jest orientacja elementów siatki, to proporcja wymiarów elementów przecinanych przez falę uderzeniową jest optymalna, jeśli fala przecina elementy wzdłuż przekątnej. Wynika to stąd, że wówczas wymiar elementu w kierunku  $n = (n_1, n_2)$  prostopadłym do nieciągłości jest minimalny przy ustalonej liczbie elementów na jednostkę długości fali, zaś wymiar ten decyduje o ostrości przybliżania nieciągłości. Jeśli rozważymy interpolację gwałtownie zmieniającej się funkcji modelującej nieciągłość postaci  $f(x_1 n_1 + n_2 x_2)$ , wówczas jej drugie pochodne występujące w definicji kierunkowych wskaźników błędów (5.2.9) mają postać:

$$\left(\frac{\partial^2 f}{\partial k^2}\right)^2 = (k \cdot n)^2 f''^2, \quad \left(\frac{\partial^2 f}{\partial l^2}\right)^2 = (l \cdot n)^2 f''^2, \quad (5.3.17)$$

gdzie  $k$  i  $l$  są wersorami boków elementu. Wobec tego warunki (5.2.9) na optymalizację błędów interpolacji w normie  $L^2$  i  $H^1$  implikują następujące relacje wymiarów elementu:

$$h_k^4 (k \cdot n)^4 = h_l^4 (l \cdot n)^4 \quad \text{i} \quad h_k^2 (k \cdot n)^4 = h_l^2 (l \cdot n)^4 \quad (5.3.18)$$

Możemy zauważyć, że pierwszy ze wzorów odpowiada sytuacji optymalnej, tj. gdy fala uderzeniowa jest równoległa do przekątnej,  $h_k |k \cdot n| = h_l |l \cdot n|$ .

Ostatecznie więc parametr  $b$  w warunku optymalności (5.2.10) przyjmujemy jako  $b = 2^{l-1} = 4$ , tj. dla błędów interpolacji w normie  $L^2$  na elementach biliniowych,  $p = 1$ .

Podkreślimy, że dyskutowany powyżej wybór normy do pomiaru błędów interpolacji wpływa jedynie na kierunek podziałów elementów a tym samym na proporcję ich wymiarów. Nie ma natomiast wpływu na sam fakt podziału elementów – zależy on wyłącznie od wielkości *residualnych* wskaźników błędów. Tak więc adaptacja zapewnia

redukcję residuum a sposób dokonywania kierunkowych podziałów wpływa na szybkość tej redukcji.

Dodatковым czynnikiem, który należy wziąć pod uwagę w strategii kierunku podziałów, jest fakt, że rozwiązanie ma cztery składowe. Decydujemy się na relatywnie konserwatywną (ostrożną) zasadę wprowadzania elementów wydłużonych: element jest dzielony kierunkowo tylko wówczas, gdy jest to wskazywane przez wartość proporcji  $\eta_k/\eta_l$  dla wszystkich czterech składowych rozwiązania, w przypadku przeciwnym podziały są izotropowe.

W końcu ostatnim elementem procedury adaptacji jest możliwość łączenia na powrót niektórych elementów wynikłych z poprzednich podziałów. Konieczność uwzględnienia takiej procedury wiąże się z występowaniem istotnych przesunięć w rozkładach błędów towarzyszących uaktualnianiu rozwiązania.

Algorytm adaptacji oparty na przedstawionych powyżej zasadach możemy ostatecznie przedstawić następująco:

1. Na danej siatce całkujemy równania przepływu w czasie do osiągnięcia stanu ustalonego.
2. Dla wszystkich elementów  $K$  obliczamy residualne wskaźniki błędów  $e_K$  oraz kierunkowe wskaźniki błędów interpolacji  $\eta_{k,K}^i, \eta_{l,K}^i, i = 1, \dots, 4$ , gdzie dodatkowy indeks  $i$  odpowiada numeracji składowych rozwiązania. Znajdujemy  $e_{max} = \max_K e_K$ . Przerwywamy obliczenia, jeśli dokładność rozwiązania jest zadowalająca.
3. Łączymy na powrót elementy wynikłe z poprzednich podziałów, jeśli dla takich elementów  $(\sum_K e_K^2)^{1/2} \leq \alpha_1 e_{max}$ , gdzie sumowanie odbywa się po czwórkach lub parach elementów otrzymanych przez podział wspólnego „elementu-ojca”.
4. Właściwa adaptacja:  
Dzielimy elementy  $K$ , dla których  $e_K \geq \alpha_2 e_{max}$ . Sposób podziału jest wybierany następująco:
  - jeśli  $\eta_{k,K}^i/\eta_{l,K}^i \leq b^{-1/2}$  dla  $i = 1, \dots, 4$ , podział kierunkowy redukujący  $h_l$ ,
  - jeśli  $\eta_{k,K}^i/\eta_{l,K}^i \geq b^{+1/2}$  dla  $i = 1, \dots, 4$ , podział kierunkowy redukujący  $h_k$ ,
  - jeśli żaden z powyższych warunków nie jest spełniony – podział izotropowy.
5. Kontynuujemy procedurę od pkt. 1.

W powyższym  $\alpha_1, \alpha_2$  są przyjmowanymi przez użytkownika programu parametrami,  $0 < \alpha_1 < \alpha_2 < 1$ .

Oddzielnym problemem jest estymowanie drugich pochodnych rozwiązania (potrzebnych do obliczania  $\eta_{k,K}^i, \eta_{l,K}^i$ ) dla aproksymacji kawałkami dwuliniowej. Można tego dokonać stosując odpowiednie ilorazy różnicowe rozciągające się na elementy sąsiednie danego elementu. Inną możliwością jest zastosowanie matematycznego post-processingu do otrzymania ciągłych pierwszych pochodnych, na przykład poprzez rzutowanie  $L^2$  obliczonych bezpośrednio nieciągłych pochodnych  $\nabla U_h$  na przestrzeń ciągłych funkcji kształtu. Niestety obydwie z wymienionych procedur są matematycznie uzasadnione tylko dla problemów eliptycznych przybliżanych na równomiernych siatkach (zob. [90],[20],[75]), tym niemniej w praktyce są często stosowane i dają satysfakcjonujące wyniki.

## 5.4 Adaptacja w warstwie przyściennej

Niezwykle małe wartości stałych lepkości i przewodnictwa cieplnego  $\mu$  i  $\kappa$  sprawiają, że nawet jeśli wartości rozwiązania skończenie elementowego  $U$  na danej siatce są względnie dokładne, to gradienty rozwiązania w warstwie przyściennej mogą różnić się od wartości dokładnych (eksperymentalnych) o całe rzędy wielkości. Dokładna symulacja strumieni lepkich (proporcjonalnych do gradientów rozwiązania) wymaga znacznie drobniejszej siatki w warstwach przyściennych niż jest to konieczne do aproksymacji  $U$ . Często stosowana metoda otrzymania wiarygodnych strumieni polega na wygenerowaniu siatki z wprowadzoną *a priori* cienką warstwą bardzo małych elementów wzdłuż brzegu opływającego obiektu. Podejście takie wymaga względnie dokładnego oszacowania grubości warstwy przyściennej, co nie zawsze jest możliwe.

W naszym podejściu chcemy wykorzystać unikalną możliwość adaptacyjnej metody elementów skończonych do dopasowywania siatki do nieregularności rozwiązania i zastosować ją w warstwie przyściennej. Akceptujemy założenie, że rozwiązanie otrzymane w globalnym etapie adaptacji (w całym obszarze obliczeniowym) jest wystarczająco dokładne. Naszym celem jest dokonanie adaptacji siatki w warstwie przyściennej prowadzące do otrzymania zadowalająco dokładnych i wiarygodnych pochodnych rozwiązania w tym obszarze.

Stawiając problem w ten sposób chcielibyśmy sformułować kryteria, które powinny być spełnione dla zaakceptowania rozwiązania w warstwie przyściennej. Czy też, myśląc bardziej realistycznie, kryteria, których niespełnienie pozwalałoby dyskwalifikować rozwiązanie. Zakładamy, że błędy strumieni lepkich, tarcia powierzchniowego i strumienia ciepła wzdłuż ściany obiektu, są wynikiem wyłącznie niedokładności gradientów rozwiązania. Tarcie powierzchniowe i strumień ciepła wzdłuż ściany ciała wyrażają się następująco:

$$\tau = \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) n_j s_i = \mu \frac{1}{\rho} \frac{\partial m_i}{\partial n} s_i \quad (5.4.19)$$

i

$$q = \kappa \frac{\partial \theta}{\partial n} = \frac{\kappa}{C_v} \left( \frac{1}{\rho} \frac{\partial e}{\partial n} - \frac{e}{\rho^2} \frac{\partial \rho}{\partial n} \right), \quad (5.4.20)$$

gdzie  $n = (n_i)$  oznacza wektor normalny do ściany,  $s = (s_i)$  - wektor styczny. W wprowadzeniu wzorów (5.4.19) i (5.4.20) wzięto też pod uwagę, że prędkości  $u_i = m_i/\rho$ , temperatura  $\theta = [e - (m_1^2 + m_2^2)/(2\rho)]/(C_v\rho)$  oraz że  $m_i = 0$ ,  $i = 1, 2$ , wzdłuż ściany. Wzory (5.4.19) i (5.4.20) implikują, że błędy wielkości  $\tau$  i  $q$ ,  $\Delta\tau$  i  $\Delta q$ , można wyrazić poprzez błędy gradientów zmiennych konserwatywnych,  $\Delta(\partial\rho/\partial n)$ ,  $\Delta(\partial m_i/\partial n)$ ,  $\Delta(\partial e/\partial n)$  następująco:

$$\Delta\tau = \mu \frac{1}{\rho} \Delta \left( \frac{\partial m_i}{\partial n} \right) |s_i| \quad (5.4.21)$$

$$\Delta q = \frac{\kappa}{C_v} \left( \frac{1}{\rho} \Delta \left( \frac{\partial e}{\partial n} \right) + \frac{e}{\rho^2} \Delta \left( \frac{\partial \rho}{\partial n} \right) \right) \quad (5.4.22)$$

Oszacowania dokładności pochodnych  $\rho$ ,  $m_i$ ,  $e$  względem  $n$  możemy dokonać na podstawie elementarnego jednowymiarowego oszacowania interpolacyjnego dla aproksymacji liniowej:

$$\left| \frac{\partial v}{\partial n} \right| \leq \frac{h}{\Delta x} \left| \frac{\partial^2 v}{\partial x^2} \right| \quad (5.4.23)$$

wynikającego z rozwinięcia Taylora:

$$v(h) - v(0) = h \frac{\partial v}{\partial n}(0) + \frac{h^2}{2} \frac{\partial^2 v}{\partial n^2}(\xi), \text{ dla } \xi \in [0, h] \quad (5.4.24)$$

gdzie  $h$  oznacza wymiar elementu mierzony w kierunku  $n$ ,  $\partial v_h / \partial n = (v_h(h) - v_h(0)) / h$ , zaś  $v$  oznacza którąkolwiek ze składowych  $U$ ,  $v_h$  jej liniowy interpolant. Oczywiście rozwiązanie skończenie elementowe nie pokrywa się z interpolantem rozwiązania ścisłego. Tym niemniej duże wartości wskaźnika:

$$\eta_v = \frac{h}{2} \left| \frac{\partial^2 v}{\partial n^2} \right|_{\infty, [0, h]} \quad (5.4.25)$$

są argumentem przekonującym o słabej dokładności  $\partial v / \partial n$ : niemożność dokładnej interpolacji rozwiązania na danej siatce sugeruje też małą dokładność aproksymacji skończenie elementowej. Dlatego też przyjmujemy  $\eta_v$  jako wskaźnik błędów dla  $\partial v / \partial n$  z  $\partial^2 v / \partial n^2$  estymowanym przez odpowiedni iloraz różnicowy zastosowany do  $v_h$ .

Strategia adaptacji, której podstawą jest dążenie do zredukowania zdefiniowanego wyżej wskaźnika błędów dotyczy elementów przylegających do ściany obiektu. Składa się z następujących kroków:

1. Całkujemy równania przepływu na danej siatce do osiągnięcia stanu ustalonego.
2. Dla wszystkich elementów  $K$  przylegających do ściany ciała szacujemy błędy strumieni lepkich  $\Delta \tau_K, \Delta q_K$  stosując wzory (5.4.21) i (5.4.22) z

$$\Delta \left( \frac{\partial U_i}{\partial n} \right) = \frac{h}{2} \left| \frac{\partial^2 U_i}{\partial n^2} \right|_{\infty, [0, h]}, \quad i = 1, \dots, 4 \quad (5.4.26)$$

Przerywamy obliczenia, gdy osiągnięta zostanie zadana dokładność.

3. Dzielimy elementy  $K$ , dla których błąd względny strumieni lepkich przekracza dany próg:

$$\frac{\Delta \tau_K}{\tau_{ref}} \geq a \quad \text{lub} \quad \frac{\Delta q_K}{q_{ref}} \geq a \quad (5.4.27)$$

gdzie  $a$  jest zadany błąd względny (powiedzmy  $a = 0.01$ ),  $\tau_{ref}, q_{ref}$  są przyjętymi przez użytkownika tarciami i strumieniem odniesienia (np.  $\tau_{ref} = \max_K \tau_K$  etc.).

4. Kontynuujemy proces od pkt. 1.

W ogólności elementy są dzielone w kierunku równoległym do brzegu. Dopuszczamy jednak podziały izotropowe, jeśli podział kierunkowy oddalałby  $\eta_k / \eta_l$  (dla normy  $H^1$ ) od wartości optymalnej, tj. 1. Ponieważ przedstawiona strategia adaptacji dzieli wyłącznie elementy przyległe do brzegu, generuje tym samym siatkę, której wielkość oczka zmniejsza się geometrycznie o czynnik 1/2 przy zbliżaniu się do brzegu. Z tego powodu nawet gdy błędy  $\Delta \tau_K$  i  $\Delta q_K$  zostaną zredukowane poniżej pewnej małej wartości, nie będziemy mieć gwarancji, czy dokładność aproksymacji w pozostałej części warstwy brzegowej jest wystarczająca. Aby zapobiec takiej ewentualności możemy rozszerzyć przedstawioną

procedurę adaptacji o możliwość dodatkowych podziałów elementów warstwy niekoniecznie przylegających do brzegu. Podziały te miałyby na celu ekwidystribucję residualnych wskaźników błędu oraz wskaźników interpolacyjnych będących uogólnieniem (5.4.25):

$$\eta = \max \left( \frac{h_k}{2} \max_{i=1,\dots,4} \left| \frac{\partial^2 U_i}{\partial k^2} \right|, \frac{h_l}{2} \max_{i=1,\dots,4} \left| \frac{\partial^2 U_i}{\partial l^2} \right| \right) \quad (5.4.28)$$

Oczywiście strategia takich podziałów jest zupełnie analogiczna do poprzednio dyskutowanej adaptacji w całym obszarze. Tym razem jednak błędy w elementach najmniejszych przylegających do brzegu wyznaczałyby graniczne wskaźniki błędów w pozostałych elementach. Przykładowo, pkt. 3. procedury adaptacji należałoby rozszerzyć o następujący krok

3'. Dla każdego elementu  $K_b$  przyległego do brzegu rozważamy pozostałe elementy  $K$  warstwy przyściennej napotykaną przy oddalaniu się od  $K_b$  w głąb warstwy. Jeśli  $\eta_K \geq \eta_{K_b}$ , dzielimy element  $K$ .

Dodajmy że procedura taka wymaga dodatkowo zdefiniowania zasięgu warstwy przyściennej (np. przez badanie wielkości strumienia ciepła lub sił tarcia).

Wreszcie ostatnim aspektem adaptacji w warstwie przyściennej jest dążenie do redukcji parametrów regularyzujących SUPG,  $\tau_a, \nu_d$ . Chcielibyśmy zmniejszyć je poniżej rzeczywistej lepkości. I w tym wypadku postać procedury adaptacji będzie wersją opisaną powyżej z zredukowanymi parametrami  $\tau_a, \nu_d$  „w roli” błędów.

# Rozdział 6

## Iteracyjne algorytmy rozwiązywania równań

### 6.1 Wprowadzenie

Istotnym składnikiem algorytmów metody elementów skończonych do rozwiązywania problemów o dużych rozmiarach jest procedura rozwiązywania równań liniowych. Duża liczba operacji rozwiązywania równań oraz znaczny rozmiar pamięci wymagany przez programy rozwiązujące równania sprawia, iż jakość i efektywność tych algorytmów ma bardzo duży wpływ na rozmiar zagadnień, których rozwiązanie jest możliwe na danym sprzęcie komputerowym.

Wśród metod rozwiązywania równań liniowych możemy wyróżnić metody bezpośrednie (skończone), typu metody eliminacji Gaussa, oraz metody iteracyjne. Metody pierwszej grupy prowadzą do rozwiązania ścisłego (w dokładnej arytmetyce) po skończonej liczbie operacji arytmetycznych, podczas gdy metody iteracyjne generują ciąg rozwiązań zbieżnych do rozwiązania ścisłego. Istotną cechą procedur rozwiązujących jest możliwość optymalizacji obliczeń w przypadku macierzy rzadkich o znanej strukturze, w przypadku dyskretyzacji skończenie elementowej – macierzy o budowie zbliżonej do pasmowej. Wspomniana optymalność polega na możliwości pominięcia składowania zerowych wyrazów macierzy w pamięci i pominięcia operacji związanych z tymi wyrazami.

W klasie metod bezpośrednich dla macierzy rzadkich rozwinięto szereg efektywnych algorytmów ściśle związanych ze strukturą macierzy, na przykład metody pasmowe czy typu *skyline*. Algorytmy te znajdują szczególnie zastosowanie w programach komercyjnych, których użytkownicy mają większe zaufanie do rozwiązań bezpośrednich.

Wśród algorytmów iteracyjnych optymalne traktowanie macierzy rzadkich jest względnie łatwe do osiągnięcia. Podstawową operacją tych procedur jest mnożenie „macierz razy wektor”, gdzie macierz jest bądź oryginalną bądź w pewien sposób przekształconą macierzą układu równań. Dzięki temu omijanie operacji na zerach jest łatwo zrealizować. Jednak wadą procedur iteracyjnych jest duża zależność ich efektywności (rzędu zbieżności) od problemu oraz fakt, iż szybkość otrzymania dokładnego rozwiązania silnie zależy od przybliżenia początkowego.

Skuteczność metod iteracyjnych objawia się szczególnie przy rozwiązywaniu zadań niestacjonarnych i przy adaptacyjnym rozwiązywaniu problemu. Wynika to w głównej mierze stąd, że mamy wówczas bardzo dobre przybliżenie szukanego rozwiązania: jest nim rozwiązanie z poprzedniego kroku czasowego lub rozwiązanie otrzymane na poprzedniej



siatce. Doświadczenie praktyczne wskazuje, że zastosowanie metod iteracyjnych we wspomnianych sytuacjach daje możliwość otrzymywania kolejnych rozwiązań przy stałej liczbie iteracji niezależnej od rozmiaru zadania.

Macierze sztywności w adaptacyjnej metodzie elementów skończonych są pasmowe podobnie jak dla siatek równomiernych. Jednakże ze względu na dużą złożoność siatek adaptacyjnych i związaną z tym trudność optymalnego ponumerowania elementów, szerokość pasma jest w siatkach adaptacyjnych istotnie większa niż dla siatek równomiernych. Można oszacować, że szerokość pasma siatki adaptacyjnej stanowi taki sam ułamek rozmiaru pełnej macierzy jak dla równomiernej siatki wyjściowej, zaś dla równomiernego wzbogacania siatki proporcja ta spada. Ten niekorzystny aspekt aproksymacji na siatkach adaptacyjnych przestaje odgrywać rolę przy zastosowaniu procedur iteracyjnych, co jest dodatkowym argumentem za ich stosowaniem.

Powodem, który także może zdecydować o wyborze metody iteracyjnej, mimo iż nie dysponujemy dobrym pierwszym przybliżeniem, jest niemożność rozwiązania zadania metodą bezpośrednią ze względu na ograniczenie pamięci. Napotykamy je bardzo szybko na przykład w zadaniach trójwymiarowych, dla których macierz sztywności zachowuje wprawdzie pasmowość, jednak szerokość pasma rośnie znacznie szybciej wraz z rozmiarem zadania niż w przypadku płaskim.

Przytoczone wyżej argumenty sugerują, że efektywny algorytm iteracyjny jest istotnym elementem procedury symulacji przepływów z zastosowaniem adaptacyjnej metody elementów skończonych. W tym kontekście krytycyzm musi budzić główny aspekt metodologii proponowanej w tej pracy – zastosowanie elementów silnie wydłużonych. Elementy takie wpływają bowiem niekorzystnie na uwarunkowanie macierzy sztywności, które z kolei determinuje szybkość zbieżności procedur iteracyjnego rozwiązywania równań. Sposobem na obejście problemu z tego uwarunkowania jest zastosowanie odpowiedniego operatora uwarunkowania wstępnego, tj. operatora, który złożony z daną macierzą daje w wyniku macierz dobrze uwarunkowaną. W następnych rozdziałach przedstawimy konstrukcję algorytmu iteracyjnego z uwarunkowaniem wstępnym, którego zbieżność jest niezależna od obecności elementów silnie wydłużonych w siatce. Algorytm ten to połączenie metody iteracyjnej typu gradienty sprzężone (w przypadku symetrii) lub procedury GMRES (przy braku symetrii) z uwarunkowaniem wstępnym za pomocą iteracji prostej w postaci metody blokowej Jacobiego lub Gaussa–Seidela.

## 6.2 Metody iteracyjne

Wśród metod iteracyjnych rozwiązywania równań liniowych

$$Ax = b \quad (6.2.1)$$

wyróżnić można tzw. iteracje proste (lub metody stacjonarne pierwszego rzędu) zdefiniowane wzorem:

$$x^{n+1} = Gx^n + k \quad (6.2.2)$$

gdzie  $G$  jest ustaloną macierzą,  $k$  ustalonym wektorem, oraz metody bardziej złożone wyrażające  $x^{n+1}$  za pomocą większej liczby poprzednich przybliżeń i w sposób niekoniecznie liniowy z ustalonymi  $G$  i  $k$  jak w (6.2.2). Do pierwszej grupy zaliczyć można znane algorytmy Jacobiego, Gaussa–Seidela wraz z ich wersjami blokowymi, metodę Richardsona, metodę nadrelaksacji (SOR, *successive overrelaxation*) czy symetryczną metodę nadrelaksacji (SSOR). Metody drugiego rodzaju to algorytm Czebyszewa, metoda gradientów

sprzężonych, metoda GMRES, Orthomin, BCG (*biconjugate gradient*) i wiele innych. Metody te są na ogół istotnie szybciej zbieżne od iteracji prostych, gdyż zawierają pewien mechanizm optymalizujący proces zbieżności. Poniżej omawiamy pokrótce warunki zbieżności metod iteracji prostej, gradientów sprzężonych i GMRES. (Definicję procedur gradientów sprzężonych, gradientów sprzężonych z uwarunkowaniem wstępnym i metody GMRES podajemy w dodatku A.5).

Warunek wystarczający zbieżności iteracji prostej ma postać (zob. np. [45])

$$\|G\| < 1, \quad (6.2.3)$$

gdzie  $\|\cdot\|$  jest normą zdefiniowaną jak dla odwzorowań liniowych, tj.

$$\|G\| = \sup_{x \neq 0} \frac{\|Gx\|}{\|x\|} \quad (6.2.4)$$

przy jakiegokolwiek normie dla wektorów  $\|x\|$  (inaczej: norma macierzy jest zgodna z daną normą dla wektorów).

Szybkość zbieżności iteracji jest wyznaczana właśnie przez wielkość  $\|G\|$ :

$$\frac{\|x_{n+1} - x\|}{\|x_n - x\|} \leq \|G\| \quad (6.2.5)$$

Dla metody gradientów sprzężonych oszacowanie zbieżności ma następującą postać [71]

$$\frac{\|x - x_n\|_A}{\|x - x_0\|_A} \leq 2 \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^n \quad (6.2.6)$$

gdzie norma wektorowa jest zdefiniowana jako  $\|x\|_A^2 := x^T A x$ , macierz układu równań  $A$  jest z założenia symetryczna i dodatnio określona,  $\kappa$  oznacza uwarunkowanie macierzy.

(Przez uwarunkowanie macierzy,  $\kappa(A)$ , rozumiemy wielkość

$$\kappa = \|A\| \cdot \|A^{-1}\|, \quad (6.2.7)$$

gdzie norma dla macierzy jest zgodna z pewną normą dla wektorów. Najpopularniejszym wyborem jest tu norma euklidesowa,  $\|x\| = (\sum_i x_i^2)^{1/2}$ , przy której dla macierzy symetrycznych dodatnio półokreślonych norma  $\|A\|$  jest maksymalną wartością własną  $A$ . W tym też przypadku

$$\kappa(A) = \lambda_{max} / \lambda_{min}, \quad (6.2.8)$$

gdzie  $\lambda_{max}$  i  $\lambda_{min}$  oznaczają odpowiednio maksymalną i minimalną wartość własną  $A$ .

W przypadku algorytmu GMRES zbieżność jest zależna od minimalnej wartości własnej symetrycznej części macierzy układu równań  $C_A$  oraz od normy macierzy  $C_A$

$$c_A = \inf_{x \neq 0} \frac{(x, Ax)}{(x, x)} \quad \text{i} \quad C_A = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad (6.2.9)$$

gdzie  $(\cdot, \cdot)$  jest danym iloczynem skalarnym,  $\|x\| := (x, x)^{1/2}$ . Jeśli  $c_A > 0$ , to algorytm GMRES zbiega i po  $m$  krokach norma residuum jest ograniczona przez:

$$\|r_m\| \leq \left( 1 - \frac{c_A^2}{C_A^2} \right)^{m/2} \|r_0\| \quad (6.2.10)$$

Wynik ten cytujemy za [41]. <http://rcin.org.pl>

### 6.3 Idea uwarunkowania wstępnego za pomocą iteracji prostej

Efektywność metody gradientów sprzężonych i GMRES spada wraz z pogarszaniem się uwarunkowania układu równań. Jak już wspomniano, jednym ze sposobów zapobieżenia temu zjawisku jest zastosowanie uwarunkowania wstępnego, tj. operatora liniowego oznaczonego zwyczajowo przez  $M^{-1}$  o tej własności, że jeśli pomnożymy przezeń oryginalny układ  $Ax = b$ , to macierz  $M^{-1}A$  nowego układu,

$$M^{-1}A = M^{-1}b \quad (6.3.11)$$

ma istotnie lepsze uwarunkowanie niż  $A$ .

Dobry algorytm uwarunkowania wstępnego powinien mieć następujące cechy:

1. Jego zastosowanie nie powinno być kosztowne.
2. Operator  $M^{-1}$  powinien być symetryczny, jeśli oryginalny układ jest symetryczny.
3. Uwarunkowanie  $M^{-1}A$  powinno być jak najbliższe 1.

Warunki 1. i 3. są wzajemnie przeciwstawne: idealnym kandydatem spełniającym 3. jest  $M^{-1} = A^{-1}$ , gdyż  $\kappa(A^{-1}A) = \kappa(I) = 1$ , który to wybór jest nie do przyjęcia ze względu na koszt. Pomysł rozwiązania opartego na tej obserwacji polega na przyjmowaniu  $M^{-1}$  jako różnych form przybliżonego operatora odwrotnego do  $A$ . Jedną z realizacji tej idei to wykorzystanie iteracji prostej typu (6.2.2) i formalne zdefiniowania  $M^{-1}$  jako  $M^{-1} := (I - G)A^{-1}$ , co pozwala przedstawić macierz z uwarunkowaniem wstępnym jako:

$$M^{-1}A = I - G \quad (6.3.12)$$

Tak zdefiniowany operator uwarunkowania wstępnego ma następujące własności:

1. Jeśli  $\|G\| < 1$ , co jest warunkiem zbieżności iteracji prostej, to możemy zgrubnie oszacować  $\kappa(M^{-1}A)$  jako:

$$\kappa(I - G) \leq \frac{1 + \|G\|}{1 - \|G\|} \quad (6.3.13)$$

(Wynika to z następującego rachunku:

$$\text{i) } \|I - G\| \leq \|I\| + \|G\| = 1 + \|G\|,$$

$$\text{ii) } \|(I - G)^{-1}\| = \|(I - G)^{-1}(I - G + G)\| = \|I + (I - G)^{-1}G\| \leq \|I\| + \|(I - G)^{-1}\| \cdot \|G\| \implies \|(I - G)^{-1}\| \leq (1 - \|G\|)^{-1}.$$

Iloczyn tych dwu oszacowań daje (6.3.13).)

Jeśli więc iteracja prosta jest szybko zbieżna, to uwarunkowanie  $I - G$  jest dobre - bliskie 1.

2. Warunek  $\|G\| < 1$  implikuje zbieżność algorytmu GMRES. Mamy bowiem:

$$c_A = \inf_{\mathbf{x}} \frac{(\mathbf{x}, (I - G)\mathbf{x})}{(\mathbf{x}, \mathbf{x})} \sup_{\mathbf{y}} \frac{(\mathbf{y}, G\mathbf{y})}{(\mathbf{y}, \mathbf{y})} \quad (6.3.14)$$

oraz

$$\sup \frac{(\mathbf{x}, \mathbf{G}\mathbf{x})}{(\mathbf{x}, \mathbf{x})} \leq \sup \frac{(\mathbf{x}, \mathbf{x})^{1/2} (\mathbf{G}\mathbf{x}, \mathbf{G}\mathbf{x})^{1/2}}{(\mathbf{x}, \mathbf{x})} = \|\mathbf{G}\| < 1 \quad (6.3.15)$$

gdzie wykorzystano nierówność Cauchy'ego-Schwarza i definicję normy operatora. Ograniczoność  $C_A$  jest oczywista:

$$C_A = \|\mathbf{I} - \mathbf{G}\| \leq 1 + \|\mathbf{G}\| < 2 \quad (6.3.16)$$

Powyższe warunki wraz z (6.2.10) dowodzą zbieżności procedury GMRES.

3. Obliczanie iloczynu „macierz razy wektor” oraz residuum dla zadanego wektora  $\mathbf{z}$  dla układu z uwarunkowaniem wstępnym przyjmuje postać:

$$\mathbf{y} = \mathbf{M}^{-1}\mathbf{A}\mathbf{z} = (\mathbf{I} - \mathbf{G})\mathbf{z} = \mathbf{z} - \mathbf{G}\mathbf{z} \quad (6.3.17)$$

$$\mathbf{r} = \mathbf{M}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{z}) = (\mathbf{I} - \mathbf{G})\mathbf{A}^{-1}\mathbf{b} - (\mathbf{I} - \mathbf{G})\mathbf{z} = (\mathbf{k} + \mathbf{G}\mathbf{z}) - \mathbf{z} \quad (6.3.18)$$

gdzie  $\mathbf{k} = (\mathbf{I} - \mathbf{G})\mathbf{A}^{-1}\mathbf{b}$ , co wynika z konsystencji (zgodności) iteracji (6.2.2), tj. z warunku  $\mathbf{x} = \mathbf{G}\mathbf{x} + \mathbf{k}$  przy  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ .

Powyższe wzory wskazują, że obydwie wspomniane operacje można wykonać przez wywołanie procedury iteracji prostej (w pierwszym wypadku dla układu jednorodnego,  $\mathbf{k} = 0$ ) i wzięcie różnicy wyniku i wyjściowego wektora  $\mathbf{z}$ . Jest to bardzo korzystne, gdyż algorytm iteracji prostej może być zastosowany jako „czarna skrzynka”, istotny jest wyłącznie wynik jego działania na danym  $\mathbf{z}$ .

W kontekście metody elementów skończonych opracowano bardzo wiele algorytmów uwarunkowania wstępnego, które mają zastosowanie do różnych zagadnień i rodzajów aproksymacji. Metody iteracji prostej, które w podejściu skończenie elementowym znajdują szczególnie przejrzystą i intuicyjną interpretację, to metoda blokowa Jacobiego i Gaussa-Seidela oraz metody do nich zbliżone. Analizę tego rodzaju procedur zainicjował w XIX wieku A.H. Schwarz, stąd też określane bywają jako metody Schwarza. Inne spotykane określenie to metody oparte o dekompozycję obszaru na zachodzące na siebie podobszary (*overlapping domain decomposition methods*). Ta nazwa wiąże się z konstrukcją algorytmów: bloki stopni swobody definiujące procedury Jacobiego i Gaussa-Seidela wybiera się jako stopnie swobody związane z ustalonymi zachodzącymi na siebie podobszarami pokrywającymi siatkę.

Argumentem za wyborem metod typu Schwarza do uwarunkowania wstępnego macierzy sztywności dla elementów silnie wydłużonych jest wynik teoretyczny wskazujący, że wypadkowe uwarunkowanie  $\kappa(\mathbf{M}^{-1}\mathbf{A})$  zależy wyłącznie od geometrii podziału siatki na wspomniane podobszary. Wynik ten otrzymano dla regularnych zadań eliptycznych drugiego rzędu dla siatek quasi-równomiernych. W dalszej części pokażemy jego uogólnienie dla siatek z elementami wydłużonymi i dla klasy problemów obejmujących sformułowanie dla przepływów.

## 6.4 Metody Schwarza

W tym podrozdziale przedstawimy podsumowanie wyników teoretycznych dla metod Schwarza według pracy [39] Dryja i Widłunda.

Rozważmy eliptyczny problem brzegowy drugiego rzędu w sformułowaniu słabym postaci:

Znaleźć  $\mathbf{u} \in V \subset (H^1(\Omega))^k$  takie że:

$$a(\mathbf{u}, \mathbf{v}) = l(\mathbf{v}), \quad \forall \mathbf{v} \in V, \quad (6.4.19)$$

gdzie  $k$  jest liczbą składowych rozwiązań,  $\Omega \subset R^2$  jest obszarem problemu,  $a(\cdot, \cdot)$  i  $l(\cdot)$  są formą dwuliniową i funkcjonałem liniowym:

$$a(\mathbf{u}, \mathbf{v}) = \int_{\Omega} (\nabla \mathbf{v})^T \mathbf{a} \nabla \mathbf{u} \, dx + \int_{\Omega} \mathbf{v}^T \mathbf{b} \mathbf{u} \, dx, \quad l(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, dx + \int_{\partial \Omega_i} \mathbf{g} \cdot \mathbf{v} \, dS \quad (6.4.20)$$

z  $\mathbf{a}, \mathbf{b}, \mathbf{f}, \mathbf{g}$  będącymi algebraicznymi formami dwu- i liniowymi,  $\partial \Omega_i$  częścią brzegu z naturalnymi warunkami brzegowymi. Zakładamy, że  $\mathbf{a}$  jest symetryczne jednostajnie dodatnio określone w  $\Omega$ ,  $\mathbf{b} = \mathbf{b}^T \geq 0$ . Zakładamy też, że postać warunków brzegowych typu Dirichleta jest taka, że forma  $a(\cdot, \cdot)$  jest ciągła i koercywna względem półnormy  $H^1$ .

Problem jest aproksymowany metodą elementów skończonych. W tym celu definiujemy standardową przestrzeń skończenie elementową  $V^h \subset V$  i definiujemy przybliżenie Galerkin problemu (6.4.19):

Znaleźć  $\mathbf{u}_h \in V^h$  takie że:

$$a(\mathbf{u}_h, \mathbf{v}_h) = l(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in V^h \quad (6.4.21)$$

Oznaczając odpowiednio przez  $\mathbf{U}, \mathbf{A}$  i  $\mathbf{L}$  wektor rozwiązania, macierz sztywności i wektor obciążenia

$$\mathbf{u}_h = \sum_{i=1}^M U_i \psi_i(\mathbf{x}), \quad A_{ij} = a(\psi_i, \psi_j), \quad L_i = l(\psi_i), \quad i, j = 1, \dots, M \quad (6.4.22)$$

z  $\psi_i$ ,  $i = 1, \dots, M$  będącymi globalnymi bazowymi funkcjami kształtu,  $M$  liczbą stopni swobody, możemy zapisać (6.4.21) jako układ algebraicznych równań liniowych:

$$\mathbf{A} \mathbf{U} = \mathbf{L} \quad (6.4.23)$$

Naszym celem jest rozwiązanie (6.4.23) metodą blokową Jacobiego z blokami niewiadomych związanymi z zachodzącymi na siebie podobszarami  $\Omega_i$ ,  $i = 1, \dots, N$  pokrywającymi cały obszar  $\Omega$ . Zakładamy że brzegi  $\partial \Omega_i$  przebiegają wzdłuż brzegów elementów.

Wprowadźmy następujące oznaczenia:

$V_i^h = \{\mathbf{u} \in V^h : \mathbf{u} = 0 \text{ w } \Omega \setminus \Omega_i\}$ ,  $i = 1, \dots, N$ , zbiór przestrzeni funkcji kształtu o nośnikach ograniczonych do  $\Omega_i$ .

$M_i$ ,  $i = 1, \dots, N$  - wymiary przestrzeni  $V_i^h$ .

$\mathbf{U}_i \in R^{M_i}$  - wektor stopni swobody funkcji  $\mathbf{u} \in V_i^h$  odpowiadających  $M_i$  bazowym funkcjom kształtu przestrzeni  $V_i^h$  wybranych spośród wszystkich funkcji kształtu  $\{\psi_i\}_{i=1}^M$ .

$\mathbf{R}_i^T \in R^{M \times M_i}$  - macierz rozszerzająca (przez uzupełnienie zerami) stopnie swobody podobszaru  $\Omega_i$  do globalnego wektora stopni swobody (macierz ta jest reprezentacją iniekcji  $\mathbf{u} \in V_i^h$  w  $\mathbf{u} \in V^h$ );  $k$ -ty wiersz  $\mathbf{R}_i^T$  zawiera zera oprócz  $l$ -tego wyrazu równego 1, jeśli  $l$ -ta funkcja bazowa  $V_i^h$  jest jednocześnie  $k$ -tą funkcją bazową w  $V^h$ .

$\mathbf{R}_i \in R^{M_i \times M}$  - macierz pozwalająca wybrać z globalnego wektora stopni swobody stopnie związane z  $\Omega_i$  (macierz restrykcji  $\mathbf{R}_i : V^h \ni \mathbf{u} \rightarrow \mathbf{u} \in V_i^h$ ).

$\mathbf{A}_i \in R^{M_i \times M_i}$ ,  $i = 1, \dots, N$ , macierze sztywności związane z wewnętrznymi stopniami swobody podobszarów  $\Omega_i$ .

Rozważmy następujący schemat Jacobiego rozwiązania (6.4.23):

Dane jest przybliżenia  $\mathbf{u}^n \in V^h$  <http://im.ORG.pl>

1. Dla  $i = 1, \dots, N$  znaleźć  $\mathbf{u}_i^{n+1} \in V_i^h + \mathbf{u}^n$ , takie że:

$$a(\mathbf{u}_i^{n+1}, \mathbf{v}) = l(\mathbf{v}), \quad \forall \mathbf{v} \in V_i^h \quad (6.4.24)$$

2. Przyjąć:

$$\mathbf{u}^{n+1} := \mathbf{u}^n + \sum_{i=1}^N (\mathbf{u}_i^{n+1} - \mathbf{u}^n) \quad (6.4.25)$$

Reprezentację macierzową powyższego algorytmu znajdziemy zauważając, że (6.4.24) jest problemem z warunkiem brzegowym Dirichleta na  $\partial\Omega_i$ :  $\mathbf{u}_i^{n+1} = \mathbf{u}^n$ . Dlatego też formalnie (6.4.24) jest równoważne zadaniu Dirichleta z jednorodnym warunkiem brzegowym  $a(\Delta \mathbf{u}_i, \mathbf{v}) = -a(\mathbf{u}^n, \mathbf{v}) + l(\mathbf{v})$ , gdzie  $\Delta \mathbf{u}_i = (\mathbf{u}_i^{n+1} - \mathbf{u}^n) \in V_i^h$ . Rozwiązanie zadania jest następujące:

1. Wyznaczamy stopnie swobody  $\Delta U$ , funkcji  $\Delta \mathbf{u}_i$ :

$$\Delta U_i = \mathbf{A}_i^{-1} \mathbf{R}_i (-A U^n + \mathbf{L}), \quad i = 1, \dots, N \quad (6.4.26)$$

2. Przyjmujemy

$$U^{n+1} = U^n + \sum_{i=1}^N \mathbf{R}_i^T \Delta U_i \quad (6.4.27)$$

Ostatecznie, łącząc (6.4.26) i (6.4.27) otrzymujemy:

$$U^{n+1} = (I - \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i A) U^n + \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i \mathbf{L} \quad (6.4.28)$$

Powyższy wzór definiuje więc iterację prostą typu (6.2.2) z

$$G = I - \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i A \quad \text{ i } \quad \mathbf{k} = \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i \mathbf{L} \quad (6.4.29)$$

Podobnie jak algorytm Jacobiego definiujemy metodę Gaussa-Seidela:

Mając dane przybliżenia  $\mathbf{u}^n \in V^h$ :

1. Przyjąć  $\mathbf{u}_0^{n+1} := \mathbf{u}^n$ .

2. Dla  $i = 1, 2, \dots, N$ :

Znaleźć  $\mathbf{u}_i^{n+1} \in V_i^h + \mathbf{u}_{i-1}^{n+1}$  takie że:

$$a(\mathbf{u}_i^{n+1}, \mathbf{v}) = l(\mathbf{v}) \quad \forall \mathbf{v} \in V_i^h \quad (6.4.30)$$

3.  $\mathbf{u}^{n+1} := \mathbf{u}_N^{n+1}$ .

Postać macierzową  $i$ -tego kroku znajdziemy podobnie jak dla algorytmu Jacobiego:

Wyznaczamy

$$\Delta U_i^{i+1} = \mathbf{A}_i^{-1} \mathbf{R}_i (-A U_{i-1}^{n+1} + \mathbf{L}) \quad (6.4.31)$$

gdzie  $\Delta U_{i-1}^{n+1}$  są stopniami swobody funkcji  $\mathbf{u}_i^{n+1} - \mathbf{u}_{i-1}^{n+1}$ . Stąd zaś znajdujemy

$$U_i^{n+1} = U_{i-1}^{n+1} + \mathbf{R}_i^T \Delta U_i^{i+1} = \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i \mathbf{L} - (\mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i A - I) U_{i-1}^{n+1} \quad (6.4.32)$$

Po uwzględnieniu, że  $L = AU$ , możemy przedstawić różnicę  $U_i^{n+1} - U$  jako:

$$U_i^{n+1} - U = (I - R_i^T A_i^{-1} R_i A)(U_{i-1}^{n+1} - U) \quad (6.4.33)$$

co, przez indukcję, implikuje z kolei, że:

$$U^{n+1} - U = \prod_{i=1}^N (I - R_i^T A_i^{-1} R_i A)(U^n - U) \quad (6.4.34)$$

czyli ostatecznie, że:

$$U^{n+1} = \prod_{i=1}^N (I - R_i^T A_i^{-1} R_i A)U^n + [I - \prod_{i=1}^N (I - R_i^T A_i^{-1} R_i A)]A^{-1}L \quad (6.4.35)$$

Powyższy wzór definiuje iterację prostą typu (6.2.2).

### 6.4.1 Globalny krok iteracji

Algorytmy Jacobiego i Gaussa-Seidela korzystnie jest uzupełnić o globalny krok w iteracjach, tj. o procedurę, w której uaktualniane są wszystkie stopnie swobody siatki choć na wybranych podobszarach zmiany te nie są niezależne lecz związane z niewielką liczbą wyróżnionych węzłów siatki. Węzły te tworzą pewną siatkę zgrubną w stosunku do siatki aktualnej. Intuicyjnie mówiąc globalny krok iteracji ma umożliwić przepływ informacji o uaktualnianym rozwiązaniu w całym obszarze  $\Omega$  jednocześnie. Teoretyczne uzasadnienie korzyści wynikających z iteracji na zgrubnej siatce znajdujemy w teorii metod wielosiatkowych (*multigrid methods*), gdzie pokazuje się, że składowe residuum (lub błąd) o wolnej zmienności przestrzennej korzystnie jest tłumić za pomocą iteracji na grubych siatkach.

Globalny krok iteracji Jacobiego czy Gaussa-Seidela zdefiniujemy analogicznie do kroków lokalnych na podobszarach  $\Omega_i$ . W tym celu definiujemy dodatkową przestrzeń skończonej elementową  $V_0^h \subset V^h$ , która podobnie jak  $V_i^h$  ma wymiar istotnie mniejszy niż  $V^h$ , lecz składa się z funkcji o nośnikach pokrywających cały obszar  $\Omega$ . Uaktualnienie rozwiązania na grubej siatce znajdujemy następująco:

Znaleźć  $u_0^{n+1} \in V_0^h + u_{ref}$  takie że:

$$a(u_0^{n+1}, v) = l(v), \quad \forall v \in V_0^h, \quad (6.4.36)$$

gdzie dla algorytmu Jacobiego  $u_{ref} = u^n$  zaś dla procedury Gaussa-Seidela  $u_{ref}$  jest najbardziej aktualnym przybliżeniem rozwiązania a więc zależy od pozycji od kroku globalnego w sekwencji pozostałych relaksacji. Jeśli krok globalny umieścimy na samym początku, wówczas także  $u_{ref} = u^n$  (co w dalszej części będziemy zakładać). Macierzowa postać kroku globalnego jest zupełnie analogiczna do wersji lokalnych, wyraża się także przez macierz sztywności  $A_0$  odpowiadającą wybranej bazie przestrzeni  $V_0^h$ , przez macierz iniekcji  $R_0^T$  z  $V_0^h$  do  $V^h$  i restrykcji  $R_0$ :  $i$ -ty wiersz  $R_0^T$  zawiera współczynniki rozwinięcia  $i$ -tej funkcji bazowej przestrzeni  $V_0^h$  jako kombinacji liniowej bazy  $V^h$ , tj.  $\{\psi_i\}_{i=1}^M$ . Zapis macierzowy globalnej relaksacji jest więc następujący:

$$\Delta U_0 = A_0^{-1} R_0 (-AU^n + L) \quad (6.4.37)$$

gdzie  $\Delta U_0$  jest wektorem stopni swobody dla  $u_0^{n+1} - u^n \in V_0^h$ , zaś ostateczna postać  $U^{n+1}$  wyraża się wzorami identycznymi z (6.4.28) i (6.4.35) z zakresem wskaźnika  $i$  rozszerzonym o wartość 0 odpowiadającą wyrazom kroku globalnego.

## 6.4.2 Modyfikacja operatora uwarunkowania wstępnego

Operator uwarunkowania wstępnego  $M^{-1}$  odpowiadający iteracji prostej (6.4.28) ma postać:

$$M^{-1} = \sum_{i=0}^N R_i^T A_i^{-1} R_i \quad (6.4.38)$$

W zastosowaniach praktycznych możliwe jest zastąpienie oryginalnej formy dwuliniowej  $a(\cdot, \cdot)$  służącej do obliczania macierzy sztywności  $A_i$  przez równoważną formę dwuliniową  $b(\cdot, \cdot)$  (tj. taką że normy indukowane przez  $a$  i  $b$  są równoważne). Celem takiej zamiany jest osiągalna w wielu przypadkach redukcja kosztów obliczeń związana np. z przyjęciem stałych współczynników w definicji  $b(\cdot, \cdot)$  czy z zastosowania szczególnie efektywnych algorytmów odwracania macierzy  $A_i$ , gdy odpowiadają one np. równaniu Laplace'a na równomiernej siatce. Formalnie osiągamy to przez zastąpienie macierzy  $A_i$  w (6.4.38) przez macierze  $B_i$  odpowiadające formie  $b(\cdot, \cdot)$ :

$$M^{-1} = \sum_{i=0}^N R_i^T B_i^{-1} R_i \quad (6.4.39)$$

Globalna macierz układu z uwarunkowaniem wstępnym przybiera postać:

$$T = M^{-1} A = \sum_{i=0}^N R_i^T B_i^{-1} R_i A \quad (6.4.40)$$

i może być rozumiana formalnie jako suma operatorów  $T_i : V^h \rightarrow V^h$  zdefiniowanych następująco: dla danej funkcji  $\mathbf{u} \in V^h$  znajdujemy  $\mathbf{u}_i \in V_i^h$ , takie że

$$b_i(\mathbf{u}_i, \mathbf{v}) = a(\mathbf{u}, \mathbf{v}), \quad \forall \mathbf{v} \in V_i^h, \quad (6.4.41)$$

i definiujemy  $T_i \mathbf{u} := \mathbf{u}_i$  jako element  $V^h$ . Postać macierzowa operatora  $T_i$  wynikająca z wyliczenia stopni swobody  $\mathbf{u}_i$  z (6.4.41) i przedstawienia ich za pomocą stopni swobody przestrzeni  $V^h$  jest więc następująca:

$$T_i = R_i^T B_i^{-1} R_i A \quad (6.4.42)$$

Zauważmy, że macierze  $T_i$  (a więc i  $T$ ) są symetryczne względem iloczynu skalarnego definiwanego przez  $A$  (tj.  $A$ -symetryczne):  $(T\mathbf{x})^T A \mathbf{y} = \mathbf{x}^T A (T\mathbf{y})$ ,  $\forall \mathbf{x}, \mathbf{y}$ , czyli że  $T^T A = A T$ . Podobnie symetryczna jest macierz  $T^{-1}$ .

Operator uwarunkowania wstępnego typu Gaussa-Seidela może zostać zmodyfikowany według identycznych reguł jak przedstawione powyżej.



## 6.5 Oszacowanie wartości własnych macierzy z warunkowaniem wstępnym

Z definicji

$$\kappa(M^{-1}A) = \|T\| \cdot \|T^{-1}\| \quad (6.5.43)$$

Dla symetrycznych względem  $A$  macierzy  $T$  i  $T^{-1}$  ich normy operatorowe indukowane przez normę wektorową  $\|x\|_A := (x^T A x)^{1/2}$  możemy wyrazić przez ekstremalne wartości własne macierzy  $T$ :

$$\|T\|_A = \lambda_{\max}(T); \quad i \quad \|T^{-1}\|_A = \lambda_{\min}^{-1}(T) \quad (6.5.44)$$

W pracy [40] przedstawiono następujące oszacowania wartości własnych  $\lambda_{\max}(T)$  i  $\lambda_{\min}(T)$ .

### Lemat 6.5.2

Załóżmy, że istnieje taka stała  $C_0$ , że dla każdego  $u \in V^h$  istnieje przedstawienie  $u = z_0 + z_1 + \dots + z_N$ ,  $z_i \in V_i^h$ ,  $i = 0, 1, \dots, N$ , o tej własności, że

$$\sum_{i=0}^N b_i(z_i, z_i) \leq C_0 a(u, u) \quad (6.5.45)$$

Wówczas

$$\lambda_{\min}^{-1} \leq C_0 \quad (6.5.46)$$

### Twierdzenie 6.5.2

Załóżmy, że istnieją takie stałe  $0 \leq \varepsilon_{ij} \leq 1$ ,  $i, j = 1, \dots, N$ , że

$$a(u_i, u_j) \leq \varepsilon_{ij} a(u_i, u_i)^{1/2} a(u_j, u_j)^{1/2}, \quad \forall u_i \in V_i^h, \quad u_j \in V_j^h, \quad i, j = 1, \dots, N \quad (6.5.47)$$

(Nierówność tę nazywa się wzmocnioną nierównością Cauchy'ego-Schwarza przez analogię do klasycznej nierówności Cauchy'ego-Schwarza,  $a(u_i, u_j) \leq a(u_i, u_i)^{1/2} a(u_j, u_j)^{1/2}$ .)

Niech  $\varrho(\varepsilon)$  będzie promieniem spektralnym macierzy  $\varepsilon$ .

Niech  $\omega_+$  będzie stałą występującą w definicji równoważności form  $a$  i  $b_i$ , tj.

$$a(u_i, u_i) \leq \omega_+ b_i(u_i, u_i), \quad \forall u_i \in V_i^h, \quad i = 0, \dots, N \quad (6.5.48)$$

Wówczas

$$\lambda_{\max} \leq \omega_+ (1 + \varrho(\varepsilon)) \quad (6.5.49)$$

Parametry  $C_0$  i  $\varrho(\varepsilon)$  występujące w powyższych oszacowaniach zależą od konstrukcji podprzestrzeni  $V_i^h$ , tj. od geometrii podziału obszaru  $\Omega$  na podobszary  $\Omega_i$ . Zależność tę przedstawiono w pracy [109].

### Twierdzenie 6.5.3

#### Założenia:

1. Siatka skończenie elementowa jest quasi-równomierna ze średnicą elementów  $h$ .
2. Liczba sąsiadów podobszaru  $\Omega_i$ , tj. podobszarów  $\Omega_j$ , takich że  $\text{meas}(\Omega_i \cap \Omega_j) > 0$ , wynosi  $k_i$ .

3. Średnica podobszaru  $\Omega_i$  wynosi  $H_i$ , zaś  $\delta_i$  oznacza minimalną grubość części wspólnej  $\Omega_i$  i jego sąsiadów,  $i = 1, \dots, N$ .
4. Normy  $b_i(\cdot, \cdot)^{1/2}$  są ograniczone przez półnormę Sobolewa  $H^1$ :

$$b_i(\mathbf{u}_i, \mathbf{u}_i) \leq \frac{1}{\omega_-} \int_{\Omega_i} |\nabla \mathbf{u}_i|^2 dx, \quad \forall \mathbf{u}_i \in V_i^h, \quad i = 0, \dots, N \quad (6.5.50)$$

5. Przestrzeń  $V_0^h$  zdefiniowana na zgrubnej siatce spełnia następujący warunek: Istnieje operator liniowy  $Q_H : V^h \rightarrow V_0^h$ , taki że

$$\int_{\Omega_i} |\nabla Q_H \mathbf{v}|^2 dx + \frac{1}{H_i^2} \int_{\Omega_i} |\mathbf{v} - Q_H \mathbf{v}|^2 dx \leq C \int_{\Omega_i} |\nabla \mathbf{v}|^2 dx \quad (6.5.51)$$

dla  $i = 1, \dots, N$  i dla każdego  $\mathbf{v} \in V^h$ .

6. Forma  $a(\cdot, \cdot)$  jest eliptyczna i ciągła względem półnormy  $H^1$ , tj.

$$a(\mathbf{v}, \mathbf{v}) \geq \alpha \int_{\Omega} |\nabla \mathbf{v}|^2 dx, \quad \forall \mathbf{v} \in V \text{ z } \alpha > 0. \quad (6.5.52)$$

$$a(\mathbf{u}, \mathbf{v}) \leq M |\mathbf{u}|_{1,\Omega} |\mathbf{v}|_{1,\Omega} \quad \forall \mathbf{u}, \mathbf{v} \in V \quad (6.5.53)$$

7. Każdy punkt  $\mathbf{x} \in \Omega$  należy do co najwyżej  $k_0$  podobszarów  $\Omega_i$ .

#### Teza:

Istnieją stałe  $C_1, C_2$  niezależne od  $\mathbf{a}, \mathbf{b}, H_i, \delta_i, h$  takie że:

$$\rho(\varepsilon) \leq \max_{i \geq 1} k_i \quad (6.5.54)$$

$$C_0 \leq \frac{k_0 C_1}{\alpha \omega_-} \left( 1 + C_2 \max_{i \geq 1} \frac{H_i}{\delta_i} \right) \quad (6.5.55)$$

A zatem, zgodnie z (6.5.46) i (6.5.49), uwarunkowanie macierzy  $\mathbf{T}$  można oszacować następująco:

$$\kappa(\mathbf{T}) \leq \frac{k_0 C_1 \omega_+}{\alpha \omega_-} \left( 1 + C_2 \max_{i \geq 1} \frac{H_i}{\delta_i} \right) \left( 1 + \max_{i \geq 1} k_i \right) \quad (6.5.56)$$

**Uwaga 6.5.1** Istotnym elementem przedstawionej teorii jest operator  $Q_H$  mający spełniać (6.5.51). Jeśli założyć wspólne ograniczenie  $H_i/h \leq C$ ,  $i = 1, \dots, N$  niezależnie od  $h$ , i jeśli podobszary  $\Omega_i$  tworzą zgrubną siatkę skończenie elementową definiującą  $V_0^h$ , to warunek (6.5.51) spełnia operator interpolacji  $I_H : V^h \rightarrow V_0^h$ .

Jeśli wspomniane wyżej wspólne ograniczenie nie jest spełnione, to rolę operatora  $Q_H$  może pełnić rzutowanie  $L^2$  z  $V^h$  do  $V_0^h$ . W tym wypadku jednak siatka zgrubna musi być quasi-równomierna z rozmiarem elementów  $H$ , zaś warunek (6.5.51) jest spełniony w wersji globalnej

$$\int_{\Omega} |\nabla Q_H \mathbf{v}|^2 dx + \frac{1}{H^2} \int_{\Omega} |\mathbf{v} - Q_H \mathbf{v}|^2 dx \leq C \int_{\Omega} |\nabla \mathbf{v}|^2 dx, \quad \forall \mathbf{v} \in V^h \quad (6.5.57)$$

(co jest wystarczającym założeniem dla rozważanego twierdzenia). Dowód własności (6.5.57) dla elementów trójkątnych można znaleźć w pracy Bramble i Xu [19]. Jego uogólnienie

na siatki elementów czworokątnych z więzami podaje praca autora [88], gdzie pokazano też warunek lokalny (6.5.51) dla operatora  $I_H$ .

Jeśli pominąć krok globalny operatora uwarunkowania wstępnego, to wartość  $C_0$  (a zatem i  $\kappa(\mathbf{T})$ ) staje się zależna od rozmiaru podobszarów [101]

$$C_0 \leq C \left( 1 + C' \max_i \frac{1}{H_i^2} \frac{H_i}{\delta_i} \right) \quad (6.5.58)$$

## 6.6 Oszacowanie zbieżności metody Gaussa–Seidela

Analizę multiplikatywnych metod Schwarza w przypadku, gdy oryginalne sformułowanie jest symetryczne prezentuje praca [18] Bramble'a *et al.* Zakładając, że algorytm Gaussa–Seidela został skonstruowany za pomocą dekompozycji obszaru na zachodzące na siebie podobszary i że rozważany jest problem o sformułowaniu słabym (6.4.19), warunek zbieżności iteracji Gaussa–Seidela jest związany z możliwością konstruowania „dekompozycji o niskiej energii” jak dla metod addytywnych, tj.  $\mathbf{u} = \sum_{i=0}^N \mathbf{z}_i$ ,  $\mathbf{z}_i \in V_i^h$  oraz:

$$\sum_{i=0}^N a(\mathbf{z}_i, \mathbf{z}_i) \leq C_0 a(\mathbf{u}, \mathbf{u}) \quad (6.6.59)$$

gdzie  $a$  jest formą dwuliniową zdefiniowaną w (6.4.20). Jeśli warunek (6.6.59) jest spełniony z  $C_0$  niezależnym od  $H_i$ , to

$$\|G_s\|_a \leq \gamma < 1, \quad (6.6.60)$$

( $\|\cdot\|_a = a(\cdot, \cdot)^{1/2}$ ), jeśli zaś pomijamy krok globalny w iteracjach (tzn.  $C_0$  jest dane wzorem (6.5.58)), to

$$\|G_s\|_a \leq 1 - \frac{H^2}{Cn^2} \quad (6.6.61)$$

gdzie  $H = \min_i H_i$ ,  $H_i = \text{dia}(\Omega_i)$ ,  $n = \max_i n_i$ , przy  $n_i$  będącym liczbą podobszarów  $\Omega_j$  takich że  $\text{meas}(\Omega_i \cap \Omega_j) > 0$ . Widzimy więc, że wykazanie zbieżności iteracji Gaussa–Seidela (a zatem i procedury GMRES z uwarunkowaniem wstępnym w postaci tej iteracji) sprowadza się, podobnie jak w przypadku addytywnym, do pokazania nierówności (6.6.59).

## 6.7 Generacja podobszarów

Istotnym problemem praktycznym w stosowaniu opisanych poprzednio algorytmów jest zdefiniowanie podobszarów  $\Omega_i$ . W przypadku siatek równomiernych (bez adaptacji) można to zrobić poprzez utworzenie w pliku danych list elementów składających się na poszczególne podobszary. Formalnie podejście to można zastosować także dla siatek adaptacyjnych: zdefiniować podobszary jako zbiory elementów otrzymanych z podziałów elementów siatki początkowej składających się na podobszary opisane we wspomnianych listach. Metoda taka miałaby jednak istotną wadę. W przypadku zastosowania bardzo wielu podziałów (co jest przecież istotą metod adaptacyjnych) liczba elementów zawartych w poszczególnych podobszarach, a więc i rozmiary problemów lokalnych, wzrastałyby do trudnych do zaakceptowania wielkości. Nawet jeśli pominąć ograniczenia pamięci przeznaczonej do ich rozwiązywania, można przypuszczać, że optymalna praca algorytmu zależy od rozmiaru podobszarów. Dlatego też istotne jest zaprojektowanie procedury generującej

podział siatki adaptacyjnej na zachodzące na siebie podobszary z możliwością kontroli ich rozmiarów.

Zanim przedstawimy propozycję takiej procedury, rozważmy wersję algorytmu Jacobiego dla siatek adaptacyjnych, którą opisano w [29]. W podejściu tym podobszary są zdefiniowane jako nośniki globalnych liniowych bazowych funkcji kształtu, tj. zbiory elementów posiadających wspólny faktyczny węzeł narożny (faktyczny w rozumieniu aproksymacji z więzami). Wadą algorytmu jest to, że podobszary zawierają bardzo niewiele elementów (co jest w pewnym sensie przeciwieństwem w stosunku do metody wspomnianej na początku) oraz fakt, iż w przypadku stosowania kierunkowych podziałów elementów podobszary mogą być silnie wydłużone (jak i same elementy), co wpływa na zwiększenie proporcji  $H_i/\delta_i$  z Twierdzenia 6.5.2. Faktycznie, doświadczenia numeryczne wskazują, że algorytm uwarunkowania wstępnego oparty na opisanych zasadach staje się bezużyteczny przy rozwiązywaniu zadań z bardzo wydłużonymi elementami.

Metoda generacji podobszarów, którą proponujemy, jest pewnym uogólnieniem opisanej procedury, nie ma jednak jej wad. Główna idea jest następująca:

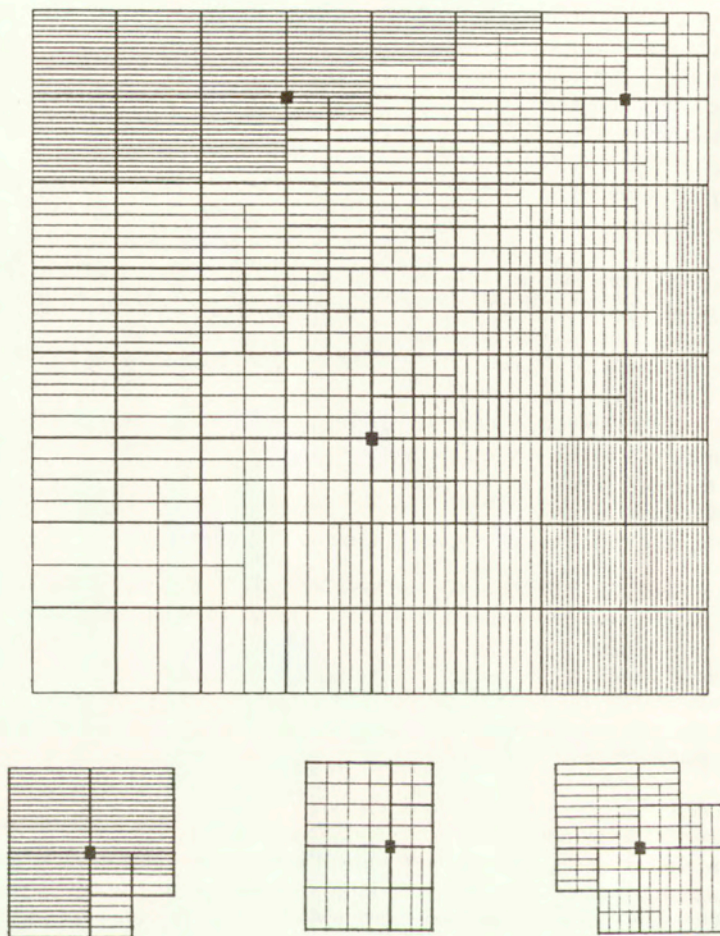
Wyobraźmy sobie, że na oryginalnej siatce początkowej oprócz faktycznie istniejącej siatki adaptacyjnej generujemy siatkę pomocniczą. Jest ona otrzymywana przez dokonywanie wyłącznie izotropowych podziałów elementów poczynając od elementów pierwotnych. Dla odróżnienia powstających w ten sposób elementów od faktycznych elementów siatki adaptacyjnej nazywamy je elementami-podobszarami. Opisanych podziałów dokonujemy zachowując 1-nieregularność powstającej siatki. W procesie tym otrzymujemy więc elementy-podobszary o wydłużeniu rzędu 1, jak w siatce pierwotnej. Podejmujemy próby łamania wszystkich pojawiających się elementów-podobszarów tak długo, aż jeden z następujących warunków zostanie spełniony:

- maksymalny rozmiar elementu aktualnej siatki zawartego w danym elemencie-podobszarze osiągnie wyznaczony ułamek  $q_1 = 1/2^i$ ,  $i = 0, 1, 2, \dots$  rozmiaru elementu-podobszaru,
- minimalny rozmiar elementu aktualnej siatki zawartego w danym elemencie-podobszarze osiągnie wyznaczony ułamek  $q_2 = 1/2^i$ ,  $i = 0, 1, 2, \dots$  rozmiaru elementu-podobszaru,

Jako efekt końcowy otrzymujemy 1-nieregularną siatkę pomocniczą elementów-podobszarów. Możemy zdefiniować teraz zachodzące na siebie podobszary  $\Omega_i$  analogicznie jak w dyskutowanej poprzednio metodzie z pracy [29], używając jednak elementów-podobszarów zamiast faktycznych elementów: podobszar  $\Omega_i$  definiujemy jako nośnik globalnej funkcji bazowej *siatki pomocniczej* związany z jej  $i$ -tym faktycznym węzłem wierzchołkowym. Rys. 6.1 przedstawia przykład siatki adaptacyjnej, odpowiadającej jej siatki pomocniczej oraz kilka podobszarów  $\Omega_i$ . W przypadku siatki równomiernej (być może z wydłużonymi elementami) siatka pomocnicza jest też równomierna i podobszary są identyczne.

Poniżej przedstawiamy kilka szczegółów implementacji algorytmu. Należy podkreślić, że dla danej siatki jest on wywoływany tylko jednokrotnie, więc czas jego wykonania jest w praktyce zaniedbywalny.

1. Dane definiujące siatkę pomocniczą są tworzone w oddzielnej strukturze danych analogicznej do struktury danych programu adaptacyjnego. Jednakże aby ograniczyć rozmiar tablic używanych w typach *siatki pomocnicza* jest generowana oddzielnie



Rysunek 6.1: Siatka adaptacyjna wraz z siatką zgrubną oraz przykładowe podobszary.

na każdym elemencie siatki pierwotnej. Następnie dokonuje się całościowego sprawdzenia, czy zasada „2-1” jest spełniona wzdłuż brzegów elementów pierwotnych. Jeśli nie – dokonuje się odpowiednich korekt siatki pomocniczej.

2. Rzeczywiste siatki mogą być krzywoliniowe. Dlatego też przez rozmiary elementów-podobszarów i faktycznych elementów rozumiemy tu ich rozmiary na kwadratowym elemencie wzorcowym odpowiadającym elementowi pierwotnemu. Co więcej, aby uniknąć porównań zmiennych rzeczywistych wprowadzamy na elemencie wzorcowym współrzędne całkowite:  $ix, iy \in \{0, 1, \dots, 2^l\}$ , gdzie  $l$  przekracza maksymalny przewidywany poziom podziałów. Podczas łamania elementów-podobszarów ich pozycja jest określana przez 4 liczby całkowite:  $ix_{min}, ix_{max}, iy_{min}, iy_{max}$ . Także dla rzeczywistych elementów siatki adaptacyjnej informacje zawarte w strukturze danych dostarczają ich lokalizacji na elemencie pierwotnym  $ixe_{min}, ixe_{max}, iye_{min}, iye_{max}$ .
3. Algorytm generujący siatkę pomocniczą na danym elemencie początkowym jest pokazany na rys. 6.2, gdzie przyjęto oznaczenia:

$listwait(\star)$  - lista podobszarów będących kandydatami do następnych podziałów,

$listsub(\star)$  - lista ostatecznych elementów podobszarów, które już nie będą dzielone.

## 6.8 Definicja $V_0^h$

Przedstawiona metoda generacji podobszarów  $\Omega_i$  pozwala zdefiniować przestrzeń skończenie elementową  $V_0^h \subset V^h$ . Definiujemy  $V_0^h$  jako przestrzeń rozpiętą przez globalne bazowe funkcje kształtu  $\theta_i$  odpowiadające siatce pomocniczej – dalej zwanej siatką zgrubną. Dokładniej mówiąc,  $\theta_i$  otrzymamy przez złożenie związanych biliniowych funkcji kształtu elementów-podobszarów, które mają wspólny rzeczywisty węzeł  $\mathbf{x}_i$  siatki zgrubnej. Oczywiście  $supp(\theta_i) = \Omega_i$ ,  $i = 1, \dots, N$ .

## 6.9 Istnienie rozkładu o niskiej energii

Założenie o quasi-równomierności siatki, które może nie być spełnione przez siatki otrzymane drogą podziałów anizotropowych, jest wykorzystywane w dowodzie drugiego rezultatu Twierdzenia 6.5.2 mówiącego o istnieniu „rozkładu o niskiej energii” (6.5.45). W niniejszym podrozdziale przedstawiamy szkic oryginalnego rozumowania przedstawionego w pracy Dryi i Widlunda [39] (a także Lionsa [64] i Nepomnyaschikh [76]) oraz wskazujemy odpowiednie uogólnienia dla elementów wydłużonych i z więzami.

Dla podobszarów  $\Omega_i$  zdefiniowanych na siatce zgrubnej proporcja (rozmiar  $H_i$  podobszaru / grubość  $\delta_i$  zachodzenia na siebie  $\Omega_i$  i jego sąsiadów) zmienia się w ustalonych granicach, np. dla siatek równomiernych wynosi 2. Dlatego też w naszym uogólnieniu Twierdzenia 6.5.2 będziemy starać się pokazać tylko ograniczoność prawej strony wzoru (6.5.55) bez badania wpływu  $H_i/\delta_i$ .

### 6.9.1 Siatki quasi-równomierne

Chcemy skonstruować rozkład dowolnej funkcji  $\mathbf{u} \in V^h$ ,  $\mathbf{u} = \mathbf{z}_0 + \mathbf{z}_1 + \dots + \mathbf{z}_N$ ,  $\mathbf{z}_i \in V_i^h$ , dla którego  $\sum_{i=0}^N b_i(\mathbf{z}_i, \mathbf{z}_i) \leq C_{0d}(\mathbf{u}, \mathbf{u})$ , „rozkład o niskiej energii”. Konstrukcja

1. Weź następnego kandydata do złamania  $K$  z  $listwait(*)$  (o ile taki kandydat jeszcze istnieje); określ jego rozmiar  $id_K$ .
2. Znajdź minimalny i maksymalny rozmiar rzeczywistych elementów zawartych w  $K$ ,  $id_{min}$  i  $id_{max}$ .
3. IF( $id_{max}/id_K = q_1$  lub  $id_{min}/id_K = q_2$ )  
 THEN  
     przenieś  $K$  z  $listwait$  na  $listsub$ ,  
     skocz do 1.  
 ELSE  
     IF( $K$  ma „dużego sąsiada”  $L$ )  
     THEN  
         IF( $L$  jest na  $listsub$  (tzn. nie może być dzielony))  
         THEN  
             przenieś  $K$  z  $listwait$  do  $listsub$ ,  
             skocz do 1.  
         ELSE  
             skocz do 1. ( $L$  ciągle może być złamany)  
         ENDIF  
     ELSE  
         złam  $K$  na 4 „synów”,  
         zapisz ich w  $listwait$ ,  
         skocz do 1.  
     ENDIF  
 ENDIF  
 ENDIF

Rysunek 6.2: Algorytm generowania pomocniczej siatki zgrubnej.

jest dwustopniowa. Po pierwsze rzutujemy  $\mathbf{u}$  na  $V_0^h$  stosując operator  $Q_H$ , otrzymując  $\mathbf{z}_0 := Q_H \mathbf{u}$ . Następnie rozkładamy pozostałą część  $\mathbf{u}$ ,  $\mathbf{w} := \mathbf{u} - Q_H \mathbf{u}$ . W tym celu wykorzystujemy rozkład jedności związany z funkcjami kształtu siatki zgrubnej:

$$1 = \sum_{i=1}^N \theta_i(\mathbf{x}) \quad (6.9.62)$$

Funkcje  $\theta_i$  spełniają warunki  $|\theta_i| \leq 1$  i  $|\nabla \theta_i|_{\infty, \Omega_i} \leq C/H_i$ ,  $i = 1, \dots, N$ , z  $H_i$  oznaczającym rozmiar  $\Omega_i = \text{supp}(\theta_i)$ ,  $C > 0$  jest niezależne od  $i$ . Definiujemy rozkład następująco:

$$\mathbf{z}_i = I_h(\theta_i \mathbf{w}), \quad i = 1, \dots, N, \quad (6.9.63)$$

gdzie  $I_h$  jest klasycznym lagranżowskim operatorem interpolacji (dla drobnej siatki). Oczywiście  $\text{supp}(\theta_i \mathbf{w}) \subset \Omega_i$  zaś  $I_h$  został zastosowany dla otrzymania  $\mathbf{z}_i$  jako funkcji kształtu, czyli aby  $\mathbf{z}_i \in V_i^h$ . W obecnej wersji dowodu zakładamy, że siatka jest regularna (bez więzów). Zdefiniowana wyżej dekompozycja jest poprawna:  $\mathbf{z}_0 + \dots + \mathbf{z}_N = \mathbf{z}_0 + \sum_{i=1}^N I_h(\theta_i \mathbf{w}) = \mathbf{z}_0 + I_h(\sum_{i=1}^N \theta_i \mathbf{w}) = \mathbf{z}_0 + I_h(\mathbf{u} - \mathbf{z}_0) = \mathbf{u}$  gdzie wykorzystano (6.9.62) i liniowość  $I_h$ .

Aby pokazać (6.5.45) udowodnimy w pierw wersję tej nierówności z formami  $a(\cdot, \cdot)$  i  $b(\cdot, \cdot)$  zastąpionymi przez równoważną im formę półnormy  $H^1$ . Oznaczmy przez  $\bar{\theta}_i$  średnią wartość  $\theta_i$  w elemencie  $K \subset \Omega_i$ . Możemy napisać, że  $\mathbf{z}_i = I_h(\theta_i \mathbf{w}) = \bar{\theta}_i \mathbf{w} + I_h((\theta_i - \bar{\theta}_i) \mathbf{w})$ , gdyż  $I_h$  jest liniowy i  $\mathbf{w} \in V^h$ . Rozkład taki pociąga nierówność:

$$\int_K |\nabla \mathbf{z}_i|^2 dx \leq 2\bar{\theta}_i^2 \int_K |\nabla \mathbf{w}|^2 dx + 2 \|\nabla I_h((\theta_i - \bar{\theta}_i) \mathbf{w})\|_{L^2(K)}^2 \quad (6.9.64)$$

Oszacujemy drugi z wyrazów po prawej stronie (6.9.64). Element  $K$  możemy rozważać jako obraz elementu wzorcowego  $\hat{K} = [0, 1]^2$  w odwzorowaniu  $\phi$ ,  $K = \phi(\hat{K})$ . Jeśli spełnia ono założenia quasi-równomierności (4.1.5), to możemy napisać

$$\int_K (\nabla I_h((\theta_i - \bar{\theta}_i) \mathbf{w}))^2 dx = \quad (6.9.65)$$

$$\begin{aligned} & \int_{\hat{K}} \sum_{j=1}^2 \left\{ \sum_{k=1}^2 \frac{\partial}{\partial \xi_k} [I_h((\theta_i - \bar{\theta}_i) \mathbf{w}) \circ \phi] \frac{\partial \xi_k}{\partial x_j} \right\}^2 \frac{\partial(x_1, x_2)}{\partial(\xi_1, \xi_2)} d\xi \leq \\ & C \|\nabla_{\xi} [I_h((\theta_i - \bar{\theta}_i) \mathbf{w}) \circ \phi]\|_{L^2(\hat{K})}^2 \leq \\ & C \|I_h((\theta_i - \bar{\theta}_i) \mathbf{w}) \circ \phi\|_{L^2(\hat{K})}^2 \leq \\ & C \|(\theta_i - \bar{\theta}_i) \mathbf{w} \circ \phi\|_{L^\infty(\hat{K})}^2 \leq \\ & C \frac{1}{h^2} \|\theta_i - \bar{\theta}_i\|_{L^\infty(K)}^2 \cdot h^2 \|\mathbf{w} \circ \phi\|_{L^2(\hat{K})}^2 \leq \\ & C \|\nabla \theta_i\|_{L^\infty(K)}^2 \cdot \|\mathbf{w}\|_{L^2(K)}^2 \leq \frac{C}{H_i^2} \|\mathbf{w}\|_{L^2(K)}^2 \end{aligned}$$

Pierwsza nierówność wynika z własności (4.1.5)  $\phi$ , druga z równoważności półnormy  $H^1$  i normy  $L^2$  na przestrzeni elementowych funkcji kształtu  $V^h(K)$ , różnych od stałej (tj.  $V^h(K)/R$ ). Trzecia nierówność jest konsekwencją definicji  $I_h$ :  $I_h: (I_h \mathbf{u}) \circ \phi = \sum_{i=1}^m I_h(\mathbf{u}(\mathbf{x}_i)) \psi_i$ , gdzie  $\psi_i$  są elementowymi funkcjami kształtu na  $\hat{K}$ ,  $\mathbf{x}_i$  węzłami elementu  $K$ . W końcu czwarty krok wynika z równoważności normy  $L^2$  i  $L^\infty$  na przestrzeni funkcji kształtu  $V^h(\hat{K})$  a ostatnie przejście oparte jest na ograniczeniu  $|\nabla \theta_i|_{\infty, \Omega_i} \leq C/H_i$ .



Sumując nierówności (6.9.64) po wszystkich elementach  $K$  składających się na podobszar  $\Omega_i$ , wykorzystując (6.9.65) i  $|\theta_i| \leq 1$  otrzymujemy

$$\int_{\Omega_i} |\nabla z_i|^2 dx \leq 2 \int_{\Omega_i} |\nabla w|^2 dx + \frac{2C}{H_i^2} \int_{\Omega_i} |w|^2 dx, \quad (6.9.66)$$

Ponieważ zaś  $w = u - Q_H u$  mamy dalej

$$\begin{aligned} \int_{\Omega_i} |\nabla z_i|^2 dx &\leq 2 \int_{\Omega_i} |\nabla(u - Q_H u)|^2 dx + \frac{C}{H_i^2} \int_{\Omega_i} |u - Q_H u|^2 dx \leq \\ &4 \int_{\Omega_i} |\nabla u|^2 dx + \max(4, C) \left( \int_{\Omega_i} |\nabla Q_H u|^2 dx + \frac{1}{H_i^2} \int_{\Omega_i} |u - Q_H u|^2 dx \right) \end{aligned} \quad (6.9.67)$$

skąd po zastosowaniu założonej własności (6.5.51) operatora  $Q_H$  otrzymujemy:

$$\int_{\Omega_i} |\nabla z_i|^2 dx \leq C \int_{\Omega_i} |\nabla u|^2 dx \quad (6.9.68)$$

Ponieważ  $\int_{\Omega_i} |\nabla z_0|^2 = \int_{\Omega_i} |\nabla Q_H u|^2 dx \leq C \int_{\Omega_i} |\nabla u|^2 dx$  (ze względu na (6.5.51)) oraz biorąc pod uwagę założoną równoważność form  $a, b$  i półnormy  $H^1$ , możemy napisać, że:

$$\sum_{i=0}^N b_i(z_i, z_i) \leq \frac{1}{\omega_-} \sum_{i=0}^N \int_{\Omega_i} |\nabla z_i|^2 dx \leq \frac{C}{\omega_-} \sum_{i=0}^N \int_{\Omega_i} |\nabla u|^2 dx \leq \frac{Ck_0}{\alpha\omega_-} a(u, u) \quad (6.9.69)$$

( $\Omega_0 \equiv \Omega$ ) co kończy dowód.

W ostatnim kroku rozważań mogliśmy też użyć globalnej wersji założenia (6.5.51):

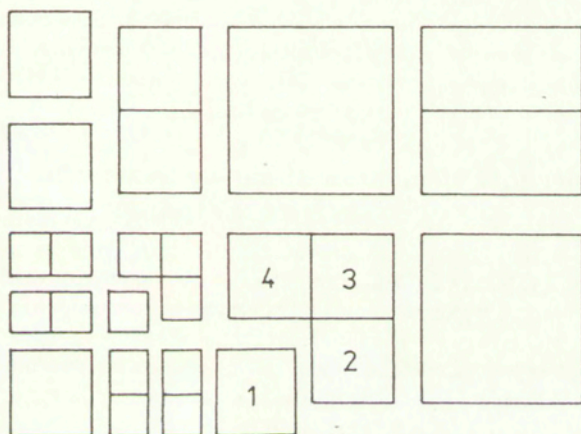
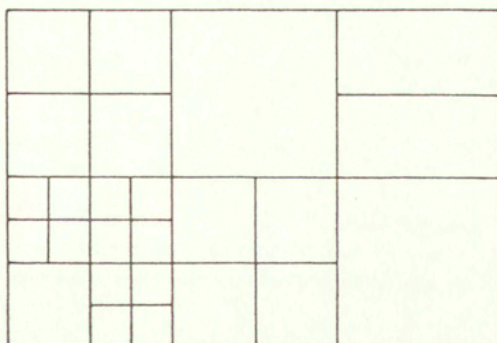
$$\int_{\Omega} |\nabla Q_H v|^2 dx + \frac{1}{H^2} \int_{\Omega} |v - Q_H v|^2 dx \leq C \int_{\Omega} |\nabla v|^2 dx, \quad \forall v \in V^h \quad (6.9.70)$$

z  $H = \max_i H_i$ . Stosując (6.9.70) i zakładając dodatkowo, że siatka zgrubna jest quasi-równomierna, tj.  $H_i \geq CH$ , otrzymujemy:

$$\begin{aligned} \sum_{i=0}^N b_i(z_i, z_i) &\leq \frac{1}{\omega_-} \left( \int_{\Omega} |\nabla z_0|^2 dx + \sum_{i=1}^N \int_{\Omega_i} |\nabla z_i|^2 dx \right) \leq \\ &\frac{1}{\omega_-} \left\{ C \int_{\Omega} |\nabla u|^2 dx + \right. \\ &\left. \sum_{i=1}^N \left[ 4 \int_{\Omega_i} |\nabla u|^2 dx + C_1 \left( \int_{\Omega_i} |\nabla Q_H u|^2 dx + \frac{1}{H_i^2} \int_{\Omega_i} |u - Q_H u|^2 dx \right) \right] \right\} \leq \\ &\frac{1}{\omega_-} \left\{ C \int_{\Omega} |\nabla u|^2 dx + k_0 \left[ 4 \int_{\Omega} |\nabla u|^2 dx + \right. \right. \\ &\quad \left. \left. C_1 \max(1, C_0^{-2}) \left( \int_{\Omega} |\nabla Q_H u|^2 dx + \frac{1}{H^2} \int_{\Omega} |u - Q_H u|^2 dx \right) \right] \right\} \leq \\ &\frac{1}{\omega_-} (C + k_0 C_2) \int_{\Omega} |\nabla u|^2 dx \leq \frac{C + k_0 C_2}{\omega_- \alpha} a(u, u) \end{aligned} \quad (6.9.71)$$

## 6.9.2 Siatki z więzami

Obecność więzów ingeruje w sposób wykazania nierówności (6.9.65). Aby uogólnić tę argumentację, musimy rozważyć uogólniony operator interpolacji  $I_h^c$  (rozdział 4.1), co jest



$$J_1 = \{2, 3, 4\}$$

Rysunek 6.3: Grupy elementów z więzami.

konieczne, aby zagwarantować, że  $\mathbf{z}_i = I_h^c(\theta_i, \mathbf{w}) \in V^h$ . Musimy też rozważyć wyrażenie (6.9.64) nie dla poszczególnych elementów lecz dla grup elementów z więzami,  $\Sigma := \cup_{K \in J_k} K$ , które przylegają do wspólnych dużych elementów-sąsiadów, rys. 6.3. Przyczyną tego jest fakt, że  $I_h^c$  wykorzystuje wartości interpolowanej funkcji w faktycznych węzłach, z których część może być położona poza elementem – na jego wydłużonych bokach. Przez rozważanie wspomnianych grup elementów sprawiamy, że interpolant funkcji  $v$ ,  $v_I = I_h^c v$  na grupie elementów zależy wyłącznie od wartości  $v$  w  $\Sigma$ , jak dla regularnych siatek.

Z powyższych względów zapiszmy (6.9.64) dla  $\Sigma$ :

$$\sum_{K \in J_k} \int_K |\nabla \mathbf{z}_i|^2 dx \leq 2 \sum_{K \in J_k} \int_K |\nabla \mathbf{w}|^2 dx + 2 \sum_{K \in J_k} \|\nabla I_h^c((\theta_i - \bar{\theta}_i)\mathbf{w})\|_{L^2(K)}^2, \quad (6.9.72)$$

z  $\bar{\theta}_i$  będącym średnią wartością  $\theta_i$  w  $\Sigma$ ,  $J_k$  oznacza tu listę elementów składających się na  $\Sigma$ . Jak i poprzednio chcemy oszacować drugi wyraz z prawej strony (6.9.72). Elementy tworzące  $\Sigma$  są wynikiem podziału ich wspólnego „elementu-ojca”  $K_0$ . Dlatego też odwzorowanie  $\phi: K_0 = \phi(\hat{K})$ , z  $\hat{K} = [0, 1]^2$  może zostać użyte do zdefiniowania geometrii wszystkich elementów z listy  $J_k$ . Jediną różnicą jest to, że elementy  $K \in J_k$  są obrazami odpowiednich ćwiartek (lub połówek)  $\hat{K}$ ,  $\phi$  jednak zachowuje własności (4.1.5).

Sekwencja nierówności (6.9.65) jest prawdziwa też w obecnej sytuacji z  $K$  zastąpionym przez  $\Sigma$ . Wynika to z własności  $\phi$  jak również z równoważności różnych norm na skończenie wymiarowej przestrzeni funkcji kształtu na  $\Sigma$ , z definicji  $I_h^c$  zawierającej wartości interpolowanej funkcji na  $\Sigma$ . Jedyne wartości stałych w oszacowaniu ulega zmianie, gdyż obszarami odniesienia dla elementów  $K$  są odpowiednie części elementu  $[0, 1]^2$  i ponieważ związane funkcje kształtu zastępują zwyczajnie. Tak więc nierówność (6.9.65) jest spełniona a pozostała część rozumowania pozostaje bez zmian.

### 6.9.3 Obecność elementów silnie wydłużonych

Jak mogliśmy zauważyć, otrzymanie oszacowania (6.9.65) było możliwe dzięki wykorzystaniu klasycznych własności odwzorowania  $\phi$  z  $\hat{K}$  do  $K$ , które są spełnione dla siatek z ograniczonym wydłużeniem elementów. Poniżej otrzymamy to oszacowanie dla szerszej klasy odwzorowań  $\phi$  opisanych dwoma niezależnymi parametrami  $h_1$  i  $h_2$  – wymiarami elementu w dwu kierunkach. W pierwszym kroku rozumowanie ograniczymy do odwzorowań afinicznych, następnie przedstawimy uogólnienie dla odwzorowań nieliniowych.

Zalóżmy że element  $K$  i odwzorowanie  $\phi$  są dane następująco:

$$\begin{aligned} \phi: [0, 1] \times [0, 1] &\rightarrow [0, h_1] \times [0, h_2] \\ x_1 &= h_1 \xi_1, \quad x_2 = h_2 \xi_2 \end{aligned} \quad (6.9.73)$$

Oznaczając  $\beta = \theta_i - \bar{\theta}_i$  możemy przedstawić lewą stronę (6.9.65) następująco:

$$\begin{aligned} \int_K |\nabla I_h(\beta \mathbf{w})|^2 dx &= \int_{\hat{K}} \sum_{i=1}^2 \left[ \frac{1}{h_i} \frac{\partial}{\partial \xi_i} (I_h(\beta \mathbf{w}) \circ \phi) \right]^2 h_1 h_2 d\xi = \\ &= \frac{h_2}{h_1} \int_{\hat{K}} \left[ \frac{\partial}{\partial \xi_1} \hat{I}_h(\hat{\beta} \hat{\mathbf{w}}) \right]^2 d\xi + \frac{h_1}{h_2} \int_{\hat{K}} \left[ \frac{\partial}{\partial \xi_2} \hat{I}_h(\hat{\beta} \hat{\mathbf{w}}) \right]^2 d\xi, \end{aligned} \quad (6.9.74)$$

gdzie  $\hat{\beta} = \beta \circ \phi$ ,  $\hat{\mathbf{w}} = \mathbf{w} \circ \phi$ ,  $\hat{I}_h \hat{\mathbf{u}}(\xi) = \sum_{i=1}^4 \hat{\mathbf{u}}(\xi_i) \hat{\psi}_i(\xi)$  z  $\xi_i$  będącymi węzłami,  $\hat{\psi}_i$  bilingwijnymi funkcjami kształtu na  $\hat{K}$ . Weźmy pod uwagę pierwszy składnik prawej strony

(6.9.74),  $I_1$ . Moglibyśmy postąpić jak w (6.9.65) pozbywając się różniczkowania i  $\hat{I}_h$  z analizowanego wyrażenia opierając się na argumentach jak poprzednio. Tym razem jednak otrzymalibyśmy proporcję  $h_1/h_2$  w końcowym wyniku. Aby tego uniknąć nasze rozumowanie będzie bardziej subtelne. Zastąpimy różniczkowanie przez iloraz różnicowy, który jest tu ścisły, gdyż różniczkowana funkcja jest liniowa:

$$\begin{aligned} \frac{\partial}{\partial \xi_1} \hat{I}_h(\hat{\beta} \hat{w}) &= \hat{\beta}(1, \cdot) \hat{w}(1, \cdot) - \hat{\beta}(0, \cdot) \hat{w}(0, \cdot) = \\ &= \frac{1}{2}(\hat{\beta}(1, \cdot) + \hat{\beta}(0, \cdot))(\hat{w}(1, \cdot) - \hat{w}(0, \cdot)) + \frac{1}{2}(\hat{\beta}(1, \cdot) - \hat{\beta}(0, \cdot))(\hat{w}(1, \cdot) + \hat{w}(0, \cdot)) \end{aligned} \quad (6.9.75)$$

co przy zastosowaniu oznaczeń  $\hat{w}(a, b) = \hat{w}_{ab}$ ,  $\hat{\beta}(a, b) = \hat{\beta}_{ab}$ , daje nierówność:

$$\begin{aligned} I_1 \leq & \quad (6.9.76) \\ & h_1 h_2 \int_{\hat{K}} \left\{ \left[ \frac{(\hat{\beta}_{11} - \hat{\beta}_{01})^2}{h_1^2} (\hat{w}_{11} + \hat{w}_{01})^2 + (\hat{\beta}_{11} + \hat{\beta}_{01})^2 \frac{(\hat{w}_{11} - \hat{w}_{01})^2}{h_1^2} \right] \xi_2^2 + \right. \\ & \left. \left[ \frac{(\hat{\beta}_{10} - \hat{\beta}_{00})^2}{h_1^2} (\hat{w}_{10} + \hat{w}_{00})^2 + (\hat{\beta}_{10} + \hat{\beta}_{00})^2 \frac{(\hat{w}_{10} - \hat{w}_{00})^2}{h_1^2} \right] (1 - \xi_2)^2 \right\} d\xi \end{aligned}$$

Ponieważ  $|(\hat{\beta}(1, \cdot) - \hat{\beta}(0, \cdot))/h_1| \leq \|\partial \beta / \partial x_1\|_{L^\infty(K)}$  a także  $|(\hat{w}(1, \cdot) - \hat{w}(0, \cdot))/h_1| \leq \|\partial w / \partial x_1\|_{L^\infty(K)}$  i  $|\hat{\beta}| \leq 1$  z (6.9.76) wynika, że

$$I_1 \leq C \left( \left\| \frac{\partial \beta}{\partial x_1} \right\|_{L^\infty(K)}^2 \|w\|_{L^2(K)}^2 + \left\| \frac{\partial w}{\partial x_1} \right\|_{L^2(K)}^2 \right) \quad (6.9.77)$$

Pisząc analogiczne oszacowanie dla pozostałego składnika (6.9.74) znajdujemy, że

$$\int_K |\nabla I_h(\beta w)|^2 dx \leq C (\|\nabla \theta_i\|_{L^\infty(K)}^2 \|w\|_{L^2(K)}^2 + \|\nabla w\|_{L^2(K)}^2) \quad (6.9.78)$$

co z kolei możemy wykorzystać w oszacowaniu (6.9.64) oryginalnego dowodu

$$\int_K |\nabla z_i|^2 dx \leq C_1 \int_K |\nabla w|^2 dx + \frac{C_2}{H_i^2} \|w\|_{L^2(K)}^2 \quad (6.9.79)$$

Rozważmy obecnie sytuację ogólną, z wydłużonymi elementami  $K$  będącymi rezultatem anizotropowych podziałów elementu pierwotnego  $K_0 = F([0, 1]^2)$ , z nieliniowym odwzorowaniem  $F$ . W tym przypadku  $K = F(\hat{K})$  gdzie  $\hat{K} \subset [0, 1]^2$  jest prostokątnym wydłużonym elementem. Wobec tego, jeśli przyjąć oznaczenie  $\hat{v} = v \circ F$ , to (6.9.78) można zastosować dla elementu  $\hat{K}$ :

$$\|\nabla \hat{I}_h(\hat{\beta} \hat{w})\|_{L^2(\hat{K})}^2 \leq C (\|\nabla \hat{\theta}_i\|_{L^\infty(\hat{K})}^2 \|\hat{w}\|_{L^2(\hat{K})}^2 + \|\nabla \hat{w}\|_{L^2(\hat{K})}^2) \quad (6.9.80)$$

Zakładane ograniczenie pochodnych  $F$ ,  $\|\nabla F\|_\infty \leq C_1 H$ ,  $\|\nabla F^{-1}\|_\infty \leq C_2/H$  pociąga następujące nierówności:

$$\|\nabla \hat{v}\|_{L^\infty(\hat{K})} \leq C H \|\nabla v\|_{L^\infty(K)} \quad (6.9.81)$$

$$\|\hat{v}\|_{L^2(\hat{K})} \leq C/H \|v\|_{L^2(K)} \quad (6.9.82)$$

$$c \|\nabla \hat{v}\|_{L^2(\hat{K})} \leq \|\nabla v\|_{L^2(K)} \leq C \|\nabla \hat{v}\|_{L^2(\hat{K})} \quad (6.9.83)$$

dla danej funkcji  $v$  z  $W^{1,\infty}(K)$ ,  $L^2(K)$  lub  $H^1(K)$ . Stosując (6.9.81), (6.9.82) i (6.9.83) w (6.9.80) otrzymujemy:

$$\begin{aligned} \|\nabla I_h(\beta \mathbf{w})\|_{L^2(K)}^2 &\leq C(\|\nabla \hat{\theta}_i\|_{L^\infty(\hat{K})}^2 \|\hat{\mathbf{w}}\|_{L^2(\hat{K})}^2 + \|\nabla \hat{\mathbf{w}}\|_{L^2(\hat{K})}^2) \leq & (6.9.84) \\ &C(\|\nabla \theta_i\|_{L^\infty(K)}^2 \|\mathbf{w}\|_{L^2(K)}^2 + \|\nabla \mathbf{w}\|_{L^2(K)}^2) \end{aligned}$$

czyli pożądane oszacowanie (6.9.77).

**Uwaga 6.9.2** Przedstawione wyżej rozumowanie stosuje się także do przypadku ogólnego,  $N$ -wymiarowych wydłużonych elementów (spłaszczonych itd.), prostopadłościanów o wymiarach  $h_1, \dots, h_N$  takich że  $K = \phi(\hat{K})$ ,  $\hat{K} = [0, 1]^N$ .

$$\phi : [0, 1]^N \rightarrow [0, h_1] \times [0, h_2] \times \dots \times [0, h_N], \quad x_i = h_i \xi_i, \quad i = 1, \dots, N \quad (6.9.85)$$

Dla takich elementów z aproksymacją  $N$ -liniową składnik  $I_1$  wzoru (6.9.74) ma postać:

$$I_1 = \frac{h_1 \dots h_N}{h_1^2} \int_{\hat{K}} \left\{ \sum_I [\hat{\beta}(1, \boldsymbol{\xi}^I) \hat{\mathbf{w}}(1, \boldsymbol{\xi}^I) - \hat{\beta}(0, \boldsymbol{\xi}^I) \hat{\mathbf{w}}(0, \boldsymbol{\xi}^I)] \hat{\chi}_I(\boldsymbol{\xi}) \right\}^2 d\boldsymbol{\xi} \quad (6.9.86)$$

gdzie  $I$  jest wielowskazywaniem  $I = (i_2, i_3, \dots, i_N)$ ,  $i_k = 0, 1$  dla  $k = 2, \dots, N$ , oraz

$$\begin{aligned} \hat{\beta}(\cdot, \boldsymbol{\xi}^I) &= \hat{\beta}(\cdot, \xi_2^{i_2}, \dots, \xi_N^{i_N}), & (6.9.87) \\ \hat{\mathbf{w}}(\cdot, \boldsymbol{\xi}^I) &= \hat{\mathbf{w}}(\cdot, \xi_2^{i_2}, \dots, \xi_N^{i_N}), \quad \hat{\chi}_I(\boldsymbol{\xi}) = \hat{\chi}_{i_2} \dots \hat{\chi}_{i_N}(\xi_N). \end{aligned}$$

z  $\hat{\chi}_0(\xi) = 1 - \xi$ ,  $\hat{\chi}_1(\xi) = \xi$  oznaczającymi jednowymiarowe funkcje kształtu. Zastosowanie (6.9.75) do funkcji podcałkowej (6.9.86) (z  $\hat{\beta}$ ,  $\hat{\mathbf{w}}$  zastąpionymi przez  $\hat{\tilde{\beta}}$ ,  $\hat{\tilde{\mathbf{w}}}$ ) prowadzi do oszacowania (6.9.77).

## 6.10 Uogólnienie dla formy dwuliniowej symetrycznego sformułowania zadania przepływu

Forma dwuliniowa (3.1.6) odpowiadająca zadaniu przepływu różni się istotnie od rozważanej dotąd formy (6.4.20) regularnego symetrycznego operatora eliptycznego:

1. Forma dwuliniowa  $\mathbf{a}_{ij} := \Delta t(\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij} \mathbf{I})$  jest niesymetryczna,  $\mathbf{a}_{ij} \neq \mathbf{a}_{ji}^T$ .
2. Forma  $\mathbf{a}_{ij}$  nie jest jednostajnie dodatnio określona w znaczeniu, że nie jest spełniony warunek:  $\sup_{\mathbf{v}_i} \mathbf{u}_i^T \mathbf{a}_{ij} \mathbf{v}_j / (\mathbf{v}_i^T \mathbf{v}_i)^{1/2} > \alpha (\mathbf{u}_i^T \mathbf{u}_i)^{1/2} \forall \mathbf{u}_i$ , dla ustalonego  $\alpha > 0$ . Przyczyną tego jest fakt, że rząd macierzy  $\mathbf{K}$  złożonej z bloków  $\mathbf{K}_{ij}$  (tj.  $\nabla \mathbf{U}^T \mathbf{K} \nabla \mathbf{V} = \mathbf{U}_i^T \mathbf{K}_{ij} \mathbf{V}_j, \forall \mathbf{U}_i, \mathbf{V}_j$ ) jest równy 5 w przestrzeni  $R^8$ . Poza tym wektory  $\sum_{i=1}^2 \mathbf{A}_i \mathbf{U}_i$  mogą zniknąć dla pewnych  $\mathbf{U}_i \neq 0$ . Co więcej, współczynniki  $\tau_a$  i  $\nu_d$  wyrazów stabilizujących mogą być dowolnie bliskie zera – w istocie chcielibyśmy, by ich wartość spadła znacznie poniżej rzędu wielkości  $\mathbf{K}_{ij}$ .

Powyższe własności nie pozwalają automatycznie uogólnić przedstawionych wcześniej wyników dotyczących form regularnych operatorów eliptycznych na obecny przypadek. Nie możemy też skorzystać z teorii Cai i Widlunda dla problemów niesymetrycznych [22],[23], gdyż rozważa się w niej wyłącznie niesymetrię wywołaną obecnością wyrazów 1 rzędu. O ile nam wiadomo, przypadek niesymetrii  $\mathbf{a}_{ij}$  nie był dotychczas rozważany. Posługując się istniejącymi wynikami możemy podać analizę algorytmu uwarunkowania wstępnego dla alternatywnego do (3.1.2) symetrycznego sformułowania metody SUPG w zmiennych entropowych. W tym przypadku trudnością, która pozostaje, jest brak jednostajnej dodatniej określoności składnika odpowiadającego wyrazom 2 rzędu.

Równania dyskretne dla zmiennych entropowych otrzymuje się analogicznie do (3.1.4). Forma dwuliniowa tego sformułowania ma postać:

$$B(\mathbf{U}, \mathbf{V}) = \int_{\Omega} \mathbf{V}^T \mathbf{A}_0 \mathbf{U} dx + \Delta t \int_{\Omega} \mathbf{V}_i^T (\bar{\mathbf{K}}_{ij} + \bar{\mathbf{A}}_i^T \bar{\tau} \bar{\mathbf{A}}_j + \bar{\nu}_d \delta_{ij} \mathbf{A}_0) \mathbf{U}_j dx \quad (6.10.88)$$

W (6.10.88) macierze lepkości  $\bar{\mathbf{K}}_{ij}$  i jacobiany  $\bar{\mathbf{A}}_i$  odpowiadają strumieniom lepkim i eulerowskim równań Naviera–Stokesa w zmiennych entropowych (zob. rozdział 2.2). Symetryczny dodatnio określony operator  $\bar{\tau}$  i parametr  $\bar{\nu}_d$  zależą od siatki skończenie elementowej i aktualnego rozwiązania (szczegółowo można znaleźć w pracy [98]).

W obecnym zadaniu ograniczymy się do pokazania „dekompozycji o niskiej energii” w postaci:

$$\sum_{i=1}^N B(u_i, u_i) \leq C_0 B(u, u) \quad (6.10.89)$$

tj. bez kroku globalnego i bez modyfikacji formy dwuliniowej uwarunkowania wstępnego. Konstrukcja rozkładu jest analogiczna do rozważanej poprzednio z pominięciem składnika globalnego  $\mathbf{z}_0$ :

$$\mathbf{U} = \sum_{i=1}^N \mathbf{U}_i, \quad \mathbf{U}_i = I_h(\theta_i \mathbf{U}), \quad i = 1, \dots, N \quad (6.10.90)$$

gdzie  $I_h$  jest operatorem interpolacji,  $\theta_i(\mathbf{x})$  są bazowymi funkcjami kształtu na siatce grubej.

W przedstawianej analizie wykorzystamy następujące własności macierzy definiujących  $B$  w (6.10.88):

1. Symetryzator  $\mathbf{A}_0 = \mathbf{A}_0^T$  jest jednostajnie ograniczony z góry i jest jednostajnie dodatnio określony w  $\Omega$ .
2. Macierz  $\bar{\tau} = \bar{\tau}^T \geq 0$  i współczynnik  $\bar{\nu}_d \geq 0$  są jednostajnie ograniczone w  $\Omega$ .
3. Macierz  $\bar{K} \in R^{8 \times 8}$  zdefiniowana przez  $\nabla U^T \bar{K} \nabla V = U_i^T \bar{K}_{ij} V_j$  dla wszystkich  $U_i, V_j$  jest symetryczna dodatnio półokreślona,  $\dim(\ker(\bar{K})) = 3$  w  $R^8$ . Dla wektorów  $W \in (\ker(\bar{K}))^\perp$  forma kwadratowa  $W^T \bar{K} W$  jest jednostajnie dodatnio określona, tzn. że najmniejsza niezerowa wartość własna  $\bar{K}$  jest jednostajnie ograniczona od zera. (W sformułowaniu stosującym zmienne naturalne, gęstość, prędkości i temperaturę, na  $\bar{K}$  składają się współczynniki operatorów liniowej teorii sprężystości i operatora Laplace'a z parametrami  $\mu$  i  $\kappa$  jednostajnie ograniczonymi w  $\Omega$ .)

Rozważmy osobno nierówność (6.10.89) dla poszczególnych form dwuliniowych składających się na  $B(\cdot, \cdot)$ . Oszacowanie

$$\sum_i \int_{\Omega_i} I_h(\theta_i U)^T \mathbf{A}_0 I_h(\theta_i U) dx \leq C \int_{\Omega} U^T \mathbf{A}_0 U dx, \quad \forall U \in V^h \quad (6.10.91)$$

wynika z elementarnych własności operatora  $I_h$ :

Ze względu na własności  $\mathbf{A}_0$  i fakt, że każdy punkt  $\mathbf{x}$  należy do ograniczonej liczby podobszarów  $\Omega_i$ , (6.10.91) będzie wykazane, jeśli nierówność tę udowodnimy dla pojedynczego elementu  $K$  i normy  $L^2$ , tj. jeśli

$$\int_K |I_h(\theta_i U)|^2 dx \leq C \int_K |U|^2 dx \quad (6.10.92)$$

Aby wykazać powyższe oszacowanie, weźmy pod uwagę, że element  $K$  jest częścią pewnego elementu siatki pierwotnej  $K_0$ . Ponieważ  $K$  powstał w wyniku dzielenia  $K_0$ , można przyjąć, iż  $K$  jest obrazem prostokąta  $\hat{K} = [a_1, b_1] \times [a_2, b_2] \subset [0, 1]^2 = \hat{K}_0$  w odwzorowaniu  $F: \hat{K}_0 \rightarrow K_0$ , które spełnia warunki (4.1.4). Z kolei  $\hat{K} = \phi(\hat{K})$ ,  $\hat{K} = [0, 1]^2$ ,  $\phi: \hat{\zeta} \rightarrow \xi$ ,  $\xi_i = a_i + (b_i - a_i)\zeta_i$ ,  $i = 1, 2$ . Zamieniając obszar całkowania w całce po lewej stronie (6.10.92) z  $K$  na  $\hat{K}$  a następnie na  $\hat{K}$  otrzymujemy:

$$\begin{aligned} \int_K |I_h(\theta_i U)|^2 dx &= \quad (6.10.93) \\ \int_K \left| \sum_{j=1}^4 \theta_i(\mathbf{x}_j) U(\mathbf{x}_j) \psi_j(\mathbf{x}) \right|^2 dx &\leq C_1 H^2 \int_{\hat{K}} \left| \sum_{j=1}^4 \hat{\theta}_i(\xi_j) \hat{U}(\xi_j) \hat{\psi}_j(\xi) \right|^2 d\xi \leq \\ C_1 H^2 (b_1 - a_1)(b_2 - a_2) \int_{\hat{K}} \left| \sum_{j=1}^4 \check{\theta}_i(\zeta_j) \check{U}(\zeta_j) \check{\psi}_j(\zeta) \right|^2 d\zeta \end{aligned}$$

gdzie wykorzystano definicję  $I_h(\mathbf{x}_i, \xi_i, \zeta_i)$  są węzłami), własność (4.1.4) funkcji  $F$ , i oznaczenia:  $\hat{z} := z \circ F$ ,  $\check{z} := z \circ \phi = z \circ F \circ \phi$  dla wszelkich funkcji  $z$ . Ponieważ  $0 \leq \theta_i \leq 1$ , ostatnią całkę można ograniczyć przez wyrażenie

$$\begin{aligned} \int_{\hat{K}} 4 \sum_{j=1}^4 (\check{\theta}_i(\zeta_j) \check{\psi}_j(\zeta))^2 |\check{U}(\zeta_j)|^2 d\zeta &\leq \quad (6.10.94) \\ \max_j |\check{U}(\zeta_j)|^2 \int_{\hat{K}} 4 \sum_{j=1}^4 (\check{\theta}_i(\zeta_j) \check{\psi}_j(\zeta))^2 d\zeta &\leq C \|\check{U}\|_{L^\infty(\hat{K})}^2 \leq C \|\check{U}\|_{L^2(\hat{K})}^2 \end{aligned}$$

gdzie wykorzystaliśmy równoważność norm  $L^\infty$  i  $L^2$  na skończenie wymiarowej przestrzeni funkcji kształtu. Ostatni wynik wraz z powrotną zamianą obszaru całkowania z  $\tilde{K}$  na  $K$  prowadzi do oszacowania

$$\int_K |I_h(\theta_i \mathbf{U})|^2 dx \leq C(b_1 - a_1)(b_2 - a_2) \int_{\tilde{K}} |\tilde{\mathbf{U}}|^2 dx \leq C \int_K |\mathbf{U}|^2 dx \quad (6.10.95)$$

Rozważmy następnie kontrybucję formy kwadratowej  $\tilde{K}_{ij}$ . Używając notacji z  $\tilde{K}$  chcemy pokazać, że:

$$\sum_i \int_{\Omega_i} \nabla I_h(\theta_i \mathbf{U})^T \tilde{K} \nabla I_h(\theta_i \mathbf{U}) dx \leq C \left( \int_{\Omega} \nabla \mathbf{U}^T \tilde{K} \nabla \mathbf{U} dx + \int_{\Omega} \mathbf{U}^T \mathbf{A}_0 \mathbf{U} dx \right). \quad (6.10.96)$$

ze stałą  $C$  niezależną od siatki skończenie elementowej. Nierówność tego rodzaju udowodniliśmy poprzednio w przypadku, gdy pierwszy składnik prawej strony (6.10.96) był równoważny półnormie  $H^1$ . Główna trudność polega obecnie na tym, że jądro macierzy  $\tilde{K}$  jest niepuste. Dlatego też nie jest oczywiste, czy wyrażenie zawierające pochodne  $\nabla \mathbf{U} \in \ker(\tilde{K})$  z lewej strony (6.10.96) da się oszacować przez wyłącznie  $\|\mathbf{U}\|_{L^2(\Omega)}$ . Co więcej, poszukiwane ograniczenie ma być niezależne od wydłużenia elementów.

Na początek rozbijmy każdą z całek po  $\Omega$ , na kontrybucje z elementów  $K$ . W każdym elemencie  $K$  wyrażmy funkcję podcałkową jako  $I_h(\theta_i \mathbf{U}) = \tilde{\theta}_i \mathbf{U} + I_h(\beta \mathbf{U})$ , gdzie  $\tilde{\theta}_i$  jest średnią wartością  $\theta_i$  w  $K$ ,  $\beta = \theta_i - \tilde{\theta}_i$ , skąd otrzymamy:

$$\int_K \nabla I_h(\theta_i \mathbf{U})^T \tilde{K} \nabla I_h(\theta_i \mathbf{U}) dx \leq \quad (6.10.97)$$

$$2\tilde{\theta}_i^2 \int_K \nabla \mathbf{U}^T \tilde{K} \nabla \mathbf{U} dx + 2 \int_K \nabla I_h(\beta \mathbf{U})^T \tilde{K} \nabla I_h(\beta \mathbf{U}) dx$$

Naszym celem jest oszacowanie drugiego składnika po prawej stronie (6.10.97) przez sumę  $\int_K \nabla \mathbf{U}^T \tilde{K} \nabla \mathbf{U} dx$  i  $\|\mathbf{U}\|_{L^2(K)}^2$ . Uprościmy tu dyskusję rozważając wyłącznie elementy prostokątne o wymiarach  $h_1$  i  $h_2$  z nieograniczoną wartością  $h_1/h_2$ . Ponieważ  $I_h(\beta \mathbf{U})$  jest funkcją biliniową, jej pochodną można przedstawić w sposób ścisły przez iloraz różnicowy. Oznaczając przez  $\mathbf{U}_{11}, \mathbf{U}_{12}, \mathbf{U}_{21}, \mathbf{U}_{22}$  wartości węzłowe  $\mathbf{U}(0, 0), \mathbf{U}(0, h_2), \mathbf{U}(h_1, 0), \mathbf{U}(h_1, h_2)$  w lokalnych współrzędnych, takich że  $K = [0, h_1] \times [0, h_2]$ , i podobnie dla  $\beta_{ij}$ , znajdujemy w ten sposób, że  $\nabla I_h(\beta \mathbf{U}) = \mathbf{W}_1 + \mathbf{W}_2$ , gdzie

$$\mathbf{W}_1 = \begin{bmatrix} \frac{\mathbf{U}_{21} - \mathbf{U}_{11}}{h_1} \frac{\beta_{21} - \beta_{11}}{2} (1 - \xi_2) + \frac{\mathbf{U}_{22} - \mathbf{U}_{12}}{h_1} \frac{\beta_{22} - \beta_{12}}{2} \xi_2 \\ \frac{\mathbf{U}_{12} - \mathbf{U}_{11}}{h_2} \frac{\beta_{12} - \beta_{11}}{2} (1 - \xi_1) + \frac{\mathbf{U}_{22} - \mathbf{U}_{21}}{h_2} \frac{\beta_{22} - \beta_{21}}{2} \xi_1 \end{bmatrix} \quad (6.10.98)$$

$$\mathbf{W}_2 = \begin{bmatrix} \frac{\mathbf{U}_{21} + \mathbf{U}_{11}}{2} \frac{\beta_{21} - \beta_{11}}{h_1} (1 - \xi_2) + \frac{\mathbf{U}_{22} - \mathbf{U}_{12}}{2} \frac{\beta_{22} - \beta_{12}}{h_1} \xi_2 \\ \frac{\mathbf{U}_{12} + \mathbf{U}_{11}}{2} \frac{\beta_{12} - \beta_{11}}{h_2} (1 - \xi_1) + \frac{\mathbf{U}_{22} - \mathbf{U}_{21}}{2} \frac{\beta_{22} - \beta_{21}}{h_2} \xi_1 \end{bmatrix} \quad (6.10.99)$$

$\xi_i = x_i/h_i$ , kolejne wiersze  $\mathbf{W}_1$  i  $\mathbf{W}_2$  odpowiadają  $\partial I_h(\beta \mathbf{U})/\partial x_i$ ,  $i = 1, 2$ . Podobnie

$$\nabla \mathbf{U} = \begin{bmatrix} \frac{\mathbf{U}_{21} - \mathbf{U}_{11}}{h_1} (1 - \xi_2) + \frac{\mathbf{U}_{22} - \mathbf{U}_{12}}{h_1} \xi_2 \\ \frac{\mathbf{U}_{12} - \mathbf{U}_{11}}{h_2} (1 - \xi_1) + \frac{\mathbf{U}_{22} - \mathbf{U}_{21}}{h_2} \xi_1 \end{bmatrix} \quad (6.10.100)$$

(Podkreślmy, że w ogólności  $\mathbf{W}_1$  nie jest równoległy do  $\nabla \mathbf{U}$ , gdyż współczynniki skalujące  $\beta_{21} + \beta_{11}$  i  $\beta_{22} + \beta_{12}$  w (6.10.98) są liczbami przeciwnymi: ich suma znika, gdyż  $\beta_{ij}$  są wartościami węzłowymi funkcji dwuwymiarowej o zerowej wartości średniej).



Przedstawiony wyżej rozkład  $\nabla I_h(\beta U)$  i postać  $W_2$  implikuje, że

$$\int_K \nabla I_h(\beta U)^T \tilde{K} \nabla I_h(\beta U) dx \leq 2 \int_K W_2^T \tilde{K} W_2 dx + 2 \int_K W_1^T \tilde{K} W_1 dx \leq C \|\nabla \beta\|_{L^\infty(K)}^2 \|U\|_{L^2(K)}^2 + 2 \int_K W_1^T \tilde{K} W_1 dx \quad (6.10.101)$$

Rozważmy składnik związany z wektorem  $W_1$  w (6.10.101). Gdyby  $\tilde{K}$  było jednostajnie dodatnio określone, moglibyśmy napisać:

$$\int_K W_1^T \tilde{K} W_1 dx \leq C \|\beta\|_{L^\infty(K)}^2 \|\nabla U\|_{L^2(K)}^2 \leq C \|\beta\|_{L^\infty(K)}^2 \int_K \nabla U^T \tilde{K} \nabla U dx \quad (6.10.102)$$

kończąc dowód. Jednak, jeśli  $\tilde{K} \geq 0$  (półokreśloność) to ostatni krok nie jest prawdziwy dla  $U$  takich że  $\nabla U \in \ker(\tilde{K})$ .

**Uwaga 6.10.3** Gdyby siatka była quasi-równomierna, moglibyśmy przedstawić  $W_1$  jako

$$W_1 = \begin{bmatrix} -(U_{21} - U_{11})(1 - \xi_2) + (U_{22} - U_{12})\xi_2 \frac{h_2}{h_1} \frac{\partial \beta}{\partial y} \\ -(U_{12} - U_{11})(1 - \xi_1) + (U_{22} - U_{21})\xi_1 \frac{h_1}{h_2} \frac{\partial \beta}{\partial x} \end{bmatrix} \quad (6.10.103)$$

gdzie pochodne  $\beta$  są obliczone w środku  $K$ . To wraz z jednostajnymi ograniczeniami  $c_1 \leq h_1/h_2 \leq c_2$  dla pewnych  $0 < c_1 \leq c_2$  implikuje, że

$$\int_K W_1 \tilde{K} W_1 dx \leq C \|\nabla \beta\|_{L^\infty(K)}^2 \|U\|_{L^2(K)}^2 \quad (6.10.104)$$

co wraz z (6.10.101) wystarczałoby do pokazania (6.10.89).

Aby móc kontynuować analizę dla przypadku ogólnego (brak ograniczenia na  $h_1/h_2$ ) musimy rozpatrzyć szczegółowo własności  $\tilde{K}$ . Za pracą [51] (Appendix A) znajdujemy, że  $\ker(\tilde{K})$  jest rozpięte na następujących wektorach w  $R^8$ :

$$\begin{aligned} v_1 &= (1, 0, 0, 0, 0, 0, 0, 0)^T, \quad v_2 = (0, 0, 0, 0, 1, 0, 0, 0)^T, \\ v_3 &= (0, 0, 1, 0, 0, -1, 0, 0)^T \end{aligned} \quad (6.10.105)$$

gdzie pierwsze cztery składowe odpowiadają różniczkowaniu względem  $x_1$ , pozostałe względem  $x_2$ . Załóżmy wpraw, że  $\nabla U = c_i v_i$ ,  $i = 1, 2$ . Wówczas oczywiście  $W_1 = c_2 v_i$  i (6.10.102) jest trywialnie spełnione. Rozważmy przypadek z  $\nabla U = c v_3$ . Zgodnie z postacią  $v_3$  w (6.10.105) porównujemy trzeci i szósty wiersz (6.10.100) (obydwa wektorowe wiersze (6.10.100) składają się z 4 składowych), otrzymując:

$$\frac{u_{21}^3 - u_{11}^3}{h_1} (1 - \xi_2) + \frac{u_{22}^3 - u_{12}^3}{h_1} \xi_2 = - \left( \frac{u_{12}^2 - u_{11}^2}{h_2} (1 - \xi_1) + \frac{u_{22}^2 - u_{21}^2}{h_2} \xi_1 \right), \quad (6.10.106)$$

dla wszystkich  $\xi_1, \xi_2 \in [0, 1]$ , gdzie górne wskaźniki oznaczają wektorowe składowe  $U$ . (6.10.106) jest spełnione, gdy

$$u_{21}^3 - u_{11}^3 = u_{22}^3 - u_{12}^3, \quad u_{12}^2 - u_{11}^2 = u_{22}^2 - u_{21}^2, \quad \frac{u_{21}^3 - u_{11}^3}{h_1} = - \frac{u_{12}^2 - u_{11}^2}{h_2} \quad (6.10.107)$$

Pierwsze dwa warunki oznaczają, że  $U$  jest liniową (nie biliniową) funkcją  $x_i$ , tj. że  $\nabla U$  jest stałe w  $K$ . Wykorzystując (6.10.107) możemy przedstawić  $W_1$  następująco

$$W_1 = \quad (6.10.108)$$

$$\begin{bmatrix} 0 \\ 0 \\ \frac{u_{21}^3 - u_{11}^3}{h_1} \left[ \frac{\beta_{21} + \beta_{11}}{2} (1 - \xi_2) + \frac{\beta_{22} + \beta_{12}}{2} \xi_2 \right] \\ 0 \\ 0 \\ -\frac{u_{31}^3 - u_{11}^3}{h_1} \left[ \frac{\beta_{12} + \beta_{11}}{2} (1 - \xi_1) + \frac{\beta_{22} + \beta_{21}}{2} \xi_1 \right] \\ 0 \\ 0 \end{bmatrix} = a v_3 + \begin{bmatrix} 0 \\ 0 \\ -(u_{12}^2 - u_{11}^2) \frac{\partial \beta}{\partial y} (\xi_1 - \frac{1}{2}) \\ 0 \\ 0 \\ (u_{21}^3 - u_{11}^3) \frac{\partial \beta}{\partial x} (\xi_2 - \frac{1}{2}) \\ 0 \\ 0 \end{bmatrix} \quad (6.10.109)$$

gdzie  $a = (u_{21}^3 - u_{11}^3)(\beta_{11} + \beta_{12} + \beta_{21} + \beta_{22})/(4h_1)$  i gdzie wzięliśmy pod uwagę fakt, że  $\beta$  jest funkcją dwuliniową a jej pochodne są obliczone w środku  $K$ . Ponieważ pierwszy składnik  $W_1$  jest w  $\ker(\tilde{K})$  rozkład (6.10.108) implikuje, że

$$\int_K W_1^T \tilde{K} W_1 dx \leq C \|\nabla \beta\|_{L^\infty(K)}^2 \|U\|_{L^2(K)}^2 \quad (6.10.110)$$

Ostatecznie łączymy (6.10.97) i (6.10.101), nierówność (6.10.102), która zachodzi, jeśli  $\nabla U \in \ker(\tilde{K})^\perp$ , i powyższy wynik dla  $\nabla \tilde{K} \in \ker(\tilde{K})$  otrzymując

$$\int_K \nabla I_h(\theta, U)^T \tilde{K} \nabla I_h(\theta, U) dx \leq C \left( \int_K \nabla U^T \tilde{K} \nabla U dx + \frac{1}{H_1^2} \|U\|_{L^2(K)}^2 \right) \quad (6.10.111)$$

gdzie wykorzystano ograniczenie  $\|\nabla \theta_i\|_{\infty, \Omega_i} \leq C/H_i$ .

Ostatnim krokiem analizy jest rozważenie kontrybucji form kwadratowych  $\nabla U^T \tilde{A} \nabla U := U^T \tilde{A}_i^T \tilde{\tau} \tilde{A}_j U_j$  do nierówności (6.10.89). Forma  $\tilde{A}$  jest co najmniej dodatnio półokreślona. Nasza argumentacja może być podobna jak dla macierzy  $\tilde{K}$ . Tym razem jednak nie możemy użyć argumentu z rozkładem  $\nabla U$  na składniki z  $\ker(\tilde{A})$  i  $\ker(\tilde{A})^\perp$ , gdyż niezzerowe wartości własne  $\tilde{A}$  mogą być dowolnie bliskie zera. Ominiemy tę trudność przez użycie formy  $\tilde{K}$  zamiast  $\tilde{A}$  w ograniczającym wyrażeniu. Rozważmy wektor  $W_1$  zdefiniowany jak (6.10.98). Niech  $\Lambda > 0, \lambda > 0$  będą zdefiniowane następująco:

$$\Lambda = \sup_{\nabla U \neq 0} \frac{\int_K \nabla U^T \tilde{A} \nabla U dx}{\|\nabla U\|_{L^2(K)}^2}, \quad \lambda = \inf_{\nabla U \in \ker(\tilde{K})^\perp} \frac{\int_K \nabla U^T \tilde{K} \nabla U dx}{\|\nabla U\|_{L^2(K)}^2} \quad (6.10.112)$$

Wówczas dla każdego  $U$ , takiego że  $\nabla U \in \ker(\tilde{K})^\perp$ , możemy napisać

$$\int_K W_1^T \tilde{A} W_1 dx \leq \quad (6.10.113)$$

$$C \Lambda \|\beta\|_{L^\infty(K)}^2 \|\nabla U\|_{L^2(K)}^2 \leq C \frac{\Lambda}{\lambda} \|\beta\|_{L^\infty(K)}^2 \int_K \nabla U^T \tilde{K} \nabla U dx$$

co odpowiada (6.10.102). W warstwach przysciennych, gdzie wprowadza się elementy silnie wydłużone, wielkość wyrazów stabilizujących  $\tilde{A}_i^T \tilde{\tau} \tilde{A}_j$  jest porównywalna z  $\tilde{K}_{ij}$ . Dlatego  $\Lambda/\lambda$  jest względnie niewielką liczbą.

Weźmy z kolei  $\nabla U \in \ker(\bar{K})$ . Dla  $\nabla U = v_i$ ,  $i = 1, 2$  w (6.10.105) bezpośrednie przeliczenie pokazuje, że oszacowanie (6.10.102) z  $\bar{A}$  zastępującym  $\bar{K}$  jest spełnione. Zbadajmy bowiem wielkość

$$v_1^T \bar{A} v_1 = (1, 0, 0, 0, 0, 0, 0, 0) [\bar{A}_1, \bar{A}_2]^T \bar{\tau} [\bar{A}_1, \bar{A}_2] (1, 0, 0, 0, 0, 0, 0, 0)^T = z^T \bar{\tau} z \quad (6.10.114)$$

gdzie  $z = \bar{A}_1 (1, 0, 0, 0)^T$ . Widzimy, że  $z$  jest pierwszą kolumną macierzy  $\bar{A}_1$ , identyczną z jej pierwszym wierszem (bo  $\bar{A}_1 = \bar{A}_1^T$ ). Ponieważ  $\bar{A}_1 = A_1 A_0$ , zaś pierwszy wiersz  $A_1$  ma postać  $(0, 1, 0, 0)$ , znajdujemy że

$$z = (0, 1, 0, 0) A_0 = \left( -\psi \frac{u}{\iota}, \frac{1}{\iota} (1 + \varrho \frac{u^2}{\iota}), \varrho \frac{uv}{\iota^2}, -\varrho \frac{u}{\iota^2} \right) \quad (6.10.115)$$

gdzie  $\psi = \varrho(u^2 + v^2)/(2\iota)$ ,  $u, v$  są składowymi prędkościami (postać  $A_1$  i  $A_0$  podano w dodatku A.2). Przynajmniej drugi wyraz  $z$  jest zawsze dodatni, co przy jednostajnej dodatniej określoności  $\bar{\tau}$  gwarantuje dodatniość  $z^T \bar{\tau} z$  i spełnianie (6.10.102) (z  $\bar{A}$  zamiast  $\bar{K}$ ).

Wektor  $v_3$  jądra  $\ker(\bar{K})$  należy też do  $\ker(\bar{A})$ , dlatego oszacowanie analogiczne do (6.10.110) jest spełnione. W końcu izotropowa kontrybucja w  $B$  może być uwzględniona w nierówności (6.10.89) według zasad z poprzedniego podrozdziału.

## 6.11 Uwarunkowanie wstępne dla sformułowania niesymetrycznego

Zweryfikowanie, czy metody Schwarza są skuteczne do uwarunkowania wstępnego niesymetrycznego sformułowania przepływów (3.1.4) napotyka na trudności wymienione na początku poprzedniego podrozdziału. Istnieje jednak możliwość skonstruowania algorytmu uwarunkowania wstępnego dla problemu niesymetrycznego w oparciu o taką procedurę dla części symetrycznej danego operatora, możliwość taką przedstawiają Xu i Cai w [113]. W niniejszym podrozdziale prezentujemy zastosowanie tej teorii do niesymetrycznych równań przepływów.

Rozważmy układ równań liniowych w skończenie wymiarowej przestrzeni Hilberta  $V$  z iloczynem skalarnym  $(\cdot, \cdot)$  i związaną z nim normą  $\| \cdot \| = (\cdot, \cdot)^{1/2}$ . Układ ma postać:

$$A_N u = f \quad (6.11.116)$$

gdzie  $A_N$  jest sumą dwu operatorów liniowych,  $A_N = A + N$ . Zakładamy, że istnieje jednoznaczne rozwiązanie (6.11.116). Zakładamy też, że spełnione są następujące założenia:

A.1. Odwzorowanie  $A : V \rightarrow V$  jest symetryczne i dodatnio określone względem  $(\cdot, \cdot)$ . Przez  $(\cdot, \cdot)_A$  oznaczamy dodatkowy iloczyn skalarny dany przez  $(A \cdot, \cdot)$ ,  $\| \cdot \|_A$  oznacza normę  $(\cdot, \cdot)_A^{1/2}$ . Zakładamy, że istnieje takie  $c_1 > 0$ , że:

$$\| u \| \leq c_1 \| u \|_A, \quad \forall u \in V \quad (6.11.117)$$

A.2. Odwzorowanie  $N : V \rightarrow V$  może być niesymetryczne i spełnia warunek

$$(Nu, v) \leq c_2 \| u \| \| v \|_A, \quad \forall u, v \in V \quad (6.11.118)$$

gdzie  $c_2 > 0$  jest stałą.

A.3. Istnieje podprzestrzeń  $V_0 \subset V$  taka że dla każdego  $u \in V$  istnieje jednoznaczne  $u_0 \in V_0$  spełniające związek:

$$(A_N u_0, v_0) = (A_N u, v_0), \quad \forall v_0 \in V_0 \quad (6.11.119)$$

Definiuje to operator  $P_0 : V \rightarrow V_0$  taki że  $P_0 u := u_0$ . Z operatorem  $P_0$  wiążemy parametr  $\delta_0$ :

$$\delta_0 := \sup_{v \in V} \frac{\| (I - P_0)v \|}{\| v \|_A} \quad (6.11.120)$$

gdzie  $I$  jest operatorem identyczności w  $V$ . Dodatkowo definiujemy operatory liniowe  $A_0 : V_0 \rightarrow V_0$  i  $Q_0 : V \rightarrow V_0$  takie że:

$$(A_0 u_0, v_0) = (A_N u_0, v_0), \quad \forall u_0, v_0 \in V_0 \quad (6.11.121)$$

$$(Q_0 u, v_0) = (u, v_0), \quad \forall u \in V, \quad \forall v_0 \in V_0 \quad (6.11.122)$$

Możemy sprawdzić, że

$$A_0 P_0 = Q_0 A_N \quad (6.11.123)$$

gdyż (6.11.119) i (6.11.121) z  $u_0 = P_0 u$  implikują, że  $(A_0 P_0 u, v_0) = (A_N u, v_0)$ , zaś (6.11.122) dla  $u := A_N u$ , że  $(Q_0 A_N u, v_0) = (A_N u, v_0)$ .

Korzystając z definicji  $A_0$  i  $Q_0$  autorzy proponują operator uwarunkowania wstępnego dla  $A_N$  w postaci

$$\text{http://rcin.org.pl} \quad (6.11.124)$$

gdzie  $B$  jest danym symetrycznym dodatnio określonym operatorem uwarunkowania wstępnego dla symetrycznej części  $A_N$ , tj.  $A, \beta > 0$  jest ustalonym parametrem zależnym od  $B$ . Oryginalny operator  $A_N$  z uwarunkowaniem wstępnym ma więc postać

$$B_N A_N = (A_0^{-1} Q_0 + \beta B) A_N = P_0 + \beta B A_N, \quad (6.11.125)$$

gdzie wykorzystano (6.11.123). Własności operatora  $B_N$  opisuje następujące

#### Twierdzenie 6.11.4

Istnieją dodatnie stałe  $\varepsilon, \alpha, \beta$  i  $\mu$ , które zależą od  $\lambda_0$  i  $\lambda_1$ , minimalnej i maksymalnej wartości własnej  $BA$ , takie że, jeśli  $\delta_0 \leq \varepsilon$ , to

$$\sup_{u \neq 0} \frac{\|B_N A_N u\|_A}{\|u\|_A} = \mu, \quad \inf_{u \neq 0} \frac{(B_N A_N u, u)_A}{(u, u)_A} = \alpha. \quad (6.11.126)$$

Postać wymienionych parametrów jest następująca

$$\varepsilon^2 = \frac{1}{4c_2^2(c_1^2 c_2^2 + 1)(\lambda_1/\lambda_0)^2}, \quad \beta = \frac{\lambda_0}{2c_1^2 c_2^2 \lambda_1^2} \quad (6.11.127)$$

$$\alpha = \frac{\lambda_0^2}{8c_1^2 c_2^2 \lambda_1^2}, \quad \mu = 1 + c_2 \delta_0 + \beta \lambda_1 (1 + c_1 c_2)$$

**Uwaga 6.11.4** W przypadku, gdy (6.11.116) odpowiada dyskretyzacji skończonej elementowej i jeśli stałe  $c_1$  i  $c_2$  oraz wartości własne  $\lambda_0, \lambda_1$  są niezależne od siatki, to także parametry  $\alpha$  i  $\mu$  są niezależne od dyskretyzacji. Gwarantuje to, zgodnie z (6.2.10), zbieżność algorytmu GMRES dla zadania niesymetrycznego z uwarunkowaniem wstępnym  $B_N$ .

Aby zastosować przedstawioną teorię do niesymetrycznych równań przepływów wykorzystamy ideę symetryzacji równań Naviera–Stokesa za pomocą macierzy  $A_0$  (2.2.32). W tym celu w formie dwuliniowej sformułowania niesymetrycznego

$$B(U, W) = \int_{\Omega} W^T U dx + \int_{\Omega} W_{,i}^T \alpha_{ij} U_{,j} dx \quad (6.11.128)$$

gdzie

$$\alpha_{ij} = \Delta t (K_{ij} + \tau_a A_i A_j + \nu_d I \delta_{ij}), \quad (6.11.129)$$

zastosujemy podstawienie:

$$U := I_h(A_0 V), \quad (6.11.130)$$

gdzie  $V$  jest funkcją z przestrzeni skończonej elementowej  $V^h$ , zaś operator interpolacji  $I_h$  zastosowano, aby także  $U \in V^h$  (funkcja  $A_0(\mathbf{x})V(\mathbf{x})$  na ogół nie należy do  $V^h$ ). Podstawienie (6.11.130) jest równoważne zamianie stopni swobody w każdym z węzłów  $\mathbf{x}_i$  zgodnie ze wzorem  $U^{(i)} = A_0(\mathbf{x}_i)V^{(i)}$ , gdzie  $U^{(i)}$  i  $V^{(i)}$  są czwórkami starych i nowych stopni swobody w  $\mathbf{x}_i$ . Zakładając iloczyn skalarny w postaci

$$(U, W) := \int_{\Omega} W^T U dx$$

definiujemy  $A_N$  następująco

$$(A_N V, W) := B(I_h(A_0 V), W) = B(A_0 V, W) - B((I - I_h)(A_0 V), W) = \int_{\Omega} W^T A_0 V dx + \int_{\Omega} W_{,i}^T \tilde{\alpha}_{ij} V_{,j} dx + \int_{\Omega} W_{,i}^T \alpha_{ij} A_{0,j} V dx - B((I - I_h)(A_0 V), W) \quad (6.11.131)$$

gdzie  $\bar{a}_{ij} = \Delta t(K_{ij} \mathbf{A}_0 + \tau_a \mathbf{A}_i \mathbf{A}_0 \mathbf{A}_j^T + \delta_{ij} \mathbf{A}_0) = \mathbf{a}_{ij} \mathbf{A}_0$ . Zauważmy, że pierwsze dwa składniki po prawej stronie (6.11.131) są symetryczne i odpowiadają formie dwuliniowej rozpatrywanej w poprzednim podrozdziale dla sformułowania symetrycznego, o ile przyjąć  $\tau = \tau_a \mathbf{A}_0^{-1}$  (co faktycznie nie odpowiada symetrycznemu sformułowaniu SUPG, lecz takie  $\tau$  jest też symetryczne dodatnio określone i jednostajnie ograniczone w  $\Omega$ ). Dla takiej formy, jak pokazano w poprzednim podrozdziale, odpowiednim symetrycznym algorytmem uwarunkowania wstępnego jest iteracja prosta typu Jacobiego.

Powyższa definicja  $A_N$  pozwala na dekompozycję tego operatora na część symetryczną  $A$  i niesymetryczną  $N$

$$\begin{aligned} (AV, W) &:= \int_{\Omega} \mathbf{W}^T \mathbf{A}_0 \mathbf{V} dx + \int_{\Omega} \mathbf{W}_j^T \bar{a}_{ij} \mathbf{V}_j dx & (6.11.132) \\ (NV, W) &:= \int_{\Omega} \mathbf{W}_i^T \mathbf{a}_{ij} \mathbf{A}_{0j} \mathbf{V}_j dx - B((I - I_h)(\mathbf{A}_0 \mathbf{V}), \mathbf{W}) \\ &\equiv (N_1 V, W) + (N_2 V, W) \end{aligned}$$

Naszym celem jest zweryfikowanie założeń A.1, A.2 i A.3.

Założenie A.1. Jest spełnione trywialnie, gdyż

$$\|U\|^2 = \int_{\Omega} U^T U dx \leq c_1 \int_{\Omega} U^T \mathbf{A}_0 U dx \leq c_1 (AU, U) = c_1 \|U\|_A^2 \quad (6.11.133)$$

Założenie A.2.

Sprawdzenia dokonamy dla obydwu składników  $N$  oddzielnie.

Składnik  $N_1$ . Stosując nierówność Cauchy'ego-Schwarza możemy napisać:

$$\begin{aligned} (N_1 V, W) &= \int_{\Omega} \mathbf{W}_i^T \bar{a}_{ij} \mathbf{A}_0^{-1} \mathbf{A}_{0j} \mathbf{V} dx \leq & (6.11.134) \\ &\left( \int_{\Omega} \mathbf{W}_i^T \bar{a}_{ij} \mathbf{W}_j dx \right)^{1/2} \left( \int_{\Omega} (\mathbf{A}_0^{-1} \mathbf{A}_{0i} \mathbf{V})^T \bar{a}_{ij} (\mathbf{A}_0^{-1} \mathbf{A}_{0j} \mathbf{V}) dx \right)^{1/2} \leq \\ &c_0 \|W\|_A \|V\| \end{aligned}$$

gdzie

$$c_0^2 := \sup_{v \neq 0} \frac{\int_{\Omega} (\mathbf{A}_0^{-1} \mathbf{A}_{0i} \mathbf{V})^T \bar{a}_{ij} (\mathbf{A}_0^{-1} \mathbf{A}_{0j} \mathbf{V}) dx}{(V, V)} \quad (6.11.135)$$

Składnik  $N_2$ . Jego postać jest następująca

$$(N_2 V, W) = \int_{\Omega} \mathbf{W}^T (I - I_h)(\mathbf{A}_0 \mathbf{V}) dx + \int_{\Omega} \mathbf{W}_i^T \mathbf{a}_{ij} [(I - I_h)(\mathbf{A}_0 \mathbf{V})]_j dx \quad (6.11.136)$$

Rozpatrzmy kontrybucje do  $N_2$  z poszczególnych elementów  $K$ . Dla pierwszego składnika wzoru (6.11.136) mamy

$$\begin{aligned} \int_K \mathbf{W}^T (I - I_h)(\mathbf{A}_0 \mathbf{V}) dx &\leq \|W\|_{L^2(K)} \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{L^2(K)} \\ &\leq c_1 \|W\|_{A,K} \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{L^2(K)} \end{aligned} \quad (6.11.137)$$

gdzie  $\|\cdot\|_{A,K}$  oznacza normę funkcji o nosniku ograniczonym do  $K$ . Rozpatrzmy drugi czynnik prawej strony (6.11.137). Załóżmy na początek, że wydłużenie elementów jest

ograniczone, tj. że każdy element charakteryzuje się parametrem  $h$ . Wówczas wspomniane wyrażenie możemy zapisać następująco:

$$\| (I - I_h)(\mathbf{A}_0 \mathbf{V}) \|_{L^2(K)}^2 \leq \quad (6.11.138)$$

$$Ch^2 \int_K (\hat{I} - \hat{I}_h)(\hat{\mathbf{A}}_0 \hat{\mathbf{V}}) d\xi \leq Ch^2 \int_K \sum_i \left| \frac{\partial^2 (\hat{\mathbf{A}}_0 \hat{\mathbf{V}})}{\partial \xi_i^2} \right|^2 d\xi \leq$$

$$Ch^2 \int_K \left[ |\hat{\mathbf{V}}|^2 \sum_i \left| \frac{\partial^2 \hat{\mathbf{A}}_0}{\partial \xi_i^2} \right|^2 + \left( \sum_i \left| \frac{\partial \hat{\mathbf{A}}_0}{\partial \xi_i} \right|^4 \right)^{1/2} \left( \sum_i \left| \frac{\partial \hat{\mathbf{V}}}{\partial \xi_i} \right|^4 \right)^{1/2} \right] d\xi$$

W powyższym a także w dalszej części symbol  $|\cdot|$  oznacza normę euklidesową wektorów w  $R^4$  bądź w  $R^8$  lub zgodną z nią normę dla macierzy  $(\hat{\mathbf{A}}_0, \partial \hat{\mathbf{A}}_0 / \partial \xi_i, \dots)$ . Pierwszy krok w (6.11.138) wynika z własności (4.1.5) odwzorowania  $\phi: \hat{K} \rightarrow K$ , w następnym zastosowano klasyczne oszacowanie błędu interpolacji w normie  $L^2$ . W ostatnim kroku zastosowano nierówność Cauchy'ego-Schwarza. Stosując dalej tę nierówność dla iloczynu skalarnego w  $L^2$  otrzymujemy:

$$\| (I - I_h)(\mathbf{A}_0 \mathbf{V}) \|_{L^2(K)}^2 \leq \quad (6.11.139)$$

$$Ch^2 \left[ \left( \int_K |\hat{\mathbf{V}}|^4 d\xi \right)^{1/4 \cdot 2} \left( \int_K \sum_i \left| \frac{\partial^2 \hat{\mathbf{A}}_0}{\partial \xi_i^2} \right|^4 \right)^{1/2} + \right.$$

$$\left. \left( \int_K \sum_i \left| \frac{\partial \hat{\mathbf{A}}_0}{\partial \xi_i} \right|^4 \right)^{1/2} \left( \int_K \sum_i \left| \frac{\partial \hat{\mathbf{V}}}{\partial \xi_i} \right|^4 \right)^{1/4 \cdot 2} \right]$$

Ponieważ wszystkie normy na skończonej wymiarowej przestrzeni są równoważne, występujące wyżej wyrazy  $\| \hat{\mathbf{V}} \|_{L^4(\hat{K})}$  i  $\| \partial \hat{\mathbf{V}} / \partial \xi_i \|_{L^4(\hat{K})}$  można oszacować przez  $\| \hat{\mathbf{V}} \|_{L^2(\hat{K})}$ . Przechodząc dodatkowo z elementu wzorcowego  $\hat{K}$  na rzeczywisty  $K$  otrzymujemy:

$$\| (I - I_h)(\mathbf{A}_0 \mathbf{V}) \|_{L^2(K)}^2 \leq \quad (6.11.140)$$

$$Ch^2 \left[ \left( \int_K |\mathbf{V}|^2 \frac{1}{h^2} dx \right)^{1/2 \cdot 2} \left( \int_K \sum_{i,j,k} \left| \frac{\partial^2 \mathbf{A}_0}{\partial x_j \partial x_k} \frac{\partial x_k}{\partial \xi_i} \frac{\partial x_j}{\partial \xi_i} \right|^4 \frac{1}{h^2} dx \right)^{1/2} + \right.$$

$$\left. \left( \int_K |\mathbf{V}|^2 \frac{1}{h^2} dx \right)^{1/2 \cdot 2} \left( \int_K \sum_{i,j} \left| \frac{\partial \mathbf{A}_0}{\partial x_j} \frac{\partial x_j}{\partial \xi_i} \right|^4 \frac{1}{h^2} \right)^{1/2} \right] \leq \alpha_{1,K} \| \mathbf{V} \|_{L^2(K)}$$

gdzie

$$\alpha_{1,K} = C \left[ h^4 \left( \frac{1}{h^2} \int_K \sum_{j,k} \left| \frac{\partial^2 \mathbf{A}_0}{\partial x_j \partial x_k} \right|^4 dx \right)^{1/2} + h^2 \left( \frac{1}{h^2} \int_K \sum_j \left| \frac{\partial \mathbf{A}_0}{\partial x_j} \right|^4 \right)^{1/2} \right] \quad (6.11.141)$$

Dla elementów w postaci silnie wydłużonych równoległoboków rozumowanie jest podobne

$$\| (I - I_h)(\mathbf{A}_0 \mathbf{V}) \|_{L^2(K)}^2 \leq \sum_i h_k h_l \int_K \left( \left| \frac{\partial^2 \hat{\mathbf{A}}_0}{\partial \xi_i^2} \right|^2 |\hat{\mathbf{V}}|^2 + \left| \frac{\partial \hat{\mathbf{A}}_0}{\partial \xi_i} \right|^2 \left| \frac{\partial \hat{\mathbf{V}}}{\partial \xi_i} \right|^2 \right) d\xi \leq$$

$$C \sum_i h_k h_l \left[ \left( \int_{\hat{K}} |\mathbf{V}|^4 d\xi \right)^{1/4.2} \left( \int_{\hat{K}} \left| \frac{\partial^2 \hat{\mathbf{A}}_0}{\partial \xi_i^2} \right|^4 d\xi \right)^{1/2} + \right. \\ \left. \left( \int_{\hat{K}} \left| \frac{\partial \hat{\mathbf{A}}_0}{\partial \xi_i} \right|^4 d\xi \right)^{1/2} \left( \int_{\hat{K}} \left| \frac{\partial \hat{\mathbf{V}}}{\partial \xi_i} \right|^4 \right)^{1/4.2} \right] \quad (6.11.142)$$

Powyższa argumentacja różni się od poprzedniej możliwością prowadzenia osobnych oszacowań dla kierunków  $\xi_1$  i  $\xi_2$ . Ze względu na równoważność norm na przestrzeni skończenie wymiarowej występujące powyżej wyrazy  $\|\hat{\mathbf{V}}\|_{L^4(\hat{K})}$  i  $\|\partial \hat{\mathbf{V}}/\partial \xi_i\|_{L^2(\hat{K})}$  można oszacować przez  $\|\hat{\mathbf{V}}\|_{L^2(\hat{K})}$ , co wraz z powrotem do obszaru całkowania  $K$  pociąga oszacowanie

$$\| (I - I_h)(\mathbf{A}_0 \mathbf{V}) \|_{L^2(K)}^2 \leq \quad (6.11.143) \\ C \sum_{p=k,l} h_k h_l \left[ \left( \int_K |\mathbf{V}|^2 \frac{1}{h_k h_l} dx \right)^{1/2.2} \left( \int_K \left| \frac{\partial^2 \mathbf{A}_0}{\partial p^2} \right|^4 h_p^3 \frac{1}{h_k h_l} dx \right)^{1/2} + \right. \\ \left. \left( \int_K \left| \frac{\partial \mathbf{A}_0}{\partial p} \right|^4 h_p^4 \frac{1}{h_k h_l} dx \right)^{1/2} \left( \int_K |\mathbf{V}|^2 \frac{1}{h_k h_l} dx \right)^{1/2.2} \right] \leq \alpha'_{1,K} \|\mathbf{V}\|_{L^2(K)}^2$$

gdzie

$$\alpha'_{1,K} = C \sum_{p=k,l} \left[ h_p^4 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial^2 \mathbf{A}_0}{\partial p^2} \right|^4 dx \right)^{1/2} + h_p^2 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial \mathbf{A}_0}{\partial p} \right|^4 dx \right)^{1/2} \right] \quad (6.11.144)$$

Przejdźmy obecnie do drugiego składnika oszacowania operatora  $N_2$  w (6.11.136)

$$\int_K \mathbf{W}_{,i}^T \mathbf{a}_{ij} [(I - I_h)(\mathbf{A}_0 \mathbf{V})]_{,j} dx = \int_K \mathbf{W}_{,i}^T \tilde{\mathbf{a}}_{ij} \mathbf{A}_0^{-1} [(I - I_h)(\mathbf{A}_0 \mathbf{V})]_{,j} dx \leq \\ \left( \int_K \mathbf{W}_{,i}^T \tilde{\mathbf{a}}_{ij} \mathbf{W}_{,j} dx \right)^{1/2} \\ \left( \int_K \{ \mathbf{A}_0^{-1} [(I - I_h)(\mathbf{A}_0 \mathbf{V})]_{,i} \}^T \tilde{\mathbf{a}}_{ij} \{ \mathbf{A}_0^{-1} [(I - I_h)(\mathbf{A}_0 \mathbf{V})]_{,j} \} dx \right)^{1/2} \leq \\ c_3 \left( \int_K \mathbf{W}_{,i}^T \tilde{\mathbf{a}}_{ij} \mathbf{W}_{,j} dx \right)^{1/2} \left( \int_K |\nabla [(I - I_h)(\mathbf{A}_0 \mathbf{V})]|^2 dx \right)^{1/2} \quad (6.11.145)$$

W powyższym zastosowaliśmy nierówność Cauchy'ego-Schwarza oraz przyjęliśmy

$$c_3 = \sup_{z \in \Omega} |\mathbf{r}_i^T \mathbf{A}_0^{-1} \tilde{\mathbf{a}}_{ij} \mathbf{A}_0^{-1} \mathbf{r}_j| \quad (6.11.146)$$

gdzie  $\mathbf{r}_i \in R^{4 \times 8}$  są macierzami restrykcji  $\mathbf{r}_1 \{a_1, \dots, a_8\}^T = \{a_1, \dots, a_4\}^T$ ,  $\mathbf{r}_2 \{a_1, \dots, a_8\}^T = \{a_5, \dots, a_8\}^T$ . Pierwszy czynnik prawej strony (6.11.145) jest ograniczony przez  $\|\mathbf{W}\|_A$ . Problemem jest oszacowanie drugiego czynnika przez  $\|\mathbf{V}\|_{L^2(K)}$ , co nie jest trywialne ze względu na obecność w nim pochodnych. Tak i poprzednio rozważmy najpierw elementy



o ograniczonym wydłużeniu. Stosując klasyczne oszacowanie interpolacyjne dla półnormy  $H^1$  otrzymujemy:

$$\int_K |\nabla[(I - I_h)(\mathbf{A}_0 \mathbf{V})]|^2 dx \leq C \int_K \sum_i \left| \frac{\partial^2(\hat{\mathbf{A}}_0 \hat{\mathbf{V}})}{\partial \xi_i^2} \right|^2 d\xi \quad (6.11.147)$$

co z dokładnością do czynnika  $h^2$  odpowiada rozważanemu poprzednio oszacowaniu (6.11.138) dla błędu w normie  $L^2$ . Ostateczna postać oszacowania jest więc też identyczna z (6.11.140) z pominięciem czynnika  $h^2$ , tj.

$$\|\nabla(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{L^2(K)}^2 \leq \alpha_{2,K} \|\mathbf{V}\|_{L^2(K)}^2 \quad (6.11.148)$$

gdzie

$$\alpha_{2,K} = C \left[ h^2 \left( \frac{1}{h^2} \int_K \sum_{j,k} \left| \frac{\partial^2 \mathbf{A}_0}{\partial x_j \partial x_k} \right|^4 dx \right)^{1/2} + \left( \frac{1}{h^2} \int_K \sum_j \left| \frac{\partial \mathbf{A}_0}{\partial x_j} \right|^4 \right)^{1/2} \right] \quad (6.11.149)$$

W końcu weźmy pod uwagę powyższe oszacowanie w przypadku elementów o nieograniczonym wydłużeniu. W tym celu musimy wykorzystać wynik interpolacyjny (4.3.60) z rozdziału 4.3.

$$\int_K |\nabla[(I - I_h)(\mathbf{A}_0 \mathbf{V})]|^2 dx \leq A_1 + A_2 + A_3 + A_4 \quad (6.11.150)$$

gdzie

$$\begin{aligned} A_1 &= Ch_k^2 \int_K \left| \frac{\partial^2(\mathbf{A}_0 \mathbf{V})}{\partial k^2} \right|^2 dx, \quad A_2 = Ch_l^2 \int_K \left| \frac{\partial^2(\mathbf{A}_0 \mathbf{V})}{\partial l^2} \right|^2 dx \\ A_3 &= Ch_k^4 \int_K \left| \frac{\partial^3(\mathbf{A}_0 \mathbf{V})}{\partial k^2 \partial l} \right|^2 dx, \quad A_4 = Ch_l^4 \int_K \left| \frac{\partial^3(\mathbf{A}_0 \mathbf{V})}{\partial l^2 \partial k} \right|^2 dx \end{aligned} \quad (6.11.151)$$

Pierwsze dwa składniki oszacowania,  $A_1, A_2$ , są znów identycznie z rozważanymi wcześniej wyrażeniami dla błędu w normie  $L^2$ , (6.11.142), pominawszy czynniki  $h_k^2$  i  $h_l^2$ , przykładowo

$$A_1 \leq h_k h_l \int_K \frac{1}{h_k^2} \left( \left| \frac{\partial^2 \hat{\mathbf{A}}_0}{\partial \xi_1} \right|^2 |\hat{\mathbf{V}}|^2 + \left| \frac{\partial \hat{\mathbf{A}}_0}{\partial \xi_1} \right|^2 \left| \frac{\partial \hat{\mathbf{V}}}{\partial \xi_1} \right|^2 \right) d\xi \quad (6.11.152)$$

Wobec tego ostateczne oszacowanie tych wyrazów jest identyczne z (6.11.143) z modyfikacją o wspomniane czynniki

$$A_1 + A_2 \leq \alpha'_{2,K} \|\mathbf{V}\|_{L^2(K)}^2 \quad (6.11.153)$$

gdzie

$$\alpha'_{2,K} = \sum_{p=k,l} \left[ h_p^2 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial^2 \mathbf{A}_0}{\partial p^2} \right|^4 dx \right)^{1/2} + \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial \mathbf{A}_0}{\partial p} \right|^4 dx \right)^{1/2} \right] \quad (6.11.154)$$

Rozważmy wreszcie pozostałe składniki oszacowania (6.11.150). Dla wyrażenia  $A_3$  możemy napisać

$$A_3 \leq h_k^4 \cdot h_k h_l \cdot \frac{1}{h_k^4 h_l^2} \int_K \left| \frac{\partial^3(\hat{\mathbf{A}}_0 \hat{\mathbf{V}})}{\partial^2 \xi_1 \partial \xi_2} \right|^2 d\xi \leq \dots$$

$$\frac{h_k}{h_l} \left\{ \left( \int_{\hat{K}} \left| \frac{\partial^3 \hat{A}}{\partial \xi_1^2 \partial \xi_2} \right|^4 d\xi \right)^{\frac{1}{2}} \left( \int_{\hat{K}} |\hat{V}|^4 d\xi \right)^{\frac{1}{2}} + \left( \int_{\hat{K}} \left| \frac{\partial^2 \hat{A}_0}{\partial \xi_1^2} \right|^4 d\xi \right)^{\frac{1}{2}} \left( \int_{\hat{K}} \left| \frac{\partial \hat{V}}{\partial \xi_2} \right|^4 \right)^{\frac{1}{2}} + \right. \\ \left. \left( \int_{\hat{K}} \left| \frac{\partial^2 \hat{A}_0}{\partial \xi_1 \partial \xi_2} \right|^4 d\xi \right)^{\frac{1}{2}} \left( \int_{\hat{K}} \left| \frac{\partial \hat{V}}{\partial \xi_1} \right|^4 \right)^{\frac{1}{2}} + \left( \int_{\hat{K}} \left| \frac{\partial \hat{A}_0}{\partial \xi_1} \right|^4 d\xi \right)^{\frac{1}{2}} \left( \int_{\hat{K}} \left| \frac{\partial^2 \hat{V}}{\partial \xi_1 \partial \xi_2} \right|^4 \right)^{\frac{1}{2}} \right\} \quad (6.11.155)$$

Ponieważ wszystkie normy na przestrzeni skończenie wymiarowej są równoważne, odpowiednie całki z  $\hat{V}$  i jego pochodnych można oszacować przez  $\|\hat{V}\|_{L^2(\hat{K})}$ , co wraz z zmianą obszaru całkowania z  $\hat{K}$  na  $K$  prowadzi do oszacowania:

$$A_3 \leq \alpha_{2,K}'' \|\mathbf{V}\|_{L^2(K)}^2 \quad (6.11.156)$$

gdzie

$$\alpha_{2,K}'' = \left\{ h_k^4 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial^3 \mathbf{A}_0}{\partial k^2 \partial l} \right|^4 dx \right)^{\frac{1}{2}} + h_k^2 \left( \frac{h_k}{h_l} \right)^2 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial^2 \mathbf{A}_0}{\partial k^2} \right|^4 dx \right)^{\frac{1}{2}} + \right. \\ \left. h_k^2 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial^2 \mathbf{A}_0}{\partial k \partial l} \right|^4 dx \right)^{\frac{1}{2}} + \left( \frac{h_k}{h_l} \right)^2 \left( \frac{1}{h_k h_l} \int_K \left| \frac{\partial \mathbf{A}_0}{\partial k} \right|^4 dx \right)^{\frac{1}{2}} \right\} \quad (6.11.157)$$

Analogiczne oszacowanie obowiązuje dla  $A_4$ .

Otrzymawszy komplet oszacowań dla elementowych kontrybucji wzoru (6.11.136) znajdujemy ostateczne oszacowanie dla operatora  $N_2$ . Stosując (6.11.139) i (6.11.145) otrzymujemy:

$$(\mathbf{N}_2 \mathbf{V}, \mathbf{W}) \leq c_1 \sum_K \|\mathbf{W}\|_{A,K} \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{L^2(K)} + \quad (6.11.158) \\ c_3 \sum_K \|\mathbf{W}\|_{A,K} \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{H^1(K)} \leq \\ \|\mathbf{W}\|_A \left( c_1 \left( \sum_K \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{L^2(K)}^2 \right)^{1/2} + \right. \\ \left. c_3 \left( \sum_K \|(I - I_h)(\mathbf{A}_0 \mathbf{V})\|_{H^1(K)}^2 \right)^{1/2} \right)$$

Biorąc pod uwagę oszacowania wyrazów ostatniego czynnika (6.11.158), wzory (6.11.140) i (6.11.148) dla elementów o małym wydłużeniu i wzory (6.11.143), (6.11.150), (6.11.153) i (6.11.156) dla elementów silnie wydłużonych otrzymujemy, że

$$(\mathbf{N}_2 \mathbf{V}, \mathbf{W}) \leq (c_1 \max_K \alpha_{1,K}^{1/2} + c_3 \max_K \alpha_{2,K}^{1/2}) \|\mathbf{W}\|_A \|\mathbf{V}\|, \quad (6.11.159)$$

w pierwszym przypadku oraz

$$(\mathbf{N}_2 \mathbf{V}, \mathbf{W}) \leq (c_1 \max_K \alpha_{1,K}^{1/2} + c_3 \max_K (\alpha_{1,K}'' + \alpha_{2,K}'')^{1/2}) \|\mathbf{W}\|_A \|\mathbf{V}\|, \quad (6.11.160)$$

w drugim. Ostatecznie, biorąc dodatkowo pod uwagę nierówność (6.11.134), otrzymujemy poszukiwane oszacowanie dla  $N = N_1 + N_2$

$$(N\mathbf{V}, \mathbf{W}) \leq c_2 \| \mathbf{W} \|_A \| \mathbf{V} \| \quad (6.11.161)$$

gdzie

$$c_2 = c_0 + c_1 \max_K \alpha_{1,K}^{1/2} + c_3 \max_K \alpha_{2,K}^{1/2} \quad (6.11.162)$$

dla elementów niezbyt wydłużonych oraz

$$c_2 = c_0 + c_1 \max_K \alpha_{1,K}^{1/2} + c_3 \max_K (\alpha'_{2,K} + \alpha''_{2,K} + \alpha'''_{2,K})^{1/2} \quad (6.11.163)$$

dla elementów silnie wydłużonych.

Aby zweryfikować niezależność parametru  $c_2$  od aproksymacji rozważmy oddzielnie przypadek rozwiązania gładkiego i rozwiązania z falą uderzeniową. W przypadku rozwiązania gładkiego można założyć, że  $\mathbf{U}^n$  i jego pierwsze pochodne są asymptotycznie dowolnie bliskie rozwiązaniu ścisłemu a stąd, że są jednostajnie ograniczone. Pozwala to na oszacowanie zależności parametrów  $c_0, c_1, \alpha_{1,K}, \dots$  od charakterystyk dyskretyzacji  $\Delta t$  i  $h, h_k, h_l$ . Odpowiednie wyrażenia definiujące  $\alpha_{1,K}, \dots, \alpha'''_{2,K}$  przedstawiono w takiej postaci, aby uwidocznili ograniczoną wyrazów zależnych od  $\mathbf{A}_0$  przy  $h, h_k, h_l \rightarrow 0$ . Biorąc pod uwagę, że współczynniki  $\nu_d$  i  $\tau_a$  metody SUPG są rzędu  $h$  zaś macierze rzeczywistej lepkości są rzędu  $\mu$  (por. (3.1.3)), wielkość współczynników definiujących  $c_2$  możemy oszacować następująco:

$$c_0 \sim \sqrt{\Delta t \max(\mu, h_{max})}, \quad c_1 \sim 1, \quad c_3 \sim \sqrt{\Delta t h_{max}}, \quad \alpha_{1,K} \sim h_K^2, \quad \alpha_{2,K} \sim 1 \quad (6.11.164)$$

A zatem

$$c_2 \sim (\Delta t \max(\mu, h_{max}))^{1/2} + h_{max} + (\Delta t \max(\mu, h_{max}))^{1/2} \sim h_{max} \quad (6.11.165)$$

gdzie  $h_{max} = \max_K h_K$  i gdzie założyliśmy, że krok czasowy  $\Delta t$  jest proporcjonalny do  $\min_K h_K$ , co jest założeniem przyjmowanym w praktyce.

Dla rozwiązania z falą uderzeniową musimy założyć, że  $|\nabla \mathbf{U}^n|$  nie jest jednostajnie ograniczone, lecz że jego rząd wielkości zmienia się wraz z rozmiarem siatki jak  $1/h$  – grubość aproksymowanych fal uderzeniowych jest proporcjonalna do  $h$ . w tym wypadku oszacowania wielkości tworzących  $c_2$  są następujące:

$$c_0 \sim (\Delta t / h_{min})^{1/2}, \quad c_1 \sim 1, \quad c_3 \sim h_{max} \Delta t, \quad \alpha_{1,K} \sim 1, \quad \alpha_{2,K} \sim 1 / h_K^2 \quad (6.11.166)$$

a zatem

$$c_2 \sim \left( \frac{\Delta t}{h_{min}} \right)^{1/2} + 1 + (h_{max} \Delta t)^{1/2} \frac{1}{h_{min}} \sim 1 \quad (6.11.167)$$

o ile wzdłuż fal uderzeniowych założymy jednostajną ograniczoną  $h_{max}/h_{min}$ , co faktycznie ma miejsce w praktyce.

Przyjrzyjmy się jeszcze zależności  $c_2$  od parametrów siatek z elementami wydłużonymi, ograniczając się do przypadku rozwiązania gładkiego. W dużej części argumentacja jest tu podobna. Istotną różnicą jest pojawienie się w definicji  $\alpha_{2,K}$  proporcji  $h_k/h_l$ . Jest to

niekorzystne, gdyż zakładając, że  $h_k = \text{const}$ ,  $h_l \rightarrow 0$ ,  $\alpha_{2,K}''$  może rosnąć nieograniczenie. Aby uniknąć takiej sytuacji, musimy narzucić dodatkowe założenie wiążące rozmiary siatki z funkcją  $A_0(U^n(\mathbf{x}))$ . Łatwo zauważyć, że wystarczającym założeniem jest

$$\left| \frac{\partial U^n}{\partial k} \right| \leq c \frac{h_l}{\max(h_k, h_l)} |\nabla U^n|, \quad (6.11.168)$$

gdź implikuje to dla ostatniego składnika  $\alpha_{2,K}''$

$$\frac{h_k}{h_l} \left| \frac{\partial A_0}{\partial k} \right| \leq \frac{h_k}{h_l} \left| \frac{\partial A_0}{\partial U} \right| \left| \frac{\partial U^n}{\partial k} \right| \leq c \left| \frac{\partial A_0}{\partial U} \right| |\nabla U^n| \quad (6.11.169)$$

i podobnie dla drugiego, że

$$h_k \frac{h_k}{h_l} \left| \frac{\partial^2 A_0}{\partial k^2} \right| \leq c^2 h_l \left| \frac{\partial^2 A_0}{\partial U^2} \right| |\nabla U^n|^2 \quad (6.11.170)$$

Analogiczne ograniczenia należy postawić dla  $|\partial U^n / \partial l|$ . Ograniczenia te wydają się łatwe i naturalne do spełnienia w praktyce: silne wydłużenie elementów zachodzi wzdłuż warstwy przyściennej a tym samym jest prostopadle do gradientu rozwiązania.

**Założenie A.3.** Jako podprzestrzeń  $V_0$  przyjmujemy przestrzeń skończenie elementową odpowiadającą zgrubnej siatce. Warunkiem wystarczającym rozwiązalności (6.11.119) jest np. spełnianie nierówności

$$\begin{aligned} \sup_{\mathbf{v}_0 \in V_0} \frac{(A_N \mathbf{u}_0, \mathbf{v}_0)}{\|\mathbf{v}_0\|_A} &\geq \alpha \|\mathbf{u}_0\|_A, \quad \forall \mathbf{u}_0 \in V_0, \quad \alpha > 0 \\ \sup_{\mathbf{v}_0 \in V_0} |(A_N \mathbf{u}_0, \mathbf{v}_0)| &> 0, \quad \forall \mathbf{v}_0 \in V_0. \end{aligned} \quad (6.11.171)$$

(założenia uogólnionego twierdzenia Laxa-Milgrima). Rozpatrzmy pierwszą z nich.

$$\sup_{\mathbf{v}_0 \in V_0} \frac{(A_N \mathbf{u}_0, \mathbf{v}_0)}{\|\mathbf{v}_0\|_A} \geq \frac{(A_N \mathbf{u}_0, \mathbf{u}_0)}{\|\mathbf{u}_0\|_A} = \frac{(A \mathbf{u}_0, \mathbf{u}_0) + (N \mathbf{u}_0, \mathbf{u}_0)}{\|\mathbf{u}_0\|_A} \geq (1 - c_1 c_2) \|\mathbf{u}_0\|_A \quad (6.11.172)$$

Nierówność ta jest więc spełniona dla dostatecznie małego  $c_2$  (analogiczny rachunek obowiązuje dla drugiej z nierówności). Dla rozwiązań gładkich zachodzi to dla dostatecznie drobnej siatki. Rozstrzygnięcie problemu dla rozwiązań z falą uderzeniową wymaga dodatkowej analizy.

### 6.11.1 Uwagi dodatkowe

Przedstawiony algorytm uwarunkowania wstępnego jest bardzo bliski procedurze Jacobiego zastosowanej bezpośrednio do sformułowania niesymetrycznego. Aby to unaościć zapiszmy w postaci macierzowej operator  $B_N A_N$  dany przez (6.11.125) z macierzą symetrycznego operatora uwarunkowania wstępnego Jacobiego

$$\mathbf{T} = \mathbf{R}_0^T (\mathbf{A}_N)_0^{-1} \mathbf{R}_0 \mathbf{A}_N + \beta \left( \sum_{i=1}^N \mathbf{R}_i^T \mathbf{A}_i^{-1} \mathbf{R}_i \right) \mathbf{A}_N \quad (6.11.173)$$

gdzie  $\mathbf{A}_N$  jest macierzą operatora  $\mathbf{A}_N \{ \mathbf{A}_i \}_i = (\mathbf{A}_i \mathbf{v}_i, \mathbf{v}_i)$  oznacza symetryczną macierz sztywności dla podobszaru  $\Omega_i$ , zaś pierwszy składnik (6.11.173) jest macierzą operatora  $P_0$

obliczoną jak w (6.4.37). Zgodnie z prezentowaną analizą uwarunkowanie  $T$  jest niezależne od siatki. Modyfikując operator uwarunkowania wstępnego poprzez zastąpienie formy  $(A \cdot, \cdot)$  formą  $(A_N \cdot, \cdot)$  otrzymujemy zmienioną macierz z uwarunkowaniem wstępnym

$$T^* = R_0^T (A_N)_0^{-1} R_0 A_N + \beta \left( \sum_{i=1}^N R_i^T (A_N)_i^{-1} R_i \right) A_N \quad (6.11.174)$$

Zamianę zmiennych (6.11.130) można przedstawić za pomocą macierzy  $S$ ,  $U = SV$ . Analogiczna relacja obowiązuje dla wektorów stopni swobody  $U_i$  i  $V_i$  związanych z podobszarem  $\Omega_i$ , tj. istnieją takie macierze  $S_i$ , że  $U_i = S_i V_i$ ,  $i = 1, \dots, N$ . Z definicji macierzy  $R_i^T$  wynika, że

$$S^{-1} R_i^T = R_i^T S_i^{-1} \quad \text{i} \quad S R_i^T = R_i^T S_i, \quad i = 1, \dots, N \quad (6.11.175)$$

(Powyższe własności macierzy  $S$  są rezultatem faktu, że podstawienie (6.11.130) wiąże ze sobą stopnie swobody danego węzła lokalnie, niezależnie od węzłów pozostałych.) Ponadto macierz sztywności oryginalnego sformułowania niesymetrycznego  $B = \{B_{ij}\}$ ,  $B_{ij} = B(\psi_i, \psi_j)$  a także jej podmacierze  $B_i$  odpowiadające podobszaram  $\Omega_i$  są związane z  $A_N$  i  $(A_N)_i$  następująco

$$B S = A_N \quad \text{i} \quad B_i S_i = (A_N)_i, \quad i = 1, \dots, N \quad (6.11.176)$$

Mnożąc równanie (6.11.174) z lewej strony przez  $S$  i z prawej przez  $S^{-1}$  (co odpowiada powrotnej zamianie stopni swobody  $V = S^{-1}U$ ) otrzymujemy macierz o identycznym spektrum jak  $T^*$

$$S T^* S^{-1} = S R_0^T (A_N)_0^{-1} R_0 A_N S^{-1} + \beta \sum_{i=1}^N R_i^T B_i^{-1} R_i B \quad (6.11.177)$$

Jest to więc z dokładnością do pierwszego składnika i czynnika  $\beta$  algorytm uwarunkowania wstępnego typu Jacobiego dla oryginalnej niesymetrycznej macierzy  $B$ .

Interesujący jest problem, czy możliwe jest pominięcie kroku globalnego  $P_0$ . Oznaczałoby to dopuszczenie  $\delta_0 = c_1$  (por. (6.11.117) i (6.11.120)) i granicznej wartości  $\varepsilon^2 \geq c_1^2$ , (6.11.127), tj.

$$\varepsilon^2 = \frac{(\lambda_0/\lambda_1)^2}{4c_2^2(c_1^2 c_2^2 + 1)} \geq c_1^2 \quad (6.11.178)$$

Przy ustalonych  $\lambda_0, \lambda_1, c_1$  jest to możliwe dla odpowiednio małego  $c_2$ , tj. jeśli składnik niesymetryczny  $N$  operatora  $A_N$  jest odpowiednio mały. Możliwość taką sugeruje wcześniejsza analiza wielkości  $c_2$ , zwłaszcza dla przypadku rozwiązań gładkich, gdy  $c_2 \sim h_{max}$ . W praktyce okazuje się, że istotnie, algorytm (6.11.177) z pominięciem kroku globalnego prowadzi do efektywnego uwarunkowania wstępnego.

# Rozdział 7

## Przykłady numeryczne

### 7.1 Uwagi wstępne

Przedstawioną w poprzednich rozdziałach technikę rozwiązywania przepływów naddźwiękowych z małą lepkością ilustrujemy kilkoma przykładami takiej analizy. Pierwsze dwa z nich to porównawcze zadania (tzw. *benchmark problems*), które bywają stosowane do weryfikowania wiarygodności programów obliczeniowych dynamiki płynów. Są to przepływy z dużą liczbą Macha (hiperdźwiękowe), z liczbą Reynoldsa rzędu  $Re = 2 \cdot 10^5$  (odniesioną do wymiaru obiektu), w których dochodzi do interakcji fal uderzeniowych z warstwą przyścienną. Przepływy te były badane eksperymentalnie, istnieje więc możliwość porównania wyników symulacji numerycznej z fizyczną rzeczywistością. Eksperymenty przeprowadzono w laboratorium NASA Langley Research Center w USA i są one związane bardziej z techniką lotów kosmicznych niż z klasyczną aerodynamiką lotniczą – parametry przepływów odpowiadają warunkom lotu na bardzo dużej wysokości z dużą liczbą Macha. Pozostałe dwa przykłady to przepływy wokół walca z liczbą Macha  $M = 8$  i liczbą Reynoldsa odpowiednio  $Re = 2 \cdot 10^5$  i  $Re = 2 \cdot 10^6$ . Dla tych przepływów możliwość weryfikacji daje porównanie z wynikami otrzymanymi na podstawie uproszczonej teorii cienkich laminarnych warstw przyściennych.

Prezentowane przykłady uchodzą za względnie trudne. Wybierając je wychodziliśmy z założenia, że tylko zadania o odpowiedniej skali trudności dają możliwość faktycznej weryfikacji przydatności proponowanych metod. Taki wybór wiązał się jednak z dużym nakładem pracy obliczeniowej. Wspomnijmy tu tylko, że zbiegnięcie się zadania do stanu ustalonego wymagało kilkuset lub nawet kilku tysięcy kroków czasowych, zaś każdy z nich to zbudowanie i rozwiązanie układu równań o 20–30 tys. stopni swobody, przy czym przez większość czasu powstawania tej pracy mogliśmy dysponować mocą obliczeniową 2 Mflops i 16 MB pamięci. Różne pomysły *ad hoc* omińcía tej bariery, w rodzaju zwiększenia kroku czasowego czy rozwiązywania układu równań z ograniczoną dokładnością, najczęściej zawodziły w myśl żartobliwie wyrażanej prawdy, iż „w zadaniach przepływu stabilne jest tylko to, że są one niestabilne”, tj. powodując właśnie utratę stabilności. Prowadzenie bardzo wielu doświadczeń numerycznych dotyczących skuteczności testowanych procedur jest w takich warunkach bardzo uciążliwe.

Kilka uwag ogólnych dotyczących sposobu rozwiązywania zadań. Krok czasowy  $\Delta t$  przyjmowano tak, że odpowiadał liczbie Couranta–Friedrichsa–Leviego  $CFL$  z przedziału od 0.5 do 5.0 (por. podroz. 3.2), tylko dla przepływu nad płytą możliwe było przyjęcie  $CFL = 25$  (w końcowej fazie).

Dokładność rozwiązywania układu równań dyskretnych kontrolowano zakładając redukcję residuum o czynnik  $10^{-10}$ . Liczba iteracji między restartami algorytmu GMRES wynosiła od 6 do 10 (odpowiada to liczbie przechowywanych w pamięci wyników kolejnych iteracji prostych), redukcja residuum przypadająca na jedną iterację typu Gauss-Seidel wykonywaną w procedurze GMRES wahała się w granicach 0.2-0.5. Siatkę zgrubną generowano dla parametrów  $q_1 = q_2 = 1$ , tzn. że pokrywała się ona z siatką oryginalną poza obszarami z elementami wydłużonymi, przy czym krok globalny był w iteracjach pomijany. Taki wybór metody rozwiązywania był uwarunkowany ograniczeniem pamięci. Zadania były rozwiązywane bez zapisywania na dysk, w pamięci, wymagało to zarezerwowania dla algorytmu GMRES macierzy 4.5 mln liczb. Obliczenia prowadzono w podwójnej precyzji, tj. z dokładnością ok. 15 cyfr znaczących.

Zbieżność zadań do stanu ustalonego badano odnosząc maksymalną różnicę rozwiązań z dwu kolejnych kroków czasowych  $\delta$  do przedziału czasowego  $\Delta t$ . Dla siatek z kilkoma poziomami podziałów wielkość  $\delta/\Delta t$  była redukowana do wartości poniżej 1. Dla siatek z ekstremalnie małymi elementami pojawiały się punkty w warstwie przyściennej, gdzie  $\delta/\Delta t$  utrzymywało się na poziomie 10 nie ulegając redukcji. Nie znamy przyczyny tego zjawiska. Wielkość  $\delta/\Delta t$  w pierwszym kroku po dokonaniu adaptacji w warstwie przyściennej sięgała  $10^5$ .

Maksymalne wydłużenie elementów w warstwie przyściennej sięgało 1000.

Wszystkie wyniki numeryczne przedstawione są w zmiennych bezwymiarowych (zob. roz. 2.)

## 7.2 Przepływ wokół naroża wklęsłego, $M = 14.1$ , $Re = 72000/ft$

W problemie tym gaz porusza się poziomo od lewej strony napotykając na swej drodze płytę, która jest płaska na długości 1.44 ft, po czym ulega zagięciu o kąt  $24^\circ$  (*flow over a  $24^\circ$  compression corner*). Przepływ ten był badany doświadczalnie przez M.S. Holdena [50] i analizowany numerycznie w pracach [111],[103] i [95]. Rys. 7.1 przedstawia geometrię problemu i jednocześnie ostateczną siatkę adaptacyjną. Dane dla zadania są następujące

liczba Macha	$M = 14.1$
liczba Reynoldsa	$Re = 72000 \text{ ft}^{-1}$
temperatura gazu	$\theta_\infty = 88.9^\circ \text{ K}$
temperatura płyty	$\theta_{\text{wall}} = 294.4^\circ \text{ K}$

( $Re$  jest tu jednostkową liczbą Reynoldsa (*unit Reynolds number*), tj. odniesioną do jednostki długości, np. 1 ft czy 1 m; jednostkowa liczba Reynoldsa jest oznaczana identycznie jak liczba Reynoldsa odniesiona do długości  $L$ , ma jednak wymiar, np.  $\text{ft}^{-1}$  czy  $\text{m}^{-1}$ .) Zadanie rozwiązano stosując anizotropową strategię adaptacyjną typu  $h$ . Siatka początkowa składała się z  $23 \times 11$  elementów. Zastosowano 4 poziomy adaptacji do otrzymania rozwiązania o uznanej za wystarczającą dokładności we wnętrzu obszaru obliczeniowego. Wykorzystano na tym etapie strategię adaptacji z podrozdziału 5.3 opartą na oszacowaniu residualnym (na pojedynczych elementach) z podziałami kierunkowymi. W drugim etapie procesu adaptacji zastosowano strategię ulepszania rozwiązania w warstwie przyściennej (podrozdział 5.4). Dla zredukowania wskaźników błędów strumieni lepkich do ok. 1% dokonano 11 dodatkowych, w większości kierunkowych, podziałów elementów w

warstwie przyściennej. Tym sposobem elementy przyległe do płyty poddano 15 poziomom podziałów, w wyniku czego minimalny rozmiar elementów osiągnął  $2^{-15}h \approx 1/32000h$ , gdzie  $h = 0.01$  jest wysokością elementów siatki pierwotnej.

W procesie rozwiązywania zadania napotkaliśmy na następujące trudności:

1. Rozwiązanie zachowywało się niestabilnie w punkcie zastoju (wierzchołek płyty). Ominęliśmy ten problem przez dodatkowe zastosowanie sztucznej lepkości Löhnera-Morgana-Peraire [70] w elementach bezpośrednio stykających się z wierzchołkiem.
2. Rozmiar elementów w otoczeniu wierzchołka płyty istotnie wpływał na położenie fal uderzeniowych. Stabilizację tego położenia udało się otrzymać przez dodatkową istotną redukcję rozmiaru elementów w bezpośredniej bliskości wierzchołka.
3. Fala uderzeniowa odbita od nachylonej części płyty jest szczególnie silna w obszarze interakcji z warstwą przyścinną, skok ciśnienia  $p_2/p_1$  w tej fali przekracza 100. Byliśmy zmuszeni zwiększyć sztuczną lepkość  $\nu_d$  daną wzorem (3.1.3) o czynnik 1.5, aby zapobiec oscylacjom ciśnienia przed falą uderzeniową prowadzącym do jego ujemnych, нефизycznych wartości. Zwiększenie  $\nu_d$  niestety odbiło się na zmniejszeniu ostrości rozwiązania w całym obszarze obliczeniowym. Bliższa analiza wzoru (3.1.3) wskazuje, że dla istotnie wydłużonych elementów stosunek  $\nu_d$  do wymiarów elementów w kierunku prostopadłym do fali jest  $\sqrt{2}$  razy mniejszy niż dla elementów niezbyt wydłużonych. Sugeruje to konieczność badań nad zakresem stosowalności wzoru (3.1.3).

Rys. 7.2 przedstawia mapy wybranych charakterystyk rozwiązania w obszarze warstwy przyściennej w pobliżu jej interakcji z falą uderzeniową. Siatka skończenie elementowa jest pokazana z zastosowaniem logarytmicznego przeskalowania w kierunku prostopadłym do brzegu w postaci  $y' = \log_2(y + \varepsilon)$ ,  $\varepsilon = 2^{-12}h$ . Takie przeskalowanie odwzorowuje oryginalną siatkę z rozmiarem elementów zmieniającym się wg postępu geometrycznego w siatkę bliską równomiernej w kierunku pionowym, dzięki czemu prezentacja jest możliwa. Rozważany obszar zajmuje 1/8 rozmiaru elementów pierwotnych w kierunku pionowym i połowę długości nachylonej płyty. Rys. 7.2 (a-h) przedstawiają rozkłady następujących wielkości:

- (a) lokalną liczbę Macha (dla wskazania zasięgu warstwy przyściennej)
- (b) residualne wskaźniki błędu
- (c) interpolacyjne wskaźniki błędu dla  $\nabla U$  (por.(5.4.28))
- (d) parametr  $\tau_a$  metody SUPG (por.(3.1.3))
- (e,f) parametr  $\nu_d$  oraz współczynnik sztucznej lepkości  $\mu$  przeskalowane w taki sposób, aby reprezentowana przez nie lepkość mogła być porównana z lepkością wyrazu z  $\tau_a$

$$\bar{\nu}_d = \frac{\nu_d}{(c + |\mathbf{u} \cdot \boldsymbol{\beta}|)^2} \quad \text{and} \quad \bar{\mu} = \frac{\mu/\rho}{(c + |\mathbf{u} \cdot \boldsymbol{\beta}|)^2} \quad (7.2.1)$$

(powyższe przeskalowanie zaczęliśmy z [4]).



(g,h) następujące symetryczne formy dwuliniowe obliczone dla  $\nabla U$

$$a_1 = \tau_a \frac{\partial U^T}{\partial x_i} \mathbf{A}_i^T \mathbf{A}_0^{-1} \mathbf{A}_j \frac{\partial U}{\partial x_j} \quad \text{and} \quad a_2 = \frac{\partial U^T}{\partial x_i} \mathbf{A}_0^{-1} \mathbf{K}_{ij} \frac{\partial U}{\partial x_j} \quad (7.2.2)$$

(wielkości te można zinterpretować fizycznie jako lokalny wkład do produkcji entropii wynikający z obecności lepkości sztucznej  $\tau_a \mathbf{A}_i \mathbf{A}_j$  i naturalnej  $\mathbf{K}_{ij}$ ).

W podrozdziale 5.4 sugerowaliśmy możliwość dokonywania dodatkowych podziałów elementów w warstwie przyściennej prowadzących do redukcji parametrów stabilizujących  $\tau_a$  i  $\nu_d$ . Drugi z nich wydaje się być zanedbywalnie mały w porównaniu z naturalną lepkością  $\mu$ , por. rys. 7.2 (e) i (f). Nie jest tak w przypadku  $\tau_a$ , które nawet przekracza  $\mu$ . Z drugiej jednak strony wpływ wyrazu  $\tau_a$  i lepkości naturalnej mierzony za pomocą form dwuliniowych  $a_1$  i  $a_2$  wskazuje na wystarczającą redukcję mechanizmu sztucznej stabilizacji. Zauważmy też, że residualne wskaźniki błędu i wskaźniki interpolacyjne nie zostały zredukowane wewnątrz warstwy przyściennej do poziomu w elementach przyległych do płyty. W procesie adaptacji jest także możliwość redukcji tych wskaźników (przez podział elementów, w których przekraczają one poziom przy brzegu). W pewnym sensie jest to nieco niespodziewane, że, mimo iż nie przeprowadzono takiej adaptacji skupiając się wyłącznie na redukcji wskaźników błędu w elementach przyległych do płyty, dokładność strumieni lepkich jest zaskakująco dobra. Rys. 7.3-7.5 przedstawiają współczynniki tarcia lepkiego

$$C_\tau = \frac{\tau}{1/2 \rho_\infty V_\infty^2}, \quad (7.2.3)$$

współczynnik przepływu ciepła

$$C_h = \frac{q}{\rho_\infty V_\infty \gamma C_v (\theta_{0,\infty} - \theta_{wall})} \quad (7.2.4)$$

(gdzie  $\theta_{0,\infty}$  jest temperaturą spiętrzenia,  $\theta_{0,\infty} = \theta_\infty (1 + (\gamma - 1)/2 \cdot M^2)$ ) i współczynnik ciśnienia

$$C_p = \frac{p}{1/2 \rho_\infty V_\infty^2} \quad (7.2.5)$$

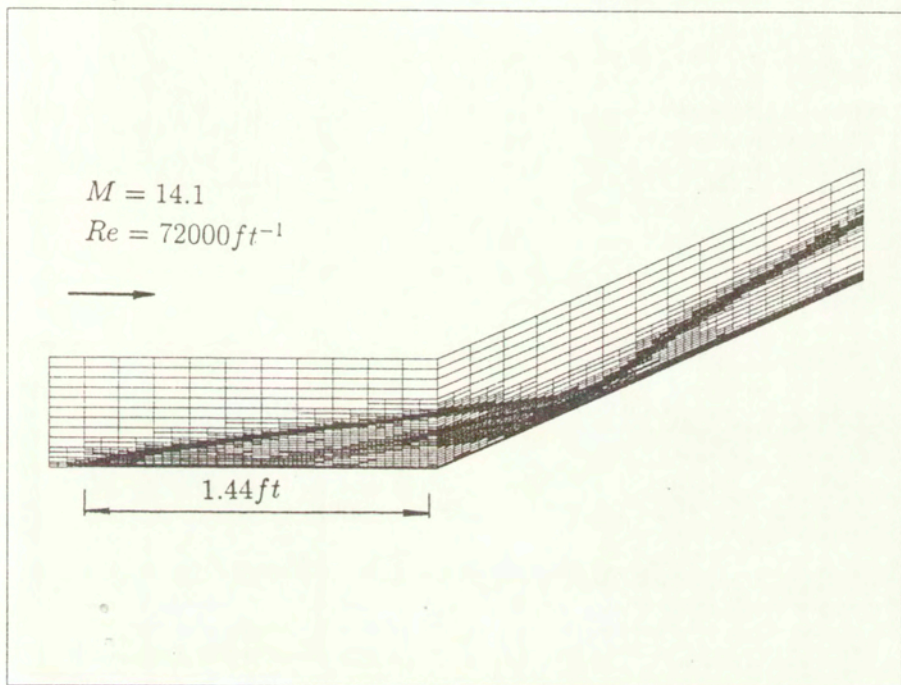
wzdłuż płyty wraz z pomiarami doświadczalnymi Holdena [50]. Strumienie lepkie są zgodne z doświadczeniem, mimo że ciśnienie jest nieco zaniżone sugerując konieczność dalszego adaptacyjnego poprawiania rozwiązania jako całości. Rys. 7.6 przedstawia ostateczny wskaźnik błędu dla strumieni lepkich,  $\max(\Delta\tau_K/\tau_{ref}, \Delta q_K/q_{ref})$  z  $\tau_{ref} = \max_K \tau_K$ ,  $q_{ref} = \max_K q_K$  otrzymane po dokonaniu 15 poziomów adaptacji.

Rys. 7.7 i 7.8 przedstawiają rozkład gęstości  $\rho$  i pionowej składowej prędkości  $v$ . Trójwymiarowe wykresy gęstości i ciśnienia przedstawiono na rys. 7.9 i 7.10.

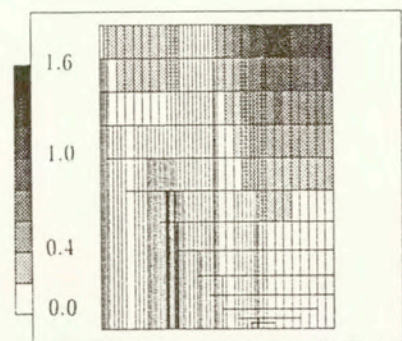
Końcowa siatka adaptacyjna składała się z 6700 elementów i 6200 węzłów.

### 7.3 Przepływ wokół walca z interferencją fal uderzeniowych, $M = 8$ , $Re = 2 \cdot 10^5$

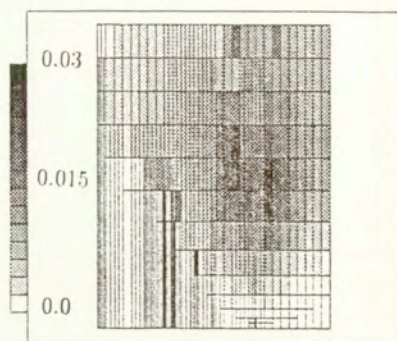
Wymieniony w tytule problem modeluje spotykane w lotnictwie zjawisko zderzenia się fali uderzeniowej z zakrzywioną częścią powierzchni samolotu naddźwiękowego. Dochodzi wówczas do złożonej interferencji „zewnątrznej” fal uderzeniowej z falą uderzeniową



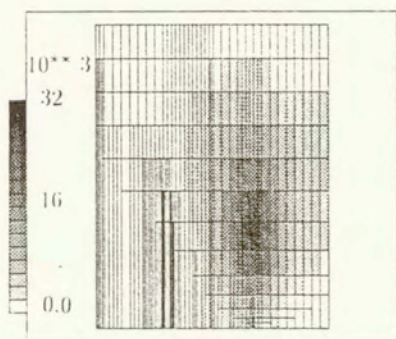
Rysunek 7.1: Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ .  
Siatka adaptacyjna i definicja problemu.



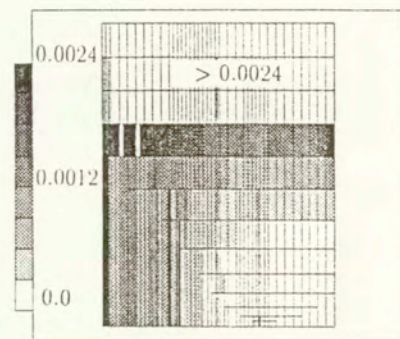
a



b

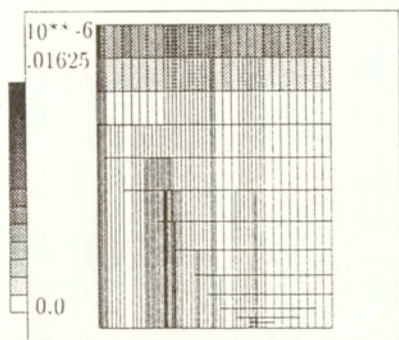


c

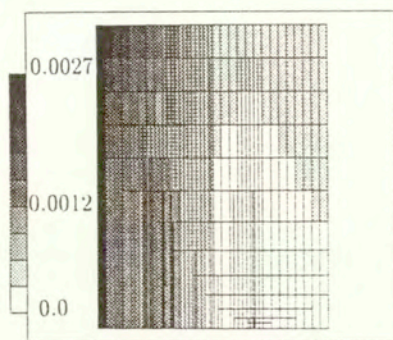


d

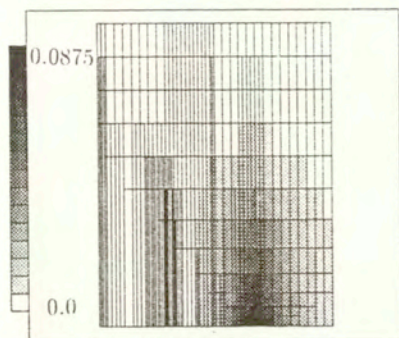
Rysunek 7.2 (a-d) Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ . Niektóre charakterystyki rozwiązania w obszarze interakcji. Pokazany obszar zajmuje połowę nachylonej płyty począwszy od naroża i  $1/8$  wysokości elementów siatki początkowej  $h$  od brzegu (por. rys. 7.1), współrzędną pionową przeskalowano wg wzoru  $y' = \log_2(y + \epsilon)$ ,  $\epsilon = 2^{-12}h$  (szczegółowy opis na str. 99).



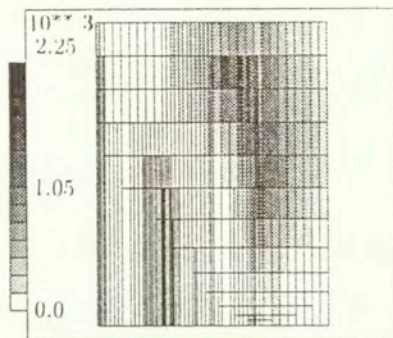
e



f

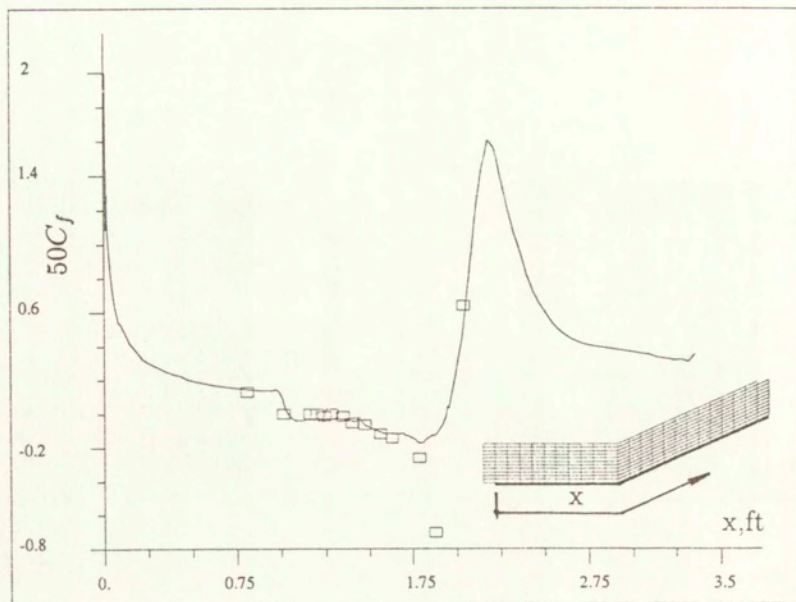


g

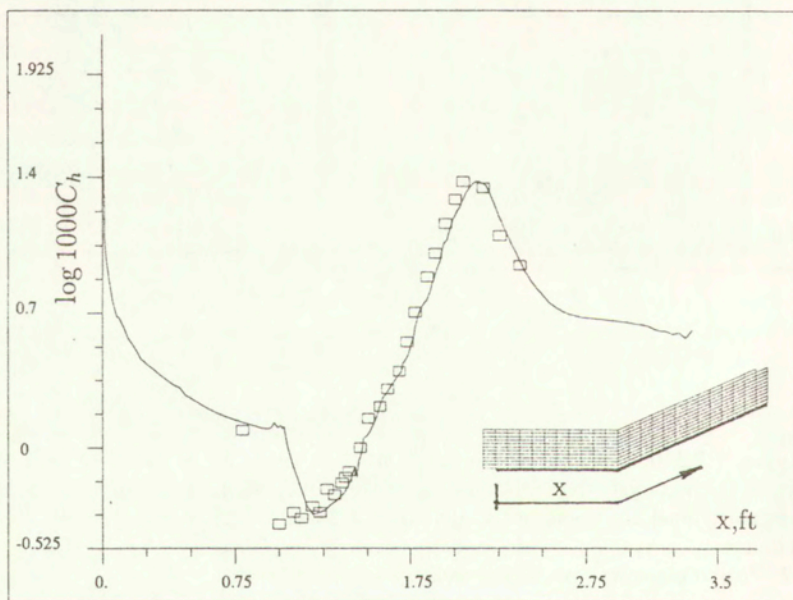


h

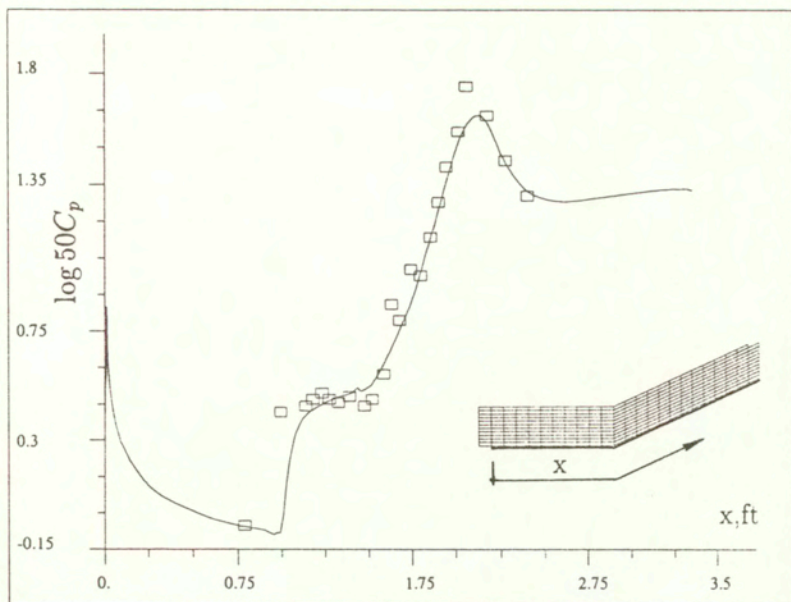
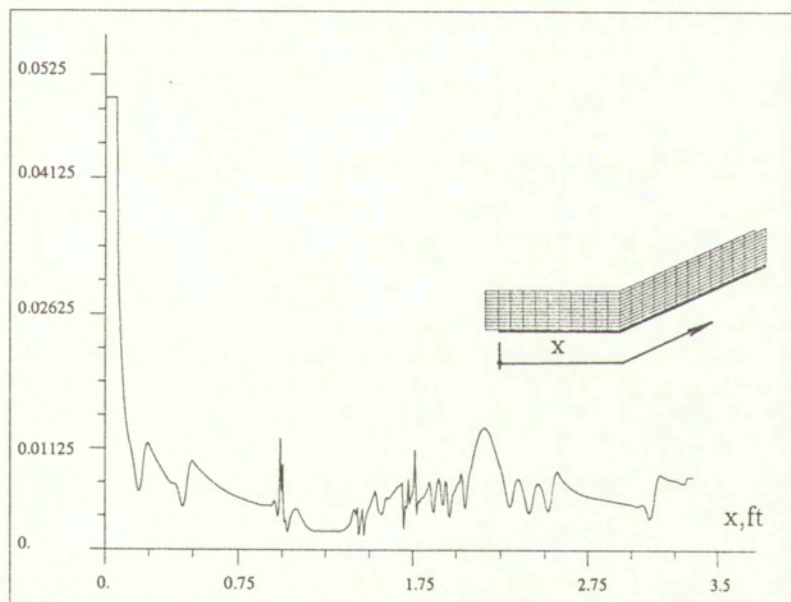
Rysunek 7.2 (e-h) Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ . Niektóre charakterystyki rozwiązania w obszarze interakcji. Pokazany obszar zajmuje połowę nachylonej płyty począwszy od naroża i  $1/8$  wysokości elementów siatki początkowej  $h$  od brzegu (por. rys. 7.1), współrzędną pionową przeskalowano wg wzoru  $y' = \log_2(y + \varepsilon)$ ,  $\varepsilon = 2^{-12}h$  (szczegółowy opis na str. 99).



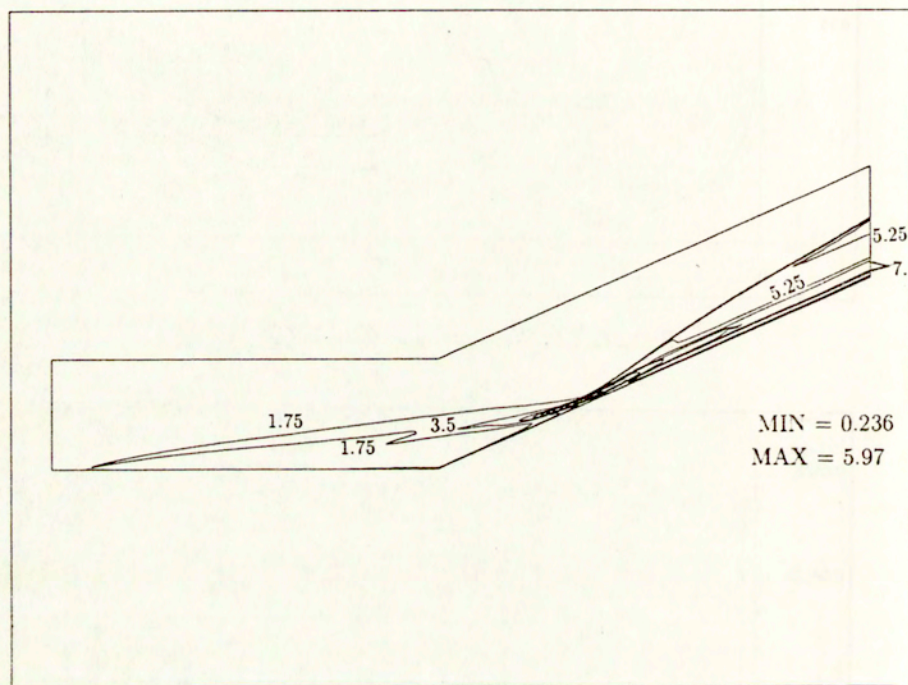
Rysunek 7.3: Współczynnik tarcia powierzchniowego  $C_f$ .



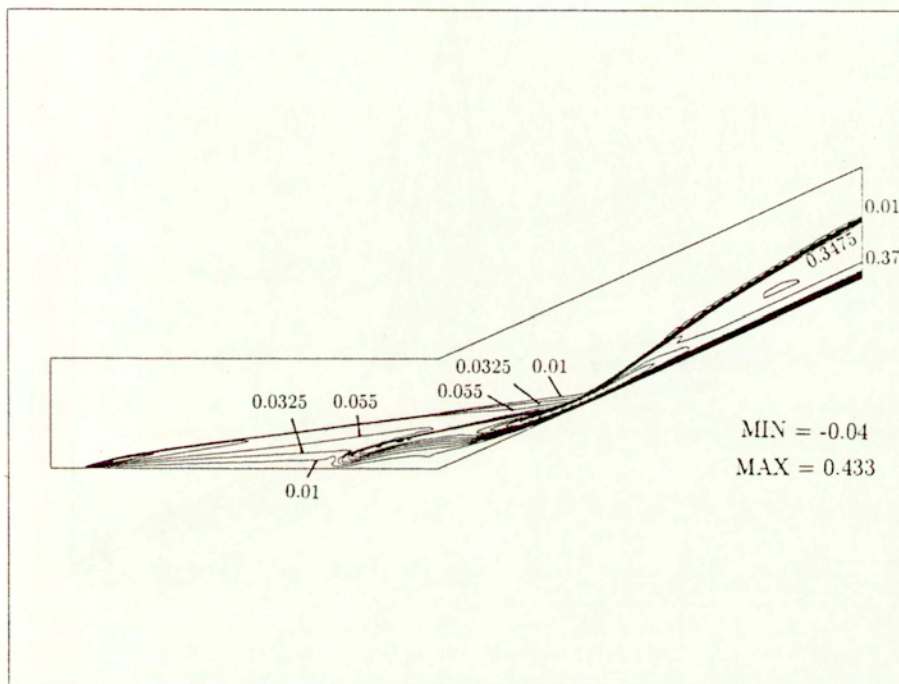
Rysunek 7.4: Współczynnik przepływu ciepła  $C_h$ .

Rysunek 7.5: Współczynnik ciśnienia  $C_p$ .

Rysunek 7.6: Wskaźnik błędów dla strumieni lepkich.

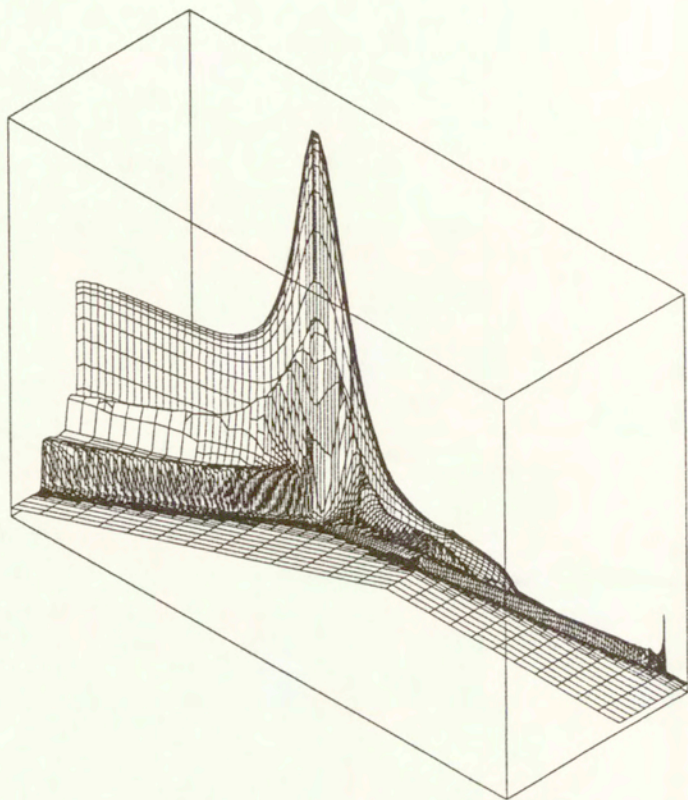


Rysunek 7.7: Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ .  
Rozkład gęstości.

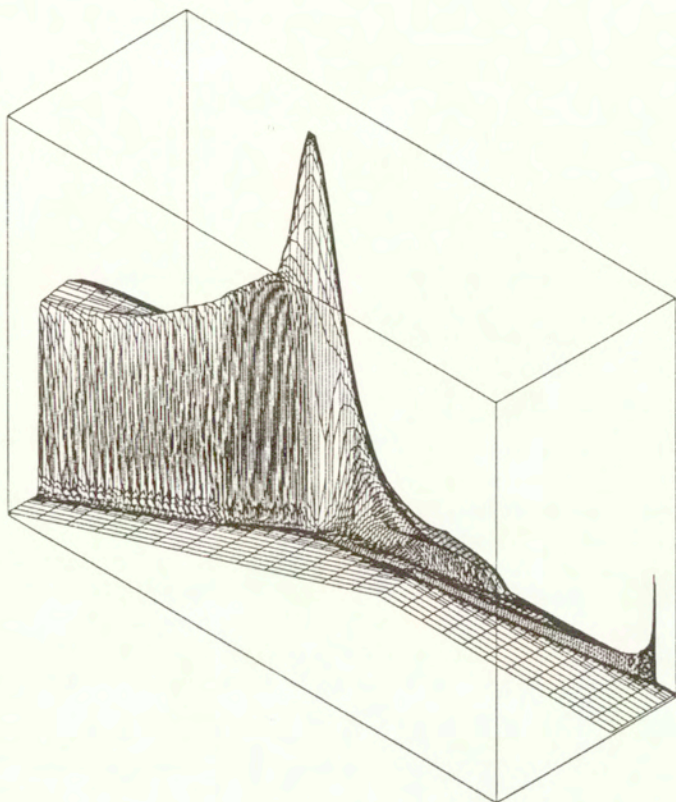


Rysunek 7.8: Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ .  
Pionowa składowa prędkości.





Rysunek 7.9: Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ .  
Wykres gęstości.



Rysunek 7.10: Przepływ wokół naroża wklęsłego,  $M = 14.1$ ,  $Re = 72000/ft$ .  
Wykres ciśnienia.

otaczającą walec w niezaburzonym przepływie. W wyniku tej interferencji pojawia się też interakcja fali uderzeniowej z warstwą przyścienną przy powierzchni walca, co wiąże się z silnym skoncentrowanym wzrostem ciśnienia i nagrzewania w obszarze interakcji. Ostatni z wymienionych efektów może być przyczyną przepalenia poszycia samolotu (zob. np. [6] str. 322). Opisany przepływ był badany eksperymentalnie przez A.R. Wietinga i M.S. Holdena [110]. Zewnętrzna fala uderzeniowa wywoływana była przez wstawienie do niezaburzonego przepływu płytki nachylonej pod kątem  $12.5^\circ$ .

Dane do zadania są następujące:

liczba Reynoldsa	$Re = 2.01 \cdot 10^5$
Temperatura walca	$\theta_{wall} = 194.4^\circ K$
Obszar A (nad falą uderzeniową)	
liczba Macha	$M = 8.03$
temperatura gazu	$\theta_\infty = 122.2^\circ K$
Obszar B (pod falą uderzeniową)	
liczba Macha	$M = 5.25$
temperatura gazu	$\theta_\infty = 260.6^\circ K$

(liczba Reynoldsa jest odniesiona do promienia walca  $r$ ). Do symulacji zjawiska przyjęto obszar obliczeniowy pokazany na rys. 7.11. Fala uderzeniowa przecina obszar obliczeniowy w odległości  $0.4r$  poniżej osi symetrii walca. Niestety dla serii wyników doświadczalnych z [110] tej lokalizacji fali uderzeniowej nie podano, w pracy tej „zmienną niezależną” jest położenie maksimum ciśnienia czy strumienia ciepła na obwodzie walca. Dlatego też otrzymane przez nas wyniki możemy porównać jedynie na podstawie interpolacji wyników doświadczalnych z różnymi położeniami wspomnianych maksimów.

Siatka początkowa składała się z  $16 \times 8$  elementów. Dodatkowo wprowadzono w niej od samego początku trzy poziomy podziały elementów przyległych do walca dla zapewnienia dokładnej aproksymacji geometrii walca (jeden element przypada na ok.  $1.5^\circ$  łuku).

Zgodnie z ogólną zasadą strategii adaptacji w pierwszym jej etapie dokonano 4 poziomów adaptacyjnych podziałów w całym obszarze obliczeniowym otrzymując rozwiązanie, o którym założyliśmy, iż jest wystarczająco dokładne (stosowano residualne oszacowanie błędu wg Ainswortha–Odena). W następnym kroku dokonano adaptacji w warstwie przyściennej, ograniczając się do podziałów elementów przylegających do walca. Dla zredukowania wskaźników błędu strumieni lepkich poniżej 1% konieczne było wprowadzenie 19 poziomów podziału elementów w obszarze największej koncentracji  $\tau, q$  i  $p$ , co odpowiada redukcji rozmiaru elementu o czynnik  $1/500000$ . Na rys. 7.12 i 7.13 pokazano aksonometrię trójwymiarowego wykresu funkcji gęstości  $q$ . Na szczególną uwagę zasługuje tu „pionowa ściana” tego wykresu wzdłuż powierzchni walca, która odpowiada gwałtownej zmienności  $q$  w warstwie przyściennej. Na rys. 7.14 przedstawiamy wykres  $q$  wzdłuż osi symetrii obszaru obliczeniowego, na którym wyraźnie odróżniamy skok rozwiązania w fali uderzeniowej i gwałtowną zmienność w warstwie przyściennej. Tę ostatnią przedstawiamy dodatkowo w 16-krotnym powiększeniu umożliwiającym obserwację zmienności  $q$ . Na rys. 7.15 pokazano rozkład ciśnienia wraz z powiększeniem obszaru interakcji. Rys. 7.16 i 7.17 przedstawiają aksonometrie funkcji ciśnienia. Obciążenia aerotermiczne, ciśnienie, strumień ciepła i tarcie powierzchniowe zostały przedstawione na rys. 7.18–7.20.

Dla ciśnienia możliwe było porównanie wyników numerycznych i pomiarów doświadczalnych z [110], gdyż położenie maksimum  $p$  w przedstawionym rozwiązaniu i w pomiarach jest zbliżone ( $-13.7^{\circ}$  i  $-16.7^{\circ}$ ). W przypadku strumieni ciepła różnica położenia maksimum jest bardziej istotna:  $-13.7^{\circ}$  i  $-18.8^{\circ}$  (rozejście się doświadczalnych położenia maksimum  $p$  i  $q$  jest nieco zaskakujące). Wartość maksymalnego  $q$  dla jego lokalizacji otrzymanej w obliczeniach można znaleźć korzystając z krzywej regresji dla wyników doświadczalnych  $q_{max}(\alpha_0)$ , gdzie  $\alpha_0$  jest lokalizacją maksimum (rys. 10. w [110]). Wyniki te mają jednak nieco niejednoznaczną interpretację:

- Z krzywej regresji odczytujemy dla  $\alpha_0 = -13.7^{\circ}$ ,  $q_{max} = 0.054$ .
- Punktem doświadczalnym najbliższym wartości  $-13.7^{\circ}$  jest  $\alpha_0 = -16.5^{\circ}$ . Dla tego położenia  $q_{max}(\alpha_0) = 0.042$ .
- Wartość  $q_{max}(-13.7^{\circ})$  można ekstrapolować ( $-13.7^{\circ}$  leży poza szeregiem punktów doświadczalnych) wychodząc z powyższego wyniku i zakładając tendencję zmienności  $q_{max}(\alpha_0)$  identyczną jak w krzywej regresji, otrzymujemy wówczas  $q_{max} = 0.035$

Z obliczeń numerycznych otrzymaliśmy  $q_{max} = 0.0315$ , co różni się od powyższych wartości odpowiednio o 42%, 25% i 10%. Trudno jest ocenić, jak daleki jest faktycznie wynik numeryczny od fizycznej rzeczywistości. Autorzy niestety nie podali zasad, którymi kierowali się tworząc krzywą regresji, zaznaczyli w jej pobliżu tylko 5 punktów numerycznych, z których 4 są poniżej krzywej a 1 powyżej. Dla punktu doświadczalnego wymienionego w b) krzywa regresji różni się od pomiaru o 36%. Wydaje się, że za realistyczny można uznać wniosek, iż wynik numeryczny ma dokładność w granicach 10–25%

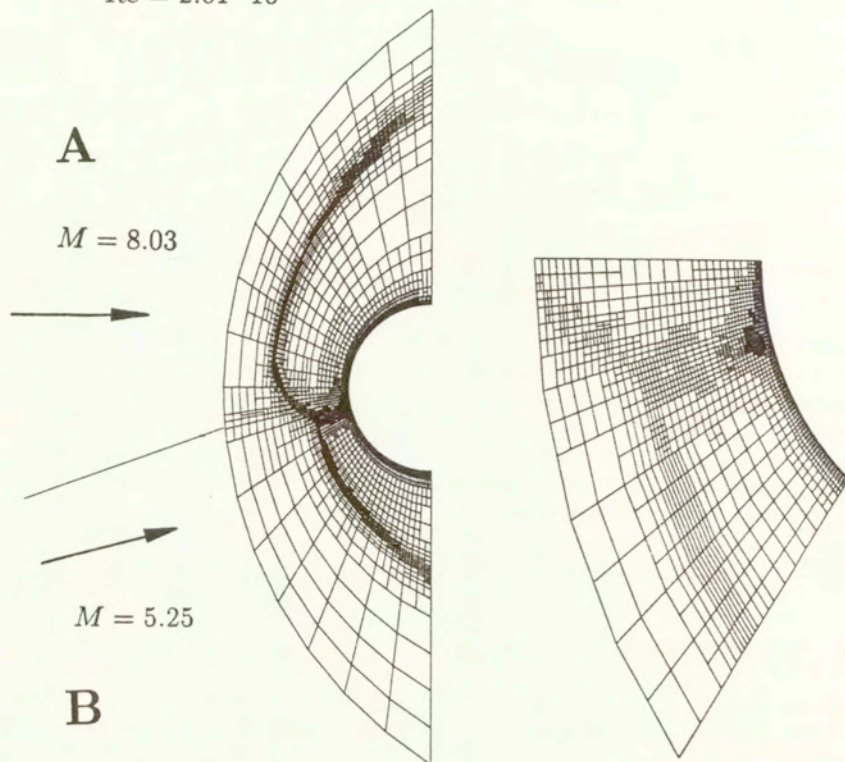
Końcowa siatka adaptacyjna dla tego zadania składała się z 6460 elementów i 6260 węzłów.

## 7.4 Optym wokół walca, $M = 8$ , $Re = 2 \cdot 10^5$

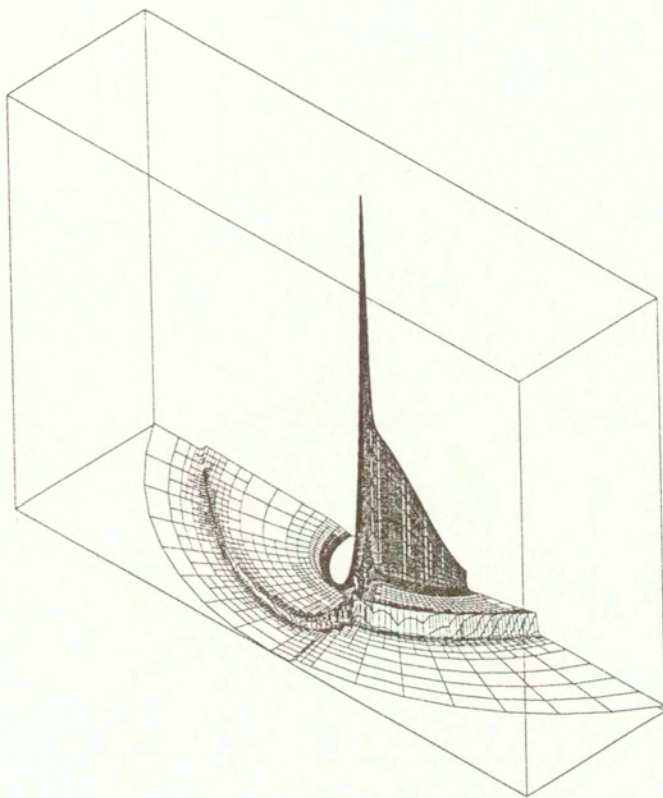
Poprzednie zadanie rozwiązano także dla przepływu niezaburzonego scharakteryzowanego danymi jak dla obszaru A. Dzięki symetrii problemu można się było ograniczyć do symulacji w górnej połowie obszaru obliczeniowego. Rys. 7.22 przedstawia ten obszar wraz z siatką adaptacyjną złożoną z 2620 elementów i podobnej liczby węzłów. W warstwie przyściennej zastosowano 14 poziomów podziałów elementów dla zredukowania wskaźników błędu poniżej 1%. Rys. 7.24 przedstawia aksonometrię trójwymiarowego wykresu gęstości  $\rho$ . Podobnie jak poprzednio prócz skoku rozwiązania wzdłuż fali uderzeniowej obserwujemy „pionową ścianę” wzdłuż powierzchni walca – gwałtownie zmieniające się  $\rho$  w warstwie przyściennej. Na rys. 7.25 pokazano wykres  $\rho$  wzdłuż osi symetrii wraz z 16-krotnym powiększeniem warstwy przyściennej. Podobny wykres dla temperatury pokazano na rys. 7.29. Na rys. 7.23 przedstawiono mapę ciśnienia. Na kolejnych rys. 7.26–7.28 przedstawiono wykresy ciśnienia, strumienia ciepła i tarcia powierzchniowego wzdłuż powierzchni walca. Obliczony strumień ciepła możemy porównać z wartością wynikającą z następującego wzoru otrzymanego dla punktu zastoju na podstawie teorii laminarnych warstw przyściennych (wzór dla sfery podali Fay i Riddell [42], jego modyfikację dla cylindra podaje praca [38])

$$q = 0.57 \left( \frac{\rho \cdot c_p \cdot U_{\infty}^3}{Re} \right)^{1/2} \cdot \sqrt{\frac{du}{dx}} \cdot \phi(\theta_e - \theta_w) \quad (7.4.6)$$

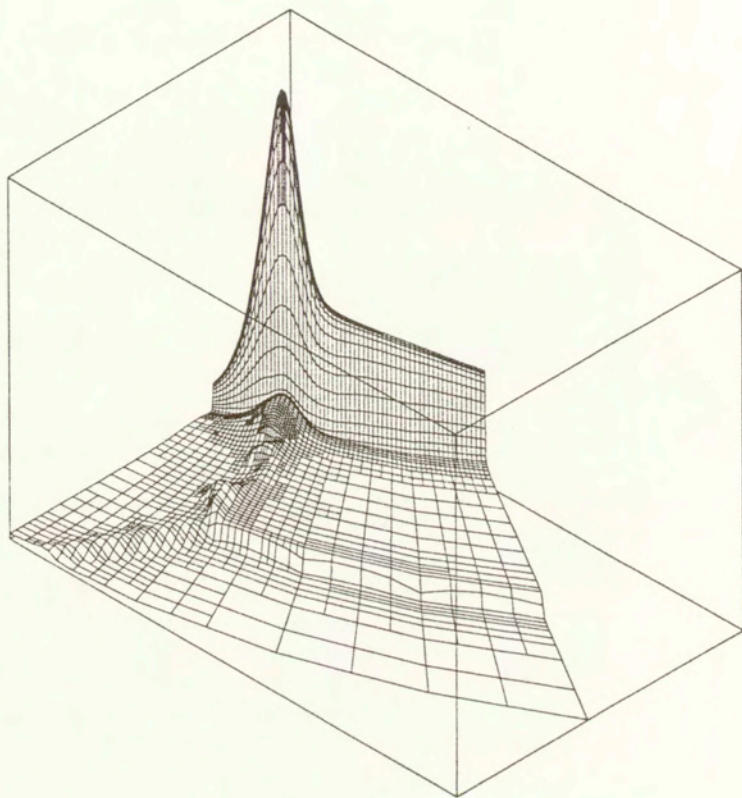
$$Re = 2.01 \cdot 10^5$$



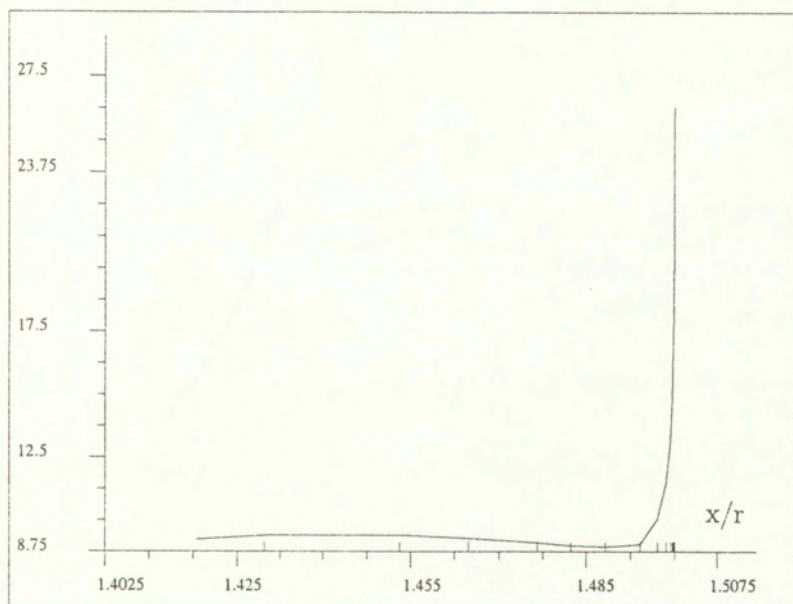
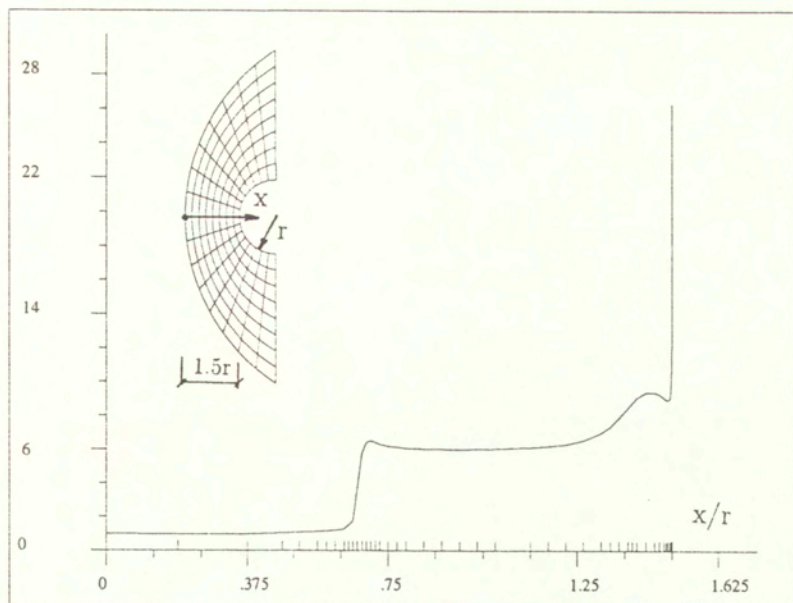
Rysunek 7.11: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Siatka adaptacyjna i definicja zadania.



Rysunek 7.12: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Wykres gęstości w obszarze obliczeniowym.

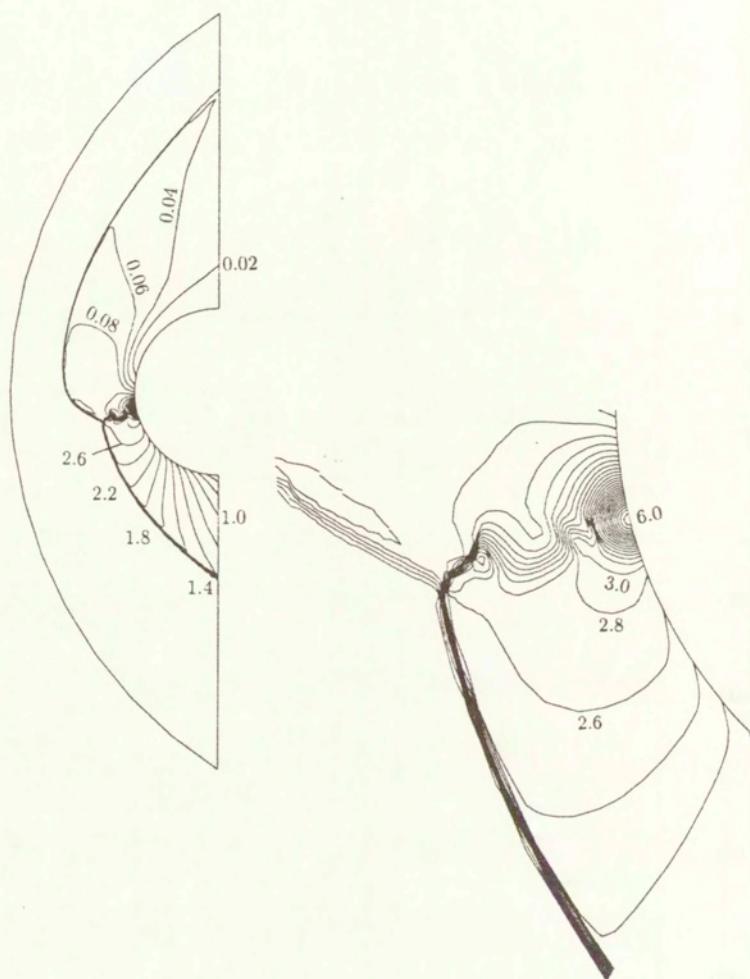


Rysunek 7.13: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Wykres gęstości w obszarze obliczeniowym - powiększenie obszaru interakcji.

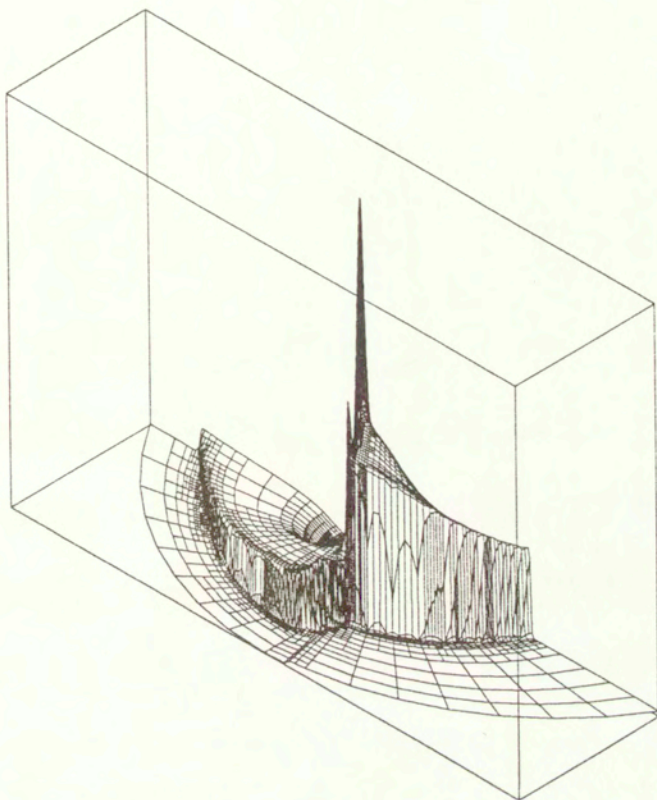


Rysunek 7.14: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Wykres gęstości wzdłuż osi symetrii obszaru obliczeniowego.

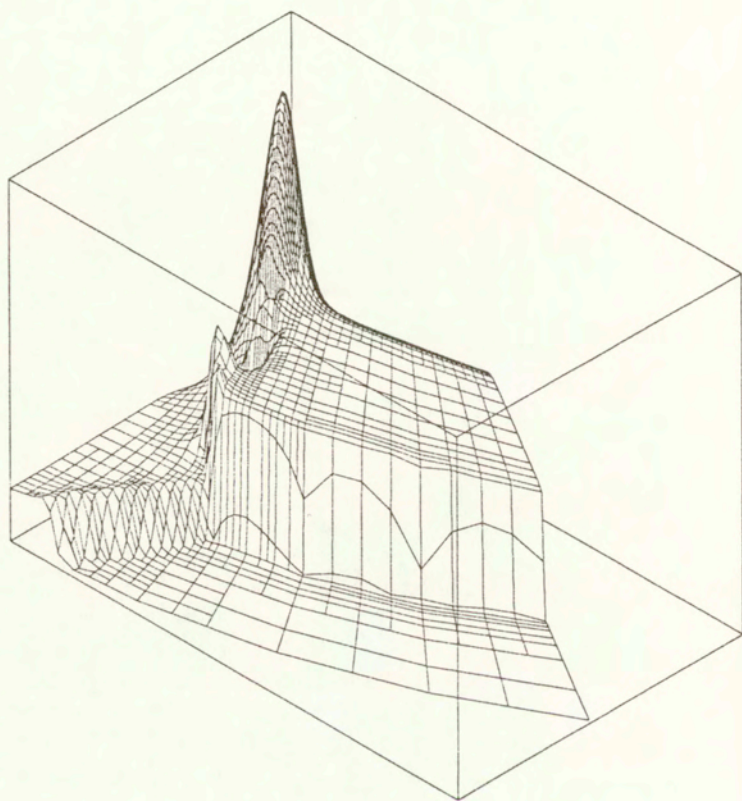




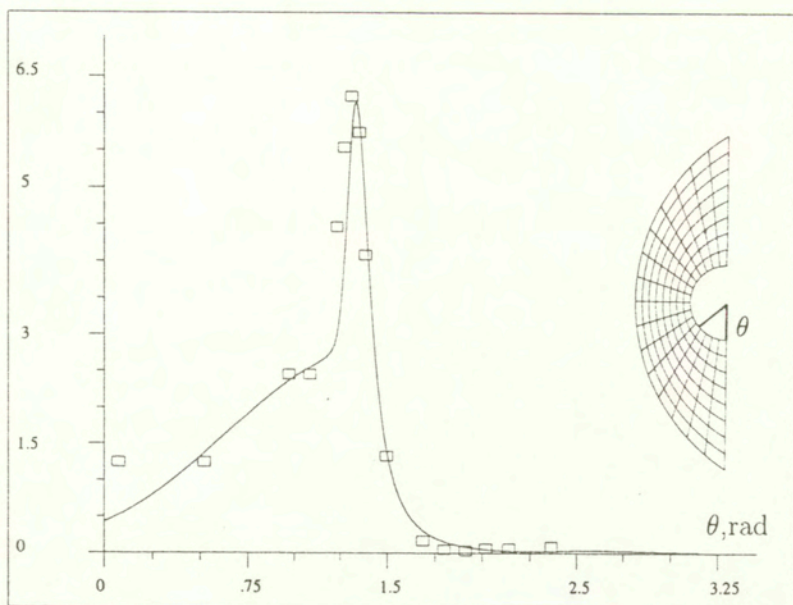
Rysunek 7.15: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Mapa rozkładu ciśnienia.



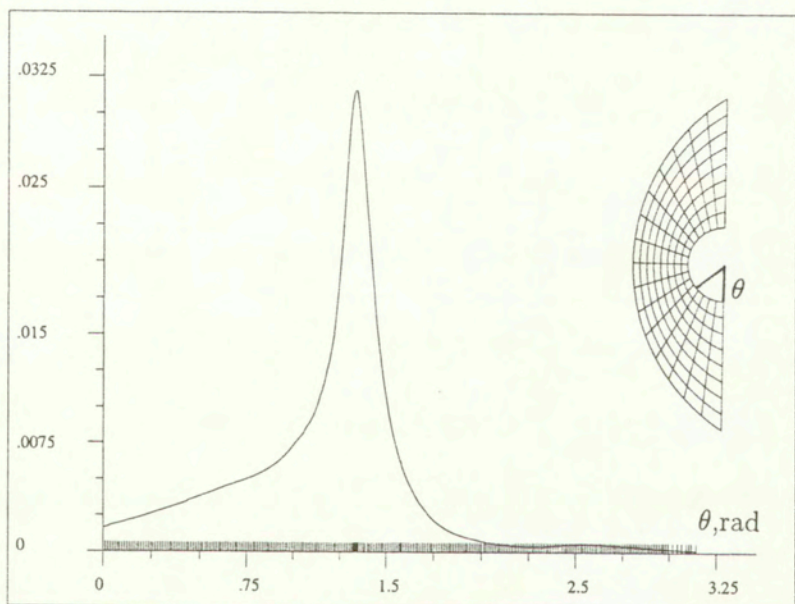
Rysunek 7.16: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Trójwymiarowy wykres ciśnienia.



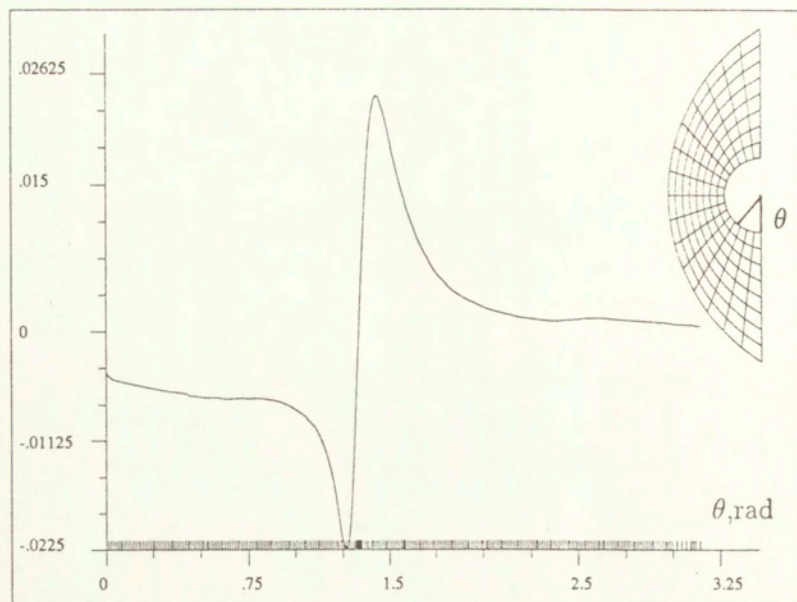
Rysunek 7.17: Przepływ wokół walca z interferencją fal uderzeniowych,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Trójwymiarowy wykres ciśnienia - powiększenie dla obszaru interakcji.



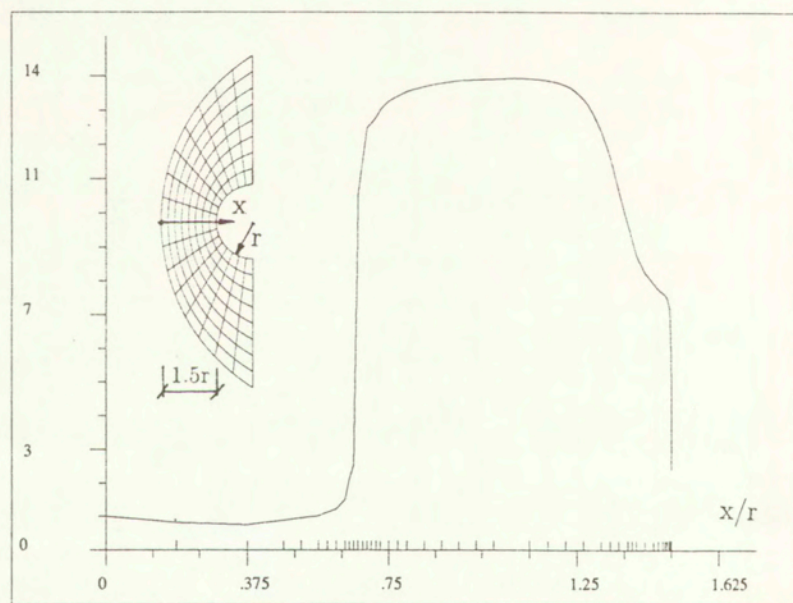
Rysunek 7.18: Wykres ciśnienia wzdłuż obwodu walca.



Rysunek 7.19: Wykres strumienia ciepła wzdłuż obwodu walca.



Rysunek 7.20: Wykres tarcia powierzchniowego wzdłuż obwodu walca.



Rysunek 7.21: Wykres temperatury wzdłuż osi symetrii obszaru obliczeniowego.

gdzie  $\rho_e, \mu_e$  są gęstością i współczynnikiem lepkości na granicy warstwy przyściennej,  $u_e$  prędkością w kierunku równoległym do warstwy,  $\theta_e, \theta_w$  temperaturami odpowiednio na granicy warstwy przyściennej i na powierzchni ściany. Wielkość  $du_e/dx$  oblicza się na podstawie wzoru

$$\frac{du_e}{dx} = \frac{1}{r} \sqrt{\frac{2(p_e - p_\infty)}{\rho_e}} \quad (7.4.7)$$

z  $r$  oznaczającym promień walca,  $p_e$  ciśnienie na krawędzi warstwy,  $p_\infty$  ciśnienie swobodnego strumienia gazu. Wielkości występujące we wzorze (7.4.6) otrzymujemy z danych do problemu oraz odczytujemy z rozwiązania numerycznego:

$$\begin{aligned} Pr &= 0.72 \text{ (liczba Prandtla)} \\ \rho_e &= 5.56 \\ \mu_e &= 0.331 \cdot 10^{-4} \\ r &= 1 \\ p_e &= 0.921 \\ p_\infty &= 1/(\gamma M^2) = 0.01107 \\ \theta_e &= 13.96 \\ \theta_w &= 2.41 \\ C_p &= 1/[(\gamma - 1)M^2] = 0.0388 \end{aligned}$$

Obliczone na tej podstawie  $q = 0.00326$  i wynik symulacji numerycznej  $q = 0.00305$  zgadzają się z dokładnością 6.5%.

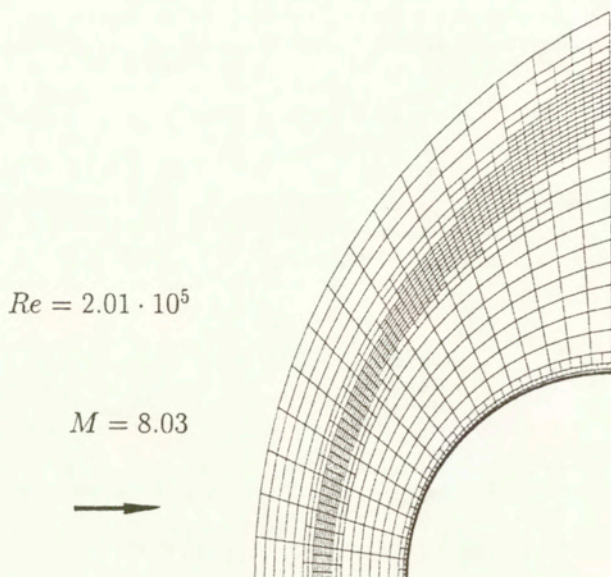
Zauważmy, że we wzorze (7.4.6) stosujemy wielkości otrzymane z symulacji numerycznej,  $p_e, \rho_e$  i  $\theta_e$ , jednakże są to wielkości „eulerowskie”, tj. niezależne od pochodnych rozwiązania. Dlatego też uznajemy je za wielkości „ścisle” w myśl hipotezy, że rozwiązanie problemu bez lepkości jest znane z dużą dokładnością.

W opisywanym zadaniu obserwujemy niekorzystny wpływ obszaru zastoju, gdzie przepływ jest bliski nieściśliwemu (oscylacje pochodnych rozwiązania widoczne w wykresie  $q$ ). Poszukiwanie algorytmów pracujących równie efektywnie dla przepływów nieściśliwych i ściśliwych jest jednym ze współczesnych kierunków badań.

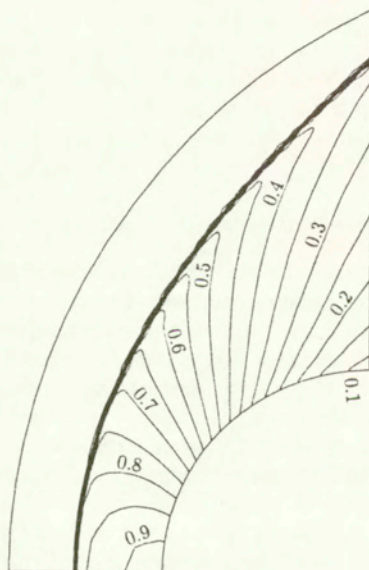
## 7.5 Optyw wokół walca, $M = 8$ , $Re = 2 \cdot 10^6$

Zadanie opływu walca rozwiązano też dla liczby Reynoldsa o rząd wyższej,  $Re = 2.01 \cdot 10^6$ . Obszar obliczeniowy i siatka będąca rezultatem adaptacji w całym obszarze była identyczna jak w zadaniu poprzednim. Jednakże położenia węzłów elementów w warstwie przyściennej zostały zmodyfikowane w taki sposób, aby odpowiadały siatce regularnej we współrzędnych biegunowych. W poprzednich zadaniach warunek ten spełniały tylko elementy siatki pierwotnej, położenia węzłów elementów wynikłych z podziałów otrzymywano przez interpolację biliniową. W obecnym zadaniu grubość warstwy przyściennej zmalała około trzykrotnie w stosunku do poprzedniego. Przypuszczamy, że niewystarczająca gładkość siatki z interpolowanymi węzłami mogła być przyczyną pewnego rodzaju „blokowania” się rozwiązania: strumień  $q$  nie wzrastał do oczekiwanych wartości mimo dokonywania podziałów elementów. Przedstawione „wygładzenie” siatki wyeliminowało to zjawisko.

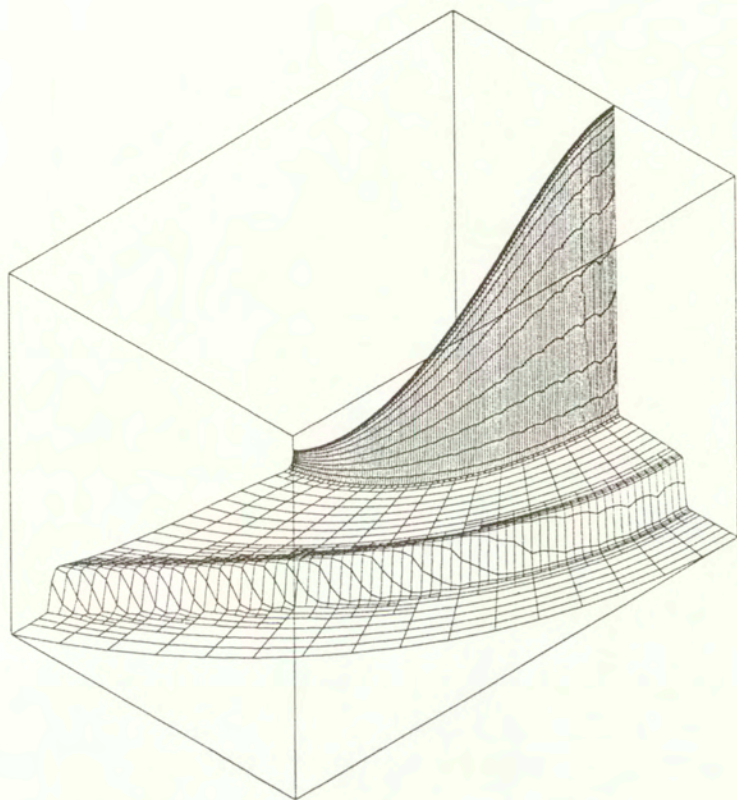
Dodatkową różnicą w stosunku do poprzedniego zadania było adaptacyjne dzielenie nie tylko elementów przyległych do ściany lecz także pozostałych, tak aby zrównać wartość



Rysunek 7.22: Przepływ wokół walca,  $M = 8$ ,  $Re = 2 \cdot 10^5$ .  
Siatka adaptacyjna i definicja zadania.

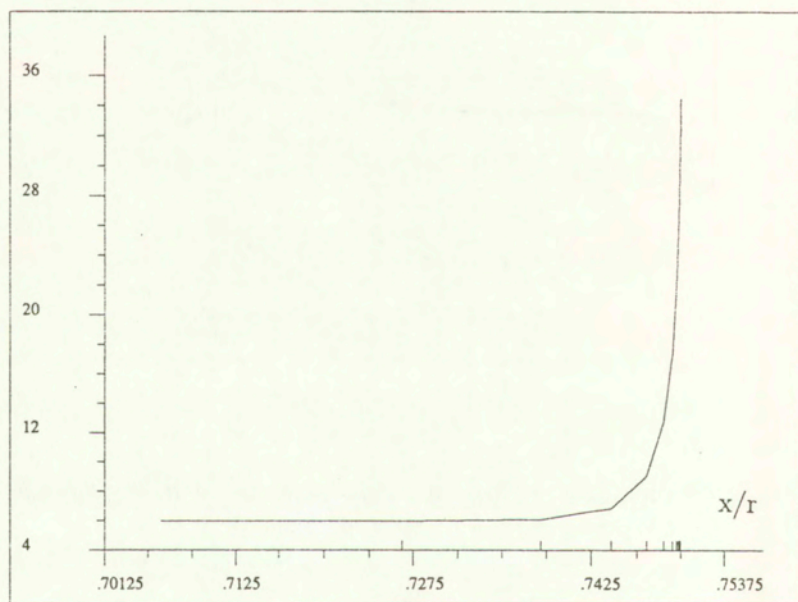
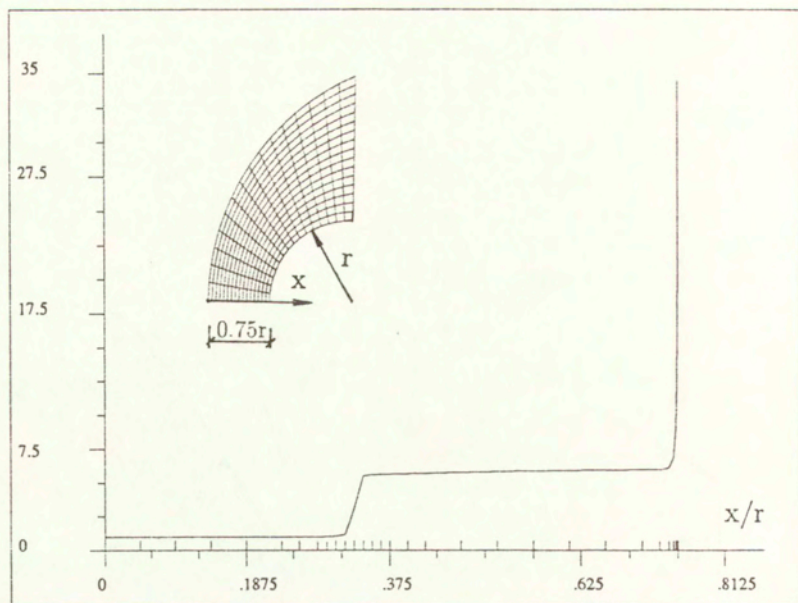


Rysunek 7.23: Przepływ wokół walca,  $M = 8$ ,  $Re = 2 \cdot 10^5$ . Mapa rozkładu ciśnienia.

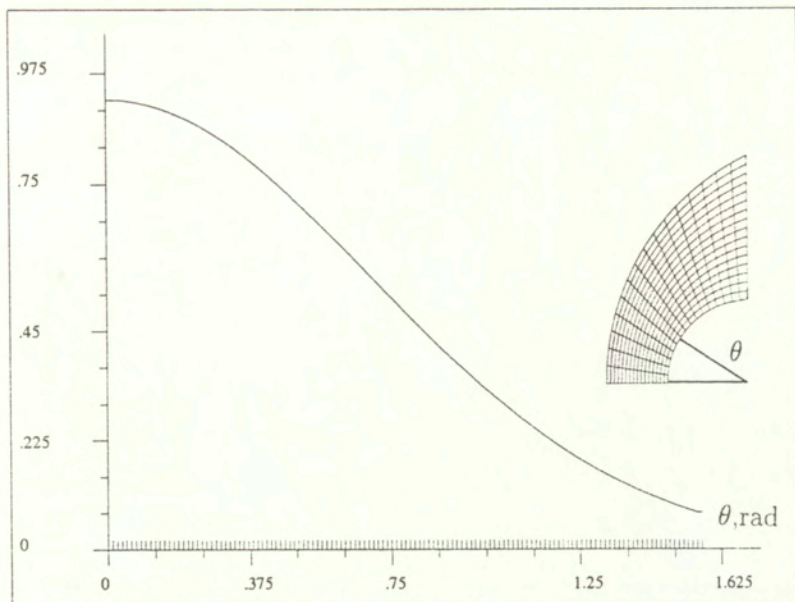


Rysunek 7.24: Przepływ wokół walca,  $M = 8$ ,  $Re = 2 \cdot 10^5$ .  
Wykres gęstości w obszarze obliczeniowym.

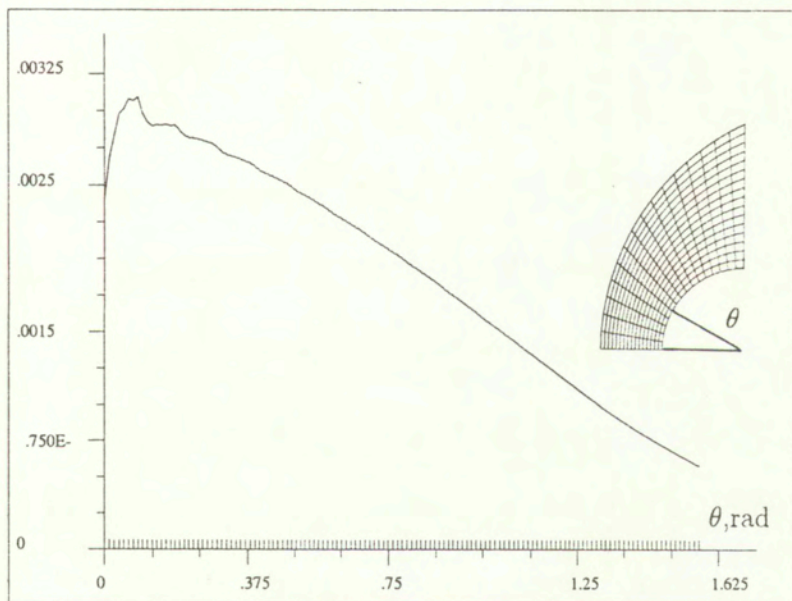




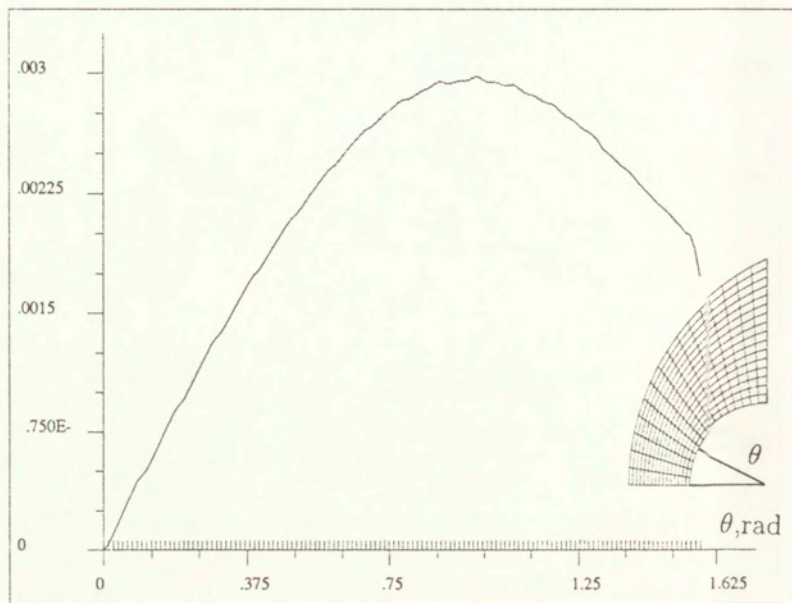
Rysunek 7.25: Przepływ wokół walca,  $M = 8$ ,  $Re = 2 \cdot 10^5$ .  
Wykres gęstości wzdłuż osi symetrii walca.



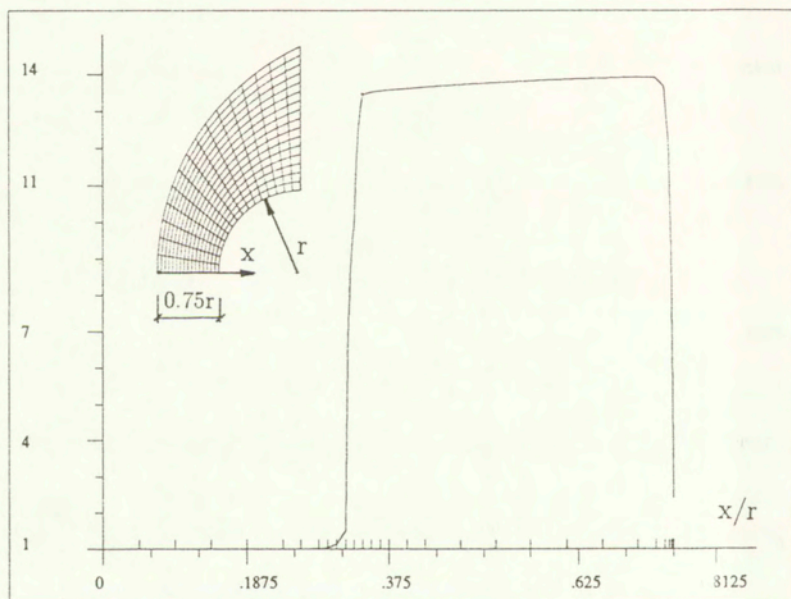
Rysunek 7.26: Wykres ciśnienia wzdłuż obwodu walca.



Rysunek 7.27: Wykres strumienia ciepła wzdłuż obwodu walca.



Rysunek 7.28: Wykres tarcia powierzchniowego wzdłuż obwodu walca.



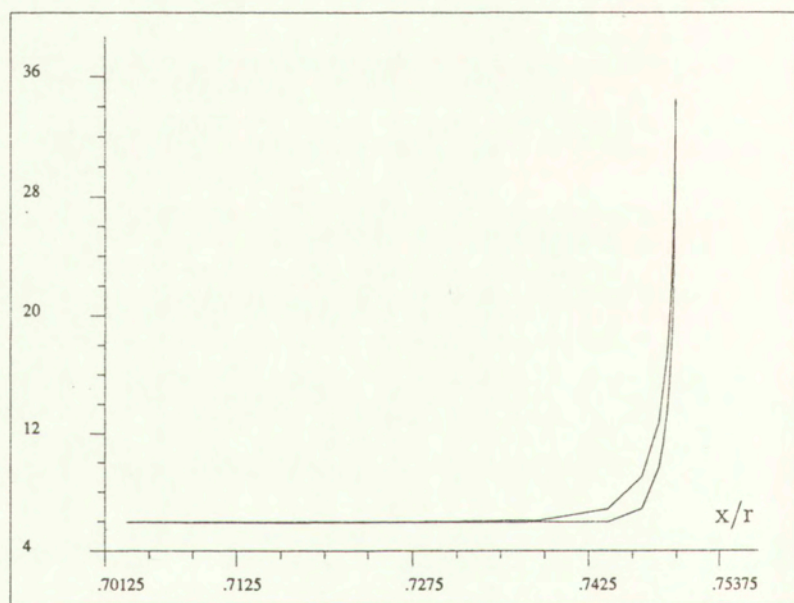
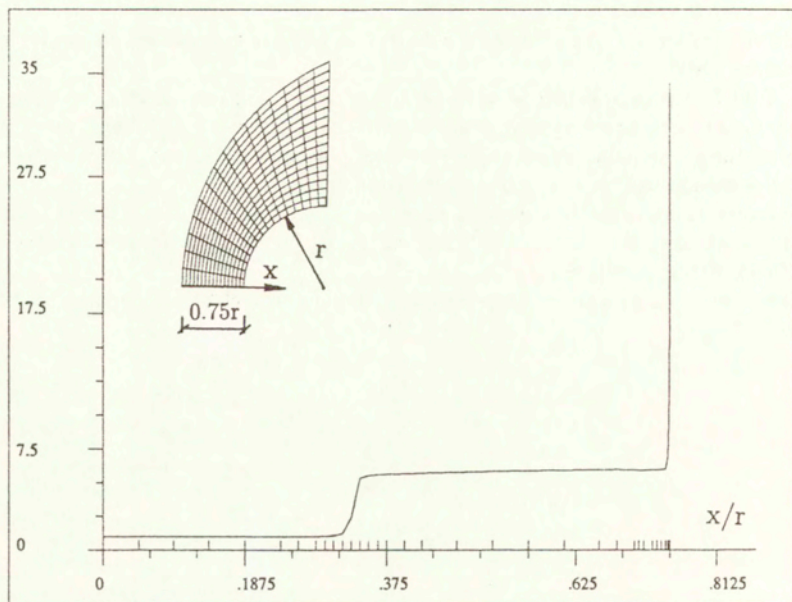
Rysunek 7.29: Wykres temperatury wzdłuż osi symetrii walca

interpolacyjnych wskaźników błędu (5.4.28). Prowadziło to faktycznie do dodatkowego podziału elementów w 3 i 4 warstwie od powierzchni walca, a tym samym do dwukrotnego zagęszczenia siatki.

Rys. 7.30-7.33 przedstawiają poszczególne charakterystyki rozwiązania omówione już poprzednio. Przekrój przez warstwę przyścienną pokazano razem z podobnym przekrojem dla poprzedniego zadania. Obserwujemy istotne pocienienie warstwy przyściennej. Na rys. 7.34 pokazano wykres wskaźnika błędu strumieni lepkich.

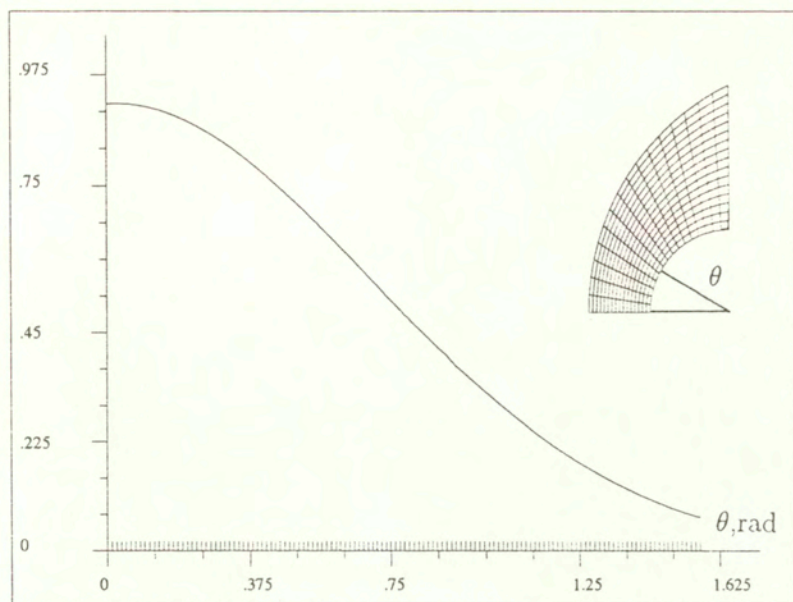
Wartość strumienia ciepła obliczona ze wzoru (7.4.6) wynosi  $q = 0.00103$  (zmianie uległa tylko wielkość  $\mu_e = 0.331 \cdot 10^{-5}$ ) zaś wartość otrzymana w symulacji numerycznej  $q = 0.00096$ . Strumienie te różnią się o ok. 7%.

Zadanie rozwiązano na siatce adaptacyjnej złożonej z 4070 elementów z podobną liczbą węzłów.

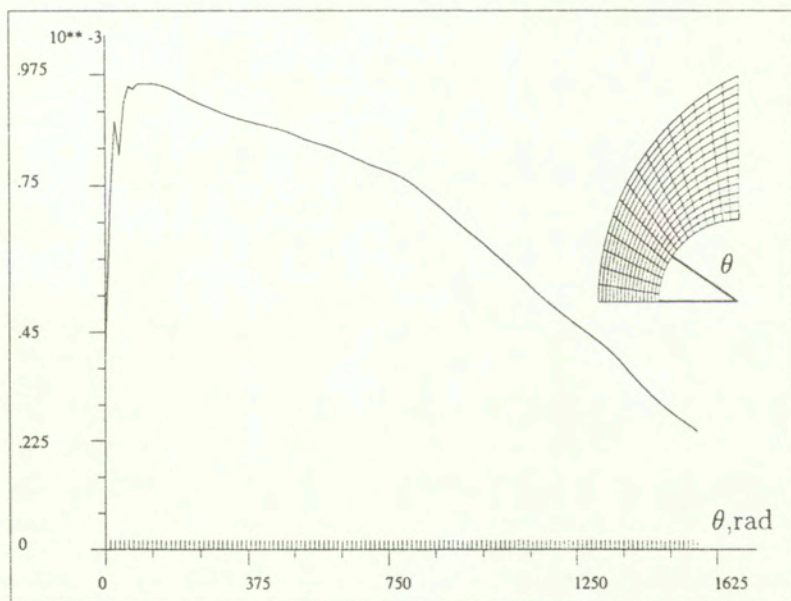


Rysunek 7.30: Przepływ wokół walca,  $M = 8$ ,  $Re = 2 \cdot 10^6$ .

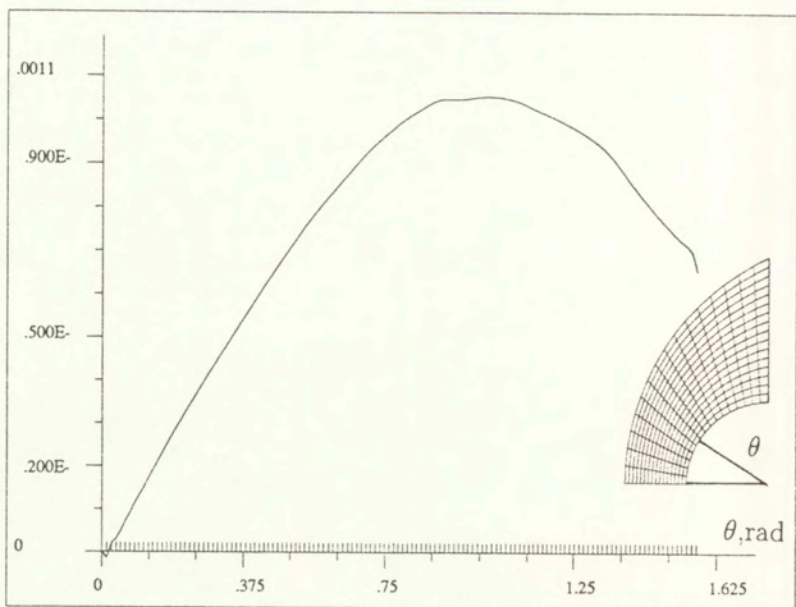
Wykres gęstości wzdłuż osi symetrii walca.



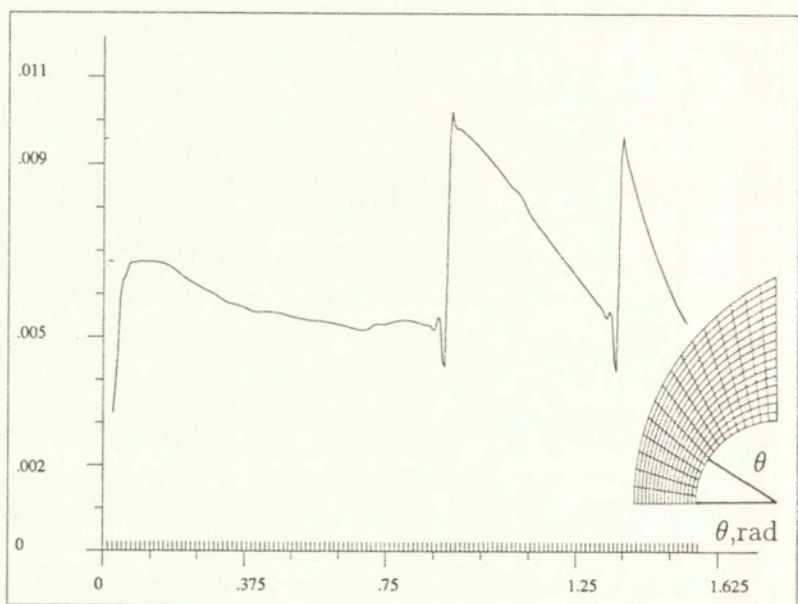
Rysunek 7.31: Wykres ciśnienia wzdłuż obwodu walca.



Rysunek 7.32: Wykres strumienia ciepła wzdłuż obwodu walca.



Rysunek 7.33: Wykres tarcia powierzchniowego wzdłuż obwodu walca.



Rysunek 7.34: Wykres wskaźnika błędów instrumentu lepkich wzdłuż obwodu walca.

## Rozdział 8

# Zakończenie

Rozwiązanie z powodzeniem zadań z poprzedniego rozdziału potwierdza efektywność i przydatność proponowanej techniki symulacji przepływów. Wydaje się, że anizotropowa adaptacyjna metoda elementów skończonych jest praktycznym i w miarę ogólnym urzeczywistnieniem intuicyjnej idei zastosowania elementów silnie wydłużonych w warstwach przyściennych. Wiele aspektów tego podejścia wymaga dalszych udoskonaleń, jednakże jego główna myśl została zrealizowana.

Najważniejsze składniki metody, w opracowaniu których osiągnięto w dużym stopniu sukces, to strategia adaptacji siatki, oszacowanie błędu rozwiązania, w tym zgrubne oszacowanie błędu strumieni lepkich, oraz efektywne procedury iteracyjne rozwiązywania równań. Dla tych algorytmów przedstawiono też teorię matematyczną, choć w części niepełną. Znaczenie zadowalającego opracowania wymienionych zagadnień możemy przybliżyć wspominając tu pominięte w pracy opisy niektórych porażek obliczeniowych, jak np. oszacowania błędu, które „nie widziały” błędu tam, gdzie należało go oczekiwać, strategię adaptacji kierunkowej wg wskaźników błędu w półnormie  $H^1$ , która „stawiała” elementy w poprzek fali uderzeniowej, czy standardowe algorytmy iteracyjne wykazujące w rozwiązywanych zadaniach rząd zbieżności 0.99. Można powiedzieć, że brak satysfakcjonującego rozwiązania któregoś ze składników metody stawiałby pod znakiem zapytania cały jej sens.

Powinniśmy wspomnieć tu także o tych aspektach metody, które istotnie utrudniały jej stosowanie i które wymagają ulepszeń i dalszego opracowania. Na pierwszym miejscu należy wymienić tu sam algorytm symulacji przepływu. Kierując się doświadczeniem obliczeniowym wybraliśmy metodę SUPG w zmiennych konserwatywnych jako algorytm o wielu korzystnych cechach na tle innych schematów. Metodę tę zaprojektowano jednak w zasadzie dla siatek quasi-równomiernych, podczas gdy stosowaliśmy ją na siatkach znacznie bardziej złożonych (z więzami, z dużą dysproporcją rozmiarów elementów, z elementami silnie wydłużonymi). W tych warunkach jakość algorytmu spadała, wymagał np. zwiększenia sztucznej lepkości czy dość sztywnego ograniczenia kroku czasowego. Celowe wydaje się zarówno przystosowanie metody do siatek adaptacyjnych jak i zastosowanie bardziej wyrafinowanego schematu całkowania w czasie. Powinien on w szczególności pozwalać na pewien rodzaj różnicowania kroku czasowego w zależności od rozmiaru elementów. Ogromne różnice w tych rozmiarach sprawiają bowiem, że przy  $\Delta t$  dopasowanym do najmniejszych elementów rozwiązanie w dużych zostaje praktycznie zamrożone i prowadzenie obliczeń dla tych obszarów w każdym z drobnych kroków czasowych jest, intuicyjnie mówiąc, zbędne. Wspomniane różnicowanie kroku czasowego mogłoby zdecydowanie przyspieszyć metodę a ponadto umożliwić rozwiązywanie zadań niestacjonarnych



(obecnie jest to niemożliwe ze względu na opisane „zamrożenie”).

Drugą z istotnych trudności, którą napotkaliśmy nieco nieoczekiwanie w praktycznych obliczeniach, była konieczność wygładzenia siatki przy brzegu krzywoliniowym (por. ostatnie zadanie). Można oczekiwać, że faktycznym rozwiązaniem tego problemu byłoby wprowadzenie odpowiednio gładkiego (z ciągłością lepszą niż  $C^0$ ) odwzorowania obszaru wzorcowego z siatką równomierną w krzywoliniowy obszar obliczeniowy, przynajmniej w pobliżu brzegu.

W końcu problemem, który wymaga głębszego zbadania, jest ustalenie kryterium optymalności siatki w warstwie przyściennej oraz związanej z tym strategii podziału elementów. Warunek redukcji i wyrównywania interpolacyjnego wskaźnika błędów strumieni lepkich, który stosowaliśmy, wyklucza wprawdzie zastosowanie zbyt dużych elementów tuż przy brzegu. Jednakże nie wiemy, czy zapewnia zbliżoną do oczekiwanej dokładność strumieni lepkich (choć w pewnym stopniu wskazują na to doświadczenia numeryczne) ani też czy prowadzi do optymalnych siatek.

Przyszłe uogólnienia metody, które byłyby ukierunkowane na poszerzenie jej praktycznych zastosowań, to wersja trójwymiarowa, uwzględnienie niestacjonarności rozwiązań oraz możliwości przemian chemicznych towarzyszących przepływom hiperdźwiękowym lub też wywołanych spalaniem.

# Dodatki

## A.1 Optymalność oszacowania błędu interpolacji

Oszacowanie błędu interpolacji  $|u - u_I|_{1,K}^2 \leq Ch^2 \int_K (u_{,xx}^2 + u_{,yy}^2) dx$  jest optymalne w znaczeniu opisywanym przez następujący

### Lemat A.1.3

Istnieje takie  $c > 0$ , że dla każdego  $u \in C^3(\bar{\Omega})$  i dla każdego punktu  $\mathbf{x} \in \Omega$ , w którym  $u_{,xx}^2 + u_{,yy}^2 > 0$ , istnieje  $h_0 > 0$ , takie że

$$h^2 \int_K (u_{,xx}^2 + u_{,yy}^2) dx \leq c \int_K |\nabla(u - u_I)|^2 dx \quad \forall h < h_0. \quad (\text{A.1.1})$$

gdzie  $u_I$  jest biliniowym interpolantem  $u$  na kwadratowym elemencie  $K = [x - h/2, x + h/2] \times [y - h/2, y + h/2]$ .

### Dowód.

Ponieważ  $u \in C^3(\bar{\Omega})$ , z rozwinięcia  $u$  w szereg Taylora wynika, że istnieje wielomian stopnia drugiego  $u^* \in P_2$  spełniający oszacowania

$$|u - u^*| \leq Ch^3, \quad |u_{,i} - u_{,i}^*| \leq Ch^2, \quad |u_{,ii} - u_{,ii}^*| \leq Ch, \quad i = 1, 2 \quad (\text{A.1.2})$$

Poza tym, jeśli  $u_I^*$  jest biliniowym interpolantem  $u^*$  na  $K$ , to

$$|\nabla(u_I - u_I^*)| \leq Ch^2 \quad (\text{A.1.3})$$

Oszacowanie to wynika z definicji  $u_I$  i  $u_I^*$

$$u_I = \sum_{i=1}^4 u(\mathbf{x}_i) \psi_i(\mathbf{x}), \quad u_I^* = \sum_{i=1}^4 u^*(\mathbf{x}_i) \psi_i(\mathbf{x}) \quad (\text{A.1.4})$$

(gdzie  $\psi_i$  są funkcjami kształtu,  $\mathbf{x}_i$  węzłami) oraz z następującego rachunku

$$\begin{aligned} |\nabla(u_I - u_I^*)| &= \frac{1}{h} |\nabla_\xi(\hat{u}_I - \hat{u}_I^*)| = \frac{1}{h} \left| \sum_i (\hat{u}(\xi_i) - \hat{u}^*(\xi_i)) \nabla \hat{\psi}_i(\xi) \right| \\ &\leq \frac{1}{h} \max_i |\hat{u}(\xi_i) - \hat{u}^*(\xi_i)| \sum_i |\nabla_\xi \hat{\psi}_i| \leq Ch^2 \end{aligned} \quad (\text{A.1.5})$$

W powyższym  $\hat{u}(\xi) = a(\mathbf{x}(\xi))$ , tj. oznacza funkcje argumentu  $\xi$  na elemencie wzorcowym  $\hat{K} = [0, 1]^2$ .  $\xi \in \hat{K}$ , zaś w ostatnim kroku zastosowano (A.1.2)<sub>1</sub>.

Nierówność (A.1.1) jest spełniona dla wszystkich wielomianów stopnia drugiego  $u^*$

$$h^2 \int_K [(u^*_{,xx})^2 + (u^*_{,yy})^2] dx \leq C \int_K |\nabla(u^* - u_j^*)|^2 dx \quad (\text{A.1.6})$$

Wynika to stąd, że zarówno lewa jak i prawa strona tej nierówności znikają wyłącznie dla funkcji biliniowej  $u^*$ , co pozwala znaleźć stałą  $C$

$$C = \sup_{\hat{u}} \int_{\hat{K}} [(\hat{u}^*_{,\xi\xi})^2 + (\hat{u}^*_{,\eta\eta})^2] dx \Big/ \int_{\hat{K}} |\nabla_\xi(\hat{u}^* - \hat{u}_j^*)|^2 dx = \\ \sup_{\hat{u} \in M} 1 / \int_{\hat{K}} |\nabla_\xi(\hat{u}^* - \hat{u}_j^*)|^2 dx \quad (\text{A.1.7})$$

gdzie  $M = \{\hat{u} = a_0 + a_1\xi + a_2\eta + b_1\xi^2 + b_2\eta^2 \mid \int_{\hat{K}} [\hat{u}^2_{,\xi\xi} + \hat{u}^2_{,\eta\eta}] dx = 1\}$ . Zbiór  $M$  jest bowiem zwarty, zaś ułamek po prawej stronie (A.1.7) jest funkcją ciągłą (jego mianownik jest dodatni dla każdego  $\hat{u}^* \in M$ ). Czynniki  $h^2$  wynika z przejścia z elementu wzorcowego  $\hat{K}$  na rzeczywisty  $K$ .

Ostatnim krokiem rozumowania jest pokazanie, że funkcja  $u$  różni się od  $u^*$  na tyle mało, że także spełnia nierówność typu (A.1.6). W tym celu zapisujemy lewą stronę (A.1.1) następująco

$$h^2 \int_K (u^2_{,xx} + u^2_{,yy}) dx \leq 2h^2 \int_K [(u^*_{,xx})^2 + (u^*_{,yy})^2] dx + 2h^2 \int_K [(u - u^*)^2_{,xx} + (u - u^*)^2_{,yy}] dx. \quad (\text{A.1.8})$$

Dla prawej strony (A.1.6) mamy z kolei

$$\int_K |\nabla(u^* - u_j^*)|^2 dx = \int_K |\nabla[(u - u_j) + (u^* - u) + (u_j - u_j^*)]|^2 dx \leq \quad (\text{A.1.9}) \\ 3 \int_K |\nabla(u - u_j)|^2 dx + 3 \int_K |\nabla(u^* - u)|^2 dx + 3 \int_K |\nabla(u_j - u_j^*)|^2 dx$$

Zastosowanie oszacowań (A.1.8) i (A.1.9) w (A.1.6) prowadzi do nierówności

$$\frac{1}{2} h^2 \int_K (u^2_{,xx} + u^2_{,yy}) dx - \left\{ h^2 \int_K [(u - u^*)^2_{,xx} + (u - u^*)^2_{,yy}] dx + \quad (\text{A.1.10}) \right. \\ \left. 3 \int_K |\nabla(u^* - u)|^2 dx + 3 \int_K |\nabla(u_j - u_j^*)|^2 dx \right\} \leq 3C \int_K |\nabla(u - u_j)|^2 dx$$

Widoczne jest, że wymagane oszacowanie (A.1.1) będzie spełnione, jeśli pokażemy, że wyrażenie w nawiasie klamrowym  $\{\cdot\}$  jest mniejsze od np.  $1/4h^2 \int_K (u^2_{,xx} + u^2_{,yy}) dx$  dla dostatecznie małych  $h$ . Jest to prawda, gdyż, zgodnie z (A.1.2), kolejne wyrażenia w nawiasie klamrowym szacują się przez  $Ch^6$ , podczas gdy

$$h^2 \int_K (u^2_{,xx} + u^2_{,yy}) dx \geq Ch^4 \quad (\text{A.1.11})$$

o ile tylko  $u^2_{,xx} + u^2_{,yy} > 0$  w  $\mathbf{x}$  i wyrażenie to jest funkcją ciągłą.

Nadmieśmy jeszcze, że możliwe jest otrzymanie oszacowania ostrzejszego niż powyżej przez stosowanie nierówności  $(a+b)^2 \leq (1+\varepsilon)a^2 + (1+1/\varepsilon)b^2$  z odpowiednio dobraćym  $\varepsilon > 0$  zamiast nierówności  $(a+b)^2 \leq 2a^2 + 2b^2$ , z której korzystaliśmy dla uproszczenia.

## A.2 Macierze jacobianowe równań Naviera–Stokesa

Jakobiany eulerowskie:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{\gamma-1}{2}q^2 - u^2 & (3-\gamma)u & -(\gamma-1)v & \gamma-1 \\ -uv & v & u & 0 \\ u(\frac{\gamma-2}{2}q^2 - \frac{c^2}{\gamma-1}) & \frac{c^2}{\gamma-1} + \frac{q^2}{2} - (\gamma-1)u^2 & -(\gamma-1)uv & \gamma u \end{bmatrix} \quad (\text{A.2.12})$$

$$A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -uv & v & u & 0 \\ \frac{\gamma-1}{2}q^2 - v^2 & -(\gamma-1)u & (3-\gamma)v & \gamma-1 \\ v(\frac{\gamma-2}{2}q^2 - \frac{c^2}{\gamma-1}) & -(\gamma-1)uv & \frac{c^2}{\gamma-1} + \frac{q^2}{2} - (\gamma-1)v^2 & \gamma v \end{bmatrix} \quad (\text{A.2.13})$$

Macierze lepkości:

$$K_{11} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\frac{4}{3}\mu \frac{m_1}{\rho^2} & \frac{4}{3}\mu \frac{1}{\rho} & 0 & 0 \\ -\mu \frac{m_2}{\rho^2} & 0 & \frac{\mu}{\rho} & 0 \\ -\frac{4}{3}\mu \frac{m_1^2}{\rho^3} - \mu \frac{m_2^2}{\rho^3} + \frac{\kappa}{C_v} \left( -\frac{e}{\rho^2} + \frac{m_1^2 - m_2^2}{\rho^3} \right) & \left( \frac{4}{3}\mu - \frac{\kappa}{C_v} \right) \frac{m_1}{\rho^2} & \left( \mu - \frac{\kappa}{C_v} \right) \frac{m_2}{\rho^2} & \frac{\kappa}{C_v} \frac{1}{\rho} \end{bmatrix} \quad (\text{A.2.14})$$

$$K_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{2}{3}\mu \frac{m_2}{\rho^2} & 0 & -\frac{2}{3}\mu \frac{1}{\rho} & 0 \\ -\mu \frac{m_1}{\rho^2} & \frac{\mu}{\rho} & 0 & 0 \\ -\frac{\mu}{3} \frac{m_1 m_2}{\rho^3} & \mu \frac{m_2}{\rho^2} & -\frac{2}{3}\mu \frac{m_1}{\rho^2} & 0 \end{bmatrix} \quad (\text{A.2.15})$$

$$K_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\mu \frac{m_2}{\rho^2} & 0 & \frac{\mu}{\rho} & 0 \\ \frac{2}{3}\mu \frac{m_1}{\rho^2} & -\frac{2}{3}\mu \frac{1}{\rho} & 0 & 0 \\ -\frac{\mu}{3} \frac{m_1 m_2}{\rho^3} & -\frac{2}{3}\mu \frac{m_2}{\rho^2} & \mu \frac{m_1}{\rho^2} & 0 \end{bmatrix} \quad (\text{A.2.16})$$

$$K_{22} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -\mu \frac{m_1}{\rho^2} & \frac{\mu}{\rho} & 0 & 0 \\ -\frac{4}{3}\mu \frac{m_2}{\rho^2} & 0 & \frac{4}{3}\mu \frac{1}{\rho} & 0 \\ -\frac{4}{3}\mu \frac{m_1^2}{\rho^3} - \mu \frac{m_2^2}{\rho^3} + \frac{\kappa}{C_v} \left( -\frac{e}{\rho^2} + \frac{m_1^2 - m_2^2}{\rho^3} \right) & \left( \mu - \frac{\kappa}{C_v} \right) \frac{m_1}{\rho^2} & \left( \frac{4}{3}\mu - \frac{\kappa}{C_v} \right) \frac{m_2}{\rho^2} & \frac{\kappa}{C_v} \frac{1}{\rho} \end{bmatrix} \quad (\text{A.2.17})$$

Symetryzator równań Naviera–Stokesa

$$A_0 = \begin{bmatrix} \frac{1}{\rho}(\psi^2 + \gamma) & \psi \frac{u}{\iota} & \psi \frac{v}{\iota} & \frac{1}{\iota}(\psi - 1) \\ \psi \frac{u}{\iota} & \frac{1}{\iota} \left( 1 + \rho \frac{u^2}{\iota} \right) & \rho \frac{uv}{\iota^2} & -\rho \frac{v}{\iota^2} \\ \psi \frac{v}{\iota} & \rho \frac{uv}{\iota^2} & \frac{1}{\iota} \left( 1 + \rho \frac{v^2}{\iota} \right) & -\rho \frac{u}{\iota^2} \\ \frac{1}{\iota}(\psi - 1) & -\rho \frac{u}{\iota^2} & -\rho \frac{v}{\iota^2} & \frac{\rho}{\iota^2} \end{bmatrix} \quad (\text{A.2.18})$$

gdzie  $\psi = \rho(u^2 + v^2)/(2\iota)$ ,  $u, v$  są składowymi prędkościami a  $\iota = e - (m_1^2 + m_2^2)/(2\rho)$  jest energią wewnętrzną na jednostkę objętości.

## A.3 Warunki brzegowe dla aproksymacji SUPG równań Naviera-Stokesa

### A.3.1 Uwagi ogólne

Postać warunków brzegowych dla sformułowania słabego SUPG (3.1.4) jest związana z równaniem różniczkowym równoważnym temu sformułowaniu

$$-\Delta t \left[ (\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}) \mathbf{U}_j^{n+1} \right]_{,i} + \mathbf{U}^{n+1} = \mathbf{U}^n - \Delta t \mathbf{A}_i \mathbf{U}_i^n + \mathbf{f}(\mathbf{U}^n), \quad (\text{A.3.19})$$

gdzie przez  $\mathbf{f}(\mathbf{U}^n)$  rozumiemy siły uogólnione odpowiadające pracy wirtualnej wyrazu  $\Delta t \sum_K \int_K \tau_a \mathbf{W}_k^T \mathbf{A}_k (\mathbf{K}_{ij} \mathbf{U}_j^n)_{,i} dx$ . Równanie to otrzymujemy całkując przez części wyrazy (3.1.4) zawierające pochodne  $\mathbf{W}_k$  i żądając spełnienia otrzymanej równości dla wszystkich  $\mathbf{W} = 0$  na  $\partial\Omega$ . Wracając z powrotem do sformułowania słabego, poprzez pomnożenie (A.3.19) przez funkcję testową  $\mathbf{W}$  i scałkowanie przez części wyrazu z drugimi pochodnymi, otrzymujemy oczywiście znów (3.1.4) wraz z jawną postacią kontrybucji brzegowej, dla której należy ustalić odpowiednie ograniczenia. Wspomniana kontrybucja, jeśli ją przenieść na prawą stronę sformułowania słabego, ma postać

$$\Delta t \int_{\partial\Omega} \mathbf{W}^T n_i (\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}) \mathbf{U}_j^{n+1} dS \quad (\text{A.3.20})$$

co odpowiada dotychczas nieokreślonego wyrazowi  $\int \mathbf{W}^T \mathbf{H} dS$  w (3.1.4). Warunki brzegowe należy sformułować bądź przez zadanie wartości wyrazu (A.3.20), bądź przez przyjęcie wartości  $\mathbf{U}^{n+1}$  na części brzegu i ograniczenie się do  $\mathbf{W} = 0$  na tej części  $\partial\Omega$ .

Poniżej przedstawiamy sformułowanie warunków brzegowych spełniające te zasady, które zaproponowano w pracach [30] i [107]. Prezentowane warunki brzegowe zostały skonstruowane na podstawie przesłanek fizycznych oraz wymagania, aby ich postać umożliwiała zachowanie liniowej stabilności problemu.

### A.3.2 Warunki na brzegu otwartym

Jest to wspomniany w rozdziale 3.3 przypadek sztucznego brzegu powstałego przez wycięcie (dla celów obliczeniowych) z całej przestrzeni ograniczonego obszaru obliczeniowego. O postaci warunków brzegowych decydują wartości własne macierzy  $\mathbf{A}_n = n_i \mathbf{A}_i$  odpowiadającej drugiemu składnikowi (A.3.20)

$$\lambda_1 = u_n - c, \quad \lambda_{2,3} = u_n, \quad \lambda_4 = u_n + c \quad (\text{A.3.21})$$

gdzie  $u_n = \mathbf{u} \cdot \mathbf{n}$  jest składową prędkości prostopadłą do brzegu,  $\mathbf{n}$  wektorem normalnej zewnętrznej. Wektor  $\Delta t \tau_a \mathbf{A}_j \mathbf{U}_j^{n+1}$  dla drugiego składnika (A.3.20) definiuje się przez następującą ogólną zależność

$$\mathbf{U}^{n+1} \cdot \mathbf{b}_i + \varepsilon^{(i)} \Delta t \tau_a \mathbf{A}_j \mathbf{U}_j^{n+1} \cdot \mathbf{b}_i = \mathbf{u}_0^{(i)} \cdot \mathbf{b}_i, \quad i = 1, \dots, 4, \quad (\text{A.3.22})$$

gdzie  $\mathbf{b}_i$  są prawymi wektorami własnymi macierzy  $\mathbf{A}_n$ . Postać  $\varepsilon^{(i)}$  oraz  $\mathbf{u}_0^{(i)}$  zależy od znaków  $\lambda_i$  i jest następująca

1. Jeśli  $\lambda_i < 0$ ,  $i = 1, \dots, 4$  (przypadek *naddźwiękowego wpływu*) to:

$\varepsilon^{(i)}$  jest małą liczbą (parametrem kary), np  $\varepsilon^{(i)} = 10^{-7}$  przy precyzji obliczeń  $10^{-14}$ ,

$\mathbf{u}_0^{(i)} = \{\mathbf{U}_\infty\}_i$ ,  $i = 1, \dots, 4$ , gdzie  $\mathbf{U}_\infty$  jest wektorem zmiennych konserwatywnych gazu w strumieniu swobodnym.

2. Jeśli przynajmniej jedna wartość własna jest dodatnia, wówczas wyróżnia się następujące przypadki:

$u_n < 0$  - poddźwiękowy wpływ,

$u_n > 0$ ,  $\lambda_1 < 0$ , poddźwiękowy wpływ,

$\lambda_i > 0$ ,  $i = 1, \dots, 4$ , naddźwiękowy wpływ.

Postać  $\varepsilon^{(i)}$ ,  $\mathbf{u}_0^{(i)}$  jest następująca:

jeśli  $\lambda_i > 0$  to  $\varepsilon^{(i)} = 1/\tau_a$ ,  $\mathbf{u}_0^{(i)} = \{\mathbf{U}^n\}_i$ ,

jeśli  $\lambda_i < 0$  to  $\varepsilon^{(i)} = |\lambda_i|/(2\tau_a)$ ,  $\mathbf{u}_0^{(i)} = \{\mathbf{U}^n\}_i$ ,

Możemy zauważyć, że dla naddźwiękowego wpływu realizowany jest (za pomocą funkcji kary) warunek brzegowy typu Dirichleta  $\mathbf{U}^{n-1} = \mathbf{U}_\infty$ , wobec czego sposób określenia pozostałych wyrazów (A.3.20) jest nieistotny. W pozostałych przypadkach strumienie związane z naturalną i sztuczną lepkością definiuje się jako wartości równe 0 lub równe tym strumieniom w poprzednim kroku czasowym. Jest to praktycznie mało istotne, choć drugie rozwiązanie sugeruje się jako lepsze w pobliżu brzegu ciała stałego.

### A.3.3 Warunki brzegowe na ścianie ciała stałego

Rozróżniamy tu dwa przypadki: ścianę adiabatyczną i ścianę o zadanej temperaturze. W obydwu prędkość przepływu  $\mathbf{u}$  na brzegu znika.

#### Ściana adiabatyczna

Przyrównujemy wyraz (A.3.20) do wyrażenia następującej postaci

$$\frac{1}{\varepsilon} \int_{\partial\Omega} \mathbf{W}^T \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{U}^{n+1} dS \quad (\text{A.3.23})$$

gdzie  $\varepsilon$  jest parametrem kary. Powoduje to wymuszenie warunku kinematycznego  $\mathbf{m}^{n-1} = 0$  oraz, w sposób słaby, jednorodnego naturalnego warunku brzegowego na czwartą składową strumienia. Jeśli założycie zaniedbywalną wartość sztucznej lepkości, oznacza to, że  $(\mathbf{F}_i^v)_4 = \sum \tau_{ij} u_j + q_i = q_i = 0$

#### Ściana o ustalonej temperaturze

W tym przypadku prócz wymuszenia znikania prędkości chcemy narzucić zadaną temperaturę  $\theta_{wall}$  na powierzchni ściany. Nie można zrobić tego jak dla prędkości, gdzie znikanie  $\mathbf{u}$  jest równoważne zerowaniu się  $\mathbf{m}$  - zmiennej podlegającej bezpośrednio aproksymacji. Mamy bowiem ustalić wielkość  $e^{n+1}/\varrho^{n+1} = \theta_{wall} C_v$ , nieliniową funkcję dwu zmiennych konserwatywnych. Proponowana metoda polega na zapisaniu zlinearyzowanej wokół  $e^n$ ,  $\varrho^n$  postaci tego warunku

$$\theta_{wall} C_v = \frac{e^n}{\varrho^n} + \frac{1}{\varrho^n} (e^{n+1} - e^n) - \frac{e^n}{(\varrho^n)^2} (\varrho^{n+1} - \varrho^n). \quad (\text{A.3.24})$$

Jest to równoważne zależności

$$e^{n+1} - \frac{e^n}{\varrho^n} (\varrho^{n+1} - \varrho^n) = \theta_{wall} C_v e^n \quad (\text{A.3.25})$$

Jeśli w rozwiązaniu początkowym zapewnimy, że  $e^0/\varrho^0 = \theta_{wall}C_v$ , to spełnianie (A.3.25) oznaczać będzie, że także  $e^n/\varrho^n = \theta_{wall}C_v$ ,  $n = 1, 2, \dots$ . Warunek (A.3.25) oraz znikanie prędkości możemy wymusić funkcją kary poprzez przyrównanie wyrazu (A.3.20) do następującego wyrażenia

$$\frac{1}{\varepsilon} \int_{\partial\Omega} \mathbf{W}^T \left( \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -e^n/\varrho^n & 0 & 0 & 1 \end{bmatrix} \mathbf{U}^{n+1} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hat{e} \end{bmatrix} \right) dS \quad (\text{A.3.26})$$

gdzie  $\hat{e} = \varrho^n \theta_{wall} C_v - e^n$ .

### A.3.4 Warunek brzegowy na osi symetrii przepływu

Warunek ten ma zastosowanie, gdy ze względu na symetrię problemu możemy ograniczyć się do rozpatrywania przepływu w jednej z dwu symetrycznych części obszaru obliczeniowego. Gdyby osią symetrii była prosta OX układu współrzędnych, to mielibyśmy:  $\varrho(y) = \varrho(-y)$ ,  $m_1(y) = m_1(-y)$ ,  $m_2(y) = -m_2(-y)$ ,  $e(y) = e(-y)$ . Związki te implikują znikanie pochodnych  $\varrho$ ,  $m_2$  i  $e$  po  $y$  oraz warunek  $m_2(0) = 0$ , co oznacza też, że  $\tau_{12} = 0$  i  $q_2 = 0$ . W ogólności, jeśli  $\mathbf{n}$  oznacza wektor normalny do brzegu, to wspomniane warunki wymusimy przyrównując wyraz (A.3.20) do wyrażenia

$$\frac{1}{\varepsilon} \int_{\partial\Omega} \mathbf{W}^T \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & n_1 n_1 & n_1 n_2 & 0 \\ 0 & n_2 n_1 & n_2 n_2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{W}^{n+1} dS \quad (\text{A.3.27})$$

Warunek brzegowy symetrii jest podobny do ściany adiabatycznej z tą różnicą, że dopuszcza poślizg wzdłuż brzegu.

### A.3.5 Warunek ustalonego wypływu

Przez powyższy warunek rozumiemy sytuację na brzegu otwartym z wypływającym gazem, gdy przepływ ustala się w taki sposób, że  $\partial\mathbf{U}/\partial n = 0$ . Zakładając, że  $\mathbf{s}$  jest wektorem stycznym do brzegu,  $\mathbf{s} = (s_1, s_2) = (-n_2, n_1)$ , możemy przedstawić pochodne  $\mathbf{U}_{,i}^{n+1}$ ,  $i = 1, 2$ , przez  $\partial\mathbf{U}^{n+1}/\partial n$ ,  $\partial\mathbf{U}^{n+1}/\partial s$

$$\mathbf{U}_j^{n+1} = \frac{\partial\mathbf{U}^{n+1}}{\partial n} n_j + \frac{\partial\mathbf{U}^{n+1}}{\partial s} s_j \quad (\text{A.3.28})$$

Stosując takie przedstawienie  $\mathbf{U}_j^{n+1}$  w wyrażeniu (A.3.20) z  $\partial\mathbf{U}^{n+1}/\partial n = 0$ , i dodatkowo zauważając, że  $\partial\mathbf{U}^{n+1}/\partial s = \mathbf{U}_{,k}^{n+1} s_k$ , otrzymujemy wyraz (A.3.20) w postaci

$$\Delta t \int_{\partial\Omega} \mathbf{W}^T n_i (\mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}) \mathbf{U}_{,k}^{n+1} s_k s_j dS \quad (\text{A.3.29})$$

Przedstawiony warunek brzegowy jest alternatywnym sposobem opisu wypływu na brzegu otwartym.

## A.4 Koercywność formy dwuliniowej niesymetrycznego sformułowania SUPG

Przez koercywność formy dwuliniowej  $B(\cdot, \cdot)$  rozumiemy spełnianie warunku

$$\sup_{W \neq 0} \frac{|B(U, W)|}{\|W\|_{X_2}} \geq \alpha \|U\|_{X_1}, \quad \forall U \quad (\text{A.4.30})$$

przy pewnym  $\alpha > 0$  i ustalonych normach  $\|\cdot\|_{X_1}$  i  $\|\cdot\|_{X_2}$ . Koercywność wraz z ciągłością formy względem tych norm oraz dodatkowym warunkiem  $\sup_U |B(U, W)| > 0 \quad \forall W \neq 0$  są założeniami uogólnionego twierdzenia Laxa–Milgrima o istnieniu rozwiązania sformułowania słabego z formą  $B$ . Problem polega na wyborze odpowiednich norm, dla których warunki te byłyby spełnione. Pozostawiając chwilowo wybór wspomnianych norm otwarty, jako punkt wyjścia dla sprawdzenia koercywności niesymetrycznej formy dwuliniowej  $B(\cdot, \cdot)$  (3.1.6) weźmy następującą oczywistą nierówność

$$\sup_{W \neq 0} \frac{|B(U, W)|}{\|W\|_{X_2}} \geq \frac{|B(U, A_0^{-1}U)|}{\|A_0^{-1}U\|_{X_2}} \quad (\text{A.4.31})$$

gdzie  $A_0$  jest symetryzatorem równań Naviera–Stokesa zdefiniowanym w (2.2.32). Licznik wyrażenia (A.4.31) ma postać

$$B(U, A_0^{-1}U) = \int_{\Omega} U^T A_0^{-1}U dx + \Delta t \int_{\Omega} U_{,i}^T \tilde{a}_{ij} U_j dx + \Delta t \int_{\Omega} U^T A_{0,i}^{-1} A_0 \tilde{a}_{ij} U_j dx \quad (\text{A.4.32})$$

gdzie  $\tilde{a}_{ij} = A_0^{-1} K_{ij} + \tau_a A_i^T A_0^{-1} A_j + \nu_d A_0^{-1} \delta_{ij}$  jest formą symetryczną dodatnio półokreśloną, tj.  $\tilde{a}_{ij}^T = \tilde{a}_{ji}$  i  $V_i \tilde{a}_{ij} V_j \geq 0 \quad \forall V_1, V_2 \in R^4$ . Ostatni składnik w (A.4.32) możemy oszacować korzystając z nierówności Cauchy'ego–Schwarza i nierówności „ $2ab \leq a^2 + b^2$ ” następująco

$$\Delta t \left| \int_{\Omega} U^T A_{0,i}^{-1} A_0 \tilde{a}_{ij} U_j dx \right| \leq \frac{\Delta t}{2} \int_{\Omega} (A_0 A_{0,i}^{-1} U)^T \tilde{a}_{ij} (A_0 A_{0,j}^{-1} U) dx + \frac{\Delta t}{2} \int_{\Omega} U_{,i}^T \tilde{a}_{ij} U_j dx \quad (\text{A.4.33})$$

Założmy, że prawdziwe jest oszacowanie

$$\frac{\Delta t}{2} \int_{\Omega} (A_0 A_{0,i}^{-1} U)^T \tilde{a}_{ij} (A_0 A_{0,j}^{-1} U) dx \leq \frac{1}{2} \int_{\Omega} U^T A_0^{-1} U dx \quad \forall U \quad (\text{A.4.34})$$

Wówczas (A.4.32) i (A.4.33) pozwala oszacować  $B(U, A_0^{-1}U)$  od dołu następująco

$$B(U, A_0^{-1}U) \geq \frac{1}{2} \left( \int_{\Omega} U^T A_0^{-1} U dx + \Delta t \int_{\Omega} U_{,i} \tilde{a}_{ij} U_j dx \right) \quad (\text{A.4.35})$$

Przejdźmy obecnie do mianownika wyrażenia (A.4.31) i zdefiniujmy normę  $\|\cdot\|_{X_2}$

$$\|V\|_{X_2}^2 = \int_{\Omega} V^T A_{0,i} V_{,i} + \int_{\Omega} V_i \tilde{a}_{ij} V_j dx, \quad (\text{A.4.36})$$



gdzie  $\hat{a}_{ij} = K_{ij} \mathbf{A}_0 + \tau_\alpha \mathbf{A}_i \mathbf{A}_0 \mathbf{A}_j^T + \nu_d \mathbf{A}_0 \delta_{ij} = \mathbf{A}_0 \bar{a}_{ij} \mathbf{A}_0$ , jest podobnie jak  $\bar{a}_{ij}$  formą symetryczną dodatnio półokreśloną. Dla mianownika w (A.4.31) mamy więc

$$\begin{aligned} \|\mathbf{A}_0^{-1} \mathbf{U}\|_{X_2}^2 = & \quad (A.4.37) \\ & \int_{\Omega} (\mathbf{A}_0^{-1} \mathbf{U})^T \mathbf{A}_0 \mathbf{A}_0^{-1} \mathbf{U} dx + \Delta t \int_{\Omega} (\mathbf{A}_{0,i}^{-1} \mathbf{U} + \mathbf{A}_0^{-1} \mathbf{U}_{,i})^T \hat{a}_{ij} (\mathbf{A}_{0,i}^{-1} \mathbf{U} + \mathbf{A}_0^{-1} \mathbf{U}_{,i}) dx \leq \\ & \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx + \\ & 2\Delta t \int_{\Omega} (\mathbf{A}_{0,i}^{-1} \mathbf{U})^T \hat{a}_{ij} (\mathbf{A}_{0,j}^{-1} \mathbf{U}) dx + 2\Delta t \int_{\Omega} (\mathbf{A}_0^{-1} \mathbf{U}_{,i})^T \hat{a}_{ij} \mathbf{A}_0^{-1} \mathbf{U}_{,j} dx = \\ & \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx + 2\Delta t \int_{\Omega} (\mathbf{A}_0 \mathbf{A}_{0,i}^{-1} \mathbf{U})^T \bar{a}_{ij} (\mathbf{A}_0 \mathbf{A}_{0,j}^{-1} \mathbf{U}) dx + 2\Delta t \int_{\Omega} \mathbf{U}_{,i}^T \bar{a}_{ij} \mathbf{U}_{,j} dx \end{aligned}$$

gdzie zastosowaliśmy nierówność „ $(a+b, a+b) \leq 2(a, a) + 2(b, b)$ ” dla iloczynu skalarnego zdefiniowanego przez drugi składnik lewej strony (A.4.37). Korzystając z przyjętego wcześniej założenia (A.4.34) powyższe oszacowanie możemy przedstawić w następującej postaci

$$\|\mathbf{A}_0^{-1} \mathbf{U}\|_{X_2}^2 \leq 3 \left( \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx + \Delta t \int_{\Omega} \mathbf{U}_{,i}^T \bar{a}_{ij} \mathbf{U}_{,j} dx \right) \quad (A.4.38)$$

Ostatecznie, wobec (A.4.35) i powyższego, możemy napisać, że

$$\frac{B(\mathbf{U}, \mathbf{A}_0^{-1} \mathbf{U})}{\|\mathbf{A}_0^{-1} \mathbf{U}\|_{X_2}} \geq \frac{\sqrt{3}}{2} \|\mathbf{U}\|_{X_1} \quad (A.4.39)$$

o ile normę  $\|\cdot\|_{X_1}$  zdefiniujemy jako

$$\|\mathbf{V}\|_{X_1}^2 = \int_{\Omega} \mathbf{V}^T \mathbf{A}_0^{-1} \mathbf{V} dx + \Delta t \int_{\Omega} \mathbf{V}_{,i}^T \bar{a}_{ij} \mathbf{V}_{,j} dx \quad (A.4.40)$$

Nierówność (A.4.39) wraz z (A.4.31) dowodzi koercywności formy  $B$ .

Nierówność (A.4.34), która jest podstawą przeprowadzonego rozumowania, dla ustalonego  $\mathbf{A}_0 = \mathbf{A}_0(\mathbf{U}^n(\mathbf{x}))$  jest spełniona dla odpowiednio małego  $\Delta t$ , gdyż minimalna wartość własna  $\mathbf{A}_0$  jest jednostajnie ograniczona od dołu od zera.

Sprawdzenie ciągłości  $B(\cdot, \cdot)$  w normach  $\|\cdot\|_{X_1}$  i  $\|\cdot\|_{X_2}$  nie przedstawia trudności. Mamy bowiem

$$\begin{aligned} \int_{\Omega} \mathbf{W}^T \mathbf{U} dx = \int_{\Omega} \mathbf{W}^T \mathbf{A}_0 \mathbf{A}_0^{-1} \mathbf{U} dx = \int_{\Omega} (\mathbf{A}_0 \mathbf{W})^T \mathbf{A}_0^{-1} \mathbf{U} dx \leq & \quad (A.4.41) \\ & \left( \int_{\Omega} (\mathbf{A}_0 \mathbf{W})^T \mathbf{A}_0^{-1} (\mathbf{A}_0 \mathbf{W}) dx \right)^{1/2} \cdot \left( \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx \right)^{1/2} = \\ & \left( \int_{\Omega} \mathbf{W}^T \mathbf{A}_0 \mathbf{W} dx \right)^{1/2} \left( \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx \right)^{1/2} \end{aligned}$$

oraz

$$\begin{aligned} \int_{\Omega} \mathbf{W}_{,i}^T \bar{a}_{ij} \mathbf{U}_{,j} dx = \int_{\Omega} (\mathbf{A}_0 \mathbf{W}_{,i})^T \bar{a}_{ij} \mathbf{U}_{,j} dx \leq & \quad (A.4.42) \\ & \left( \int_{\Omega} (\mathbf{A}_0 \mathbf{W}_{,i})^T \bar{a}_{ij} (\mathbf{A}_0 \mathbf{W}_{,j}) dx \right)^{1/2} \cdot \left( \int_{\Omega} \mathbf{U}_{,i}^T \bar{a}_{ij} \mathbf{U}_{,j} dx \right)^{1/2} = \\ & \left( \int_{\Omega} \mathbf{W}_{,i}^T \hat{a}_{ij} \mathbf{W}_{,j} dx \right)^{1/2} \left( \int_{\Omega} \mathbf{U}_{,i}^T \bar{a}_{ij} \mathbf{U}_{,j} dx \right)^{1/2} \end{aligned}$$

(gdzie  $\mathbf{a}_{ij} = \mathbf{K}_{ij} + \tau_a \mathbf{A}_i \mathbf{A}_j + \nu_d \delta_{ij}$ ). Definicja  $B(\mathbf{U}, \mathbf{W})$  (3.1.6) oraz powyższe nierówności implikują, że

$$|B(\mathbf{U}, \mathbf{W})| \leq \left( \int_{\Omega} \mathbf{W}^T \mathbf{A}_0 \mathbf{W} dx + \Delta t \int_{\Omega} \mathbf{W}_{,i}^T \hat{\mathbf{a}}_{ij} \mathbf{W}_{,j} dx \right)^{1/2}. \quad (\text{A.4.43})$$

$$\left( \int_{\Omega} \mathbf{U}^T \mathbf{A}_0^{-1} \mathbf{U} dx + \Delta t \int_{\Omega} \mathbf{U}_{,i}^T \hat{\mathbf{a}}_{ij} \mathbf{U}_{,j} dx \right)^{1/2} = \|\mathbf{W}\|_{X_2} \|\mathbf{U}\|_{X_1} \quad (\text{A.4.44})$$

## A.5 Metody gradientów sprzężonych i GMRES

### A.5.1 Metoda gradientów sprzężonych

Rozwiązujemy układ równań  $\mathbf{Ax} = \mathbf{b}$  z macierzą  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , symetryczną dodatnio określoną. Procedura iteracyjna ma postać [45]

DANE  $\mathbf{x}_0$

$k = 0$ ;  $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$

WHILE ( $\mathbf{r}_k \neq 0$ )

$k = k + 1$

IF ( $k = 1$ )

$\mathbf{p}_1 = \mathbf{r}_0$

ELSE

$\beta_k = \mathbf{r}_{k-1}^T \mathbf{r}_{k-1} / \mathbf{r}_{k-2}^T \mathbf{r}_{k-2}$

$\mathbf{p}_k = \mathbf{r}_{k-1} + \beta_k \mathbf{p}_{k-1}$

ENDIF

$\alpha_k = \mathbf{r}_{k-1}^T \mathbf{r}_{k-1} / \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k$

$\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{p}_k$

$\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A} \mathbf{p}_k$

END

$\mathbf{x} = \mathbf{x}_k$

### A.5.2 Metoda gradientów sprzężonych z uwarunkowaniem wstępnym

Idea metody polega na zastosowaniu powyższego („standardowego”) algorytmu do zmodyfikowanego układu równań, równoważnego układowi wyjściowemu

$$\tilde{\mathbf{A}} \tilde{\mathbf{x}} = \tilde{\mathbf{b}} \quad (\text{A.5.45})$$

z  $\tilde{\mathbf{A}} = \mathbf{C}^{-1} \mathbf{A} \mathbf{C}^{-1}$ ,  $\tilde{\mathbf{x}} = \mathbf{C} \mathbf{x}$ ,  $\tilde{\mathbf{b}} = \mathbf{C}^{-1} \mathbf{b}$ , gdzie  $\mathbf{C} \in \mathbb{R}^{n \times n}$  jest macierzą symetryczną dodatnio określoną. Można pokazać, że procedurę taką da się sprowadzić do algorytmu pokazanego na rys. A.1, jeśli powrócić w otrzymanej wersji do macierzy  $\mathbf{A}$ ,  $\mathbf{x}$  i  $\mathbf{b}$  [45]. Macierz  $\mathbf{M} = \mathbf{C}^2$  w tej procedurze jest macierzą symetryczną dodatnio określoną. Widoczne jest, że algorytm stosuje wyłącznie macierz  $\mathbf{M}$ , macierz  $\mathbf{C}$  służy wyłącznie do konstrukcji metody.

Zauważmy, że układ równań (A.5.45) w starych zmiennych

$$\mathbf{A} \mathbf{x} = \mathbf{b} \quad (\text{A.5.46})$$

```

DANE  $\mathbf{x}_0$ 
 $k = 0$ ;  $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ 
WHILE ( $\mathbf{r}_k \neq 0$ )
   $\mathbf{z}_k = \mathbf{M}^{-1}\mathbf{r}_k$ 
   $k = k + 1$ 
  IF ( $k = 1$ )
     $\mathbf{p}_1 = \mathbf{z}_0$ 
  ELSE
     $\beta_k = \mathbf{r}_{k-1}^T \mathbf{z}_{k-1} / \mathbf{r}_{k-2}^T \mathbf{z}_{k-2}$ 
     $\mathbf{p}_k = \mathbf{z}_{k-1} + \beta_k \mathbf{p}_{k-1}$ 
  ENDIF
   $\alpha_k = \mathbf{r}_{k-1}^T \mathbf{z}_{k-1} / \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k$ 
   $\mathbf{x}_k = \mathbf{x}_{k-1} + \alpha_k \mathbf{p}_k$ 
   $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k \mathbf{A} \mathbf{p}_k$ 
END
 $\mathbf{x} = \mathbf{x}_k$ 

```

Rysunek A.1: Algorytm gradientów sprzężonych z uwarunkowaniem wstępnym.

jest równoważny układowi

$$\mathbf{M}^{-1} \mathbf{A} \mathbf{x} = \mathbf{M}^{-1} \mathbf{b} \quad (\text{A.5.47})$$

oraz, że spektrum macierzy  $\tilde{\mathbf{A}}$  jest identyczne jak macierzy  $\mathbf{M}^{-1} \mathbf{A}$  (gdyż  $\mathbf{C}^{-1} \mathbf{A} \mathbf{C}^{-1} = \mathbf{C}(\mathbf{M}^{-1} \mathbf{A})\mathbf{C}^{-1}$ , zaś operacja ta odpowiada zmianie bazy nie wpływającej na własności geometryczne operatora  $\mathbf{M}^{-1} \mathbf{A}$ ). Macierz  $\mathbf{M}^{-1} \mathbf{A}$  jest symetryczna względem iloczynu skalarnego  $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}$

$$(\mathbf{M}^{-1} \mathbf{A} \mathbf{x}, \mathbf{y}) = (\mathbf{M}^{-1} \mathbf{A} \mathbf{x})^T \mathbf{A} \mathbf{y} = \mathbf{x}^T \mathbf{A} (\mathbf{M}^{-1} \mathbf{A} \mathbf{y}) = (\mathbf{x}, \mathbf{M}^{-1} \mathbf{A} \mathbf{y}). \quad (\text{A.5.48})$$

Normy zdefiniowane za pomocą macierzy  $\tilde{\mathbf{A}}$  i  $\mathbf{A}$  są ze sobą związane:

$$\|\tilde{\mathbf{x}}\|_{\tilde{\mathbf{A}}}^2 = \tilde{\mathbf{x}}^T \tilde{\mathbf{A}} \tilde{\mathbf{x}} = (\mathbf{C} \mathbf{x})^T (\mathbf{C}^{-1} \mathbf{A} \mathbf{C}^{-1}) (\mathbf{C} \mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \|\mathbf{x}\|_{\mathbf{A}}^2 \quad (\text{A.5.49})$$

Wskazuje to np. że zbieżność ciągu  $\mathbf{x}_n$  w normie  $\|\cdot\|_{\mathbf{A}}$  dla metody gradientów sprzężonych z uwarunkowaniem wstępnym jest identyczna jak zbieżność ciągu  $\tilde{\mathbf{x}}_n$  w normie  $\|\cdot\|_{\tilde{\mathbf{A}}}$  dla metody standardowej zastosowanej do układu (A.5.45)

$$\frac{\|\mathbf{x} - \mathbf{x}_n\|_{\mathbf{A}}}{\|\mathbf{x} - \mathbf{x}_0\|_{\mathbf{A}}} = \frac{\|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_n\|_{\tilde{\mathbf{A}}}}{\|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_0\|_{\tilde{\mathbf{A}}}} \leq 2 \left( \frac{\sqrt{\kappa(\tilde{\mathbf{A}})} - 1}{\sqrt{\kappa(\tilde{\mathbf{A}})} + 1} \right)^n \quad (\text{A.5.50})$$

### A.5.3 Metoda GMRES

W metodzie GMRES (*Generalized Minimal Residual*) [96] macierz układu równań  $\mathbf{A}$  może być niesymetryczna. Idea metody polega na poszukiwaniu poprawki  $\mathbf{z}$  dla danego przybliżenia rozwiązania  $\mathbf{x}_0$  w  $m$ -wymiarowej przestrzeni Kryłowa

$$K_m(\mathbf{r}_0) = \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \dots, \mathbf{A}^{m-1}\mathbf{r}_0\}, \text{ przy } \mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0 \quad (\text{A.5.51})$$

która minimalizowałaby residuum dla poprawionego rozwiązania, tj. byłaby rozwiązaniem problemu

$$\min_{z \in K_m(r_0)} \|b - A(x_0 + z)\| \quad (\text{A.5.52})$$

gdzie  $\|x\|^2 = x^T x$ .

Algorytm metody jest następujący

DANE  $x_0$

$$r_0 = b - Ax_0, \quad v_1 = r_0 / \|r_0\|$$

FOR  $j = 1, \dots, m$

$$H_{ij} = (Av_j)^T v_i, \quad i = 1, \dots, j$$

$$\hat{v}_{j+1} = Av_j - \sum_{i=1}^j H_{ij} v_i$$

$$H_{j-1,j} = \|\hat{v}_{j-1}\|$$

$$v_{j+1} = \hat{v}_{j-1} / H_{j-1,j}$$

ENDFOR  $j$

$$x_m = x_0 + z, \quad z = V_m y_m$$

gdzie  $y_m \in R^m$  minimalizuje  $\|\beta e_1 - Hy_m\|^2$ ,  $\beta = \|r_0\|$ ,  $e_1 = (1, 0, \dots, 0)^T \in R^n$ ,  $V_m \in R^{n \times m}$  jest macierzą, której kolumnami są wektory  $v_i$ ,  $i = 1, \dots, m$ . Jeśli występujące w algorytmie dzielenie jest niewykonalne, gdyż  $\|\hat{v}_{j+1}\| = 0$ , to oznacza, że dla aktualnej wartości  $j$  poprawione rozwiązanie jest dokładne (tzw. przypadek „lucky crash”). Jeśli  $m = n$ , to algorytm daje rozwiązanie ściśle. Jednak, ponieważ procedura wymaga pamiętania  $m$  wektorów  $v_i$ , w praktyce  $m$  przyjmuje się znacznie mniejsze niż  $n$  a poprawione rozwiązanie  $x_m$  jest punktem wyjściowym do następnego wywołania algorytmu z  $x_0 := x_m$ . Powtarza się to wielokrotnie aż do wymaganej redukcji residuum. Tę wersję algorytmu nazywamy restartowaną procedurą GMRES( $m$ ) (*restared GMRES(m)*).

# Literatura

- [1] M. Ainsworth and J.T. Oden, A posteriori error estimates for second order elliptic systems: Part 2. An optimal order process for calculating self equilibrating fluxes, *Comput. Math. Appl.* 26 (1993) 75-87
- [2] M. Ainsworth and J.T. Oden, A procedure for a posteriori error estimation for h-p finite element methods, *Comput. Methods Appl. Mech. Engrg.* 101 (1992) 73-96
- [3] M. Ainsworth and J.T. Oden, A unified approach to a posteriori error estimation using element residual methods, *Numer. Math.* 65 (1993) 23-50
- [4] S.K. Aliabadi, S.E. Ray and T.E. Tezduyar, SUPG finite element computation of viscous compressible flows based on the conservation and entropy variables formulations, *Comput. Mechanics*, 11 (1993) 300-312
- [5] R.C. Almeida and A.C. Galeao, An adaptive Petrov-Galerkin formulation for the compressible Euler and Navier-Stokes equations, *Comput. Meth. Appl. Mech. Engrg.* 129 (1996) 157-176
- [6] J.D. Anderson Jr., *Hypersonic and High Temperature Gas Dynamics*, McGraw-Hill, 1989
- [7] D.A. Anderson, J.C. Tannehill and R.H. Fletcher, *Computational Fluid Mechanics and Heat Transfer*, McGraw-Hill, New York, 1984
- [8] J. Argyris, I.S. Doltsinis and H. Friz, Hermes space shuttle: Exploration of reentry aerodynamics, *Comput. Methods Appl. Mech. Engrg.* 73 (1990) 1-51
- [9] I. Babuška, Miller, The post-processing approach in the finite element method – Part 1: Calculation of displacements, stresses and other higher derivatives of the displacements, *Internat. J. Numer. Methods in Engrg.* 20 (1984) 1085-1109
- [10] I. Babuška and W.C. Rheinboldt, Error estimates for adaptive finite element computations, *SIAM J. Numer. Anal.* 15 (1978) 736-754
- [11] I. Babuška and W.C. Rheinboldt, A posteriori error estimates for the finite element method, *Internat. J. Numer. Methods Engrg.* 12 (1978) 1597-1615
- [12] I. Babuška and W.C. Rheinboldt, On the reliability and optimality of the finite element method, *Computers & Structures*, 10 (1979)
- [13] I. Babuška and W.C. Rheinboldt, Analysis of optimal finite element meshes in  $R^1$ , *Math. Comput.* 33 (1979) 435-463

- [14] K. Banaś and L. Demkowicz, Entropy controlled finite element simulations for compressible gas flow, *J. Comput. Phys.* 126 (1996) 181-201
- [15] K. Banaś and J. Płażek, Dynamic load balancing for the preconditioned GMRES solver in a parallel, adaptive finite element Euler code, in *Proc. of 3rd ECCOMAS Computational Fluid Dynamics Conf. 1996*, (J.-A. Desideri, G.Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 1025-1031
- [16] C.E. Baumann, M.A. Stori and S.R. Idelsohn, A Petrov-Galerkin technique for the solution of transonic and supersonic flows, *Comput. Meth. Appl. Mech. Engrg.* 95 (1992) 47-70
- [17] G.J. Le Beau, S.E. Ray, S.K. Aliabadi and T.E. Tezduyar, SUPG finite element computation of compressible flows with the entropy and conservation variables formulations, *Comput. Methods Appl. Mech. Engrg.* 104 (1993) 397-422
- [18] J.H. Bramble, J.E. Pasciak, J. Wang and J. Xu, Convergence estimates for product iterative methods with applications to domain decomposition, *Math. Comput.* 57 (1991) 1-21
- [19] J. Bramble and J. Xu, Some estimates for a weighted  $L^2$  projection, *Math. Comput.* 56 (1991) 463-476
- [20] J.H. Bramble, J.A. Nitsche and A.H. Schatz, Maximum norm interior estimates for Ritz-Galerkin methods, *Math. Comput.* 29 (1975) 677-688
- [21] M.O. Bristeau, R. Glowinski, B. Maute, T. Périaux and G. Rogé, Self adaptive finite element method for 3-D, compressible Navier-Stokes simulation in aerospace engineering, in *Proc. 11th Conf. Numerical Methods in Fluid Dynamics*, Williamsburg, PA, Springer, Berlin, 1988
- [22] X.-C. Cai and O.B. Widlund, Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems, *SIAM J. Numer. Anal.* 30 (1993) 936-952
- [23] X.-C. Cai and O.B. Widlund, Domain decomposition algorithms for indefinite elliptic problems, *SIAM J. Sci. Stat. Comput.* 13 (1992) 243-258
- [24] M.J. Castro-Diaz and H. Hecht, Error interpolation minimization and anisotropic mesh generation, in *Proceedings of the 9th International Conference on Finite Elements in Fluids, New Trends and Applications, Venezia 1995* (M. Morandi Cecchi, K. Morgan, J. Periaux, B.A. Schrefler, and O.C. Zienkiewicz, eds.), 1139-1148
- [25] M.J. Castro-Diaz, H. Bouchaki, P.L. George, F. Hecht and B. Mohammadi, Anisotropic adaptive mesh generation in two dimensions for computational fluid dynamics, in *Proceedings of 3rd ECCOMAS Computational Fluid Dynamics Conference 1996* (J.-A. Desideri, G.Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 181-186
- [26] P.G. Ciarlet, *The finite element method for elliptic problems*, North Holland, 1978
- [27] P.G. Ciarlet and P.A. Raviart, Interpolation theory over curved elements, *Comput. Methods Appl. Mech. Engrg.* 1 (1972) 217-249

- [28] P. Dechaumphai, Evaluation of an adaptive unstructured remeshing technique for integrated fluid-thermal-structural analysis, AIAA paper 90-0556, 28th Aerospace Sciences Meeting, Reno, Nevada, 8-11 January 1990
- [29] L. Demkowicz, J.T. Oden and W. Rachowicz, A new finite element method for solving compressible Navier-Stokes equations based on an operator splitting method and  $h$ - $p$  adaptivity, *Comput. Methods Appl. Mech. Engrg.* 84 (1990) 275-326
- [30] L. Demkowicz, J.T. Oden, W. Rachowicz and O. Hardy, An  $h$ - $p$  Taylor-Galerkin finite element method for compressible Euler equations, *Comput. Methods Appl. Mech. Engrg.* 88 (1991) 363-396
- [31] L. Demkowicz, J.T. Oden and T. Strouboulis, Adaptive finite element methods for flow problems with moving boundaries. Part I: Variational principles and a-posteriori estimates, *Comput. Methods Appl. Mech. Engrg.* 46 (1984) 217-251
- [32] L. Demkowicz, P. Devloo and J.T. Oden, On an  $h$ -type mesh-refinement strategy based on minimization of interpolation errors, *Comput. Methods Appl. Mech. Engrg.* 53 (1985) 67-89
- [33] L. Demkowicz, J.T. Oden, W. Rachowicz and O. Hardy, Towards a universal  $h$ - $p$  adaptive finite element strategy, Part I. Constrained approximation and data structure, *Comput. Methods Appl. Mech. Engrg.* 77 (1989) 79-112
- [34] L. Demkowicz, P. Devloo and J.T. Oden, On an  $h$ -type mesh refinement strategy based on minimization of interpolation errors, *Comput. Methods Appl. Mech. Engrg.* 53 (1985) 67-89
- [35] Ph. Devloo, J.T. Oden and P. Pattani, An  $h$ - $p$  adaptive finite element method for the numerical simulation of compressible flow, *Comput. Methods Appl. Mech. Engrg.* 70 (1988) 203-235
- [36] P. Devloo, J.T. Oden and T. Strouboulis, Implementation of an adaptive refinement technique for the SUPG algorithm, *Comput. Methods Appl. Mech. Engrg.* 61 (1987) 339-358
- [37] I. St. Doltsinis and J. Urban, An adaptive operator technique for hypersonic flow computation, in *Proc. of 3rd ECCOMAS Computational Fluid Dynamics Conf. 1996*, (J.-A. Desideri, G. Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 220-228
- [38] E.R. van Driest, The problem of aerodynamic heating, *Aeronautical Engineering Review*, October 1956, pp. 26-41
- [39] M. Dryja and O. Widlund, An additive variant of the Schwarz alternating method for the case of many subregions, Technical Report 339, Courant Institute of Mathematical Sciences, 1987
- [40] M. Dryja and O. Widlund, Additive Schwarz methods for elliptic finite element problems in three dimensions, in *5th Conf. on Domain Decomposition Methods for Partial Differential Equations*, Norfolk 1991 (T.F. Chan, D.E. Keyes, G.A. Meurant, J.S. Scroggs, and R.G. Voigt, eds.), SIAM, Philadelphia, 1992

- [41] S.C. Eisenstat, H.C. Elman and M.H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations, *SIAM J. Numer. Anal.* 20 (1983) 345-357
- [42] J.A. Fay and F.R. Riddell, Theory of stagnation point heat transfer in dissociated air, *Journal of the Aeronautical Sciences*, 25 (1958) 73-85
- [43] M. Fortin, Y. Bourgault, W.G. Habashi, J. Dompierre and M.G. Vallet, Mesh adaption for viscous compressible flows, in *Proc. of the 9th Internat. Conf. on Finite Elements in Fluids, New Trends and Applications, Venezia 1995* (M. Morandi Cecchi, K. Morgan, J. Periaux, B.A. Schrefler, and O.C. Zienkiewicz, eds.), 1151-1160
- [44] M. Fortin, M.-G. Vallet, J. Dompierre, Y. Bourgault and W.G. Habashi, Anisotropic mesh adaption: theory, validation and applications, in *Proc. of 3rd ECCOMAS Computational Fluid Dynamics Conf. 1996* (J.-A. Desideri, G.Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 174-180
- [45] G.H. Golub, C.F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore and London, 1989
- [46] P. Hansbo and C. Johnson, Adaptive streamline diffusion methods for compressible flow using conservation variables, *Comput. Methods Appl. Mech. Engrg.* 87 (1991) 267-280
- [47] O. Hassan, K. Morgan and J. Peraire, An adaptive implicit/explicit finite element scheme for compressible viscous high speed flows, *AIAA 27th Aerospace Sciences Meeting*, Reno, Nevada, 9-12 January 1989
- [48] O. Hassan, K. Morgan and J. Peraire, An implicit finite element method for high speed flows, *AIAA paper 90-0402*, *AIAA 28th Aerospace Sciences Meeting*, Reno, Nevada, 1990
- [49] R. Hentschel and E.M. Hirschel, Self adaptive flow computations on structural grids, in *Proc. of the 2nd European Computational Fluid Dynamics Conf. Stuttgart 1994* (S. Wagner, E.H. Hirschel, J. Periaux, and R. Piva, eds.), John Wiley & Sons, 242-249
- [50] M.S. Holden and J.R. Noselle, Theoretical and experimental studies of the shock wave-boundary layer interaction on compression surfaces in hypersonic flow, *CALSPAN Report ARL 70-0002*, University of Buffalo Research Center, January 1970
- [51] T.J.R. Hughes, L.P. Franca and M. Mallet, A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics, *Comput. Methods Appl. Mech. Engrg.* 54 (1986) 223-234
- [52] T.J.R. Hughes, M. Mallet and A. Mizukami, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, *Comput. Methods Appl. Mech. Engrg.* 54 (1986) 341-355
- [53] T.J.R. Hughes and M. Mallet, A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective-diffusive systems, *Comput. Methods Appl. Mech. Engrg.* 58 (1986) 305-328



- [54] T.J.R. Hughes and M. Mallet, A new finite element formulation for computational fluid dynamics: IV. A discontinuity-capturing operator for multidimensional advective-diffusive systems, *Comput. Methods Appl. Mech. Engrg.* 58 (1986) 329-336
- [55] T.J.R. Hughes, L. Franca and M. Mallet, A new finite element formulation for computational fluid dynamics: VI. Convergence analysis of the generalized SUPG formulation for linear time-dependent multidimensional advective-diffusive systems, *Comput. Methods Appl. Mech. Engrg.* 63 (1987) 97-112
- [56] T.J.R. Hughes, L.P. Franca and G.M. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least squares method for advective-diffusive equations, *Comput. Methods Appl. Mech. Engrg.* 73 (1989) 173-189
- [57] T.J.R. Hughes and T.E. Tezduyar, Finite element methods for first-order hyperbolic systems with particular emphasis on the compressible Euler equations, *Comput. Methods Appl. Mech. Engrg.* 45 (1984) 217-284
- [58] G.S. Ianneli and A.J. Baker, An efficient solution-adaptive implicit finite element CFD Navier-Stokes algorithm, AIAA paper 90-0400, AIAA 28th Aerospace Sciences Meeting, Reno, Nevada, 1990
- [59] Z. Johan, T.J.R. Hughes, K.K. Mathur and S.L. Johnsson, A data parallel finite element method for computational fluid dynamics on the Connection Machine system. *Comput. Meth. App. Mech. Engrg.* 99 (1992) 113-134
- [60] Z. Johan, K.K. Mathur, S.L. Johnsson and T.J.R. Hughes, An efficient communications strategy for finite element methods on the Connection Machine CM-5 system. *Comput. Meth. Appl. Mech. Engrg.* 113 (1994) 363-388
- [61] C. Johnson, A. Szepessy and P. Hansbo, On the convergence of shock capturing streamline diffusion finite element methods for hyperbolic conservation laws, *Math. Comp.* 54 (1990) 82-107
- [62] J.-C. Jouhand and M. Barrel, A hierarchical adaptive mesh refinement method: application to 2D flows, in *Proc. of 3rd ECCOMAS Computational Fluid Dynamics Conf. 1996* (J.-A. Desideri, G.Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 268-274
- [63] G.H. Klopfer and H.C. Yee, Viscous hypersonic shock-on-shock interaction on blunt cowl lips, AIAA paper 88-0233, AIAA 26th Aerospace Sciences Meeting, Reno, Nevada, 11-14 January 1988
- [64] P.L. Lions, On the Schwarz alternating method I, in *1st Internat. Symposium on Domain Decomposition Methods for Partial Differential Equations* (E. Glowinski, G.H. Golub, G.A. Meurant, and J. Periaux, eds.), SIAM, Philadelphia, 1988
- [65] R. Löhner, K. Morgan and O.C. Zienkiewicz, An adaptive finite element procedure for compressible high speed flows, *Comput. Methods Appl. Mech. Engrg.* 51 (1985) 441-465

- [66] R. Löhner, K. Morgan and O.C. Zienkiewicz, Adaptive grid refinement for the Euler and compressible Navier-Stokes equations, in Adaptive Methods and Error Refinement in Finite Element Computations (I. Babuška, O.C. Zienkiewicz, J.P. de S.R. Gago, and A. de Oliveira, eds.) John Wiley and Sons, Ltd., London, 1986
- [67] R. Löhner, K. Morgan and O.C. Zienkiewicz, Adaptive grid refinement for the Euler and compressible Navier-Stokes equations, Techn. Report, C/R/470/83, University of Wales, Dec. 1983
- [68] R. Löhner, Adaptive remeshing for transient problems, Comput. Methods Appl. Mech. Engrg. 75 (1989) 195-214
- [69] R. Löhner, K. Morgan and O.C. Zienkiewicz, The solution of nonlinear hyperbolic equation systems by the finite element method, Internat. J. Numer. Methods in Fluids, 4 (1984) 1043-1063
- [70] R. Löhner, K. Morgan and T. Peraire, A simple extension to multidimensional problems of the artificial viscosity due to Lapidus, Comm. Appl. Num. Meth. 1 (1985) 141-147
- [71] D.G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison Wesley, New York, 1973
- [72] D.L. Marcum and N.P. Weatherill, A procedure for efficient generation of solution adapted unstructured grids, Proc. of the 4th Internat. Conf. Numerical Grid Generation in Computational Fluid Dynamics and Related Fields 1994 (N.P. Weatherill, P.R. Eismann, J. Häuser, and J.F. Thompson, eds.), Pineridge Press Ltd, 639-650
- [73] A. Mizukami and Th.J.R. Hughes, A Petrov-Galerkin finite element method for convection-dominated flows: an accurate upwinding technique for satisfying the maximum principle, Comput. Methods Appl. Mech. Engrg. 50 (1985) 181-193
- [74] K. Morgan, J. Peraire, J. Reiro and O. Hassan, Unstructured grid methods for high speed compressible flows, in The Mathematics of Finite Elements and Applications, Highlights 1993 (J.R. Whiteman, ed.), John Wiley & sons 1994
- [75] J.A. Nitsche and A.H. Schatz, Interior estimates for Ritz-Galerkin methods, Math. Comput. 28 (1974) 937-958
- [76] S.V. Nepomnyaschikh, Domain decomposition and Schwarz methods in a subspace for the approximate solution of elliptic boundary value problems, Ph. D. thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986
- [77] J.T. Oden, T. Strouboulis and Ph. Devloo, Adaptive finite element methods for the analysis of inviscid compressible flow: Part 1. Fast refinement/unrefinement and moving mesh methods for unstructured meshes, Comput. Methods Appl. Mech. Engrg. 59 (1986) 327-362
- [78] J.T. Oden, T. Strouboulis and Ph. Devloo, Adaptive finite element methods for high-speed compressible flows, Internat. J. Numer. Methods in Engrg. 7 (1987) 1211-1228

- [79] J.T. Oden, L. Demkowicz, T. Liszka and W. Rachowicz,  $h$ - $p$  adaptive finite element methods for compressible and incompressible flows, *Computing Systems in Engineering* 1 (1990) 523-534
- [80] J.T. Oden, T. Strouboulis, P. Devloo, L.W. Spradley and J. Price, An adaptive finite element strategy for complex flow problems, *AIAA Paper* 87-0557, 1987
- [81] J.T. Oden, L. Demkowicz, T. Strouboulis and P. Devloo, Adaptive methods for problems in solid and fluid mechanics, in *Accuracy Estimates and Adaptive Refinement in Finite Element Computations* (I. Babuška, O.C. Zienkiewicz, J. Gago and E.R. de A. Oliveira, eds.), John Wiley and Sons, 1986, 249-280
- [82] J.T. Oden, L. Demkowicz, W. Rachowicz and T. Westermann, A posteriori error analysis in finite elements: the element residual method for symmetrizable problems with application to compressible Euler and Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg.* Special Issue: Reliability in Computational Mechanics 82 (1990) 183-204
- [83] J.T. Oden and H.J. Brauchli, On the calculation of consistent stress distributions in finite element applications, *Internat. J. Numer. Methods Engrg.* 3 (1971) 317-325
- [84] J.T. Oden and G.F. Carey, *Finite Elements. Mathematical Aspects.* vol IV, Prentice-Hall, Englewood Cliffs, NJ, 1983
- [85] B. Palmerio, A two-dimensional FEM adaptive moving-node method for steady Euler simulations, *Comput. Methods Appl. Mech. Engrg.* 71 (1988) 315-340
- [86] J. Peraire, M. Vahdata, K Morgan and O.C. Zienkiewicz, Adaptive remeshing for compressible flow computations, *J. Comput. Phys.* 72 (1987) 449-466
- [87] W. Rachowicz, An anisotropic  $h$ -type mesh-refinement strategy, *Comput. Methods Appl. Mech. Engrg.* 109 (1993) 169-181
- [88] W. Rachowicz, An overlapping domain decomposition preconditioner for an anisotropic  $h$ -adaptive finite element method, *Comput. Methods Appl. Mech. Engrg.* 127 (1995) 269-292
- [89] W. Rachowicz, An anisotropic  $h$ -adaptive finite element method for compressible Navier-Stokes equations, *Comput. Meth. Appl. Mech. Engrg.* to appear (1997)
- [90] W. Rachowicz and J.T. Oden, On the accuracy and convergence of conjugate flux approximations, *Numer. Methods Partial Diff. Equations*, 5 (1989) 143-156
- [91] W. Rachowicz, J.T. Oden and L. Demkowicz, Toward a universal  $h$ - $p$  adaptive finite element strategy, Part 3. Design of  $h$ - $p$  meshes, *Comput. Methods Appl. Mech. Engrg.* 77 (1989) 181-212
- [92] R. Ramakrishnan, E.A. Thornton and A.R. Wieting, An adaptive finite element procedure for compressible flows with strong viscous-inviscid interactions. *AIAA Thermodynamics, Plasmadynamics and Lasers Conference*, San Antonio, Texas, 27-29 June 1988

- [93] R. Ramakrishnan, K.S. Bey and E.A. Thornton, An adaptive quadrilateral and triangular finite element scheme for compressible flows, AIAA 26th Aerospace Sciences Meeting, Reno, Nevada, 11-14 January 1988
- [94] R. Richter and P. Leyland, Completely parallel flow solvers on unstructured meshes for complex 3D compressible flows on the CRAY-T3D, in Proc. of 3rd ECCOMAS Computational Fluid Dynamics Conf. 1996, (J.-A. Desideri, G.Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, eds.), John Wiley & Sons, 1038-1044
- [95] D.H. Rudy, J.L. Thomas, A. Kumar, P.A. Gnoffo and S.R. Chakravarthy, A validation study of four Navier-Stokes codes for high speed flows, AIAA paper 89-1838 (1989)
- [96] Y. Saad and M. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. and Stat. Comput. 7 (1986) 856-869
- [97] H.A. Schwarz, Über einige Abbildungsaufgaben, J. für die Reine und Angewandte Mathematik, 70 (1869) 105-120
- [98] F. Shakib, T.J.R. Hughes and Z. Johan, A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier-Stokes equations. Comput. Methods Appl. Mech. Engrg. 89 (1991) 141-219
- [99] T. Strouboulis, P. Devloo and J.T. Oden, A moving grid finite element algorithm for supersonic flow interaction between moving bodies, Antabes FEM in Fluids V, 1986
- [100] J. Szmelter, M.J. Marchant, A. Evans and N.P. Weatherill, Two-dimensional Navier-Stokes equations with adaptivity on structured meshes, in Proc. of the Second Workshop on Reliability in Computational Mechanics, 1991 (L. Demkowicz, J.T. Oden and I. Babuska, eds.), North-Holland, 1992
- [101] P. Le Tallec, Domain decomposition methods in computational mechanics, in Computational Mechanics Advances (J.T. Oden, ed.), North-Holland, Amsterdam 1994
- [102] T.E. Tezduyar and J. Liou, Adaptive implicit-explicit finite element algorithms for fluid mechanics problems, Comput. Methods Appl. Mech. Engrg. 78 (1990) 165-179
- [103] R.R. Thareja, R.K. Prabhu, K. Morgan, J. Peraire, J. Peiro and S. Soltani, Applications of an adaptive unstructured solution algorithm to the analysis of high speed flows, AIAA paper 90-0395 (1990)
- [104] E.A. Thornton, J.T. Oden W.W. Tworzydło and S.K. Youn, Thermo-visco-plastic analysis of hypersonic structures subjected to severe aerodynamic heating, Journal of Aircraft, 27 (1991) 826-835
- [105] E.A. Thornton, P. Dechaumphai and G. Vemaganti, A finite element approach for prediction of aerothermal loads, AIAA Paper 86-1050. AIAA/ASME 4th Fluid Mechanics, Plasma Dynamics and Lasers Conference, San Antonio, Texas, 27-29 June 1988

- [106] R. Tilch, K. Morgan and N.P. Weatherill, Transient problems with moving boundaries using adaptive unstructured meshes, in Proc. of the 2nd European Computational Fluid Dynamics Conf. Stuttgart 1994 (S. Wagner, E.H. Hirschel, J. Periaux. and R. Piva, eds.), John Wiley & Sons, 273-280
- [107] W.W. Tworzydło, J.T. Oden and E.A. Thornton, Adaptive implicit/explicit finite element method for compressible viscous flows, *Comput. Methods Appl. Mech. Engrg.* 95 (1992) 397-440
- [108] W.W. Tworzydło, C.Y. Huang and J.T. Oden, An adaptive implicit/explicit finite element methods for axisymmetric viscous turbulent flows with moving boundaries. *Comput. Methods Appl. Mech. Engrg.* 97 (1992) 245-288
- [109] O. Widlund, Some Schwarz Methods for Symmetric and Nonsymmetric Elliptic Problems, in Proc. of the 5th Internat. Symposium on Domain Decomposition Methods for Partial Differential Equations, Norfolk, May 1991, (T.Chan, D.Keyes, G.Meurant, S.Scroggs, R.Voigt, eds.), SIAM, Philadelphia. 1992
- [110] A.R. Wieting and M.S. Holden, Experimental shock-wave interference heating on a cylinder at Mach 6 and 8, *AIAA Journal*, 27 (1989) 1557-1565
- [111] G. Vemaganti and A. Wieting, Application of a finite element algorithm for high speed viscous flows using structured and unstructured meshes, AIAA paper 90-1648, AIAA 21st Fluid Dynamics, Plasma Dynamics and Lasers Conference, Seattle, 18-20 June 1990
- [112] R. Verfürth, A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations, *Math. Comput.* 62 (1994) 445-475
- [113] J. Xu and X.-C. Cai, A preconditioned GMRES method for nonsymmetric or indefinite problems, *Math. Comput.* 59 (1992) 311-319
- [114] D.M. Young, Iterative solution of large linear systems, New York, Academic Press, 1971
- [115] O.C. Zienkiewicz, R. Löhner and K. Morgan, High speed inviscid compressible flow by the finite element method, Tech. Report C/R/477/84, University of Wales, May 1986

Przedstawiono adaptacyjną metodę elementów skończonych typu  $h$  do rozwiązywania lepkich przepływów naddźwiękowych. Jako procedurę symulacji przepływu zastosowano schemat SUPG. Główną ideą metody jest zastosowanie elementów silnie wydłużonych w warstwach przyściennych tworzących się w płynie przy brzegu opływanego ciała stałego. Generowanie takich elementów odbywa się automatycznie i jest częścią ogólnej strategii adaptacji siatki. Polega ona na kierunkowym dzieleniu elementów w oparciu o rozkład residualnych wskaźników błędu, oszacowanie błędu interpolacji rozwiązania oraz oszacowanie błędu strumieni lepkich w warstwie przyściennej.

Przedstawiono iteracyjne algorytmy rozwiązywania równań, których zbieżność jest niezależna od wydłużenia elementów siatki. Algorytmy te to metoda gradientów sprzężonych i GMRES z uwarunkowaniem wstępnym w postaci blokowych wersji iteracji Jacobiego i Gaussa-Seidela. Pokazano zbieżność tych procedur.

Adaptacyjną metodę elementów skończonych zastosowano do rozwiązania kilku zadań testowych z liczbą Macha 8 i 14.1 i z liczbą Reynoldsa z zakresu  $2 \cdot 10^5 - 2 \cdot 10^6$ . Wśród nich są problemy z interakcją fali uderzeniowej i warstwy przyściennej. Otrzymane wyniki numeryczne są zgodne z danymi doświadczalnymi i z rozwiązaniami analitycznymi.

## An $h$ -Adaptive Finite Element Method for Solving Compressible Navier-Stokes Equations

### Summary

An  $h$ -adaptive finite element method for solving viscous supersonic flows is presented. The SUPG scheme is used as a flow solver. The main idea of the method consists in using significantly stretched elements in boundary layers which develop in a fluid along a boundary of a solid obstacle. Generation of such elements is automatic and it is a part of a general adaptive mesh-refinement strategy. It consists in directional subsetting of elements based on distribution of residual error indicators, an interpolation error estimate for the solution, and an error estimate for the viscous fluxes in boundary layers.

Iterative solvers with the rate of convergence independent of the aspect ratio of elements are presented. The algorithms are the preconditioned conjugate gradient and the GMRES methods with the block-Jacobi and Gauss-Seidel iterations used as preconditioners. Proofs of convergence are given.

The adaptive method was used to solve a few test problems with the Mach number 8 and 14.1 and the Reynolds number ranging from  $2 \cdot 10^5$  to  $2 \cdot 10^6$ . The problems include simulations of the shock-boundary layer interaction. The numerical results match experimental data and analytic solutions.

## Eine h-adaptive Finite-Element-Methode zur Lösung der kompressible Navier-Stokes-Gleichungen

### Zusammenfassung

Eine h-adaptive Finite-Element-Methode zur Lösung der reibungsbehafteten supersonischen Strömungen wurde in dieser Arbeit vorgestellt. Als diskrete Modell der Strömungen ist ein SUPG Schema genommen. Der Schwerpunkt der vorgestellten Methode sind erheblich ausgedehnte Elemente in Grenzschichten, die sich beim Rande des umgeflossenen Körpers bilden. Automatische Generierung solcher Elementen ist ein Teil des gesamten Netzadaptionalgorithmus. Die Elementen werden richtungsweise geteilt, je nach Werten der residualen Fehlersindikatoren, Interpolationsfehlerabschätzungen und Fehlerabschätzungen der reibungsbehafteten Strömungen in der Grenzschichten.

Iterative Verfahren zur Lösung der geltenden Gleichungssystemen, deren Konvergenz unabhängig von Ausdehnung der Elementen ist, wurden vorgeschlagen. Hierbei wurden als Löser die Methode der adjungierten Gradienten und GMRES mit verschiedenen Vorkonditionierungen, wie die Jacobi- bzw. die Gauss-Seidel-Blockiterationen, verwendet. Eine Konvergenz dieser Algorithmen wurde demonstriert.

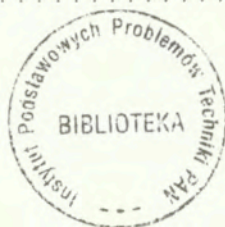
Dieses adaptive FE-Verfahren wurde zur Berechnung einiger Testproblemen mit dem Machzahl 8 oder 14.1 und mit dem Reynoldszahl von  $2 \cdot 10^5$  bis  $2 \cdot 10^6$  angewandt. Die Aufgaben mit einer Interaktion von eine Stoßwelle und eine Grenzschichte wurden auch eingeschlossen. Der Vergleich der Simulationsergebnissen mit Messungen und exakten Formeln zeigt eine gute Übereinstimmung.

# Spis rzeczy

<b>1</b>	<b>Wstęp</b>	<b>3</b>
1.1	Tematyka pracy . . . . .	3
1.2	Dotychczasowe doświadczenia z zastosowaniem adaptacji do przepływów ściśliwych . . . . .	5
1.3	Motywacja . . . . .	7
1.4	Efektywność adaptacji . . . . .	9
1.5	Uwagi o terminologii i oznaczeniach . . . . .	14
<b>2</b>	<b>Równania Naviera–Stokesa dla przepływów ściśliwych</b>	<b>15</b>
2.1	Bezwymiarowa postać równań Naviera–Stokesa . . . . .	17
2.2	Symetryzacja równań Naviera–Stokesa . . . . .	19
2.3	Fale uderzeniowe . . . . .	19
2.4	Warstwy przyścienne . . . . .	20
<b>3</b>	<b>Dyskretyzacja skończenie elementowa równań przepływów metodą SUPG</b>	<b>23</b>
3.1	Metoda SUPG . . . . .	23
3.2	Zbieżność i stabilność metody SUPG . . . . .	25
3.3	Technika rozwiązywania problemów przepływów . . . . .	30
<b>4</b>	<b>Adaptacyjna metoda elementów skończonych</b>	<b>33</b>
4.1	Metoda siatek 1–nieregularnych . . . . .	33
4.2	Residualne oszacowania błędu a posteriori . . . . .	38
4.3	Anizotropowe oszacowania błędu interpolacji w normach $L^2$ i $H^1$ . . . . .	42
<b>5</b>	<b>Strategia anizotropowej adaptacji typu <math>h</math></b>	<b>49</b>
5.1	Ogólna idea adaptacji w problemach przepływów . . . . .	49
5.2	Kryteria optymalności dla anizotropowych siatek adaptacyjnych . . . . .	51
5.3	Strategia adaptacji . . . . .	53
5.4	Adaptacja w warstwie przyściennej . . . . .	56
<b>6</b>	<b>Iteracyjne algorytmy rozwiązywania równań</b>	<b>59</b>
6.1	Wprowadzenie . . . . .	59
6.2	Metody iteracyjne . . . . .	60
6.3	Idea uwarunkowania wstępnego za pomocą iteracji prostej . . . . .	62
6.4	Metody Schwarza . . . . .	63
6.4.1	Globalny krok iteracji . . . . .	66
6.4.2	Modyfikacja operatora uwarunkowania wstępnego . . . . .	67
6.5	Oszacowanie wartości własnych macierzy z uwarunkowaniem wstępnym . . . . .	68



6.6	Oszacowanie zbieżności metody Gaussa-Seidela . . . . .	70
6.7	Generacja podobszarów . . . . .	70
6.8	Definicja $V_0^h$ . . . . .	73
6.9	Istnienie rozkładu o niskiej energii . . . . .	73
6.9.1	Siatki quasi-równomierne . . . . .	73
6.9.2	Siatki z więzami . . . . .	76
6.9.3	Obecność elementów silnie wydłużonych . . . . .	78
6.10	Uogólnienie dla formy dwuliniowej symetrycznego sformułowania zadania przepływu . . . . .	81
6.11	Uwarunkowanie wstępne dla sformułowania niesymetrycznego . . . . .	87
6.11.1	Uwagi dodatkowe . . . . .	95
<b>7</b>	<b>Przykłady numeryczne</b> . . . . .	<b>97</b>
7.1	Uwagi wstępne . . . . .	97
7.2	Przepływ wokół naroża wklęsłego, $M = 14.1$ , $Re = 72000/ft$ . . . . .	98
7.3	Przepływ wokół walca z interferencją fal uderzeniowych, $M = 8$ , $Re = 2 \cdot 10^5$ 100	
7.4	Optyw wokół walca, $M = 8$ , $Re = 2 \cdot 10^5$ . . . . .	111
7.5	Optyw wokół walca, $M = 8$ , $Re = 2 \cdot 10^6$ . . . . .	121
<b>8</b>	<b>Zakończenie</b> . . . . .	<b>131</b>
	<b>Dodatki</b> . . . . .	<b>133</b>
A.1	Optymalność oszacowania błędu interpolacji . . . . .	133
A.2	Macierze jacobianowe równań Naviera-Stokesa . . . . .	135
A.3	Warunki brzegowe dla aproksymacji SUPG równań Naviera-Stokesa . . . . .	136
A.3.1	Uwagi ogólne . . . . .	136
A.3.2	Warunki na brzegu otwartym . . . . .	136
A.3.3	Warunki brzegowe na ścianie ciała stałego . . . . .	137
A.3.4	Warunek brzegowy na osi symetrii przepływu . . . . .	138
A.3.5	Warunek ustalonego wypływu . . . . .	138
A.4	Koercywność formy dwuliniowej niesymetrycznego sformułowania SUPG . . . . .	139
A.5	Metody gradientów sprzężonych i GMRES . . . . .	141
A.5.1	Metoda gradientów sprzężonych . . . . .	141
A.5.2	Metoda gradientów sprzężonych z uwarunkowaniem wstępnym . . . . .	141
A.5.3	Metoda GMRES . . . . .	142



56552