

Den Endelige Element Metode

Carsten Jørgensen

august 1995

Forord

Dette speciale markerer afslutningen på mit matematik studium ved Matematisk Institut, Københavns Universitet. Specialet er skrevet i perioden december 1994 til august 1995 og tager sit udgangspunkt i kurset “Numerisk løsning af differentiaalligninger” udbudt i forårssemesteret 1995 af Jens Hugger.

Et legat fra ERASMUS muliggjorde et studieophold ved University of Leeds, England i studieåret 1992–93. Jeg vil gerne takke Niels Grønbæk og Kjeld Bagger Laursen for at have hjulpet mig med de praktiske forberedelser forud for opholdet. Igennem Valérie Guimet stiftede jeg her bekendtskab med den endelige element metode, som er emnet for dette speciale.

Set i bagspejlet viste matematik studiet sig at være det helt rigtige for mig. Mit studievalg skyldes ikke mindst Søren Antonius som hermed takkes.

Endelig vil jeg gerne takke min vejleder Jens Hugger for at have delt sin viden om endelige element metoder med mig og for at have besvaret utallige spørgsmål.

Noter

Afsnittet om historien bag den endelige element metode er koncist men ufuldstændigt. Dette skal forstås således, at der ikke gives nogle referencer til artikler, hvori ideerne er præsenteret første gang. Årsagen til ovenstående valg skyldes (historiske) vanskeligheder med at datere oprindelse for den endelige element metode samt et, fra min side, ønske om at undgå kontroverser. For en mere fyldestgørende redegørelse for historien bag den endelige element metode henvises til artiklerne [7], [37] og [44].

Af tidsmæssige årsager var det desværre ikke muligt at implementere en endelig element løser eller dele heraf. I stedet har jeg ved at studere programmerne FEM2D, FEMLAB-ODE [20], NFEARS [35] og PLTMG [14] fået et indblik i, hvordan endelig element software fungerer. Jeg har desuden haft adgang til dele af koden for nogle af ovennævnte programmer og har på den måde haft mulighed for at studere strukturen i sådanne programmer.

Så godt som al eksisterende litteratur om endelige element metoder er engelsk sprog. Dette giver af og til vanskeligheder med at finde passende danske udtryk. I de situationer hvor det har været specielt svært, har jeg valgt at angive det engelske udtryk i parentes efter det danske udtryk. En del af eksemplerne i dette speciale tager udgangspunkt i fysikkens verden. Min ikke overvældende fysiske indsigt gør, at jeg ligeledes her har valgt at angive det engelske fagudtryk i parentes efter det danske udtryk. På den måde skulle eventuelle unøjagtigheder kunne modvirkes.

Indhold

Forord	i
Noter	ii
1 Introduktion	1
1.1 Historiske og bibliografiske noter	1
1.2 Galerkins metode	2
1.3 Et eksempel	3
1.4 Den endelige element metode	5
2 Variationsformer	7
2.1 Problemstilling	7
2.2 Fordele ved variationsformuleringen	8
2.3 “Rimelige krav”	8
2.3.1 “Fornuftige” krav	8
2.3.2 Matematiske krav	9
2.4 Randbetingelser	11
2.4.1 Klassificering af randbetingelser	11
2.5 Konstruktionsregler	13
2.6 Eksistens og entydigheds sætninger	13
2.7 Eksistens og entydighed for det diskrete problem	16
2.8 Eksempel på konstruktion af variationsform	18
3 Endelige elementer og endelige element rum	20
3.1 Notation	20
3.2 Endelige element rum	21
3.3 Triangulering af $\bar{\Omega}$	21
3.3.1 Lokale endelige element rum	22
3.4 Polynomiumsrummene P_k	23
3.5 Lagrange endelige elementer	24
3.5.1 Triangulering af $\bar{\Omega}$ med simplekser	28

3.6	Endelige element rum hørende til n -simplexer af type (k) . . .	29
3.7	Basisfunktioner for endelige element rum	30
3.8	Hermite endelige elementer	31
3.9	Generelle endelige elementer	33
3.9.1	Typen af frihedsgrader	34
3.10	Affine familier af endelige elementer	35
3.11	Generelle endelige element rum	37
4	Polynomiumsinterpolation	39
4.1	A priori fejlligningen	39
4.2	Lokal Taylor polynomiums interpolation i et rektangel i \mathbb{R}^2 . .	40
5	Netgenerering	53
5.1	Sammenhæng mellem net tætheds funktioner og følger af net .	53
5.2	Netgenerering i to dimensioner	60
5.2.1	Konstruktion af næsten optimale net størrelses funktioner for netinddelinger med rektangulære elementer .	61
5.2.2	Bestemmelse af det næsten optimale antal af elementer for rektangulære netinddelinger	65
5.2.3	Netinddelings algoritmen	65
5.2.4	Teoretiske resultater for algoritmen	74
6	Fejlestimering	77
6.1	Generel teori om fejlestimering	77
6.1.1	Beregningsmæssige- og modelleringsfejl	77
6.1.2	Forskellige typer af fejlestimatorer	78
6.1.3	Forskellige former for fejlestimering	79
6.1.4	Definitioner og begreber	79
6.2	Zienkiewicz–Zhu fejl estimatoren	81
6.2.1	Model problem	81
6.2.2	Forskellige metoder til bestemmelse af σ^*	84
6.2.3	Teoretisk analyse af Z^2 fejl estimatoren	90

Figurer

1.1	Geometrisk fortolkning af den endelige element løsning	4
3.1	Trekant af type (1) og tetraede af type (1)	27
3.2	Trekant af type (2) og tetraede af type (2)	27
3.3	Trekant af type (3) og tetraede af type (3)	27
3.4	Eksempel på tilladt triangulering	29
3.5	Eksempel på ikke-tilladt triangulering	29
3.6	Illustration for situationen $n = 2$ og trekant af type (2)	30
3.7	Eksempel på støtte for basisfunktioner hørende til n -simplexer af type (2)	31
3.8	Argyris' trekant	33
3.9	Eksempel på affin afbildning, hvor et element afbildes på et typisk reference element	35
3.10	Hjørner i et typisk reference n -simplex	36
3.11	Typisk enheds 2-simplex	36
4.1	Grafisk illustration af u opsplittet i $T_{pq}\{u\}$ og e_{pq}^T	41
4.2	Illustration af u i tilfældet $K = 4\pi$ og $\square = [0, 1] \times [0, 1]$	43
5.1	Eksempler på β -domæner	56
5.2	Net og element notation	61
5.3	Eksempel på tilladt netinddeling	67
5.4	Datastruktur for netinddelingen i figur 5.3	67
5.5	(a) Mulige forfininger af et element og (b) den resulterende data struktur	68
5.6	Søskende sæt	69
5.7	Tilladte forgrovnings samt et eksempel på en ikke-tilladt for- grovnings	71
6.1	Eksempel på celle	80
6.2	Eksempler på element bælter omkring en samlingspunkts knu- de markeret med \bullet	86

6.3	Eksempel på kvadratiske elementer med bi-lineære basisfunktioner. Punktet mærket \bigcirc er samlingspunkts knuden. De lokale basisfunktioner er bestemt ved værdien i punkterne mærket \circ/\bullet . Det ligeledes lineære polynomium σ_p^* er defineret over hele element bæltet og bestemt ved værdien i Gauss punkterne mærket \triangle . I punktet mærket \bullet har vi udregnet en approksimeret værdi af stresset ved at indsætte passende koordinater i udtrykket for σ_p^*	88
6.4	Eksempel på kvadratiske elementer med trunkerede bi-kvadratiske basisfunktioner. Punktet mærket \bigcirc er samlingspunkts knuden. De lokale basisfunktioner er bestemt ved værdien i punkterne mærket \circ/\bullet . Det kvadratiske polynomium σ_p^* er defineret over hele element bæltet og bestemt ved værdien i Gauss punkterne mærket \triangle . I punkterne mærket \bullet har vi udregnet en approksimeret værdi af stresset ved at indsætte passende koordinater i udtrykket for σ_p^*	89
6.5	Eksempel med kvadratiske elementer og bi-kvadratiske basisfunktioner. \triangle Gauss punkt. \bullet Approksimeret værdi af stresset. \bigcirc Samlingspunkts knude.	91
6.6	Eksempel med kvadratiske elementer og bi-kvadratiske basisfunktioner. \triangle Gauss punkt. \bullet Approksimeret værdi af stresset. \bigcirc Samlingspunkts knude.	91

Tabeller

4.1	Værdierne for p_k , q_k og α_k i formen for e_{pq}^T	46
4.2	Eksempler på værdier for $C_{k,0,r,p}^T$ for normerne \mathcal{L}^1 og \mathcal{L}^2	50
4.3	Eksempler på værdier for $C_{k,0,r,p}^T \cdot g(1/2, 1/2)$ for normerne \mathcal{L}^1 og \mathcal{L}^2	50
4.4	Eksempler på værdier for $C_{m,r,p}^T$ for normerne \mathcal{L}^1 , $\mathcal{H}^1 = \mathcal{W}^{1,2}$ og $\mathcal{W}^{1,4}$	51
4.5	Eksempler på værdier for $C_{m,r,p}^T \cdot g(1/2, 1/2)$ for normerne \mathcal{L}^1 , $\mathcal{H}^1 = \mathcal{W}^{1,2}$ og $\mathcal{W}^{1,4}$	51

Algoritmer

5.1	Netinddelings algoritme byggende på net tætheds funktion . .	59
5.2	Forgrovnings delen af algoritmen	73
5.3	Forfinings delen af algoritmen	73

Kapitel 1

Introduktion

Der kan gives mange eksempler fra fx fysik og økonomi på problemer, der resulterer i et ønske om at løse differentialligninger af typen

$$\mathcal{D}u = f, \tag{1.1}$$

hvor \mathcal{D} er en given differentialoperator, muligvis med rand- eller begyndelsesværdibetingelser, f en given funktion og u den ukendte løsning til problemet. Ofte kan man ikke klare sig med analytiske metoder til løsning af ligningen (1.1), og man vil derfor anvende numeriske metoder. En af disse metoder kaldes Galerkins metode. I dette kapitel skal vi give en overordnet beskrivelse af denne, men først et par historiske og bibliografiske noter.

1.1 Historiske og bibliografiske noter

I forbindelse med numerisk løsning af partielle differentialligninger begyndte ingeniører fra flyindustrien i USA i slutningen af 50'erne at anvende en metode, der ligner den metode, vi i dag kender som den endelige element metode. Da man i midten af 60'erne begyndte at studere den matematiske teori bag metoden, blev det klart, at der var tale om en generel metode til numerisk løsning af differential- og integralligninger. Op gennem 70'erne blev den matematiske teori for især h -metoden grundlagt, og i 80'erne kom så p -metoden til. I h -metoden forfines inddelingen af domænet yderligere, såfremt en given fejltolerance ikke nås. I p -metoden fastholder man derimod netinddelingen og øger i stedet graden af de polynomier, som man forsøger at approksimere løsningen med. Desuden var og er fejlestimation en stadig kilde til forskning. I de senere år har man yderligere arbejdet med hp -metoden, dvs en kombination af h - og p -metoden. Ved udgangen af 80'erne var hovedparten af

teorien for endimensionelle problemer på plads, mens der stadigvæk er åbne spørgsmål for to- og tredimensionelle problemer.

En standard lærebog om den endelige element metode er [19]. Kapitlet om variations former er inspireret af denne bog samt af [40]. Endelige elementer og endelige element rum er ligeledes godt beskrevet i [19]. Funktionsinterpolation er gennemgået meget generelt i [19]. Vi har brug for et mere specielt tilfælde, som er bedre beskrevet i [29]. Netinddelings algoritmen vi skal studere senere er kun en enkelt algoritme blandt mange. Artiklen [8] indeholder referencer til andre algoritmer. Fejlestimering er et område med en del forskningsaktivitet. I [9, 10, 11] findes omfattende litteraturhenvisninger til artikler om dette emne.

1.2 Galerkins metode

For at løse ligningen (1.1) vælger vi en række lineært uafhængige løsningsfunktioner (eng. trial functions) ϕ_1, \dots, ϕ_m og forsøger at tilnærme u med en linearkombination af disse, dvs.

$$u = \sum_{j=1}^m \xi_j \phi_j. \quad (1.2)$$

Indsættes denne approksimation i differentialligningen (1.1) fås et residual givet ved

$$\mathcal{D}\left(\sum_{j=1}^m \xi_j \phi_j\right) - f. \quad (1.3)$$

For at få så god en approksimation til u som muligt, ønsker vi, at dette residual skal være så lille som muligt. Vi vil derfor kræve, at integralet af residualet vægtet med en række af test funktioner (eng. test functions) er 0. Ideen bag dette krav er hentet fra indre produkter på funktionsvektorrum. I \mathcal{L}^2 har vi det indre produkt $\langle f, g \rangle = \int_{\Omega} f(x)g(x) dx$ for to funktioner f og g . Som bekendt siger vi, at f og g er ortogonale, hvis $\langle f, g \rangle = 0$. Den geometriske fortolkning er her, at vi, som approksimeret løsning, vælger den løsning, der ligger "tættest" på testrummet (dvs. den løsning som er ortogonal på testrummet).

Approksimationen (1.2) indeholder m ubekendte konstanter ξ_1, \dots, ξ_m . Ved at vælge m lineært uafhængige test funktioner w_1, \dots, w_m fås et $m \times m$ lineært ligningssystem

$$\sum_{j=1}^m \xi_j \int \mathcal{D}\phi_j w_i dx = \int f w_i dx, \quad 1 \leq i \leq m, \quad (1.4)$$

hvorfra ξ_1, \dots, ξ_m kan bestemmes. Denne metode er kendt som den vægtede residual metode. Vælges ϕ_1, \dots, ϕ_m som test funktioner kaldes den ovenfor beskrevne fremgangsmåde for Galerkins metode.

1.3 Et eksempel

Som et konkret eksempel betragter vi følgende model problem: Bestem u således at

$$\begin{aligned} -u'' + u &= f, & \text{for } x \in]0, 1[, \\ u(0) &= 0, & u(1) = 0, \end{aligned} \quad (1.5)$$

hvor f er en given funktion. Vælges test og løsnings funktionerne som kontinuerte stykkevise polynomier kaldes Galerkins metode også for den endelig element metode. Vi skal i dette eksempel benytte kontinuerte stykkevis lineære funktioner som test og løsnings funktioner.

Lad $\mathcal{T}_h : 0 = x_0 < x_1 < \dots < x_{m+1} = 1$ være en inddeling af $I =]0, 1[$ i elementer $I_j =]x_{j-1}, x_j[$ med længde $h_j = x_j - x_{j-1}$, og lad $\mathcal{X}_{0,fe}$ være vektorrummet bestående af kontinuerte stykkevis lineære funktioner på \mathcal{T}_h , som er 0 i punkterne $x = 0$ og $x = 1$. Intervallerne $I_j =]x_{j-1}, x_j[$, $j = 1, \dots, m+1$ kaldes for endelige elementer. (Vi skal senere give en præcis definition.) Defineres ϕ_j , $j = 1, \dots, m$ som kontinuerte stykkevis lineære funktioner på \mathcal{T}_h opfyldende

$$\phi_j(x_i) = \begin{cases} 1 & \text{hvis } i = j, \\ 0 & \text{ellers,} \end{cases} \quad (1.6)$$

er det oplagt, at $\phi_j \in \mathcal{X}_{0,fe}$, og at $\{\phi_j\}_{j=1}^m$ er en basis for $\mathcal{X}_{0,fe}$, thi enhver funktion i $\mathcal{X}_{0,fe}$ er entydig bestemt ved dens værdi i de indre punkter x_j , $j = 1, \dots, m$.

For at anvende Galerkins metode giver vi (1.5) en variationel formulering: Find en funktion $u \in \mathcal{H}_0^1([0, 1])$, således at

$$\int_I (u'v' + uv) dx = \int_I f v dx \quad (1.7)$$

for alle funktioner $v \in \mathcal{H}_0^1([0, 1])$. Denne ligning er fremkommet ved at multiplicere (1.5) med funktionen v og dernæst foretage en partiel integration over I for integralet

$$-\int_I u''v dx = [-u'v]_0^1 + \int_I u'v' dx = \int_I u'v' dx \quad (1.8)$$

idet $v(0) = v(1) = 0$.

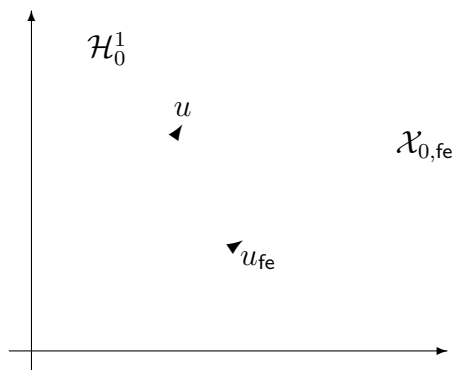
En (analog) endelig dimensional formulering af (1.7) giver os den endelige element metode for (1.5): Find $u_{\text{fe}} \in \mathcal{X}_{0,\text{fe}}$ således at

$$\int_I (u'_{\text{fe}} v' + u_{\text{fe}} v) dx = \int_I f v dx, \quad \text{for alle } v \in \mathcal{X}_{0,\text{fe}}. \quad (1.9)$$

Ved nu at bemærke, at $\mathcal{X}_{0,\text{fe}} \subset \mathcal{H}_0^1([0, 1])$ ses, at (1.7) gælder for $v \in \mathcal{X}_{0,\text{fe}}$. Ved at subtrahere (1.9) fra (1.7) fås følgende relation for $v \in \mathcal{X}_{0,\text{fe}}$

$$\int_I ((u' - u'_{\text{fe}})v' + (u - u_{\text{fe}})v) dx = 0. \quad (1.10)$$

Ligningen (1.10) genkender vi som det indre produkt på \mathcal{H}_0^1 af $u - u_{\text{fe}}$ og v . Med andre ord fejlen $u - u_{\text{fe}}$ er ortogonal på $\mathcal{X}_{0,\text{fe}}$ med hensyn til det indre produkt på \mathcal{H}_0^1 . Dette kan vi også udtrykke som følger: Den endelige element løsning u_{fe} er projektionen (med hensyn til det indre produkt på \mathcal{H}_0^1) af den eksakte løsning u på $\mathcal{X}_{0,\text{fe}}$, dvs. u_{fe} er den funktion i $\mathcal{X}_{0,\text{fe}}$, som ligger tættest på u målt i \mathcal{H}_0^1 normen. Symbolsk kan dette illustreres som vist i figur 1.1, hvor \mathcal{H}_0^1 er repræsenteret som hele planen, og den rette linie gennem centrum repræsenterer $\mathcal{X}_{0,\text{fe}}$.



Figur 1.1: Geometrisk fortolkning af den endelige element løsning

Da $u_{\text{fe}} \in \mathcal{X}_{0,\text{fe}}$, kan u_{fe} skrives på formen

$$u_{\text{fe}} = \sum_{j=1}^m \xi_j \phi_j. \quad (1.11)$$

Koefficienterne ξ_j , $j = 1, \dots, m$ kan bestemmes ved i (1.9) på skift at vælge $v = \phi_i$, $i = 1, \dots, m$. Herved fås et $m \times m$ lineært ligningssystem

$$\mathbf{A}\xi = b, \quad (1.12)$$

hvor \mathbf{A} er en $m \times m$ matrix med elementer

$$a_{ij} = \int_I (\phi_j' \phi_i' + \phi_j \phi_i) dx, \quad (1.13)$$

og b en vektor med elementer

$$b_j = \int_I f \phi_j dx, \quad (1.14)$$

med ξ som løsning. Matricen \mathbf{A} kaldes en stivheds (eng. stiffness) matrix, og b kaldes en kraft (eng. load) vektor. Vi bemærker, at \mathbf{A} i dette tilfælde er symmetrisk, positiv definit og tyndt besat (eng. sparse).

Den grundlæggende ide i den endelige element metode er altså at søge en approksimeret løsning i et løsningsrum bestående af polynomier defineret på en given netinddeling. Den approksimerede løsning vælges nu således, at den er ortogonal på basisfunktioner for et givet testrum af polynomier defineret på samme netinddeling. På denne måde fås et lineært ligningssystem, hvor løsningen er koefficienterne i den linearkombination af løsningsfunktioner, der approksimerer den eksakte løsning.

1.4 Den endelige element metode

Eksemplet i forrige afsnit illustrerer 3 vigtige aspekter ved den endelige element metode. For det første valg af løsnings- og testfunktioner. For det andet valg af inddeling af intervallet og for det tredje omformuleringen af differenttialligningen til et variationsproblem. En mere skematisk oversigt over den endelige element metode lyder

1. Omformulering af differenttialligningen til et variationsproblem.
2. Valg af endelige elementer til brug for inddelingen af domænet.
3. Valg af løsnings- og testrum også kaldet endelige element rum.
4. Opskrivning og løsning af det tilhørende lineære ligningssystem.

Selve løsningen, af det til den endelige element metode hørende lineære ligningssystem, skal vi ikke beskæftige os nærmere med. Læseren formodes at være bekendt med metoder hertil fra kurser i lineær algebra samt kurser i numerisk analyse. Derimod skal vi i de næste 2 kapitler studere variationsformer og endelige element rum.

Andre mindst lige så vigtige aspekter i den endelige element metode er funktionsinterpolation, automatisk netinddeling og fejlestimering. Den netinddelings algoritme, som vi skal studere i kapitel 5, benytter en ganske bestemt a priori fejlligning, som derfor vil blive verificeret i kapitel 4 om polynomsinterpolation. Fejlestimerings metoder for den endelige element metode kan groft opdeles i to kategorier: residuum metoder og udglatnings metoder. Kapitel 6 indeholder en gennemgang af en fejlestimator af sidstnævnte type.

Kapitel 2

Variationsformer

Et karakteristikum ved den endelige element metode er, at man forsøger at omformulere et givet differentiallyignings problem til et variations problem. Vi skal i dette kapitel blandt andet forklare, hvilke fordele der er ved at gå fra et differentiallyignings problem til et variations problem. Da en sådan omformulering langt fra er entydig, skal vi desuden forsøge at opstille nogle “rimelige” krav til en sådan. De opstillede krav til variationsformuleringen skal blandt andet sikre, at der eksisterer en entydig løsning til problemet. I afsnit 2.6 findes beviser, der godtgør at de opstillede krav rent faktisk sikrer eksistens og entydighed. I [40] kan man finde en række af eksempler, hvor differentiallyignings problemer bliver omformuleret til variations problemer.

2.1 Problemstilling

Betragt problemet: Find $u_{\text{cl}} \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ så

$$(D) \quad \mathcal{D}u_{\text{cl}} = f,$$

hvor \mathcal{D} er en anden ordens, lineær differentialoperator med rand- eller begyndelsesværdibetingelser og f en given funktion. Dette problem vil vi så gerne omformulere til: Find $u_{\text{ex}} \in \mathcal{X}(\Omega) = \mathcal{X}$ så

$$(V) \quad B(u_{\text{ex}}, v) = F(v) \text{ for alle } v \in \mathcal{Y}(\Omega) = \mathcal{Y}.$$

Her er \mathcal{X} og \mathcal{Y} normerede lineære vektorrum, B en passende bilinear form defineret på $\mathcal{X} \times \mathcal{Y}$, og F en passende lineær afbildning defineret på \mathcal{Y} . I første omgang skal vi ikke præcisere, hvordan \mathcal{X} og \mathcal{Y} skal vælges. Vi skal dog forudsætte, at $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega}) \subset \mathcal{X}$, således at der er muligt at tale om $B(u_{\text{cl}}, \cdot)$.

2.2 Fordele ved variationsformuleringen

Når man forsøger at omformulere et differentiaallignings problem til et variations problem skyldes det naturligvis, at man opnår visse fordele derved som fx

- Vi skal som hovedregel vælge B , F , \mathcal{X} og \mathcal{Y} på en sådan måde, at problemet (V) er en generalisering af (D), dvs. (V) kan have løsninger, selvom (D) ikke har det.
- Når man søger løsninger til (D) i $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ skyldes dette traditionelle/klassiske løsningsmetoder. Fysiske problemstillinger har dog ofte løsninger, der ikke tilhører $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$, men vi er alligevel interesseret i at løse disse problemer. Dette er måske muligt i en mere generel formulering af problemet. I [40] findes eksempler, hvor den generaliserede løsning til (V) kan ses som grænseværdien for en følge af klassiske løsninger til (mere komplicerede) versioner af (D).
- Der er ofte lettere at påvise eksistens og entydighed af en løsning for problemer givet ved (V) fremfor (D).
- Fejlestimation er nemmere for problemer givet ved (V) end for (D).

2.3 “Rimelige krav”

Som tidligere nævnt kan en omformulering af problemet (D) til (V) konstrueres på forskellige måder. Vi skal i dette afsnit forsøge at opstille nogle “rimelige” krav til bilinear formen B , der vil sikre en fornuftig omformulering af problemet. Man bør skelne mellem to typer af krav til bilinear formen B

1. “Fornuftige” krav.
2. Matematiske krav.

2.3.1 “Fornuftige” krav

Følgende krav bør være opfyldt for at sikre, at vores fysiske og intuitive forestillinger om problemets opførsel holder stik

1. Hvis $u \in \mathcal{X}$, og $B(u, v) = 0$ for alle $v \in \mathcal{Y}$, da er $u = 0$.
2. Hvis der findes en løsning u_{cl} til (D), da skal $u_{cl} \in \mathcal{X}(\Omega)$, og $B(u_{cl}, v) = F(v)$ for alle $v \in \mathcal{Y}$.

3. For en vis klasse \mathcal{F} af lineære afbildninger F defineret på \mathcal{Y} , skal der for alle $F \in \mathcal{F}$, findes et $u_{\text{ex}} \in \mathcal{X}$, så $B(u_{\text{ex}}, v) = F(v)$ for alle $v \in \mathcal{Y}$. Klassen \mathcal{F} skal indeholde $F = 0$ for at sikre, at det homogene problem har en løsning.
4. Hvis u_{ex} er en løsning til (V), gælder $\|u_{\text{ex}}\| \leq \alpha \|F\|_{\mathcal{Y}^*}$ for alle $F \in \mathcal{F}$, hvor α er en positiv konstant uafhængig af F , og \mathcal{Y}^* er \mathcal{Y} 's duale rum.

2.3.1 BEMÆRKNING Betingelse (1) følger af (4) ved at vælge $F = 0$, men er nævnt som en selvstændig betingelse, idet (1) og (4) har forskellige fysiske fortolkninger.

For at retfærdiggøre ovenstående krav skal vi nu forklare, hvilke praktiske konsekvenser kravene har:

- ad 1). Denne betingelse sikrer entydighed af den generaliserede løsning. Antag at $B(u_1, v) = F(v)$, og $B(u_2, v) = F(v)$ for alle $v \in \mathcal{Y}$, da vil $B(u_1 - u_2, v) = 0$ for alle $v \in \mathcal{Y}$. Altså må $u_1 = u_2$.
- ad 2). Denne betingelse sikrer, at når der findes en løsning u_{cl} til (D), da er denne løsning også en generaliseret løsning til (V).
- ad 3). Denne betingelse sikrer generaliseret eksistens og entydighed af variations problemet. Dette skal forstås således, at hvis der ændres lidt på F (dvs. på data), skal der stadig eksistere en løsning til problemet. Hvis denne betingelse ikke er opfyldt, kan vi risikere, at der ikke findes en løsning til det generaliserede problem, såfremt data (dvs. F) er behæftet med en vis usikkerhed.
- ad 4). Denne betingelse sikrer, at løsningen afhænger kontinuert af data. Antag at $B(u_1, v) = F_1(v)$, og $B(u_2, v) = F_2(v)$ for alle $v \in \mathcal{Y}$, da er $B(u_1 - u_2, v) = (F_1 - F_2)(v)$ for alle $v \in \mathcal{Y}$, så $\|u_1 - u_2\| \leq \alpha \|F_1 - F_2\|_{\mathcal{Y}^*}$. Altså hvis F ændres lidt, vil løsningen også kun ændres lidt. Hvis denne betingelse ikke er opfyldt, kan vi risikere at få ustabile numeriske løsninger, hvis F fx kun er kendt med usikkerhed.

2.3.2 Matematiske krav

For at sikre at omformuleringen fra (D) til (V) opfylder ovenstående betingelser 1.-4., må vi stille en række matematiske krav til bilinear formen B og rummene \mathcal{X} og \mathcal{Y} . Vi skal her præcisere disse krav samt se, at betingelserne 1.-4. så vil være opfyldt.

Betingelsen (1) følger som nævnt tidligere af (4) ved at vælge $F = 0$.

Betingelsen (2) kan man ikke sikre ved matematiske krav. Vi må derfor være påpasselige, når vi udleder (V) fra (D) og holde øje med, at denne betingelse er opfyldt.

Betingelsen (3) kan opfyldes på to forskellige måder. I begge situationer skal vi antage, at B og F er begrænsede.

3.1 Er $\mathcal{X} = \mathcal{Y}$, \mathcal{X} et Hilbert rum, og er B \mathcal{X} -elliptisk, dvs. $B(u, u) \geq \alpha \|u\|_{\mathcal{X}}^2$, $\forall u \in \mathcal{X}$, hvor α er en positiv konstant, følger eksistens og entydighed af variations problemet af Lax-Milgrams lemma, se afsnit 2.6.

3.2 Såfremt betingelserne

- \mathcal{X} er et Banach rum og \mathcal{Y} er et refleksivt Banach rum.
- $\inf_{\|u\|_{\mathcal{X}}=1} \sup_{\|v\|_{\mathcal{Y}}=1} |B(u, v)| \geq C > 0$.
- For alle $v \in \mathcal{Y} \setminus \{0\}$ findes et $u_v \in \mathcal{X}$ så $|B(u_v, v)| > 0$.

er opfyldte vil eksistens og entydighed være sikret jvf. sætning 2.6.2 i afsnit 2.6. Betingelsen (b) kaldes også inf-sup betingelsen eller Babuška-Brezzi betingelsen.

At betingelse 3.2 rent faktisk er stærkere end betingelse 3.1 er indeholdt i følgende lemma.

2.3.2 LEMMA *Betingelse 3.1 \Rightarrow 3.2.*

BEVIS Da $\mathcal{X} = \mathcal{Y}$ er et Hilbert rum, er det klart at, (a) i 3.2 er opfyldt. Da B er \mathcal{X} -elliptisk vil (b) hhv. (c) i 3.2 være opfyldt ved at vælge $u = v$ hhv. $u_v = v$.

2.3.3 BEMÆRKNING Det er klart, at 3.2 $\not\Rightarrow$ 3.1, idet der let kan gives eksempler på Banach rum, der ikke er Hilbert rum.

Betingelsen (4) kan ligeledes opfyldes på to forskellige måder

4.1 Hvis $\mathcal{X} = \mathcal{Y}$ og B er \mathcal{X} -elliptisk, dvs. $B(u, u) \geq \alpha \|u\|_{\mathcal{X}}^2$, da vil $\|u_{\text{ex}}\|_{\mathcal{X}}^2 \leq \frac{1}{\alpha} B(u_{\text{ex}}, u_{\text{ex}}) = \frac{1}{\alpha} F(u_{\text{ex}}) \leq \frac{1}{\alpha} \|F\|_{\mathcal{Y}^*} \|u_{\text{ex}}\|_{\mathcal{X}}$, så $\|u_{\text{ex}}\|_{\mathcal{X}} \leq \frac{1}{\alpha} \|F\|_{\mathcal{Y}^*}$.

4.2 Hvis der for alle $u \in \mathcal{X} \setminus \{0\}$, findes et $v_u \in \mathcal{Y} \setminus \{0\}$, så en af følgende betingelser er opfyldt

- (a) $\|v_u\|_{\mathcal{Y}} \leq c_1 \|u\|_{\mathcal{X}}$ og $|B(u, v_u)| \geq c_2 \|u\|_{\mathcal{X}}^2 \geq \frac{c_2}{c_1} \|u\|_{\mathcal{X}} \|v_u\|_{\mathcal{Y}}$.
- (b) $|B(u, v_u)| \geq c_3 \|u\|_{\mathcal{X}} \|v_u\|_{\mathcal{Y}}$.

Da gælder $\|u_{\text{ex}}\|_{\mathcal{X}}\|v_u\|_{\mathcal{Y}} \leq \frac{1}{c}|B(u_{\text{ex}}, v_u)| = \frac{1}{c}F(v_u) \leq \frac{1}{c}\|F\|_{\mathcal{Y}^*}\|v_u\|_{\mathcal{Y}}$,
dvs.

$\|u_{\text{ex}}\|_{\mathcal{X}} \leq \frac{1}{c}\|F\|_{\mathcal{Y}^*}$, hvor $c \in \{\frac{c_2}{c_1}, c_3\}$.

2.3.4 BEMÆRKNING Betingelse (b) i 3.2 er ækvivalent med betingelse (b) i 4.2.

2.3.5 BEMÆRKNING Rummene $\mathcal{H}_k^p(I)$, $p \neq 1, \infty$ er refleksive.

2.4 Randbetingelser

Vi skal nu se nærmere på, hvad randbetingelser betyder ved omformuleringen til variationsproblemet. Det er svært at sige noget om den generelle situation pga. de meget forskelligeartede randbetingelser. Vi vil derfor nøjes med et eksempel, der illustrerer de forhold, man skal tage højde for. Vi skal dog se, at man overordnet (i variationsformulering sammenhæng) kan skelne mellem to typer at randbetingelser.

2.4.1 Klassificering af randbetingelser

2.4.1 DEFINITION *Randbetingelser, der kan indpasses i variationsformuleringen via restriktioner på løsnings- og/eller testrummet, kaldes stærke (eng. essential) randbetingelser. Randbetingelser, der ikke kan indpasses via restriktioner, kaldes svage (eng. natural).*

2.4.2 BEMÆRKNING Stærke randbetingelser kaldes også for væsentlige randbetingelser, ligesom man til tider kalder svage randbetingelser for naturlige randbetingelser.

2.4.3 EKSEMPEL Lad os som et eksempel betragte Poissons ligning

$$-\Delta u + u = f \quad \text{i } \Omega \subset \mathbb{R}^2, \quad (2.1)$$

i to dimensioner med randbetingelserne

$$\frac{\partial u}{\partial \mathbf{n}} = g_1 \quad \text{på } \Gamma_1, \quad (2.2)$$

$$u = g_2 \quad \text{på } \Gamma_2, \quad (2.3)$$

hvor $\Gamma_1 \cup \Gamma_2 = \partial\Omega$. Lad os et kort øjeblik se bort fra randbetingelser. Det checkes da let, at en klassisk løsning til (2.1) også er en løsning til variations

problemet:

$$\text{Find } u \in \mathcal{H}^1(\Omega) : \iint_{\Omega} (\nabla u \cdot \nabla v + uv) dx dy = \iint_{\Omega} f v dx dy, \quad \forall v \in \mathcal{H}_0^1(\Omega). \quad (2.4)$$

Randbetingelsen $u = g_2$ på Γ_2 er en stærk randbetingelse, idet den kan indpasses i løsningsrummet $\mathcal{H}^1(\Omega)$ via restriktionen $u|_{\Gamma_2} = g_2$. Randbetingelsen $\frac{\partial u}{\partial \mathbf{n}} = g_1$ på Γ_1 er derimod en svag randbetingelse (idet $\frac{\partial u}{\partial \mathbf{n}}$ ikke er defineret på kurver i \mathcal{H}^1), så den må indpasses i selve variationsformen via leddet

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} v ds \quad (2.5)$$

tilføjet til højresiden af (2.4).

Vi kunne vælge at bruge samme test- og løsningsrum, men betingelsen $v|_{\Gamma_2} = g_2$ på testrummet er upraktisk, idet det giver “komplicerede” basisfunktioner. Som testfunktioner vil vi derfor vælge de $v \in \mathcal{H}^1(\Omega)$, hvor $v|_{\Gamma_2} = 0$, hvilket giver os følgende variationsformulering af (2.1): Find $u \in \mathcal{H}^1(\Omega)$ med $u|_{\Gamma_2} = g_2$ så

$$\iint_{\Omega} (\nabla u \cdot \nabla v + uv) dx dy = \iint_{\Omega} f v dx dy + \int_{\Gamma_1} \frac{\partial u}{\partial \mathbf{n}} v ds \quad (2.6)$$

for alle $v \in \mathcal{H}^1(\Omega)$, $v|_{\Gamma_2} = 0$

Af hensyn til teoretiske og praktiske situationer vælger man ofte samme test- og løsningsrum. Vi skal nu se, hvorledes dette kan opnås for ovenstående eksempel. Sæt $u_0 = u - w$, hvor $w \in \mathcal{C}^2(\bar{\Omega})$ er valgt således, at $w|_{\Gamma_2} = g_2$. Da vil $u_0|_{\Gamma_2} = 0$ og

$$\begin{aligned} -\Delta u_0 + u_0 &= -\Delta(u - w) + (u - w) \\ &= -\Delta u + u + \Delta w - w \\ &= f - (-\Delta w + w). \end{aligned} \quad (2.7)$$

Vi kan nu formulere variationsproblemet som: Find $u_0 \in \mathcal{H}^1(\Omega)$ med $u_0|_{\Gamma_2} = 0$ så

$$\iint_{\Omega} (\nabla u_0 \cdot \nabla v + u_0 v) dx dy = \iint_{\Omega} f_0 v dx dy + \int_{\Gamma_1} \frac{\partial u}{\partial \mathbf{n}} v ds \quad (2.8)$$

for alle $v \in \mathcal{H}^1(\Omega)$ med $v|_{\Gamma_2} = 0$. Her er $f_0 = f - (-\Delta w + w)$.

Som det fremgår af definitionen og eksemplet bør man i så høj grad som muligt indpasse randbetingelser som stærke randbetingelser, idet disse har den umiddelbare fordel, at de nedsætter dimensionen af løsnings- og/eller testrummene.

2.5 Konstruktionsregler

Man kan ikke give præcise retningslinier for, hvordan man udleder (V) fra (D). Vi skal dog forsøge at angive nogle generelle konstruktionsregler for, hvordan B , F , \mathcal{X} og \mathcal{Y} kan udledes, således at man får en generaliseret formulering (V) af (D).

1. Udled B og F fra (D) udelukkende ved at bruge operationer som er "i orden" for den klassiske løsning til (D). Typiske eksempler er multiplikation med en testfunktioner og integration over det betragtede domæne.
2. Check at B og F er lineære og begrænsede.
3. Overvej $\mathcal{H}^p(\Omega)$ hhv. $\mathcal{H}^q(\Omega)$ som løsnings- hhv. testrum, hvor p og q er valgt så tæt på hinanden som muligt. Derved minimeres $\max(p, q)$. Disse rum har flere fordele. For det første er $\mathcal{C}^p \subset \mathcal{H}^p$, og for det andet er disse rum Hilbert rum. Ved at vælge p og q så små som muligt begrænses glathedskravet til løsnings- og testfunktioner mest muligt.
4. Har differentiaalligningen randbetingelser, forsøg da at indpasse disse i form af restriktioner på løsnings- og/eller testrummet jvf. diskussionen om stærke randbetingelser.
5. Check at betingelse 3 og 4 fra afsnit 2.3 er opfyldt.

2.6 Eksistens og entydigheds sætninger

Lad os først vise, at betingelse 3.1 fra side 10 sikrer eksistens og entydighed.

2.6.1 SÆTNING (LAX-MILGRAMS LEMMA) *Lad \mathcal{X} være et Hilbert rum, $B : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ en begrænset \mathcal{X} -elliptisk bilinear form, og lad $F : \mathcal{X} \rightarrow \mathbb{R}$ være en begrænset lineær afbildning. Da har variations problemet: Bestem et element $u \in \mathcal{X}$ så*

$$B(u, v) = F(v) \quad \forall v \in \mathcal{X}, \quad (2.9)$$

én og kun én løsning.

BEVIS Da B er begrænset findes en konstant M så

$$|B(u, v)| \leq M \|u\| \|v\| \quad \text{for alle } u, v \in V. \quad (2.10)$$

For ethvert $u \in \mathcal{X}$ er funktionalen $L : \mathcal{X} \rightarrow \mathbb{R}$ givet ved $L(v) = B(u, v)$ kontinuert og lineær. Altså findes et entydigt bestemt element $x_u \in \mathcal{X}$, se fx [39, theorem 4.2], så

$$L(v) = \langle v, x_u \rangle, \quad (2.11)$$

hvor $\langle \cdot, \cdot \rangle$ er det indre produkt i \mathcal{X} . Et indre produkt er specielt en lineær funktional, hvorfor der findes et entydigt bestemt element $f_{x_u} \in \mathcal{X}^*$, så $\langle v, x_u \rangle = f_{x_u}(v)$, hvor \mathcal{X}^* er det duale rum til \mathcal{X} . Lad os for ethvert $u \in \mathcal{X}$ definere $Au = f_{x_u} \in \mathcal{X}^*$, så vil

$$B(u, v) = L(v) = \langle v, x_u \rangle = f_{x_u}(v) = Au(v). \quad (2.12)$$

Med $\|\cdot\|^*$ betegnes normen i \mathcal{X}^* . Vi har nu fra (2.10) og (2.12)

$$\|Au\|^* = \sup_{v \in V \setminus \{0\}} \frac{|Au(v)|}{\|v\|} \leq M\|u\|, \quad (2.13)$$

hvilket viser at afbildningen $A : \mathcal{X} \rightarrow \mathcal{X}^*$ er kontinuert med $\|A\| \leq M$.

I følge Riesz' repræsentations sætning findes en afbildning $\tau : \mathcal{X}^* \rightarrow \mathcal{X}$ så

$$F(v) = \langle \tau F, v \rangle \quad \text{for alle } F \in \mathcal{X}^* \text{ og alle } v \in \mathcal{X}. \quad (2.14)$$

Vi har nu for $u \in \mathcal{X}$ og for alle $v \in \mathcal{X}$

$$B(u, v) = F(v) \Leftrightarrow Au(v) = F(v) \Leftrightarrow \langle \tau Au, v \rangle = \langle \tau F, v \rangle \Leftrightarrow \tau Au - \tau F = 0 \quad (2.15)$$

At løse variations problemet (2.9) er derfor ækvivalent med at løse ligningen $\tau Au = \tau F$.

For at vise at denne ligning har én og kun én løsning, skal vi se, at afbildningen

$$v \in \mathcal{X} \rightarrow v - \rho(\tau Av - \tau F) \in \mathcal{X} \quad (2.16)$$

er en kontraktion for passende valgt ρ . Da \mathcal{X} er et Hilbert rum, er \mathcal{X} specielt et fuldstændigt metrisk rum, og som bekendt har en kontraktion defineret på et fuldstændigt metrisk rum netop et fixpunkt.

Bemærkes nu at

$$\langle \tau Av, v \rangle = Av(v) = B(v, v) \geq \alpha\|v\|^2, \quad (2.17)$$

$$\|\tau Av\| = \|Av\|^* \leq \|A\|\|v\| \leq M\|v\|, \quad (2.18)$$

hvor α skyldes \mathcal{X} -ellipticitet, følger kontraktionen let thi

$$\|v - \rho\tau Av\|^2 = \|v\|^2 - 2\rho\langle \tau Av, v \rangle + \rho^2\|\tau Av\|^2 \quad (2.19)$$

$$\leq (1 - 2\rho\alpha + \rho^2 M^2)\|v\|^2. \quad (2.20)$$

Afbildningen defineret ved (2.16) er derfor en kontraktion for $\rho \in]0, 2\alpha/M^2[$. Altså findes et element $v \in V$, så $v - \rho(\tau Av - \tau F) = v$, dvs $\tau Av = \tau F$ som ønsket.

Betingelse 3.2 fra side 10 sikrer ligeledes eksistens og entydighed jvf. nedenstående sætning

2.6.2 SÆTNING *Lad $B : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ være en begrænset bilinear form, der opfylder betingelserne 3.2 (b) og (c) fra side 10, og $F : \mathcal{Y} \rightarrow \mathbb{R}$ er en begrænset lineær afbildning. Her er \mathcal{X} et Banach rum, og \mathcal{Y} er et refleksivt Banach rum. Da har variations problemet: Bestem et element $u \in \mathcal{X}$ så*

$$B(u, v) = F(v) \quad \forall v \in \mathcal{Y}, \quad (2.21)$$

én og kun én løsning.

BEVIS Lad $R(u) : v \in \mathcal{Y} \rightarrow R(u)v \in \mathbb{R}$ være en lineær begrænset funktional tilhørende \mathcal{Y}^* defineret ved

$$R(u)v = B(u, v) \quad \forall v \in \mathcal{Y}. \quad (2.22)$$

Lad tilsvarende $F \in \mathcal{Y}^*$ være givet ved

$$Fv = F(v) \quad \forall v \in \mathcal{Y}. \quad (2.23)$$

Målet er at vise, at $R(\mathcal{X}) = \mathcal{Y}^*$, idet vi så kan anvende Banachs sætning, se fx [16, s. 34], til at få det ønskede resultat. Lad os først vise, at $R(\mathcal{X})$ er afsluttet i \mathcal{Y}^* . Lad $\{R(u_n)\}$ være en Cauchy følge i \mathcal{Y}^* . Da er $\{u_n\}$ en Cauchy følge i \mathcal{X} , thi

$$\begin{aligned} \|R(u_n) - R(u_m)\|_{\mathcal{Y}^*} &= \|R(u_n - u_m)\|_{\mathcal{Y}^*} \\ &= \|R(\frac{u_n - u_m}{\|u_n - u_m\|_{\mathcal{X}}})\|_{\mathcal{Y}^*} \|u_n - u_m\|_{\mathcal{X}} \\ &= \sup_{\|v\|_{\mathcal{Y}=1}} |R(\frac{u_n - u_m}{\|u_n - u_m\|_{\mathcal{X}}})v| \|u_n - u_m\|_{\mathcal{X}} \\ &\geq \inf_{\|u\|_{\mathcal{X}=1}} \sup_{\|v\|_{\mathcal{Y}=1}} |B(u, v)| \|u_n - u_m\|_{\mathcal{X}} \\ &\geq C \|u_n - u_m\|_{\mathcal{X}}. \end{aligned} \quad (2.24)$$

Konstanten C kommer fra inf-sup betingelsen 3.2(b). Antag nu at $\lim_{n \rightarrow \infty} u_n = u$ i \mathcal{X} . Vi finder så

$$\begin{aligned} \|R(u_n) - R(u)\|_{\mathcal{Y}^*} &= \sup_{\|v\|_{\mathcal{Y}=1}} |B(u_n - u, v)| \\ &\leq M \|u_n - u\|_{\mathcal{X}}, \end{aligned} \quad (2.25)$$

da B er begrænset. Altså vil $\lim_{n \rightarrow \infty} R(u_n) = R(u)$ i \mathcal{Y}^* , hvorfor $R(\mathcal{X})$ er afsluttet i \mathcal{Y}^* . Da R er lineær, vil $R(\mathcal{X})$ være et afsluttet underrum af \mathcal{Y}^* . Vi vil nu vise, at $R(\mathcal{X}) = \mathcal{Y}^*$. Antag omvendt at $R(\mathcal{X}) \neq \mathcal{Y}^*$. Da $R(\mathcal{X})$ er et underrum af \mathcal{Y}^* , vil $\mathcal{Y}^* \setminus R(\mathcal{X})$ indeholde mindst et underrum af dimension 1. Altså findes der et $g \in \mathcal{Y}^{**} \setminus \{0\}$ så $g(R(u)) = 0$ for alle $u \in \mathcal{X}$. Da \mathcal{Y} er reflektiv, er den kanoniske afbildning $\iota : \mathcal{Y} \rightarrow \mathcal{Y}^{**}$ defineret ved

$$\iota v(v^*) = v^*(v), \quad \forall v^* \in \mathcal{Y}^*, v \in \mathcal{Y}, \quad (2.26)$$

surjektiv. Derfor findes for alle $h \in \mathcal{Y}^{**}$, et $v_h \in \mathcal{Y}$ så $h = \iota v_h$, dvs.

$$h(v^*) = \iota v_h(v^*) = v^*(v_h) \quad \forall v^* \in \mathcal{Y}^*. \quad (2.27)$$

Vi bemærker specielt, at

$$v_h = 0 \Rightarrow v^*(v_h) = 0 \quad \forall v^* \in \mathcal{Y}^* \Rightarrow h(v^*) = 0 \quad \forall v^* \in \mathcal{Y}^* \Rightarrow h = 0. \quad (2.28)$$

Da $g \in \mathcal{Y}^{**} \setminus \{0\}$ findes et $v_g \in \mathcal{Y} \setminus \{0\}$ så $g(R(u)) = \iota v_g(R(u)) = R(u)v_g = 0 \quad \forall u \in \mathcal{X}$. Med andre ord $B(u, v_g) = 0 \quad \forall u \in \mathcal{X}$ i modstrid med betingelse 3.2(c), hvorfor $R(\mathcal{X}) = \mathcal{Y}^*$.

Vi har nu vist, at $R : \mathcal{X} \rightarrow \mathcal{Y}^*$ er en lineær, begrænset og surjektiv operator, og det følger så af Banachs sætning, at R har en lineær begrænset invers $R^{-1} : \mathcal{Y}^* \rightarrow \mathcal{X}$. Vælges nu $u = R^{-1}(F)$ fås

$$B(u, v) = R(R^{-1}(F))v = Fv = F(v) \quad \forall v \in \mathcal{Y}, \quad (2.29)$$

hvilket skulle vises.

2.7 Eksistens og entydighed for det diskrete problem

Ovenfor har vi kun beskæftiget os med eksistens og entydighed for det kontinuerte problem. Vi skal nu vise, at hvis man har eksistens og entydighed for det kontinuerte problem, da vil eksistens og entydighed af det diskrete problem følge automatisk.

Den diskrete version af problemet (V) lyder: Find $u \in \mathcal{X}_{\text{fe}}(\Omega) = \mathcal{X}_{\text{fe}}$ så

$$\tilde{B}(u, v) = \tilde{F}(v), \quad \text{for alle } v \in \mathcal{Y}_{\text{fe}}(\Omega) = \mathcal{Y}_{\text{fe}}. \quad (2.30)$$

Her er \tilde{B} og \tilde{F} diskrete versioner af B hhv. F , og \mathcal{X}_{fe} hhv. \mathcal{Y}_{fe} er endelige element rum indeholdt i \mathcal{X} hhv. \mathcal{Y} . Vi skal i næste kapitel give en formel definition af endelige element rum. For tiden er det tilstrækkeligt at opfatte

\mathcal{X}_{fe} hhv. \mathcal{Y}_{fe} som endelig-dimensionale underrum af \mathcal{X} hhv. \mathcal{Y} . Hvis $\{\psi\}_{i=1}^n$ og $\{\phi\}_{j=1}^m$ betegner baser for \mathcal{X}_{fe} hhv. \mathcal{Y}_{fe} , er det kendt fra den lineære algebra, at \tilde{B} og \tilde{F} kan skrives på formen

$$\tilde{B}(u, v) = b^t K a \quad \text{og} \quad \tilde{F}(v) = b^t q, \quad (2.31)$$

hvor $u = \sum_{i=1}^n a_i \psi_i$, $v = \sum_{j=1}^m b_j \phi_j$, $K_{ij} = \tilde{B}(\psi_i, \phi_j)$. Vi skal betragte tilfældet, hvor $\dim \mathcal{X} = \dim \mathcal{Y}$, dvs. K er kvadratisk.

2.7.1 SÆTNING *Lad K være som ovenfor. Følgende betingelser er da ækvivalente*

1. *Det diskrete problem $\tilde{B}(u, v) = \tilde{F}(v)$ har én og kun én løsning.*
2. *K er invertibel.*
3. $\inf_{\|u\|_{\mathcal{X}_{\text{fe}}}=1} \sup_{\|v\|_{\mathcal{Y}_{\text{fe}}}=1} |\tilde{B}(u, v)| \geq C > 0$.

BEVIS Det er klart at vi har $1. \Leftrightarrow 2.$. Lad os vise $2. \Leftrightarrow 3.$. Vi har:

$$\begin{aligned} & \inf_{\|u\|_{\mathcal{X}_{\text{fe}}}=1} \sup_{\|v\|_{\mathcal{Y}_{\text{fe}}}=1} |\tilde{B}(u, v)| \geq C > 0 \\ & \Updownarrow \\ & \forall u \in \mathcal{X}_{\text{fe}} \setminus \{0\} \exists v_u \in \mathcal{Y}_{\text{fe}} \setminus \{0\} : |\tilde{B}(u, v_u)| \geq C \|u\|_{\mathcal{X}_{\text{fe}}} \|v_u\|_{\mathcal{Y}_{\text{fe}}} \\ & \Updownarrow \\ & \forall u \in \mathcal{X}_{\text{fe}} \setminus \{0\} \exists v_u \in \mathcal{Y}_{\text{fe}} \setminus \{0\} : |\tilde{B}(u, v_u)| > 0 \\ & \Updownarrow \\ & \forall \xi \in \mathbb{R}^n \setminus \{0\} : K\xi \neq 0 \\ & \Updownarrow \\ & K \text{ er invertibel} \end{aligned}$$

2.7.2 BEMÆRKNING Man kan vise, at betingelsen

$$4. \inf_{\|v\|_{\mathcal{Y}_{\text{fe}}}=1} \sup_{\|u\|_{\mathcal{X}_{\text{fe}}}=1} |\tilde{B}(u, v)| > 0.$$

er ækvivalent med betingelserne 3. fra sætning 2.7.1.

2.7.3 BEMÆRKNING Som i det kontinuerte tilfælde, er betingelserne

$$\forall u \in \mathcal{X} \setminus \{0\} \exists v_u \in \mathcal{Y} \setminus \{0\} : |\tilde{B}(u, v_u)| \geq \|u\|_{\mathcal{X}} \|v_u\|_{\mathcal{Y}} \quad (2.32)$$

$$\forall v \in \mathcal{Y} \setminus \{0\} \exists u_v \in \mathcal{X} \setminus \{0\} : |\tilde{B}(u_v, v)| \geq \|u_v\|_{\mathcal{X}} \|v\|_{\mathcal{Y}} \quad (2.33)$$

ækvivalente med hhv. betingelse 3. fra sætning 2.7.1 og betingelse 4. fra bemærksetning 2.7.2.

2.8 Eksempel på konstruktion af variationsform

2.8.1 EKSEMPEL Som et konkret eksempel på anvendelsen af ovenstående principper skal vi betragte følgende homogene Dirichlet problem: Bestem $u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ så

$$\begin{aligned} -\Delta u + au &= f \quad \text{i } \Omega, \\ u &= 0 \quad \text{på } \partial\Omega = \Gamma, \end{aligned} \tag{2.34}$$

hvor Ω er et begrænset, sammenhængende område i \mathbb{R}^2 med Lipschitz kontinuert rand Γ . Funktionener a og f antages at tilhøre hhv. $\mathcal{L}^\infty(\Omega)$ og $\mathcal{L}^2(\Omega)$. Desuden antages a at være ikke-negativ næsten overalt i Ω .

Variationsformuleringen af (2.34) fremkommer ved at multiplicere (2.34) med en testfunktion og dernæst foretage en partiel integration over Ω . Derved fås følgende formulering af (2.34): Bestem $u \in \mathcal{X}$ så

$$B(u, v) = F(v) \quad \forall v \in \mathcal{Y}, \tag{2.35}$$

hvor

$$B(u, v) = \int_{\Omega} (\nabla u \nabla v + auv) dx \tag{2.36}$$

og

$$F(v) = \int_{\Omega} f v dx \tag{2.37}$$

Jvf. ovenstående diskussion vil det her være naturligt at vælge $\mathcal{X} = \mathcal{Y} = \mathcal{H}_0^1(\Omega)$. Det er oplagt, at B er symmetrisk og bilinear, samt at F er lineær. For at påvise eksistens og entydighed af en løsning til (2.35), vil det derfor være tilstrækkeligt at redegøre for, at B er begrænset og $\mathcal{H}_0^1(\Omega)$ -elliptisk samt at F er begrænset. Da $\mathcal{X} = \mathcal{Y} = \mathcal{H}_0^1(\Omega)$ er et Hilbert rum vil eksistens og entydighed så følge af Lax-Milgrams lemma.

Lad os starte med at vise, at B er begrænset. For $u, v \in \mathcal{H}_0^1(\Omega)$ have

$$\begin{aligned} |B(u, v)| &\leq \|\nabla u\|_{\mathcal{W}^{0,2}(\Omega)} \|\nabla v\|_{\mathcal{W}^{0,2}(\Omega)} + \|a\|_{\mathcal{W}^{0,\infty}(\Omega)} \|u\|_{\mathcal{W}^{0,2}(\Omega)} \|v\|_{\mathcal{W}^{0,2}(\Omega)} \\ &\leq \max\{1, \|a\|_{\mathcal{W}^{0,\infty}(\Omega)}\} \|u\|_{\mathcal{W}^{1,2}(\Omega)} \|v\|_{\mathcal{W}^{1,2}(\Omega)}, \end{aligned} \tag{2.38}$$

her betegner $\|\cdot\|_{\mathcal{W}^{0,2}(\Omega)}$, $\|\cdot\|_{\mathcal{W}^{0,\infty}(\Omega)}$ og $\|\cdot\|_{\mathcal{W}^{1,2}(\Omega)}$ normerne på $\mathcal{L}^2(\Omega)$, $\mathcal{L}^\infty(\Omega)$ og $\mathcal{H}^1(\Omega)$. Altså er B begrænset. At B er $\mathcal{H}_0^1(\Omega)$ -elliptisk følger af følgende ulighed for $v \in \mathcal{H}_0^1(\Omega)$

$$B(v, v) \geq \int_{\Omega} (\nabla v)^2 dx = |v|_{\mathcal{W}^{1,2}(\Omega)}^2. \quad (2.39)$$

Her er $|\cdot|_{\mathcal{W}^{1,2}(\Omega)}$ \mathcal{H}_0^1 's seminorm. Vi mangler nu kun at vise, at F er begrænset. Lad $v \in \mathcal{H}_0^1(\Omega)$, da vil

$$|f(v)| \leq \|f\|_{\mathcal{W}^{0,2}(\Omega)} \|v\|_{\mathcal{W}^{0,2}(\Omega)} \leq \|f\|_{\mathcal{W}^{0,2}(\Omega)} \|v\|_{\mathcal{W}^{1,2}(\Omega)}, \quad (2.40)$$

som ønsket

Med ovenstående har vi fået opfyldt krav 3 og 4 fra afsnit 2.3. At kravene 1 og 2 er opfyldt verificeres let.

Kapitel 3

Endelige elementer og endelige element rum

Når man skal foretage en inddeling af et domæne i planen i endelige elementer, vælger man af beregningsmæssige hensyn ofte at opdele området i simple geometriske figurer såsom trekanter eller firkanter. Er området ikke-polygonalt kan man enten vælge at benytte kurvede elementer (se fx [19, 31]) eller at foretage en finere inddeling af området langs randen i mindre trekanter/firkanter. Endeligt kan man også vælge at afbilde domænet bijektivt på et “pænere” domæne. Hvilken fremgangsmåde man vælger er ofte en smags-sag. Traditionelt har man arbejdet med inddelinger i trekanter (eller mere generelt n -simplexer), men firkanter er ved at vinde indpas. I dette kapitel, hvor vi formelt skal definere endelige elementer og endelige element rum, vil udgangspunktet være trekanter/ n -simplexer. I et senere kapitel, hvor vi skal studere en netinddelings algoritme, vil rektangler blive anvendt i trianguleringen. Definitionen af nogle af begreberne i dette kapitel varierer fra forfatter til forfatter. Fremgangsmåden, vi her skal anvende, kaldes for en konformerende endelig element metode og er inspireret af [19].

3.1 Notation

Vi skal kort redegøre for notationen vedrørende afledede af funktioner $v : \mathbb{R}^n \rightarrow \mathbb{R}$. Lad e_i , $1 \leq i \leq n$ være den sædvanlige kanoniske basis for \mathbb{R}^n . De

partielle afledede af v vil da blive betegnet med

$$\partial_i v(a) = Dv(a)e_i, \quad (3.1)$$

$$\partial_{ij} v(a) = D^2 v(a)(e_i, e_j), \quad (3.2)$$

$$\partial_{ijk} v(a) = D^3 v(a)(e_i, e_j, e_k), \quad (3.3)$$

osv.

Til tider skal vi benytte en multi-indeks notation. Lad $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ og sæt $|\alpha| = \sum_{i=1}^n \alpha_i$. Ved $\partial^\alpha v(a)$ vil vi da forstå resultatet af at anvende den $|\alpha|$ 'te afledede $D^{|\alpha|} v(a)$ på en vilkårlig $|\alpha|$ -vektor i $(\mathbb{R}^n)^{|\alpha|}$, hvor basisvektoren e_i forekommer α_i gange $1 \leq i \leq n$. For eksempel for $n = 3$ er $\partial^{(1,0,0)} v(a) = \partial_1 v(a)$, $\partial^{(1,1,1)} v(a) = \partial_{123} v(a)$ og $\partial^{(3,0,0)} v(a) = \partial_{111} v(a)$. Endelig skal vi også erstatte de kanoniske basisvektorer e_i med vilkårlige faste vektorer ξ_i for at opnå vilkårlige retningsafledede.

3.2 Endelige element rum

En væsentlig del af den endelige element metode består af konstruktionen af endelige dimensionale underrum \mathcal{X}_{fe} og \mathcal{Y}_{fe} af løsningsrummet \mathcal{X} hhv. testrummet \mathcal{Y} . Fremgangsmåden ved konstruktionen af \mathcal{X}_{fe} og \mathcal{Y}_{fe} er selvfølgelig identisk, og vi skal fremover blot betragte \mathcal{X}_{fe} . Denne konstruktion kan karakteriseres ved 3 aspekter:

1. En triangulering \mathcal{T}_h af $\bar{\Omega}$.
2. For ethvert $K \in \mathcal{T}_h$ skal rummet $P_K = \{v|_K \mid v \in \mathcal{X}_{\text{fe}}\}$ indeholde polynomier.
3. Der skal findes en basis for \mathcal{X}_{fe} , bestående af funktioner med så "lille" støtte som mulig.

Grunden til at vi forlanger, at rummene P_K skal indeholde polynomier, der kan udtrykkes i en basis bestående af funktioner med så "lille" støtte som mulig, er et ønske om dels at få simple beregninger af elementerne i (1.12-1.14) dels at opnå en matrix, der er tyndt besat. Vi vil nu studere de tre aspekter i detaljer.

3.3 Triangulering af $\bar{\Omega}$

Man tillader naturligvis ikke vilkårlige inddelinger af domænet. For at få en sammenhængende teori har følgende definition vist sig at være fornuftig.

3.3.1 DEFINITION Lad $\bar{\Omega}$ betegne det domæne, hvori en given differential-ligning søges løst. Ved en triangulering \mathcal{T}_h af $\bar{\Omega}$ forstås en inddeling af $\bar{\Omega}$ i et endeligt antal delmængder K på en sådan måde, at følgende betingelser er opfyldt:

\mathcal{T}_h 1. $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} K$.

\mathcal{T}_h 2. For alle $K \in \mathcal{T}_h$ gælder $K = \bar{K}$ og $K^\circ \neq \emptyset$.

\mathcal{T}_h 3. For alle $K_1, K_2 \in \mathcal{T}_h$ hvor $K_1 \neq K_2$ gælder $K_1^\circ \cap K_2^\circ = \emptyset$.

\mathcal{T}_h 4. For alle $K \in \mathcal{T}_h$ er ∂K Lipschitz kontinuert.

Vi skal senere tilføje en femte betingelse, der skal være med til at sikre kontinuitet over randene for naboelementer.

3.3.2 BEMÆRKNING Ordet “endelig” i den endelige element metode går på opdelingen af $\bar{\Omega}$ i et *endeligt* antal delmængder.

3.3.1 Lokale endelige element rum

For et givet endelig element rum \mathcal{X}_{fe} (vi skal senere vende tilbage til en præcis definition) defineres det endelige dimensionale rum P_K som

$$P_K = \{v|_K \mid v \in \mathcal{X}_{fe}\}. \quad (3.4)$$

Som nævnt i kapitel 2 vil \mathcal{X} ofte være $\mathcal{H}^1(\Omega)$, $\mathcal{H}^2(\Omega)$ eller lignende. Vi skal nu vise en sætning, der giver betingelser, som vil sikre at $\mathcal{X}_{fe} \subset \mathcal{H}^1(\Omega)$.

3.3.3 SÆTNING Lad Ω være en åben og begrænset delmængde af \mathbb{R}^n . Antag at inklusionen $P_K \subset \mathcal{H}^1(K)$ holder for alle $K \in \mathcal{T}_h$, og at $\mathcal{X}_{fe} \subset \mathcal{C}^0(\bar{\Omega})$. Da vil

$$\mathcal{X}_{fe} \subset \mathcal{H}^1(\Omega), \quad (3.5)$$

$$\mathcal{X}_{0,fe} = \{v \in \mathcal{X}_{fe} \mid v = 0 \text{ på } \partial\Omega\} \subset \mathcal{H}_0^1(\Omega). \quad (3.6)$$

BEVIS Lad $v \in \mathcal{X}_{fe}$ være givet. Da $\bar{\Omega}$ er begrænset, og da $v \in \mathcal{C}^0(\bar{\Omega})$, vil $v \in \mathcal{L}^2(\Omega)$. Ifølge definitionen af $\mathcal{H}^1(\Omega)$ vil $v \in \mathcal{H}^1(\Omega)$, såfremt vi for $i = 1, \dots, n$ kan finde en funktion $v_i \in \mathcal{L}^2(\Omega)$, så der for alle $\phi \in \mathcal{D}_c^\infty(\Omega)$ ($\mathcal{D}_c^\infty(\Omega)$ er mængden af ∞ -gange differentiable funktioner med kompakt støtte i Ω) gælder

$$\int_{\Omega} v_i \phi \, dx = - \int_{\Omega} v \partial_i \phi \, dx. \quad (3.7)$$

Sæt nu $v_i = \partial_i(v|_K)$ for ethvert $K \in \mathcal{T}_h$, dvs. v_i er restriktionen til ethvert endeligt element K af $\partial_i v$. Da ethvert endeligt element K har en Lipschitz kontinuert rand ∂K , kan vi anvende Green's sætning til at få

$$\int_K \partial_i(v|_K) \phi \, dx = - \int_K v|_K \partial_i \phi \, dx + \int_{\partial K} v|_K \phi \nu_{i,K} \, d\gamma \quad \text{for alle } K \in \mathcal{T}_h, \quad (3.8)$$

hvor $\nu_{i,K}$ betegner den i 'te komponent i enheds normal vektoren til ∂K . Ved nu at summere over alle endelig elementer fås

$$\int_{\Omega} v_i \phi \, dx = - \int_{\Omega} v \partial_i \phi \, dx + \sum_{K \in \mathcal{T}_h} \int_{\partial K} v|_K \phi \nu_{i,K} \, d\gamma \quad (3.9)$$

Vi bemærker nu, at sidste led på højre siden i ovenstående relation er 0, thi enten vil en del af ∂K være en del af Ω 's rand Γ , og så vil $\phi = 0$, eller også vil bidraget fra to nabo elementer være nul, pga. modsat omløbsretning. Altså haves den første inklusion. Da randen Γ er Lipschitz kontinuert, følger den anden inklusion af

$$\mathcal{X}_{0,\text{fe}} = \{v \in \mathcal{X}_{\text{fe}} \mid v = 0 \text{ på } \gamma\} \subset \{v \in \mathcal{H}^1(\Omega) \mid v = 0 \text{ på } \Gamma\} = \mathcal{H}_0^1(\Omega). \quad (3.10)$$

Beviset for følgende sætning minder meget om beviset for forrige sætning. Vi vil derfor nøjes med at angive sætningen.

3.3.4 SÆTNING *Lad Ω være en åben og begrænset delmængde af \mathbb{R}^n . Antag at inklusionen $P_K \subset \mathcal{H}^2(K)$ holder for alle $K \in \mathcal{T}_h$, og at $\mathcal{X}_{\text{fe}} \subset \mathcal{C}^1(\overline{\Omega})$. Da vil*

$$\mathcal{X}_{\text{fe}} \subset H^2(\Omega), \quad (3.11)$$

$$\mathcal{X}_{0,\text{fe}} = \{v \in \mathcal{X}_{\text{fe}} \mid v = 0 \text{ på } \partial\Omega\} \subset \mathcal{H}^2(\Omega) \cap \mathcal{H}_0^1(\Omega), \quad (3.12)$$

$$\mathcal{X}_{00,\text{fe}} = \{v \in \mathcal{X}_{\text{fe}} \mid v = \partial_\nu v = 0 \text{ på } \partial\Omega\} \subset \mathcal{H}_0^2(\Omega), \quad (3.13)$$

3.4 Polynomiumsrummene P_k

Vi skal nu definere det polynomiumsrum, der skal bruges i definitionen af en klasse af endelige elementer.

3.4.1 DEFINITION *For ethvert $k \geq 0$ defineres P_k^n som rummet af polynomier med grad $\leq k$ i de n variable x_1, \dots, x_n , dvs. polynomierne $p \in P_k^n$ er afbildninger $p: \mathbb{R}^n \rightarrow \mathbb{R}$ givet ved*

$$p(x) = p(x_1, \dots, x_n) = \sum_{\sum_{i=1}^n \alpha_i \leq k} \gamma_{\alpha_1 \dots \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \quad (3.14)$$

for passende konstanter $\gamma_{\alpha_1 \dots \alpha_n}$.

3.4.2 BEMÆRKNING Vi vil som regel skrive P_k istedet for P_k^n , idet dimensionen af rummet vil fremgå af sammenhængen.

3.4.3 BEMÆRKNING Vi skal til tider benytte os af en multi-indeks notation og udtrykke polynomier $p \in P_k$ som

$$p(x) = \sum_{|\alpha| \leq k} \gamma_\alpha x^\alpha. \quad (3.15)$$

3.4.4 SÆTNING Dimensionen af polynomiumsrummet P_k^n er givet ved

$$\dim P_k^n = \binom{n+k}{k}. \quad (3.16)$$

BEVIS Det er oplagt, at $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, hvor $\alpha_1 + \cdots + \alpha_n \leq k$, $\alpha_i \geq 0$, er en basis for P_k^n . Vi skal så bestemme antallet af måder hvorpå vi kan vælge α 'erne så $\alpha_1 + \cdots + \alpha_n = i \leq k$ med $0 \leq i \leq k$. Dette svarer til valg af i elementer ud af n elementer med tilbagelægning uden hensyn til ordning. Altså er der

$$\sum_{i=0}^k \binom{n-1+i}{i} = \sum_{i=0}^k \binom{n-1+i}{n-1} = \binom{n+k}{k} \quad (3.17)$$

muligheder.

3.4.5 DEFINITION For en delmængde $A \subset \mathbb{R}^n$ sættes

$$P_k(A) = \{p|_A \mid p \in P_k\}. \quad (3.18)$$

3.4.6 BEMÆRKNING Hvis $A^\circ \neq \emptyset$, er $\dim P_k = \dim P_k(A)$.

3.5 Lagrange endelige elementer

3.5.1 DEFINITION Ved et (ikke degenereret) n -simplex forstås det konvekse hylster K af $n+1$ punkter $a_j = (a_{ij})_{i=1}^n \in \mathbb{R}^n$, også kaldet hjørner, der ligger således, at matricen

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n+1} \\ a_{21} & a_{22} & \cdots & a_{2,n+1} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n+1} \\ 1 & 1 & \cdots & 1 \end{pmatrix} \quad (3.19)$$

er invertibel.

3.5.2 BEMÆRKNING Som bekendt er K givet ved

$$K = \left\{ x = \sum_{j=1}^{n+1} \lambda_j a_j \mid 0 \leq \lambda_j \leq 1, 1 \leq j \leq n+1, \sum_{j=1}^{n+1} \lambda_j = 1 \right\}. \quad (3.20)$$

3.5.3 BEMÆRKNING Vi bemærker, at et 2-simplex er en trekant, og et 3-simplex er en tetraæde.

3.5.4 DEFINITION For ethvert tal $m \in \mathbb{N}_0$ med $0 \leq m \leq n$ defineres en m -side af et n -simplex K som et m -simplex, hvis $m+1$ hjørner også er hjørner i K . Specielt kaldes en $(n-1)$ -side for en side, og en 1-side kaldes en kant.

3.5.5 SÆTNING Et polynomium $p \in P_1$ givet ved $p : x \rightarrow \sum_{|\alpha| \leq 1} \gamma_\alpha x^\alpha$ er entydig bestemt af dets værdier i de $n+1$ hjørner a_j i et vilkårligt n -simplex K i \mathbb{R}^n .

BEVIS Polynomiet p kan skrives på formen $p(x) = \gamma_0 + \gamma_1 x_1 + \cdots + \gamma_n x_n$, hvor vi af notationsmæssige årsager skal kalde γ_0 for γ_{n+1} . Lad værdierne $p(a_j) = \mu_j$, $1 \leq j \leq n+1$ være givet. Vi skal da vise, at koefficienterne γ_j er entydig bestemt, dvs. at ligningssystemet

$$\mu_j = \gamma_1 a_{1j} + \cdots + \gamma_n a_{nj} + \gamma_{n+1}, \quad 1 \leq j \leq n+1 \quad (3.21)$$

har en entydig løsning. Sættes $\mu = (\mu_1, \dots, \mu_{n+1})^t$ og $\gamma = (\gamma_1, \dots, \gamma_{n+1})^t$ kan vi skrive ligningssystemet som

$$\mu = \mathbf{A}^t \gamma. \quad (3.22)$$

Dette system har imidlertid en entydig løsning, da K er et n -simplex.

3.5.6 BEMÆRKNING Som en konsekvens af forrige sætning ses, at hvis vi sætter $\lambda_i(a_j) = \delta_{ij}$, kan ethvert polynomium $p \in P_1$ repræsenteres som

$$p = \sum_{i=1}^{n+1} p(a_i) \lambda_i. \quad (3.23)$$

Vi kan nu definere den simpleste form for et endelig element, som vi skal kalde et n -simplex af type (1) som værende et tripel (K, P_K, Σ_K) bestående af n -simplexet K med hjørner a_j , rummet $P_K = P_1(K)$, samt mængden Σ_K af frihedsgrader som består af de parametre, der entydigt definerer en funktion i rummet P_K , dvs. $\Sigma_K = \{p(a_j) \mid 1 \leq j \leq n+1\}$.

Betegnes midtpunkterne af kanterne i et n -simplex med $a_{ij} = (a_i + a_j)/2$, kan man efter samme ide som i beviset for sætning 3.5.5 vise, at polynomier i $P_K = P_2(K)$ har repræsentationen

$$p = \sum_{i=1}^{n+1} p(a_i) \lambda_i (2\lambda_i - 1) + \sum_{\substack{i,j=1 \\ i < j}}^{n+1} p(a_{ij}) 4\lambda_i \lambda_j, \quad (3.24)$$

hvor $\lambda_k(a_{ij}) = (\delta_{ki} + \delta_{kj})/2$, $1 \leq i < j \leq n+1$, $1 \leq k \leq n+1$. Heraf fås definitionen for endnu et endeligt element, kaldet et n -simplex af type (2), bestående af K som før, rummet $P_K = P_2(K)$ samt $\Sigma_K = \{p(a_i) \mid 1 \leq i \leq n+1\} \cup \{p(a_{ij}) \mid 1 \leq i < j \leq n+1\}$.

Lad K være et n -simplex med hjørner a_1, \dots, a_{n+1} . Sættes $a_{iij} = (2a_i + a_j)/3$ for $i \neq j$ og $a_{ijk} = (a_i + a_j + a_k)/3$ for $i < j < k$ kan man (igen på tilsvarende måde som i sætning 3.5.5) vise, at polynomier i $P_3(K)$ har formen

$$\begin{aligned} \forall p \in P_3, \quad p = \sum_i \frac{\lambda_i(3\lambda_i - 1)(3\lambda_i - 2)}{2} p(a_i) \\ + \sum_{i \neq j} \frac{9\lambda_i \lambda_j (3\lambda_i - 1)}{2} p(a_{iij}) + \sum_{i < j < k} 27\lambda_i \lambda_j \lambda_k p(a_{ijk}). \end{aligned} \quad (3.25)$$

Dette giver os et n -simplex af type (3). Figurerne 3.1, 3.2 og 3.3 viser eksempler på nogle af de mest anvendte endelige elementer. For at kunne give en generel definition har vi brug for nedenstående sætning. Beviset for sætningen er ret teknisk og ikke interessant i sig selv, hvorfor vi vil undlade det.

3.5.7 SÆTNING *Lad K være et n -simplex med hjørner a_j , $1 \leq j \leq n+1$. Da vil et vilkårlig polynomium $p \in P_k$ være entydigt bestemt af dets værdier i punkterne givet*

$$\begin{aligned} L_k(K) = \left\{ x = \sum_{j=1}^{n+1} \gamma_j a_j \mid \sum_{j=1}^{n+1} \gamma_j = 1, \right. \\ \left. \gamma_j \in \left\{ 0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1 \right\}, 1 \leq j \leq n+1 \right\}. \end{aligned} \quad (3.26)$$

BEVIS Se fx [36].

3.5.8 BEMÆRKNING Mængden $L_k(K)$ kaldes for hovedgitteret af orden k for n -simplexet K .

Figur 3.1: Trekant af type (1) og tetraede af type (1)



Figur 3.2: Trekant af type (2) og tetraede af type (2)



Figur 3.3: Trekant af type (3) og tetraede af type (3)

Vi kan nu give en generel definition af en bestemt klasse af endelige elementer også kaldet n -simplexer af type (k) .

3.5.9 DEFINITION Ved et n -simplex af type (k) forstås et triple (K, P_K, Σ_K) bestående af et n -simplex K med hjørner a_j , $1 \leq j \leq n+1$, polynomiumsrummet $P_K = P_k(K)$ samt mængden af frihedsgrader $\Sigma_K = \{p(x) \mid x \in L_k(K)\}$.

3.5.10 DEFINITION Endelige elementer, hvor frihedsgraderne udelukkende er punktværdier, kaldes for Langrange endelige elementer.

3.5.1 Triangulering af $\bar{\Omega}$ med simplexer

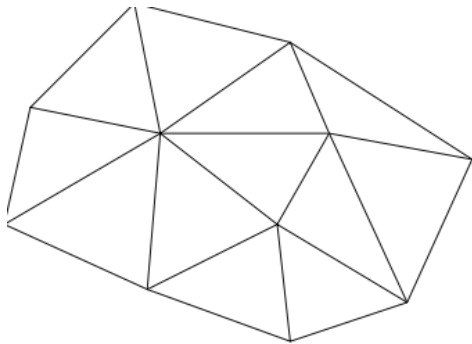
Vi skal nu se nærmere på, hvorledes man konstruerer en triangulering af Ω med de endelige elementer, der er blevet defineret i afsnit 3.5. Som nævnt i afsnit 3.3 vil vi tilføje et femte krav til de fire fra definition 3.3.1. Da ikke-degenererede n -simplexer er afsluttede, har et ikke tomt indre og en Lipschitz kontinuert rand, vil betingelserne $\mathcal{T}_h2.$ og $\mathcal{T}_h4.$ altid være opfyldte. Vi behøver derfor kun at koncentrere os om at opdele Ω , så betingelserne $\mathcal{T}_h1.$ og $\mathcal{T}_h3.$ er opfyldt. Vi skal nu stille endnu et krav til en triangulering af Ω . Dette krav skal sikre, at $\mathcal{X}_{fe} \subset \mathcal{C}^0(\bar{\Omega})$, og senere hen at $\mathcal{X}_{fe} \subset \mathcal{C}^1(\bar{\Omega})$.

3.5.11 DEFINITION For at sikre en fornuftig definition af endelige element rum har vi brug for følgende ekstra krav til trianguleringen af $\bar{\Omega}$

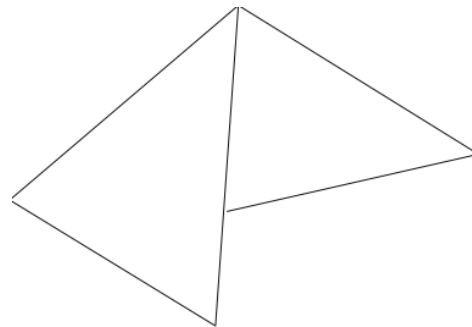
$\mathcal{T}_h5.$ En vilkårlig side i et vilkårlig n -simplex K_1 i en given triangulering er enten en delmængde af randen $\partial\Omega$ eller en side i et andet n -simplex K_2 .

3.5.12 BEMÆRKNING Betingelsen \mathcal{T}_h5 er teoretisk ideel, idet den er med til at sikre kontinuitet i den endelige element løsning over randene mellem de enkelte elementer. I praktiske situationer er den ikke nødvendig, for at den endelige element metode fungerer. Måske er man end ikke interesseret i en løsning, der er kontinuert over randene. Man kan også få kontinuitet over kanterne uden denne betingelse, hvis blot man er mere påpasselig med at vælge polynomiumsrummene P_K i de enkelte elementer K . I figurerne 3.4 og 3.5 ses eksempler på en tilladt triangulering og en ikke-tilladt triangulering.

3.5.13 BEMÆRKNING Har to n -simplexer K_1 og K_2 en fælles side, kaldes K_1 og K_2 naboelementer (eng. adjacent).



Figur 3.4: Eksempel på tilladt triangulering



Figur 3.5: Eksempel på ikke-tilladt triangulering

3.6 Endelige element rum hørende til n -simplexer af type (k)

Vi er nu i stand til at give en definition af endelige element rum hørende til n -simplexer af type (k) .

3.6.1 DEFINITION Lad \mathcal{T}_h være en triangulering af $\overline{\Omega}$. Ved et endeligt element rum \mathcal{X}_{fe} hørende til n -simplexer af type (k) forstås en mængde af funktioner $v : \overline{\Omega} \rightarrow \mathbb{R}$ opfyldende følgende krav

1. restriktionen af v til K tilhører $P_k(K)$ for alle $K \in \mathcal{T}_h$.
2. funktionen v er entydigt bestemt ved dens værdier i punkterne

$$\Sigma_h = \{v(x) \mid x \in \cup_{K \in \mathcal{T}_h} L_k(K)\}. \quad (3.27)$$

Mængden Σ_h kaldes mængden af frihedsgrader for det endelige element rum \mathcal{X}_{fe} .

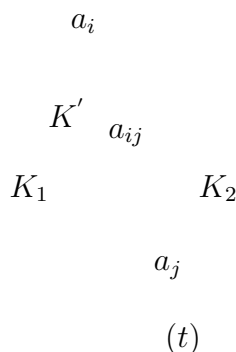
3.6.2 BEMÆRKNING På grund af betingelse $\mathcal{T}_h 5.$ er der ingen flertydighed i definitionen af frihedsgraderne.

3.6.3 SÆTNING Lad \mathcal{X}_{fe} være det endelige element rum hørende til n -simplexer af type (k) . Da vil

$$\mathcal{X}_{\text{fe}} \subset \mathcal{C}^0(\overline{\Omega}) \cap \mathcal{H}^1(\Omega). \quad (3.28)$$

BEVIS Lad K_1 og K_2 være to n -simplexer af type (k) med fælles side K' , hvor K_1 er bestemt af punkterne a_1, \dots, a_{n+1} og K_2 af punkterne b_1, \dots, b_{n+1} . Da

K_1 og K_2 har en fælles side K' vil n af punkterne a_1, \dots, a_{n+1} være identiske med n af punkterne b_1, \dots, b_{n+1} . Disse n sammenfaldende punkter er præcis de punkter, der udspænder $(n-1)$ -simplexet K' . Vi bemærker, at $L_k(K') \subset L_k(K_1)$ og $L_k(K') \subset L_k(K_2)$. For en given funktion $v \in \mathcal{X}_{\text{fe}}$ betragtes nu funktionerne $v|_{K_1}$ og $v|_{K_2}$ på den fælles side K' . Vi indfører koordinatsystemet $t = (t_1, \dots, t_{n-1})$ på K' betragtet som hyperplan. Betragtet som funktion af t er $v|_{K_1}$ og $v|_{K_2}$ k 'te grads polynomier på K' , der stemmer overens i de $\binom{n+k}{k}$ punkter i $L_k(K')$. Altså må $v|_{K_1}$ og $v|_{K_2}$ være identiske på K' ifølge sætning 3.5.7, og dermed $\mathcal{X}_{\text{fe}} \subset \mathcal{C}^0(\bar{\Omega})$. Inklusionen $\mathcal{X}_{\text{fe}} \subset \mathcal{H}^1(\Omega)$ følger af sætning 3.3.3. Figur 3.6 illustrerer tilfældet, hvor $n = 2$ og trekanter af type (2)



Figur 3.6: Illustration for situationen $n = 2$ og trekant af type (2)

3.7 Basisfunktioner for endelige element rum

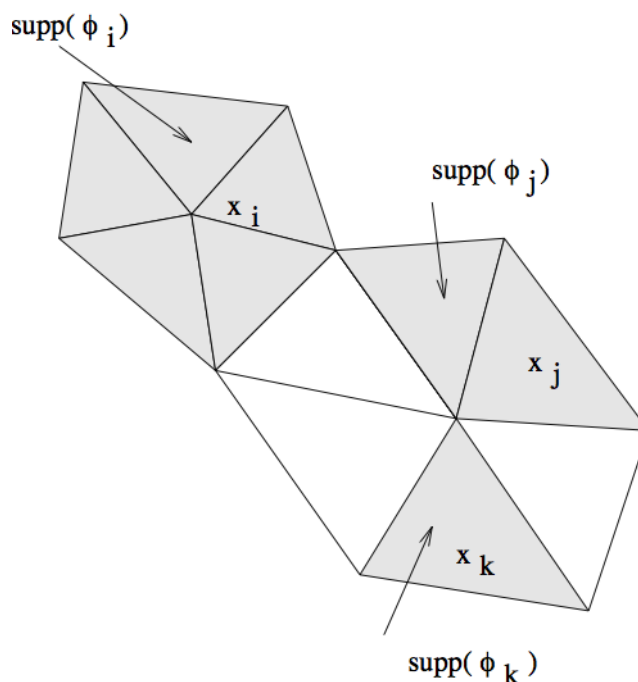
I afsnit 3.2 pointerede vi vigtigheden af, at basisfunktionerne for endelige element rum havde så “lille” støtte som muligt. For endelige element rum hørende til n -simplexer af type (k) er mængden af frihedsgrader Σ_h givet ved

$$\Sigma_h = \{v(x_j) \mid 1 \leq j \leq M\}, \quad (3.29)$$

hvor punkterne x_j , $j = 1, \dots, m$ er punkterne i $\cup_{K \in \mathcal{T}_h} L_k(K)$. Det er oplagt, at funktionerne ϕ_i , $1 \leq i \leq M$ defineret ved betingelserne

$$\phi_i \in \mathcal{X}_{\text{fe}} \quad \text{og} \quad \phi_i(x_j) = \delta_{ij}, \quad 1 \leq i, j \leq M, \quad (3.30)$$

udgør en basis for \mathcal{X}_{fe} , og at støtten for disse funktioner er “lille”. Figur 3.7 indeholder et eksempel på støtten for basisfunktioner hørende til n -simplexer af type (2).



Figur 3.7: Eksempel på støtte for basisfunktioner hørende til n -simplexer af type (2)

3.8 Hermite endelige elementer

Vi skal nu se et eksempel på et endelig element, hvor frihedsgraderne ikke udelukkende er punktværdier, men også retningsafledede.

3.8.1 SÆTNING *Lad K være en trekant med hjørner a_i , $i = 1, 2, 3$, hvor sidernes midtpunkter betegnes med $a_{ij} = (a_i + a_j)/2$, $1 \leq i < j \leq 3$. Da vil ethvert polynomium af grad 5 være entydig bestemt ved dets værdier i mængden af de 21 frihedsgrader*

$$\Sigma_K = \left\{ \partial^\alpha p(a_i), |\alpha| \leq 2, 1 \leq i \leq 3; \frac{\partial p(a_{ij})}{\partial \mathbf{n}}, 1 \leq i < j \leq 3 \right\}, \quad (3.31)$$

hvor $\frac{\partial}{\partial \mathbf{n}}$ betegner den normal afledede langs randen af K .

BEVIS Lad $p \in P_5(K)$. Da dimensionen af $P_5(K)$ er lig med antallet af frihedsgrader ($= 21$), er det nok at vise, at hvis værdien i alle frihedsgraderne er 0, da vil $p \equiv 0$. Hvis s betegner retningen af siden a_2a_3 vil

$$p(a_i) = \frac{\partial p}{\partial s}(a_i) = \frac{\partial^2 p}{\partial s^2}(a_i) = 0, \quad i = 2, 3. \quad (3.32)$$

Da p er et polynomium af grad højst 5 på siden a_2a_3 , må $p \equiv 0$ på a_2a_3 . På a_2a_3 er $\frac{\partial p}{\partial \mathbf{n}}$ et polynomium af grad højst 4 og

$$\frac{\partial p}{\partial \mathbf{n}}(a_{23}) = \frac{\partial p}{\partial \mathbf{n}}(a_i) = \frac{\partial}{\partial s} \left(\frac{\partial p}{\partial \mathbf{n}} \right) (a_i) = 0, \quad i = 2, 3 \quad (3.33)$$

hvilket kun er muligt, hvis $\frac{\partial p}{\partial \mathbf{n}} = 0$ på a_2a_3 . Altså er både p og $\frac{\partial p}{\partial \mathbf{n}}$ identisk 0 på a_2a_3 , hvilket betyder, at vi kan faktorisere $(\lambda_1(x))^2$ ud af p . Vi har derfor

$$p(x) = (\lambda_1(x))^2 p_3(x), \quad (3.34)$$

hvor $p_3 \in P_3(K)$. På tilsvarende vis ses, at vi kan faktorisere $(\lambda_2(x))^2$ og $(\lambda_3(x))^2$ ud, hvorfor

$$p = \gamma \lambda_1^2 \lambda_2^2 \lambda_3^2, \quad (3.35)$$

hvor γ er en passende konstant. Da $p \in P_5(K)$ er denne faktorisering kun mulig såfremt $\gamma = 0$, og dermed er $p \equiv 0$ på K .

3.8.2 DEFINITION *Det endelige element $(K, P_5(K), \Sigma_K)$, hvor K og Σ_K er som i sætning 3.8.1, kaldes også Argyris' trekant, se figur 3.8.*

3.8.3 SÆTNING *Lad \mathcal{X}_{fe} være det endelig element rum associeret med Argyris' trekant. Da vil*

$$\mathcal{X}_{fe} \subset \mathcal{C}^1(\bar{\Omega}) \cap \mathcal{H}^2(\Omega). \quad (3.36)$$

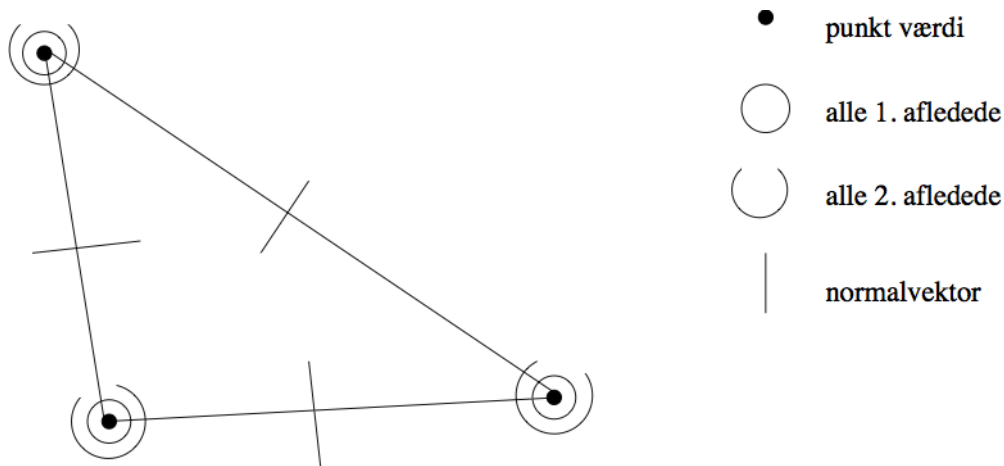
BEVIS Ifølge sætning 3.3.4 er det nok at vise $X_h \subset \mathcal{C}^1(\bar{\Omega})$. Lad K_1 og K_2 være to nabo trekanter med fælles side K' og lad p_i , $i = 1, 2$ være to funktioner i \mathcal{X}_{fe} med

$$\partial^\alpha p_1 = \partial^\alpha p_2 \text{ i } K' \text{'s endepunkter, } |\alpha| \leq 2, \quad (3.37)$$

$$\frac{\partial p_1}{\partial \mathbf{n}} = \frac{\partial p_2}{\partial \mathbf{n}} \text{ i } K' \text{'s midtpunkt,} \quad (3.38)$$

hvor $\frac{\partial}{\partial \mathbf{n}}$ betegner differentiation i normal retningen til K' . Det følger nu af (3.32) og (3.33) fra sætning 3.8.1, at vi for differensen $p = p_1 - p_2$ har

$$p \equiv \frac{\partial p}{\partial n} \equiv 0 \text{ på } K', \quad (3.39)$$



Figur 3.8: Argyris' trekant

og da $p \equiv 0$ på K' må

$$\frac{\partial p}{\partial s} \equiv 0, \quad (3.40)$$

hvor $\frac{\partial}{\partial s}$ betegner differentiation i retningen tangentielt til K' . Det følger nu af (3.39) og (3.40), at funktionerne p_i samt deres første afledede varierer kontinuert over K' , så $\mathcal{X}_{fe} \subset \mathcal{C}^1(\bar{\Omega})$.

3.8.4 DEFINITION *Endelige elementer, hvor frihedsgraderne består af punktværdier og retningsafledede, kaldes Hermite endelige elementer.*

3.9 Generelle endelige elementer

For at kunne give en generel definition af endelige elementer har vi brug for følgende definition, der kan ses som en generalisering af begrebet frihedsgrader.

3.9.1 DEFINITION *Lad F være en mængde af funktioner med reelle værdier defineret på en mængde A , og lad Θ være en mængde af lineært uafhængige lineære afbildninger θ_i , $1 \leq i \leq N$ defineret på F . Θ kaldes da F -unisolvent, hvis der for givne reelle tal α_i , $1 \leq i \leq N$ findes en entydig funktion $f \in F$, som opfylder*

$$\theta_i(f) = \alpha_i, \quad 1 \leq i \leq N. \quad (3.41)$$

3.9.2 DEFINITION *Et endelig element i \mathbb{R}^n er et tripel (K, P, Σ) , hvor*

1. K er en afsluttet delmængde af \mathbb{R}^n med et ikke tomt indre og en Lipschitz kontinuert rand.
2. P er en mængde af funktioner med reelle værdier defineret på K .
3. Σ er en P -unisolvent mængde af lineært uafhængige, lineære afbildninger ϕ_i ,
 $1 \leq i \leq N$ defineret på P .

3.9.3 BEMÆRKNING På grund af betingelse (3) i definition 3.9.2 findes der funktioner p_i , $1 \leq i \leq N$, som opfylder

$$\phi_j(p_i) = \delta_{ij}, \quad 1 \leq j \leq N, \quad (3.42)$$

så polynomier $p \in P$ kan skrives på formen

$$p = \sum_{i=1}^N \phi_i(p) p_i. \quad (3.43)$$

Altså er funktionsrummet P endelig dimensionalt, og $\dim P = N$.

3.9.4 DEFINITION De lineære afbildninger ϕ_i , $1 \leq i \leq N$ kaldes frihedsgraderne for det endelige element (K, P, Σ) , og funktionerne p_i , $1 \leq i \leq N$ kaldes basisfunktionerne for det endelige element.

Hvis (K, P, Σ) er et endeligt element, skal vi undertiden referere til K som et endeligt element, hvis det af sammenhængen fremgår, hvad P og Σ er, eller hvis P og Σ er uden betydning.

3.9.1 Typer af frihedsgrader

Indtil videre har vi betragtet eksempler på endelige elementer, hvor frihedsgraderne har været punktværdier eller retningsafledede af typerne

$$p(a_i^0), \quad 1 \leq i \leq N_0, \quad (3.44)$$

$$Dp(a_i^1) \xi_{ik}^1, \quad 1 \leq i \leq N_1, \quad (3.45)$$

$$Dp(a_i^2)^2(\xi_{ik}^2, \xi_{il}^2), \quad 1 \leq i \leq N_2, \quad (3.46)$$

hvor punkterne a_i^r , $r = 0, 1, 2$ tilhører det endelige element. Punkterne a_i^r , $r = 0, 1, 2$ kaldes for knudepunkterne for det endelige element og betegnes også med $\mathcal{N}_h = \{a_i^r\}$. Man kan selvfølgelig også benytte partielle afledede af højere orden som frihedsgrader, men det ses sjældent i praksis. Endelig kan man også forestille sig frihedsgrader, der ikke er knyttet til punkter, se fx [19, s. 211].

3.10 Affine familier af endelige elementer

I dette afsnit skal vi præsentere en ide, der har afgørende indflydelse i praktiske situationer.

3.10.1 DEFINITION Lad (K, P, Σ) og $(\hat{K}, \hat{P}, \hat{\Sigma})$ være to endelige elementer, hvor frihedsgraderne har formen (3.44-3.46). (K, P, Σ) og $(\hat{K}, \hat{P}, \hat{\Sigma})$ kaldes da *affin ækvivalente*, såfremt der findes en invertibel affin afbildning $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ givet ved

$$F(\hat{x}) = \mathbf{B}\hat{x} + b, \quad (3.47)$$

hvor $\mathbf{B} \in \mathbb{R}_n^n$ og $b \in \mathbb{R}^n$, og følgende relationer er opfyldt

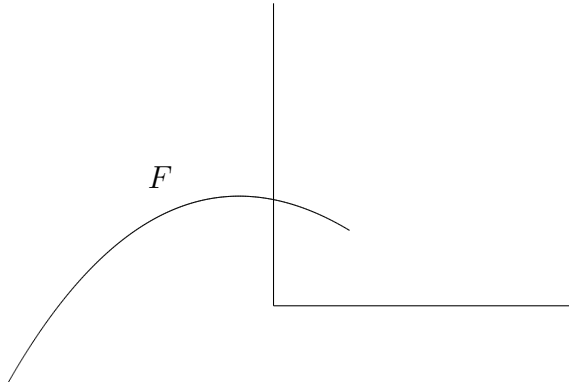
$$K = F(\hat{K}), \quad (3.48)$$

$$P = \{p : K \rightarrow \mathbb{R} \mid p = \hat{p} \cdot F^{-1}, \hat{p} \in \hat{P}\}, \quad (3.49)$$

$$a_i^r = F(\hat{a}_i^r), \quad r = 0, 1, 2, \quad (3.50)$$

$$\xi_{ik}^1 = B\hat{\xi}_{ik}^1, \quad \xi_{ik}^2 = B\hat{\xi}_{ik}^2, \quad \xi_{il}^2 = B\hat{\xi}_{il}^2, \quad (3.51)$$

hvor a_i^r hhv. \hat{a}_i^r og vektorerne ξ_{ik}^1, ξ_{ik}^2 og ξ_{il}^2 hhv. $\hat{\xi}_{ik}^1, \hat{\xi}_{ik}^2$, og $\hat{\xi}_{il}^2$ er som i (3.44-3.46).



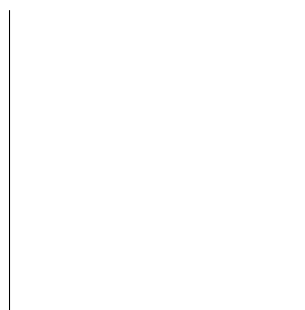
Figur 3.9: Eksempel på affin afbildning, hvor et element afbildes på et typisk reference element

3.10.2 DEFINITION En familie af endelige elementer (K, P_K, Σ_K) kaldes en *affin familie*, hvis alle dens endelige elementer er affin ækvivalente med det samme endelige element (K, P, Σ) . (K, P, Σ) kaldes da for et *reference endelig element* for familien.

3.10.3 BEMÆRKNING Definitionen af affine familier af endelige elementer kræver ikke, at reference endelig elementet tilhører familien.

Hvis man har med en affin familie af n -simplexer af type (k) at gøre, vælges K i reference endelig elementet oftets som enheds n -simplexet, se figurerne 3.10 og 3.11.

$$\begin{aligned} a_1 &= (1, 0, \dots, 0) \\ a_2 &= (0, 1, 0, \dots, 0) \\ &\vdots \\ &\vdots \\ a_n &= (0, \dots, 0, 1) \\ a_{n+1} &= (0, 0, \dots, 0) \end{aligned}$$



Figur 3.10: Hjørner i et typisk reference n -simplex

Figur 3.11: Typisk enheds 2-simplex

Konceptet med affine familier af endelige elementer er essentielt i praktiske situationer, bl.a. af følgende årsager:

1. I praktiske beregningssituationer vil alle udregningerne af koefficienterne til det lineære ligningssystem (1.12) blive udført ved hjælp af et reference endelig element. Ofte er koefficienterne i (1.12) udtrykt ved integraler, der skal evalueres numerisk. Har man med en affin familie af endelige elementer af gøre, kan man afbilde de endelige elementer på reference endelige elementet og derefter bruge samme kvadraturregel.
2. For affine familier af endelige elementer findes en “pæn” interpolations-teori, se fx [19, afsnit 3.1]

3.11 Generelle endelige element rum

Det er meget svært at give en præcis beskrivelse af konstruktionen af endelige element rum for en vilkårlig familie af endelige elementer. Dette skyldes vanskeligheder med at

1. give en præcis definition af “sider” i vilkårlige endelige elementer.
2. angive betingelser for frihedsgrader bestemt af retningsaffledede beliggende på en fælles “side” K' for to nabo endelige elementer K_1 og K_2 , således at en funktion i det endelige element rum restringeret til K_1 hhv. K_2 har samme værdi i frihedsgraderne beliggende på K' .

For at undgå ovennævnte vanskeligheder skal vi betragte situationen, hvor de endelige elementer er polygoner. Dermed bliver $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} K$ også polygonal. Vi skal også antage, at frihedsgraderne er af Lagrange typen, samt at hvert polygon i trianguleringen \mathcal{T}_h har et ikke tomt indre, og at det indre af polygonerne \mathcal{T}_h er parvis disjunkte. Dette sikrer, at betingelserne $\mathcal{T}_h 1. - \mathcal{T}_h 4.$ er opfyldt. (Vi husker på, at et polygonalt domæne altid har en Lipschitz kontinuert rand og er afsluttet.)

3.11.1 DEFINITION *En delmængde K' af randen af et polygonalt endelige element K kaldes en side af K , såfremt K' er en maksimalt sammenhængende delmængde af en affin hyperplan \mathcal{P} i \mathbb{R}^n med et ikke tomt indre (relativt til \mathcal{P}).*

3.11.2 DEFINITION *For at undgå problemer med entydigheden af elementerne i endelige element rum skal vi (i analog til definition 3.5.11) antage at triangulering \mathcal{T}_h af $\bar{\Omega}$ opfylder:*

$\mathcal{T}_h 5.$ *En vilkårlig side i et vilkårligt polygonalt endelige element K_1 er enten en side i et andet endeligt element K_2 , eller en delmængde af randen.*

Såfremt to endelige elementer har en fælles side, kaldes de nabo endelig elementer. For at sikre entydighed af elementerne i de endelige element rum, skal vi kræve, at frihedsgraderne $\Sigma_{K_j} = \{p(a_i^j) | 1 \leq i \leq N_j\}$ for to nabo endelige elementer $(K_j, P_{K_j}, \Sigma_{K_j})$, $j = 1, 2$ opfylder

$$\left(\bigcup_{i=1}^{N_1} \{a_i^1\}\right) \cap K_2 = \left(\bigcup_{i=1}^{N_2} \{a_i^2\}\right) \cap K_1. \quad (3.52)$$

3.11.3 DEFINITION *For en familie af endelige elementer (K, P_K, Σ_K) sættes*

$$\mathcal{N}_h = \cup_{K \in \mathcal{T}_h} \mathcal{N}_K \quad (3.53)$$

hvor \mathcal{N}_K betegner mængden af knudepunkter for $K \in \mathcal{T}_h$. For $b \in \mathcal{N}_h$ vil vi bruge betegnelsen K_λ , $\lambda \in \Lambda(b)$ om de endelige elementer, hvor b er et knudepunkt.

Vi er nu i stand til at give en definition af et endeligt element rum hørende til de endelige elementer (K, P_K, Σ_K) .

3.11.4 DEFINITION *Det endelige element rum \mathcal{X}_{fe} hørende til de endelige elementer (K, P_K, Σ_K) defineres som underrummet*

$$\mathcal{X}_{\text{fe}} = \left\{ v = (v_K)_{K \in \mathcal{T}_h} \in \prod_{K \in \mathcal{T}_h} P_K \mid \forall b \in \mathcal{N}_h, \forall \lambda, \mu \in \Lambda(b), v_{K_\lambda}(b) = v_{K_\mu}(b) \right\} \quad (3.54)$$

af produktrummet $\prod_{K \in \mathcal{T}_h} P_K$.

3.11.5 BEMÆRKNING Vi ser, at elementer $v \in \mathcal{X}_{\text{fe}}$ er entydig bestemt af værdien i punkterne

$$\Sigma_h = \{v(b) \mid b \in \mathcal{N}_h\} \quad (3.55)$$

som også kaldes mængden af frihedsgrader for det endelige element rum.

3.11.6 BEMÆRKNING Elementerne i \mathcal{X}_{fe} er generelt ikke funktioner defineret på $\bar{\Omega}$, idet elementerne ikke nødvendigvis har en entydig definition på en, for to nabo endelige elementer, fælles side. Det følger dog at definition 3.11.4, at elementer i \mathcal{X}_{fe} er “kontinuerte” i knudepunkter fælles for to nabo endelige elementer.

Kapitel 4

Polynomiumsinterpolation

Den endelige element løsning er som bekendt en linearkombination af en række polynomier fra et vist løsningsrum. Når man skal udføre fejlestimering, er det derfor vigtigt at vide, hvor godt man kan approksimere funktioner med polynomier fra det valgte løsningsrum. Ofte er fejlestimatorer og dermed polynomiumsinterpolation drivkraften bag netinddelings algoritmer. I næste kapitel skal vi studere en netinddelings algoritme, der gør brug af en bestemt fejlligning, som vi derfor skal verificere i dette kapitel. Det skal bemærkes, at vi ikke er interesseret i den bedste polynomiums interpolant, men derimod mere interesseret i at opnå en fejlligning på en bestemt form. Interpolationsteori, hvor man forsøger at få den bedst mulige approksimation, er fx behandlet i [17, 18]. Generel interpolationsteori for vilkårlige Sobolev rum kan fx findes i [19].

4.1 A priori fejlligningen

Som nævnt ovenfor er vi interesseret i en ganske bestemt fejlligning. Mere præcist ønsker vi at udtrykke interpolations fejlen på formen

$$\|e\|^q = \sum_{i=1}^k \phi_i h_x^{\gamma_{1,i}} h_y^{\gamma_{2,i}} + \text{h. o. l.} \quad (4.1)$$

Vi skal studere tilfældet, hvor det to dimensionale fysiske domæne er blevet afbildet bijektivt på et passende antal kvadrater, og hvor de endelige elementer (i kvadraterne) består af akseparallelle rektangler med sidelængde h_x og h_y . Funktionerne ϕ_i må gerne afhænge af den funktion, vi ønsker at interpolere, men ikke af de endelige elementers geometri. Konstanterne q , $\gamma_{j,i}$ og k er passende norm afhængige konstanter, og h. o. l. betegner højere ordens led.

Lad $\square = [a, a + h_x] \times [b, b + h_y]$ være et rektangel i \mathbb{R}^2 . Vi skal da betragte approksimationer til funktioner $u : \square \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ med polynomier fra produktrummet Q_{pq} givet ved

$$Q_{pq} = \text{span}\{\varphi_{ij} : (x, y) \rightarrow x^i y^j \in \mathbb{R} \mid 0 \leq i \leq p, 0 \leq j \leq q\}. \quad (4.2)$$

Vi skal antage, at u er tilstrækkelig differentiabel til, at vi kan foretage en Taylor udvikling af $(p + 2) \times (q + 2)$ 'te orden med restled i ethvert punkt $(x_0, y_0) \in \square$, dvs.

$$u(x, y) = \sum_{i=0}^{p+2} \sum_{j=0}^{q+2} \phi_{ij}(x, y) + R_{p+2, q+2}(x, y), \quad (4.3)$$

hvor

$$\phi_{ij}(x, y) = \frac{1}{i!j!} \frac{\partial^{i+j} u(x_0, y_0)}{\partial x^i \partial y^j} (x - x_0)^i (y - y_0)^j, \quad (4.4)$$

$$R_{p+2, q+2}(x, y) = \sum_{k=0}^{q+3} \mathcal{O}_{h_x, h_y \rightarrow 0}(h_x^{p+3} h_y^k) + \sum_{k=0}^{p+3} \mathcal{O}_{h_x, h_y \rightarrow 0}(h_x^k h_y^{p+3}). \quad (4.5)$$

Notationen i (4.5) indikerer, at vi hovedsageligt er interesseret i det asymptotiske tilfælde, hvor $h_x \rightarrow 0$ og $h_y \rightarrow 0$.

4.2 Lokal Taylor polynomiums interpolation i et rektangel i \mathbb{R}^2

Med notationen fra forrige afsnit skal vi definere Taylor polynomiums interpolanten af u af grad p i x og af grad q i y som

$$T_{pq}\{u\}(x, y) = \sum_{i=0}^p \sum_{j=0}^q \phi_{ij}(x, y). \quad (4.6)$$

Fejlen, vi begår ved at approksimere u med $T_{pq}\{u\}$, defineres som

$$e_{pq}^T = u - T_{pq}\{u\}. \quad (4.7)$$

I figur 4.1 er en grafisk illustration af u opsplittet i $T_{pq}\{u\}$ og e_{pq}^T . For at undgå udartede rektangler skal vi antage, at der er en sammenhæng mellem sidebredden og -længden. Mere formelt skal vi forudsætte, at der findes en global konstant H , så der for alle rektangler i netinddelingen findes et $s \in [1/H, H]$, så

$$h_x = h, \quad h_y = sh. \quad (4.8)$$

ϕ_{00}	\cdots	ϕ_{p0}	$\phi_{p+1,0}$	$\phi_{p+2,0}$	$\mathcal{O}(h_x^{p+3}h_y^0)$
\vdots		\vdots	\vdots	\vdots	\vdots
ϕ_{0q}	\cdots	ϕ_{pq}	$\phi_{p+1,q}$	$\phi_{p+2,q}$	$\mathcal{O}(h_x^{p+3}h_y^q)$
$\phi_{0,q+1}$	\cdots	$\phi_{p,q+1}$	$\phi_{p+1,q+1}$	$\phi_{p+2,q+1}$	$\mathcal{O}(h_x^{p+3}h_y^{q+1})$
$\phi_{0,q+2}$	\cdots	$\phi_{p,q+2}$	$\phi_{p+1,q+2}$	$\phi_{p+2,q+2}$	$\mathcal{O}(h_x^{p+3}h_y^{q+2})$
$\mathcal{O}(h_x^0h_y^{q+3})$	\cdots	$\mathcal{O}(h_x^p h_y^{q+3})$	$\mathcal{O}(h_x^{p+1}h_y^{q+3})$	$\mathcal{O}(h_x^{p+2}h_y^{q+3})$	$\mathcal{O}(h_x^{p+3}h_y^{q+3})$

$T_{pq}\{u\}$	
	e_{pq}^T

Figur 4.1: Grafisk illustration af u opsplittet i $T_{pq}\{u\}$ og e_{pq}^T

Vi skal senere udnytte, at dette krav medfører $h_x \approx h_y$ for $h_x \rightarrow 0$ og $h_y \rightarrow 0$. Ved nu at bemærke, at $\phi_{ij} = \mathcal{O}_{h_x, h_y \rightarrow 0}(h_x^i h_y^j) = \mathcal{O}_{h \rightarrow 0}(h^{i+j})$ for vilkårlige ikke-negative tal i og j fås følgende relation for e_{pq}^T

$$e_{pq}^T = \begin{cases} \{\phi_{p+1,0} + \phi_{0,q+1}\} + \text{h. o. l.} & \text{for } p = q, \\ \{\phi_{p+1,0}\} + \{\phi_{p+2,0} + \phi_{p+1,1} + \phi_{0,q+1}\} + \text{h. o. l.} & \text{for } p = q - 1, \\ \{\phi_{0,q+1}\} + \{\phi_{0,q+2} + \phi_{1,q+1} + \phi_{p+1,0}\} + \text{h. o. l.} & \text{for } p = q + 1, \\ \{\phi_{p+1,0}\} + \{\phi_{p+2,0} + \phi_{p+1,1}\} + \text{h. o. l.} & \text{for } p < q - 1, \\ \{\phi_{0,q+1}\} + \{\phi_{0,q+2} + \phi_{1,q+1}\} + \text{h. o. l.} & \text{for } p > q + 1. \end{cases} \quad (4.9)$$

Vi har her brugt betegnelsen h. o. l. om højere ordens led i h (eller i h_x og h_y). For at forklare betydningen af de understregede led i (4.9) skal vi betragte følgende eksempel:

4.2.1 EKSEMPEL Betragt funktionen $u(x, y) = \cos Kx \cos y$ for et passende $K \gg 1$, og lad $\square = [0, h_x] \times [0, h_y]$, $(x_0, y_0) = (h_x/2, h_y/2)$ samt $p = q + 1 = 1$. Vi har valgt $p > q$, da u varierer hurtigere i x -retningen end i y -retningen. Fra (4.9) fås

$$e_{pq}^T = \{\phi_{0,q+1}\} + \{\phi_{0,q+2} + \phi_{1,q+1} + \phi_{p+1,0}\} + \text{h. o. l.}, \quad (4.10)$$

hvor vi i dette tilfælde finder

$$\phi_{0,q+1} = -\cos(Kh_x/2) \sin(h_y/2)(y - h_y/2), \quad (4.11)$$

$$\phi_{0,q+2} = -\frac{1}{2} \cos(Kh_x/2) \cos(h_y/2)(y - h_y/2)^2, \quad (4.12)$$

$$\phi_{1,q+1} = K \sin(Kh_x/2) \sin(h_y/2)(x - h_x/2)(y - h_y/2), \quad (4.13)$$

$$\phi_{p+1,0} = -\frac{K^2}{2} \cos(Kh_x/2) \cos(h_y/2)(x - h_x/2)^2, \quad (4.14)$$

Udnyttes nu, at $|\sin t| \leq 1$, $|\cos t| \leq 1$, $|x - h_x/2| \leq h_x/2$ og $|y - h_y/2| \leq h_y/2$, fås

$$|e_{pq}^T| \leq \frac{1}{2}h_y + \frac{1}{2 \cdot 4}h_y^2 + \frac{K}{2 \cdot 2}h_x h_y + \frac{K^2}{2 \cdot 4}h_x^2 + \text{h. o. l.} \quad (4.15)$$

Som nævnt ovenfor er vores mål at have $h_x \approx h_y$. Ved at indsætte $h_x = h$ og $h_y = sh$ i (4.15) ses, at

$$|e_{pq}^T| \leq \frac{1}{2}h + \frac{s^2}{2 \cdot 4}h^2 + \frac{Ks}{2 \cdot 2}h^2 + \frac{K^2}{2 \cdot 4}h^2 + \text{h. o. l.} \quad (4.16)$$

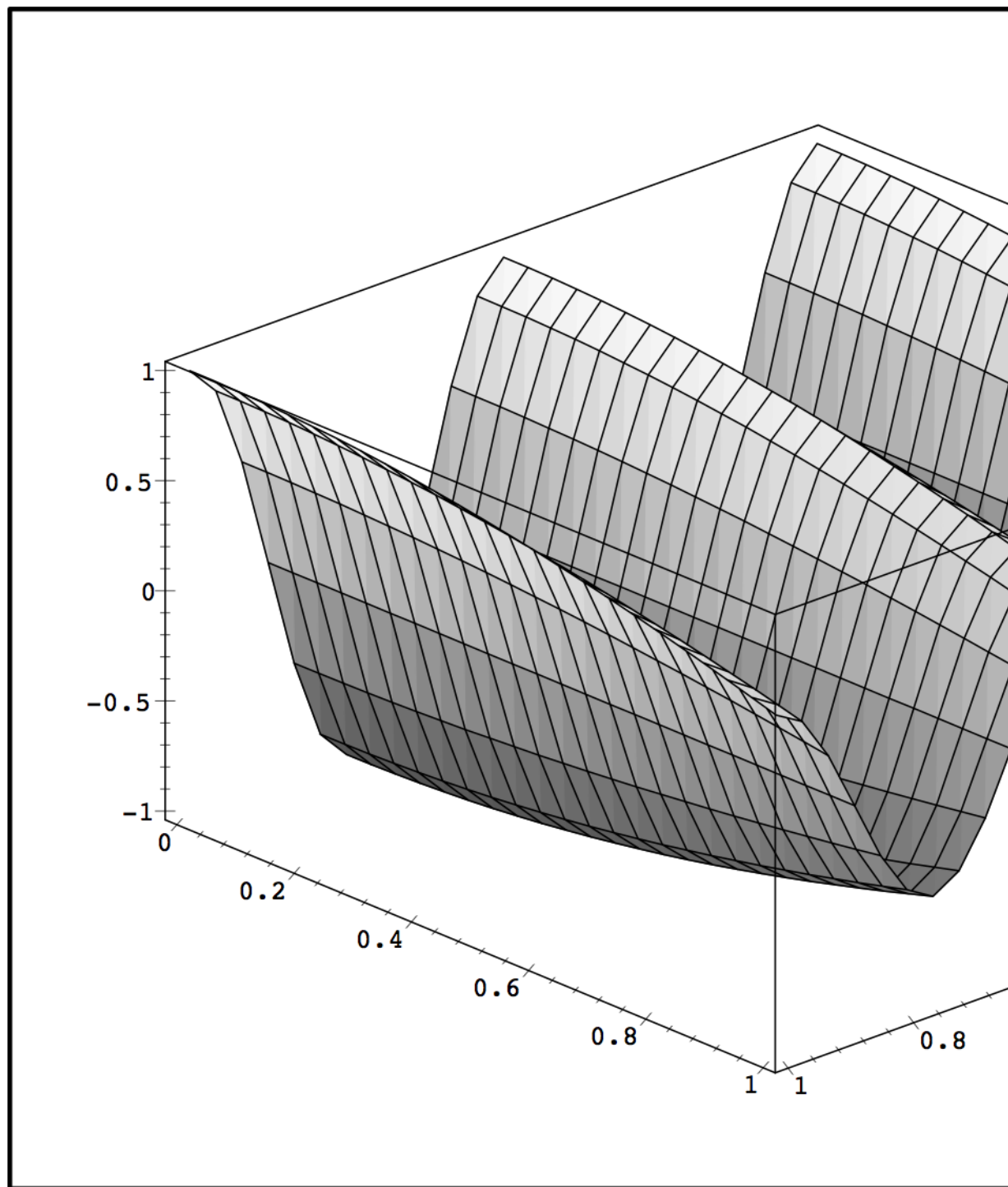
I denne ligning er det klart, at leddet af størrelsesorden h dominerer for $h \rightarrow 0$, og idet $K \gg 1$ vil det sidste led dominere de andre h^2 led. Eller sagt med andre ord: De dominerede led i e_{pq}^T er $\phi_{0,q+1}$ og $\phi_{p+1,0}$, dvs. de understregede led.

For et eksempel i tilfældet $p = q - 1 = 1$ henvises til [29]. Disse eksempler giver os følgende formodning.

4.2.2 KONJEKTUR *Vælges p og q forskelligt på grund af en hurtigere variation i u i enten x -retningen eller i y -retningen, og har vi som overordnet mål at opnå $h_x = h \approx h_y = sh$, da vil de dominerende led i e_{pq}^T for $h \rightarrow 0$, være de understregede led i ligningen (4.9).*

Ved kun medtage de dominerende led fra ligning (4.9) fås følgende udtryk for e_{pq}^T .

$$e_{pq}^T = \begin{cases} \phi_1 h_x^{p+1} + \phi_2 h_y^{q+1} + \text{h. o. l.} & \text{for } p = q, \\ \phi_1 h_x^{p+1} + \phi_2 h_y^{q+1} + \text{f. m. l.} & \text{for } p = q - 1, \\ \phi_1 h_y^{q+1} + \phi_2 h_x^{p+1} + \text{f. m. l.} & \text{for } p = q + 1, \\ \phi_1 h_x^{p+1} + \phi_2 h_x^{p+1} h_y + \text{f. m. l.} & \text{for } p < q - 1, \\ \phi_1 h_y^{q+1} + \phi_2 h_x h_y^{q+1} + \text{f. m. l.} & \text{for } p > q + 1. \end{cases} \quad (4.17)$$



Figur 4.2: Illustration af u i tilfældet $K = 4\pi$ og $\square = [0, 1] \times [0, 1]$.

Funktionerne ϕ_1 og ϕ_2 afhænger hverken af h_x eller af h_y , men er ikke nødvendigvis identiske i de enkelte tilfælde. I overensstemmelse med konjektur 4.2.2 betegner f. m. l. formodet mindre led.

For at opskrive $\mathcal{W}^{m,r}$ Sobolev normen af fejlen genkalder vi os først definitionen af denne

$$\|v\|_{\mathcal{W}^{m,r}(\square)}^r = \sum_{\substack{i,j \geq 0 \\ i+j \leq m}} \int_{\square} \left| \frac{\partial^{i+j} v(x,y)}{\partial x^i \partial y^j} \right|^r dx dy. \quad (4.18)$$

Som sædvanlig er v en passende funktion, m et ikke negativt tal, og r et positivt tal. Ved at anvende denne definition på de enkelte tilfælde i (4.17) fås for $m = 0$

$$\|e_{pq}^T\|_{\mathcal{W}^{0,r}(\square)}^r = \begin{cases} (\phi_1 h_x^{p+1} + \phi_2 h_y^{q+1})^r h_x h_y + \text{h. o. l.} & \text{for } p = q, \\ (\phi_1 h_x^{p+1} + \phi_2 h_y^{q+1})^r h_x h_y + \text{f. m. l.} = \\ \phi_1 h_x^{(p+1)r+1} h_y + \phi_2 h_x^{(p+1)(r-1)+1} h_y^{q+2} + \text{f. m. l.} & \text{for } p = q - 1, \\ (\phi_1 h_y^{q+1} + \phi_2 h_x^{p+1})^r h_x h_y + \text{f. m. l.} = \\ \phi_1 h_x h_y^{(q+1)r+1} + \phi_2 h_x^{p+2} h_y^{(q+1)(r-1)+1} + \text{f. m. l.} & \text{for } p = q + 1, \\ (\phi_1 h_x^{p+1} + \phi_2 h_x^{p+1} h_y)^r h_x h_y + \text{f. m. l.} = \\ \phi_1 h_x^{(p+1)r+1} h_y + \phi_2 h_x^{(p+1)r+1} h_y^2 + \text{f. m. l.} & \text{for } p < q - 1, \\ (\phi_1 h_y^{q+1} + \phi_2 h_x h_y^{q+1})^r h_x h_y + \text{f. m. l.} = \\ \phi_1 h_x h_y^{(q+1)r+1} + \phi_2 h_x^2 h_y^{(q+1)r+1} + \text{f. m. l.} & \text{for } p > q + 1. \end{cases} \quad (4.19)$$

For at undgå for tung notation har vi her skrevet ϕ_1 hhv. ϕ_2 istedet ϕ_1^r hhv. ϕ_2^r . Når $m \geq 1$ er det ikke nok kun at betragte de understregede led i (4.9). Lad os som eksempel betragte tilfældet $m = 1$ og $p = q + 1$. Her er

$$\begin{aligned} \|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r &= (\phi_1 h_y^{q+1} + \phi_2 h_x^p)^r h_x h_y \\ &\quad + (\phi_3 h_y^q + \phi_4 h_y^{q+1} + \phi_5 h_x h_y^q)^r h_x h_y \\ &\quad + (\phi_6 h_y^{q+1} + \phi_7 h_y^{q+2} + \phi_8 h_x h_y^{q+1} + \phi_9 h_x^{p+1})^r h_x h_y + \text{h. o. l.} \\ &= \phi_1 h_x h_y^{qr+1} + \phi_2 h_x^2 h_y^{qr+1} + \text{f. m. l.} \end{aligned} \quad (4.20)$$

Her har vi igen “misbrugt” notationen forstået på den måde, at ϕ_1 hhv. ϕ_2 i det endelige resultat, ikke er det samme ϕ_1 og ϕ_2 som i den første ligning. De øvrige tilfælde fås på lignende vis.

$$\|e_{pq}^T\|_{\mathcal{W}^{1,r}(\square)}^r = \begin{cases} \phi_1 h_x^{pr+1} h_y + \phi_2 h_x h_y^{qr+1} + \text{h. o. l.} & \text{for } p = q \\ \phi_1 h_x^{pr+1} h_y + \phi_2 h_x^{pr+1} h_y^2 + \text{f. m. l.} & \text{for } p = q - 1 \\ \phi_1 h_x h_y^{qr+1} + \phi_2 h_x^2 h_y^{qr+1} + \text{f. m. l.} & \text{for } p = q + 1 \\ \phi_1 h_x^{pr+1} h_y + \phi_2 h_x^{pr+1} h_y^2 + \text{f. m. l.} & \text{for } p < q - 1 \\ \phi_1 h_x h_y^{qr+1} + \phi_2 h_x^2 h_y^{qr+1} + \text{f. m. l.} & \text{for } p > q + 1 \end{cases} \quad (4.21)$$

For $2 \leq m \leq 1 + \min\{p, q\}$ kan man (igen på tilsvarende måde) få

$$\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r = \begin{cases} \phi_1 h_x^{(p+1-m)r+1} h_y + \\ \quad \phi_2 h_x h_y^{(q+1-m)r+1} + \text{h. o. l.} & \text{for } p = q \\ h_x^{(p+1-m)r+1} (\phi_1 h_y + \phi_2 h_y^2) + \text{f. m. l.} & \text{for } p = q - 1 \\ h_y^{(q+1-m)r+1} (\phi_1 h_x + \phi_2 h_x^2) + \text{f. m. l.} & \text{for } p = q + 1 \\ h_x^{(p+1-m)r+1} (\phi_1 h_y + \phi_2 h_y^2) + \text{f. m. l.} & \text{for } p < q - 1 \\ h_y^{(q+1-m)r+1} (\phi_1 h_x + \phi_2 h_x^2) + \text{f. m. l.} & \text{for } p > q - 1 \end{cases} \quad (4.22)$$

Sammenfattes ovenstående resultater fås følgende sætning.

4.2.3 SÆTNING *Lad p og q være positive tal, og antag at $u : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er glat nok til at have en Taylor udvikling af orden $(p+2) \times (q+2)$ med restled i ethvert punkt i et rektangel $\square \subset \mathbb{R}^2$ med sidelængder h_x og h_y . Vi har da følgende asymptotiske fejlligning for Taylor interpolations polynomier i $Q_{pq}(\square)$ når $h_x \rightarrow 0$ og $h_y \rightarrow 0$:*

Der findes funktioner ϕ_1 og ϕ_2 og tal s_1, s_2, t_1 og t_2 uafhængige af h_x og h_y men afhængige af u, p og q så

$$e_{pq}^T = \phi_1 h_x^{s_1} h_y^{s_2} + \phi_2 h_x^{t_1} h_y^{t_2} + \text{f. m. l.}, \quad (4.23)$$

og

$$\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r = \phi_1 h_x^{s_1} h_y^{s_2} + \phi_2 h_x^{t_1} h_y^{t_2} + \text{f. m. l.} . \quad (4.24)$$

Her er m og r tal, som opfylder $0 \leq m \leq 1 + \min\{p, q\}$. Funktionerne ϕ_1 og ϕ_2 er ikke nødvendigvis identiske i (4.23) og (4.24), hvilket også gælder for s_1, s_2, t_1 og t_2 . Værdierne af s_1, s_2, t_1 og t_2 er givet fra ligningerne (4.17), (4.19), (4.21) og (4.22). Er $p = q$ erstattes f. m. l. med h. o. l..

Sætning 4.2.3 giver ikke mange oplysninger om funktionerne ϕ_1 og ϕ_2 . Vi kan få mere præcise oplysninger om disse fra ligning (4.9). Det er klart, at vi kan skrive e_{pq}^T på formen

$$e_{pq}^T = \sum_{k=1}^6 \alpha_k \phi_{p_k, q_k} + \text{h. o. l.} , \quad (4.25)$$

hvor konstanterne α_k , $k = 1, \dots, 6$ enten har værdien 0 eller 1 afhængig af, om den tilhørende funktion ϕ_{p_k, q_k} indgår i ligning (4.9). Værdierne for p_k , q_k og α_k ses i tabel 4.1. Ved at anvende (4.18), (4.25), multinomial formlen

		$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
	p_k	$p + 1$	$p + 1$	$p + 2$	0	1	0
	q_k	0	1	0	$q + 1$	$q + 1$	$q + 2$
$p = q$	α_k	1	0	0	1	0	0
$p = q - 1$	α_k	1	1	1	1	0	0
$p = q + 1$	α_k	1	0	0	1	1	1
$p < q - 1$	α_k	1	1	1	0	0	0
$p > q + 1$	α_k	0	0	0	1	1	1

Tabel 4.1: Værdierne for p_k , q_k og α_k i formen for e_{pq}^T

samt følgende relation mellem normen og seminormen i $\mathcal{W}^{m,r}$

$$\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r = |e_{pq}^T|_{\mathcal{W}^{m,r}(\square)}^r + \text{h. o. l.} \quad (4.26)$$

hvor

$$|v|_{\mathcal{W}^{m,r}(\square)}^r = \sum_{\substack{i,j \geq 0 \\ i+j=m}} \int_{\square} \left| \frac{\partial^{i+j} v(x,y)}{\partial x^i \partial y^j} \right|^r dx dy \quad (4.27)$$

ses, at

$$\begin{aligned} \|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r &= \left\| \sum_{k=1}^6 \alpha_k \phi_{p_k, q_k} \right\|^r + \text{h. o. l.} \\ &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \int_{\square} \left| \frac{\partial^{i+j} \sum_{k=1}^6 \alpha_k \phi_{p_k, q_k}(x,y)}{\partial x^i \partial y^j} \right|^r dx dy + \text{h. o. l.} \\ &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \int_{\square} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \alpha_k \frac{\partial^{i+j} \phi_{p_k, q_k}(x,y)}{\partial x^i \partial y^j} \right|^{r_k} dx dy + \text{h. o. l.} \end{aligned} \quad (4.28)$$

Defineres F som

$$\begin{aligned} F(x,y) &= r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \alpha_k \frac{\partial^{i+j} \phi_{p_k, q_k}(x,y)}{\partial x^i \partial y^j} \right|^{r_k} \\ &= r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \frac{\alpha_k}{p_k! q_k!} \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \frac{\partial^i (x-x_0)^{p_k}}{\partial x^i} \frac{\partial^j (y-y_0)^{q_k}}{\partial y^j} \right|^{r_k}, \end{aligned} \quad (4.29)$$

kan vi skrive (4.28) som

$$\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r = \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} \int_{\square} F(x, y) dx dy + \text{h. o. l.} \quad (4.30)$$

Integralet i ligning (4.30) kan udregnes ved hjælp af (4.29) og giver

$$\begin{aligned} & \|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r \\ &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} \\ & \quad \int_{\square} r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \frac{\alpha_k}{p_k! q_k!} \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \frac{\partial^i u(x - x_0)^{p_k}}{\partial x^i} \frac{\partial^j u(y - y_0)^{q_k}}{\partial y^j} \right|^{r_k} dx dy + \text{h. o. l.} \\ &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \frac{\alpha_k}{p_k! q_k!} \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \right|^{r_k} \\ & \quad \cdot \int_{\square} \prod_{k=1}^6 \left| \frac{\partial^i u(x - x_0)^{p_k}}{\partial x^i} \frac{\partial^j u(y - y_0)^{q_k}}{\partial y^j} \right|^{r_k} dx dy + \text{h. o. l.} \\ &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} r! \prod_{k=1}^6 \frac{1}{r_k!} \left| \frac{\alpha_k}{p_k! q_k!} \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \right|^{r_k} \\ & \quad \cdot \int_a^{a+h_x} \prod_{k=1}^6 \left| \frac{\partial^i u(x - x_0)^{p_k}}{\partial x^i} \right|^{r_k} dx \int_b^{b+h_y} \prod_{k=1}^6 \left| \frac{\partial^j u(y - y_0)^{q_k}}{\partial y^j} \right|^{r_k} dy + \text{h. o. l.} \end{aligned} \quad (4.31)$$

I det sidste led regner vi på det ene integrale og finder (for $1 \leq i \leq p_k$,

$$1 \leq j \leq q_k \text{ og } r_k > 0)$$

$$\begin{aligned}
& \int_a^{a+h_x} \prod_{k=1}^6 \left| \frac{\partial^i (x - x_0)^{p_k}}{\partial x^i} \right|^{r_k} dx \\
&= \int_a^{a+h_x} \prod_{k=1}^6 \left| \prod_{s=0}^{i-1} (p_k - s)(x - x_0)^{p_k-i} \right|^{r_k} dx \\
&= \prod_{k=1}^6 \prod_{s=0}^{i-1} (p_k - s)^{r_k} \int_a^{a+h_x} \prod_{k=1}^6 |x - x_0|^{(p_k-i)r_k} dx \\
&= \prod_{k=1}^6 \prod_{s=0}^{i-1} (p_k - s)^{r_k} \int_a^{a+h_x} |x - x_0|^{\sum_{k=1}^6 (p_k-i)r_k} dx \\
&= \prod_{k=1}^6 \prod_{s=0}^{i-1} (p_k - s)^{r_k} \left(\left[\frac{1}{1 + \sum_{k=1}^6 (p_k - i)r_k} |x - x_0|^{1 + \sum_{k=1}^6 (p_k-i)r_k} \right]_{x_0}^{a+h_x} \right. \\
&\quad \left. - \left[\frac{1}{1 + \sum_{k=1}^6 (p_k - i)r_k} |x - x_0|^{1 + \sum_{k=1}^6 (p_k-i)r_k} \right]_a^{x_0} \right)
\end{aligned} \tag{4.32}$$

Indføres notation

$$\kappa_{1,k} = \begin{cases} 1 & \text{for } i = 0 \text{ eller } r_k = 0 \\ \prod_{s=0}^{i-1} (p_k - s)^{r_k} & \text{for } 1 \leq i \leq p_k \text{ og } r_k > 0, \\ 0 & \text{for } i > p_k \text{ og } r_k > 0 \end{cases} \tag{4.33}$$

$$\kappa_{2,k} = \begin{cases} 1 & \text{for } j = 0 \text{ eller } r_k = 0 \\ \prod_{s=0}^{j-1} (q_k - s)^{r_k} & \text{for } 1 \leq j \leq q_k \text{ og } r_k > 0, \\ 0 & \text{for } j > q_k \text{ og } r_k > 0 \end{cases} \tag{4.34}$$

$$\gamma_1 = 1 + \sum_{k=1}^6 (p_k - i)r_k, \quad \gamma_2 = 1 + \sum_{k=1}^6 (q_k - j)r_k, \tag{4.35}$$

$$g(\alpha, \beta) = (\alpha^{\gamma_1} + (1 - \alpha)^{\gamma_1}) \cdot (\beta^{\gamma_2} + (1 - \beta)^{\gamma_2}), \tag{4.36}$$

$$(x_0, y_0) = (a + \alpha h_x, b + \beta h_y) \text{ dvs } (\alpha, \beta) \in [0, 1]^2. \tag{4.37}$$

kan (4.31) mere overskueligt skrives som

$$\begin{aligned}
\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r &= \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} \prod_{k=1}^6 \left\{ \frac{\kappa_{1,k} \kappa_{2,k}}{r_k!} \left(\frac{\alpha_k}{p_k! q_k!} \right)^{r_k} \left| \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \right|^{r_k} \right\} \\
&\quad \cdot \frac{r!}{\gamma_1 \gamma_2} g(\alpha, \beta) h_x^{\gamma_1} h_y^{\gamma_2} + \text{h. o. l.}, \tag{4.38}
\end{aligned}$$

Da normen af fejlen afhænger af funktionen g , vil det være på sin plads med et par bemærkninger om denne. Vi ser, at $(1/2)^{\gamma_1+\gamma_2-2} \leq g(\alpha, \beta) \leq 1$ for $(\alpha, \beta) \in [0, 1]^2$. Funktionen g 's minimum antages såfremt (x_0, y_0) er rektanglets centrum, dvs $(\alpha, \beta) = (1/2, 1/2)$, mens dens maksimum antages i rektanglets hjørner, dvs $\alpha, \beta \in \{0, 1\}$.

Vi vil nu betragte det mest almindelige special tilfælde, hvor polynomiumsgraden er den samme i x -retningen som i y -retningen, dvs $p = q$. Fra tabel 4.1 ses, at vi i denne situation har, $\alpha_1 = \alpha_4 = 1$, og $\alpha_2 = \alpha_3 = \alpha_5 = \alpha_6 = 0$. Fra ligning (4.30) ses, at vi kun får bidrag til fejlen såfremt $r_2 = r_3 = r_5 = r_6 = 0$, og $r_1 + r_4 = r$. For at forenkle notationen vil vi for vilkårligt $k \in \{0, \dots, r\}$ sætte $r_1 = k$ og $r_4 = r - k$. Før vi ser, hvad (4.30) reducerer til i dette special tilfælde, bemærker vi, at $p_1 = q_4 = p + 1$ og $q_1 = p_4 = 0$, hvorfor minimum af g er

$$\min_{(\alpha, \beta) \in [0, 1]^2} g(\alpha, \beta) = \left(\frac{1}{2}\right)^{(p+1-m)r}. \quad (4.39)$$

Altså er g 's minimum uafhængig af k men afhængig af m . Lad os først betragte tilfældet $m = 0$, da vil $i = j = 0$, og ligningen (4.30) simplificerer til

$$\begin{aligned} & \|e_{pp}^T\|_{\mathcal{W}^{0,r}(\square)}^r \\ &= \sum_{\substack{r_1, r_4 \geq 0 \\ r_1 + r_4 = r}} \frac{r!}{r_1! r_4! \{(p+1)!\}^r} \left| \frac{\partial^{p+1} u(x_0, y_0)}{\partial x^{p+1}} \right|^{r_1} \left| \frac{\partial^{p+1} u(x_0, y_0)}{\partial y^{p+1}} \right|^{r_4} \\ & \cdot \int_a^{a+h_x} |x - x_0|^{(p+1)r_1} dx \int_b^{b+h_y} |y - y_0|^{(p+1)r_4} dy + \text{h. o. l.} \quad (4.40) \\ &= \sum_{k=0}^r C_{k,0,r,p}^T g(\alpha, \beta) \left| \frac{\partial^{p+1} u(x_0, y_0)}{\partial x^{p+1}} \right|^k \left| \frac{\partial^{p+1} u(x_0, y_0)}{\partial y^{p+1}} \right|^{r-k} \\ & \cdot h_x^{(p+1)k+1} h_y^{(p+1)(r-k)+1} + \text{h. o. l.} \quad \text{for } p, r \in \mathbb{N}, \end{aligned}$$

hvor

$$C_{k,0,r,p}^T = \frac{r!}{k!(r-k)!\{(p+1)!\}^r ((p+1)k+1)((p+1)(r-k)+1)} \quad \text{for } k = 0, \dots, r \text{ og } p, r \in \mathbb{N} \quad (4.41)$$

I tabel 4.2 ses eksempler på værdier af $C_{k,0,r,p}^T$ for nogle af de mest anvendte normer. I tabel 4.3 vises $C_{k,0,r,p}^T \cdot g(1/2, 1/2)$ for de samme normer. For $m > 0$ er der to tilfælde, hvor vi får et bidrag til summen i (4.30): ($i = m, j =$

norm	m	r	k	$p = 1$	$p = 2$	$p = 3$	$p = 4$
\mathcal{L}^1	0	1	0	1/6	1/24	1/120	1/720
\mathcal{L}^1	0	1	1	1/6	1/24	1/120	1/720
\mathcal{L}^2	0	2	0	1/20	1/252	1/5184	1/158400
\mathcal{L}^2	0	2	1	1/18	1/288	1/7200	1/259200
\mathcal{L}^2	0	2	2	1/20	1/252	1/5184	1/158400

Tabel 4.2: Eksempler på værdier for $C_{k,0,r,p}^T$ for normerne \mathcal{L}^1 og \mathcal{L}^2 .

norm	m	r	k	$p = 1$	$p = 2$	$p = 3$	$p = 4$
\mathcal{L}^1	0	1	0	1/24	1/192	1/1920	1/23040
\mathcal{L}^1	0	1	1	1/24	1/192	1/1920	1/23040
\mathcal{L}^2	0	2	0	1/320	1/16128	1/1327104	1/162201600
\mathcal{L}^2	0	2	1	1/288	1/18432	1/1843200	1/265420800
\mathcal{L}^2	0	2	2	1/320	1/16128	1/1327104	1/162201600

Tabel 4.3: Eksempler på værdier for $C_{k,0,r,p}^T \cdot g(1/2, 1/2)$ for normerne \mathcal{L}^1 og \mathcal{L}^2 .

$0, r_1 = r$ og $r_4 = 0$) og ($i = 0, j = m, r_1 = 0$ og $r_4 = r$). Vi finder her

$$\begin{aligned}
& \|e_{pp}^T\|_{\mathcal{W}^{m,r}(\square)}^r \\
&= \left(\frac{1}{(p+1)!}\right)^r \left[\left| \frac{\partial^{p+1}u(x_0, y_0)}{\partial x^{p+1}} \right|^r \int_a^{a+h_x} \left| \frac{d^m(x-x_0)^{p+1}}{dx^m} \right|^r dx \int_b^{b+h_y} 1 dy \right. \\
&\quad \left. + \left| \frac{\partial^{p+1}u(x_0, y_0)}{\partial y^{p+1}} \right|^r \int_a^{a+h_x} 1 dx \int_b^{b+h_y} \left| \frac{d^m(y-y_0)^{p+1}}{dy^m} \right|^r dy \right] + \text{h. o. l.} \\
&= C_{m,r,p}^T g(\alpha, \beta) \left[\left| \frac{\partial^{p+1}u(x_0, y_0)}{\partial x^{p+1}} \right|^r h_x^{(p+1-m)r+1} h_y \right. \\
&\quad \left. + \left| \frac{\partial^{p+1}u(x_0, y_0)}{\partial y^{p+1}} \right|^r h_x h_y^{(p+1-m)r+1} \right] + \text{h. o. l.} \\
&\text{for } (m, r, p \in \mathbb{N}, m \leq p+1)
\end{aligned} \tag{4.42}$$

Det checkes let, at situationen $m = r - 1 = 0$ giver samme resultat. Ovenfor

er

$$\begin{aligned}
C_{m,r,p}^T &= \left(\frac{1}{(p+1)!} \right)^r \frac{[(p+1) \cdot \dots \cdot (p+2-m)]^r}{(p+1-m)r+1} \\
&= \frac{1}{[(p+1-m)!]^r ((p+1-m)r+1)} \\
&\text{for } (m, r, p \in \mathbb{N}, m \leq p+1) \text{ eller } (p \in \mathbb{N}, m = r-1 = 0)
\end{aligned} \tag{4.43}$$

I tabel 4.4 og 4.5 ses eksempler på værdien af $C_{m,r,p}^T$ og $C_{m,r,p}^T \cdot g(1/2, 1/2)$ for nogle af de mest anvendte normer.

norm	m	r	$p = 1$	$p = 2$	$p = 3$	$p = 4$
\mathcal{L}^1	0	1	1/6	1/24	1/120	1/720
\mathcal{H}^1	1	2	1/3	1/20	1/252	1/5184
$\mathcal{W}^{1,4}$	1	4	1/5	1/144	1/16848	1/5640192

Tabel 4.4: Eksempler på værdier for $C_{m,r,p}^T$ for normerne \mathcal{L}^1 , $\mathcal{H}^1 = \mathcal{W}^{1,2}$ og $\mathcal{W}^{1,4}$

norm	m	r	$p = 1$	$p = 2$	$p = 3$	$p = 4$
\mathcal{L}^1	0	1	1/24	1/192	1/1920	1/23040
\mathcal{H}^1	1	2	1/12	1/320	1/16128	1/1327104
$\mathcal{W}^{1,4}$	1	4	1/80	1/36864	1/69009408	1/369635622912

Tabel 4.5: Eksempler på værdier for $C_{m,r,p}^T \cdot g(1/2, 1/2)$ for normerne \mathcal{L}^1 , $\mathcal{H}^1 = \mathcal{W}^{1,2}$ og $\mathcal{W}^{1,4}$

4.2.4 BEMÆRKNING Fra ligningerne 4.40 og 4.42 ses, at normen af interpolations fejlen afhænger af funktionen g . Vi genkalder, at g 's minimum antages i $(1/2, 1/2)$. I tabellerne 4.3 og 4.5 har vi derfor udregnet værdierne af $C_{k,0,r,p}^T \cdot g(1/2, 1/2)$ hhv. $C_{m,r,p}^T \cdot g(1/2, 1/2)$.

Ovenstående resultater kan sammenfattes i følgende sætning

4.2.5 SÆTNING *Lad p og q være positive tal, og antag at $u : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ er glat nok til at have en Taylor udvikling af orden $(p+2) \times (q+2)$ med restled i ethvert punkt i et rektangel $\square \subset \mathbb{R}^2$ med sidelængder h_x og h_y . Vi har da følgende asymptotiske fejlligning for Taylor interpolations polynomier i*

$Q_{pq}(\square)$ når $h_x \rightarrow 0$ og $h_y \rightarrow 0$:

$$\|e_{pq}^T\|_{\mathcal{W}^{m,r}(\square)}^r = \sum_{\substack{i,j \geq 0 \\ i+j=m}} \sum_{\substack{r_1, \dots, r_6 \geq 0 \\ r_1 + \dots + r_6 = r}} C_-^T \left\{ \prod_{k=1}^6 \left| \frac{\partial^{p_k+q_k} u(x_0, y_0)}{\partial x^{p_k} \partial y^{q_k}} \right|^{r_k} \right\} \\ \cdot h_x^{1-ir + \sum_{k=1}^6 p_k r_k} h_y^{1-jr + \sum_{k=1}^6 q_k r_k} + \text{h. o. l.} \quad (4.44)$$

for et vilkårligt punkt $(x_0, y_0) = (a + \alpha h_x, b + \beta h_y) \in \square$, og tal m og r opfyldende $0 \leq m \leq 1 + \min\{p, q\}$ og $r > 0$. Konstanten C_-^T er givet ved

$$C_-^T = \left\{ r! \prod_{k=1}^6 \frac{1}{r_k!} \left(\frac{\alpha_k}{p_k! q_k!} \right)^{r_k} \right\} \\ \cdot \int_0^1 \prod_{k=1}^6 \left| \frac{\partial^i (s - \alpha)^{p_k}}{\partial s^i} \right|^{r_k} ds \int_0^1 \prod_{k=1}^6 \left| \frac{\partial^j (t - \beta)^{q_k}}{\partial t^j} \right|^{r_k} dt \quad (4.45)$$

Er $p = q$ reducerer (4.44) og (4.45) til (4.40) og (4.41) for $m = 0$ og til (4.42) og (4.43) for $1 \leq m \leq p + 1$ eller $m = r - 1 = 0$.

4.2.6 BEMÆRKNING I artiklen [29] betragtes udover Taylor interpolation også Lagrange interpolation. For Lagrange interpolation vises et analogt resultat til sætning 4.2.5. Strukturen af fejlligningerne vises at være identiske blot afhænger interpolations fejlen i Lagrange tilfældet ikke af en funktion g . Det konkluderes i artiklen at Lagrange interpolation er bedre end Taylor interpolation i det generelle tilfælde. Ved at vælge (x_0, y_0) som rektanglets centrum (dvs. således at g 's minimum antages) vil Taylor interpolation være bedre end eller mindst lige så god som Lagrange interpolation for små værdier af m , r og p .

Kapitel 5

Netgenerering

Målet med dette kapitel er at præsentere en netinddelings metode, der giver asymptotiske optimale netinddelinger (indenfor klassen af netinddelinger med rektangulære elementer). Metoden bygger på begrebet net tætheds funktion og blev indført af Babuška og Gui i [12]. Metoden er senere videreudviklet af Hugger i artiklerne [21], [22], [23], [25], [29] og senest af Babuška, Hugger et al i [8]. Med net tætheds funktioner kan man beskrive følger af netinddelinger ved brug af simple midler, og man undgår dermed de komplicerede datastrukturer, der ofte karakteriserer netinddelings algoritmer. Ønsket om simple metoder til at beskrive netinddelinger på udspringer fra parametriserede problemer, hvor man med fordel ofte kan genanvende tidligere netinddelinger. Det vil fremgå af teorien nedenfor, at metoden med fordel også kan anvendes, selvom man kun skal løse et enkelt problem. Netinddelings algoritmen er ment som en “one shot” metode, der på baggrund af informationer fra net tætheds funktionen og den bruger bestemte fejltolerance forsøger at konstruere en netinddeling, hvor fejlen i den tilhørende endelige element løsning efter første gennemløb af algoritmen er mindre end den givne fejltolerance. Mislykkes dette, vil metoden iterere.

5.1 Sammenhæng mellem net tætheds funktioner og følger af net

Før vi definerer net tætheds funktioner, har vi brug for følgende

5.1.1 DEFINITION *Lad $\mathcal{T} = \{\square_j\}_{j=1}^m$ være et endeligt element net. Vi siger da, at \mathcal{T} er en optimal endelig element netinddeling indenfor en given klasse af netinddelinger, såfremt fejlen i den tilhørende endelige element løsning er mindre end en given fejltolerance, og såfremt de beregningsmæssige omkost-*

ninger ved at finde denne løsning er minimale. Minimum er her taget over alle netinddelinger indenfor den givne klasse.

5.1.2 DEFINITION Lad Ω være en åben sammenhængende mængde i \mathbb{R}^n . Ved mængden af endeligt ikke-glatte funktioner defineret på Ω , vil vi forstå mængden af reelle funktioner defineret på Ω , som er uendeligt ofte differentiable (C^∞) i Ω undtaget på et antal delmængder af Ω . Om disse mængder (hvor funktionen ikke tilhører C^∞) skal der gælde, at antallet af disse i enhver begrænset delmængde af Ω er endeligt. Desuden antages enhver af disse mængder at være glat og endelig i den forstand, at en sådan mængde enten er et punkt eller en k -dimensional C^∞ undermangfoldighed af \mathbb{R}^n med endeligt k -volumen, hvor $k \in \{1, \dots, n-1\}$. Vi vil betegne mængden af endeligt ikke-glatte funktioner defineret på Ω med $\mathcal{C}_{\text{fin}}^\infty(\Omega)$.

5.1.3 BEMÆRKNING Er Ω fra definition 5.1.2 begrænset, er der kun endeligt mange "singularitets" mængder.

5.1.4 BEMÆRKNING Funktionerne i $\mathcal{C}_{\text{fin}}^\infty(\Omega)$ er tilstrækkelige til at beskrive stort set alle praktiske fænomener indenfor anvendt matematik.

Ved brug af funktionerne i $\mathcal{C}_{\text{fin}}^\infty(\Omega)$ kan vi nu definere net tætheds funktioner som

5.1.5 DEFINITION En net tætheds funktion (eng. mesh density function) defineret på en åben sammenhængende mængde $\Omega \subset \mathbb{R}^n$, er en næsten overalt positiv (eventuelt ubegrænset) funktion \mathcal{D} tilhørende $\mathcal{C}_{\text{fin}}^\infty(\Omega)$, som opfylder

$$\int_{\Omega} \mathcal{D}(x) dx = 1. \quad (5.1)$$

Mængden af net tætheds funktioner betegnes med D_Ω .

5.1.6 DEFINITION Lad \square være et element i en netinddeling, og lad \mathcal{D} være en net tætheds funktion. Ved tætheden af \square forstås værdien af $\int_{\square} \mathcal{D}(x) dx$.

Net tætheds funktioner blev indført af Babuška og Gui i [12]. Deres ide var at konstruere netinddelinger på en sådan måde, at tætheden i ethvert element i netinddelingen er den samme. Principielt kunne man arbejde med netinddelinger, hvor tætheden varierer fra element til element, men kravet er praktisk set fra en implementations vinkel. Ved hjælp af ovenstående definition kan dette krav udtrykkes som

$$\int_{\square_j} \mathcal{D}(x) dx = 1/N_a \quad \text{for } j = 1, \dots, N_a, \quad (5.2)$$

for en netinddeling $\mathcal{T} = \{\square_j\}_{j=1}^{N_a}$. I praksis vil man kun kræve, at (5.2) gælder approksimativt.

Bemærk at (5.2) beskriver en hel følge af netinddelinger, en netinddeling for hvert positivt tal N_a . Net tætheds funktioner kan altså bruges til at beskrive følger af net og ikke kun et enkelt net. Artiklen [22] indeholder resultater, der i det asymptotiske tilfælde og under passende forudsætninger viser, at vælges net tætheds funktionen på fornuftig vis, da vil normen af fejlen i den endelige element løsning være minimal. De følger af netinddelinger vi fremover skal arbejde med antages at have den egenskab, at antallet af elementer i netinddelingen er ikke-aftagende og divergerer mod uendelig. En mere formel definition, der også giver os den nødvendige notation, lyder

5.1.7 DEFINITION *Lad Ω være en åben sammenhængende mængde i \mathbb{R}^n , og lad $\{\mathcal{T}_m\}_{m=1}^\infty$ være en vilkårlig følge af endelige element net på Ω . For ethvert positivt tal m betegner $N_{a,m}$ antallet af elementer i \mathcal{T}_m . Mængden af ikke-aftagende følger af endelige element net, betegnet med P_Ω , defineres da som de følger af endelige element net $\{\mathcal{T}_m\}_{m=1}^\infty$, hvor den tilhørende følge $\{N_{a,m}\}_{m=1}^\infty$ er ikke-aftagende og divergerer mod uendelig.*

Følgende definition giver en sammenhæng mellem net tætheds funktioner og ikke-aftagende følger af endelige element net.

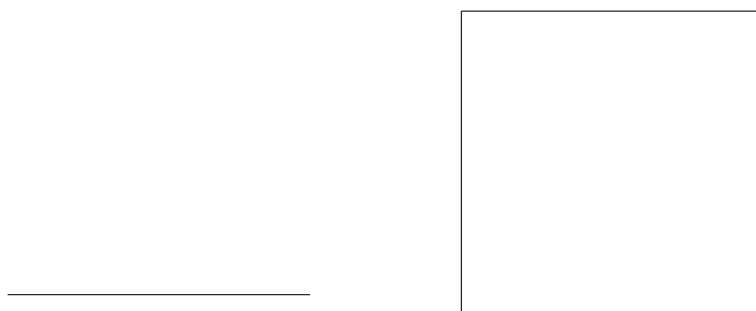
5.1.8 DEFINITION *Lad D_Ω og P_Ω være som i definition 5.1.5 og 5.1.7, og lad Ω være en åben sammenhængende mængde i \mathbb{R}^n . En net tætheds operator A (eng. mesh density operator) defineret på Ω er en invertibel afbildning*

$$A : \mathcal{D} \in D_\Omega \rightarrow \{\mathcal{T}_m\}_{m=1}^\infty \in R_A \subset P_\Omega , \quad (5.3)$$

som afbilder mængden D_Ω af alle net tætheds funktioner på en delmængde R_A (billedet af A) af mængden P_Ω af alle ikke-aftagende følger af endelige element net.

At der rent faktisk eksisterer net tætheds operatorer er ikke oplagt. I det endimensionale tilfælde er det let at bevise, se fx [23]. Et bevis i det n -dimensionale tilfælde er mere kompliceret, idet (5.2) ikke indeholder informationer nok til at bestemme en netinddeling (eller en følge af netinddelinger) entydigt. Udfra (5.2) kan vi kun få oplysninger om elementernes tæthed og ikke om deres geometriske form. Dette problem kan løses på to forskellige måder. En fremgangsmåde er at tilføje betingelser til (5.2), der specificerer elementernes geometriske form. Alternativt kan man vælge kun at benytte en bestemt type af netinddelinger med foruddefinerede geometriske former for elementerne. Vi skal her forfølge den sidstnævnte ide.

Før vi definerer tilladte netinddelinger og elementer, er det nødvendigt at præcisere, hvilken form vi vil tillade for domænet for det betragtede problem. Tilladte domæner vil blive benævnt β -domæner. Ideen er at tillade domæner i \mathbb{R}^n , der kan inddeles på en sådan måde, at hver delmængde i inddelingen kan afbildes bijektivt på en enheds n -kube $\square_n = [0, 1]^n$. Denne afbildning skal desuden bevare glatheden i den endelige element løsning i de enkelte elementer samt over randene mellem elementerne. Figur 5.1 indeholder et par eksempler på tilladte domæner. Ideerne i de følgende definitioner er udviklet i [6] hhv. [23]. En definition af β -domæner kræver lidt notation vedrørende



Figur 5.1: Eksempler på β -domæner

enheds n -kuber.

5.1.9 DEFINITION Ved en afsluttet r -side i enheds n -kuben \square_n for $r = 0, \dots, n$ forstås en delmængde af \square_n med r frie koordinater, og hvor de $n - r$ resterende koordinater enten er 0 eller 1.

Med denne definition er 0-siderne i \square_n hjørnepunkterne, og 1-siderne er kanterne. Vi kan nu give en definition af β -domæner.

5.1.10 DEFINITION En åben sammenhængende mængde $\Omega \subset \mathbb{R}^n$ kaldes et β -domæne, såfremt følgende betingelser er opfyldt

- Ω kan inddeles i et endeligt antal $l \geq 1$ af parvis disjunkte, åbne, sammenhængende delmængder $\Omega_d \subset \Omega$, $d \in \{1, \dots, l\}$ på en sådan måde, at $\overline{\Omega} = \cup_{d=1}^l \overline{\Omega}_d$.
- Der findes l invertible \mathcal{C}^∞ afbildninger β_d , $d \in \{1, \dots, l\}$, hvor afslutningen af $\beta_d(\Omega_d)$, $d \in \{1, \dots, l\}$ er enheds n -kuben \square_n .
- For to vilkårlige r -sider ($r \in \{1, \dots, n\}$) i to forskellige delmængder Ω_e og Ω_f i inddelingen af Ω , skal der gælde, at fællesmængden for r -siderne enten er tom eller en hel s -side for et passende $s \leq r$.

- Afbildningerne β_d , $d \in \{1, \dots, l\}$ skal bevare en eventuel glathed over randene mellem delmængder i inddelingen af Ω .
- Lad u_d , $d \in \{1, \dots, l\}$ være funktioner tilhørende $\mathcal{C}^{k_{\max}}(\square_n)$ defineret på enheds n -kuben. Her er k_{\max} det størst mulige tal som opfylder, at $\mathcal{C}^{k_{\max}}$ er et underrum af testrummet for den betragtede endelige element metode. Antag er der for to punkter $y_1, y_2 \in \square_n$, findes $e, f \in \{1, \dots, l\}$ med $e \neq f$, så $\beta_e(y_1) = \beta_f(y_2)$. Da skal $Du_e(y_1) = Du_f(y_2)$ for alle afledede D af orden mindre end eller lig med k_{\max} . Desuden skal funktionen U givet ved $U|_{\Omega_d} = u_d \circ \beta_d$, $d = 1, \dots, l$ være entydig defineret i hele Ω og tilhøre $\mathcal{C}^{k_{\max}}(\Omega)$.

For β -domæner skal vi definere en bestemt type af netinddelinger også kaldet β -netinddelinger.

5.1.11 DEFINITION Lad $\Omega \subset \mathbb{R}^n$ være et β -domæne, og lad notationen være som i definition 5.1.10. Lad $\mathcal{T}_d = \{\square_{j,d}\}_{j=1}^{N_{a,d}}$ være et endelig element net over Ω_d for $d \in \{1, \dots, l\}$, og lad $\mathcal{T} = \cup_{d=1}^l \mathcal{T}_d$ være netinddelingen af Ω bestående af alle elementerne i delnettene. Da defineres β -netinddelingen \mathcal{T}_β hørende til \mathcal{T} som l -tuplen $(\mathcal{T}_{1,\beta}, \dots, \mathcal{T}_{l,\beta})$ af netinddelinger over enheds n -kuben, hvor $\mathcal{T}_{d,\beta} = \{\beta_d(\square_{j,d})\}_{j=1}^{N_{a,d}}$ for $d = 1, \dots, l$. Pr. definition er elementerne i \mathcal{T}_β mængden bestående af $\beta_d(\square_{j,d})$ for $j = 1, \dots, N_{a,d}$ og $d = 1, \dots, l$. β -netinddelingen $\mathcal{T}^* = (\beta_1(\bar{\Omega}_1), \dots, \beta_l(\bar{\Omega}_l))$ kaldes for reference β -netinddelingen af Ω (hørende til inddelingen $\{\Omega_d\}_{d=1}^l$ af Ω og afbildningerne $\{\beta_d\}_{d=1}^l$).

5.1.12 DEFINITION En delmængde \square af \mathbb{R}^n siges at være n -firkantet (eng. n -quadrilateral), såfremt der findes en lineær afbildning $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ så $L(\square_n) = \square$, hvor \square_n er n -enheds kubens. Hvis kanterne (se definition 5.1.9) i \square har samme længde, og alle vinkler er rette, kaldes \square for n -kubisk.

5.1.13 DEFINITION Lad $\Omega \subset \mathbb{R}^n$ være et β -domæne, og lad notationen være som i definition 5.1.10 og 5.1.11. β -netinddelingen \mathcal{T}_β siges da at have n -firkantede elementer, hvis ethvert element \square i \mathcal{T}_β er n -firkantet. Tilsvarende siges \mathcal{T}_β at have n -kubiske elementer, såfremt ethvert element er n -kubisk.

For at påvise eksistensen af net tætheds operatorer vil vi ikke tillade vilkårlige β -netinddelinger. Vi vil begrænse os til følgende

5.1.14 DEFINITION Ved mængden af \mathcal{T}^* -tilladte β -netinddelinger med n -kubiske elementer (eng. \mathcal{T}^* -compatible β -meshes with n -cubic elements) forkortet TCC , forstås de netinddelinger, der kan konstrueres udfra følgende rekursive regel

- \mathcal{T}^* er et TCC .
- Hvis \mathcal{T} er et TCC , er den β -netinddeling, der fremkommer ved at inddele et vilkårligt element i \mathcal{T} i 2^n kongruente n -kuber igen et TCC .
- En β -netinddeling, der ikke kan konstrueres ud fra overstående, er ikke et TCC .

5.1.15 DEFINITION Lad \mathcal{T}^* være en reference β -netinddeling, \mathcal{T} en \mathcal{T}^* -tilladt β -netinddeling med n -kubiske elementer, og lad $\{\square_j\}_{j \in S}$ være en mængde af elementer i \mathcal{T} . Hvis netinddelingen \mathcal{T}' , fået ud fra \mathcal{T} ved at slå elementerne $\{\square_j\}_{j \in S}$ sammen til et element $\square_p = \cup_{j \in S} \square_j$, er en \mathcal{T}^* -tilladt β -netinddeling med n -kubiske elementer, kaldes \square_p et forældre element, og mængden $\{\square_j\}_{j \in S}$ kaldes et søskende sæt.

Vi har nu begreber og notation på plads til at kunne vise eksistensen af net tætheds operatorer.

5.1.16 SÆTNING Lad Ω være et begrænset β -domæne i \mathbb{R}^n , og lad \mathcal{T}^* være en reference β -netinddeling af Ω . I algoritmen 5.1 er $\hat{\mathcal{T}}$ en \mathcal{T}^* -tilladt β -netinddeling med n -kubiske elementer, og α er en reel, positiv parameter. Endelig er $\mathcal{D} \in D_\Omega$ en net tætheds funktion defineret på Ω . Definer nu operatoren A_n^* ved

$$A_n^*(\mathcal{D}) = \{\mathcal{T}_m\}_{m=1}^\infty, \quad (5.4)$$

hvor $\{\mathcal{T}_m\}_{m=1}^\infty$ er den følge β -netinddelinger, som man kan opnå ved at anvende algoritme 5.1 med intensiteten \mathcal{I} (den reciprokke af det "ønskede" antal elementer) varierende over alle positive tal, og derefter sorteret så den tilhørende følge $\{N_{a,m}\}_{m=1}^\infty$ er ikke-aftagende. Da er A_n^* en net tætheds operator. Betegnes med $R_{A_n^*} = A_n^*(D_\Omega)$ billedet af A_n^* , er den inverse afbildning $(A_n^*)^{-1}$ til A_n^* givet ved

$$(A_n^*)^{-1} : \{\mathcal{T}_m\}_{m=1}^\infty \in R_{A_n^*} \rightarrow \mathcal{D} \in D_\Omega, \quad (5.5)$$

hvor

$$\mathcal{D}(x) = \alpha \lim_{k \rightarrow \infty} \frac{\mathcal{I}_{m_k}}{V_n(\square_{m_k}(x))} \quad \text{næsten overalt i } \Omega. \quad (5.6)$$

I definitionen (5.6) af net tætheds funktionen er $\square_m(x)$ det element i $\mathcal{T}_m = \{\square_{j,m}\}_{m=1}^{N_{a,m}}$, der indeholder x . Bemærk at elementet $\square_m(x)$ kun er næsten overalt entydig. $V_n(\square)$ er n -volumenet af \square , og \mathcal{I}_m er infimum over alle de intensiteter \mathcal{I} fra algoritme 5.1, der resulterede i β -netinddelinger \mathcal{T}_m . Endelig er $\{m_k\}_{k=1}^\infty$ en delfølge af indices valgt således, at elementet $\square_{m_k}(x)$ fra netinddelingen \mathcal{T}_{m_k} i den efterfølgende netinddeling $\mathcal{T}_{m_{k+1}}$ er blevet inddelt i 2^n kongruente elementer.

I. Forgrovning

For alle relevante søskende sæt $\{\square_j\}_{j \in S}$ i $\hat{\mathcal{T}}$ gøres følgende:

- (1) Udregn $\kappa(\square_p) = (1/\mathcal{I}) \int_{\square_p} \mathcal{D}(x) dx$, hvor $\square_p = \cup_{j \in S} \square_j$
- (2) Hvis $\kappa(\square_p) \geq \alpha$ foretages ingenting
- (3) Hvis $\kappa(\square_p) < \alpha$ konstrueres en ny β -netinddeling

ved

at erstatte $\{\square_j\}_{j \in S}$ med \square_p

II. Forfining

For alle elementerne \square i β -netinddelingen som blev konstrueret i forgrovnings delen gøres følgende:

- (1) Udregn $\kappa(\square) = (1/\mathcal{I}) \int_{\square} \mathcal{D}(x) dx$, (hvis det ikke allerede er gjort)
- (2) Hvis $\kappa(\square) \leq \alpha$ forbliver \square uændret
- (3) Hvis $\kappa(\square) > \alpha$
 - (a) Konstruer en ny β -netinddeling ved at

inddele \square i

2^n kongruente n -kuber

- (b) For alle underelementerne \square_s af \square gå til

II.(1)

Algoritme 5.1: Netinddelings algoritme byggende på net tætheds funktion

5.1.17 BEMÆRKNING I algoritme 5.1 udføres forgrovning og forfining kun en enkelt gang. Dette er valgt for at undgå muligheder for uendelige løkker. Forgrovning udføres før forfining for at sikre en mere effektiv udnyttelse af computer lagerplads. Ved at udføre forgrovning og forfining i nævnte rækkefølge kan man nemlig udnytte den lagerplads, der er blevet frigjort i forbindelse med forgrovning til at gemme informationer om forfining i. For flere detaljer om algoritmen henvises til afsnit 5.2.3, hvor vi konstruerer en mere generel algoritme. Den initiale netinddeling $\hat{\mathcal{T}}$ er bruger bestemt og kan eventuelt vælges som \mathcal{T}^* .

BEVIS Det er klart, at $A_n^*(\mathcal{D})$ er en tilladt ikke-aftagende netinddeling af Ω for enhver net tætheds funktion \mathcal{D} . Dette følger af den måde, hvorpå netinddelinger i $A_n^*(\mathcal{D})$ er sorteret samt af, at algoritmen 5.1 er konstrueret således, at ethvert element på et eller andet tidspunkt bliver inddelt i mindre elementer når intensiteten aftager. Altså vil det samlede antal elementer gå mod uendelig, når intensiteten går mod nul. For at vise at det resterende udsagn

i sætningen er korrekt, skal vi vise, at \mathcal{D} defineret ved (5.5) og (5.6) opfylder $A_n^*(\mathcal{D}) = \{\mathcal{T}_m\}_{m=1}^\infty$, hvor \mathcal{T}_m er netinddelingen fået fra algoritme 5.1. Da $\{\mathcal{T}_m\}_{m=1}^\infty$ antages at tilhøre billedet af A_n^* , findes der en net tætheds funktion $\tilde{\mathcal{D}}$, så $A_n^*(\tilde{\mathcal{D}}) = \{\mathcal{T}_m\}_{m=1}^\infty$. Lad S_0 betegne den mængde, hvor $\tilde{\mathcal{D}}$ er diskontinuert eller har singulariteter. Mængden S_0 har mål 0. Bemærk at mængden $S \equiv S_0 \cup (\cup_{m=1}^\infty \cup_{j=1}^{N_{a,m}} \partial \square_{j,m})$ ligeledes har mål 0. Sæt nu $a_m = \int_{\square_m(x)} \tilde{\mathcal{D}} dx$, og lad \mathcal{I}_m være et vilkårligt tal. Ved at indsætte \mathcal{I}_m for \mathcal{I} i algoritmen 5.1 fås en netinddeling $\{\mathcal{T}_m\}_{m=1}^\infty$. Ifølge algoritmen vil en forfining af $\square_m(x)$ kun ske såfremt $a_m(x)/\mathcal{I}_m > \alpha$. Men ifølge konstruktionen af algoritmen og definitionen af a_m vil $a_m(x)/\mathcal{I}_m \leq \alpha$. Fra definitionen af indeks delfølgen $\{m_k\}_{k=1}^\infty$ følger det nu, at $a_{m_k}/\mathcal{I}_{m_k} = \alpha$ for alle positive heltal k . Vi har nu

$$\begin{aligned} \tilde{\mathcal{D}}(x) &= \lim_{m \rightarrow \infty} \frac{a_{m_k}(x)}{V_n(\square_{m_k}(x))} \\ &= \frac{a_{m_k}(x)}{\mathcal{I}_{m_k}} \lim_{m \rightarrow \infty} \frac{\mathcal{I}_{m_k}}{V_n(\square_{m_k}(x))} \\ &= \alpha \lim_{m \rightarrow \infty} \frac{\mathcal{I}_{m_k}}{V_n(\square_{m_k}(x))} \quad \text{for } x \in \bar{\Omega} \setminus S. \end{aligned} \tag{5.7}$$

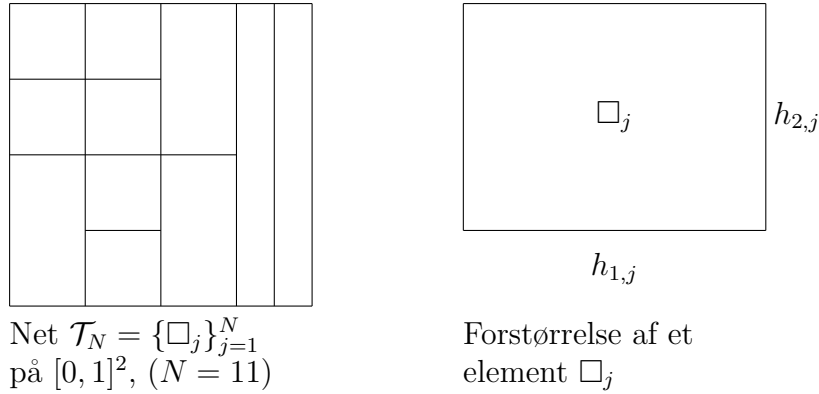
Da S har mål 0 følger det af (5.7), at \mathcal{D} givet ved (5.6) er den rigtige net tætheds funktion, da den er identiske med $\tilde{\mathcal{D}}$ næsten overalt.

5.2 Netgenerering i to dimensioner

Vi skal fremover antage, at domænet er en delmængde af \mathbb{R}^2 , samt at dette kan afbildes bijektivt på et passende antal enheds kvadrater $[0, 1]^2$. For at forenkle notationen skal vi som hovedregel kun betragte et enkelt generisk enheds kvadrat. Elementerne i de endelige element net $\mathcal{T}_N = \{\square_j\}_{j=1}^N$, som vi skal arbejde med består udelukkende af rektangulære elementer, se figur 5.2 for et eksempel og notation. Vi tillader altså her mere generelle elementer end i sætning 5.1.16 fra forrige afsnit. Netinddelings algoritmen, som vi skal betragte, kræver, at vi har en a priori fejligning på formen

$$\|e\|_{\square_j}^q = \phi_1^q h_{1,j}^{\gamma+1} h_{2,j} + \phi_2^q h_{1,j} h_{2,j}^{\gamma+1} + \text{h. o. l.} \quad \text{for } j = 1, \dots, N. \tag{5.8}$$

I ligningen (5.8) er \square_j de rektangulære elementer, som det endelige element net over det afbildede domæne består af. Sidelængderne af \square_j betegnes med $h_{1,j}$ og $h_{2,j}$ i x -retningen hhv. y -retningen. Vi skriver h. o. l. for højere ordens led i $h_{1,j}$ og $h_{2,j}$. Funktionerne ϕ_1 og ϕ_2 er kendte men ikke-beregnelige funktioner, der afhænger af den funktion, vi approksimerer, men er uafhængige



Figur 5.2: Net og element notation

af netinddelingen. Kendes den funktion, som vi forsøger at approksimere, er ϕ_1 og ϕ_2 beregnelige. Endelig er q et positivt heltal, og γ er en reel parameter afhængig af polynomiums graden p (på de polynomier som vi approksimerer med) og normen $\|\cdot\|$. Det er underforstået, at den globale fejl opfylder følgende summations princip

$$\|e\|_{\Omega}^q = \sum_{j=1}^N \|e\|_{\square_j}^q. \quad (5.9)$$

I special tilfældet hvor alle elementer er kvadratiske, reduceres fejlligningen (5.8) til $\|e\|_{\square_j}^q = 2\phi^q h_j^{\gamma+2} + \text{h. o. l. for } j = 1, \dots, N$, hvor $2\phi^q = \phi_1^q + \phi_2^q$. Vi så i kapitel 4, at det er muligt at udlede en a priori fejlligning opfyldende ovenstående krav.

5.2.1 Konstruktion af næsten optimale net størrelses funktioner for netinddelinger med rektangulære elementer

Betragt nu en følge af net $\{\mathcal{T}_N\}_{N=1}^{\infty}$, og lad $\mathcal{T}_N = \{\square_j\}_{j=1}^N$ være et fast generisk net i følgen svarende til et fast valgt N . For dette net kan vi opfatte længden af siderne $h_{1,j}$ og $h_{2,j}$ i elementet \square_j som værdien af stykkevis konstante, ikke-negative (net størrelses) funktioner h_1 og h_2 . Det er oplagt, at disse opfylder

$$\int_{\Omega} \frac{1}{N h_1 h_2} dx dy = \frac{1}{N} \sum_{j=1}^N \int_{\square_j} \frac{1}{h_1 h_2} dx dy = 1. \quad (5.10)$$

Fjernes betingelsen “stykkevis konstant” fås følgende generelle definition af net størrelses funktioner.

5.2.1 DEFINITION For et givet endeligt element net $\mathcal{T}_N = \{\square_j\}_{j=1}^N$ defineres net størrelse funktioner h_1 og h_2 som værende to ikke-negative, næsten overalt glatte funktioner opfyldende (5.10) og

$$\int_{\square_j} \frac{1}{h_1} dx dy = h_{2,j} \quad \text{og} \quad \int_{\square_j} \frac{1}{h_2} dx dy = h_{1,j} \quad \text{for } j = 1, \dots, N. \quad (5.11)$$

5.2.2 BEMÆRKNING Definitionen af net størrelses funktioner kan umiddelbart generaliseres til højere dimensioner.

5.2.3 BEMÆRKNING Ifølge definition 5.1.5 er den næsten overalt glatte funktion $\mathcal{D} = \frac{1}{Nh_1h_2}$ en net tætheds funktion. I tilfældet med lutter kvadratiske elementer er $h_1 \equiv h_2 \equiv h$, hvorfor $\mathcal{D} = \frac{1}{Nh^2}$.

5.2.4 BEMÆRKNING Kravene i ligning (5.11) implicerer, at $h_i(x, y)$ op til højere ordens led er identisk med $h_{i,j}$ for ethvert punkt $(x, y) \in \square_j$ (og $i = 1, 2, j = 1, \dots, N$).

5.2.5 BEMÆRKNING Har vi givet to net størrelses funktioner h_1 og h_2 beskrivende en eller flere netinddelinger, kan vi ved brug af (5.11) rekonstruere disse netfølger.

I praktiske beregningssituationer forventer man ikke, at kunne konstruere net opfyldende (5.11). Man vil derfor normalt tillade approksimationer af typen

$$k \leq \frac{1}{h_{2,j}} \int_{\square_j} \frac{1}{h_1} dx dy \leq K \quad \text{og} \quad k \leq \frac{1}{h_{1,j}} \int_{\square_j} \frac{1}{h_2} dx dy \leq K, \quad (5.12)$$

for $j = 1, \dots, N$ og passende konstanter k og K med $k \leq K$.

Vi er interesseret i at konstruere net størrelses funktioner for asymptotisk optimale net følger $\{\hat{\mathcal{T}}_N\}_{N=1}^\infty$, hvor $\hat{\mathcal{T}}_N$ er det net med den mindst mulige fejl $\|\hat{e}\|_\Omega$ blandt alle net med N elementer. Fejlen er her beskrevet ved $\|\hat{e}\|_\Omega$ som fås fra (5.8) ved at udelade højere ordens led (bemærk at dette intet ændrer asymptotisk). Vi skal nu forklare, hvorledes man kan konstruere asymptotiske optimale net størrelse funktioner \hat{h}_1 og \hat{h}_2 (og dermed en asymptotisk optimal net tætheds funktion \mathcal{D}) hørende til $\hat{\mathcal{T}}_N$. Hertil skal vi benytte ligningerne (5.10) og (5.8), hvor vi i ligning (5.8), skal erstatte $h_{i,j}$ med h_i i hvert element \square_j . Vi skal desuden se bort fra højere ordens led.

Ved som sagt at erstatte $h_{i,j}$ med h_i i (5.8) og summere over $j = 1, \dots, N$ fås

$$\|\hat{e}\|_\Omega^q = \int_\Omega (\phi_1^q h_1^\gamma + \phi_2^q h_2^\gamma) dx dy. \quad (5.13)$$

Bemærk at vi benytter skrivemåden \hat{e} for at indikere, at vi ser bort fra højere ordens led. Arbejder vi med kvadratiske elementer fås istedet $\|\hat{e}\|_{\Omega}^q = \int_{\Omega} 2\phi^q h^{\gamma} dx dy$.

For at bestemme \hat{h}_1 og \hat{h}_2 skal vi minimere $\|\hat{e}\|_{\Omega}^q$ for fastholdt N , dvs vi skal løse følgende minimeringsproblem:

$$\text{Bestem de net størrelses funktioner } \hat{h}_1 \text{ og } \hat{h}_2 \quad (5.14)$$

$$\text{som minimerer udtrykket } \int_{\Omega} (\phi_1^q h_1^{\gamma} + \phi_2^q h_2^{\gamma}) dx dy$$

$$\text{under bibetingelse } \int_{\Omega} \frac{1}{N h_1 h_2} dx dy = 1, \text{ hvor } N \text{ er fast.}$$

Dette minimeringsproblem kan løses fx ved hjælp af Lagrange optimering som følger: Definer Lagrange funktionen L som

$$L(h_1, h_2, \lambda) = \int_{\Omega} (\phi_1^q h_1^{\gamma} + \phi_2^q h_2^{\gamma}) dx dy + \lambda \left(\int_{\Omega} \frac{1}{N h_1 h_2} dx dy - 1 \right). \quad (5.15)$$

De afledede af L med hensyn til h_1 og h_2 er

$$\begin{aligned} \delta L(h_1, h_2, \lambda) &= \int_{\Omega} \gamma (\phi_1^q h_1^{\gamma-1} \delta h_1 + \phi_2^q h_2^{\gamma-1} \delta h_2) dx dy \\ &\quad - \frac{\lambda}{N} \int_{\Omega} \left(\frac{1}{h_1^2 h_2} \delta h_1 + \frac{1}{h_1 h_2^2} \delta h_2 \right) dx dy. \end{aligned} \quad (5.16)$$

Sættes δL til 0 for vilkårlige variationer δh_1 og δh_2 fås følgende optimalitets betingelser

$$\phi_1^q \hat{h}_1^{\gamma+1} \hat{h}_2 = \frac{\lambda}{\gamma N} = C, \quad (5.17)$$

$$\phi_2^q \hat{h}_1 \hat{h}_2^{\gamma+1} = \frac{\lambda}{\gamma N} = C. \quad (5.18)$$

Hvor det sidste lighedstegn i begge ovenstående ligninger definerer den globale konstant C .

Ved nu at anvende ligningerne (5.8), (5.17) og (5.18) samt det faktum, at h_i og $h_{i,j}$ op til højere ordens led er identiske i \square_j fås

$$\begin{aligned} \|\hat{e}\|_{\square_j}^q &= \phi_1^q \hat{h}_1^{\gamma+1} \hat{h}_2 + \phi_2^q \hat{h}_1 \hat{h}_2^{\gamma+1} + \text{h. o. l.} \\ &= 2C + \text{h. o. l.} \quad \text{for } j = 1, \dots, N. \end{aligned} \quad (5.19)$$

Heraf ses, at den asymptotisk optimale fejl $\|\hat{e}\|_{\square}$ er uafhængig af elementet \square , dvs. fejlen er ækvidistribueret.

Værdien af konstanten C kan bestemmes fra bibetingelsen $\int_{\Omega} \frac{1}{N\hat{h}_1\hat{h}_2} dx dy = 1$ på følgende måde: multiplicer ligningerne (5.17) og (5.18), opløft til $1/(\gamma + 2)$ 'te potens, isoler $\frac{1}{N\hat{h}_1\hat{h}_2}$ på den ene side, integrer over Ω og sæt resultatet lig med en dvs.

$$\begin{aligned}
& (\phi_1\phi_2)^q \hat{h}_1^{\gamma+2} \hat{h}_2^{\gamma+2} = C^2 \\
& \Downarrow \\
& (\phi_1\phi_2)^{\frac{q}{\gamma+2}} \hat{h}_1 \hat{h}_2 = C^{\frac{2}{\gamma+2}} \\
& \Downarrow \\
& \frac{1}{N} C^{\frac{\gamma+2}{2}} \int_{\Omega} (\phi_1\phi_2)^{\frac{q}{\gamma+2}} dx dy = \int_{\Omega} \frac{1}{N\hat{h}_1\hat{h}_2} dx dy \\
& \Downarrow \\
& \frac{1}{N} C^{\frac{\gamma+2}{2}} \int_{\Omega} (\phi_1\phi_2)^{\frac{q}{\gamma+2}} dx dy = 1 \\
& \Downarrow \\
& C = \left(\frac{1}{N} \int_{\Omega} (\phi_1\phi_2)^{q/(\gamma+2)} dx dy \right)^{\frac{\gamma+2}{2}}.
\end{aligned}$$

Da konstanten C nu er kendt, kan vi ved hjælp af (5.17) og (5.18) bestemme \hat{h}_1 og \hat{h}_2 . Isoleres \hat{h}_2 i (5.17), og indsættes resultatet så i (5.18), kan vi bestemme \hat{h}_1 til

$$\begin{aligned}
\hat{h}_1 &= (C^{\gamma} \phi_1^{-q(\gamma+1)} \phi_2^q)^{\frac{1}{\gamma^2+2\gamma}} \\
&= \frac{\overline{C}}{\sqrt{N}} (\phi_1^{-q(\gamma+1)} \phi_2^q)^{\frac{1}{\gamma^2+2\gamma}}.
\end{aligned} \tag{5.20}$$

Her er konstanten \overline{C} uafhængig af nettet og givet ved

$$\overline{C} = \left(\int_{\Omega} (\phi_1\phi_2)^{\frac{q}{\gamma+2}} dx dy \right)^{\frac{1}{2}}. \tag{5.21}$$

På tilsvarende vis bestemmes \hat{h}_2

$$\begin{aligned}
\hat{h}_2 &= (C^{\gamma} \phi_1^q \phi_2^{-q(\gamma+1)})^{\frac{1}{\gamma^2+2\gamma}} \\
&= \frac{\overline{C}}{\sqrt{N}} (\phi_1^q \phi_2^{-q(\gamma+1)})^{\frac{1}{\gamma^2+2\gamma}}.
\end{aligned} \tag{5.22}$$

Er elementerne kvadratiske fås $\hat{h} = \frac{\overline{C}}{\sqrt{N}} \phi^{-q/(\gamma+2)}$ med $\overline{C} = \left(\int_{\Omega} \phi^{2q/(\gamma+2)} dx dy \right)^{1/2}$.

I begge situationer er minimums værdien for (5.14) givet ved

$$\int_{\Omega} (\phi_1^q \hat{h}_1^\gamma + \phi_2^q \hat{h}_2^\gamma) dx dy = \frac{2\overline{C}^{\gamma+2}}{N^{\frac{\gamma}{2}}}. \quad (5.23)$$

Endelig bemærkes at $\hat{h}_1\sqrt{N}$ og $\hat{h}_2\sqrt{N}$ er uafhængige af netinddelingen. Det betyder, at man alene ved at gemme de asymptotisk optimale net størrelses funktioner kan rekonstruere hele netfølgen.

5.2.2 Bestemmelse af det næsten optimale antal af elementer for rektangulære netinddelinger

I forrige afsnit så vi, hvordan man kan bestemme asymptotisk optimale net størrelses funktioner til beskrivelse af den asymptotisk optimale netinddeling. I dette afsnit skal vi se, hvordan vi bestemmer antallet af elementer i dette. I praktiske situationer er vi interesseret i tilfældet, hvor brugeren kræver en løsning med en fejl mindre end en vis given tolerance τ , dvs. brugeren kræver $\|e\|_{\Omega} \leq \tau$. Ved hjælp af \hat{h}_1 og \hat{h}_2 kan vi bestemme det minimale antal elementer \hat{N} , der kræves, for at den asymptotisk optimale netinddeling opfylder de specificerede krav til fejlen. Det kan gøres ved at løse følgende minimerings problem:

$$\begin{aligned} &\text{Bestem det mindste tal } \hat{N} \text{ som opfylder} \\ &\int_{\Omega} (\phi_1^q \hat{h}_1^\gamma + \phi_2^q \hat{h}_2^\gamma) dx dy \leq \tau^q. \end{aligned} \quad (5.24)$$

Er elementerne kvadratiske lyder minimerings problemet istedet: Bestem det mindste \hat{N} så $\int_{\Omega} 2\phi^q \hat{h}^\gamma dx dy \leq \tau^q$. Løsningen til dette problem findes let ved brug af formel (5.23) og er givet ved

$$\hat{N} = \left\lceil \left(\frac{2\overline{C}^{\gamma+2}}{\tau^q} \right)^{\frac{2}{\gamma}} \right\rceil. \quad (5.25)$$

For kvadratiske elementer fås samme formel, men her er $\overline{C} = (\int_{\Omega} \phi^{2q/(\gamma+2)} dx dy)^{1/2}$.

5.2.3 Netinddelings algoritmen

Vi er nu i stand til at udlede algoritmen til konstruktion af den optimale netinddeling. Dette vil blive gjort udfra de informationer, som vi har til rådighed fra de asymptotisk optimale net størrelses funktioner samt det asymptotisk optimale antal af elementer. Af notationsmæssige hensyn vil vi kun betragte

netinddelinger af et enkelt generisk kvadrat i rektangulære elementer. For at få en simpel netinddelings algoritme vil vi ikke tillade vilkårlige rektangler. Dette leder til følgende definition

5.2.6 DEFINITION *Mængden af tilladte netinddelinger (eng. the class of admissible meshes) defineres som følgende:*

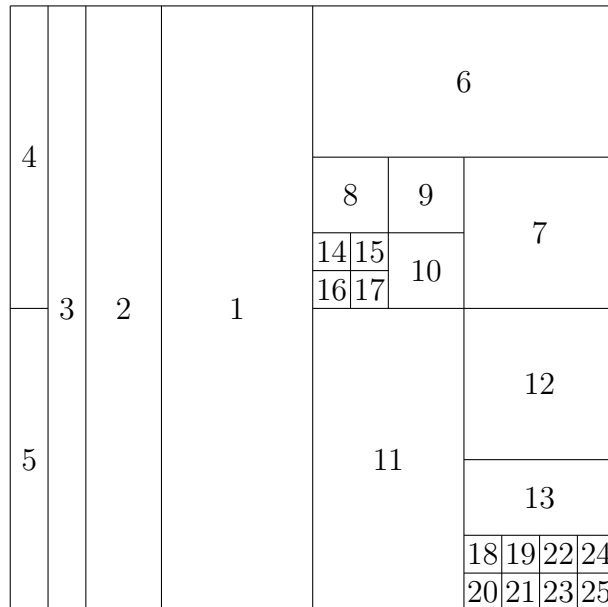
- *Enheds kvadratet er en tilladt netinddeling.*
- *Hvis \mathcal{T}' er konstrueret udfra en tilladt netinddeling \mathcal{T} ved at opdele et element i \mathcal{T} i to kongruente rektangler, er \mathcal{T}' en tilladt netinddeling.*
- *Ingen andre netinddelinger er tilladte.*

5.2.7 BEMÆRKNING I artiklen [8] er algoritmen testet på en række eksempler. I disse numeriske test har man arbejdet med endnu en betingelse i definitionen af tilladte netinddelinger. Der henvises til [8] for detaljer. Dette ekstra krav har ingen betydning for de teoretiske resultater, som vi skal vise i afsnit 5.2.4 og er udelukkende medtaget af implementationsmæssige årsager.

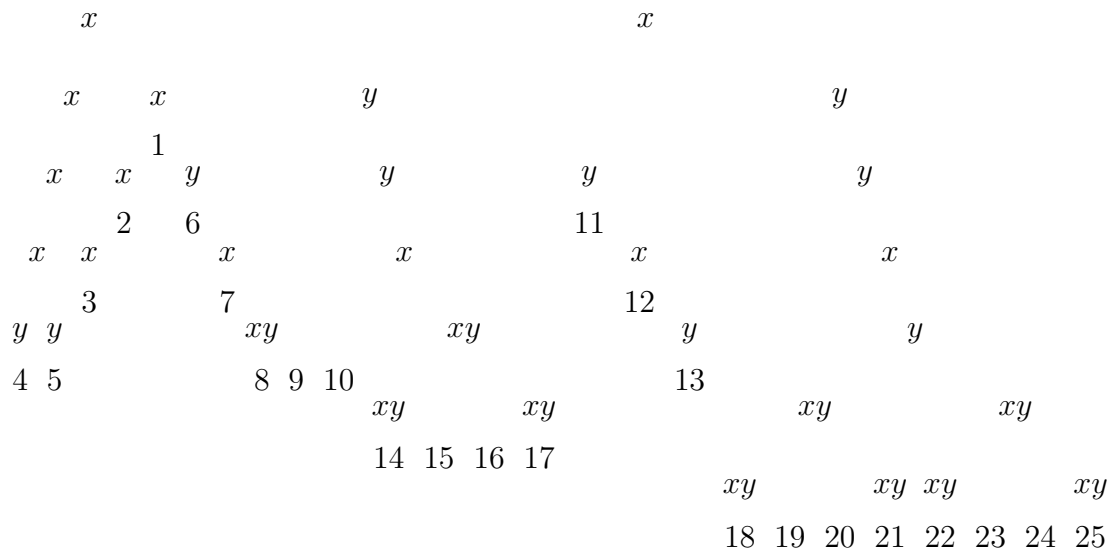
I diskussionen fremover vil det være hensigtsmæssigt at have fastlagt en datastruktur for tilladte netinddelinger. I figur 5.3 findes et eksempel på en tilladt netinddeling. Den tilhørende datastruktur kan repræsenteres ved et træ, se figur 5.4

Et vilkårligt element kan forfines på tre forskellige måder. Disse vil blive betegnet med x , y og xy inddeling. Figur 5.5 indeholder en forklaring af notationen. I forbindelse med forgrovnings vil vi kun betragte en bestemt type af søskende sæt med 2 eller 4 elementer. For disse 2 eller 4 elementer skal der gælde, at de udgør samtlige blade i en given gren i træet, og at denne gren ikke har nogle undergrene. Der findes tre muligheder for søskende sæt af ovennævnte type betegnet med hhv. type x , y og xy , se figur 5.5(a). For et konkret eksempel henvises til figur 5.3. Her udgør elementerne 4 og 5 et søskende sæt af type y , mens elementerne 8 og 9 ikke er et søskende sæt i ovenstående forstand. Den valgte datastruktur er ikke entydig, idet identiske netinddelinger kan have forskellig træ repræsentation afhængig af, hvordan de er opnået. Der henvises til [8] for et eksempel. Da diskussionen her er teoretisk har det ingen betydning.

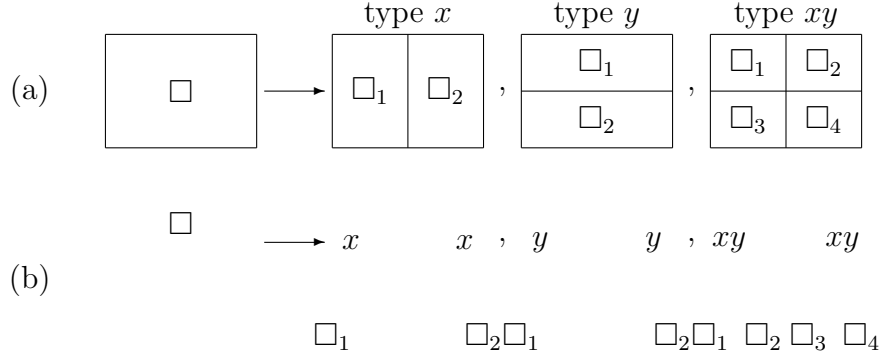
I overensstemmelse med ligning (5.11) vil netinddelings algoritmen blive ba-



Figur 5.3: Eksempel på tilladt netinddeling



Figur 5.4: Datastruktur for netinddelingen i figur 5.3



Figur 5.5: (a) Mulige forfininger af et element og (b) den resulterende data struktur

seret på funktionerne

$$\kappa_1(\square) = \frac{h_1}{\hat{h}_1(\bar{x}, \bar{y})} = \frac{1}{h_2} \int_{\square} \frac{1}{\hat{h}_1} dx dy + \text{h. o. l.}, \quad (5.26)$$

$$\kappa_2(\square) = \frac{h_2}{\hat{h}_2(\bar{x}, \bar{y})} = \frac{1}{h_1} \int_{\square} \frac{1}{\hat{h}_2} dx dy + \text{h. o. l.}, \quad (5.27)$$

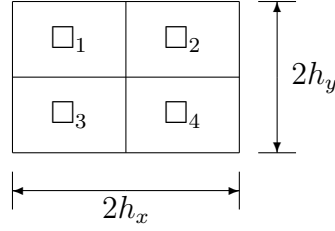
hvor \square er et vilkårligt element med sidelængde h_1 i x -retningen og sidelængde h_2 i y -retningen. De asymptotisk optimale net størrelses funktioner betegnes som sædvanligt med \hat{h}_1 og \hat{h}_2 . Punktet (\bar{x}, \bar{y}) er et vilkårligt punkt i \square . I de numeriske test i [8] er (\bar{x}, \bar{y}) valgt som centrum af \square .

I netinddelings algoritmen har vi brug for nedenstående simple observation. For et søskende sæt af type xy (med notation som i figur 5.6) gælder følgende relation

$$\begin{aligned} \kappa_1(\square_1 \cup \square_3) &= \frac{1}{2h_y} \int_{\square_1 \cup \square_3} \frac{1}{\hat{h}_1} dx dy + \text{h. o. l.} \\ &= \frac{1}{2} \left(\frac{1}{h_y} \int_{\square_1} \frac{1}{\hat{h}_1} dx dy + \frac{1}{h_y} \int_{\square_3} \frac{1}{\hat{h}_1} dx dy \right) + \text{h. o. l.} \\ &= \frac{1}{2} \left(\kappa_1(\square_1) + \kappa_1(\square_3) \right) + \text{h. o. l.} \end{aligned} \quad (5.28)$$

Tilsvarende ses nedenstående relationer at være opfyldt

$$\begin{aligned}
\kappa_1(\square_1 \cup \square_2) &= \kappa_1(\square_1) + \kappa_1(\square_2) + \text{h. o. l.} \\
\kappa_1(\square_3 \cup \square_4) &= \kappa_1(\square_3) + \kappa_1(\square_4) + \text{h. o. l.} \\
\kappa_2(\square_1 \cup \square_3) &= \kappa_2(\square_1) + \kappa_2(\square_3) + \text{h. o. l.} \\
\kappa_2(\square_2 \cup \square_4) &= \kappa_2(\square_2) + \kappa_2(\square_4) + \text{h. o. l.} \\
\kappa_1(\square_1 \cup \square_2 \cup \square_3 \cup \square_4) &= \tfrac{1}{2}(\kappa_1(\square_1) + \kappa_1(\square_2) + \kappa_1(\square_3) + \kappa_1(\square_4)) + \text{h. o. l.} \\
\kappa_1(\square_1 \cup \square_3) &= \tfrac{1}{2}(\kappa_1(\square_1) + \kappa_1(\square_3)) + \text{h. o. l.} \\
\kappa_1(\square_2 \cup \square_4) &= \tfrac{1}{2}(\kappa_1(\square_2) + \kappa_1(\square_4)) + \text{h. o. l.} \\
\kappa_2(\square_1 \cup \square_2 \cup \square_3 \cup \square_4) &= \tfrac{1}{2}(\kappa_2(\square_1) + \kappa_2(\square_2) + \kappa_2(\square_3) + \kappa_2(\square_4)) + \text{h. o. l.} \\
\kappa_2(\square_1 \cup \square_2) &= \tfrac{1}{2}(\kappa_2(\square_1) + \kappa_2(\square_2)) + \text{h. o. l.} \\
\kappa_2(\square_3 \cup \square_4) &= \tfrac{1}{2}(\kappa_2(\square_3) + \kappa_2(\square_4)) + \text{h. o. l.}
\end{aligned} \tag{5.29}$$



Figur 5.6: Søskende sæt

Målet med netinddelings algoritmen er at konstruere en tilladt netinddeling, hvor $\kappa_1(\square) = \kappa_2(\square) = 1$ for alle elementer \square , dvs. alle elementer har den “rigtige” tæthed. Fra ligningerne (5.26) og (5.27) ses, at en forfining eller forgrovn i x -retningen (dvs. en forøgelse eller formindskning af h_1) forøger eller formindsker κ_1 uden af der (op til højere ordens led) ændres på κ_2 . Tilsvarende overvejelser gør sig gældende om forfining/forgrovn i y -retningen.

For at udlede en heuristik for netinddelings algoritmen skal vi se bort fra højere ordens led samt antage at net størrelse funktionerne \hat{h}_1 og \hat{h}_2 er konstante. Dette svarer til situationen, hvor net størrelses funktionerne (og dermed net tætheds funktionen) er glatte og uden singulariteter.

Netforfining

Som vist i figur 5.5(a) kan et element \square med sidelængder $h_1 \times h_2$ forfines på tre forskellige måder.

Er $1 < \kappa_1(\square) < 2$ skal h_1 formindskes, og er $\kappa_1(\square) > 2$, skal h_1 halveres. Lignende observation gør sig gældende for κ_2 og h_2 . Dette giver os følgende heuristik.

- Hvis $\kappa_1(\square) > \beta$ halveres h_1
- Hvis $\kappa_2(\square) > \beta$ halveres h_2

hvor $1 \leq \beta \leq 2$. Sættes $\beta = 1$, er det muligt, at vi får en over-forfining af netinddelingen, idet vi så halverer h_1 (h_2) selvom $\kappa_1(\square)$ (hhv. $\kappa_2(\square)$) kun er en anelse større end 1. Tilsvarende kan $\beta = 2$ føre til en under-forfining, idet vi her kun halverer h_1 (hhv. h_2), når $\kappa_1(\square)$ (hhv. $\kappa_2(\square)$) er dobbelt så stor, som den burde være.

Netforgrovning

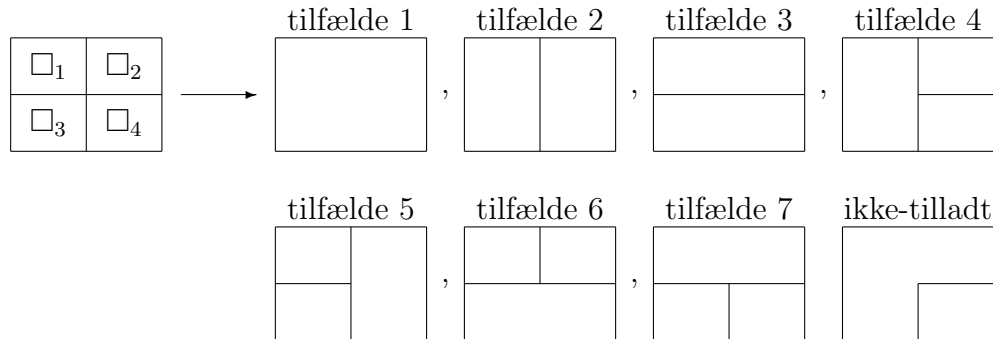
Som vi nu skal se, er situationen med netforgrovning en anelse mere kompliceret. Igen er der tre tilfælde svarende til søskende sæt af henholdsvis type x , y og xy . Notationen er som i figur 5.5(a).

Lad os starte med type x . Er $\kappa_1(\square_1 \cup \square_2) < 1$ slås \square_1 og \square_2 sammen til et element. Har vi derimod $\kappa_1(\square_1) > 1$ og $\kappa_1(\square_2) > 1$ eller ækvivalent $\kappa_1(\square_1) + \kappa_1(\square_2) > 2$ (\hat{h}_1 er antaget konstant) gør vi ingenting. Ovenstående diskussion kan gentages for søskende sæt af type y , hvor κ_1 er erstattet med κ_2 . Denne diskussion kan sammenfattes i heuristikken

- For søskende sæt (\square_1, \square_2) af type x
Hvis $\kappa_1(\square_1) + \kappa_1(\square_2) < \alpha$ fordobles h_1
- For søskende sæt (\square_1, \square_2) af type y
Hvis $\kappa_2(\square_1) + \kappa_2(\square_2) < \alpha$ fordobles h_2

Her er $1 \leq \alpha \leq 2$. Vælges $\alpha = 1$, er det muligt med over-forfining, idet vi kun forgrover netinddelingen, når det er absolut nødvendigt. For $\alpha = 2$ kan vi risikere en under-forfining, idet vi i denne situation forgrover, selvom elementerne kun er en anelse for små.

Vi mangler nu kun søskende sæt af type xy . Dette er den mest komplicerede situation på grund af de mange muligheder for forgrovning. I figur 5.7 ses alle tilladte forgrovninger samt et eksempel på en ikke-tilladt netinddeling. Ved en tilladt forgrovning forstås en forgrovning, der resulterer i en tilladt netinddeling.



Figur 5.7: Tilladte forgrovninger samt et eksempel på en ikke-tilladt forgrovning

I tilfælde 1 slås alle 4 elementer i søskende sættet sammen til et element. Dette gøres når $\kappa_i(\square_1 \cup \square_2 \cup \square_3 \cup \square_4) < 1$ for $i = 1, 2$. Er $\kappa_i(\square_1)$, $\kappa_i(\square_2)$, $\kappa_i(\square_3)$, $\kappa_i(\square_4) > 1$ for $i = 1, 2$ foretages intet. Anvendes (5.29), samt det faktum at $\kappa_i(\square_j)$ er uafhængig af j for konstant \hat{h}_i , fås første del af netforgrovningsheuristikken for søskende sæt af type xy til

- For søskende sæt $(\square_1, \square_2, \square_3, \square_4)$ af type xy
Hvis $\sum_{j=1}^4 \kappa_i(\square_j) < 2\alpha$ for $i = 1, 2$ slå
da alle 4 elementer sammen til 1 element

Som ovenfor er $1 \leq \alpha \leq 2$.

I tilfældene 2–7 skal vi være mere påpasselige for at undgå ikke-tilladte forgrovninger. Ideen er her forgrove de elementer, hvor det er mest nødvendigt, og derefter forsætte med at forgrove så længe dette ikke giver ikke-tilladte netinddelinger.

Fra figur 5.6 ses, at eventuelle forgrovning i disse tilfælde kontrolleres af tallene $\kappa_1(\square_1) + \kappa_1(\square_2)$, $\kappa_1(\square_3) + \kappa_1(\square_4)$, $\kappa_2(\square_1) + \kappa_2(\square_3)$ og $\kappa_2(\square_2) + \kappa_2(\square_4)$. Vi definerer nu $\kappa^{(1)}$ som værende det største af disse fire tal, og $\kappa^{(2)}$ som værende det tal med samme indeks på κ som $\kappa^{(1)}$. Er $\kappa^{(1)} < \alpha$ slås de to involverede elementer sammen. Dette vil altid give et tilladt net. Er derimod $\kappa^{(1)} > \alpha$ lader vi de to involverede elementer være uændrede, og forgrovningsprocessen stoppes. Ellers fortsættes med $\kappa^{(2)}$. Har vi $\kappa^{(2)} < \alpha$ forgrover vi og ellers foretages intet. Dette giver ligeledes en tilladt netinddeling. Er $\kappa^{(1)}$ valgt som $\kappa_2(\square_1) + \kappa_2(\square_3)$ eller $\kappa_2(\square_2) + \kappa_2(\square_4)$ fås et af tilfældene 2, 4 eller 5. Er $\kappa^{(1)}$ derimod valgt som $\kappa_1(\square_1) + \kappa_1(\square_2)$ eller $\kappa_1(\square_3) + \kappa_1(\square_4)$ haves et af tilfældene 3, 6 eller 7. Dette giver den resterende del af heuristikken for

forgrovning af søskende sæt af type xy . Enhver yderligere forgrovning vil føre til en ikke-tilladt netinddeling (eller til tilfælde 1, hvilket allerede er dækket).

- For søskende sæt $(\square_1, \square_2, \square_3, \square_4)$ af type xy
 Lad $\kappa^{(1)}$ og $\kappa^{(2)}$ være som ovenfor, og antag
 at $\kappa^{(1)} = \kappa_{i_1}(\square_{j_1}) + \kappa_{i_1}(\square_{j_2})$ og $\kappa^{(2)} = \kappa_{i_1}(\square_{j_3}) + \kappa_{i_1}(\square_{j_4})$
 (Her er $i_1 \in \{1, 2\}$, $\{j_1, j_2, j_3, j_4\} = \{1, 2, 3, 4\}$).
 Hvis $\kappa^{(1)} < \alpha$ slå
 da \square_{j_1} og \square_{j_2} sammen til et element
 Hvis $\kappa^{(2)} < \alpha$ slå
 da \square_{j_3} og \square_{j_4} sammen til et element

Igen er $1 \leq \alpha \leq 2$. Som ovenfor kan man også her foretage betragtninger om over- og under-forfining afhængig af valget af α .

Selve algoritmen

Ovenfor har vi gennemgået alle tilladte netforfininger og -forgrovninger. Lad os sammenfatte resultaterne i følgende netinddelings algoritme. Som input har algoritmen et tilladt net \mathcal{T} , to net størrelses funktioner \hat{h}_1 og \hat{h}_2 samt \hat{N} , som er det optimale antal af elementer, algoritmen skal forsøge at inddele domænet i. Desuden skal værdien af de reelle parametre $\alpha, \beta \in [1, 2]$ specificeres.

Gentag følgende indtil det ikke er muligt
at foretage flere forgrovnings-

For alle søskende sæt gøres følgende:

For søskende sæt (\square_1, \square_2) af type x
Hvis $\kappa_1(\square_1) + \kappa_1(\square_2) < \alpha$ slås \square_1 og \square_2 sammen til et
element

For søskende sæt (\square_1, \square_2) af type y
Hvis $\kappa_2(\square_1) + \kappa_2(\square_2) < \alpha$ slås \square_1 og \square_2 sammen til et
element

For søskende sæt $(\square_1, \square_2, \square_3, \square_4)$ af type xy
Hvis $\sum_{j=1}^4 \kappa_i(\square_j) < 2\alpha$ for $i = 1, 2$
slå da alle 4 elementer sammen til 1 element

Ellers
Lad $\kappa^{(1)}$ og $\kappa^{(2)}$ være som ovenfor, og antag
at $\kappa^{(1)} = \kappa_{i_1}(\square_{j_1}) + \kappa_{i_1}(\square_{j_2})$ og $\kappa^{(2)} = \kappa_{i_1}(\square_{j_3}) + \kappa_{i_1}(\square_{j_4})$
(Her er $i_1 \in \{1, 2\}$, $\{j_1, j_2, j_3, j_4\} = \{1, 2, 3, 4\}$).
Hvis $\kappa^{(1)} < \alpha$
slå da \square_{j_1} og \square_{j_2} sammen til et element
Hvis $\kappa^{(2)} < \alpha$
slå da \square_{j_3} og \square_{j_4} sammen til et element

Algoritme 5.2: Forgrovnings delen af algoritmen

Gentag følgende indtil det ikke er muligt
at foretage flere forfininger:

For alle elementer \square i \mathcal{T}
Hvis $\kappa_1(\square) > \beta$ halveres h_1
Hvis $\kappa_2(\square) > \beta$ halveres h_2

Algoritme 5.3: Forfinings delen af algoritmen

Benyttes kvadratiske elementer, er netinddelingen (i lighed med (5.26) og (5.27)) kontrolleret af funktionen

$$\kappa(\square) = \frac{h}{\hat{h}(\bar{x}, \bar{y})} = \frac{1}{h} \int_{\square} \frac{1}{\hat{h}} dx dy + \text{h. o. l.} \quad (5.30)$$

I analogi med tidligere er målet her at opnå $\kappa(\square) = 1$ for alle \square i netinddelingen. Da netforfining og -forgrovnning skal give lovlige netinddelinger, er vi i

det kvadratiske tilfælde nødsaget til kun at bruge forfininger og forgrovnings af type xy , men ellers kan algoritmen ovenfor anvendes uændret.

5.2.4 Teoretiske resultater for algoritmen

I artiklen [8] findes omfattende numeriske test af algoritmen. Fra et matematisk synspunkt er det dog ikke tilfredsstillende blot at have gode praktiske erfaringer med en given algoritme. Vi skal derfor i dette afsnit redegøre for nogle teoretiske resultater vedrørende algoritmens virkemåde.

Lad \hat{N} være det teoretisk optimale antal elementer i netinddelingen, der skal til for at en given forudbestemt brugertolerance nås (se afsnit 5.2.2), og lad N være det faktiske antal elementer opnået ved at anvende algoritmen. Hvis $N > \hat{N}$ siger vi at algoritmen har resulteret i en over-forfining. Har vi omvendt $N < \hat{N}$ taler vi om en under-forfining. Den optimale situation er selvfølgelig at have $N = \hat{N}$, men dette ses sjældent i praksis. Det er derfor vigtigt at have nedre og øvre grænser for N/\hat{N} , hvilket nedenstående sætning blandt andet giver

5.2.8 SÆTNING (Over- og underforfining) *Antag at a priori fejlligningen (5.8) holder, og at summations princippet (5.9) gælder. Med \hat{h}_1 og \hat{h}_2 betegnes de asymptotisk optimale net størrelses funktioner givet ved ligningerne (5.20) og (5.22). Funktionerne \hat{h}_1 og \hat{h}_2 antages at have højest algebraisk singulariteter, at være kontinuerte samt at have begrænsede afledede. Lad \hat{N} være det minimale antal af elementer, der skal til for at opfylde kravet til den på forhånd givne fejltolerance. Den optimale net tætheds funktion betegnes så med $\hat{D} = 1/(\hat{N}\hat{h}_1\hat{h}_2)$. For ethvert rektangulært element \square med sidelængder h_1 og h_2 i hhv. x og y retningen i en tilladt netinddeling defineres $\kappa_1(\square) = h_1/\hat{h}_1(\bar{x}, \bar{y})$ og $\kappa_2(\square) = h_2/\hat{h}_2(\bar{x}, \bar{y})$ som i ligningerne (5.26) og (5.27). Punktet (\bar{x}, \bar{y}) er her et vilkårligt punkt i \square . Endelig er $\alpha, \beta \in [1, 2]$ de parametre, der anvendes i algoritmerne (5.2) hhv. (5.3), og N er det antal elementer algoritmerne resulterer i (med ovennævnte input). Da gælder følgende relationer*

$$\frac{\beta}{2} \leq \kappa_1(\square), \quad \kappa_2(\square) \leq \beta, \quad (\text{for } h_1, h_2 \rightarrow 0), \quad (5.31)$$

og

$$\frac{\beta^2}{4} \leq \hat{N} \int_{\square} \hat{D}(x, y) dx dy \leq \beta^2, \quad (\text{for } h_1, h_2 \rightarrow 0), \quad (5.32)$$

$$\frac{1}{4} \leq N \int_{\square} \hat{D}(x, y) dx dy \leq 4, \quad (\text{for } h_1, h_2 \rightarrow 0), \quad (5.33)$$

$$\frac{1}{\beta^2} \leq \frac{N}{\hat{N}} \leq \frac{4}{\beta^2} \quad (\text{for } h_1, h_2 \rightarrow 0). \quad (5.34)$$

5.2.9 BEMÆRKNING Arbejdes der med kvadratiske elementer \square med sidelængde h , har vi naturligvis kun en asymptotisk optimal netstørrelses funktion \hat{h} , en κ -funktion $\kappa(\square) = h/\hat{h}(\bar{x}, \bar{y})$, og den asymptotiske net tætheds funktion er så givet ved $\hat{D} = 1/(\hat{N}\hat{h}^2)$. Da gælder tilsvarende resultater til ovenstående sætning. Eneste ændring er i ulighed (5.31), som skal erstattes af

$$\frac{\beta}{2} \leq \kappa(\square) \leq \beta, \quad (\text{for } h \rightarrow 0). \quad (5.35)$$

BEVIS Der henvises til artiklen [23], hvor sætningen er vist for kvadratiske elementer.

Sætning 5.2.8 viser, at forholdet N/\hat{N} asymptotisk er under kontrol.

5.2.10 BEMÆRKNING Ligningen (5.31) svarer til ligningen (5.12) med $k = \beta/2$ og $K = \beta$.

Sammenhængen mellem algoritmen og den globale fejl er indeholdt i følgende sætning.

5.2.11 SÆTNING (*Den globale fejl*) Antag at forudsætningerne fra sætning 5.2.8 gælder, samt at notationen er uændret. Da findes følgende globale grænser for fejlen

$$(\beta/2)^\gamma \cdot \tau^q \leq \|e\|_\Omega^q \leq \beta^\gamma \cdot \tau^q, \quad (\text{for } h_1, h_2 \rightarrow 0). \quad (5.36)$$

I denne ligning er τ den brugerbestemte fejltolerance, der også optræder i formel (5.25) for \hat{N} .

BEVIS Fra ligningerne (5.8) og (5.12) findes

$$\begin{aligned} \|e\|_{\square_j}^q &= \phi_1^q h_{1,j}^{\gamma+1} h_{2,j} + \phi_2^q h_{1,j} h_{2,j}^{\gamma+1} \\ &= \phi_1^q \frac{h_{1,j}^\gamma}{h_1^\gamma} h_{1,j} h_{2,j} h_1^\gamma + \phi_2^q \frac{h_{2,j}^\gamma}{h_2^\gamma} h_{1,j} h_{2,j} h_2^\gamma \\ &\leq \phi_1^q h_1^\gamma h_{1,j} h_{2,j} K^\gamma + \phi_2^q h_2^\gamma h_{1,j} h_{2,j} K^\gamma \\ &= \int_{\square_j} (\phi_1^q h_1^\gamma + \phi_2^q h_2^\gamma) dx dy K^\gamma. \end{aligned} \quad (5.37)$$

Ved nu at anvende betingelse (5.24), bemærkning 5.2.10 samt summationsprincippet (5.9) fås sætningens ene ulighed. Den anden følger på tilsvarende vis.

Følgende lokale version af ovenstående sætning viser, at algoritmen genererer asymptotisk ækvilibrerede netinddelinger ud fra asymptotisk optimale net størrelses funktioner. Ved en ækvilibreret netinddeling forstås en netinddeling, hvor den lokale fejl er ens i alle elementerne i netinddelingen.

5.2.12 SÆTNING (*Den lokale fejl*) Antag at forudsætningerne fra sætning 5.2.8 gælder, samt at notationen er uændret. Da haves følgende lokale grænser for fejlen

$$(\beta/2)^\gamma \cdot 2C \leq \|e\|_\square^q \leq \beta^\gamma \cdot 2C, \quad (\text{for } h \rightarrow 0), \quad (5.38)$$

hvor værdien af konstanten C er givet ved formel (5.21).

BEVIS Den eneste forskel fra beviset for forrige sætning er, at man her anvender ligning (5.19) istedet for (5.12).

Kapitel 6

Fejlestimering

Når man omformer uendelig-dimensionelle problemer i form af differentiallyigninger til et sæt af endeligt mange ligninger med endeligt mange ubekendte, vil der uundgåeligt opstå nogle fejl. En stor del af forskningsindsatsen indenfor den endelige element metode består i at udvikle og dokumentere metoder til evaluering af den fejl, der opstår i denne proces. I dette kapitel skal vi se nærmere på en hjørne af fejlestimering.

6.1 Generel teori om fejlestimering

Før vi betragter en egentlig metode til fejlestimering, vil vi ganske kort gøre nogle generelle betragtninger omkring fejlestimering.

6.1.1 Beregningsmæssige- og modelleringsfejl

Den fejl, som man begår ved at omformulere et differentiallyignings problem til et variations problem, kan opdeles i

- beregningsmæssige fejl.
- modelleringsfejl.

For at konkretisere disse fejl typer skal vi betragte en matematiske model

$$\mathcal{D}u = f, \tag{6.1}$$

hvor \mathcal{D} er en differentialoperator (muligvis med begyndelsesværdi- eller randbetingelser) defineret på et vist domæne Ω , f en given funktion, og u er den

ukendte løsning til problemet. Den matematiske model består her af \mathcal{D} og f . Udgangspunktet for den matematiske model er typisk en fysisk model¹

$$\tilde{\mathcal{D}}\tilde{u} = \tilde{f}. \quad (6.2)$$

Her har $\tilde{\mathcal{D}}$, \tilde{u} og \tilde{f} analoge betydninger. Den fysiske model antages at have en entydig løsning, og vi skal desuden antage, at den matematiske model er konstrueret ud fra den fysiske model på en sådan måde, at der ligeledes for denne findes en entydig løsning. Lad nu u_{ex} være den eksakte løsning til (6.1), og lad u_{fe} være en approksimeret løsning til (6.1) beregnet ved hjælp af den endelige element metode. Vi vil da definere den beregningsmæssige fejl som

$$e_{\text{c}} = u_{\text{ex}} - u_{\text{fe}}. \quad (6.3)$$

Er \tilde{u}_{ph} løsningen til (6.2), defineres modelleringsfejlen som

$$e_{\text{m}} = \tilde{u}_{\text{ph}} - u_{\text{ex}}. \quad (6.4)$$

Den samlede fejl defineres da som

$$e = \tilde{u}_{\text{ph}} - u_{\text{fe}} = \tilde{u}_{\text{ph}} - u_{\text{ex}} + u_{\text{ex}} - u_{\text{fe}} = e_{\text{m}} + e_{\text{c}}, \quad (6.5)$$

dvs. som summen af modelleringsfejlen og den beregningsmæssige fejl. Vi skal her kun beskæftige os med beregningsmæssige fejl.

6.1.2 Forskellige typer af fejlestimatorer

Igennem årene er der blevet foreslået et utal af forskellige fejlestimatorer til brug i den endelige element metode, se fx [9] for litteraturhenvisninger. Fler-tallet af disse fejlestimatorer kan klassificeres i en af følgende to kategorier:

- Udglatnings metoder (eng. smoothening/flux-projection).
- Residuum metoder.

Vi skal i afsnit 6.2 betragte en bestemt udglatnings metode. Traditionelt set har residuum metoder haft størst udbredelse, men den udglatnings metode, som vi senere skal studere, vil måske medvirke til at åbne vejen for en større udbredelse af fejlestimatorer af denne type.

¹Ved en fysisk model forstås en model, der modellerer en situation hentet fra fx økonomi, kemi eller fysik.

6.1.3 Forskellige former for fejlestimering

Man skelner normalt mellem to former for fejlestimering

- A priori fejlestimering.
- A posteriori fejlestimering.

Ved a priori fejlestimering forsøger man at vurdere fejlen mellem den eksakte løsning og den approksimerede løsning alene på grundlag af egenskaberne ved den eksakte (ukendte) løsning til problemet. I a posteriori fejlestimering forsøger man derimod at vurdere fejlen vha. den beregnede approksimerede løsning. De to former for fejlestimering ses ofte anvendt i forskellige teoretiske og praktiske situationer. A priori fejlestimering bruges typisk til at bestemme konvergens egenskaberne for den numeriske metode, mens a posteriori fejlestimering ofte anvendes i adaptive algoritmer som en del af netinddelings algoritmen og stopkriteriet.

Specifikke egenskaber ved fejlestimatorer er tæt knyttet til den type af differentiaalligninger, som de er konstrueret i tilknytning til. Af denne grund findes kun meget få generelle teoretiske resultater om fejlestimatorer. Til gengæld findes en del begreber og definitioner, som man benytter, når man skal vurdere og/eller sammenligner en eller flere fejlestimatorer.

6.1.4 Definitioner og begreber

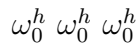
6.1.1 DEFINITION Ved en celle forstås en mindre samling af sammenhængende elementer, se figur 6.1. Bemærk at en celle godt kan bestå af et enkelt element.

6.1.2 DEFINITION Lad Ω betegne domænet, hvori vi ønsker at løse en given differentiaalligning, og lad $\mathcal{T} = \{\Delta_k\}_{k=1}^N$ være en triangulering af Ω . Ved en element fejl indikator for elementet $\Delta_k \in \mathcal{T}$ forstås en udtryk, der afhænger af u_{fe} men ikke af u_{ex} , og som vurderer fejlen i elementet Δ_k .

6.1.3 EKSEMPEL Betragtes model problemet (6.1) med endelig element løsning u_{fe} , er

$$\eta_{\Delta} = \text{meas}(\Delta) \|R_{\Delta}(u_{fe})\| \quad (6.6)$$

et meget simpelt eksempel på en element fejl indikator. Her er $R_{\Delta}(u_{fe}) = (\mathcal{D}u_{fe} - f)|_{\Delta}$ residualet restringeret til Δ , og $\text{meas}(\Delta)$ er arealet af Δ .



$$\omega_0^h \quad \omega_0^h \quad \omega_0^h$$

Figur 6.1: Eksempel på celle

6.1.4 DEFINITION Lad ω_0^h være en celle. Vi vil da – i analogi med summationsprincippet (5.9) for fejlen – definere en fejlestimator som

$$\mathcal{E}_{\omega_0^h} = \left(\sum_{\Delta \in \mathcal{T}} \eta_{\Delta}^q \right)^{1/q}. \quad (6.7)$$

Her er η_{Δ} en element fejl indikator for elementet Δ tilhørende trianguleringen \mathcal{T} af det betragtede domæne.

For at kunne vurdere hvor god en fejlestimator er, har vi brug for et mål, der definerer, hvad vi forstår ved en god fejlestimator. Et sådan mål er effektivitets indekset.

6.1.5 DEFINITION Ved effektivitets indekset for en fejlestimator $\mathcal{E}_{\omega_0^h}$ forstås

$$\kappa_{\omega_0^h} = \frac{\mathcal{E}_{\omega_0^h}}{\|e\|_{\omega_0^h}}, \quad (6.8)$$

hvor $\|\cdot\|_{\omega_0^h}$ er den norm, vi ønsker at vurdere forskellen på $e = u_{\text{ex}} - u_{\text{fe}}$ i.

Jo tættere $\kappa_{\omega_0^h}$ er på 1, jo bedre er fejlestimatoren $\mathcal{E}_{\omega_0^h}$.

Normalt forventer man ikke, at fejlen i den endelig element metode er specielt lille, såfremt netinddelingen er grov. Omvendt bør en god fejlestimator approksimere den eksakte fejl godt, såfremt netinddelingen er fin. På den baggrund skal vi nu definere, hvad vi forstår ved en asymptotisk eksakt fejlestimator.

6.1.6 DEFINITION Lad $\mathcal{E}_{\omega_0^h}$ være en fejlestimator. $\mathcal{E}_{\omega_0^h}$ siges at være asymptotisk eksakt, såfremt

$$\kappa_{\omega_0^h} \rightarrow 1 \quad \text{for } \|e\|_{\omega_0^h} \rightarrow 0 \quad (6.9)$$

for en given klasse af løsninger og en følge af netinddelinger.

6.2 Zienkiewicz–Zhu fejl estimatoren

Vi skal nu introducere en fejl estimator af udglatnings typen. Denne fejl estimator sammenkobler teori fra to områder indenfor den endelige element metode, som man normalt ikke forbinder med hinanden. Ved at sammenkoble metoder til beregning af punktværdier for de retningsafledede i den eksakte løsning ud fra den endelige element løsning med typiske ideer fra a posteriori fejl estimators, har Zienkiewicz og Zhu konstrueret en fejl estimator med pæne egenskaber. Fejl estimatoren kaldes ofte for Z^2 fejl estimatoren. Z^2 fejl estimatoren har gode konvergens egenskaber og er desuden relativ nem at implementere i allerede eksisterende endelig element software.

6.2.1 Model problem

Som nævnt ovenfor er fejl estimators ofte udviklet i forbindelse med en ganske bestemt differentia ligning. Dette er også tilfældet for Z^2 fejl estimatoren. Model problemet, vi skal betragte for denne fejl estimator, er Lamé–Navier elasticitets ligningerne i to dimensioner,

$$-(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) - \mu\Delta\mathbf{u} = \mathbf{f} \quad \text{i } \Omega \subset \mathbb{R}^2, \quad (6.10)$$

hvor \mathbf{u} er forskydnings vektoren (eng. displacement vector), f er kraften (eng. body force), og λ og μ er Lamé koefficienterne givet ved

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad (6.11)$$

og

$$\mu = \frac{E}{2(1+\nu)}. \quad (6.12)$$

I definitionen af Lamé koefficienter er E Youngs modul, og ν er Poissons forhold.

Randbetingelser for (6.10) er på en del af randen givet ved en forskydning, og på den resterende del ved en gnidningskraft (eng. tractions), dvs.

$$\mathbf{u} = \hat{\mathbf{u}} \quad \text{på } \Gamma_d, \quad (6.13)$$

$$\mathbf{H}\boldsymbol{\sigma} = \hat{\mathbf{t}} \quad \text{på } \Gamma_n, \quad (6.14)$$

hvor Γ_d og Γ_n opfylder $\Gamma = \partial\Omega = \Gamma_d \cup \Gamma_n$. Matricen \mathbf{H} er givet ved

$$\mathbf{H} = \begin{pmatrix} n_1 & 0 & n_2 \\ 0 & n_2 & n_1 \end{pmatrix}, \quad (6.15)$$

hvor $(n_1, n_2)^t$ er en udad rettet normalvektor til Γ_n , og stresset i model problemet er $\boldsymbol{\sigma} = \mathbf{D}\mathbf{S}\mathbf{u}$, hvor differentialoperatoren \mathbf{S} er defineret som

$$\mathbf{S} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix} \quad (6.16)$$

og matricen \mathbf{D} som

$$\mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1}{2}(1-2\nu) \end{pmatrix}. \quad (6.17)$$

Model problemet (6.10) kan nu formuleres som

$$\mathbf{S}^t \mathbf{D} \mathbf{S} \mathbf{u} = \mathbf{f}. \quad (6.18)$$

Indenfor elasticitets teori er \mathbf{D} den elasticitets matrix, som beskriver tøjning (eng. plain strain).

Variationsformuleringen af model problemet lyder: Bestem

$$\mathbf{u} \in \mathcal{H}_{\hat{\mathbf{u}}} : B(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \hat{\mathbf{t}}, \mathbf{v} \rangle_{\Gamma_n}, \quad \forall \mathbf{v} \in \mathcal{H}_0, \quad (6.19)$$

hvor

$$\mathcal{H}_{\hat{\mathbf{u}}} = \{\mathbf{u} \in [\mathcal{H}^1(\Omega)]^2 \mid \mathbf{u} = \hat{\mathbf{u}} \quad \text{på } \Gamma_d\}, \quad (6.20)$$

$$\mathcal{H}_0 = \{\mathbf{u} \in [\mathcal{H}^1(\Omega)]^2 \mid \mathbf{u} = 0 \quad \text{på } \Gamma_d\}, \quad (6.21)$$

$$B(\mathbf{u}, \mathbf{v}) = \int_{\Omega} (\mathbf{S}\mathbf{u})^t \mathbf{D} \mathbf{S} \mathbf{v} \, d\mathbf{x}, \quad (6.22)$$

$$(\mathbf{f}, \mathbf{v}) = \int_{\Omega} \mathbf{f}^t \mathbf{v} \, d\mathbf{x}, \quad (6.23)$$

$$\langle \hat{\mathbf{t}}, \mathbf{v} \rangle_{\Gamma_n} = \int_{\Gamma_n} \mathbf{v}^t \hat{\mathbf{t}} \, ds. \quad (6.24)$$

Ved hjælp af den endelige element metode fås en approksimeret løsning på formen

$$\mathbf{u}_h = \sum_{i=1}^n \mathbf{N}_i \mathbf{u}_i. \quad (6.25)$$

Her er \mathbf{N}_i basis funktioner konstrueret ud fra de lokale basis funktioner i de enkelte elementer. Koefficienterne \mathbf{u}_i er på standard vis bestemt fra kravet

$$B(\mathbf{u}_h, \mathbf{N}_i) = (\mathbf{f}, \mathbf{N}_i) + \langle \hat{\mathbf{t}}, \mathbf{N}_i \rangle_{\Gamma_n}, \quad i = 1, \dots, n. \quad (6.26)$$

Vi definerer nu fejlen i den approksimerede løsning som $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$. Målet er at opnå en fejl estimator med beregnelige øvre og nedre grænse for energi normen af fejlen. Per definition er stresset/gradienten i model problemet givet ved

$$\boldsymbol{\sigma} = \mathbf{D}\mathbf{S}\mathbf{u}. \quad (6.27)$$

Vi vil fremover bruge betegnelsen stress eller gradient i flæng. Det beregnelige stress er givet ved

$$\boldsymbol{\sigma}_h = \mathbf{D}\mathbf{S}\mathbf{u}_h. \quad (6.28)$$

Ved nu at bruge (6.27) og (6.28) fås følgende udtryk for energi normen for fejlen

$$\begin{aligned} \|\mathbf{e}\|^2 &= B(\mathbf{e}, \mathbf{e}) \\ &= \int_{\Omega} (\mathbf{S}\mathbf{e})^t \mathbf{D}\mathbf{S}\mathbf{e} \, d\mathbf{x} \\ &= \int_{\Omega} \mathbf{e}_{\sigma}^t \mathbf{D}^{-1} \mathbf{e}_{\sigma} \, d\mathbf{x}, \end{aligned} \quad (6.29)$$

hvor det sidste lighedstegn følger af definitionen $\mathbf{e}_{\sigma} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$. Vi har desuden udnyttet, at \mathbf{D} er symmetrisk. Ideen bag Z^2 fejl estimatoren er nu at angive en metode til at beregne en approksimation $\boldsymbol{\sigma}^*$ til $\boldsymbol{\sigma}$ og så anvende denne i (6.29). Altså vi erstatter $\boldsymbol{\sigma}$ med en approksimation $\boldsymbol{\sigma}^*$ og får

$$\mathbf{e}_{\sigma} \approx \mathbf{e}_{\sigma}^* = \boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h. \quad (6.30)$$

Dermed haves følgende fejl estimator

$$\|\mathbf{e}\| \approx \|\tilde{\mathbf{e}}\| = \left(\int_{\Omega} (\mathbf{e}_{\sigma}^*)^t \mathbf{D}^{-1} \mathbf{e}_{\sigma}^* \, d\mathbf{x} \right)^{1/2}. \quad (6.31)$$

6.2.2 Forskellige metoder til bestemmelse af σ^*

Vi skal studere fire forskellige metoder til bestemmelse af en approksimation σ^* til σ . Den første metode blev foreslået af Zienkiewicz og Zhu i artiklen [41], hvor også Z^2 fejl estimatoren første gang blev præsenteret. I [1] præsenteres to andre metoder til bestemmelse af σ^* . Den første af disse er en generalisering af metoden fra [41]. For de tre ovennævnte metoder findes i [1] en teoretisk og numerisk analyse af deres egenskaber. Vi skal nedenfor vende tilbage til disse resultater. Den fjerde metode, som vi skal betragte, blev indført i [42], hvori man også kan finde en række numeriske test, der berettiger dens eksistens overfor de tre tidligere fremgangsmåder. Der findes endnu ingen publicerede tilbunds gående teoretiske analyser af den sidstnævnte metode.

De oprindelige metoder

Den første metode, som vi skal betragte, blev indført af Zienkiewicz og Zhu i [41]. Fra (6.28) ser vi, at selvom den endelige element løsning er \mathcal{C}^0 kontinuert, vil den beregnelige stress σ_h ikke nødvendigvis være kontinuert over randene mellem de enkelte elementer. Den intuitive ide er nu at approksimere σ med en mere glat (kontinuert) stress σ^* .

Lad $\hat{\mathbf{N}}_i$, $i = 1, \dots, m$ være kontinuerte, globale basisfunktioner konstrueret (udfra de lokale basisfunktioner i hvert element) på en sådan måde, at de opfylder randbetingelsen (6.14) for stresset. Den udglattede stress σ^* defineres da som

$$\sigma^* = \sum_{i=1}^m \hat{\mathbf{N}}_i \bar{\sigma}_i^* , \quad (6.32)$$

hvor koefficienterne opfylder

$$\int_{\Omega} (\sigma^* - \sigma_h)^t \hat{\mathbf{N}}_i d\mathbf{x} = 0, \quad i = 1, \dots, m. \quad (6.33)$$

Da $\hat{\mathbf{N}}_i \in \mathcal{C}^0(\Omega)$ vil σ^* være kontinuert (også over randene mellem de enkelte elementer). Basisfunktionerne $\hat{\mathbf{N}}_i$ antages at have samme grad p som basisfunktionerne \mathbf{N}_i .

Ved at erstatte basisfunktionerne $\hat{\mathbf{N}}_i$ med basisfunktionerne \mathbf{M}_i , hvor \mathbf{M}_i opfylder de samme randbetingelser og globale kontinuitetsbetingelser som $\hat{\mathbf{N}}_i$, men hvor graden af \mathbf{M}_i er $q \neq p$, fås en anden metode til bestemmelse af en approksimation til σ . Da denne metode er en generalisering af den ovenfor nævnte, vil vi benytte samme notation og skrive

$$\sigma^* = \sum_{i=1}^m \mathbf{M}_i \bar{\sigma}_i^* , \quad (6.34)$$

hvor koefficienterne opfylder

$$\int_{\Omega} (\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h)^t \mathbf{M}_i d\mathbf{x} = 0, \quad i = 1, \dots, m. \quad (6.35)$$

Ændres kravet (6.35) til koefficienterne for $\boldsymbol{\sigma}^*$, kan man få en tredje metode til bestemmelse af en approksimation. En sådan approksimation er givet ved

$$\boldsymbol{\sigma}^D = \sum_{i=1}^m \mathbf{M}_i \bar{\boldsymbol{\sigma}}_i^D, \quad (6.36)$$

hvor koefficienterne nu er givet ved hjælp af et vægtet indre produkt

$$\int_{\Omega} (\boldsymbol{\sigma}^D - \boldsymbol{\sigma}_h)^t \mathbf{D}^{-1} \mathbf{M}_i d\mathbf{x} = 0, \quad i = 1, \dots, m. \quad (6.37)$$

Den nye metode

Ideen bag den nye metode er den samme, som den bag de oprindelige metoder. Vi ønsker at bestemme en approksimation $\boldsymbol{\sigma}^*$ til $\boldsymbol{\sigma}$, der er kontinuert over randene mellem de enkelte elementer. Som tidligere skal vi skrive $\boldsymbol{\sigma}^*$ på formen

$$\boldsymbol{\sigma}^* = \mathbf{N} \bar{\boldsymbol{\sigma}}^*, \quad (6.38)$$

hvor \mathbf{N} er de samme basisfunktioner, som dem vi brugte til at bestemme den endelige element løsning \mathbf{u}_h med.

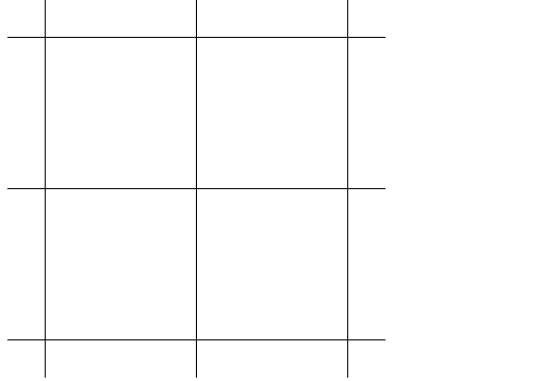
Vi skal nu redegøre for, hvordan koefficienterne $\bar{\boldsymbol{\sigma}}^*$ i (6.38) bestemmes. Koefficienterne $\bar{\boldsymbol{\sigma}}^*$ vil blive bestemt som værdien af et vist polynomium $\boldsymbol{\sigma}_p^*$ med samme orden p som basisfunktionerne i \mathbf{N} . Polynomiet $\boldsymbol{\sigma}_p^*$ vil blive defineret på et bælte af elementer omkring det knudepunkt, hvor vi er interesseret i at bestemme en approksimeret værdi af stresset i . Fremover skal vi kalde dette knudepunkt for samlingspunkts knuden (eng. patch assembly point). “Bæltet” af elementer omkring samlingspunkts knuden består af foreningsmængden af de elementer, der indeholder den betragtede samlingspunkts knude. I figur 6.2 er vist et par eksempler på typiske element bælte.

Polynomiet $\boldsymbol{\sigma}_p^*$ kan altså skrives som

$$\boldsymbol{\sigma}_p^* = \mathbf{P} \mathbf{a}, \quad (6.39)$$

hvor \mathbf{P} består af polynomielle led, og \mathbf{a} er de (indtil videre) ubekendte koefficienter til \mathbf{P} . For endimensionelle elementer har vi fx

$$\mathbf{P} = [1, x, x^2, \dots, x^p] \quad (6.40)$$



Figur 6.2: Eksempler på element bælte omkring en samlingspunkts knude markeret med •

og

$$\mathbf{a} = [a_1, a_2, \dots, a_{p+1}]. \quad (6.41)$$

For lineære polynomier i to dimensioner haves

$$\mathbf{P} = [1, x, y], \quad (6.42)$$

og for bi-lineære polynomier er

$$\mathbf{P} = [1, x, y, xy]. \quad (6.43)$$

For kvadratiske polynomier i to dimensioner er

$$\mathbf{P} = [1, x, y, x^2, xy, y^2], \quad (6.44)$$

og for bi-kvadratiske polynomier er

$$\mathbf{P} = [1, x, y, x^2, xy, y^2, x^2y, xy^2, x^2y^2]. \quad (6.45)$$

Endelig er de trunkerede bi-kvadratiske polynomier i to dimensioner bestemt ved

$$\mathbf{P} = [1, x, y, x^2, xy, y^2, x^2y, xy^2]. \quad (6.46)$$

Bestemmelse af koefficienterne \mathbf{a}

Der findes flere metoder til at bestemme koefficienterne \mathbf{a} i udtrykket (6.39). Artiklen [42] foreslår, at disse bestemmes som det sæt af koefficienter, der

minimerer funktionen F givet ved

$$\begin{aligned} F(x) &= \sum_{i=1}^n (\boldsymbol{\sigma}_h(x_i, y_i) - \boldsymbol{\sigma}_p^*(x_i, y_i))^2 \\ &= \sum_{i=1}^n (\boldsymbol{\sigma}_h(x_i, y_i) - \mathbf{P}(x_i, y_i)\mathbf{a})^2, \end{aligned} \quad (6.47)$$

hvor (x_i, y_i) er koordinaterne til en række samplings punkter, $n = mk$ er det samlede antal samplings punkter, og k er antallet af samplings punkter i elementet Ω_{m_j} , $m_j = 1, \dots, m$ i element bæltest $\Omega_s = \cup_{j=1}^m \Omega_{m_j}$. Igennem numeriske test konkluderes i [42], at approksimationen $\boldsymbol{\sigma}^*$ til $\boldsymbol{\sigma}$ bliver bedre, hvis man som samplings punkter vælger punkter, der er superkonvergente. Ved at differentiere F og derefter løse ligningen $F'(x) = 0$ ses, at \mathbf{a} er givet ved

$$\sum_{i=1}^n \mathbf{P}^t(x_i, y_i) \mathbf{P}(x_i, y_i) \mathbf{a} = \sum_{i=1}^n \mathbf{P}^t(x_i, y_i) \boldsymbol{\sigma}_h(x_i, y_i). \quad (6.48)$$

Ligningen (6.48) kan udtrykkes på matrix form som

$$\mathbf{a} = \mathbf{A}^{-1} \mathbf{b}, \quad (6.49)$$

hvor

$$\mathbf{A} = \sum_{i=1}^n \mathbf{P}^t(x_i, y_i) \mathbf{P}(x_i, y_i) \quad (6.50)$$

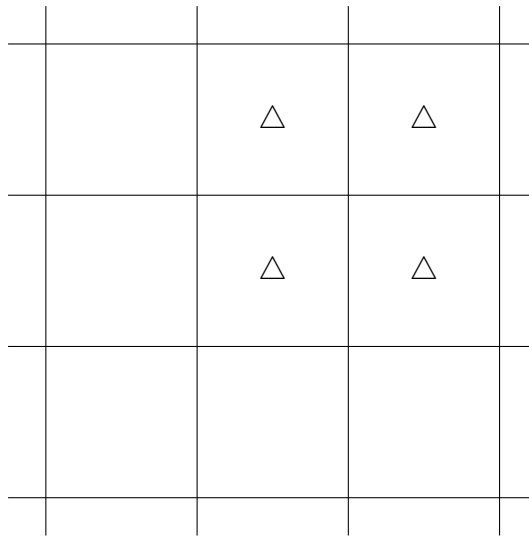
og

$$\mathbf{b} = \sum_{i=1}^n \mathbf{P}^t(x_i, y_i) \boldsymbol{\sigma}_h(x_i, y_i) \quad (6.51)$$

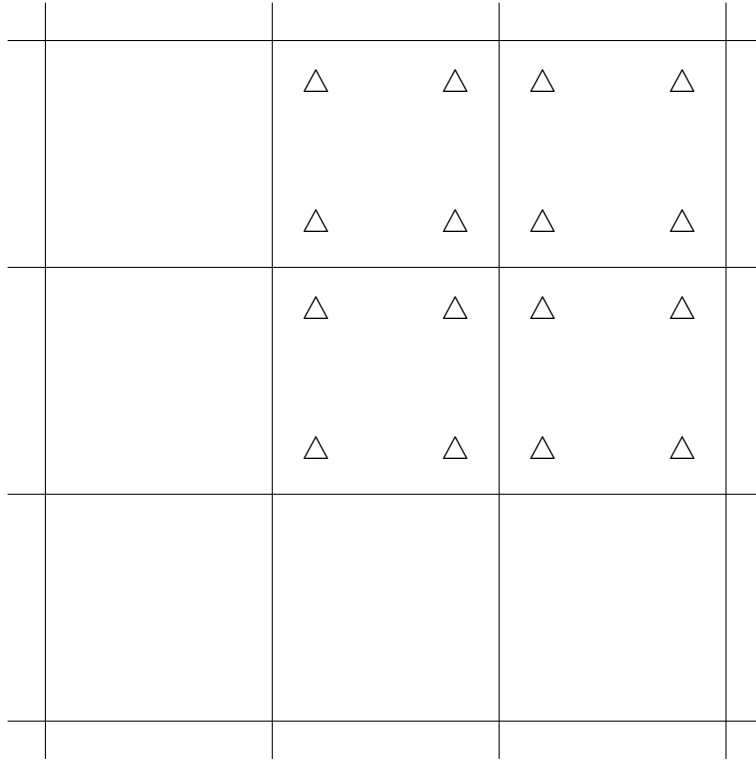
Når koefficienterne til \mathbf{a} er kendt, kan man bestemme approksimative punktværdier for stresset ved at indsætte passende koordinater i udtrykket for $\boldsymbol{\sigma}_p^*$. For at opnå de mest præcise resultater, vil vi kun bestemme værdien af $\boldsymbol{\sigma}^*$ i punkter beliggende i det indre af det betragtede element bælte. Nedenfor beskrives en undtagelse for punkter på randen af domænet. I figur 6.3 og 6.4 findes eksempler på ovenstående proces. I artiklen [42] underbygger numeriske test en formodning om, at værdierne af $\boldsymbol{\sigma}_p^*$ alle er superkonvergente.

Indre knudepunkter og randknudepunkter

For knudepunkter, der ligger i det indre af domænet, kan man risikere et vist overlap mellem element bælterne. For knudepunkter hvor man har et sådan overlap, vil der blive udregnet en approksimeret værdi for stresset



Figur 6.3: Eksempel på kvadratiske elementer med bi-lineære basisfunktioner. Punktet mærket \bigcirc er samlingspunkts knuden. De lokale basisfunktioner er bestemt ved værdien i punkterne mærket \circ/\bullet . Det ligeledes lineære polynomium σ_p^* er defineret over hele element bællet og bestemt ved værdien i Gauss punkterne mærket \triangle . I punktet mærket \bullet har vi udregnet en approksimeret værdi af stresset ved at indsætte passende koordinater i udtrykket for σ_p^*



Figur 6.4: Eksempel på kvadratiske elementer med trunkerede bi-kvadratiske basisfunktioner. Punktet mærket \bigcirc er samlingspunkts knuden. De lokale basisfunktioner er bestemt ved værdien i punkterne mærket \circ/\bullet . Det kvadratiske polynomium σ_p^* er defineret over hele element bæltet og bestemt ved værdien i Gauss punkterne mærket \triangle . I punkterne mærket \bullet har vi udregnet en approksimeret værdi af stresset ved at indsætte passende koordinater i udtrykket for σ_p^*

i det pågældende punkt fra to eller flere forskellige element bælter. Artiklen [42] foreslår, at man som den endelige approksimative værdi for stresset i sådanne punkter benytter et gennemsnit af de udregnede værdier. Dette foreslås udfra motiver om, at den ved hjælp af polynomiet σ_p^* udregnede approksimerede stress er superkonvergent. Det er derfor hurtigere/lettere at anvende ovenstående metode fremfor at benytte en metode, der undgår dette problem.

Der skal ligeledes tages specielt hensyn til randknodepunkter. For hjørnekoder vil det tilhørende element bælte ikke være stort nok (p^2 Gauss punkter) til entydigt at bestemme de $(p+1)^2$ koefficienter i \mathbf{a} . I sådanne tilfælde vil den approksimerende værdi af stresset blive udregnet for værdierne på randen for polynomiet σ_p^* fra et “indre” element bælte med samplingspunkts knude i det indre af domænet som vist i figur 6.5. For øvrige randknodepunkter og polynomier med grad større end eller lig 3 er situationen simplere, idet der i det tilhørende element bælte her vil være samplings punkter ($2p^2$) nok til at bestemme de $(p+1)^2$ koefficienter i \mathbf{a} . For polynomier med grad mindre end 3 går det galt, da man her kun har værdien i 8 Gauss punkter til at bestemme 9 koefficienter, se figur 6.6 for et eksempel. I stedet kan et fuldt indre element bælte benyttes som før hjørnepunkterne. I praksis vil denne udvej altid blive foretrukket også for polynomier med grad større end eller lig 3, da den sparer tid.

6.2.3 Teoretisk analyse af Z^2 fejl estimatoren

De oprindelige metoder

Som nævnt ovenfor er målet for Z^2 fejl estimatoren at opnå beregnelige øvre og nedre grænser for fejlen målt i energi normen. Den næste sætning giver os en øvre grænse.

6.2.1 SÆTNING *Antag at funktionen $\hat{\sigma} \in [\mathcal{H}^1(\Omega)]^2$ opfylder Neumann randbetingelsen (6.14), dvs.*

$$\mathbf{H}\hat{\sigma} = \hat{\mathbf{t}} \quad \text{på } \Gamma_n. \quad (6.52)$$

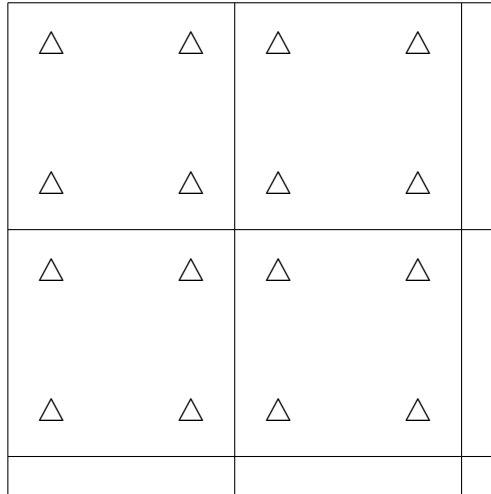
Lad $\alpha \in]0, 2[$ og $K > 0$ være reelle konstanter og definer funktionalen \hat{K} som

$$\hat{K}(\hat{\sigma}) = -\frac{1}{2} \int_{\Omega} (\hat{\sigma} - \sigma_h)^t \mathbf{D}^{-1}(\hat{\sigma} - \sigma_h) d\mathbf{x} - \frac{1}{2} \frac{h^\alpha}{K} \int_{\Omega} (\mathbf{f} + \mathbf{S}^t \hat{\sigma})^2 d\mathbf{x}. \quad (6.53)$$

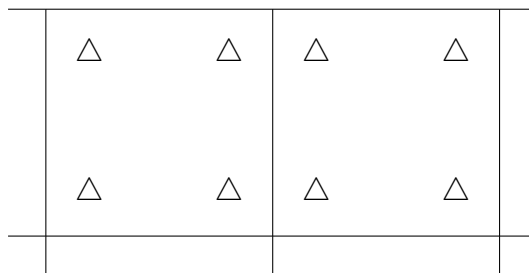
Da findes et h så

$$\|\mathbf{e}\|_E^2 \leq -2(1 + Ch^\delta) \hat{K}(\hat{\sigma}), \quad (6.54)$$

hvor C og δ er positive konstanter uafhængige af h .



Figur 6.5: Eksempel med kvadratiske elementer og bi-kvadratiske basisfunktioner. \triangle Gauss punkt. \bullet Approksimeret værdi af stresset. \bigcirc Samlingspunkts knude.



Figur 6.6: Eksempel med kvadratiske elementer og bi-kvadratiske basisfunktioner. \triangle Gauss punkt. \bullet Approksimeret værdi af stresset. \bigcirc Samlingspunkts knude.

BEVIS Se fx [1].

6.2.2 BEMÆRKNING Da basisfunktionerne \mathbf{N}_i er kontinuerte og opfylder randbetingelsen for stresset, vil det udglattede stress $\boldsymbol{\sigma}^*$ defineret ved (6.34) opfylde betingelserne i sætning 6.2.1.

6.2.3 BEMÆRKNING I lighed med forrige bemærkning ses, at den udglattede stress $\boldsymbol{\sigma}^D$ defineret ved (6.36) ligeledes opfylder betingelserne i sætning 6.2.1.

6.2.4 DEFINITION Lad \mathbf{D} være som i afsnit 6.2.1. For passende funktioner \mathbf{v} defineres $\|\cdot\|_{\mathbf{D}^{-1}}^2$ som

$$\|\mathbf{v}\|_{\mathbf{D}^{-1}}^2 = \int_{\Omega} \mathbf{v}^t \mathbf{D}^{-1} \mathbf{v} \, d\mathbf{x}. \quad (6.55)$$

6.2.5 BEMÆRKNING Det er ikke svært at vise, at $\|\cdot\|_{\mathbf{D}^{-1}}$ rent faktisk er en norm, og at man har følgende ækvivalens mellem normerne $\|\cdot\|_{\mathbf{D}^{-1}}$ og $\|\cdot\|_{\mathcal{L}^2(\Omega)}$.

$$\frac{(1+\nu)(1-2\nu)}{E} \|\mathbf{v}\|_{\mathcal{L}^2(\Omega)}^2 \leq \|\mathbf{v}\|_{\mathbf{D}^{-1}}^2 \leq \frac{2(1+\nu)}{E} \|\mathbf{v}\|_{\mathcal{L}^2(\Omega)}^2. \quad (6.56)$$

De næste to sætninger giver os nedre grænser for fejlen. Denne nedre grænse er afhængig af, om vi benytter approksimationen $\boldsymbol{\sigma}^*$ eller $\boldsymbol{\sigma}^D$ til $\boldsymbol{\sigma}$.

6.2.6 SÆTNING Lad $\boldsymbol{\sigma}^D$ være stressen defineret ved projektionen (6.36). Under antagelse af, at den eksakte stres $\boldsymbol{\sigma}$ er tilpas glat, gælder

$$\|\boldsymbol{\sigma}^D - \boldsymbol{\sigma}_h\|_{\mathbf{D}^{-1}} \leq (1 + C(\mathbf{u})h^{q-p+1}) \|\mathbf{e}\|_E. \quad (6.57)$$

BEVIS Se fx [1].

6.2.7 SÆTNING Lad $\boldsymbol{\sigma}^*$ være stressen defineret ved projektionen (6.34). Under antagelse af, at den eksakte stres $\boldsymbol{\sigma}$ er tilpas glat, gælder

$$\|\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\|_{\mathbf{D}^{-1}} \leq \sqrt{\frac{2}{1-2\nu}} (1 + C(\mathbf{u})h^{q-p+1}) \|\mathbf{e}\|_E. \quad (6.58)$$

BEVIS Se fx [1].

Sættes

$$\varepsilon^2 = \int_{\Omega} (\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_h)^t \mathbf{D}^{-1} (\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_h) \, d\mathbf{x}, \quad (6.59)$$

og

$$\Lambda^2 = \int_{\Omega} (\mathbf{f} + \mathbf{S}^t \boldsymbol{\sigma})^2 d\mathbf{x}, \quad (6.60)$$

hvor $\hat{\boldsymbol{\sigma}}$ betegner den udglattede stress opnået ved en af projektionerne (6.34) eller (6.36), kan vi sammenfatte resultaterne i sætningerne 6.2.1, 6.2.6 og 6.2.7 til

$$(1 + C_1(\mathbf{u})h^{q-p+1})^{-2} \varepsilon^2 \leq \|\mathbf{e}\|_E^2 \leq (1 + C_2(\mathbf{u})h^\delta) \left(\varepsilon^2 + \frac{h^\alpha}{K} \Lambda^2 \right) \quad (6.61)$$

for projektionen givet ved (6.36) og til

$$\frac{1-2\nu}{2} (1 + C_1(\mathbf{u})h^{q-p+1})^{-2} \varepsilon^2 \leq \|\mathbf{e}\|_E^2 \leq (1 + C_2(\mathbf{u})h^\delta) \left(\varepsilon^2 + \frac{h^\alpha}{K} \Lambda^2 \right) \quad (6.62)$$

for projektionen givet ved (6.34).

Den nye metode

Som nævnt ovenfor har man endnu ingen tilbundsgående teoretiske analyser af den nye metode til konstruktion af $\boldsymbol{\sigma}^*$. Nedenstående sætning angiver en betingelse, der giver os at $\|\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\| \approx \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|$ asymptotisk, dvs. asymptotisk er $\boldsymbol{\sigma}^*$ en god approksimation til $\boldsymbol{\sigma}$. Igennem numeriske test i [1] konkluderes, at sætningens forudsætninger er opfyldt for mange test problemer.

6.2.8 SÆTNING *Lad $\tilde{\mathbf{e}}_\sigma = \boldsymbol{\sigma} - \boldsymbol{\sigma}^*$ være fejlen, som vi begår ved at approksimere den eksakte stress $\boldsymbol{\sigma}$ med $\boldsymbol{\sigma}^*$, og lad $\mathbf{e}_\sigma = \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$ være forskellen mellem den eksakte stress og den beregnelige stress. Da er fejl estimatoren $\|\mathbf{e}_\sigma^*\| = \|\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\|$ asymptotisk eksakt hvis*

$$\frac{\|\tilde{\mathbf{e}}_\sigma\|}{\|\mathbf{e}_\sigma\|} \rightarrow 0 \quad \text{for} \quad \|\mathbf{e}_\sigma\| \rightarrow 0. \quad (6.63)$$

BEVIS Skrives fejl estimatoren $\|\mathbf{e}_\sigma^*\|$ som

$$\|\mathbf{e}_\sigma^*\| = \|\boldsymbol{\sigma}^* - \boldsymbol{\sigma}_h\| = \|(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - (\boldsymbol{\sigma} - \boldsymbol{\sigma}^*)\|, \quad (6.64)$$

fås ved brug af trekants uligheden

$$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| - \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^*\| \leq \|\mathbf{e}_\sigma^*\| \leq \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\| + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}^*\| \quad (6.65)$$

eller

$$\|\mathbf{e}_\sigma\| - \|\tilde{\mathbf{e}}_\sigma\| \leq \|\mathbf{e}_\sigma^*\| \leq \|\mathbf{e}_\sigma\| + \|\tilde{\mathbf{e}}_\sigma\| \quad (6.66)$$

Ved at dividere igennem med $\|\mathbf{e}_\sigma\|$ bliver (6.66) til

$$1 - \frac{\|\tilde{\mathbf{e}}_\sigma\|}{\|\mathbf{e}_\sigma\|} \leq \frac{\|\mathbf{e}_\sigma^*\|}{\|\mathbf{e}_\sigma\|} \leq 1 + \frac{\|\tilde{\mathbf{e}}_\sigma\|}{\|\mathbf{e}_\sigma\|}, \quad (6.67)$$

Når $\|\mathbf{e}_\sigma\| \rightarrow 0$ vil $\frac{\|\tilde{\mathbf{e}}_\sigma\|}{\|\mathbf{e}_\sigma\|} \rightarrow 0$ jvf. sætningens antagelse, og dermed vil $\frac{\|\mathbf{e}_\sigma^*\|}{\|\mathbf{e}_\sigma\|} \rightarrow 1$. Dette er ifølge definition 6.1.6 det samme som at sige, at $\|\mathbf{e}_\sigma^*\|$ er asymptotisk eksakt.

Litteratur

- [1] J. Z. Zhu, A. W. Craig, O. C. Zienkiewicz og M. Ainsworth. Analysis of the Zienkiewicz–Zhu A–Posteriori Error Estimator in the Finite Element Method. *Int. J. Numer. Methods in Engineering*, 28:2161–2174, 1989.
- [2] J.T. Oden og M. Ainsworth. A Procedure for A Posteriori Error Estimation for $h-p$ Finite Element Methods. *Computer Methods in Applied Mechanics and Engineering*, North-Holland, 101:73–96, 1992.
- [3] J.T. Oden og M. Ainsworth. A Posteriori Error Estimators for Second Order Elliptic Systems: Part 1. Theoretical Foundations and A Posteriori Error Analysis. *Computers Math. Applic.*, Vol. 25(No. 2):101–113, 1993.
- [4] J.T. Oden og M. Ainsworth. A Posteriori Error Estimators for Second Order Elliptic Systems: Part 2. An Optimal Order Process for Calculating Self-Equilibrating Fluxes. *Computers Math. Applic.*, Vol. 26(No. 9):75–87, 1993.
- [5] J.T. Oden og M. Ainsworth. A Unified Approach to A Posteriori Error Estimation using Element Residual Methods. *Numerische Mathematik, Springer-Verlag*, 65:23–50, 1993.
- [6] A. Miller og I. Babuška. A Feedback Finite Element Methode with A Posteriori Error Estimation: Part I. The Finite Element Methode and some Basic Properties of the A Posteriori Error Estimator. *Computer Methods in Applied Mechanics and Engineering*, 61:1–40, 1987.
- [7] Ivo Babuška. Courant Element: Before and After.
- [8] J. Hugger, T. Strouboulis, K. Copps, S. K. Gangaraj og I. Babuška. The Asymptotically Optimal Meshsize Function for $bi-p$ Degree Interpolation over Rectangular Element Subdivisions in Two Dimensions. Ikke-publiceret artikel, september 1994.

- [9] T. Strouboulis, C.S. Upadhyay og I. Babuška. A Model Study of the Quality of A-Posteriori Estimators for Linear Elliptic Problems. Part 1a: Error Estimation in the Interior of Patchwise Uniform Grids of Triangles. Technical Note BN-1147, Institute for Physical Science and Technology, University of Maryland, College Park , MD 20742, U.S.A, 1993.
- [10] T. Strouboulis, C.S. Upadhyay, S.K. Gangaraj, K. Copps og I. Babuška. Validation of A-Posteriori Error Estimators by Numerical Approach. Technical Note BN-1151, Institute for Physical Science and Technology, University of Maryland, College Park , MD 20742, U.S.A, 1993.
- [11] T. Strouboulis, C.S. Upadhyay, S.K. Gangaraj og I. Babuška. A Model Study of Element Residual Estimators for Linear Elliptic Problems: The Quality of the Estimators in the Interior of Meshes of Triangles and Quadrilaterals. Technical Note BN-1171, Institute for Physical Science and Technology, University of Maryland, College Park , MD 20742, U.S.A, 1993.
- [12] W. Gui og I. Babuška. Basic Principles of Feedback and Adaptive Approaches in the Finite Element Method. *Computer Methods in Applied Mechanics and Engineering*, 55:27–42, 1986.
- [13] A. Weiser og R.E. Bank. Some A Posteriori Error Estimators for Elliptic Partial Differential Equations. *Mathematics of Computations*, Vol. 44(Number 170):283–301, 1985.
- [14] R. E. Bank. *PLTMG : A Software Package for Solving Elliptic Partial Differential Equations. Users' Guide 7.0*. Nummer 15 i Frontiers in Applied Mathematics. SIAM, 1994.
- [15] R.E. Bank. Analysis of a Local A Posteriori Error Estimate for Elliptic Equations. I I. Babuška et. al., editor, *Accuracy Estimates and Adaptive Refinements in Finite Element Computations*, kapitel 7. Wiley, 1986.
- [16] Melvyn S. Berger. *Nonlinearity and Functional Analysis*. Academic Press, 1977.
- [17] Ivo Babuška og Qi Chen. Polynomial Interpolation of Real Functions. Interpolation in an Interval. Technical Note BN-1153, IPST, Inst. for Phys. Sci. and Tech., University of Maryland, College Park , MD 20742, U.S.A, 1993.
- [18] Ivo Babuška og Qi Chen. Polynomial Interpolation of Real Functions. Symmetrical Interpolation for the Triangle. Technical Note BN-1153,

- IPST, Inst. for Phys. Sci. and Tech., University of Maryland, College Park, MD 20742, U.S.A, 1993.
- [19] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company, 1978.
 - [20] R. Williams og D. Estep. *FEMLAB-ODE*, august 1994. README fil som følger med den elektroniske distribution af FEMLAB-ODE, version 3.2.
 - [21] Jens Hugger. Adaptive Recovery of Near Optimal Meshes in the Finite Element Method for Parameter Dependent Problems. *Computer Methods in Applied Mechanics and Engineering*, 101:127–141, 1992.
 - [22] Jens Hugger. Recovery and Few Parameter Representation of the Optimal Mesh Density Function for Near Optimal Finite Element Meshes. *Computer Methods in Applied Mechanics and Engineering*, 109:41–71, 1993.
 - [23] Jens Hugger. The Theory of Density Representation of Finite Element Meshes. Examples of Density Operators with Quadrilateral Elements in the Mapped Domain. *Computer Methods in Applied Mechanics and Engineering*, 109:17–39, 1993.
 - [24] Jens Hugger. An Asymptotically Exact, Pointwise, A Posteriori Error Estimator for Linear, Elliptic, One-dimensional Problems with Possible Singularities in the Solution. *Journal of Computational and Applied Mathematics*, 54:185–220, 1994.
 - [25] Jens Hugger. Recovery of the Optimal Mesh Density Function and Near Optimal Meshed with Rectangular Elements in the Mapped Domain. Pubblicazioni N. 915, Istituto di Analisi Numerica, Italy, 1994.
 - [26] Jens Hugger. An Asymptotically Exact, Pointwise, A Posteriori Error Estimator for the Finite Element Method with Super Convergence Properties. I I. Babuška, W.D. Henshaw, J.E. Flaherty, J.E. Oliger, J.E. Hopcroft og T. Tezduyar, editor, *Modelling, Mesh Generation, and Adaptive Numerical Methods for Partial Differential Equations*, volume 75 of “*IMA Volumes in Mathematics and its Applications*”, side 277–307. Springer-Verlag, 1995.
 - [27] Jens Hugger. The Elemental Equilibrium Method for Local, A Posteriori, Pointwise Estimation of the Finite Element Error. Ikke-publiceret artikel. Matematisk Institut, Københavns Universitet, 1995.

- [28] Jens Hugger. The Equilibrium Method for Local, A Posteriori, Pointwise Estimation of the Finite Element Error. Preprint Series, Nr. 2, Matematisk Institut, Københavns Universitet, 1995.
- [29] Jens Hugger. Local Polynomial Interpolation in a Rectangle. Ikke-publiceret artikel, marts 1995.
- [30] Jens Hugger. Numerisk Løsning af Differentialligninger, Forårsemesteret 1995. Noter til kursus af samme navn udbudt ved Københavns Universitet.
- [31] Claes Johnson. *Numerical Solutions of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1987.
- [32] Kenneth Erikson, Don Estep, Peter Hansbo og Claes Johnson. Introduction to Numerical Methods for Differential Equations. Department of Mathematics, Chalmers University of Technology, 1994.
- [33] D.W. Kelly. The Self Equilibration of Residuals and Complementary Error Estimates in the Finite Element Method. *International Journal of Methods in Engineering*, 20:1491–1506, 1984.
- [34] D. Leguillon og P. Ladeveze. Error Estimate Procedure in the Finite Element Method and Applications. *SIAM J. Numer. Anal.*, Vol. 20(No. 3):485–509, 1983.
- [35] W. C. Rheinboldt og C. K. Mesztenyi. *NFEARS, A Nonlinear Adaptive Finite Element Solver*. College Park , MD 20742, U.S.A.
- [36] R. A. Nicolaides. On a Class of Finite Elements Generated by Lagrange Interpolation. *SIAM J: Numer. Anal.*, 9(3):127–141, 1972.
- [37] J. Tinsley Oden. Historical Comments on Finite Elements. I Stephen G. Nash, editor, *A History of Scientific Computation*. ACM press, 1990. Conference on the History of Scientific and Numeric Computations held at Princeton University 1987.
- [38] Gert K. Pedersen. *Analysis Now*. Springer-Verlag, 1989.
- [39] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill Book Company, third edition, 1987.
- [40] Ivo Babuška og Barna Szabo. Lecture Notes on Finite Element Analysis, 2. Basic Mathematical Concepts, November 1982.

- [41] O. C. Zienkiewicz og J. Z. Zhu. A Simple Error Estimator and Adaptive Procedure for Practical Engineering Analysis. *Int. J. Numer. Methods in Engineering*, 24:337–357, 1987.
- [42] O. C. Zienkiewicz og J. Z. Zhu. The Superconvergent Patch Recovery and A Posteriori Error Estimates. Part 1: The Recovery Technique. *Int. J. Numer. Methods in Engineering*, 33:1331–1364, 1992.
- [43] O. C. Zienkiewicz og J. Z. Zhu. The Superconvergent Patch Recovery and A Posteriori Error Estimates. Part 2: Error Estimates and Adaptivity. *Int. J. Numer. Methods in Engineering*, 33:1365–1382, 1992.
- [44] O. C. Zienkiewicz. Finite Elements - the Background Story. I J. R. Whiteman, editor, *The Mathematics of Finite Elements and Applications*. Academic Press, 1972. Proceedings of the Brunel University Conference of the Institute of Mathematics and Its Applications.