

# Elementtimenetelmä

Ilari Angervuori

16.5.2018

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Science		Mathematics and Statistics	
Tekijä — Författare — Author Ilari Angervuori			
Työn nimi — Arbetets titel — Title Finite Element Method			
Oppiaine — Läroämne — Subject Mathematics			
Työn laji — Arbetets art — Level Pro gradu -thesis		Aika — Datum — Month and year May 2018	Sivumäärä — Sidoantal — Number of pages 74 p.
Tiivistelmä — Referat — Abstract <p>Purpose of the work is to give an elementary introduction to the Finite Element Method. We give an abstract mathematical formalization to the Finite Element Problem and work out how the method is a suitable method in approximating solutions of Partial Differential Equations. In Chapter 1 we give a concrete example code of Finite Element Method implementation with Matlab of a relatively simple problem. In Chapter 2 we give an abstract formulation to the problem. We introduce the necessary concepts in Functional Analysis. When Finite Element Method is interpreted in a suitable fashion, we can apply results of Functional Analysis in order to examine the properties of the solutions. We introduce the two equivalent formulations of weak formulation to differential equations: Galerkin's formulation and minimizing problem. In addition we define necessary concepts regarding to certain function spaces. For example we define one crucial complete inner product space, namely Sobolev space. In Chapter 3 we define the building blocks of the element space: meshing and the elements. Elements consists of their geometric shape and of basis functions and functionals on the basis functions. We also introduce the concepts of interpolation and construct basis functions in one, two and three dimensions. In Chapter 4 we introduce implementation techniques in a rather broad sense. We introduce the crucial concepts of stiffness matrix and load vector. We introduce a procedure for implementing Poisson's equation and Helmholtz's equation. We introduce one way of doing numerical integration by Gaussian quadrature points- and weights. We define the reference element and mathematical concepts relating to it. Reference element and Gaussian quadrature points are widely used techniques when implementing Finite Element Method with computer. In Chapter 5 we give a rigid analysis of convergence properties of Finite Element Method solution. We show that an arbitrary function in Sobolev space can be approximated arbitrarily close by a certain polynomial, namely Sobolev polynomial. The accuracy of the approximation depends on the size of the mesh and degree of the polynomial. Polynomial approximation theory in Sobolev spaces have a connection to Finite Element Methods approximation properties through Cèa's lemma. In Chapter 6 we give some examples of posteriori convergence properties. We compare Finite Element Method solution acquired with computer to the exact solution. Interesting convergence properties are found using linear- and cubic basis functions. Results seem to verify the properties acquired in Chapter 5.</p>			
Avainsanat — Nyckelord — Keywords Finite Element Method, $L^p$ -space, Poisson equation, Sobolev inequality, Helmholtz equation			
Säilytyspaikka — Förvaringsställe — Where deposited Helsinki University Library			
Muita tietoja — Övriga uppgifter — Additional information			

Tiedekunta/Osasto — Fakultet/Sektion — Faculty		Laitos — Institution — Department	
Matemaattis-luonnontieteellinen		Matematiikan ja tilastotieteen laitos	
Tekijä — Författare — Author Ilari Angervuori			
Työn nimi — Arbetets titel — Title Elementtimenetelmä			
Oppiaine — Läroämne — Subject Matematiikka			
Työn laji — Arbetets art — Level Pro gradu -tutkielma		Aika — Datum — Month and year Toukokuu 2018	Sivumäärä — Sidoantal — Number of pages 74 s.
Tiivistelmä — Referat — Abstract <p>Työssä esitellään elementtimenetelmä. Menetelmälle esitetään matemaattinen abstrakti muotoilu, sekä perustellaan matemaattisen analyysin avulla elementtimenetelmän toimivuus. Lisäksi työssä annetaan esimerkkejä menetelmän käytännön sovelluksista tietokoneelle ja tutkitaan suppenemisominaisuuksia täsmällisen ratkaisun suhteen. Luvussa 1 lähdetään liikkelle elementtimenetelmän yksinkertaisesta yksiulotteisesta sovelluksesta differentiaaliyhälöihin. Samalla esitellään menetelmän teoreettinen perusta. Luvussa annetaan toteutuksen Matlab koodi, jotta lukija pääsee vaivattomasti itse kokeilemaan menetelmän soveltamista. Luvussa 2 esitellään menetelmän matemaattinen abstrakti muotoilu. Luvussa esitellään tarvittavat funktioanalyysin käsitteitä ja tuloksia. Kun elementtimenetelmän tulkitsee sopivien funktioavaruuksien kautta, pääsemme soveltamaan siihen funktioanalyysin yleisesti tunnettuja tuloksia. Luvussa esitellään yhtäpitävät muotoilut differentiaaliyhälöiden heikolle muotoilulle: Galerkinin muotoilu ja minimointitehtävä. Lisäksi määritellään oleellisia vektoriavaruuksien ominaisuuksia ja vektoriavaruuksia, esimerkiksi eräs tärkeä täydellinen sisätuloavaruus: Sobolevin avaruus. Luvussa 3 määritellään elementtiavaruuden rakennusosat: ositus ja elementti. Elementit koostuvat paitsi geometrisesta muodosta, myös elementteihin määritellyistä kantafunktioista ja kantafunktioiden duaaliavaruudesta. Luvussa esitellään interpolaation käsite ja konstruoidaan eriasteisia kantafunktioita ensimmäisessä, toisessa ja kolmannessa ulottuvuudessa. Luvussa 4 esitellään suurpiirteisesti menetelmän toteutus tietokoneelle. Luvussa esitellään elementtikannalta oleelliset käsitteet: jäykkyysmatriisi ja kuormavektori. Luvussa käydään läpi menetelmän toteuttaminen Poissonin yhtälölle sekä Helmholtzin yhtälölle. Numeerinen integrointi Gaussin vertailupisteillä- ja painoilla esitellään. Lisäksi luvussa määritellään vertailuelementti ja käydään läpi siihen liittyviä matemaattisia käsitteitä. Vertailuelementti ja Gaussin vertailupisteet ovat tärkeitä käsitteitä menetelmän toteuttamisessa. Luvussa 5 perustellaan menetelmän toimivuus polynomisen approksimointiteorian kautta. Luvussa osoitetaan ensin, että Sobolevin avaruuksissa voidaan approksimoida mielivaltaisia funktioita mielivaltaisen tarkasti niin sanoituilla Sobolevin polynomeilla. Approksimoinnin tarkkuus riippuu elementtien koosta ja polynomien asteesta. Polynomisen approksimointiteorian tulokset liittyy elementtimenetelmän suppenemisominaisuuksiin Cèan lemman kautta. Luvussa 6 käsitellään posteriori suppenemisominaisuuksia. Luvussa vertailaan tietokoneella laskettua elementtimenetelmän ratkaisua tiedossa olevaan täsmälliseen ratkaisuun. Luvussa saadaan mielenkiintoisia suppenemisominaisuuksia ensimmäisen- ja kolmannen asteen kantafunktioilla. Tulokset tukevat luvussa 5 saatuja tuloksia.</p>			
Avainsanat — Nyckelord — Keywords Elementtimenetelmä, $L^p$ -avaruus, Poissonin yhtälö, Sobolevin epäyhtälö, Helmholtzin yhtälö			
Säilytyspaikka — Förvaringsställe — Where deposited Kumpulan tiedekirjasto			
Muita tietoja — Övriga uppgifter — Additional information			

# Sisältö

<b>1</b>	<b>Johdanto</b>	<b>4</b>
1.1	Yksinkertainen esimerkki . . . . .	5
<b>2</b>	<b>Abstrakti muotoilu</b>	<b>11</b>
2.1	Peruskäsitteet . . . . .	11
2.2	Tehtävän muotoilu . . . . .	15
2.3	Ratkaisujen olemassaolo . . . . .	18
<b>3</b>	<b>Elementtiavaruus</b>	<b>25</b>
3.1	Elementti . . . . .	26
3.2	Yleinen interpolaatio . . . . .	28
3.3	Kantafunktioiden konstruointi . . . . .	29
3.3.1	Yksiulotteisia elementtejä . . . . .	29
3.3.2	Lagrangen tyyppiset elementit . . . . .	32
<b>4</b>	<b>Tekninen toteutus</b>	<b>38</b>
4.1	Jäykkyysmatriisi ja kuormavektori Poissonin yhtälölle . . . . .	38
4.2	Jäykkyysmatriisi ja kuormavektori Helmholtzin yhtälölle . . . . .	40
4.3	Numeerinen integrointi . . . . .	41
4.4	Vertailuelementti . . . . .	43
4.4.1	Matemaattisia käsitteitä . . . . .	46
<b>5</b>	<b>A priori suppenemisominaisuudet</b>	<b>48</b>
5.1	Sobolevin kehitelmä . . . . .	49
5.2	Virhetermi . . . . .	52
5.3	Virhetermin suuruus . . . . .	57
5.4	Interpolaation virhe . . . . .	61
<b>6</b>	<b>Posteriori suppenemisominaisuuksia</b>	<b>70</b>

# Luku 1

## Johdanto

Elementtimenetelmä (Finite Element Method), mihin jatkossa viittaamme lyhyesti myös kirjaimilla ”FEM”, on eräs *osittaisdifferentiaaliyhtälöiden* ratkaisumenetelmä. Eräs FEMin sovellus voisi esimerkiksi kuvailla tehtävän

$$(1.1) \quad \begin{aligned} &\text{etsi } u \in W, \text{ jolla} \\ &Lu + au = f, \\ &\text{missä } L : W \rightarrow W \text{ on lineaarinen operaattori,} \\ &a \in \mathbb{C} \text{ ja } W \text{ on sopiva funktioavaruus,} \\ &\text{algebrallisena yhtälönä} \end{aligned}$$

$$(1.2) \quad \begin{aligned} &\text{etsi } u \in \mathbb{R}^n, \text{ jolla} \\ &Au = b, \\ &\text{missä } A \in \mathbb{R}^{n \times n} \text{ ja } b \in \mathbb{R}^n \text{ on annettuja.} \end{aligned}$$

Funktiot  $f, u : U \rightarrow \mathbb{C}$  on määritelty alueessa  $U \subset \mathbb{R}^n$ . Lisäksi määritellään reunaehdot joukon  $U$  reunalle  $\partial U$ . FEM on tietenkin likimääräinen menetelmä: jatkuva tehtävä (1.1) tulkitaan diskreettinä tehtävänä (1.2). Tämä seuraa olennaisesti alueen  $U$  jakamisesta äärellisiin osiin – siitä englannin kielinen nimi *Finite Element Method*. Kutsumme alueen  $U$  jakamista osiin *osittamiseksi* (meshing). Tulemme näyttämään, että elementtimenetelmä on toimiva menetelmä likimääräisen ratkaisun löytämiseksi siinä mielessä, että  $u$  tehtävässä (1.2) lähestyy tehtävän (1.1) ratkaisua, kun  $n$  lähestyy ääretöntä. Jälkimmäinen tarkoittaa käytännössä, että suurin elementin koko alueen  $U$  osittamisessa lähestyy nollaa. Vektori  $u \in \mathbb{R}^n$  tulee tulkita sopivasti, jotta voimme verrata sitä funktioon  $u \in W$ ; elementtimenetelmässä vektorin  $u$  koordinaatit ovat käytännössä elementtikohtaisia vakiota *interpoloitaville kantafunktioille*.

Kolme ydinkäsitettä liittyen elementtimetodin soveltamiseen ovat:

- Ratkaisujen olemassolo.
- FEM mallin teoreettinen rakentaminen.
- Toteutus tietokoneella.

## 1.1 Yksinkertainen esimerkki

Luvussa esitetään yksinkertainen esimerkki elementtimenetelmän toteutuksesta. Ratkaistaan reuna-arvot tehtävä

$$(1.3) \quad \begin{aligned} &\text{etsi } u \in C^2(0, 2\pi), \text{ jolla} \\ &u''(x) = \sin(x) \text{ alueessa } x \in [0, 2\pi], \\ &\text{reunaehdoilla } u(0) = 0, \quad u(2\pi) = 0, \end{aligned}$$

missä merkinnällä  $C^2(0, 2\pi)$  tarkoitetaan suljetulla välillä  $[0, 2\pi]$  jatkuvasti kaksi kertaa deirivoituvien funktioiden avaruutta. On helppo todeta, että tehtävän täsmällinen ratkaisu on funktio  $x \mapsto -\sin(x)$ . Tehtävän yhtälö (1.3) mallintaa tilanteita ainakin jäykkyyyslaskennassa ja lämmön siirtymisessä yksiulotteisissa ”tangoissa” ([8]).

Niin sanottu *heikko muotoilu* tehtävälle (1.3) saadaan kertomalla yhtälö puolittain luvulla  $f(x)$  ja laskemalla integraali koko alueen yli. Näin päädytään tehtävään

$$(1.4) \quad \begin{aligned} &\text{etsi } u \in C^2(0, 2\pi), \text{ jolla} \\ &\int_0^{2\pi} u''(x)f(x)dx = \int_0^{2\pi} \sin(x)f(x)dx \text{ kaikilla } f \in C^2(0, 2\pi), \\ &\text{reunaehdoilla } u(0) = 0, \quad u(2\pi) = 0. \end{aligned}$$

Muotoilua (1.4) kutsutaan *heikoksi muotoiluksi*. Sen motivaatio on, että tehtävän (1.3) ratkaisu toteuttaa myös heikon muotoilun.

Elementtimenetelmässä muotoilussa ratkaisua etsitään jatkuvien funktioiden aliavaruudesta  $C_h \subset C(0, 2\pi)$ , esimerkiksi jostakin paloittain lineaaristen funktioiden avaruudesta, jolloin päädytään muotoiluun

$$(1.5) \quad \begin{aligned} &\text{etsi } u \in C_h(0, 2\pi), \text{ jolla} \\ &\int_0^{2\pi} u''(x)f(x)dx = \int_0^{2\pi} \sin(x)f(x)dx \text{ kaikilla } f \in C_h(0, 2\pi), \\ &\text{reunaehdoilla } u(0) = 0, \quad u(2\pi) = 0. \end{aligned}$$

Kunhan avaruus  $C_h$  valitaan sopivasti, muotoilun (1.5) ratkaisu approksimoi muotoilun muotoilun (1.4) ratkaisua – tulemme osoittamaan tämän työssä hyvin yleisessä tapauksessa. On tärkeää huomata, että  $u \in C_h$  ei ole välttämättä kaikissa pisteissä derivoituva. Jättäen muodollisen käsittelyn myöhemmäksi, voidaan ajatella, että kyseiset pisteet jätetään huomioimatta, sillä integraali tehtävässä (1.5) ei riipu yksittäisistä pisteistä.

Elementtimenetelmän teoreettisen tarkastelun ytimessä on oletus, että avaruus  $C_h$  on *äärellisulotteinen*; tämä mahdollistaa diskreetin tehtävän muodostamisen. Määritellään avaruuden  $C_h$  kanta: olkoon  $\{\phi_1, \dots, \phi_n\}$  joukko jatkuvia funktioita, jotka on määritelty alueella  $U$ . Tällöin funktiot  $\{\phi_1, \dots, \phi_n\}$  *virittää* jonkin jatkuvien funktioiden avaruuden  $C_h$ . Toisin sanoen jokainen  $u \in C_h$  voidaan kirjoittaa muodossa

$$u = \sum_{i=1}^n c_i \phi_i,$$

missä  $c_i \in \mathbb{R}$ .

*Osittaisintegroimalla* yhtälö (1.5), ja merkitsemällä  $f' = \sum a_i \phi'_i$ , saamme tehtävän

etsi  $u \in C_h$ , jolle

$$\int_0^{2\pi} u'(x) \phi'_i(x) dx = - \int_0^{2\pi} \sin(x) \phi_i(x) dx \text{ kaikilla } i = 1, \dots, n,$$

missä reunaehdot otettiin huomioon osittaisintegroinnissa. Edelleen, merkitsemällä  $u = \sum u_j \phi_j$  saadaan

Etsi  $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ , jolle

$$\sum_{j=1}^n u_j \int_0^{2\pi} \phi'_j(x) \phi'_i(x) dx = - \int_0^{2\pi} \sin(x) \phi_i(x) dx \text{ kaikilla } i = 1, \dots, n.$$

Lineaarialgebran kielellä tehtävä voidaan esittää muodossa

etsi  $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ , jolle

(1.6)  $Au = b$ , missä

$$A_{ji} = \int_0^{2\pi} \phi'_j(x) \phi'_i(x) dx \text{ ja } b_i = - \int_0^{2\pi} \sin(x) \phi_i(x) dx.$$

Seuraavaksi muotoilemme konkreettisesti avaruuden  $C_h$  kantafunktiot. Periaateessa kantafunktiot voisi olla mitä vain, mutta usein käytetään *kompaktikantajaisia* funktioita. Esimerkin vuoksi käytämme *paloittain lineaarisia* ”telttafunktioita”. Kyseiset funktiot virittävät (erään) paloittain lineaaristen jatkuvien funktioiden avaruuden.

Jaetaan ensin väli  $[0, 2\pi]$  tasaisesti  $n - 1$  osaan  $n$  solmun välille. Olkoon

$$\{N_i\}_{i=1,\dots,n} = \left\{ \frac{2\pi(i-1)}{(n-1)} \right\}_{i=1,\dots,n}$$

solmupisteet. Määritetään kantafunktiot paloittain

$$\phi_i(x) = \begin{cases} 1, & \text{solmussa } N_i, \\ 0, & \text{kun } x \leq N_{i-1} \text{ ja } N_{i+1} \leq x, \\ \text{lineaarinen välissä,} \end{cases}$$

indekseille  $i = 1, \dots, n$ .

Geometrinen tarkastelu antaa suoraan derivaatoiksi

$$\phi'_i(x) = \begin{cases} (n-1)/2\pi, & \text{kun } N_{i-1} < x < N_i, \\ -(n-1)/2\pi, & \text{kun } N_i < x < N_{i+1}, \\ 0, & \text{muulloin,} \end{cases}$$

missä kulmapisteet on jätetty huomioimatta aiemmin mainitussa mielessä. FEM mallimme on nyt valmis, ja voimme koota matriisin  $A$  ja vektorin  $b$ . Matriisi  $A$  on symmetrinen neliömatriisi ja sillä on käänteismatriisi  $A^{-1}$ , joten voimme algebrallisesti ratkoa  $u = A^{-1}b$ . Lisäksi  $A$  is *harva* (sparse), sillä  $A_{ij}$  saa nollasta poikkeavia arvoja vain kun  $|i - j| \leq 1$ . Tämä ominaisuus tekee elementtimenetelmästä verrattain vähän muistitilaa vievän menetelmän, kunhan matriisi  $A$  säilötään sopivaan tietorakenteeseen; esimerkiksi matlabissa rutiini *sparse(A)* säilöo matriisin  $A$  pienempään muistitilaan.

Seuraava Matlab koodi ratkaisee esimerkkitehtävän.

```
%Vertailuvalin [0,1] Gaussin vertailupisteet -
%ja painot vektoreissa X ja W.
X= [0.0469; 0.2308;0.5000; 0.7692;0.9531];
W= [0.1185; 0.2394; 0.2844; 0.2393; 0.1185];

%Solmupisteiden maara.
n=6
%Kantafunktiot vertailuelementissa [0,1].
C_1 = @(x) x;
C_2 = @(x) 1-x;

%Vertailuelementtien derivaatat vertailuelementissa.
DC_1 = @(x) 1+ 0.*x;
DC_2 = @(x) -1 + 0.*x;
```



```

%Skaalatut derivaatat.
dc_1 = @(x) ((n-1)/(2*pi))*(1 + 0.*x);
dc_2 = @(x) ((n-1)/(2*pi))*(-1 + 0.*x);

%Kuvaus F: [0,1],k,n ==>
%[2*pi*(k-1)/(n-1), 2*pi*(k/(n-1))].
%Kuvaus F kuvaa valin [0,1] solmujen k ja k+1
%valiselle valille.
%Muuttujan x tulee sijaita valilla [0,1].
F = @(x,k,n) ((2*pi*k)/(n-1) ...
- (2*pi*(k-1))/(n-1))*x ...
+ (2*pi*(k-1))/(n-1);
%Kaanteiskuvaus iF kuvaa nodejen k ja k+1
%valisen valin valille [0,1].
%Muuttuja x tulee sijaita solmujen
%k ja k+1 valilla.
iF = ...
@(x,k,n) (-1/(((2*pi*k)/(n-1))+2*pi*(k-1)/(n-1)))*x ...
+ 2*pi*(k-1)/((n-1)*(-2*pi*k/(n-1)+2*pi*(k-1)/(n-1)));

%Kuormafunktio.
f = @(x) sin(x);

A = zeros(n);
L = zeros(n, 1);

for k = 1: n
    if(k ==1)
    elseif(k==n)
    else
        %Pisteet solmusta "taakse".
        Xg1 = F(X,k-1,n);
        %Pisteet solmusta "eteen".
        Xg2 = F(X,k,n);
        tempL1 = (2*pi/(n-1))*f(Xg1).*C_1(X);
        tempL2 = (2*pi/(n-1))*f(Xg2).*C_2(X);
        L(k) = -(tempL1'*W+ tempL2'*W);
    end
end

```

```

    end
    if (k<n && k>1)
        A(k-1,k) = (2*pi/(n-1))*(dc_1(X).*dc_2(X))'*W;
        A(k,k) = (2*pi/(n-1))*((dc_1(X).*dc_1(X))'*W + ...
            (2*pi/(n-1))*((dc_2(X).*dc_2(X))'*W);
        A(k+1,k) = (2*pi/(n-1))*(dc_2(X).*dc_1(X))'*W;
    end
end
end
%Reunasolmut pois:
A(1,:) = [];
A(:,1) = [];
L(1) = [];

A(n-1,:) = [];
A(:,n-1) = [];
L(n-1) = [];

%Ratkaistaan vakiot kantafunktioille.
u = inv(A)*L;

%Lisataan reunaehdot.
U = zeros(length(u)+2,1);
U(2:length(U)-1) = u;

%Piirra vastaus.
kkk = linspace(0,2*pi,length(U));
t = linspace(0,2*pi,1000);
f = @(x) -sin(x);
figure(1)
plot(kkk,U); hold;
plot(t,f(t),'—');
title('Elementtien_maara:5')
legend('FEM-ratkaisu_lineaarisilla_kantafunktioilla',...
'Tasmallinen_ratkaisu', 'Location','southeast')

```

Koodissa käytetään niin sanottuja *vertailuelementtiä* ja *paikallisia kantafunktioita*. Vertailuväli  $[0, 1]$  kuvataan *globaaliksi* väliksi  $[2\pi(k-1)/(n-1), 2\pi k/(n-1)]$  lineaarikuvauksella muotoa  $F = x \mapsto ax + b$ . Kuvaus  $F$  on suoraviivaista löytää ratkaisemalla arvot  $a$  ja  $b$  yhtälöparista

$$\begin{cases} (F(0) =) & b = 2\pi(k-1)/(n-1) \\ (F(1) =) & a + b = 2\pi k/(n-1). \end{cases}$$

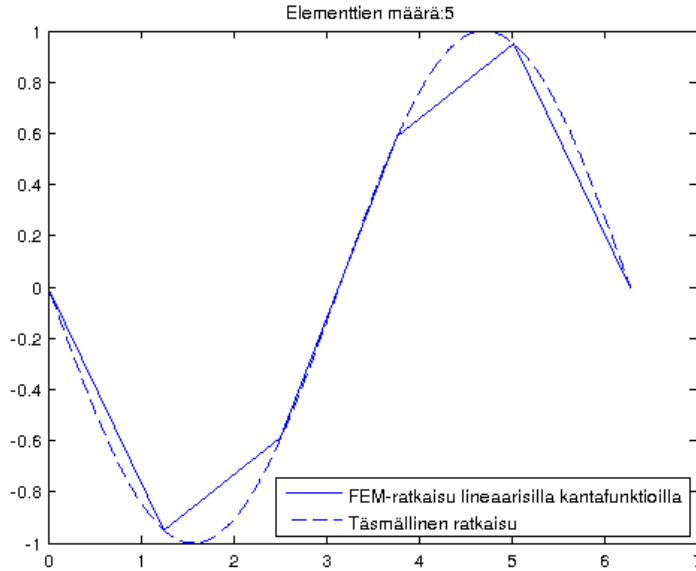
Integraali lasketaan *Gaussin vertailupisteiden ja sijoitusmenetelmän* avulla, eli

$$\int_a^b f(x)dx = \int_0^1 f(\phi(x))\phi'(x)dx \approx \frac{2\pi}{(n-1)} \sum_{i=1}^n w_i f(\phi(x_i)),$$

missä arvot  $\{w_i\}$  ovat pistekohtaiset painotukset pisteille  $x_i \in [0, 1]$ . Kantafunktiot määritellään siten, että mielivaltaiselle välille  $[a, b]$  määrittelylle kantafunktiolle  $\tilde{\phi}$  pätee  $\tilde{\phi}(x) := \phi(F^{-1}(x))$  kaikilla  $x \in [a, b]$ , jossa  $\phi$  on vertailuelementtiin määritelty kantafunktio. Tällöin derivaatalle pätee ketjusäännön nojalla

$$\frac{d}{dx} \tilde{\phi}(x) = D_x F^{-1}(x) \cdot D_{F^{-1}(x)} \phi(F^{-1}(x)),$$

mikä on hyvä muistaa menetelmää toteuttaessa. Vertailuelementin ja Gaussin vertailupisteiden käyttö ovat yleisiä menetelmiä FEM algoritmia toteuttaessa.



Kuva 1.1: Tehtävän (1.5) ratkaisu

# Luku 2

## Abstrakti muotoilu

Luvussa käsitellään elementtimetodin matemaattista perustaa modernin analyysin käsitteillä. Moderni analyysi tarjoaa hyvät työkalut ratkaisuiden olemassaolon ja suppenemisominaisuuksien tutkimukseen. Oleellista on heikon muotoilun ratkaisun tulkinta *jakaumaksi* (distribution), eli yleistetyksi funktioksi. Näin pääsemme käyttämään jakaumille kehitettyä teoriaa suoraan FEM malliimme.

### 2.1 Peruskäsitteet

Työssä *skalaarien kunta*  $\mathbb{K}$  tarkoittaa käytännössä kompleksilukuja  $\mathbb{C}$  tai reaalilukuja  $\mathbb{R}$ . Funktiot ymmärretään niin, että enintään *nollamittaisesti* toisistaan poikkeavat funktiot samaistetaan keskenään. Täsmälleen ottaen puhutaan siis funktio-tyyppien *kategorioista*, eikä yksittäisistä funktioista. Esimerkiksi *nollavektori* ”0” on funktio  $f : U \rightarrow \mathbb{K}$ , jolla  $f(x) = 0$  *melkein kaikilla*  $x \in U$ . Jatkuvista funktioista puhuessa tarkoitetaan, että käsiteltävä *kategoria* sisältää jatkuvan edustajan. Integraalimerkinnällä  $\int_U dx$ , tai vain lyhyesti  $\int_U$ , viitataan perinteiseen *Lebesguen integraaliin* yli alueen  $U \subset \mathbb{R}^n$ .

**Määritelmä 2.1** (Lineaarikuvaus). Olkoot  $W$  ja  $V$  *vektoriavaruuksia*. Lineaarikuvaukselle  $\mathcal{L} : W \rightarrow V$  pätee

$$\mathcal{L}(x + by) = \mathcal{L}(x) + b\mathcal{L}(y)$$

kaikilla  $b \in \mathbb{K}$ .

**Määritelmä 2.2** (Duaaliavaruus). Olkoon  $V$  vektoriavaruus. Avaruuden  $V$  *duaaliavaruus*  $V'$  koostuu *jatkuvista* lineaarisista kuvauksista

$$L : V \rightarrow \mathbb{R}.$$

Esimerkiksi jatkuvien funktioiden avaruuden eräs duaaliavaruuden alkio  $L$  voidaan identifioida integraalilla

$$L = f \mapsto \int_U f(x)v(x)dx,$$

missä  $v$  on jokin jatkuva funktio. Voidaan itseasiassa osoittaa, että jokainen duaaliavaruuden alkio voidaan tietyissä funktioavaruuksissa identifioida edellisen kaltaisella integraalilla, kunhan funktiot  $v$  valitaan sopivasti. Esimerkiksi *Lebesguen avaruuden*  $L^p$  duaali voidaan identifioida avaruuden  $L^q$  alkiolla, missä  $1/p + 1/q = 1$  ([5]).

Eräs tapa tunnistaa jatkuva lineaarinen funktio on huomio, että jatkuvuuden funktio on myös rajoitettu ja päinvastoin.

**Määritelmä 2.3** (Rajoitettu kuvaus). Olkoot  $W$  ja  $V$  vektoriavaruuksia. Lineaarikuvaus  $\mathcal{L} : W \rightarrow V$  on rajoitettu jos on olemassa  $C > 0$ , jolle

$$\|\mathcal{L}(x)\|_V \leq C\|x\|_W$$

kaikilla  $x \in W$ .

**Lause 2.4.** *Lineaarinen kuvaus  $\mathcal{L} : W \rightarrow V$  on jatkuva jos ja vain jos se on rajoitettu.*

*Todistus.* Rajoitettu lineaarinen kuvaus on jopa *Lipschitz jatkuva*, sillä

$$\|\mathcal{L}(x - y)\| = \|\mathcal{L}(x) - \mathcal{L}(y)\| \leq C\|x - y\|.$$

Toisaalta jatkuvuuden nojalla  $\|\mathcal{L}(z)\| \leq 1$ , kun  $\|z\| \leq \delta$ . Olkoon  $y \in W$  ja  $z = \frac{\delta}{\|y\|}y$ , jolloin lineaarisuuden nojalla

$$\|\mathcal{L}(z)\| = \frac{\delta}{\|y\|}\|\mathcal{L}(y)\| \leq 1,$$

mistä seuraa  $\|\mathcal{L}(y)\| \leq 1/\delta\|y\|$ , eli  $\mathcal{L}$  on rajoitettu. □

Erityisesti voidaan valita  $V = \mathbb{R}$ , jolloin lauseen 2.4 tulos pätee duaaliavaruuden alkiolle.

**Huomautus.** *Kaikki lineaariset funktionaalit eivät suinkaan ole jatkuvia. Esimerkiksi jatkuvien funktioiden  $g : \mathbb{R} \rightarrow \mathbb{R}$  avaruudessa varustettuna normilla  $\|g\| = \int_{\mathbb{R}} |g(x)|dx$  Diracin deltafunktio  $\delta = g \mapsto g(0)$  on lineaarinen, mutta ei jatkuva. Esimerkiksi Gaussin käyristä koostuvalle funktiojonolle*

$$(g_n(x))_{n \in \mathbb{N}} := \left( \frac{n}{\sqrt{\pi}} e^{-(xn)^2} \right)_{n \in \mathbb{N}}$$

arvo  $\delta(g_n) = g_n(0)$  saa mielivaltaisen suuria arvoja, mutta  $\|g_n\| = 1$  kaikille  $n$ . Näin ollen Diracin deltafunktio ei ole rajoitettu. Myöhemmin esiteltävä Sobolevin lemma kuitenkin takaa jatkuvuuden tietyissä avaruuksissa.

Määritellään seuraavaksi klassisen derivaatan yleistys *heikko derivaatta*. Heikon derivaatan käsite tarjoaa työkalut puhua funktioiden derivaatoista niissäkin tapauksissa, joissa klassista derivaattaa ei ole määritelty kaikissa pisteissä. Elementtimenetelmän kannalta tämä on oleellista jo siinä mielessä, että kantafunktiot eivät usein ole derivoituvia kaikkialla.

**Määritelmä 2.5** (Heikko derivaatta). Määritellään *multi-indeksi*

$$\alpha = (\alpha_1, \dots, \alpha_n),$$

jolle  $|\alpha| = \alpha_1 + \dots + \alpha_n$ . Merkitään seuraavassa funktion  $\lambda$  kertaluokan  $\alpha$  *klassista osittaisderivaattaa* merkinnällä

$$\frac{\partial^{|\alpha|} \lambda}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_n} x_n} = \lambda^{(\alpha)}.$$

Tällöin funktio  $u$  on funktion  $v$  kertaluokan  $\alpha$  heikko derivaatta, jos

$$\int_U v \lambda^{(\alpha)} = (-1)^{|\alpha|} \int_U u \lambda$$

kaikilla *kompaktikantaisilla* äärettömän monta kertaa (klassisesti) derivoituvilla funktioilla  $\lambda \in C_0^\infty(U)$ . Kompaktikantaisuus tarkoittaa, että  $\lambda$  saa nollasta poikkeavia arvoja vain jossakin *kompaktissa* alueen  $U$  osajoukossa.

*Esimerkki 1.* Olkoon avaruus  $V = C^1([-1, 1]) \rightarrow \mathbb{R}$  kerran jatkuvasti derivoituvien funktioiden avaruus suljetulla välillä  $[-1, 1]$ . Haetaan heikko derivaatta funktiolle

$$f(x) = \begin{cases} -x, & \text{kun } x < 0 \\ x, & \text{kun } x \geq 0. \end{cases}$$

Selvästi funktion  $f$  klassiselle derivaatalle pätee

$$f'(x) = \begin{cases} -1, & \text{kun } x < 0 \\ 1, & \text{kun } x > 0, \end{cases}$$

mutta derivaattaa pisteessä  $x = 0$  ei ole hyvin määritelty. Kyseessä on kuitenkin pisteenä nollamittainen joukko, joten heikon derivaatan kannalta arvo pisteessä  $x = 0$  voidaan määritellä mielivaltaisesti. Osoitetaan täten, että funktion  $f$  heikko derivaatta on esimerkiksi

$$u(x) = \begin{cases} -1, & \text{kun } x < 0 \\ 1, & \text{kun } x \geq 0. \end{cases}$$

Olkoon  $\lambda \in C_0^\infty([-1, 1])$ , jolloin *osittaisintegroimalla* saadaan

$$\begin{aligned} \int_{-1}^1 u\lambda &= \int_{-1}^0 -\lambda + \int_0^1 \lambda' \\ &= - \int_{-1}^0 x\lambda + \int_{-1}^0 x\lambda' + \int_0^1 x\lambda - \int_0^1 x\lambda' \\ &= \int_{-1}^0 x\lambda' - \int_0^1 x\lambda' = - \int_{-1}^1 f\lambda', \end{aligned}$$

joten määritelmän 2.5 ehdot täyttyy.

*Esimerkki 2.* Aiemmin mainittu Diracin deltafunktio tulee esille tutkiessa *Heavisiden funktion* heikkoa derivaattaa

$$H(x) = \begin{cases} 0, & \text{kun } -1 \leq x < 0 \\ 1, & \text{kun } 0 \leq x \leq 1. \end{cases}$$

Heavisiden funktion heikolle derivaatalle  $u$  pitäisi päteä

$$\begin{aligned} - \int_{-1}^1 H\lambda' &= - \int_0^1 \lambda' = - \int_0^1 \lambda = \lambda(0) \\ (2.1) \qquad \qquad \qquad &= \int_{-1}^1 u\lambda, \end{aligned}$$

kaikilla  $\lambda \in C_0^\infty$ . Heavisiden funktion heikko derivaatta näyttäisi funktiolta  $u$ , mikä kuvaa integraalin sisällä jokaisen funktion  $\lambda$  sen arvolle pisteessä 0, eli  $\int u\lambda = \lambda(0)$ . Mikään funktio  $u : \mathbb{R} \mapsto \mathbb{R}$  ei toteuta kaavaa 2.1 perinteisen Lebesguen integraalin mielessä, mutta voidaan sanoa, että Heavisiden funktion *distribuutioderivaatta* on kuvaus

$$\delta = \lambda \mapsto - \int_{-1}^1 H\lambda' = \lambda(0).$$

**Propositio 2.6.** *Heikko derivaatta yhtenee klassisen derivaatan kanssa, jos se on olemassa.*

*Todistus.* Olkoon funktio  $\frac{\partial v}{\partial x_i}$  funktion  $v$  osittaisderivaatta klassisessa mielessä muuttujan  $x_i$  suhteen. Tällöin yleistetyn osittaisintegrointikaavan nojalla

$$\begin{aligned}\int_U \frac{\partial \lambda}{\partial x_i} v dU &= \int_{\partial U} \lambda v \hat{v}_i dA - \int_U u \frac{\partial \lambda}{\partial x_i} \quad (\lambda = 0 \text{ reunalla } \partial u) \\ &= - \int_U \lambda \frac{\partial v}{\partial x_i},\end{aligned}$$

missä  $\hat{v}$  on *pinnan normaali*. Tulos seuraa tästä induktiivisesti kaikille multiindekseille  $\alpha$ .  $\square$

Tästä lähtien derivaatasta puhuessa tarkoitetaan aina heikkoa derivaattaa, ellei toisin mainita. Esimerkiksi merkinnällä  $v^{(\alpha)}$  tarkoitetaan funktion  $v$  kertaluokan  $\alpha$  heikkoa derivaattaa.

E erityisesti heikko derivaatta on tulkittavissa duaaliavaruuden  $(C_0^\infty(U))'$  alkioiksi, sillä se voidaan nähdä lineaarikuvauksena

$$u : C_0^\infty(U) \rightarrow \mathbb{K}, u = \lambda \mapsto (-1)^{|\alpha|} \int_U v \lambda^{(\alpha)}.$$

Lisää jakaumista Helsingin yliopiston luentomonisteissa ([9]).

## 2.2 Tehtävän muotoilu

Lyhyesti muotoiltuna, työn ongelma kuuluu

$$\begin{aligned}\text{Olkoon } A : V &\rightarrow V' \text{ ja } L \in V'. \text{ Etsi } u \in V, \text{ jolla} \\ Au &= L.\end{aligned}$$

Työn kannalta tämä on aikalailla yliabstrahoitu muotoilu, joka kuitenkin ehkä selventää miksi duaaliavaruudet on keskeisessä osassa elementtimenetelmän matemaattisessa analyysissä. Täsmällisempi muotoilu löytyy kohdasta 2.10.

**Määritelmä 2.7** (Bilineaarinen muoto). Olkoon  $V$  vektoriavaruus. Kuvaus  $a : V \times V \rightarrow \mathbb{K}$ ,  $a = (v, w) \mapsto a(v, w)$  on bilineaarinen muoto, jos

- $a(v, w) = a(w, v)$
- $\lambda a(v, w) = a(v, \lambda w)$
- $a(v, w) + a(z, w) = a(v + z, w)$ .

**Määritelmä 2.8** (Koersiivinen muoto). Bilineaarinen muoto on *koersiivinen*, tai *elliptinen*, jos on olemassa  $c > 0$ , jolle

$$|a(w, w)| \geq c \|w\|^2$$

kaikilla  $w \in V$ .



**Määritelmä 2.9** (Rajoitettu muoto). Bilineaarinen muoto on *rajoitettu* jos on olemassa  $C > 0$ , jolle

$$|a(w, v)| \leq C \|w\| \|v\|$$

kaikilla  $w, v \in V$ .

Osittaisdifferentiaaliyhtälön (ODY) heikko muoto voidaan muotoilla seuraavasti. Muotoa kutsutaan myös *Galerkinin muotoiluksi* ja *variaatiotehtäväksi*.

**Määritelmä 2.10** (Galerkinin muotoilu).

Etsi  $v \in V$ , jolle

$$a(v, w) = L(w) \text{ kaikilla } w \in V.$$

Vastaava minimointitehtävä kuuluu

**Määritelmä 2.11.** Etsi  $v \in V$ , jolle

$$J(v) = \frac{1}{2}a(v, v) - L(v) \leq \frac{1}{2}a(w, w) - L(w) = J(w)$$

kaikilla  $w \in V$ . Tämä vastaa fysiikan minimienergia periaatetta.

**Korollaari 2.12.** *Määritelmät 2.10 ja 2.11 ovat yhtäpitävät.*

*Todistus.* Oletetaan, että  $v$  ratkaisee tehtävän 2.10. Tällöin kaikille  $w \in W$  pätee, että  $J(v + w) \geq J(v)$ , sillä

$$\begin{aligned} (2.2) \quad J(v + w) &= \frac{1}{2}a(v, v) + a(v, w) + \frac{1}{2}a(w, w) - L(w) - L(v) \\ &\geq \frac{1}{2}a(v, v) - L(v) = -\frac{1}{2}a(v, v), \end{aligned}$$

missä viimeinen yhtäpitävyys seuraa suoraan oletuksesta. Edelleen (2.2) pätee jos ja vain jos

$$a(v, v) + a(v, w) + \frac{1}{2}a(w, w) - L(w) - L(v) = \frac{1}{2}a(w, w) \geq 0,$$

mikä pitää koersiivisuuden nojalla.

Toisaalta jos  $u$  on *funktionaalin minimi*, niin funktionaalin  $J$  derivaatta saa arvon 0 pisteessä  $u$ . Tällöin pisteille  $v \in W$  pätee

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \left[ \frac{d}{d\epsilon} J(u + \epsilon v) \right] &= \lim_{\epsilon \rightarrow 0} \left[ \frac{d}{d\epsilon} \left( \frac{1}{2}a(u, u) + a(u, \epsilon v) + \frac{1}{2}a(\epsilon v, \epsilon v) - L(\epsilon v) - L(u) \right) \right] \\ &= \lim_{\epsilon \rightarrow 0} \left[ \frac{d}{d\epsilon} \left( \epsilon a(u, v) + \frac{\epsilon^2}{2}a(v, v) - \epsilon L(v) \right) \right] \\ &= \lim_{\epsilon \rightarrow 0} [a(u, v) + \epsilon a(v, v) - L(v)] \\ &= a(u, v) - L(v) = 0, \end{aligned}$$

mikä osoittaa, että ratkaisu  $u$  toteuttaa tehtävän 2.10.

□

Elementtimenetelmän teoriassa ratkaistaan jompikumpi edeltävistä tehtävistä avaruudessa  $V_h \subset V$  – tässä työssä Galerkinin muotoilu 2.10. Avaruuden  $V_h$  virittää äärellinen määrä kantafunktioita, mitkä on määritelty äärellisessä määrässä elementtejä, joten avaruus  $V_h$  on äärellisulotteinen. Sen sijaan avaruus  $V$  on usein ääretönulotteinen.

*Esimerkki 3.* Olkoon  $V = V(U)$  kaksi kertaa heikosti derivoituvien funktioiden avaruus alueessa  $U \subset \mathbb{R}^n$ . Muodostetaan heikko muotoilu *Poissonin yhtälölle*

$$\Delta u = f$$

käyttämällä avaruuden  $V$  aliavaruutta

$$V_h := \{v \in V \mid v \text{ on lineaarinen määrättyjen solmujen } \{x_i\}_{i=1}^n \text{ välissä}\}.$$

Kyseinen paloittain lineaaristen kantafunktioiden avaruus on yleinen FEM mallien funktioavaruus.

Kerrotaan Poissonin yhtälöä funktiolla  $v_h \in V_h$ , jolloin

$$\int_U \Delta u_h v_h = \int_U f v_h.$$

Nyt voidaan määritellä  $a(u_h, v_h) := \int_U \Delta u_h v_h$  ja  $L(v_h) := \int_U f v_h$ , jolloin ratkaisulle  $u_h \in V_h$  tulee päteä

$$a(u_h, v_h) = L(v_h)$$

kaikilla  $v_h \in V_h$ .

Ratkaisua Galerkinin muotoilulle ei ole välttämättä olemassa, vaan on asetettava tiettyjä ehtoja funktioavaruudelle  $V$  ja alueen  $U \subset \mathbb{R}^n$  muodolle. Työssä palataan kyseisiin ehtoihin. Esimekriksi esimerkissä 3 tullaan jo ongelmaan, että paloittain lineaarisille funktioille ei löydy toista heikkoa derivaattaa funktiona. Tullaan kuitenkin huomaamaan, että ongelma korjaantuu *Greenin kaavalla*.

**Lause 2.13** (Greenin kaava ([11])). *Olkoon  $v$  kaksi kertaa jatkuvasti klassisessa mielessä derivoitua ja  $u$  kerran jatkuvasti derivoitua. Tällöin*

$$\int_U u \Delta v dx = - \int_U \nabla u \cdot \nabla v dx + \int_{\partial U} u (\nabla v \cdot \mathbf{n}) dS,$$

missä  $\mathbf{n}$  on pinnan  $S$  normaali.

**Lemma 2.14** (Galerkinin ortogonaalisuus). *Oletetaan, että  $u_h \in V_h \subset V$  toteuttaa yhtälön (2.10) avaruudessa  $V_h$ , ja  $u$  toteuttaa vastaavasti yhtälön avaruudessa  $V$ , tällöin*

$$a(u - u_h, w_h) = 0$$

kaikilla  $w_h \in V_h$ .

*Todistus.* Erityisesti pätee  $a(u, w_h) = L(w_h)$  ja  $a(u_h, w_h) = L(w_h)$ , jolloin

$$a(u, w_h) - a(u_h, w_h) = a(u - u_h, w_h) = 0$$

kaikilla  $w_h \in V_h$ . □

Lemma 2.14 tarkoittaa, että tehtävän (2.10) ratkaisu avaruudessa  $W_h$  on *ortogonaalinen projektio* avaruudesta  $W$  avaruuteen  $W_h$ , jos määritellään sisätuloksi  $(v, w) \mapsto a(v, w)$ .

**Korollari 2.15** (Ratkaisun yksikäsitteisyys). *Jos  $v_h$  ja  $u_h$  on ratkaisuja tehtävälle 2.10 avaruudessa  $V_h$ , niin  $v_h = u_h$ .*

*Todistus.* Galerkinin ortogonaalisuuden nojalla

$$a(u - u_h, w_h) - a(u - v_h, w_h) = a(v_h - u_h, w_h) = 0,$$

mikä pätee erityisesti kun  $w_h = v_h - u_h$ , joten

$$a(v_h - u_h, v_h - u_h) = 0.$$

Tämä pätee kuvauksen  $a$  koersiivisuuden nojalla vain jos  $v_h - u_h = 0$ . □

## 2.3 Ratkaisujen olemassaolo

Elementtimenetelmän abstraktissa matemaattisessa muotoilussa ollaan kiinnostuneita seuraavista asioista:

- ratkaisun olemassaolo,
- suppenemisominaisuudet.

Ratkaisuiden olemassaolon analyysissä käytetään funktioanalyysin tuloksia ja aihe on sinänsä erillinen elementtimenetelmästä. Oleellista on todeta tarkan ratkaisun olemassaolo tietyillä reunaehdoilla, sillä tällöin voidaan varmistaa, että FEM

ratkaisumme todella approksimoi jotakin funktiota, toisin sanoen, että tehtävänäsettely on ylipäättänsä järkevä. Myös FEM mallin ratkaisun olemassaolo on kiinnostavaa, sillä on turha rakentaa malli, jolle ei löydy ratkaisua. FEM ratkaisujen suppenemisominaisuuksien tutkiminen, toisin sanoen miten nopeasti FEM mallin ratkaisu lähestyy täsmällistä ratkaisua, on myös kiinnostavaa. On teoreettisesti osoitettavissa, että FEM todellakin toimii (tietyissä puitteissa), mutta myös ”kokeellinen” suppenemisominaisuuksien tutkiminen on usein hyödyllistä. Käytännössä tämä tarkoittaa, että tiedämme täsmällisen ratkaisun vähintään hyvin tarkasti, ja vertaamme saamaamme FEM ratkaisua tähän. Kokeellisella metodilla voimme saada käsitystä erityyppisiin tehtäviin liittyvistä suppenemisominaisuuksista tihentämällä alueen ositusta tai muuttamalla kantafunktioita. Näin ollen teoreettinen ja kokeellinen tarkastelu menee usein käsi kädessä. Niin sanottu *posteriori* virheanalyysi perustuu ”kokeellisesti” saatuun FEM ratkaisuun, mitä kautta voidaan hakea eksplisiittinen virhearvio FEM ratkaisulle ([1]). Luvussa 5 esitellään *a priori* arvio, mikä antaa lähinnä kvalitatiivisen arvion virheelle, mutta asettaa menetelmän matemaattisesti vankalle pohjalle. Luvussa 6 tutkimme numeerista virhettä täsmällisen ratkaisun suhteen.

Ratkaisuja etsiessä on asetettava tiettyjä rajoituksia käsiteltäville funktioavaruuksille. Määriteltyjen lisäehtojen tarpeellisuus tulee esille myöhemmässä analyysissä.

**Määritelmä 2.16** (Lebesguen avaruus). Funktioavaruutta kutsutaan Lebesguen avaruudeksi  $L^p(U)^n$  jos kaikkien alkioiden  $v : U \rightarrow (\mathbb{K}_1, \dots, \mathbb{K}_n)$ ,  $v = x \mapsto (v_1(x), \dots, v_n(x))$  normille  $\|\cdot\|_{L^p(U)}$  pätee

$$\|v_i\|_{L^p(U)} := \left( \int_U |v_i(x)|^p dx \right)^{1/p} < \infty$$

kaikilla  $i = 1, \dots, n$ .

**Määritelmä 2.17** (Hilbertin avaruus). *Täydellistä normiavaruuta*  $V$  kutsutaan Hilbertin avaruudeksi, jos avaruuteen on määritelty *sisätulo*

$$\langle \cdot, \cdot \rangle_V : V \times V \rightarrow \mathbb{K},$$

mikä toteuttaa ehdot:

- $\langle u, v \rangle_V = \overline{\langle v, u \rangle_V}$ , missä yläviiva merkitsee *kompleksikonjugaattia*,
- $\langle \alpha u, v \rangle_V = \alpha \langle u, v \rangle_V$ ,  $\alpha \in \mathbb{K}$ ,
- $\langle v, u + w \rangle_V = \langle v, u \rangle_V + \langle v, w \rangle_V$ ,
- $\langle v, v \rangle_V \neq 0$  kaikilla  $v \in W$ ,

-  $\langle v, v \rangle_V = 0$  jos ja vain jos  $v = 0$ .

Hilbertin avaruudessa  $V$  normi  $\|\cdot\|_V : V \mapsto \mathbb{R}$  määrittyy luonnollisesti sisätulolla, jolloin

$$\|\cdot\|_V = x \mapsto \sqrt{\langle x, x \rangle_V}.$$

*Sobolevin avaruudet* ovat funktioavaruuksia, mitkä sisältävät määrätyn kertaluvun verran heikosti derivoituvat funktiot. Seuraavassa esitellään Sobolev avaruuksien yleinen määritelmä.

**Määritelmä 2.18** (Sobolevin avaruus). Sobolevin avaruus  $W^{p,k}(U)^n$  sisältää funktiot  $v : U \rightarrow (\mathbb{K}_1, \dots, \mathbb{K}_n)$ ,  $v = x \mapsto (v_1(x), \dots, v_n(x))$ , joiden normille  $\|\cdot\|_{W^{p,k}(U)}$  pätee

$$\|v_i\|_{W^{p,k}(U)} := \sum_{|\alpha| \leq p} \|v_i^{(\alpha)}\|_{L^k(U)} = \sum_{|\alpha| \leq p} \left( \int_U |v_i^{(\alpha)}(x)|^k dx \right)^{1/k} < \infty$$

kaikilla  $i = 1, \dots, n$ .

Seuraavat avaruudet ovat usein luonnollisia valintoja ratkaisuvaryuksiksi. Lebesguen avaruus  $L^2(U)$  on sisätuloavaruus sisätulolla

$$\langle \cdot, \cdot \rangle_{L^2(U)} = u, v \mapsto \int_U u \cdot \bar{v},$$

missä piste ” $\cdot$ ” merkitsee *pistetuloa*  $(v, u) \mapsto v_1 u_1 + \dots + v_n u_n$ .

Jos  $u, v : U \rightarrow \mathbb{C}$ , merkitään lyhyesti  $W^{2,1}(U) = H^1(U)$ . Tämä on sisätuloavaruus sisätulolla

$$\langle \cdot, \cdot \rangle_{H^1(U)} = u, v \mapsto \int_U \nabla u \cdot \nabla \bar{v} + \int_U u \bar{v},$$

missä kuvausta  $\nabla = u \mapsto (\frac{\partial}{\partial x_1} u, \dots, \frac{\partial}{\partial x_n} u)$  kutsutaan *gradientiksi*.

Edeltävät sisätulot määräävät normit

$$\|\cdot\|_{L^2(U)} = u \mapsto \sqrt{\langle u, u \rangle_{L^2(U)}}$$

ja

$$\|\cdot\|_{H^1(U)} = u \mapsto \sqrt{\langle u, u \rangle_{H^1(U)}}.$$

Olemme puhuneet esimerkiksi funktion  $u$  reunaehdosta reunalla  $\partial U$ . Määritelyjen avaruuksien luonteen takia tätä ei ole ihan täsmällisesti määritelty tällaiseen, sillä funktiota  $u$  ei ole yleisesti hyvin määritelty nollamittaisessa joukossa  $\partial U$ . Koska vain nollamittaisessa joukossa toisistaan eroavat funktiot samaistetaan, voi  $u$  saada tässä mielessä mielivaltaisia arvoja reunalla  $\partial U$ . Asia korjaantuu käsitteellä *jälki* (trace).

**Määritelmä 2.19** (Jälki). Olkoon  $U$  rajoitettu joukko, minkä reuna on muotoa  $C^1$ , eli kerran jatkuvasti klassisessa mielessä derivoituva. Tällöin on osoitettavissa, että on olemassa lineaarinen rajoitettu operaattori

$$T : W^{1,p}(U) \rightarrow L^p(\partial U),$$

jolle

$$Tu = u|_{\partial U},$$

kun  $u \in W^{1,p}(U) \cap C^1(\bar{U})$ . Kuvausta  $Tu$  kutsutaan jäljeksi. Koska avaruus  $C^1(\bar{U})$  on tiheä avaruudessa  $W^{1,p}(U)$ , voidaan jokaista funktioita  $u \in W^{1,p}(U)$  approksimoida mielivaltaisen tarkasti funktioilla  $\{u_n \in C^1(U)\}$ , jolloin reunaehto määräytyy yksikäsitteisesti

$$u|_{\partial U} := Tu = \lim_{n \rightarrow \infty} u_n|_{\partial U}.$$

Kun puhutaan reunaehdosta  $u|_{\partial U} = g$ , tarkoitetaan tarkalleen ottaen, että  $Tu = g$ . Vastaavasti voidaan yleistää klassisia tuloksia kattamaan yleisempiä funktioita käyttäen edellisen kaltaista approksimointia. Tästä esimerkkinä työssä käytettävä Greenin kaava 2.13. Erityisesti jälki on hyvin määritelty ja jatkuva funktioille avaruudessa  $H^1(U)$  (ei kuitenkaan välttämättä derivoituva) ([4]).

On osoitettavissa, että avaruudet  $L^2(U)$  ja  $H^1(U)$  ovat täydellisiä, eli ne ovat Hilbertin avaruuksia ([5]). Tästä lyhenne  $W^{1,p}(U) = H^1(U)$ .

**Lause 2.20** (Laxin-Milgramin lause). Olkoon  $a : V \times V \rightarrow \mathbb{C}$  koersiivinen ja rajoitettu bilineaarinen kuvaus ja  $L : V \rightarrow \mathbb{C}$  duaaliavaruuden  $V'$  alkio. Lisäksi oletetaan, että  $V$  on Hilbertin avaruus. Tällöin on olemassa yksikäsitteinen  $u \in V$ , jolle

$$a(u, v) = L(v)$$

kaikilla  $v \in V$ .

*Todistus.* Todistus seuraa Avnerin kirjan ([5]) todistusta. Kiinnitetään  $u \in V$ , jolloin koersiivisuudesta seuraa  $|a(u, v)| \leq C'\|v\|$ , jossa  $C' = \|u\|$ . Muoto  $a(u, \cdot) = v \mapsto a(u, v)$  on näin ollen rajoitettu lineaarinen kuvaus (eli duaaliavaruuden  $V'$  alkio).

Rieszin esityslauseen nojalla tällöin on olemassa alkio  $z \in V$ , jolla

$$a(u, v) = \langle z, v \rangle \text{ kaikilla } v \in V,$$

missä merkintä  $\langle \cdot, \cdot \rangle$  tarkoittaa avaruudessa  $V$  määriteltyä sisätuloa.

Olkoon  $T$  sopiva lineaarikuvaus avaruudesta  $V$  avaruuteen  $V^* \subset V$ , jolloin vastaavasti Rieszin esityslauseen nojalla on olemassa  $Tu$ , jolla

$$a(u, v) = \langle Tu, v \rangle \text{ kaikilla } v \in V^*.$$

Nyt pätee

$$\|Tu\|^2 = \langle Tu, Tu \rangle = a(u, Tu) \leq C\|v\|\|Tu\|,$$

eli  $\|Tu\| \leq C\|u\|$ , siis  $T$  on rajoitettu.

Cauchy-Swarchin epäyhtälön nojalla pätee

$$c\|u\|^2 \leq |a(u, u)| = | \langle Tu, u \rangle | \leq \|Tu\|\|u\|,$$

siis  $c\|u\| \leq \|Tu\|$ , eli  $\|Tu\| = 0$  jos ja vain jos  $\|u\| = 0$ . Näin ollen lineaarisena jatkuvana kuvauksena  $T$  on bijektiivinen kuvaus, joka kuvaa avaruuden  $V$  suljetulle aliavaruudelle  $V^*$ . Osoitetaan, seuraavaksi, että  $V = V^*$ . Jos  $V^* \neq V$ , niin on olemassa nollasta poikkeava  $z \in V$ , jolle  $\langle Tu, z \rangle = 0$  kaikilla  $Tu \in V^*$ . Erityisesti pätee

$$0 = | \langle Tz, z \rangle | = |a(z, z)| \geq c\|z\|,$$

mikä on mahdollista vain jos  $z = 0$ , mikä on ristiriita. Todetaan, että  $V^* = V$ , eli  $T$  on bijektiivinen kuvaus avaruudelta  $V$  itseensä.

Olkoon  $L \in V'$ . Rieszin esityslauseen nojalla on olemassa  $z \in V$ , jolle

$$L(v) = \langle z, v \rangle$$

kaikilla  $v \in V$ . Koska  $T(V) = V$ , on olemassa  $u \in V$ , jolle  $Tu = z$ . Tämä todistaa väitteen.

Lisäksi ratkaisu  $u$  on yksikäsitteinen, sillä ratkaisuille  $u, w$  pätee

$$a(u - w, v) = 0$$

kaikilla  $v \in V$ . Tämä pätee erityisesti kun  $v = u - w$ , joten

$$a(u - w, u - w) = 0,$$

mikä pätee koersiivisuuden nojalla vain jos  $u - w = 0$ . □

Lax-Milgramin lause takaa olemassaolon ja yksikäsitteisyyden täsmälliselle ratkaisulle avaruudessa  $V$ . Vastaavasti lause takaa yksikäsitteisen ratkaisun olemassaolon äärellisulotteisessa *suljetussa* aliavaruudessa  $V_h$ .

Seuraava lause on hyödyllinen.

**Lause 2.21** (Poincarén epäyhtälö ([4])). *Olkoon  $U \subset \mathbb{R}^n$ , ja  $u \in W^{1,p}(U)$ . Lisäksi  $u = 0$  reunalla  $\partial U$ . Tällöin on olemassa  $c > 0$ , jolle*

$$\|u\|_{L^p(U)} \leq c\|\nabla u\|_{L^p(U)}.$$

*Esimerkki 4.* Osoitetaan, että *Helmholtzin yhtälölle*

$$(2.3) \quad \begin{aligned} -\Delta u + \lambda u &= f \\ u &= 0 \text{ reunalla } \partial U \end{aligned}$$

on olemassa (heikko) ratkaisu avaruudessa  $H_0^1(U)$ . Oletetaan lisäksi, että reuna  $\partial U$  on sileä, eli kerran klassisessa mielessä derivoituva. Edellä reunaehto on määritelty jälkimiessä, eli  $u|_{\partial U} = Tu$ . Vakio  $\lambda \geq 0$ .

Puolittain kertomalla funktion  $v$  kompleksikonjugaatilla  $\bar{v} \in H_0^1(U)$  ja integroimalla saadaan

$$(2.4) \quad -\int_U \Delta u \bar{v} + \lambda \int_U u \bar{v} = \int_U f \bar{v}.$$

Tässä kohtaa Cauchyn-Scwarzin epäyhtälö takaa, että integraalit ovat hyvin määriteltyjä, eli saavat äärellisiä arvoja. Tämän vuoksi myös oletus  $f \in L^2(U)$  on oleellinen. Greenin kaavan 2.13 nojalla

$$(2.5) \quad -\int_U \Delta u \bar{v} = \int_U \nabla u \cdot \nabla \bar{v} + \int_{\partial U} \bar{v}(\nabla u \cdot n) dA,$$

missä  $n$  merkitsee alueen  $U$  pinnan normaalia ja  $dA$  *pintadifferentiaalia*. Sijoittamalla kaava (2.5) kaavaan (2.4) ja ottamalla testifunktion reunaehto  $\bar{v}|_{\partial U} = 0$  huomioon saadaan

$$\int_U \nabla u \cdot \nabla \bar{v} + \lambda \int_U u \bar{v} = \int_U f \bar{v}.$$

Määritellään

$$a(u, v) := \int_U \nabla u \cdot \nabla \bar{v} + \lambda \int_U u \bar{v}$$

ja

$$L(v) := \int_U f \bar{v}.$$

Määritellyt kuvaukset ovat (bi-)lineaarisia.

Huomataan aluksi, että nyt

$$|a(u, u)| = \|\nabla u\|_{L^2(U)}^2 + \lambda \|u\|_{L^2(U)}^2.$$

Poincarén epäyhtälön 2.21 nojalla on olemassa  $c - 1 > 0$ , jolla

$$\|v\|_{L^2(U)}^2 \leq (c - 1) \|\nabla v\|_{L^2(U)}^2,$$

mistä seuraa

$$\|\nabla v\|_{L^2(U)}^2 \geq \frac{1}{c} \|v\|_{L^2(U)}^2 + \frac{1}{c} \|\nabla v\|_{L^2(U)}^2,$$



joten

$$\begin{aligned} |a(u, u)| &\geq \frac{1}{c} \|v\|_{L^2(U)}^2 + \frac{1}{c} \|\nabla v\|_{L^2(U)}^2 + \lambda \|v\|_{L^2(U)}^2 \geq \frac{1}{c} \left( \|u\|_{L^2(U)}^2 + \|\nabla u\|_{L^2(U)}^2 \right) \\ &= \frac{1}{c} \|u\|_{H^1(U)}^2, \end{aligned}$$

mikä osoittaa, että muoto  $a$  on koersiivinen. Kuvaus  $a$  on rajoitettu, sillä

$$\begin{aligned} |a(u, w)| &= | \langle \nabla u, \nabla w \rangle_{L^2(U)} + \lambda \langle u, w \rangle_{L^2(U)} | \\ &\leq | \langle \nabla u, \nabla w \rangle_{L^2(U)} | + | \lambda \langle u, w \rangle_{L^2(U)} | \\ &\leq \|\nabla u\|_{L^2(U)} \|\nabla w\|_{L^2(U)} + \lambda \|u\|_{L^2(U)} \|w\|_{L^2(U)} \\ &\hspace{15em} \text{(Cauchyn-Schwarzin epäyhtälö)} \\ &\leq \begin{cases} \|u\|_{H^1(U)} \|w\|_{H^1(U)}, & \text{jos } \lambda < 1 \\ \lambda \|u\|_{H^1(U)} \|w\|_{H^1(U)}, & \text{jos } \lambda \geq 1. \end{cases} \end{aligned}$$

Lisäksi funktionaali  $L$  on rajoitettu, sillä Cauchyn-Schwarzin epäyhtälön nojalla

$$\left| \int_U f \bar{v} \right| \leq \|f\|_{L^2(U)} \|v\|_{L^2(U)} \leq \|f\|_{H^1(U)} \|v\|_{H^1(U)}.$$

Kuvaukset  $a$  ja  $L$  toteuttavat näin Laxin-Milgramin lauseen vaatimat ehdot avaruudessa  $H_0^1(U)$ , joten ratkaisu Helmholtzin yhtälölle löytyy kyseisestä avaruudesta.

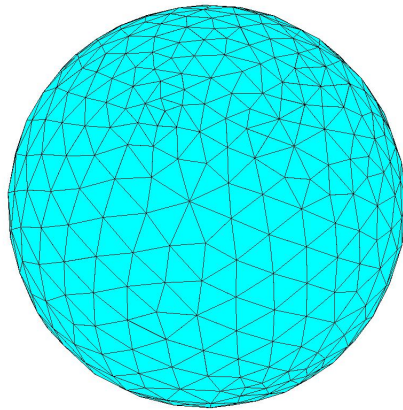
Eriyistapauksessa  $\lambda = 0$  Helmholtzin yhtälöä kutsutaan *Poissonin yhtälöksi*, ja tapauksessa  $\lambda = 0, f = 0$  *Laplacen yhtälöksi*.

**Huomautus.** *Esimerkissä 2.4 osoitettiin ratkaisun olemassaolo tapauksessa  $\lambda \geq 0$ . Tapaus  $\lambda \leq 0$  on itseasiassa Helmholtzin yhtälön yleisempi muotoilu, mutta jätämme tapauksen tässä todistamatta.*

# Luku 3

## Elementtiavaruus

Alueen  $U$  jakamista geometrisiin osiin kutsutaan osittamiseksi. Tässä työssä ei käsitellä yksityiskohtaisesti osittamisalgoritmien toimintaa, vaan lähinnä esitellään eri tyyppisiä käyttökelpoisia elementtejä ja niihin liittyvää teoriaa. Elementtien muodon valinta riippuu paljolti alueen muodosta ja ratkaistavasta tehtävästä. Mitään yksiselitteistä sääntöä tähän ei kuitenkaan ole. Kolmiomaisia elementtejä käytettäessä voidaan mahdollisesti saada täytettyä alue tarkemmin kuin neliömäisiä elementtejä käytettäessä, mutta tämäkään ei ole aivan yksiselitteistä ([7]). Tässä työssä keskitytään lähinnä kolmiomaisiin elementteihin, sillä teorian tutkiminen on käytännöllistä aloittaa niiden kautta.



Kuva 3.1: Yksikköpallon ositus Matlabin rutiinilla *generatemesh* tehtynä

Kuvan 3.1 osituksessa käytetään tetraedeja. Matlab suorittaa osittamisen naatiivina kahdessa- ja kolmessa ulottuvuudessa. Kolmiulotteisia malleja voi tuoda

Matlabiin ainakin .stl tiedostoina, eli *CAD* muodossa.

Elementtimenetelmän välitön hyöty on sen kyvyssä arvioida differentiaaliyhätlöiden ratkaisuja mielivaltaisissa alueissa. Tämä mahdollistaa monimutkaisten ongelmien tutkimisen esimerkiksi lujuuslaskennassa, akustiikassa tai vaikka kvanttimekaniikassa.

### 3.1 Elementti

Elementtien valinta riippuu oleellisesti ratkaistavasta differentiaaliyhtälöstä, käytetyistä kantafunktioista ja alueen muodosta. Elementtien ominaisuudet vaikuttavat olennaisesti suppenemisominaisuuksiin, mitä käsitellään tarkemmin luvussa 5.

Elementille on käytännöllistä antaa muodollinen määritelmä myöhemmän matemaattisen teorian muotoilun kannalta.

**Määritelmä 3.1.** Olkoon  $\{\mathcal{T}^h : h \in (0, 1]\}$  jokin alueen  $U$  avoimien alijakojen perhe. Joukko  $\mathcal{T}^h$  koostuu avoimista elementeistä  $\{T_i\}_{i=1}^n$ . Lisäksi kaikilla  $i, j = 1, \dots, n$  pätee  $T_i \cap T_j = \emptyset$  kaikilla  $h \in (0, 1]$ . Lisäksi  $\cup T_i \subset U$ .

Joukkoa  $\mathcal{T}^h$  kutsutaan alueen  $U$  ositukseksi.

Seuraava on Ciarletin määritelmä elementille ([3]).

**Määritelmä 3.2.** Elementti on kolmikko  $(T, \mathcal{P}, \mathcal{N})$ , jossa

1. (avoimella) alueella  $T \subset \mathbb{R}^n$  on paloittain sileä reuna  $\partial T$ ,
2.  $\mathcal{P}$  on äärellisulotteinen funktioavaruus alueessa  $T$ , minkä virittää kanta  $\{\phi_1, \dots, \phi_k\} \subset \mathcal{P}$  ja
3.  $\mathcal{N} = \{N_1, \dots, N_k\}$  on duaalin  $\mathcal{P}' : \mathcal{P} \rightarrow \mathbb{K}$  kanta. Lisäksi vaaditaan, että kaikilla  $i, j$  pätee  $N_i(\phi_j) = \delta_{ij}$ ,

missä *Knoeckerin delta* on määritelty

$$\delta_{ij} = \begin{cases} 1, \text{ jos } i = j \\ 0, \text{ jos } i \neq j. \end{cases}$$

Joukkoa  $\{\phi_1, \dots, \phi_k\}$  kutsutaan kannan  $\mathcal{N}$  *duaalikannaksi*.

Avaruus  $\mathcal{P}$  merkitsee työssä tietyn asteisten polynomien avaruutta. Kantafunktioiden ei kuitenkaan tarvitse olla polynomeja, vaan voivat olla periaatteessa mitä tahansa.

Käytännössä työssä kantafunktioiden  $N_i$  oletetaan olevan määritelty pisteittäin ja tuottaa funktion  $\phi_i \in \mathcal{P}$  liittyvän arvon tiettyssä solmussa. Kuvaus  $N_i$  voisi kuitenkin olla esimerkiksi integraali jonkin alueen yli. Tämän tyyppisiä duaalikantoja käytetään interpolaation määrittelyssä yleisemmille (ei jatkuville) funktioille. Työssä ei kuitenkaan käsitellä tällaisia duaalikantoja – yleisempiä tuloksia löytyy muun muassa Brennerin ja Scottin kirjasta ([2]).

**Huomautus.** *Sanalla elementti viitataan yleisesti sekä geometriseen olioon  $K$ , että kolmikkoon  $(K, \mathcal{P}, \mathcal{N})$ .*

*Esimerkki 5.* Merkitään kolmiomaista elementtiä  $K \subset \mathbb{R}^2$  matriisilla

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

missä sarakkeet viittaavat kolmion kärkipisteisiin. Olkoon  $\mathcal{P}_{lin}$  paloittain lineaaristen funktioiden avaruus. Lisäksi määritellään solmuittain  $N_i(v) = v(z_i)$ , missä piste  $z_i$  merkitsee matriisin  $\mathbf{K}$  saraketta indeksillä  $i$ . Määritellään ”jokaiselle kulle” oma kantafunktio  $v_1(x, y) = 1 - x - y$ ,  $v_2(x, y) = x$ ,  $v_3(x, y) = y$ . Funktiot  $\{v_1, v_2, v_3\}$  virittävät avaruuden  $\mathcal{P}_{lin}$ , ja lisäksi pätee  $N_i(v_j) = v_j(z_i) = \delta_{ji}$ , joten kolmikko  $(K, \mathcal{P}_{lin}, \mathcal{N})$  on elementti.

*Esimerkki 6.* Olkoon  $K$  esimerkin 5 mukainen. Tällöin kolmikko  $K, \mathcal{P} = \{1 - x - y, x, y, xy, x^2y^2/4, y^2x^2/4\}$ ,

$$\mathcal{N} = \left\{ v(z_1), v(z_2), v(z_3), \frac{\partial^2 v(z_1)}{\partial x \partial y}, \frac{\partial^3 v(z_2)}{\partial x \partial^2 y}, \frac{\partial^3 v(z_3)}{\partial^2 x \partial y} \right\}$$

määrittää (hieman epäkäytännöllisen) elementin.

**Lemma 3.3.** *Olkoon joukko  $\{N_1, \dots, N_d\}$  avaruuden  $\mathcal{P}'$  kanta ja  $v \in \mathcal{P}$ . Tällöin  $N_i(v) = 0$  kaikilla  $i = 1, \dots, d$  jos ja vain jos  $v \equiv 0$ .*

*Todistus.* Olkoon  $\{N_1, \dots, N_k\}$  jokin kanta avaruudelle  $\mathcal{P}'$ . Tällöin määritelmällisesti annetulle  $L \in \mathcal{N}$  voidaan kirjoittaa

$$L = \alpha_1 N_1 + \dots + \alpha_k N_k,$$

joten erityisesti

$$L(\phi_j) = \alpha_1 N_1(\phi_j) + \dots + \alpha_k N_k(\phi_j)$$

kaikilla  $j = 1, \dots, k$ . Tämä voidaan esittää algebrallisella yhtälöllä

$$\mathbf{N}\alpha = L,$$

millä on yksikäsitteinen ratkaisu  $\alpha$  vain jos käänteismatriisi  $\mathbf{N}^{-1}$  on olemassa.

Toisaalta; oletetaan, että annetulle  $v \in \mathcal{P}$  pätee  $N_i v = 0$ . Jokainen  $v \in \mathcal{P}$  voidaan kirjoittaa muodossa

$$v = \beta_1 \phi_1 + \cdots + \beta_k \phi_k,$$

joten kaikille  $i = 1, \dots, k$  pätee

$$(3.1) \quad N_i(v) = \beta_1 N_i(\phi_1) + \cdots + \beta_k N_i(\phi_k) = 0,$$

joten  $B_1 = \cdots = B_k = 0$ , koska  $N_i(\phi_j) = \delta_{ij}$ . Näin ollen  $v \equiv 0$ . Kaava (3.1) voidaan esittää algebrallisella yhtälöllä  $\mathbf{M}\beta = 0$ , millä on yksikäsitteinen ratkaisu vain jos käänteismatriisi  $\mathbf{M}^{-1}$  on olemassa.

Edelliset päättelyt yhdistäen käänteismatriisi  $\mathbf{N}^{-1} = (\mathbf{M}^T)^{-1} = (\mathbf{M}^{-1})^T$  on olemassa täsmälleen kun  $\mathbf{M}^{-1}$  on olemassa. Tämä todistaa väitteen.  $\square$

## 3.2 Yleinen interpolaatio

*Interpolaatio* on tärkeä käsite johtaessa *FEM* ratkaisujen suppenemisominaisuuksia, sillä se mahdollistaa virheen ylärajan selvittämisen seuraavassa kappalessa esiteltävän *Ceân lemmän* kautta.

**Määritelmä 3.4.** Olkoon  $(K, \mathcal{P}, \mathcal{N})$  määritelmän 3.2 mukainen elementti. Tällöin paikallinen interpolaatio määritellään kaavalla

$$\mathcal{I}_K v := \sum_{i=1}^d N_i(v) \phi_i.$$

*Esimerkki 7.* Olkoon  $v = \sin(xy)$ . Esimerkin 5 elementille saadaan tällöin (lineaarinen) interpolaatio

$$\begin{aligned} (\mathcal{I}_K v)(x, y) &= N_1(\sin(0))(1 - x - y) + N_2(\sin(0))x + N_3(\sin(0))y \\ &= 0. \end{aligned}$$

**Korollari 3.5.** Kuvauksen  $N_i \in \mathcal{P}'$  lineaarisuudesta seuraa suoraan *interpolaation lineaarisuus*.

**Propositio 3.6.**

$$N_i(\mathcal{I}_K(f)) = N_i(f) \text{ kaikilla } 1 \leq i \leq d.$$

*Todistus.* Lineaarisuuden nojalla

$$\begin{aligned} N_i(\mathcal{I}_K(f)) &= N_i\left(\sum_{j=1}^d N_j(f)\phi_j\right) \\ &= \sum_{j=1}^d N_j(f)N_i(\phi_j) \\ &= N_i(f), \end{aligned}$$

sillä  $N_i(\phi_j) = \delta_{ij}$ . □

**Propositio 3.7.** *Jos  $f \in \mathcal{P}$ , niin  $\mathcal{I}_k(f) = f$ .*

*Todistus.* Proposition 3.6 ja kuvauksen  $N_i$  lineaarisuuden nojalla  $N_i(\mathcal{I}_k(f) - f) = 0$  kaikilla  $i \in 1, \dots, d$ . Lemmasta 3.3 seuraa tällöin  $\mathcal{I}_K(f) - f = 0$ . □

**Määritelmä 3.8.** (Globaali interpolaatio) Olkoon jokainen elementti  $(K_i, \mathcal{P}_i, \mathcal{N}_i)$  määritelmän 3.2 mukainen, ja  $\mathcal{K}$  jokin alueen  $U$  ositus. Määritellään *globaali interpolaatio*  $\mathcal{I}_{\mathcal{K}}$  ehdolla

$$\mathcal{I}_{\mathcal{K}}f|_{K_i} = \mathcal{I}_{K_i}f$$

kaikilla  $K_i \in \mathcal{T}$ .

## 3.3 Kantafunktioiden konstruointi

Kappaleessa konstruoidaan elementtikohtaisesti muutamia kantafunktioita annetuille elementeille.

### 3.3.1 Yksiulotteisia elementtejä

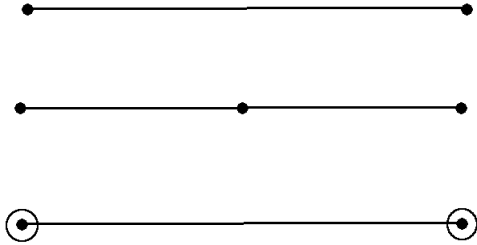
Kuvassa 3.2, ja tästä eteenpäin, tummalla pisteellä tarkoitetaan funktion arvoa kyseisessä pisteessä ja ympyrä tarkoittaa funktion derivaatan arvoa.

*Esimerkki 8.* Osoitetaan, että kuvan 3.2 toinen elementti määrää kannan (aidosti) toisen asteen polynomeille eli funktioille avaruudessa  $\mathcal{P}_2$ .

Olkoon  $x_1, x_2, x_3$  elementin solmupisteitä. Käytännössä muotoa  $\phi_i = a_i x^2 + b_i x + 1$ ,  $i = 1, 2, 3$  oleville kantafunktioille tulee elementin määritelmää mukaillen päteä

$$(3.2) \quad \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \end{pmatrix}.$$

Kuva 3.2: Ensimmäisen, toisen ja kolmannen asteen elementtejä yhdessä ulottuvuudessa



Yhtälön (3.2) matriisilla on käänteismatriisi olettaen, että *determinantti*  $(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) \neq 0$ , eli  $x_1 \neq x_2 \neq x_3$ . Näin ollen jokaiselle kantafunktiolle  $\phi_i = a_i x^2 + b_i x + 1$ ,  $i = 1, 2, 3$  löytyy sopivat vakiot  $a_i, b_i$  ja  $c_i$ . Lisäksi jokainen kantafunktio on aidosti toista astetta, sillä (ei-triviaalilla) lineaarisella kantafunktiolla voi olla tasan yksi nollakohta välillä  $[x_1, x_2]$ .

**Huomautus.** *Esimerkissä 8 ei osoitettu, että kantafunktiot virittää elementtikohdaisesti koko toisen asteen polynomien avaruuden. Tässä tapauksessa se tarkoittaa käytännössä, että konstruoidut kantafunktiot on lineaarisesti riippumattomia. Jätämme kysymyksen (ainakin vielä) auki. Asia ei kuitenkaan päde yleisessä tapauksessa kuten esimerkissä 6 konstruoitu elementti osoittaa.*

**Korollaari 3.9.** *Kuvan 3.2 kolmas elementti määrää kannan (aidosti) avaruudessa  $\mathcal{P}_3$  ja globaali interpolaatio  $I_{\mathcal{T}v}$  on klassisessa mielessä jatkuvasti derivoituva.*

*Todistus.* Vastaavasti kuin esimerkissä 8 tulee etsiä kantafunktiot muotoa  $\phi_i = a_i x^3 + b_i x^2 + c_i x + d_i$ . Saadaan yhtälö

$$\begin{pmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ 3x_1^2 & 2x_1 & 1 & 0 \\ x_2^3 & x_2^2 & x_2 & 1 \\ 3x_2^2 & 2x_2 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_i \\ b_i \\ c_i \\ d_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \delta_{4i} \end{pmatrix},$$

mikä ratkeaa täsmälleen kun determinantti  $(x_1 - x_2)^4 \neq 0$ . Lisäksi kantafunktiot ovat aidosti kolmatta astetta, sillä esimerkiksi kun  $x_1 = 0, x_2 = 1$ , yhtälöllä

$$\begin{pmatrix} x_1^2 & x_1 & 1 \\ 2x_1 & 1 & 0 \\ x_2^2 & x_2 & 1 \\ 2x_2 & 1 & 0 \end{pmatrix} \begin{pmatrix} b_i \\ c_i \\ d_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \delta_{4i} \end{pmatrix}$$

ei ole ratkaisua.

Tarvitsemme seuraavan lemmän, jotta voimme yleistää paikallisia elementtejä koskevan tuloksen koko avaruuteen.

**Lemma 3.10.** *Olkoon  $\tilde{\phi}, \phi \in C^n(\overline{K})$ . Olkoon elementit  $K_1$  ja  $K_2$  affiinisti ekvivalentteja vastavien elementtien  $\tilde{K}_1$  ja  $\tilde{K}_2$  kanssa. Tällöin jos interpolaatiot  $D^\alpha \mathcal{I}_1 v$  ja  $D^\alpha \mathcal{I}_2 v$  yhtyvät elementtien  $K_1$  ja  $K_2$  jossain yhteisissä pisteissä  $x \in \partial K_1 \cap \partial K_2$ , niin myös interpolaatit  $D^\alpha \tilde{\mathcal{I}}_1 v$  ja  $D^\alpha \tilde{\mathcal{I}}_2 v$  yhtyvät vastaavissa pisteissä  $\tilde{x} = F(x) \in \partial \tilde{K}_1 \cap \partial \tilde{K}_2$ .*

*Todistus.* Olkoon  $x \in \partial K_1 \cap \partial K_2$ , millä  $D^\alpha \mathcal{I}_1 v(x) = D^\alpha \mathcal{I}_2 v(x)$ . Olkoon  $\{\phi_1^1, \dots, \phi_1^k\}$  ja  $\{\phi_2^1, \dots, \phi_2^k\}$  elementtien  $K_1$  ja  $K_2$  kantafunktiot. Kaikille  $i = 1, \dots, k$  saadaan suoraan

$$\begin{aligned} D^\alpha \tilde{\phi}_1^i(F(x)) &= D^\alpha \phi_1^i(x) \\ &= D^\alpha \phi_2^i(x) = D^\alpha \tilde{\phi}_2^i(F(x)), \end{aligned}$$

sillä derivaatan määritelmässä affiinisti ekvivalenteille kantafunktioille pätee symbolisesti

$$\lim_{h \rightarrow 0} \frac{(\tilde{\phi} \circ F)(x+h) - (\tilde{\phi} \circ F)(x)}{h} = \lim_{h \rightarrow 0} \frac{\phi(x+h) - \phi(x)}{h}.$$

Tämä todistaa väitteen. □

Riittää osoittaa derivaatan jatkuvuus elementeille välillä  $(-1, 0)$  ja  $(0, 1)$ , jolloin yleinen tulos seuraa lemmasta 3.10. Konstruoidaan kantafunktiot elementeille valituilla väleillä. Suoraan laskemalla (esimerkiksi Matlabilla) saadaan välin  $(-1, 0)$  elementin kantafunktioille polynomien vakioiksi

$$\phi_{1,2,3,4}(x) = \begin{cases} 2x^3 + 3x^2, \\ x^3 + x^2, \\ -2x^3 - 3x^2 + 1, \\ x^3 + 2x^2 + x. \end{cases}$$

Vastaavasti välin  $(0, 1)$  elementille saadaan

$$\varphi_{1,2,3,4}(x) = \begin{cases} 2x^3 - 3x^2 + 1, \\ x^3 - 2x^2 + x, \\ -2x^3 + 3x^2, \\ x^3 - x^2. \end{cases}$$

Summaamalla kantafunktiot saadaan interpolaatioiksi väleillä  $(-1, 0)$  ja  $(0, 1)$



$$\begin{aligned}
(\mathcal{I}v_{(-1,0)})(x) &= v(-1)\phi_1(x) + Dv(-1)\phi_2(x) + v(0)\phi_3(x) + Dv(0)\phi_4(x) \\
&= [Dv(-1) + Dv(0) + 2v(-1) - 2v(0)]x^3 + \\
&\quad [Dv(-1) + 2Dv(0) + 3v(-1) - 3v(0)]x^2 + Dv(0)x + v(0)
\end{aligned}$$

ja

$$\begin{aligned}
(\mathcal{I}v_{(0,1)})(x) &= v(0)\phi_1(x) + Dv(0)\phi_2(x) + v(1)\phi_3(x) + Dv(1)\phi_4(x) \\
&= [Dv(0) + Dv(1) + 2v(0) - 2v(1)]x^3 + \\
&\quad [3v(1) - Dv(1) - 3v(0) - 2v(0)]x^2 + Dv(0)x + v(0).
\end{aligned}$$

Interpolaatioista on helppo todeta, että

$$\lim_{x \rightarrow 0^+} (\mathcal{I}v_{(-1,0)})(x) = \lim_{x \rightarrow 0^-} (\mathcal{I}v_{(0,1)})(x)$$

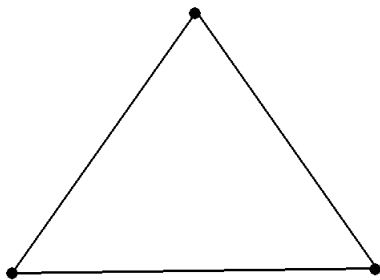
ja

$$\lim_{x \rightarrow 0^+} (D\mathcal{I}v_{(-1,0)})(x) = \lim_{x \rightarrow 0^-} (D\mathcal{I}v_{(0,1)})(x),$$

mikä todistaa väitteen. □

### 3.3.2 Lagrangen tyyppiset elementit

Kuva 3.3: Ensimmäisen asteen Lagrangen elementti kahdessa ulottuvuudessa



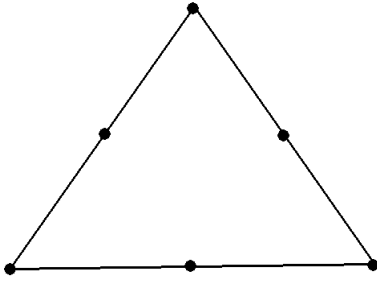
Kappaleessa johdetaan toisen asteen kantafunktiot kaksiulotteiselle Lagrangen elementille, sekä lineaariset kantafunktiot kolmiulotteiselle Lagrangen elementille.

*Esimerkki 9.* Johdetaan kantafunktiot kuvan 3.4 mukaiselle elementille.

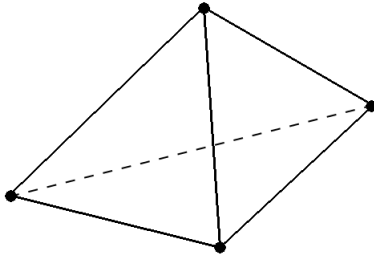
Toisen asteen polynomit kahdessa ulottuvuudessa ovat muotoa

$$(x, y) \mapsto a + bx + cy + dx^2 + ey^2 + fxy.$$

Kuva 3.4: Toisen asteen Lagrangen elementti kahdessa ulottuvuudessa



Kuva 3.5: Ensimmäisen asteen Lagrangen elementti kolmessa ulottuvuudessa



Yksikäsitteisen toisen asteen polynomin määrittämiseen tarvitaan siis yhteensä viisi vakiota  $\{a, b, c, d, e, f\}$ .

Toisen asteen Lagrangen elementti määrittää kuusi kantafunktiota (katso kuva 3.4), joita merkitsemme

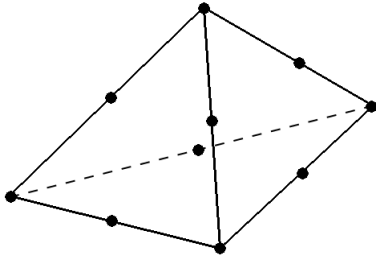
$$\{\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6\}.$$

Kantafunktioilta vaaditaan, että yhdessä solmupisteessä se saa arvon 1 ja muissa pisteissä arvon 0. Merkitsemällä kantafunktiokohtaisia vakioita alaindeksillä  $i = 1, \dots, 6$  ja solmupisteitä  $\{(x_i, y_i) : i = 1, \dots, 6\}$  saadaan matriisiyhtälö

$$(3.3) \quad \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1 y_1 \\ 1 & x_2 & y_2 & x_2^2 & y_2^2 & x_2 y_2 \\ 1 & x_3 & y_3 & x_3^2 & y_3^2 & x_3 y_3 \\ 1 & x_4 & y_4 & x_4^2 & y_4^2 & x_4 y_4 \\ 1 & x_5 & y_5 & x_5^2 & y_5^2 & x_5 y_5 \\ 1 & x_6 & y_6 & x_6^2 & y_6^2 & x_6 y_6 \end{pmatrix} \begin{pmatrix} a_i \\ b_i \\ c_i \\ d_i \\ e_i \\ f_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \delta_{4i} \\ \delta_{5i} \\ \delta_{6i} \end{pmatrix}.$$

Yhtälö 3.3 ei ratkea yleisillä pisteillä  $(x_i, y_i)$ , vaan niiden suhteelliselle sijainnille pitää asettaa varsin tiukkoja ehtoja. Esimerkiksi yli kolme pistettä ei voi sijaita

Kuva 3.6: Toisen asteen Lagrangen elementti kolmessa ulottuvuudessa



samalla suoralla, sillä toisen asteen polynomilla voi olla korkeintaan 2 nollakoh-  
taa kyseisellä suoralla. On kuitenkin osoitettavissa, että toisen asteen Lagrangen  
elementti todella aina määrää yksikäsitteiset kuusi toisen asteen kantapolynomia  
([2]). Yhtälö 3.3 ratkeaa siis ainakin, jos annettujen pisteiden  $(x_i, y_i), i = 1, \dots, 6$   
kautta pystyy ”piirtämään” kolmion, jonka jokainen sivu sisältää kolme pistettä.

Johdetaan kantafunktiot kuvan 3.4 elementille solmupisteillä

$$\begin{aligned} (x_1, y_1) &= (0, 0) & (x_2, y_2) &= (1, 0) \\ (x_3, y_3) &= (0, 1) & (x_4, y_4) &= (0, 1/2) \\ (x_5, y_5) &= (1/2, 0) & (x_6, y_6) &= (1/2, 1/2), \end{aligned}$$

eli solmupisteet on kuvan 3.4 mukaiset pisteet skaalattuna kolmioon  $K \subset \mathbb{R}^2$ , jonka  
kärkipisteet sijaitsevat pisteissä  $(0, 1), (1, 0)$  ja  $(0, 0)$ .

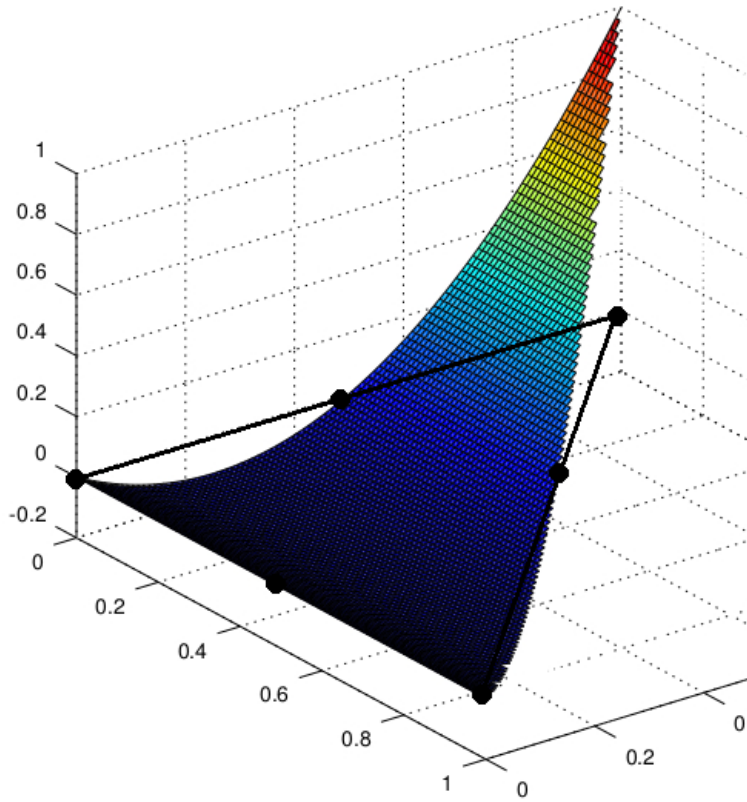
Näin saadaan yhtälö

$$(3.4) \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1/2 & 0 & 1/4 & 0 \\ 1 & 1/2 & 0 & 1/4 & 0 & 0 \\ 1 & 1/2 & 1/2 & 1/4 & 1/4 & 1/4 \end{pmatrix} \begin{pmatrix} a_i \\ b_i \\ c_i \\ d_i \\ e_i \\ f_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \delta_{4i} \\ \delta_{5i} \\ \delta_{6i} \end{pmatrix}.$$

Ratkaisemalla yhtälö (3.4) (esimerkiksi Matlabilla) saadaan kantafunktioiksi

$$\phi_{1,2,3,4,5,6}(x, y) = \begin{cases} 1 - 3x - 3y + 2x^2 + 2y^2 + 4xy, \\ -x + 2x^2, \\ -y + 2y^2, \\ 4y - 4y^2 - 4xy, \\ 4x - 4x^2 - 4xy, \\ 4xy \end{cases}$$

Kuva 3.7: Eräs toisen asteen kantafunktio

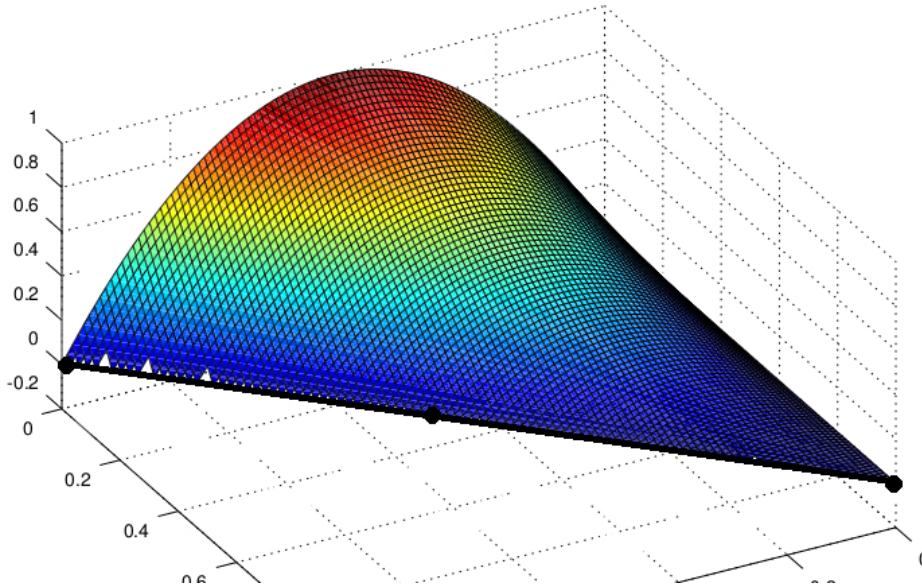


*Esimerkki 10.* Johdetaan lineaariset kantafunktiot kolmiulotteiselle Lagrangen elementille (kuva 3.5).

Oletetaan, että kärkipisteet sijaitsevat pisteissä

$$\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}.$$

Kuva 3.8: Eräs toisen asteen kantafunktio



Ensimmäisen asteen polynomit kolmessa ulottuvuudessa ovat muotoa

$$(x, y, z) \mapsto a + bx + cy + dz,$$

joten saadaan yhtälö

$$(3.5) \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_i \\ b_i \\ c_i \\ d_i \end{pmatrix} = \begin{pmatrix} \delta_{1i} \\ \delta_{2i} \\ \delta_{3i} \\ \delta_{4i} \end{pmatrix}.$$

Ratkaisemalla yhtälö (3.5) saadaan kantafunktioiksi

$$(3.6) \quad \phi_{1,2,3,4}(x, y, z) = \begin{cases} 1 - x - y - z, \\ x, \\ y, \\ z. \end{cases}$$

Käytännössä *vertailuelementtiä* ja *Gaussin vertailupisteitä* (katso luku 4) käytettäessä jokaiselle elementille ei tarvitse erikseen hakea kantafunktioita, vaan vain yhdelle vertailuelementille konstruoidut kantafunktiot riittävät. Toisaalta kantafunktioiden *olemassaolo* yleisillä solmupisteillä on sinänsä oleellinen asia.

**Propositio 3.11.** *Lagrangen elementti on  $C^0$  elementti, eli siihen liittyvä interpolaatio  $\mathcal{I}u$  on jatkuva.*

*Todistus.* Todistetaan tulos kahdessa ulottuvuudessa. Sama logiikka pätee kolmessa ulottuvuudessa.

Olkoon  $\partial K_1 \cap \partial K_2$  jana (tai taso kolmessa ulottuvuudessa), missä elementtien reunat yhtyvät. Olkoon elementit astetta  $n$ . Tällöin jana  $\partial K_1 \cap \partial K_2$  sisältää  $n + 1$  solmupistettä. Määritelmällisesti interpolaatiot molemmissa elementeissä  $K_1$  ja  $K_2$  saavat yhtenevät arvot jokaisessa solmupisteessä, mutta ainoa asteen  $n$  polynomi  $v$  millä on  $n + 1$  nollakohtaa kyseisellä janalla on  $v \equiv 0$ . Tämä todistaa väitteen.  $\square$

# Luku 4

## Tekninen toteutus

Elementtimenetelmässä variaatiotehtävän 2.10 teknisessä toteutuksessa konstruoidaan matriisi  $A$  ja vektori  $b$  tehtävässä

$$(4.1) \quad \begin{aligned} &\text{Etsi } u \in \mathbb{R}^n, \text{ jolle} \\ &Au = b, \\ &\text{missä } A \in \mathbb{R}^{n \times n} \text{ ja } b \in \mathbb{R}^n \text{ on annettu.} \end{aligned}$$

Kun tämä on tehty, voimme suoraan ratkaista  $u = A^{-1}b$  (kunhan käänteismatriisi  $A^{-1}$  tiedetään). Matriisia  $A$  kutsutaan usein historiallisesti *jäykkymatriisiksi* (stiffness matrix) ja vektoria  $b$  *kuormavektoriksi* (load vector). Nimitykset tulevat kehittelyn ensimmäisistä sovelluskohteista lujuuslaskennassa.

Menetelmän toteutuksen voi tietenkin toteuttaa monella tapaa. Tässä käymme läpi toteutuksen ehkä yleisimmin menetellyn filosofian mukaisesti. Esimerkkinä johdamme elementtimenetelmän mallin ensin Poissonin yhtälölle.

### 4.1 Jäykkymatriisi ja kuormavektori Poissonin yhtälölle

Jokaista elementtiä  $T \in \mathcal{T}$  vastaa paikka jäykkymatriisissa  $A$  ja kuormavektorissa  $b$ .

Kerrotaan Poissonin yhtälöä

$$(4.2) \quad -\Delta u = f$$

puolittain testifunktiolla  $v$  ja integroidaan alueen  $U$  yli, jolloin käyttämällä Greenin kaavaa 2.13 saadaan

$$\int_U \nabla u \cdot \nabla v = \int_U f v.$$

Olkoon  $\{\varphi_j\}$  valitut kantafunktiot. Merkitsemällä  $v = \sum b_j \varphi_j$ ,  $u = \sum u_j \varphi_j$  (vertaa esimerkki 1.1) saadaan yhtälö

$$\sum_{j=1}^n u_j \int_U \nabla \varphi_j \cdot \nabla \varphi_i = \int_U \varphi_i f \quad \text{kaikilla } i = 1, \dots, n.$$

Näin ollen jäykkyysmatriisiin indekseiksi saadaan

$$A_{ij} = \int_U \nabla \varphi_i \cdot \nabla \varphi_j,$$

Vastaavasti kuormavektorille saadaan

$$b_i = \int_U \varphi_i f.$$

Käytännössä usein, ja ainakin tässä työssä, kantafunktiot  $\varphi_i$  ovat kompaktikantaisia. Riippuen miten kantafunktiot määritellään, tämä kompakti joukko merkitsee yhtä tai useampaa elementtiä. Kolmiomaisia elementtejä käyttäessä kantafunktion kanta voi koostua esimerkiksi kolmioista (tai tetraedreista kolmessa ulottuvuudessa jne.), joilla on yhteinen solmupiste. Kappaleessa 3.3 on esitelty muutaman tällaisen kantafunktion elementtikohtainen konstruointi.

Jäykkyysmatriille ja kuormasvektorille riittää siis laskea integraali vain jossakin aidossa joukon  $U$  osajoukossa. Kolmiomaisia kantafunktioita käyttäen tämä tarkoittaa käytännössä kolmioita, jotka jakavat saman solmun. Merkitään näitä kolmioita  $K_n \cup \dots \cup K_k$ , missä indeksit  $n, \dots, k$  viittaa alueen kolmioinnissa annettuihin indekseihin. Tällöin jäykkyysmatriisille tulee Poissonin yhtälössä

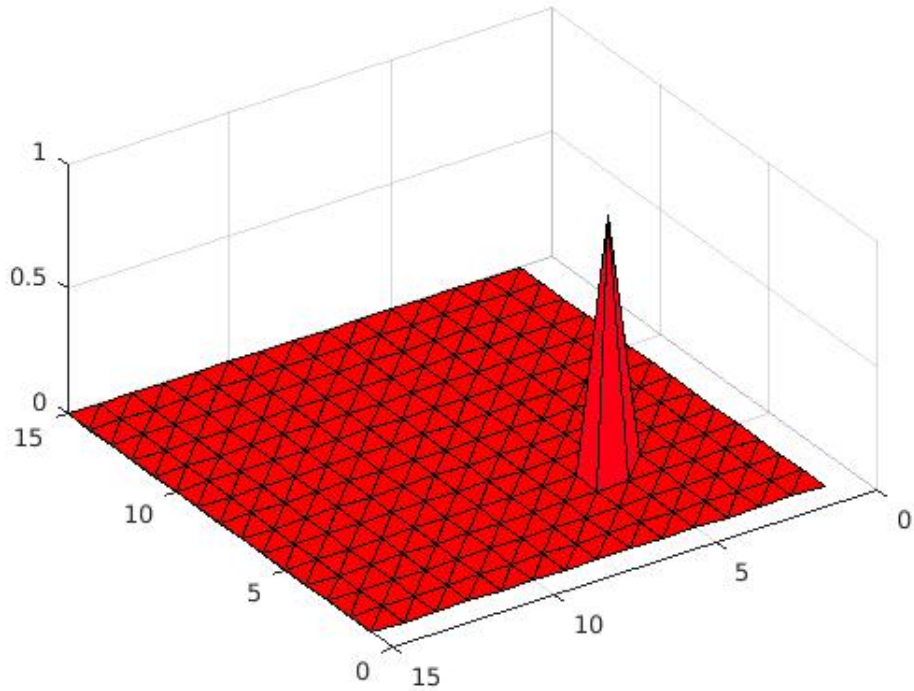
$$\begin{aligned} (4.3) \quad A_{ij} &= \int_U \nabla \varphi_i \cdot \nabla \varphi_j = \int_{K_n \cup \dots \cup K_k} \nabla \varphi_i \cdot \nabla \varphi_j, \\ &= \int_{K_n} \nabla \varphi_i \cdot \nabla \varphi_j + \dots + \int_{K_k} \nabla \varphi_i \cdot \nabla \varphi_j, \end{aligned}$$

ja kuormavektorille vastaavasti

$$(4.4) \quad b_i = \int_U \varphi_i f = \int_{K_n \cup \dots \cup K_k} \varphi_i f = \int_{K_n} \varphi_i f + \dots + \int_{K_k} \varphi_i f.$$

Tietokoneelle toteuttassa algoritmin ideana on käydä elementit  $\{K_i\}$  läpi, ja elementtikohtaisten arvojen lisääminen jäykkyysmatriisiin ja kuormavektoriin yksitellen.





Kuva 4.1: Paloittain lineaarinen kompaktikantainen kantafunktio kahdessa ulottuvuudessa

## 4.2 Jäykkymatriisi ja kuormavektori Helmholtzin yhtälölle

Johdetaan vielä jäykkymatriisi ja kuormavektori Helmholtzin yhtälölle 2.3

$$\begin{aligned} -\Delta u + \lambda u &= f, \\ u &= 0 \text{ reunalla } \partial U. \end{aligned}$$

Esimerkin 4 mukaisesti variaatiotehtävän muodoiksi saadaan

$$a(u, v) = \int_U \nabla u \cdot \nabla \bar{v} + \lambda \int_U u \bar{v}$$

ja

$$L(v) := \int_U f \bar{v}.$$

Merkitsemällä  $v_h = \sum_i a_i \varphi_i$ ,  $u_h = \sum_i u_i \varphi_i$  saadaan jäykkyyismatriisiin ja kurmavektorin indeksikohtaisiksi arvoiksi

$$A_{ij} = \int_U \nabla \varphi_i \cdot \nabla \varphi_j + \lambda \int_U \varphi_i \varphi_j$$

ja

$$b_i = \int_U f \varphi_i.$$

On tärkeää huomata, että kuormavektorista ja jäykkyyismatriisista emme voi lähteä suoraan ratkomaan

$$u = A^{-1}b,$$

sillä  $A$  ei ole edes välttämättä kääntyvä. Matriisista  $A$  ja vektorista  $b$  on vielä eroteltava reunasolmuja vastaavien kantafunktioiden indeksit ja asetettava vaaditut reunaehdot ([1]).

Käytännön toteutusta lineaarisilla kantafunktioilla on käsitelty selkeästi Hannukaisen, Juntusen ja Huhtalan luentomonisteissa ([1]); monisteissa on toteutettu Matlabille muun muassa tässä luvussa käsitellyt muodot. Hyvin edistyneitä algoritmeja on käsitelty esimerkiksi Zienkiewiczin ja Hughesin kirjassa ([7]).

Kuvassa 4.2 on esitetty Helmholtzin yhtälön FEM ratkaisu ensimmäisen asteen Lagrangen elementeillä. Alue  $U$  on ympyrä jonka säde on 10. Lähdefunktio  $f$  on kompaktikantajainen ja aaltoluvuksi on määritelty  $\lambda = -2$ . Yhtälö on ratkaistu Matlabin *pdetool* paketilla.

### 4.3 Numeerinen integrointi

Käsitlemme lyhyesti yleisesti käytetyn numeerisen integrointimenetelmän; niin kutsutut *Gaussin arviointipisteet*.

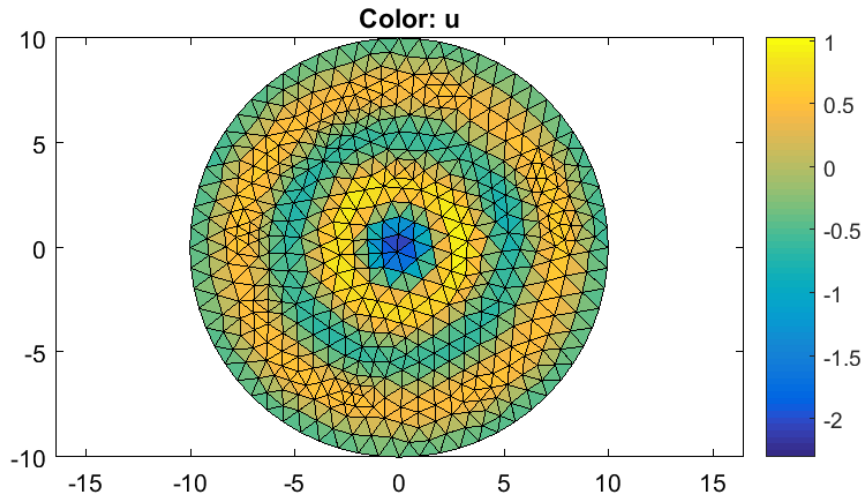
Gaussin arviointipisteiden avulla integraali

$$\int_U u(x) dx$$

saadaan laskettua summana

$$\sum_{i=0}^n a_i u(x_i),$$

missä vakioita  $\{a_i\}$  kutsutaan *painotusvakioiksi* ja pisteitä  $\{x_i\}$  kutsutaan Gaussin arviointipisteiksi. Erityisesti polynomien tapauksessa integraali saadaan jopa laskettua täsmällisesti kunhan vertailupisteitä valitaan tarpeeksi monta. Integrointimenetelmä on hyvin kätevä toteuttaessa FEM algoritmia tietokoneelle.



Kuva 4.2: Helmholtzin yhtälön FEM ratkaisu ensimmäisen asteen Lagrangen elementeillä

*Esimerkki 11.* Merkitään kolmiomaista vertailuelementtiä  $\hat{K} \subset \mathbb{R}^2$  matriisilla

$$\hat{\mathbf{K}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

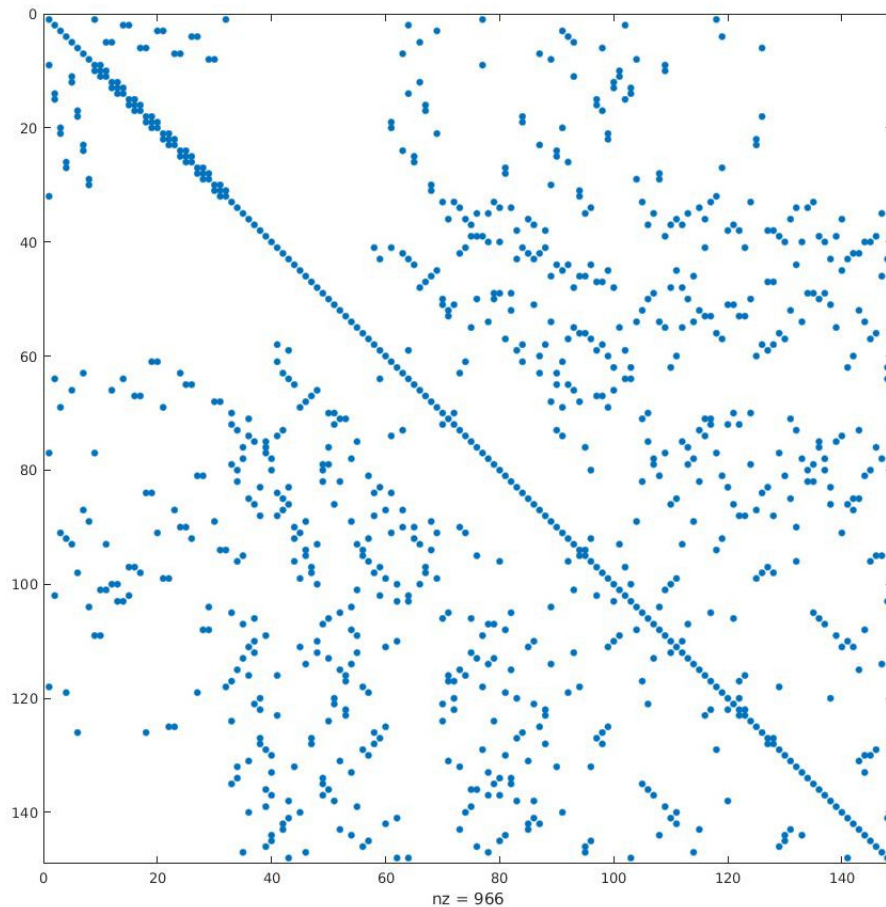
missä sarakkeet  $\hat{K}(i, \cdot)$ ,  $i = 1, 2, 3$  viittaavat solmupisteiden koordinaatteihin.

Arviointivakiot solmupisteille ovat tällöin

$$a_1, a_2, a_3 = 1/6.$$

Kyseiset arviointivakiot ja -pisteet tuottavat täsmällisen arvon integraalille ensimmäisen asteen polynomeilla ([7]).

Gaussin vertailupisteiden käyttöä on käsitelty kattavasti esimerkiksi Zienkiewiczin ja Hughesin kirjassa ([7]). Esimerkin 11 vertailukolmiota ja vertailupisteitä käytetään elementtimenetelmän toteutuksessa Matlabille Hannukaisen, Huhtalan ja Juntusen luentomonisteissa ([1]).



Kuva 4.3: Nollasta poikkeavat elementit eräässä jäykkyyismatriisissa

## 4.4 Vertailuelementti

Jäykkyyismatriisiin ja kuormavektorin arvojen laskeminen elementtikohtaisesti jokaiselle elementille erikseen voi olla käytännössä hankalaa ja laskentatehoa kuluttavaa. Siksi käytämme niin sanottua vertailuelementtiä, mikä on vakiona pysyvä kolmio (tai tetraedi kolmessa ulottuvuudessa jne.). Hyöty tästä on, että voimme laskea vertailuelementille valmiiksi *Gaussin arviointipisteet ja -vakiot*, jolloin integrointi (eli sisätulon laskeminen) onnistuu näiden ennalta laskettujen arvojen avulla jokaiselle elementille. Menetelmän idea on karkeasti seuraavanlainen:

1. Muodosta (affiini) kuvaus  $F$  globaalin elementin  $\tilde{K} \subset \mathbb{R}^n$  ja vertailuelementin  $K$  välille. (Tämä on varsin suoraviivainen tehtävä, eikä kuluta laskentaaikaa.)

2. Laske integraali vertailuelementissä  $K$  ja kuvaa arvo elementtiin  $\tilde{K}$ .

**Määritelmä 4.1.** Affiini kuvaus on muotoa  $F(x) = \mathbf{A}x + \mathbf{b}$ , missä  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ja  $\mathbf{b} \in \mathbb{R}^n$ . Lisäksi vaaditaan, että  $\mathbf{A}$  on kääntyvä.

**Määritelmä 4.2.** Olkoon  $F$  affiini kuvaus. Elementit  $(K, \mathcal{P}, \mathcal{N})$  ja  $(\tilde{K}, \tilde{\mathcal{P}}, \tilde{\mathcal{N}})$  ovat affiinisti ekvivalentteja, jos

- $F(K) = \tilde{K}$ ,
- $F^*\tilde{\mathcal{P}} = \mathcal{P}$ ,
- $F_*\mathcal{N} = \tilde{\mathcal{N}}$ ,

missä  $F^*(\tilde{f}) := \tilde{f} \circ F$  ja  $(F_*\mathcal{N})(f) := \mathcal{N}(F^*(f))$ . Merkintä  $f \circ F$  merkitsee *yhdistettyä kuvausta*, jolle  $(f \circ F)(x) := f(F(x))$ .

**Lause 4.3** (Sijoitusmenetelmä). *Olkoon  $F : K \rightarrow \mathbb{R}^n$  bilipchitzkuvaus, eli on olemassa  $C > 0$ , jolle  $\frac{1}{C}|x_1 - x_2| \leq |F(x_1) - F(x_2)| \leq C|x_1 - x_2|$  kaikilla  $x_1, x_2 \in U$ . Tällöin*

$$\int_{\hat{K}} f \circ F |\det DF| = \int_{F(K)} f.$$

*Merkintää*

$$|\det DF| = \left| \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix} \right|$$

*kutsutaan Jacobianin determinantiksi.*

**Korollaari 4.4.** *Olkoon  $x \in \tilde{K}$ . Ketjusäännön nojalla pätee*

$$\frac{\partial}{\partial x_i} \tilde{f}(x) = \nabla f(F^{-1}(x)) \cdot \left( \left( \frac{\partial}{\partial x_i} F^{-1} \right)_1(x), \dots, \left( \frac{\partial}{\partial x_i} F^{-1} \right)_n(x) \right),$$

missä  $\nabla = f \mapsto \left( \frac{\partial}{\partial x_1} f, \dots, \frac{\partial}{\partial x_n} f \right)$ .

Elementtimenetelmää toteuttaessa tulee käytännössä etsiä kuvaus  $F$ , jolle

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

missä  $\hat{K}$  on vertailuelementti ja  $K$  on jokin mielivaltainen elementti.

Matriisi  $B_K$  ja vektori  $b_K$  löytyvät esimerkiksi merkitsemällä jokaiselle solmupisteelle erikseen  $F(\hat{x}) = x$ , missä  $\hat{x} \in \mathbb{R}^n$  ja  $x \in \mathbb{R}^n$  viittaa kolmioiden  $\hat{K}$  ja  $K$  solmupisteisiin.

*Esimerkki 12.* Merkitään kolmiomaista vertailuelementtiä  $\hat{K} \subset \mathbb{R}^2$  matriisilla

$$\hat{\mathbf{K}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

missä sarakkeet  $K(i, :)$ ,  $i = 1, 2, 3$  merkitsee solmupisteiden koordinaatteja. Vastaavasti määritellään globaali elementti

$$\mathbf{K} = [\mathbf{n}_1 \quad \mathbf{n}_2 \quad \mathbf{n}_3],$$

missä  $\mathbf{n}_i \in \mathbb{R}^2$  merkitsee pystyvektoria. Nyt affiini kuvaus  $F_K$  on määriteltävissä solmupisteittäin

$$F_K(\hat{x}) = [\mathbf{n}_2 - \mathbf{n}_1 \quad \mathbf{n}_3 - \mathbf{n}_1] \hat{x} + \mathbf{n}_1,$$

sillä on helppo todeta, että jokaiselle vertailukolmion  $\hat{\mathbf{K}}$  solmupisteelle  $\hat{\mathbf{K}}(:, i)$ ,  $i = 1, 2, 3$  pätee

$$F(\hat{\mathbf{K}}(:, i)) = \mathbf{n}_i.$$

Edellä käytettiin Matlabin notaatiota, missä  $\mathbf{K}(:, i)$  tarkoittaa matriisin  $\mathbf{K}$  saraketta indeksillä  $i$ .

*Esimerkki 13.* Merkitään 3-ulotteisen tetraedin kärkipisteitä matriisilla

$$\hat{\mathbf{K}} = \begin{bmatrix} \beta & \alpha & \beta & \beta \\ \beta & \beta & \alpha & \beta \\ \beta & \beta & \beta & \alpha \end{bmatrix},$$

ja vastaavasti globaalia elementtiä

$$\mathbf{K} = [\mathbf{n}_1 \quad \mathbf{n}_2 \quad \mathbf{n}_3 \quad \mathbf{n}_4].$$

Otetaan hahmottamisen helpottamiseksi käyttöön apuelementti  $\hat{\mathbf{K}}'$  :

$$\hat{\mathbf{K}}' = \hat{\mathbf{K}} - \begin{bmatrix} \beta & \beta & \beta & \beta \\ \beta & \beta & \beta & \beta \\ \beta & \beta & \beta & \beta \end{bmatrix} = \begin{bmatrix} 0 & \alpha - \beta & 0 & 0 \\ 0 & 0 & \alpha - \beta & 0 \\ 0 & 0 & 0 & \alpha - \beta \end{bmatrix}.$$

Oletetaan, että  $\tilde{x} \in \hat{\mathbf{K}}'$ . Affiiniksi kuvaukseksi apuelementille  $\hat{\mathbf{K}}'$  on helppo todeta kuvaus

$$F_K(\hat{x}) = \frac{[\mathbf{n}_2 - \mathbf{n}_1 \quad \mathbf{n}_3 - \mathbf{n}_1 \quad \mathbf{n}_4 - \mathbf{n}_1]}{\alpha - \beta} \tilde{x} + \mathbf{n}_1.$$

Tämä on yhtäpitävää yhtälön

$$F_K(\hat{x}) = \frac{[\mathbf{n}_2 - \mathbf{n}_1 \quad \mathbf{n}_3 - \mathbf{n}_1 \quad \mathbf{n}_4 - \mathbf{n}_1]}{\alpha - \beta} \left[ \hat{x} - \begin{bmatrix} \beta \\ \beta \\ \beta \end{bmatrix} \right] + \mathbf{n}_1$$

kanssa, missä  $\hat{x} \in \hat{\mathbf{K}}$ . Saadaan

$$F_K(\hat{x}) = \frac{[\mathbf{n}_2 - \mathbf{n}_1 \quad \mathbf{n}_3 - \mathbf{n}_4 \quad \mathbf{n}_4 - \mathbf{n}_1]}{\alpha - \beta} \hat{x} - [\mathbf{n}_1 - [\frac{\mathbf{n}_2 - \mathbf{n}_1}{\alpha - \beta} \beta + \frac{\mathbf{n}_3 - \mathbf{n}_1}{\alpha - \beta} \beta + \frac{\mathbf{n}_4 - \mathbf{n}_1}{\alpha - \beta} \beta]].$$

Esimerkeissä 13 ja 12 käsitellyt elementit on yleisessä käytössä olevia vertailuelementtejä kolmiomaista ositusta käytettäessä ([7]).

#### 4.4.1 Matemaattisia käsitteitä

Esitellään muutamia elementtien muotoon liittyviä hyödyllisiä käsitteitä. Osa esiteltävistä käsitteistä ja tuloksista on aikalailla abstrakteja (ja sinänsä intuitiivisesti selviä), mutta ovat ehkä muodollisen todistuksen arvoisia kuitenkin. Menetelmän toteuttamisen kannalta seuraavat tulokset eivät siis ole välttämättä kovin oleellisia, mutta ne rakentavat vankkaa matemaattista pohjaa menetelmän teorialle. Monia määritelmiä tarvitaan myös virheanalyysissä luvussa 5.

**Määritelmä 4.5.** (Konvekssi joukko) Joukko  $U \subset \mathbb{R}^n$  on konvekssi, mikäli  $x, y \in U$  jos ja vain jos  $\alpha(x - y) + y \in U$  kaikilla  $\alpha \in [0, 1]$ . Toisin sanoen mielivaltaisia pisteitä joukossa  $U$  yhdistää aina jana, mikä sisältyy itse joukkoon  $U$ .

**Määritelmä 4.6** (Konvekssi verho). Olkoon  $A \subset \mathbb{R}^n$ . Konvekssi verho on pienin konvekssi ja suljettu joukko, mikä sisältää joukon  $A$ . Jos pisteitä on äärellinen määrä, eli  $A = \{x_1, \dots, x_k\} \subset \mathbb{R}^n$ , konvekssi verho on joukko

$$(4.5) \quad \left\{ \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k : \sum_{i=1}^k \alpha_i = 1, \alpha_i \geq 0 \right\}.$$

**Määritelmä 4.7** (Säteittäisesti konvekssi joukko). Olkoon  $B(x_0, \rho) = \{x : |x - x_0| < \rho\}$ . Joukko  $U \subset \mathbb{R}^n$  on radiaalisesti konvekssi, jos jollakin pallolla  $B(x_0, \rho) \subset U$  joukon  $B(x_0, \rho) \cup \{x\}$  konvekssi verho sisältyy joukkoon  $U$  kaikilla  $x \in U$ .

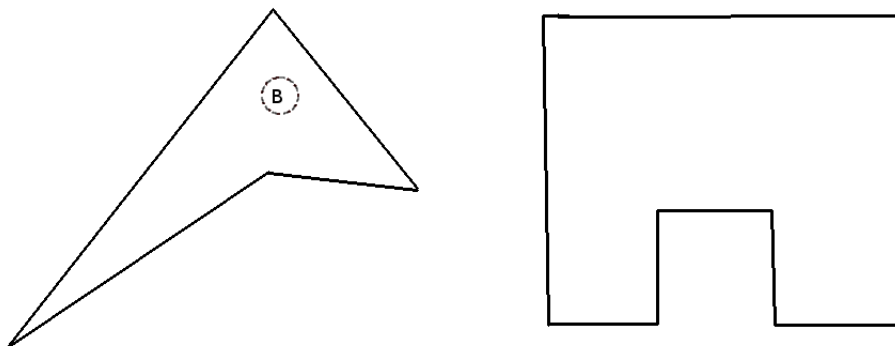
**Lemma 4.8.** *Konvekssi joukko kuvautuu konveksiksi joukoksi affiinissa kuvauksessa  $F = x \mapsto Ax + b$ .*

*Todistus.* Merkitään  $F(x) = \tilde{x}$ . Olkoon  $\alpha \in [0, 1]$ . Vektori  $\alpha(x - y) + y$  kuvaa pistettä pisteiden  $x, y$  välisellä janalla. Kuvauksen  $A$  lineaarisuuden nojalla

$$\begin{aligned} F(\alpha(x - y) + y) &= A(\alpha(x - y) + y) + b \\ &= \alpha Ax - \alpha Ay + Ay + b \\ &= \alpha(\tilde{x} - b) - \alpha(\tilde{y} - b) + \tilde{y} \\ &= \alpha(\tilde{x} - \tilde{y}) + \tilde{y}, \end{aligned}$$

joten kaikkia pisteitä  $\tilde{x}, \tilde{y} \in F(K)$  yhdistää jana, mikä sisältyy joukkoon  $F(K)$ .  $\square$

Kuva 4.4: Pallon  $B$  suhteen radiaalisesti konvekssi elementti, ja ei-radiaalisesti konvekssi elementti



**Propositio 4.9.** *Kuvaus  $Fx = Ax + b$  kuvaa radiaalisesti konveksit joukot radiaalisesti konveksiksi joukoksi.*

*Todistus.* Proposition 4.8 nojalla joukon  $B(x_0, \rho) \cup \{x\}$  konveksin verhon  $V$  kuva  $F(V)$  on konvekssi kaikilla  $x \in K$ . Kuvaus  $F$  on kääntyvä (eli kuvien dimensiot säilyvät kuvauksessa), joten voidaan valita säde  $\epsilon$ , jolla  $B(F(x_0), \epsilon) \subset F(B(x_0, \rho))$ , jolloin kuva  $F(K)$  on radiaalisesti konvekssi pallon  $B(F(x_0), \epsilon)$  suhteen, sillä jatkuvuuden nojalla alkukuva  $F^{-1}B(F(x_0), \epsilon) \subset B(x_0, \rho)$ .  $\square$

Seuraava tulos on oleellinen varsinkin tietokonealgoritmia toteuttaessa, sillä standardielementit on solmupisteidensä määrittämiä konvekseja verhoja, esimerkiksi kolmioita. Tällöin voidaan perustellusti konstruoida kuvaus  $F$  esimerkkien 12 ja 13 mukaisesti.

**Korollaari 4.10.** *Olkoon elementit  $K$  ja  $\tilde{K}$  solmupisteidensä konvekseja verhoja, eli elementit määrittyvät solmupisteidensä kautta kaavalla (4.5). Tällöin jos löytyy kuvaus  $F = x \mapsto Ax + b$ , jolla kaikilla solmupisteillä pätee  $F(n_i) = \tilde{n}_i$ ,  $i = 1, \dots, k$ , niin  $F(K) = \tilde{K}$ .*

*Todistus.* Mielivaltaisille  $\alpha_1, \dots, \alpha_k$ , joilla  $\sum \alpha_i = 1$ ,  $\alpha_i \geq 0$ , pätee

$$\begin{aligned} F(\alpha_1 n_1 + \alpha_2 n_2 + \dots + \alpha_k n_k) &= \alpha_1 A n_1 + \alpha_2 A n_2 + \dots + \alpha_k A n_k + b \\ &= \alpha_1 (\tilde{n}_1 - b) + \alpha_2 (\tilde{n}_2 - b) + \dots + \alpha_k (\tilde{n}_k - b) + b \\ &= \alpha_1 \tilde{n}_1 + \alpha_2 \tilde{n}_2 + \dots + \alpha_k \tilde{n}_k - b \sum_i^k \alpha_i + b \\ &= \alpha_1 \tilde{n}_1 + \alpha_2 \tilde{n}_2 + \dots + \alpha_k \tilde{n}_k, \end{aligned}$$

joten  $F(K)$  on tasan solmupisteiden  $\tilde{n}_1, \dots, \tilde{n}_k$  konvekssi verho  $\tilde{K}$ .  $\square$



# Luku 5

## A priori suppenemisominaisuudet

Saattaaksemme elementtimenetelmän matemaattisesti vankalle pohjalle herää vielä kysymys FEM suppenemisominaisuuksista, eli kuinka nopeasti ratkaisu lähestyy täsmällistä ratkaisua, vai lähestyykö se sitä ollenkaan. Luvussa johdamme konstruktiiivisesti *a-priori* virhearviot FEM ratkaisulle. Luvussa merkinnällä  $u$  tarkoitetaan variaatiotehtävän 2.10 ratkaisua funktioavaruudessa  $W$ , ja vastaavasti merkinnällä  $u_h$  merkitään ratkaisua aliavaruudessa  $W_h \subset W$ .

Virheanalyysi seuraa vahvasti Brennerin ja Scottin kirjaa ([2]). Kirjassa on myös käsitelty virheteoriaa vielä yleisemmin, kuin tässä työssä.

**Lemma 5.1** (Céan lemma 1). *Jos  $a : W \times W \rightarrow \mathbb{K}$  on rajoitettu ja bilineaarinen muoto, niin joillekin  $C, c > 0$  pätee*

$$\|u - u_h\|_W \leq \frac{C}{c} \|u - w_h\|_W$$

kaikilla  $w_h \in W_h$ .

*Todistus.* Galerkin ortogonaalisuuden, koersiivisuuden ja rajoittuneisuuden nojalla

$$\begin{aligned} c\|u - u_h\|^2 &\leq a(u - u_h, u - u_h) = a(u - u_h, u - w_h) + a(u - u_h, w_h - u_h) \\ &= a(u - u_h, u - w_h) \leq C\|u - u_h\|\|u - w_h\|, \end{aligned}$$

kaikilla  $w_h \in W_h$ . □

Oletetaan, että muoto  $a$  on koersiivinen. Jos määrittelemme normin  $\|\cdot\|_a = w \mapsto \sqrt{a(w, w)}$ , vastaavalla tavalla *Cauchy-Schwarzin* epäyhtälöstä seuraa suoraan:

**Lemma 5.2** (Céan lemma 2).

$$\|u - u_h\|_a \leq \|u - w_h\|_a \text{ kaikilla } w_h \in W_h .$$

Normia  $\|\cdot\|_a$  kutsutaan usein *energianormiksi*.

Tutkimalla ratkaisun  $u$  lineaarista interpolaatiota voidaan huomata suoraan Céan lemmasta, että FEM ratkaisu suppenee kohti täsmällistä ratkaisua ainakin tietyissä funktioavaruuksissa jakoa tihentäessä.

**Määritelmä 5.3.** Funktion  $u$  lineaarinen interpolaatio on *paloittain lineaarinen* funktio  $\mathcal{I}_{lin}u$ , jolle

$$(\mathcal{I}_{lin}u)(x_n) = u(x_n) \text{ kaikilla } x_n \in \mathcal{X},$$

missä  $\mathcal{X}$  on alueen  $U$  osituksen solmupisteiden joukko (käytännössä usein elementtien kärkipisteet, mutta muitakin mahdollisuuksia on ([2])).

Nyt Céan lemmasta seuraa suoraan

$$\|u - u_h\|_a \leq \|u - \mathcal{I}_{lin}u\|_a,$$

josta huomataan, että ositusta tihentäessä kun  $\mathcal{I}_{lin}u \rightarrow u$ , niin myös  $u_h \rightarrow u$  normin  $\|\cdot\|_a$  suhteen. Käsittelemme suppenemista vielä tarkemmin tässä luvussa.

*Esimerkki 14.* Johdantoluvun 1 esimerkissä saadaan täsmällisen ratkaisun ja FEM ratkaisun  $u_h$  eroksi

$$\|u - u_h\|_a = \sqrt{a(u - u_h, u - u_h)} = \sqrt{\int_0^{2\pi} (u - u_h)'(x)dx} = 0.2927,$$

ja nollavektorille vastaavasti

$$\|u - 0\|_a = 0.3545,$$

eli  $\|u - u_h\|_a \leq \|u\|_a$ , niin kuin teoria ennustaa.

## 5.1 Sobolevin kehitelmä

Kappaleessa esiteltävä *Sobolevin kehitelmä* on Taylorin polynomin yleistys Sobolevin avaruuksiin, missä funktioiden pisteittäisiä arvoja ei ole yleisesti hyvin määritelty.

**Määritelmä 5.4** (Leikkausfunktio). Olkoon  $B = \{x \in \mathbb{R}^n : |x - x_0| < \rho\}$ . Kutsumme funktiota  $\phi \in C_0^\infty(\mathbb{R}^n)$  *katkaisufunktioksi* (cut off function), jos kantajalle pätee

$$\text{supp } \phi = \overline{B},$$

ja integraalille pätee

$$\int_{\mathbb{R}^n} \phi(x)dx = 1.$$

**Määritelmä 5.5** (Sobolevin kehitelmä). Määritellään keskiarvoistettu Taylorin kehitelmä (Averaged Taylor Polynomial), tai Sobolevin kehitelmä, kaavalla

$$Q^m(x) := \int_B T_y^m(x) \phi(y) dy,$$

missä

$$T_y^m(x) := \sum_{|\alpha| < m} \frac{1}{\alpha!} D^\alpha u(y) (x - y)^\alpha.$$

Multi-indekseille on määritelty  $\alpha = (\alpha_1, \dots, \alpha_n)$ ,  $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$ ,  $\alpha! = \prod_{i=1}^n \alpha_i!$  ja  $|\alpha| = \sum_{i=1}^n \alpha_i$ .

Jos kirjoitamme

$$(x - y)^\alpha = \prod_{i=1}^n (x_i - y_i)^{\alpha_i} = \sum_{|\gamma| + |\beta| = |\alpha|} a_{(\gamma, \beta)}^\alpha x^\gamma y^\beta,$$

missä  $\beta$  ja  $\gamma$  ovat ei-negatiivisia multi-indeksejä, ja  $a_{(\gamma, \beta)}^\alpha x^\gamma$  ovat muuttujasta  $y$  riippumattomia vakioita, voidaan Sobolevin kehitelmä esittää muodossa

$$(5.1) \quad Q^m u(x) = \sum_{|\alpha| < m} \sum_{|\gamma| + |\beta| = |\alpha|} \frac{1}{\alpha!} a_{(\gamma, \beta)}^\alpha x^\gamma \int_B D^\alpha u(y) y^\beta \phi(y) dy.$$

Osittaisintegroimalla kaava (5.1) voidaan myös esittää muodossa

$$(5.2) \quad Q^m u(x) = \sum_{|\alpha| < m} \sum_{|\gamma| + |\beta| = |\alpha|} \frac{(-1)^\alpha}{\alpha!} a_{(\gamma, \beta)}^\alpha x^\gamma \int_B u(y) D^\alpha (y^\beta \phi(y)) dy,$$

joten Sobolevin kehitelmä voidaan määritellä funktioille avaruudessa  $L^1(B)$ .

**Korollaari 5.6.** *Kehitelmä  $Q^m u$  on rajoitettu kuvaus avaruudesta  $L^1(B)$  avaruuteen  $W^{\infty, k}(U)$ , kun  $k \in \mathbb{N}$ .*

*Todistus.* Arvot

$$\sup_{x \in U} |D^\alpha x^\lambda|$$

on rajoitettuja kaikilla rajoitetuilla multi-indekseillä, joten

$$\|Q^m u\|_{W^{\infty, k}(U)} \leq C_{m, n, \rho} \|u\|_{L^1(B)}.$$

□

Koska  $W^{p, k}(U) \subset L^1(U)$  kaikilla  $k, p \in \mathbb{N}$ , niin korollaari 5.6 takaa, että Sobolevin kehitelmä on yleisesti hyvin määritelty.

**Propositio 5.7.** *Sobolevin kehitelmän derivaatoille pätee*

$$D^\alpha Q^m u = Q^{m-|\alpha|} D^\alpha u \text{ kaikilla } u \in W^{1,|\alpha|}(B).$$

*Todistus.* Yhtäläisyys

$$(5.3) \quad D_x^\alpha T_y^m u(x) = T^{m-|\alpha|} D^\alpha u(x)$$

pätee, nimittäin

$$\begin{aligned} D_x^\alpha T_y^m u(x) &= \sum_{|\beta| < m} \frac{1}{\beta!} D^\beta u(y) D_x^\alpha (x-y)^\beta \\ &= \sum_{|\beta| < m} \frac{1}{\beta!} D^\beta u(y) \frac{\beta!}{(\beta-\alpha)!} (x-y)^{\beta-\alpha} \\ &= \sum_{|\beta| < m} \frac{1}{(\beta-\alpha)!} D^\beta u(y) (x-y)^{\beta-\alpha} \quad (\gamma = \beta - \alpha) \\ &= \sum_{|\gamma| < m-|\alpha|} \frac{1}{\gamma!} D^{\beta+\alpha} u(y) (x-y)^\gamma \\ &= \sum_{|\gamma| < m-|\alpha|} \frac{1}{\gamma!} D^\beta D^\alpha u(y) (x-y)^\gamma \\ &= T^{m-|\alpha|} D_x^\alpha u(x). \end{aligned}$$

Oletetaan, että  $u \in C^\infty(U) \cap W^{p,k}(U)$ , jolloin  $D_x^\alpha T_y^m$  on hyvin määritelty. Saadaan

$$\begin{aligned} D_x^{|\alpha|} Q^m u(x) &= \int_B D_x^\alpha T_y^m u(x) \phi(y) dy = \\ &= \int T_y^{m-|\alpha|} D_x^\alpha u(x) \phi(y) dy \\ &= Q^{m-|\alpha|} D^\alpha u(x). \end{aligned}$$

Käytetään hyväksi ominaisuutta, että

$$C^\infty(U) \cap W^{p,k}(U)$$

on tiheä avaruudessa  $W^{p,k}(U)$ . Todistuksen ideaa kutsutaan usein *tiheysargumentiksi* (density argument):

Tiheysominaisuuden nojalla jokaiselle alkiolle  $u \in W^{1,|\alpha|}(U)$  voidaan liittää *Cauchy jono*  $(u_j) \subset C^\infty(U) \cap W^{1,|\alpha|}(U)$ , jolle  $\lim_{j \rightarrow \infty} u_j \rightarrow u$ . Tällöin kuvauksen  $Q^m$  lineaarisuuden ja korollaarin 5.6 nojalla

$$\begin{aligned}
\|Q^{m-|\alpha|}D^\alpha u_n - Q^{m-|\alpha|}D^\alpha u_k\|_{W^{\infty,|\alpha|}(U)} &= \|D_x Q^m u_n - Q^{m-|\alpha|}D^\alpha u_k\|_{W^{\infty,|\alpha|}(U)} \\
&\leq C_{m,n,\rho} \|u_n - u_k\|_{L^1(B)} \\
&\leq C_{m,n,\rho} \|u_n - u_k\|_{W^{1,|\alpha|}(B)} < \epsilon,
\end{aligned}$$

joten kuvauksen  $Q^m : L^1(B) \rightarrow W^{\infty,|\alpha|}(U)$  jatkuvuudesta seuraa  $D_x Q^m u - Q^{m-|\alpha|}D^\alpha u = 0$ .

□

*Esimerkki 15.* Olkoon  $n = 1$  ja  $m = 3$ . Nyt

$$\begin{aligned}
DQ_y^3 u(x) &= \int_B D_x(u(y) + u'(y)(x-y)^1 + \frac{1}{2}u''(y)(x-y)^2)\phi(y)dy \\
&= \int_B (u'(y) + u''(y)(x-y))\phi(y)dy \\
&= Q^2 Du(x).
\end{aligned}$$

## 5.2 Virhetermi

Kappaleessa tutkimme Sobolevin kehittämän virhetermiä

$$R^m u(x) = u(x) - Q^m u(x).$$

Johdamme teorian säteittäisesti konveksille alueelle (katso määritelmä 4.7). Vaikka standardielementit ovat konvekseja, yleisessä tapauksessa muunkinlaiset elementit ovat käytännöllisiä – esimerkiksi reunaelementeillä voi olla kaareva reuna. ([7]).

**Lause 5.8** (Taylorin lause). *Taylorin lause avaruudessa  $f \in C^m([0, 1])$  kohdassa  $x = 1$  saa muodon*

$$f(1) = \sum_{k=0}^{m-1} \frac{1}{k!} f^{(k)}(0) + m \int_0^1 \frac{1}{m!} s^{m-1} f^{(m)}(1-s) ds.$$

*Todistus.* *Analyysin peruslauseen* ([6]) *nojalla*

$$f(1) = f(0) + \int_0^1 f'(s) ds,$$

mistä osittaisintegroimalla saadaan

$$\begin{aligned}
 f(1) &= f(0) + \int_0^1 f'(s) \\
 &= f(0) + f'(1) - \int_0^1 s f''(s) ds \quad (\text{analyysin peruslause}) \\
 &= f(0) + f'(0) + \int_0^1 f''(s) ds - \int_0^1 s f''(s) ds \\
 (5.4) \quad &= f(0) + f'(0) + \int_0^1 (1-s) f''(s) ds.
 \end{aligned}$$

Osoitetaan loput induktiolla. Oletetaan, että

$$(5.5) \quad f(1) = f(0) + \sum_{k=0}^{m-1} \frac{1}{k!} f^{(k)}(0) + m \int_0^1 \frac{1}{m!} (1-s)^{(m-1)} f^{(m)}(s) ds$$

jollakin  $m$ . Osittaisintegroimalla jäännöstermiä saadaan

$$\begin{aligned}
 &\int_0^1 \frac{1}{(m-1)!} (1-s)^{(m-1)} f^{(m)}(s) ds = - \int_0^1 \frac{1}{(m-1)!} s^{(m-1)} f^{(m)}(1-s) ds \\
 &= \int_0^1 \frac{1}{(m-1)! m} (1-s)^m f^{(m)}(s) + \int_0^1 \frac{1}{(m-1)! m} s^m f^{(m-1)}(1-s) ds \\
 &= \frac{1}{(m-1)! m} f^{(m)}(0) + \int_0^1 \frac{1}{(m-1)! m} (1-s)^m f^{(m+1)}(s) ds \\
 &= \frac{1}{m!} f^{(m)}(0) + (m+1) \int_0^1 \frac{1}{(m+1)!} (1-s)^m f^{(m+1)}(s) ds,
 \end{aligned}$$

mikä todistaa väitteen induktioperiaatteella oletuksen (5.5) ja kaavan (5.4) nojalla.  $\square$

Oletetaan  $u \in C^m(U)$ . Jos oletamme, että  $U$  on säteittäisesti konvekksi, niin  $y + s(x - y) \in U$  kaikilla  $x \in U$  ja  $y \in B$ ,  $s \in [0, 1]$ . Merkitsemällä  $f(s) = u(y + s(x - y))$ , saadaan ketjusäännön nojalla

$$(5.6) \quad \frac{1}{k!} f^{(k)}(s) = \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha u(y + s(x - y)) (x - y)^\alpha,$$

mikä sijoittamalla Taylorin lauseeseen 5.8 tuottaa kaavan

$$\begin{aligned}
 (5.7) \quad u(x) &= f(1) = \sum_{|\alpha| < m} \frac{1}{\alpha!} D^\alpha u(y) (x - y)^\alpha \\
 &+ \sum_{|\alpha|=m} (x - y)^\alpha \int_0^1 \frac{m}{\alpha!} s^{m-1} D^\alpha u(x + s(y - x)) ds \\
 &= T_y^m u(x) + m \sum_{|\alpha|=m} (x - y)^\alpha \int_0^1 \frac{1}{\alpha!} s^{m-1} D^\alpha u(x + s(y - x)) ds.
 \end{aligned}$$

Jäännöstermiä  $R^m u(x) = u(x) - Q^m u(x)$  on käytännöllistä tutkia näiden esitystapojen pohjalta.

Sobolevin polynomin määritelmästä 5.5, kaavasta 5.7 ja leikkausfunktion  $\phi$  ehdosta  $\int_U \phi = 1$  seuraa

$$\begin{aligned}
 (5.8) \quad R^m u(x) &= \int_B u(x) \phi(y) - \int_B T_y^m u(x) \phi(y) dy \\
 &= \int_B [u(x) - T_y^m u(x)] \phi(y) \\
 &= \int_B \phi(y) \left( \sum_{|\alpha|=m} (x - y)^\alpha \int_0^1 \frac{s^{m-1}}{\alpha!} D^\alpha u(x + s(y - x)) ds \right) dy
 \end{aligned}$$

**Propositio 5.9.** *Jäännöstermi  $R^m = u - Q^m$  toteuttaa*

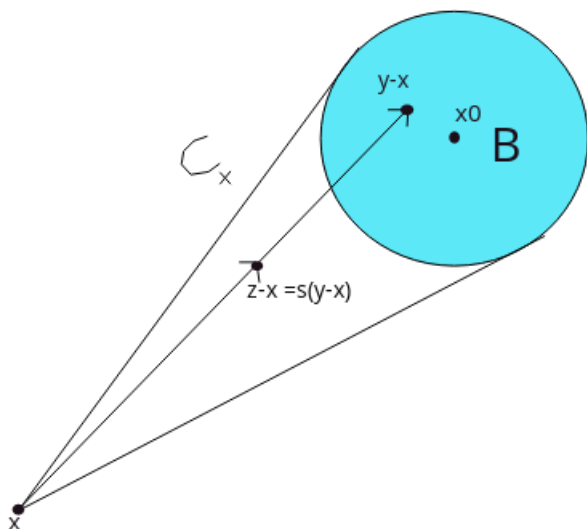
$$R^m u(x) = m \sum_{|\alpha|=m} \int_{C_x} k_\alpha(x, z) D^\alpha u(z) dz,$$

missä  $z = x + s(y - x)$ ,  $k_\alpha(x, z) = (1/\alpha!)(x - z)^\alpha k(x, z)$  ja

$$|k(x, z)| \leq C \left( 1 + \frac{1}{\rho} |x - x_0| \right)^n |z - x|^{-n}.$$

*Todistus.* Tehdään muuttujanvaihdos  $z = x + s(y - x)$ , jolloin lauseen 4.3 mukainen muodollinen differentiaalinen muutoskaava on  $ds dy = s^{-n} ds dz$ , sillä

$$\begin{aligned}
 |\det D_z ((1/s)(z - x) + y)| &= \left| \begin{bmatrix} \frac{\partial((1/s)(z_1 - x_1) + y_1)}{\partial z_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{\partial((1/s)(z_n - x_n) + y_n)}{\partial z_n} \end{bmatrix} \right| \\
 &= \left| \begin{bmatrix} s^{-1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & s^{-1} \end{bmatrix} \right| = s^{-n}.
 \end{aligned}$$



Kun muuttujat  $y$  ja  $s$  käyvät koko integroimisalueen  $B \times (0, 1]$  läpi, käy muuttuja  $z$  läpi joukon

$$C_x = \{z : (z - x) = s(y - x) \mid y \in B, s \in (0, 1]\}.$$

Uusi integroimisalue on siten joukko

$$A = \{(z, s) : s \in (0, 1], |(1/s)(z - x) + x - x_0| < \rho\}.$$

Muuttujalle  $s$  saadaan alaraja joukossa  $A$ , sillä jos  $(z, s) \in A$ , niin

$$(5.9) \quad \frac{|z - x|}{|x - x_0| + \rho} < s,$$

mikä seuraa kolmioepäyhtälöstä:

$$\begin{aligned} |z - x| &= s|y - x| = s|y - x_0 - (x - x_0)| \\ &\leq s|y - x_0| + s|x - x_0| = s\rho + s|x - x_0|. \end{aligned}$$

Toisaalta yläraja on  $s = 1$ , sillä tällöin  $z = y$ . Lisäksi

$$(5.10) \quad (x - y)^\alpha = s^{-m}(x - z)^\alpha.$$

Sijoittamalla muuttujanvaihdos ja kaava (5.10) jäännöstermin kaavaan (5.8), saadaan *Fubinin lauseen* nojalla

$$(5.11) \quad \begin{aligned} R^m u(x) &= m \sum_{|\alpha|=m} \int_{C_x} \frac{1}{\alpha!} D^\alpha u(z) (x - z)^\alpha \\ &\times \left[ \int_0^1 \chi_A(z, s) \phi(x + (1/s)(z - x)) s^{-n-1} ds \right] dz, \end{aligned}$$



missä  $\chi_A(z, s)$  on joukon  $A$  *karakteristinen funktio*. Jos määritellään  $k(x, z) = \int_0^1 \chi_A(z, s) \phi(x + (1/s)(z - x)) s^{-n-1} ds$  ja  $k_\alpha(x, z) = (1/\alpha!)(x - z)^\alpha k(x, z)$ , tiivistyy kaava (5.11) muotoon

$$m \sum_{|\alpha|=m} \int_{C_x} k_\alpha(x, z) D^\alpha u(z) dz,$$

mikä osoittaa proposition ensimmäisen osan. Vielä tulee todistaa jälkimmäinen arvio arvolle  $k(x, z)$ . Kolmioepäyhtälön nojalla

$$\begin{aligned} |k(x, z)| &= \left| \int_0^1 \chi_A(z, s) \phi(x + (1/s)(z - x)) s^{-n-1} ds \right| \\ &\leq \int_t^1 |\phi(x + (1/s)(z - x)) s^{-n-1}| ds \quad (\text{kaava (5.9)}) \\ &\leq \|\phi\|_{L^\infty(B)} \int_t^1 \frac{s^{-n}}{n} \\ &\leq (1/n) \|\phi\|_{L^\infty(B)} t^{-n} \\ &= (1/n) \|\phi\|_{L^\infty(B)} (\rho + |x - x_0|)^n |z - x|^{-n} \\ &\leq C \rho^{-n} (\rho + |x - x_0|)^n |z - x|^{-n} \quad (\text{valitaan } C \geq \max |\phi|/\rho^{-n}) \\ &:= C(1 + \frac{1}{\rho} |x - x_0|)^n |z - x|^{-n}. \end{aligned}$$

□

Tutkitaan seuraavaksi alueen  $U$  muodon vaikutusta suppenemisominaisuuksiin. Ilmenee, että osituksen elementtien muodolla on oleellinen vaikutus ratkaisun suppenemiseen.

**Määritelmä 5.10** (Paksuusparametri). Oletetaan, että  $U$  on radiaalisesti konvekksi pallon  $B(x_0, \rho)$  suhteen. Oletetaan, että joukon  $U$  *läpimitta* on  $d$ . Olkoon  $\rho_{max} = \sup\{\rho : U \text{ on radiaalisesti konvekksi jonkin } \rho \text{ säteisen pallon suhteen}\}$ . Tällöin joukon  $U$  paksuusparametri (chunkiness parameter) määritellään

$$\gamma = \frac{d}{\rho_{max}}.$$

**Korollaari 5.11.** *Pallo  $B$  voidaan valita niin, että proposition 5.9 funktio  $k(x, z)$  toteuttaa*

$$|k(x, z)| \leq C(\gamma + 1)^n |z - x|^{-n},$$

missä  $\gamma$  on alueen  $U$  paksuusparametri.

*Todistus.* Valitaan pallo  $B(x_0, \rho)$  niin, että  $U$  on radiaalisesti konvekssi pallon  $B$  suhteen, ja  $\rho > (1/2)\rho_{max}$ . Proposition 5.9 nojalla

$$\begin{aligned} |k(x, z)| &\leq C \left(1 + \frac{1}{\rho}|x - x_0|\right)^n |z - x|^{-n} \\ &\leq C \left(1 + \frac{2d}{\rho_{max}}\right)^n |z - x|^{-n} \\ &\leq C2^n \left(1 + \frac{d}{\rho_{max}}\right)^n |z - x|^{-n} \\ &:= C(1 + \gamma)^n |z - x|^{-n}. \end{aligned}$$

□

### 5.3 Virhetermin suuruus

Kappaleessa esitetään arvio virhetermille  $R^m$ . Seurauksena johdetaan myös *Sobolevin lemma*, mikä kertoo tärkeistä funktioiden jatkuvuusominaisuuksista  $W^{p,k}(U)$ -avaruuksissa.

**Lemma 5.12.** *Jos  $f \in L^p(U)$ , kun  $1 < p < \infty$  ja  $m > n/p$ , niin*

$$\int_U |x - z|^{-n+m} |f(z)| dz \leq C_p d^{m-n/p} \|f\|_{L^p(U)}.$$

*Jos  $p = 1$ , niin yhtäpitävyys pitää, kun  $m \geq n$ .*

*Todistus.* Oletetaan  $1 < p < \infty$  ja  $m > n/p$ . Olkoon  $q > 0$ , jolle  $1/p + 1/q = 1$ . Hölderin epäyhtälön nojalla

$$\begin{aligned} &\int_U |x - z|^{-n+m} |f(z)| dz \\ &\leq \left( \int_U |x - z|^{(-n+m)q} dz \right)^{1/q} \|f\|_{L^p(U)} \quad (\text{Hölderin epäyhtälö}) \\ (5.12) \quad &\leq C \left( \int_0^d r^{(-n+m)q+n-1} dr \right)^{1/q} \|f\|_{L^p(U)} \\ &= C \left( \frac{d^{(-n+m)q+n}}{(-n+m)q+n-1} \right)^{1/q} \|f\|_{L^p(U)} \\ &= C_{m,n,p} d^{m+n((1/q)-1)} \|f\|_{L^p(U)} = C_{m,n,p} d^{m-n/p} \|f\|_{L^p(U)}. \end{aligned}$$

Arviossa (5.12) peitetään alue  $U$   $d$  säteisellä kuulalla ja integroidaan kuulan yli polaarikoordinaateissa.

Oletetaan, että  $p = 1$  ja  $m \geq n$ , tällöin

$$\begin{aligned} \int_U |x - z|^{-n+m} |f(z)| dz &\leq \|(x - z)^{-n+m}\|_{L^\infty(U)} \|f\|_{L^1(U)} \\ &\leq d^{-n+m} \|f\|_{L^1(U)}. \end{aligned}$$

□

**Propositio 5.13.** *Kun  $u \in W^{p,m}(U)$ ,*

$$\|R^m u\|_{L^\infty(U)} \leq C_{m,n,\gamma} d^{m-n/p} |u|_{W^{p,m}(U)},$$

*jos  $1 < p < \infty$  ja  $m > n/p$ , tai  $p = 1$  ja  $m \geq n$ . Normi  $|\cdot|_{W^{p,m}(U)}$  on määritelty  $|u|_{W^{p,m}(U)} = \left( \sum_{|\alpha|=m} \|D^\alpha u\|_{L^p(U)}^p \right)^{1/p}$ .*

*Todistus.* Oletetaan, että  $u \in C^m(U) \cap W^{p,m}(U)$ , jolloin proposition 5.9 nojalla

$$\begin{aligned} |R^m u(x)| &= \left| m \sum_{|\alpha|=m} \int_{C_x} (1/\alpha!) (x-z)^\alpha k(x,z) D^\alpha u(z) dz \right| \\ &\leq C'_{m,\gamma,n} \sum_{|\alpha|=m} \int_U |(x-z)^\alpha| |x-z|^{-n} |D^\alpha u(z)| dz \quad (\text{korollaari 5.11}) \\ (5.13) \quad &\leq C'_{m,\gamma,n} \sum_{|\alpha|=m} \int_U |(x-z)^{m-n}| |D^\alpha u(z)| dz \\ &\leq C''_{m,n,\gamma} d^{m-n/p} \sum_{|\alpha|=m} \|D^\alpha u\|_{L^p(U)} \quad (\text{lemma 5.12}) \\ (5.14) \quad &\leq C_{m,n,\gamma} d^{m-n/p} |u|_{W^{p,m}(U)}. \end{aligned}$$

Arvio (5.13) pätee, sillä

$$\begin{aligned} |(x-z)^\alpha| &= \prod_{i=1}^n |x_i - z_i|^{\alpha_i} \\ &\leq \prod_{i=1}^n |x - z|^{\alpha_i} = |x - z|^{|\alpha|} \quad \text{kaikilla } z \in U. \end{aligned}$$

Tulos viimeistellään tiheysargumentilla. Olkoon  $u \in W^{p,m}(U)$ . Kuvaus  $R^m$  on jatkuva kuvaus avaruudelta  $W^{p,k}(U) \subset L^p(U) \subset L^1(U)$  avaruuteen  $L^\infty(U) = W^{\infty,0}(U)$ , sillä korollaarin 5.6 ja kolmioepäyhtälön nojalla

$$\|R^m u\|_{L^\infty(U)} \leq \|u\|_{L^\infty(U)} + \|Q^m u\|_{L^\infty(U)} \leq C \|u\|_{L^1(U)}.$$

Voidaan valita jono  $(u_i) \subset C^\infty(U) \cap W^{p,m}(U)$ , jolle  $\lim_{i \rightarrow \infty} u_i = u$ . Kaavan (5.14) nojalla saadaan

$$\lim_{i \rightarrow \infty} \|R^m u_i\|_{L^\infty(U)} = \|R^m u\|_{L^\infty(U)} \leq C_{m,n,\gamma} d^{m-n/p} |u|_{W^{p,m}(U)},$$

missä ensimmäisen yhtäpitävyys pätee kuvauksen  $R^m$  jatkuvuuden nojalla.  $\square$

Seuraavaksi esiteltävä Sobolevin lemma takaa hyödyllisiä jatkuvuusominaisuuksia  $W^{p,m}(U)$  avaruuksissa; tarpeeksi monen heikon derivaatan olemassaolo takaa funktion  $u \in W^{p,m}(U)$  jatkuvuuden.

**Lemma 5.14** (Sobolevin epäyhtälö). *Oletetaan, että  $U$  on läpimitaltaan  $d$ . Lisäksi  $U$  on radiaalisesti konvekssi pallon  $B$  suhteen. Jos  $u \in W^{p,m}(U)$ , jossa joko  $1 < p < \infty$  ja  $m > n/p$ , tai  $p = 1$  ja  $m \geq n$ , niin  $u$  on jatkuva alueessa  $U$  ja*

$$\|u\|_{L^\infty(U)} \leq C_{m,n,\gamma,d} \|u\|_{W^{p,m}(U)}.$$

*Todistus.* Korollarin 5.6 ja proposition 5.13 nojalla

$$\begin{aligned} \|u\|_{L^\infty(U)} &\leq \|u - Q^m u\|_{L^\infty(U)} + \|Q^m u\|_{L^\infty(U)} \\ &\leq C_{m,n,\gamma} |u|_{W^{p,m}(U)} + C_{m,n,\rho} \|u\|_{L^1(U)} \\ &\leq C_{m,n,\gamma,d} \|u\|_{W^{p,m}(U)} \end{aligned}$$

Jatkuvuus seuraa ottamalla Cauhcy jono  $(u_i) \subset C^\infty(U) \cap W^{p,m}(U)$ , jolle  $u_j \rightarrow u$ . Funktion  $u_n$  jatkuvuuden nojalla

$$\begin{aligned} |u(x) - u(y)| &\leq |u(x) - u_n(x)| + |u_n(y) - u(y)| + |u_n(x) - u_n(y)| \\ &\leq 2\|u - u_n\|_{L^\infty(U)} + \epsilon \leq C_{m,n,\gamma,d} \|u - u_n\|_{W^{p,m}(U)} + \epsilon \end{aligned}$$

kaikilla  $|x - y| < \delta$ . Ottamalla  $n \rightarrow \infty$  saadaan

$$|u(x) - u(y)| \leq \epsilon,$$

joten  $u$  on jatkuva.  $\square$

Todistetaan seuraavaksi *Bramblen-Hilbertin* lemma, mikä takaa funktioita  $u \in W^{p,m}(U)$  approksimoivan polynomin  $Q^m u$  suppenemisen kohti funktiota  $u$ , kun alueen  $U$  läpimitta lähestyy nollaa.

**Lemma 5.15.** *Olkoon  $f \in L^p(U)$ ,  $p \geq 1$  ja  $m \geq 1$ . Määritellään*

$$g(x) = \int_U |x - z|^{m-n} |f(z)| dz.$$

*Tällöin pätee*

$$\|g\|_{L^p(U)} \leq C_{m,n} d^m \|f\|_{L^p(U)}.$$

*Todistus.* Oletetaan  $1 < p < \infty$ , ja  $1/p + 1/q = 1$ .

$$\begin{aligned}
\|g\|_{L^p(U)}^p &= \int_U |g(x)|^p dx = \int_U \left( \int_U |x - z|^{m-n} |f(z)| dz \right)^p dx \\
&\leq \int_U \left[ \left( \int_U |f(z)|^p dz \right)^{1/p} \left( \int_U |x - z|^{(m-n)q} dz \right)^{1/q} \right]^p dx \text{ (Hölderin epäytälö)} \\
&\leq C_{m,n,p} \int_U \int_U |f(z)|^p |x - z|^{m-n} dz dx \text{ (lemma 5.12 ja oletus } mp > 1) \\
&\leq C_{m,n} d^{mp/q} \int_U \left( \int_U |x - z|^{m-n} dx \right) |f(z)|^p dz \text{ (Fubinin lause)} \\
&\leq C_{m,n} d^{mp} \|f\|_{L^p(U)}^p \text{ (lemma 5.12 ja oletus } p/q \leq p).
\end{aligned}$$

Kun  $p = 1$ , niin vastaavasti

$$\begin{aligned}
\|g\|_{L^1(U)} &= \int_U |g(x)| dx = \int_U \left( \int_U |x - z|^{m-n} |f(z)| dz \right) dx \\
&\leq C_{m,n} d^m \|f\|_{L^1(U)}
\end{aligned}$$

Jos  $p = \infty$

$$\begin{aligned}
|g(x)| &= \int_U |x - z|^{m-n} |f(z)| dz \leq \|f\|_{L^\infty(U)} \int_U |x - z|^{m-n} dz \\
&\leq C_{m,n} d^m \|f\|_{L^\infty(U)}.
\end{aligned}$$

□

**Lemma 5.16** (Bramblen-Hilbertin lemma). *Olkoon  $U$  radiaalisesti konvekssi pallon  $B$  suhteen, jonka säteelle pätee  $\rho > (1/2)\rho_{max}$ . Olkoon  $Q^m u$  Sobolevin polynomi liittyen palloon  $B$ , missä  $u \in W^{p,m}(U)$  ja  $p \geq 1$ . Tällöin pätee*

$$|u - Q^m u|_{W^{p,k}(U)} \leq C_{m,n,\gamma} d^{m-k} |u|_{W^{p,m}(U)} \quad k = 0, 1, \dots, m,$$

missä  $d$  on alueen  $U$  läpimitta.

*Todistus.* Olkoon  $u \in C^m(U) \cap W^{p,m}(U)$ . Erityistapauksessa  $k = m$  saadaan proposition 5.7 nojalla

$$\begin{aligned}
|u - Q^m u|_{W^{p,k}(U)} &= \left( \sum_{\alpha=m} \|D^\alpha u - D^\alpha Q^m u\|_{W^{p,k}(U)}^p \right)^{1/p} \\
&= \left( \sum_{\alpha=m} \|D^\alpha u - Q^0 D^\alpha u\|_{W^{p,k}(U)}^p \right)^{1/p} = |u|_{W^{p,m}(U)},
\end{aligned}$$

kun tulkitaan  $Q^0 \equiv 0$ .

Kun  $k = 0$ , seuraa propositiosta 5.9 ja lemmasta 5.15

$$\begin{aligned} \|u - Q^m u\|_{W^{p,k}(U)} &= \|R^m u\|_{W^{p,k}(U)} \\ &\leq C_{m,n}(1 + 1/\rho|x - x_0|)^{nm} \sum_{|\alpha|=m} \left\| \int_U |x - z|^{m-n} |D^\alpha u(z)| dz \right\|_{L^p(U)} \\ &\leq C_{m,n,\gamma} d^m |u|_{W^{p,m}(U)}, \end{aligned}$$

missä viimeinen arvio on rajoitettu, sillä oletuksesta seuraa

$$2\gamma := \frac{2d}{\rho_{max}} = \frac{d}{(1/2)\rho_{max}} > \frac{d}{\rho}.$$

Kun  $0 < k < m$ , saadaan propositioiden 5.7 ja 5.13 nojalla

$$\begin{aligned} |u - Q^m u|_{W^{p,k}(U)} &= |R^m u|_{W^{p,k}(U)} \\ &= \left( \sum_{|\alpha|=m} \|R^{m-k} D^\alpha u\|_{L^p(U)}^p \right)^{1/p} \\ &\leq C_{m,n,\gamma} d^{m-k} \left( \sum_{|\alpha|=k} |D^\alpha u|_{W_p^{m-k}(U)}^p \right)^{1/p} \\ &\leq C_{m,n,\gamma} d^{m-k} |u|_{W^{p,m}(U)}. \end{aligned}$$

Yleinen tulos funktioille  $u \in W^{p,m}(U)$  seuraa tiheysargumentilla valitsemalla jono  $(u_k) \subset C^m(U) \cap W^{p,m}(U)$ , jolle  $u_k \rightarrow u \in W^{p,m}(U)$ .

□

## 5.4 Interpolaation virhe

Tähän mennessä olemme johtaneet paikallisen virhearvion polynomiselle approksimaatiolle  $R^m = u - Q^m$ . Käytännössä tämä tarkoittaa elementtikohtaista virhettä. Tässä kappaleessa johdamme globaalin virheen koko ositettavalle alueelle. Otamme käyttöön kappaleessa 3.1 esitetyn määritelmän elementeille.

**Lemma 5.17.** *Olkoon  $(K, \mathcal{P}, \mathcal{N})$  elementti, jolle  $\mathcal{P} \subset W_\infty^m(K)$  ja  $\mathcal{N} = (C^l(\bar{K}))'$ , eli  $N_i$  sisältää korkeintaan asteen  $l$  derivaattoja. Tällöin interpolaatio-operaattori on rajoitettu kuvaus avaruudelta  $C^l(\bar{K})$  avaruuteen  $W_k^m(K)$ , kun  $1 \leq p \leq \infty$ .*

*Todistus.* Olkoon  $\mathcal{N} = \{N_1, \dots, N_k\}$  ja  $\{\phi_1, \dots, \phi_k\} \subset \mathcal{P}$ . Interpolaatio on määritelty  $\mathcal{I}u = \sum_{i=1}^k N_i(u)\phi_i$ , ja oletuksen nojalla jokainen  $\phi_i \in W_\infty^m(K) \subset W^{p,m}(K)$ . Kolmioepäyhtälön ja duaalinormin määritelmän mukaan saadaan

$$\begin{aligned} \|\mathcal{I}u\|_{W^{p,m}(K)} &\leq \sum_{i=1}^k |N_i(u)| \|\phi_i\|_{W^{p,m}(K)} \\ &\leq \left( \sum_{i=1}^k \|N_i\|_{C^l(\bar{K})} \|\phi_i\|_{W^{p,m}(K)} \right) \|u\|_{C^l(\bar{K})} \\ &= C \|u\|_{C^l(\bar{K})}. \end{aligned}$$

□

**Määritelmä 5.18.** Olkoon  $(K, \mathcal{P}, \mathcal{N})$  lemmän 5.17 ehtojen mukainen elementti. Määritellään elementtikohtaisen kuvauksen  $\mathcal{I} : C^l(\bar{K}) \rightarrow W^{p,m}(K)$  operaattorinormi

$$\sigma(K) := \sup_{v \neq 0} \frac{\|\mathcal{I}v\|_{W^{p,m}(K)}}{\|v\|_{C^l(\bar{K})}}.$$

**Määritelmä 5.19.** Olkoon alue  $K$  rajoitettu. Mekitään  $d := \text{diam } K$ . Määritellään

$$\hat{K} := \{(1/d)x : x \in K\},$$

ja  $\hat{u}(y) := u(dy)$ .

Seuraavissa todistuksissa käytetään toistuvasti niin kutsuttua *homogeenisyyssargumenttia*. Argumentti takaa, että tulokset alueelle  $\hat{K}$ , jolle  $\text{diam}(\hat{K}) = 1$ , voidaan samantien yleistää mielivaltaiselle rajoitetulle alueelle  $K$ . Argumentti perustuu seuraavaan huomioon:

**Lemma 5.20.** *Olkoon  $u \in W^{p,m}(K)$ . Tällöin*

$$|\hat{u}|_{W^{p,k}(\hat{K})} = (\text{diam } K)^{k-n/p} |u|_{W^{p,k}(K)},$$

missä  $0 \leq k \leq m$ .

*Todistus.* Koska  $\hat{u} \in W^{p,m}(\hat{K})$ , yhdistetyn derivaatan kaavalla ja sijoitusmenetel-

mällä 4.3 saadaan

$$\begin{aligned}
|\hat{u}|_{W^{p,k}(\hat{K})} &= \left( \sum_{|\alpha|=m} \int_{\hat{U}} |D^\alpha(\hat{u}(y))|^p dy \right)^{1/p} = \left( \sum_{|\alpha|=k} \int_{\hat{U}} |D^\alpha(u(dy))|^p dy \right)^{1/p} \\
&= \left( \sum_{|\alpha|=k} \int_{\hat{U}} |(D^\alpha u)(y)|^p dy \right)^{1/p} \\
&= d^{k-n/p} \left( \sum_{|\alpha|=k} \int_U |(D^\alpha u)(y)|^p dy \right)^{1/p} \\
&= d^{k-n/p} |u|_{W^{p,k}(\hat{K})}.
\end{aligned}$$

□

**Lause 5.21.** *Olkoon  $(K, \mathcal{P}, \mathcal{N})$  elementti, jolle pätee*

- (i)  *$K$  on radiaalisesti konvekksi jonkin pallon  $B$  suhteen,*
- (ii)  *$\mathcal{P}_{m-1} \subset \mathcal{P} \subset W^{\infty,m}(K)$  ja*
- (iii)  *$\mathcal{N} \subset (C^l(\bar{K}))'$ .*

*Oletetaan  $1 \leq p \leq \infty$ , ja joko  $m - l - n/p > 0$  kun  $p > 1$ , tai  $m - l - n \leq 0$  kun  $p = 1$ . Tällöin kun  $0 \leq i \leq m$  ja  $v \in W^{p,m}(K)$  saadaan*

$$|v - \mathcal{I}v|_{W^{p,i}(K)} \leq C_{m,n,\gamma,\sigma(\hat{K})} (\text{diam } K)^{m-i} |v|_{W^{p,m}(K)},$$

*missä  $\hat{K} = \{(1/\text{diam}(K))x : x \in K\}$  ja  $\gamma$  on alueen  $K$  paksuusparametri.*

*Todistus.* Oletetaan, että pallon  $B$  säteelle pätee  $\rho > (1/2)\rho_{max}$ . Olkoon  $Q^m v$  Sovolevin polynomi liittyen palloon  $B$ . Lemman 3.7 nojalla  $\mathcal{I}f = f$ , jos  $f \in \mathcal{P}$ , joten  $\mathcal{I}Q^m v = Q^m v$ , sillä  $Q^m v \in \mathcal{P}_{m-1} \subset \mathcal{P}$ . Kolmiolle  $\hat{K}$  saadaan

$$\begin{aligned}
|\hat{v} - \mathcal{I}\hat{v}|_{W^{p,i}(\hat{U})} &\leq |\hat{v} - Q^m \hat{v}|_{W^{p,i}(\hat{K})} + |Q^m \hat{v} - \mathcal{I}\hat{v}|_{W^{p,i}(\hat{K})} \\
&\leq |\hat{v} - Q^m \hat{v}|_{W^{p,i}(\hat{K})} + \sigma(\hat{K}) \|Q^m \hat{v} - \hat{v}\|_{C^i(\hat{K})} \\
&\leq (1 + \sigma(\hat{K}) C_{m,n,\gamma,1}) \|\hat{v} - Q^m \hat{v}\|_{W^{p,m}(\hat{K})} \quad (\text{Sobolevin epäytälö 5.14}) \\
&\leq C_{m,n,\gamma,\sigma(\hat{K})} |\hat{v}|_{W^{p,m}(\hat{K})} \quad (\text{Bramlben-Hilbertin lemma 5.16}).
\end{aligned}$$

Yleinen tulos mielivaltaiselle alueelle  $K$  seuraa lemmän 5.20 huomiosta. □



**Korollaari 5.22.** *Oletetaan, että  $i \leq l$ , ja olkoon elementti  $(K, \mathcal{P}, \mathcal{N})$  muuten lauseen 5.21 mukainen. Tällöin*

$$|v - \mathcal{I}v|_{W^{\infty,i}(K)} \leq C_{m,n,\gamma,\sigma(\hat{K})} (\text{diam } K)^{m-i-n/p} |v|_{W^{p,m}(K)}.$$

*Todistus.* Olkoon  $\text{diam}(K) = 1$ , tällöin

$$\begin{aligned} \|v - \mathcal{I}v\|_{W^{\infty,i}(K)} &\leq C_{m,n,\gamma} \|v - \mathcal{I}v\|_{W^{p,m}(K)} \quad (\text{Sobolevin lause 5.14}) \\ &\leq C_{m,n,\gamma,\sigma(\hat{K})} |v|_{W^{p,m}(K)} \quad (\text{lause 5.21}). \end{aligned}$$

Yleinen tulos seuraa lemmasta 5.20, nimittäin yleiselle  $K$  pätee:

$$\begin{aligned} |v - \mathcal{I}v|_{W^{\infty,i}(K)} &= d^{-i+n/\infty} |v - \mathcal{I}v|_{W^{\infty,i}(\hat{K})} \\ &\leq d^{-i} C_{m,n,\gamma,\sigma(\hat{K})} |v|_{W^{p,m}(\hat{K})} = d^{-i+m-n/p} C_{m,n,\gamma,\sigma(\hat{K})} |v|_{W^{p,m}(K)}. \end{aligned}$$

□

Johdetaan virhearviot globaalille interpolantille. Olkoon  $(K, \mathcal{P}, \mathcal{N})$  referenssielementti, mikä on affiniesti ekvivalentti elementin  $(\hat{K}, \hat{\mathcal{P}}, \hat{\mathcal{N}})$  kanssa kuvauksen  $F = x \mapsto ax + b$  kautta. Määritelmän 4.2 mukaisesti tällöin  $\hat{N} = F_*N$  ja  $\hat{\phi} = F^*\phi$ . Saadaan

$$\begin{aligned} \hat{\mathcal{I}}\hat{v}(\hat{x}) &= \sum_{\hat{N} \in \hat{\mathcal{N}}} \hat{N}(\hat{v}) \cdot \hat{\phi}_{\hat{N}}(\hat{x}) \\ &= \sum_{N \in \mathcal{N}} (F_*N)\hat{v} \cdot (F^{-1})^*\phi_N(\hat{x}). \end{aligned}$$

Seuraava tavoite on muodostaa arvio operaattorinormille  $\sigma(\hat{K})$ . Arvion tulee riippua vertailuelementistä riippuvasta vakiosta  $C_{\text{ref}} = C_{K,\mathcal{P},\mathcal{N}}$  ja jostakin jatkuvasta funktiosta  $\chi : \mathcal{S}(\mathbb{R}) \rightarrow \mathbb{R}$ , missä  $\mathcal{S}(\mathbb{R})$  viittaa kääntyvien symmetristen matriisien avaruuteen. Määritellään avaruuteen  $\mathcal{S}(\mathbb{R})$  (esimerkiksi) normi

$$\|\cdot\|_{\mathcal{S}(\mathbb{R})} = a \mapsto \max_{1 \leq i,j \leq n} |a_{ij}|.$$

**Propositio 5.23.** *Olkoon vertailuelementti  $(K, \mathcal{P}, \mathcal{N})$  affiniesti ekvivalentti elementin  $(\hat{K}, \hat{\mathcal{P}}, \hat{\mathcal{N}})$  kanssa kuvauksen  $F = x \mapsto ax + b$  kautta. Tällöin on olemassa jatkuva funktio  $\chi : \mathcal{S}(\mathbb{R}) \rightarrow \mathbb{R}$  ja vertailuelementistä riippuva vakio  $C_{\text{ref}}$ , joille*

$$\sigma(\hat{K}) \leq C_{\text{ref}} \chi(a).$$

*Todistus.* Koska  $(F_*N)\hat{v} = N(F^*\hat{v})$ , missä  $(F^*\hat{v})x = \hat{v}(Fx)$ , saadaan

$$\begin{aligned} |(F_*N)\hat{v}| &= |N(F^*\hat{v})| \\ &\leq C_N \|F^*\hat{v}\|_{C^l(\bar{K})} \quad (\text{kuvauksen } N \text{ on rajoitettu}) \\ &\leq C_{N,n,l} \left(1 + \max_{1 \leq i,j \leq n} |a_{ij}|\right)^l \|\hat{v}\|_{C^l(\bar{K})}. \end{aligned}$$

Vastaavasti voidaan ilmaista

$$\begin{aligned} \|(F^{-1})^* \phi_N\|_{W^{p,m}(\hat{K})} &\leq C'_{N,n,m} \left(1 + \max_{1 \leq i,j \leq n} |(a^{-1})_{ij}|\right)^m \times \\ &\|\phi_N\|_{W^{p,m}(K)} |\det a|^{1/p}. \end{aligned}$$

Koska  $\|\phi_N\|_{W^{p,m}(\hat{U})}$  on lemmän 5.17 nojalla rajoitettu vertailulelementissä, saadaan

$$\begin{aligned} \|\hat{\mathcal{I}}\hat{v}\|_{W^{p,m}(\hat{K})} &\leq C_{\text{ref}} \left(1 + \max_{1 \leq i,j,n} |a_{ij}|\right)^l \times \\ &\left(1 + \max_{1 \leq i,j \leq n} |(a^{-1})_{ij}|\right)^m |\det a|^{1/p} \|\hat{v}\|_{C^l(\bar{K})}, \end{aligned}$$

missä

$$C_{\text{ref}} = \dim(\mathcal{N}) \cdot \max_{N \in \mathcal{N}} \{C_{N,n,l}\} \cdot \max_{N \in \mathcal{N}} \{C'_{N,n,m}\} \cdot \max_{N \in \mathcal{N}} \{\|\phi_N\|_{W^{p,m}(U)}\}.$$

Näin ollen operaattorinormille saadaan arvio

$$\begin{aligned} \sigma(\hat{K}) &\leq C_{\text{ref}} \left(1 + \max_{1 \leq i,j,n} |a_{ij}|\right)^l \times \\ &\left(1 + \max_{1 \leq i,j \leq n} |(a^{-1})_{ij}|\right)^m |\det a|^{1/p}. \end{aligned}$$

Determinantin jatkuvuudesta seuraa kuvauksen

$$a \mapsto \left(1 + \max_{1 \leq i,j,n} |a_{ij}|\right)^l \left(1 + \max_{1 \leq i,j \leq n} |(a^{-1})_{ij}|\right)^m |\det a|^{1/p}$$

jatkuvuus, joten ollaan löydetty eräs jatkuva kuvaus  $\chi$ . □

**Määritelmä 5.24** (Rajoitettu jakoperhe). Olkoon  $\{\mathcal{T}^h\}$ ,  $0 < h \leq 1$  jakojen perhe jolle

$$\max\{\text{diam } T : T \in \mathcal{T}^h\} \leq h \text{ diam } U$$

alueessa  $U \subset \mathbb{R}^n$ . Kutsutaan tällöin perhettä  $\{\mathcal{T}^h\}$  rajoitetuksi jakoperheeksi.

**Määritelmä 5.25** (Säännöllinen jakoperhe). Rajoitettu jakoerhe  $\{\mathcal{T}^h\}$  on *säännöllinen* (quasi-uniform), jos on olemassa  $\rho > 0$ , jolle

$$\min\{\text{diam } B_T : T \in \mathcal{T}^h\} \geq \rho h \text{ diam } U$$

kaikilla  $h \in (0, 1]$ . Merkintä  $B_T$  viittaa suurimpaan palloon, mikä sisältyy elementtiin  $T$ . Toisaalta säännölliselle perheelle pätee

$$\rho \leq \frac{\min\{\text{diam } B_T : T \in \mathcal{T}^h\}}{\max\{\text{diam } T : T \in \mathcal{T}^h\}}$$

jollekin  $\rho > 0$ . Edellinen antaa ehkä intuitiivisemmän määritelmän säännöllisyydelle; suurimman ja pienimmän elementin koon suhde ei säännöllisessä jakoperheessä saa saada mielivaltaisen pieniä arvoja.

**Määritelmä 5.26** (Ei-degeneroituva jakoperhe). Perhe on *ei-degeneroituva* jos on olemassa  $\rho > 0$ , jolle kaikilla  $T \in \mathcal{T}^h$  ja kaikille  $h \in (0, 1]$  pätee

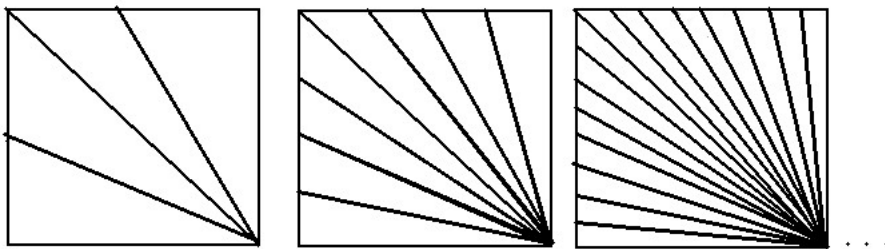
$$\text{diam } B_T \geq \rho \text{ diam } T.$$

Säännöllinen jakoperhe on ei-degeneroituva, sillä mielivaltaiselle  $T \in \{\mathcal{T}^h\}$  pätee

$$\rho h \text{ diam } T \leq \rho \text{ diam } T \leq \min\{\text{diam } B_T : T \in \mathcal{T}^h\} \leq \text{diam } B_T.$$

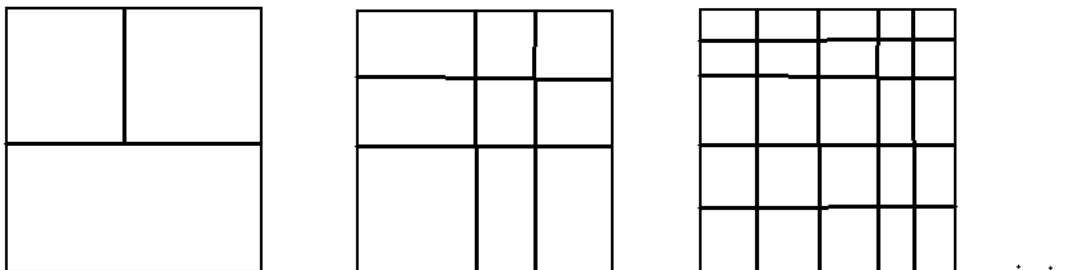
Ei-degeneroituva perhe ei kuitenkaan välttämättä ole säännöllinen; tämä on suoraviivaista osoittaa esimerkiksi konstruoimalla esimerkkitapaus välille  $(0, 1)$ . Mielivaltaisesta kolmioinnista saa säännöllisen kolmioinnin yhdistämällä kolmion sivut viivalla keskenään. Vastaavanlainen konstruktio voidaan tehdä kolmessa ulottuvuudessa ([10]).

Kuva 5.1: Degeneroituvan jakoperheen kolme ensimmäistä jäsentä



**Lause 5.27.** Olkoon  $\{\mathcal{T}^h\}$ ,  $0 < h \leq 1$ , alueen  $U \subset \mathbb{R}^n$  ei-degeneroituva alijakojen perhe. Olkoon  $(K, \mathcal{P}, \mathcal{K})$  vertailuelementti, joka toteuttaa lauseessa 5.22 määritellyt ehdot joillekin  $l$ ,  $m$  ja  $p$ . Olkoon  $(T, \mathcal{P}_T, \mathcal{N}_T)$  vertailuelementin kanssa affiniesti

Kuva 5.2: Ei-säännöllisen (mutta ei-degeneroituvan) jakoperheen kolme ensimmäistä jäsentä



ekvivalentti elementti, kun  $T \in \mathcal{T}^h, 0 < h \leq 1$ . Tällöin on olemassa vertailuelementistä, vakioista  $n, m, p$  ja vakiosta  $\rho$  riippuva vakio  $C$ , jolle kun  $0 \leq s \leq m$ , pätee

$$(5.15) \quad \left( \sum_{T \in \mathcal{T}^h} \|v - \mathcal{I}^h v\|_{W^{p,s}(T)}^p \right)^{1/p} \leq Ch^{m-s} |v|_{W^{p,s}(U)}$$

kaikilla  $v \in W^{p,m}(U)$ . Tapauksessa  $p = \infty$  tulkitaan

$$(5.16) \quad \max_{T \in \mathcal{T}^h} \|v - \mathcal{I}^h v\|_{W^{\infty,s}(T)} \leq Ch^{m-s-n/p} |v|_{W^{p,m}(U)} \text{ kaikilla } v \in W^{p,m}(U).$$

*Todistus.* Seuravaa arviota tarvitaan todistuksessa

$$(5.17) \quad \sup\{\sigma(\hat{T}) : T \in \mathcal{T}^h, 0 < h \leq 1\} = C(\rho, m, n, p, K) < \infty.$$

Elementti  $(\hat{T}, \hat{\mathcal{P}}_T, \hat{\mathcal{N}}_T)$  on affiinisti ekvivalentti elementin  $(K, \mathcal{P}, \mathcal{N})$  kanssa jonkin kuvauksen  $Ax = ax + b$  kautta, joten proposition 5.23 nojalla  $\sigma(\hat{T}) \leq C_{\text{ref}} \chi(a)$  jollain jatkuvalla kuvauksella  $\chi : \mathcal{S}(\mathbb{R}) \rightarrow \mathbb{R}$ . Tulemme osoittamaan, että kuvaus  $a \mapsto C_{\text{ref}} \chi(a)$  on rajoitettu.

Esitellään tässä käsite *simpleksi*, sillä tulemme tarvitsemaan sitä todistuksen agumentissa. Simpleksi on yksinkertaisin *monikulmio*, mitä on määriteltävissä avaruuksiin  $\mathbb{R}^n$ . Esimerkiksi avaruudessa  $\mathbb{R}^2$  tämä tarkoittaa kolmiota.

**Määritelmä 5.28** (Simpleksi). Simpleksi on joukko

$$S = \left\{ \gamma_1 e^1 + \gamma_2 e^2 + \cdots + \gamma_n e^n : \sum_{i=1}^n \gamma_i \leq t_0, \gamma_i \geq 0 \right\},$$

missä  $(e^1, \dots, e^n)$  on ortonormaali kanta avaruudelle  $\mathbb{R}^n$ . Edellä simpleksin kärkipisteiden  $e^n = [0, \dots, 1, \dots, 0]$  indeksi  $n$  saa arvon 1.

Osoitetaan, että matriisi  $a$  kuuluu joukon  $\mathcal{S}(\mathbb{R})$  *kompaktiin* osajoukkoon. Tällöin kaava (5.17) seuraa kuvauksen  $\chi$  jatkuvuudesta. Määritelmän 5.26 nojalla on olemassa  $B \subset \hat{T}$ , jolle  $\text{diam } B \geq \rho > 0$ . Seuraa

$$\begin{aligned} \text{meas } B &\leq \text{meas } \hat{T} := \int_{\hat{T}} d\hat{x} \\ &= |\det a| \int_K dx := |\det a| (\text{meas } K). \end{aligned}$$

Toisaalta pallon tilavuuden kaavan nojalla  $\text{meas } B \geq C_n \rho^n$ , joten  $0 < C_n \rho^n \leq |\det a| (\text{meas } K)$ . Toisin sanoen  $|\det a| \geq \epsilon > 0$ , missä  $\epsilon$  riippuu valituista  $\rho$ ,  $K$  ja  $n$ . Näin ollen  $a$  kuuluu suljettuun joukkoon  $\{b : |\det b| \geq \epsilon > 0\}$ . simpleksi  $S$ , mikä sisältyy siirtoa vaille elementtiin  $K$ . Tällöin  $A(te^i) = ta e^i + b \in \hat{T}$  (siirtoa vaille), kun  $0 \leq t \leq t_0$ . Näin ollen  $|ae^i| = |(a_{1i}, \dots, a_{ni})| \leq \text{diam } \hat{T}/t_0 = 1/t_0$  kaikille  $1 \leq i \leq n$ , joten  $|a_{ij}| \leq 1/t_0, i, j = 1, \dots, n$ . Nyt joukko  $a \in \{b : |\det b| \geq \epsilon > 0, |b_{ij}| \leq 1/t_0\}$  on suljettuna ja rajoitettuna joukkona kompakti. Matriisi  $a$  kuuluu siis kompaktiin joukkoon, kun  $T \in \mathcal{T}^h$  ja  $h \in (0, 1]$ , joten kuvauksen  $\rho$  jatkuvuuden nojalla kaava (5.17) pätee.

Kaava (5.16) pätee, sillä

$$\begin{aligned} &\sum_{T \in \mathcal{T}^h} \|v - \mathcal{I}_T^h v\|_{W^{p,s}(T)}^p \\ &\leq \sum_{T \in \mathcal{T}^h} C_{m,n,\gamma,\sigma(\hat{T})}^p \sum_{i=0}^s (\text{diam } T)^{p(m-i)} |v|_{W^{p,m}(T)}^p \quad (\text{lauseen 5.21 arvio}) \\ &\leq \sum_{T \in \mathcal{T}^h} C_{m,n,\gamma,\sigma(\hat{T})}^p \sum_{i=0}^s (h \text{ diam } U)^{p(m-i)} |v|_{W^{p,m}(T)}^p \quad (\text{määritelmä 5.24}) \\ &\leq \sum_{T \in \mathcal{T}^h} C_{m,n,\rho,\sigma(\hat{T})}^p \sum_{i=0}^s (h \text{ diam } U)^{p(m-i)} |v|_{W^{p,m}(T)}^p \quad (\text{määritelmä 5.26}) \\ &\leq C_{m,n,K,\rho} h^{p(m-s)} \sum_{T \in \mathcal{T}} |v|_{W^{p,m}(U)}^p \quad (\text{kaava (5.17)}) \\ &= C_{m,n,K,\rho} h^{p(m-s)} |v|_{W^{p,m}(U)}^p. \end{aligned}$$

Vastaavasti kaava (5.16) pätee, sillä lauseen 5.22 nojalla

$$\begin{aligned}
\max_{T \in \mathcal{T}^h} \|v - \mathcal{I}^h v\|_{W^{\infty,s}(T)} & \\
&\leq C_{m,n,\gamma,\sigma(T)} (\text{diam } T)^{m-i-n/p} |v|_{W^{p,m}(U)} \\
&\leq C_{m,n,\rho,\sigma(T)} (h \text{ diam } U)^{m-i-n/p} |v|_{W^{p,m}(U)} \\
&\leq C_{m,n,K,\rho} h^{m-i-n/p} |u|_{W^{p,m}(U)}.
\end{aligned}$$

□

Luvun päätelmien nojalla FEM ratkaisu suppenee kohti täsmällistä ratkaisua sekä pisteittäisessä mielessä, että  $\|\cdot\|_{W^{p,s}}$  normin suhteen, sillä Cean lemma 5.1 liittyy interpolointien suppenemistulokset FEM ratkaisun suppenemisominaisuuksiin.

# Luku 6

## Posteriori suppenemisominaisuuksia

Luvussa vertaillaan luvussa 3.3.1 johdettujen ensimmäisen asteen kantafunktioiden ja kolmannen asteen kantafunktioiden suppenemisominaisuuksia tehtävässä

$$(6.1) \quad \begin{aligned} &\text{etsi } u \in C^2(0, 2\pi), \text{ jolla} \\ &u''(x) = \sin(5x) \text{ alueessa } x \in [0, 2\pi], \\ &\text{reunaehdoilla } u(0) = 0, \quad u(2\pi) = 0. \end{aligned}$$

Vastaava heikko muotoilu kuuluu

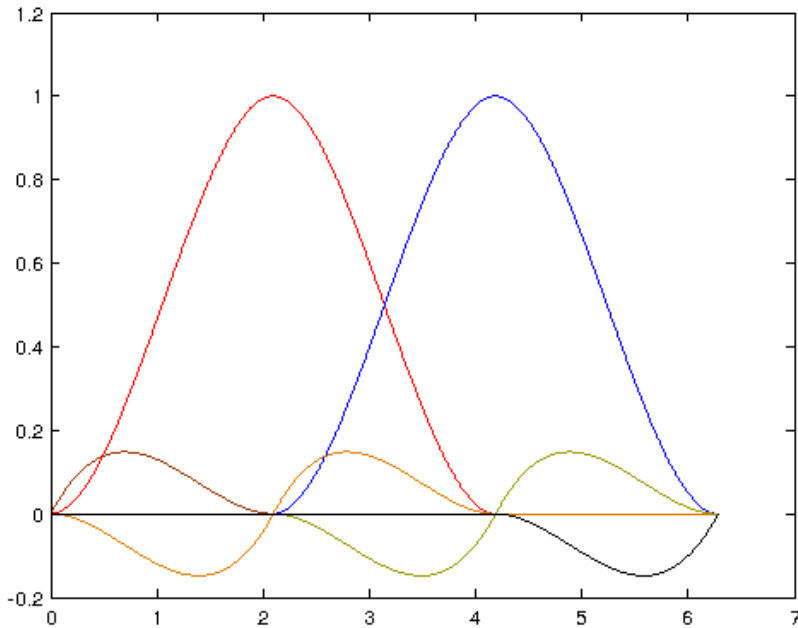
$$(6.2) \quad \begin{aligned} &\text{etsi } u_h = (u_1, \dots, u_n) \in \mathbb{R}^n, \text{ jolle} \\ &Au_h = b, \text{ missä} \\ &A_{ji} = \int_0^{2\pi} \phi_j'(x)\phi_i'(x)dx \text{ ja } b_i = - \int_0^{2\pi} \sin(x)\phi_i(x)dx, \end{aligned}$$

jossa  $\{\phi_i\}_{i=1}^n \subset C_h$  on valitut kantafunktiot.

Tehtävän 6.1 täsmällinen ratkaisu on funktio  $u = x \mapsto -\frac{1}{25}\sin(x)$ . Käytämme vertailtavina FEM avaruuksina *paloittain lineaaristen* ja *paloittain kolmannen asteen sileiden* funktioiden avaruuksia. Kantafunktiot määrittää kuvan 3.2 ensimmäinen ja viimeinen elementti.

Esimerkin 1.3 toteutus on suhteellisen suoraviivaista laajentaa kolmannen asteen kantafunktiolle. Kantafunktioiden indekseistä tulee vain sopivasti pitää kirjaa kuormavektorissa ja jäykkyyismatriisissa, sillä käyttämillämme kolmannen asteen kantafunktiolla jokaiselle solmulle liitetään kaksi kantafunktiota. Kuvassa 6.1 on merkitty solmuittaiset kantafunktiot eri väreillä, kun solmujen määrä on neljä. Huomaa, että kantafunktioiden tulee hävitä pisteissä  $x = 0$  ja  $x = 2\pi$ .

Eri jakojen tiheyksillä saadaan kuvissa 6.2 -6.7 piirrettyjä ratkaisuja. Kuvista voi päätellä, että elementtien pituuden tulee olla ainakin kaksi kertaa ratkaisun aallonpituutta pienempi, jotta elementtimenetelmä antaa hyödyllisiä ratkaisuja.



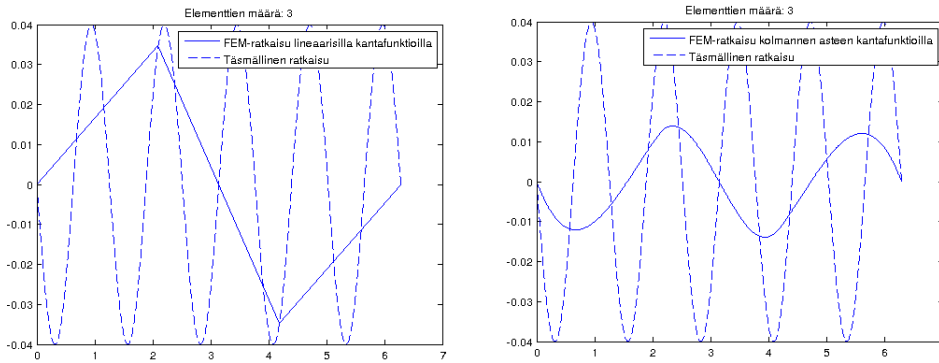
Kuva 6.1: Käytettävät kolmannen asteen kantafunktiot kolmen elementin tapauksessa

Kolmannen asteen kantafunktiot tavoittavat oskillaation jo suuremmissa elementeissä, mutta ratkaisussa on silti merkittävästi virhettä.

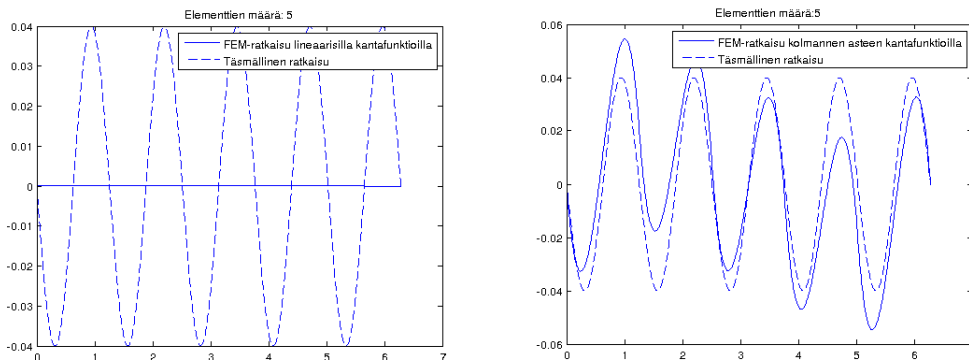
Kuvissa 6.8 ja 6.9 on vertailtu virhetermejä  $u - u_h$  energianormin  $\|\cdot\|_a$  ja  $L^2$ -normin  $\|\cdot\|_{L^2(U)}$  suhteen. Ratkaisu lineaarisilla kantafunktioilla suppenee tietyn pisteen jälkeen hyvin hitaasti kohti täsmällistä ratkaisua  $L^2$ -normissa – toisaalta energianormin suhteen suppeneminen on jopa nopeampaa kuin kolmannen asteen kantafunktioilla. Voi päätellä, että kolmannen asteen kantafunktiot antavat parempaa tietoa haettavan ratkaisun pisteittäisistä arvoista, kun lineaariset kantafunktiot tavoittavat ratkaisun derivaatan paremmin. Tämä johtunee kolmannen asteen kantafunktioilla tapahtuvasta oskillaatiosta, mikä aiheuttaa virhettä derivaatassa.

On huomattavaa, että paloittain lineaarisia kantafunktioita käyttäessä elementtimenetelmän tuottama ratkaisu näyttäisi yhtyvän täsmällisen ratkaisun lineaarisen interpolaation kanssa. Toisin sanoen lineaarinen interpolaatio olisi (tässä tapauksessa) tehtävän 6.1 täsmällisen ratkaisun ortogonaalinen projektio valittujen paloittain lineaaristen kantafunktioiden virittämään avaruuteen energianormin suhteen. Jätämme tämän kysymyksen vielä avoimeksi.

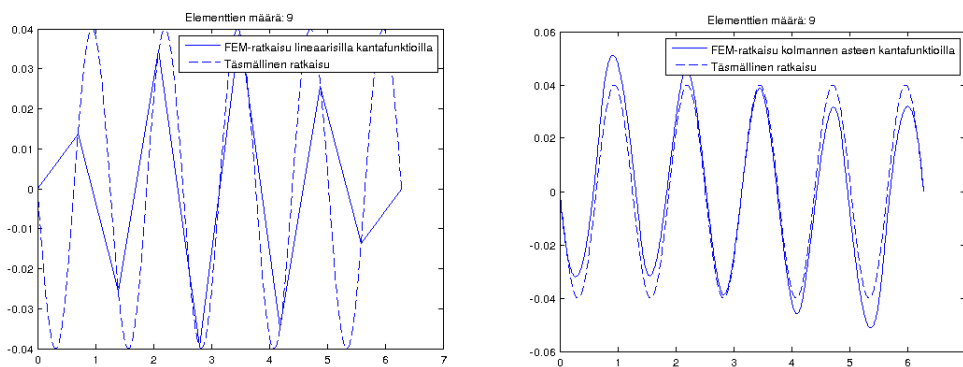




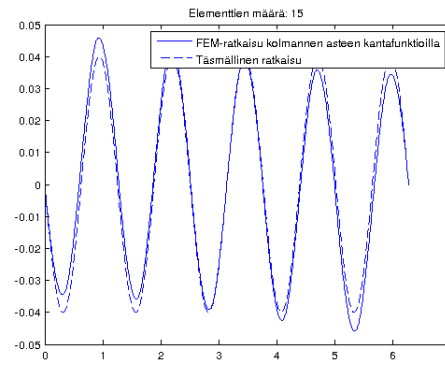
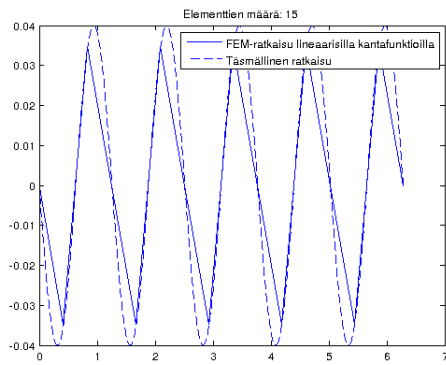
Kuva 6.2: Jakoväli on liian suuri



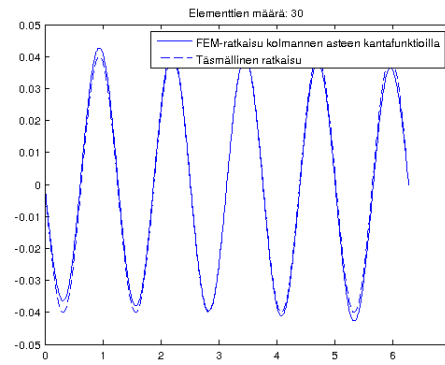
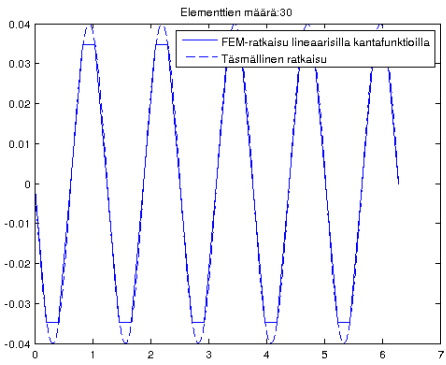
Kuva 6.3: Kolmannen asteen kantafunktiot tavoittavat jo oskillaation



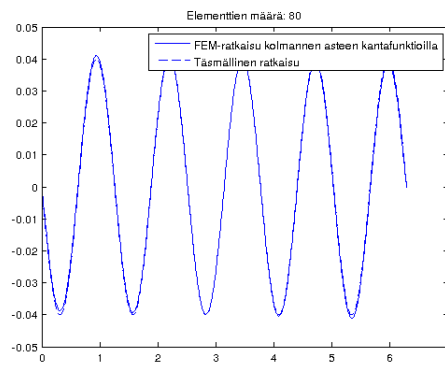
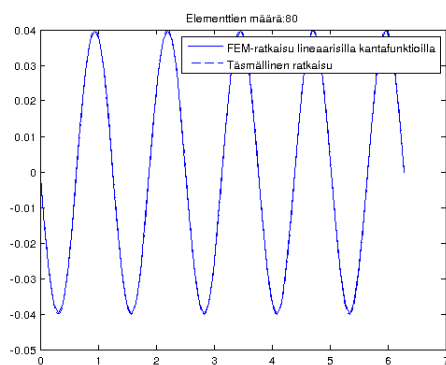
Kuva 6.4: Lineaariset kantafunktiot eivät näyttäisi vielä toimivan



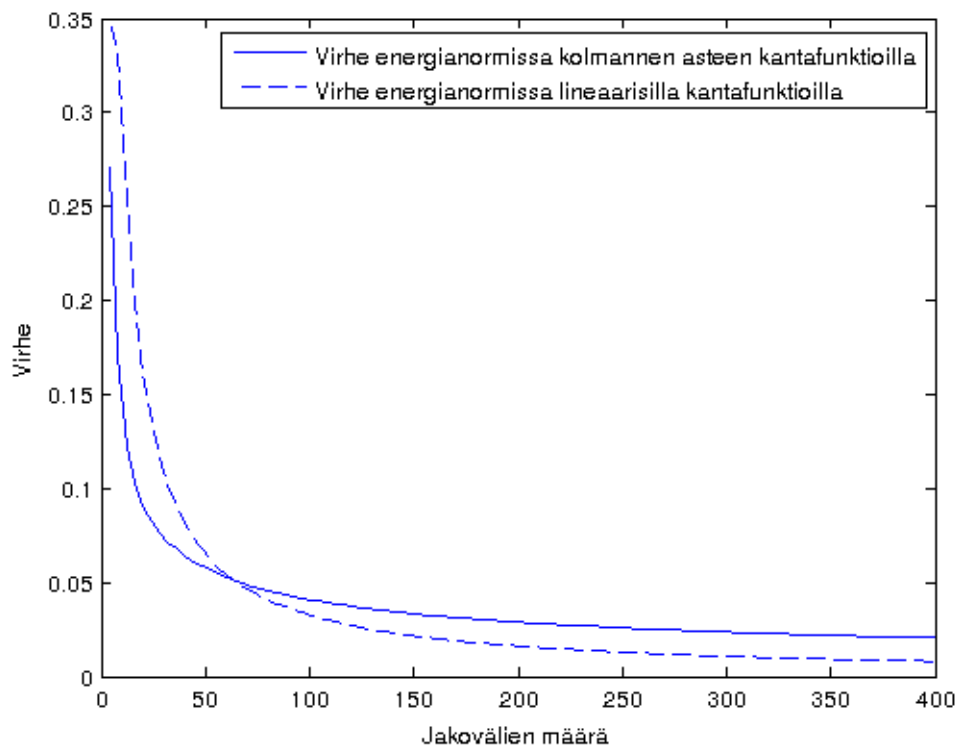
Kuva 6.5: Lineaariset kantafunktiot alkavat tavoittamaan ratkaisua



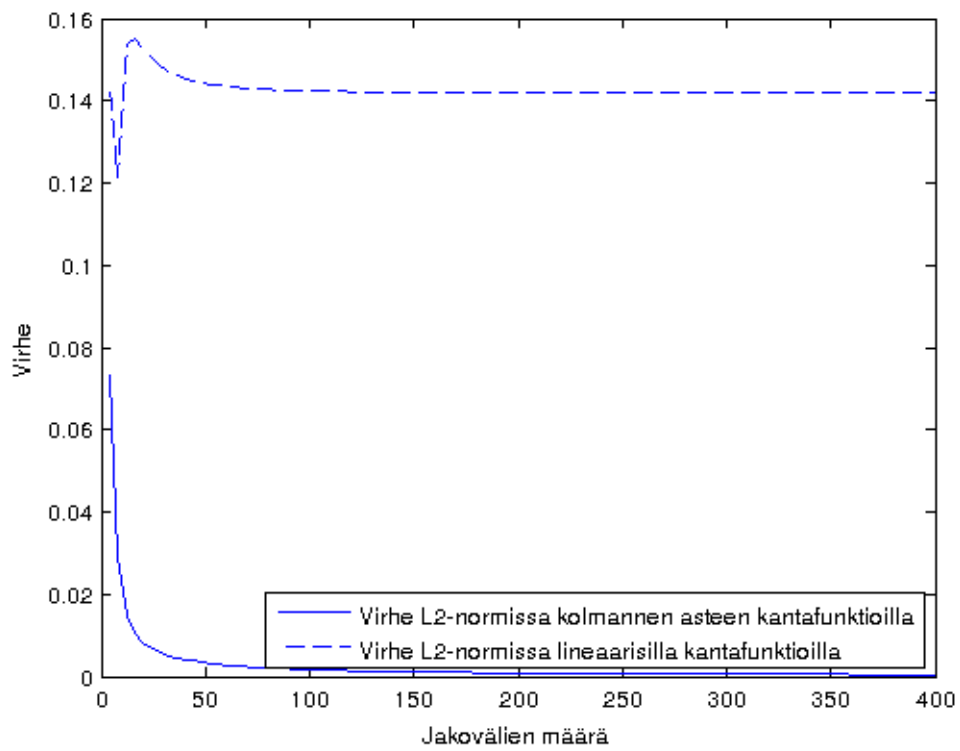
Kuva 6.6: Tarkkuus paranee



Kuva 6.7: Tarkkuus paranee



Kuva 6.8: Suppeneminen energianormin suhteen



Kuva 6.9: Suppeneminen L2-normin suhteen

# Kirjallisuus

- [1] Antti Huhtala Antti Hannukainen Mika Juntunen. *Mat-1.3650 Finite Element Methods I*. Mekaniikan laitos, Aalto yliopisto. 2015. URL: [https://mycourses.aalto.fi/pluginfile.php/417410/mod\\_resource/content/1/Lecture%20notes.pdf](https://mycourses.aalto.fi/pluginfile.php/417410/mod_resource/content/1/Lecture%20notes.pdf).
- [2] Susanne C. Brenner. *The mathematical theory of finite element methods*. Toim. L. Ridgway Scott. Texts in applied mathematics. New York: Springer, 1994.
- [3] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Studies in Mathematics and its Applications. Elsevier Science, 1978. ISBN: 9780080875255. URL: <https://books.google.fi/books?id=TpHfoXnpKvAC>.
- [4] Emmanuele DiBenedetto. *Partial differential equations*. Boston: Birkhäuser, 1995.
- [5] Avner Friedman. *Foundations of modern analysis*. New York: Dover, 1982.
- [6] kirjoittaja Harjulehto Petteri. *Analyysia reaaliluvuilla*. Turku: Petteri Harjulehto, 2014.
- [7] Thomas J. R. Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Blackwell Publishing Ltd, 1989. ISBN: 0486411818.
- [8] Claes Johnson. *Numerical solutions of partial differential equations by the finite element method*. Lisäpainokset: 2nd pr. 1992. Lund: Studentlitteratur, 1987.
- [9] Hans-Olav Tylli Kari Astala Petteri Piiroinen. *Funktionaalianalyysin peruskurssi*. Matematiikan ja tilastotieteen laitos, Helsingin Yliopisto, 2010. URL: <https://wiki.helsinki.fi/download/attachments/87265589/FAPk.pdf?version=1&modificationDate=1358837301082&api=v2>.
- [10] Zhang S. *Successive subdivisions of tetrahedra and multigrid methods on tetrahedral meshes*. Houston: J. of Math.
- [11] W.A. Strauss. *Partial Differential Equations: An Introduction*. Wiley, 2007. ISBN: 9780470054567. URL: <https://books.google.fi/books?id=PihAPwAACAAJ>.