

Keijo Mattila

Sähkömagneettisen sironnan numeerinen simulointi

Tietotekniikan
(tieteellinen laskenta)
pro gradu -tutkielma
12. tammikuuta 2004

Jyväskylän yliopisto

Tietotekniikan laitos

Jyväskylä

Tekijä: Keijo Mattila

Yhteystiedot: kemattil@st.jyu.fi

Työn nimi: Sähkömagneettisen sironnan numeerinen simulointi

Title in English: Numerical Simulation of Electromagnetic Scattering

Työ: Tietotekniikan (tieteellinen laskenta) pro gradu -tutkielma

Sivumäärä: 82

Tiivistelmä: Tässä tutkielmassa tarkastellaan, kuinka simuloidaan sähkömagneettisen aallon heijastumista yksinkertaisen kappaleen pinnalta. Pinta oletetaan täydelliseksi johteeksi. Matemaattinen malli johdetaan Maxwellin yhtälöistä. Malli diskretisoidaan äärellisten elementtien menetelmällä. Diskretisoinnissa käytetään ensimmäisen asteen Nédélec -kuutioelementtejä. Tuloksena saatava lineaarinen yhtälöryhmä ratkaistaan GMRES menetelmällä. Tutkielmassa esitettyjä menetelmiä sovelletaan käytäntöön simuloimalla muutama yksinkertainen esimerkkitapaus.

English abstract: In this thesis we study how to simulate scattering of an electromagnetic wave from a perfectly conducting obstacle with simple geometry. Mathematical model is derived from the Maxwell's equations. Discrete representation of the model is formed with Finite Element Method. In discretization first order Nédélec -cubic elements are used. Resulting linear system of equations is solved with GMRES method. Methods presented in this thesis are put in practice by simulating few modest case studies.

Avainsanat: Maxwellin yhtälöt, Silver-Müller, FEM, Nédélec -elementti, GMRES

Keywords: Maxwell's equations, Silver-Müller, FEM, Nédélec -element, GMRES

Esipuhe

Lähes vuoden työrupeama saapuu päätökseensä tämän tutkielman valmistumisen myötä. Viimeinen – samalla myös mieluisin – tehtäväni on jakaa kiitosta niille henkilöille, jotka sen mielestäni ansaitsevat. Kun tutkielmani ohjaajat filosofian tohtori Janne Martikainen ja filosofian tohtori Tuomo Rossi saattoivat tutkielman aiheen korviini noin vuosi takaperin, onnistuivat he samalla kampeamaan mielenkiintoni uuteen osoitteeseen. Aikaa myöten ilmeni, että työskentely Jannen ja Tuomon ohjauksessa on paitsi tuotteliasta myös aidosti välitöntä. Apua oli aina saatavilla. Varsinkin siitä, mutta myös kaikesta muusta hyvästä, molemmille kuuluu kiitos ja puolet kunniasta.

Osansa saavat kotiväkikin. Vanhemmiltani Pentiltä ja Mirjamilta olen vastaanottanut paljon. Ennen kaikkea olen ollut vapaa lähtemään ja saapumaan. Heille olen alati kiitoksen velkaa. Kiitoksen ansaitsevat myös siskoni ja veljeni, joiden iloisesta seurasta olen nauttinut lukemattomia kertoja. Mikäli nauru todellakin pidentää ikää, Mervin ja Markon seurassa minulle on karttunut useita lisävuosia.

Ystävääni Anssi Pennasta kiitän yhteistyöstä, jota olemme opiskelujemme edistämiseksi tehneet. Se on ollut antoisaa ajatusten vaihtoa. Jyväskylässä asumani vuodet ovat tarjonneet ikimuistoisia hetkiä myös opiskelun ulkopuolelta. Niistä hetkistä kiitän kaikkia ystäviäni ja läheisiäni tasapuolisesti.

Jyväskylä, joulukuu 2003

Keijo Mattila

Matemaattiset merkinnät

Alla esitetty lista sisältää tämän tutkielman tärkeimmät matemaattiset merkinnät. Suurin osa merkinnöistä on pyritty selittämään myös varsinaisessa tekstissä siinä yhteydessä, missä niitä ensimmäisen kerran käytetään. Käytäntönä on, että pienet kreikkalaiset aakkoset sekä kursivoidut latinalaiset aakkoset, esimerkiksi i, α ja E , edustavat kokonais-, reaali- tai kompleksilukuja, toisin sanoen skalaareita. Edellä mainitun kaltaiset aakkoset voivat edustaa myös skalaariarvoisia funktioita. Lihavoidut latinalaiset aakkoset sekä nuolella varustetut kreikkalaiset aakkoset, esimerkiksi \mathbf{q}, φ ja \mathbf{B} , edustavat vektoreita tai vektoriarvoisia funktioita. Isot latinalaiset aakkoset ilman erikoistehosteita, esimerkiksi A ja J , edustavat matriiseja.

\mathbb{N}	positiiviset kokonaisluvut, luku nolla kuuluu joukkoon
\mathbb{R}^n	n -ulotteinen reaalilukuavaruus
\mathbb{C}^n	n -ulotteinen kompleksilukuavaruus
\mathbb{K}^n	n -ulotteinen reaali- tai kompleksilukuavaruus
$\gamma \in \mathbb{R}$	γ on reaaliluku
\mathbf{i}	imaginääriyksikkö $\sqrt{-1}$
π	pii ≈ 3.142
e	Neperin luku ≈ 2.178
$\mathbf{u} \in \mathbb{C}^n$	\mathbf{u} on vektori ja sen n komponenttia ovat kompleksiarvoisia, joskus vektori määritellään kirjoittamalla $\{u_i\}_{i=1}^n$, $u_i \in \mathbb{C}$
$A \in \mathbb{K}^{m \times n}$	A on matriisi, jonka alkiot ovat reaali- tai kompleksilukuja, matriisissa on m riviä ja n saraketta; neliömatriisit määritellään joskus muodossa $\{a_{i,j}\}_{i,j=1}^n$, $a_{i,j} \in \mathbb{K}$
A^{-1}, A^T	matriisin A käänteismatriisi ja transpoosi
$\det(A)$	matriisin A determinantti
$(\mathbf{u})_i$	vektorin \mathbf{u} i :nnes komponentti, kun vektorilla on n komponenttia ja $i = 1, \dots, n$; myös merkintää u_i käytetään
$(A)_{i,j}$	matriisin A alkio rivillä i sarakkeessa j , kun $i = 1, \dots, m$ ja $j = 1, \dots, n$, missä m on matriisin rivien ja n sarakkeiden lukumäärä; toisinaan käytetään merkintää $a_{i,j}$
$[a, b],]a, b[\subset \mathbb{R}$	suljettu ja avoin reaalilukuväli, missä $a, b \in \mathbb{R}$
$\Omega \subset \mathbb{R}^n$	n -ulotteisen reaalilukuavaruuden avoin osajoukko, tässä tutkielmassa osajoukoilla tarkoitetaan yleensä alueita
$\overline{\Omega}$	avoimen joukon Ω sulkeuma
$\partial\Omega$	alueen Ω reuna
Γ	reunan osajoukko, esimerkiksi $\Gamma \subset \partial\Omega$

$\Omega_a \setminus \Omega_b$	joukko-opillinen vähennysoperaatio, luetaan " Ω_a :sta pois Ω_b "
$ w $	skalaarin w itseisarvo
$\ \mathbf{u}\ $	vektorin \mathbf{u} euklidinen normi
(\mathbf{u}, \mathbf{v}) tai $\mathbf{u} \cdot \mathbf{v}$	vektoreiden \mathbf{u} ja \mathbf{v} sisätulo
$\mathbf{u} \times \mathbf{v}$	kolmiulotteisten vektoreiden \mathbf{u} ja \mathbf{v} ristitulo
$\mathbf{g}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{C}^n$	funktio, joka kuvaa n -ulotteisen reaalilukuavaruuden pisteen n -ulotteiseen kompleksilukuavaruuteen
$\mathbf{n} \in \mathbb{R}^n$	reunan tai pinnan pisteen ulkonormaalivektori, yleensä yksikkövektori
$\mathbf{t} \in \mathbb{R}^n$	reunan tai pinnan pisteen tangenttivektori, yleensä yksikkövektori
$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_1}$	funktion \mathbf{f} osittaisderivaatta muuttujan x_1 suhteen
$\nabla f(\mathbf{x})$	skalaariarvoisen funktion f gradientti; $\nabla f = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$, kun $\mathbf{x} \in \mathbb{R}^n$
$\frac{\partial f(\mathbf{x})}{\partial \mathbf{n}}$	funktion f derivaatta ulkonormaalivektorin \mathbf{n} suuntaan, vaihtoehtoinen merkintä on $\nabla f(\mathbf{x}) \cdot \mathbf{n}$
$\nabla \cdot \mathbf{f}(\mathbf{x})$	vektoriarvoisen funktion $\mathbf{f} \in \mathbb{K}^n$ divergenssi; $\nabla \cdot \mathbf{f} = \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}$, kun $\mathbf{x} \in \mathbb{R}^n$
$\nabla \times \mathbf{f}(\mathbf{x})$	funktion $\mathbf{f} \in \mathbb{K}^3$ roottori; $\nabla \times \mathbf{f} = (\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2})$, kun $\mathbf{x} \in \mathbb{R}^3$
$\int_{\Omega} \mathbf{f}(\mathbf{x}) d\mathbf{x}$	funktion \mathbf{f} integraali yli joukon Ω muuttujan \mathbf{x} suhteen
dS	joukon S mielivaltaisen pieni osajoukko
$\oint_S \cdot dS$	integrointi yli suljetun joukon S
$\sum_{i=1}^n b_i$	lukujen b_1, \dots, b_n summa
δ_{ij}	Kroneckerin symboli, joka saa arvon 1, kun $i = j$; muulloin symboli saa arvon 0
ω	kulmataajuus
λ	aallonpituus, joissakin yhteyksissä myös ominaisarvo
α	aaltoluku
\mathbf{E}	sähkökenttä
\mathbf{B}	magneettikenttä
\mathbf{J}	Jacobin matriisi

Sisältö

Esipuhe	i
Matemaattiset merkinnät	ii
Sisältö	iv
Kuvat	vi
1 Johdanto	1
2 Sähkömagnetismi	3
2.1 Varaukset	3
2.1.1 Sähköiset voimat	3
2.1.2 Varaustiheys	4
2.2 Staattinen sähkökenttä	5
2.2.1 Sähkökentän vuo	6
2.2.2 Gaussin laki	7
2.3 Sähkövirta	9
2.3.1 Virrantiheys	10
2.3.2 Jatkuvuusyhtälö	11
2.4 Staattinen magneettikenttä	12
2.4.1 Magneettiset voimat	12
2.4.2 Magneettikentän vuo	13
2.5 Magneettikentän alkuperä	15
2.5.1 Sähkövirran aiheuttama magneettikenttä	16
2.5.2 Ampéren laki	16
2.6 Dynaaminen sähkö- ja magneettikenttä	18
2.6.1 Faradayn laki	18
2.6.2 Kenttävirta	20
2.7 Maxwellin yhtälöt	21
3 Matemaattinen malli	22
3.1 Aallon eteneminen	22
3.2 Aallon heijastuminen	24
3.3 Absorboiva reunaehto	26
3.4 Mallin kokoaminen	27

4	Heikko muoto	28
4.1	Funktioavaruudet	28
4.2	Integraaliyhtälö	30
4.3	Simulointitehtävän heikko muotoilu	32
5	Simulointitehtävän diskretisointi	34
5.1	Äärellisten elementtien menetelmä	34
5.2	Elementtimenetelmän peruskäsitteitä	35
5.2.1	Kantafunktiot	36
5.2.2	Viite-elementit	37
5.2.3	Äärellinen elementti	39
5.3	Nédélec-elementti	40
5.4	Yhtälöryhmän kokoaminen	45
6	Yhtälöryhmän ratkaiseminen	49
6.1	Iteratiiviset menetelmät	49
6.2	Projektiomenetelmät	51
6.2.1	Projektiomenetelmien luokittelu	54
6.2.2	Krylovin aliavaruusmenetelmät	56
6.3	GMRES	58
6.3.1	Arnoldin menetelmä	58
6.3.2	Algoritmi GMRES menetelmälle	59
6.3.3	Algoritmin käytännöllinen toteutus	61
7	Laskentaesimerkkejä	65
7.1	Esimerkkitapausten simulointi	65
7.2	Tulosten tarkastelu	67
7.3	Johtopäätökset	70
8	Yhteenveto	72
9	Kirjallisuutta	73

Kuvat

2.1	Pistevarauksien väliset voimat	4
2.2	Positiivisen varauksen muodostama sähkökenttä	6
2.3	Pallon pinnan läpäisevä sähkökenttä	7
2.4	Hahmotelma osa-alueen projektiosta pallon pinnalle	8
2.5	Johteessa kulkevan sähkövirran suunta	9
2.6	Johteen poikkileikkauksen läpi kulkeva sähkövirta	10
2.7	Liikkuvaan varaukseen vaikuttava magneettinen voima	13
2.8	Sähkövirran aiheuttama magneettikenttä	14
2.9	Liikkuvan varauksen aiheuttama magneettikenttä	15
2.10	Integrointipolku magneettikentän viivaintegraalin laskemiseen	17
3.1	Sähkömagneettinen aalto etenee x-akselin suuntaan	24
3.2	Sähkömagneettinen aalto heijastuu suoralta pinnalta	25
3.3	Kaksiulotteinen rajoitettu simulointialue	26
5.1	Kaksiulotteisen monikulmion jako kolmioelementteihin	37
5.2	Viite-elementin kuvautuminen alkuperäiseen koordinaatistoon	38
5.3	Nédélec-kuutioviite-elementti ja sen särmien numerointi	43
6.1	Vektorin vinoprojektio aliavaruuteen \mathcal{K}	52
6.2	Residuaalivektorin kohtisuoruus avaruutta \mathcal{L} vastaan.	53
6.3	Projektiomenetelmien luokittelukaavio	56
7.1	Harva verkko ja heijastava kuutio.	66
7.2	Portaasta heijastuneen aallon muodostama sähkökenttä.	68
7.3	Poikkileikkauskuva tiheimmällä verkolla lasketusta sähkökentästä.	69
7.4	Residuaalivektorien normin suppeneminen iteraatioaskelten edetessä.	69

1 Johdanto

Tämän tutkielman aiheena on sähkömagneettisen sironnan numeerinen simulointi. Sähkömagneettinen aalto siroaa, kun se kohtaa rajapinnan, joka erottaa kaksi eri väliainetta. Sähkömagneettisia aaltoja ovat esimerkiksi radioaallot, eriväriset laservalot tai auringonvalo. Väliaineiden rajapinnalla osa aallosta heijastuu takaisin ja osa läpäisee rajapinnan. Luonnollinen esimerkki siroamisilmiöstä on auringonvalon heijastuminen ja taivuttuminen vedenpinnalta. Ilmiö esiintyy myös monissa nykypäivän sovelluksissa. Ilmeisenä esimerkkinä mainittakoon tutka, jolla voidaan havaita vaikkapa lentokone heijastuneen radioaallon avulla.

Numeerisella simuloinnilla tarkoitetaan, että simuloitavaa ilmiötä mallinnetaan tietokoneen avulla; ilmiön käyttäytymiselle lasketaan arvio hyödyntämällä tietokoneen laskentatehokkuutta. Numeerinen simulointi etenee yleensä seuraavasti:

1. Selvitetään simuloitavan ilmiön fysikaaliset perusteet.
2. Muodostetaan ilmiötä kuvaava matemaattinen malli.
3. Muokataan matemaattista mallia niin, että se soveltuu ratkaistavaksi jollakin numeerisella menetelmällä.
4. Laaditaan muokattua mallia vastaava diskreetti tehtävä, joka usein muotoillaan lineaariseksi yhtälöryhmäksi.
5. Ratkaistaan yhtälöryhmä tilanteeseen sopivalla ratkaisumenetelmällä.
6. Lasketaan lopullinen simulointitehtävän ratkaisu, joka valtaosalle tehtävistä on jokin fysikaalinen suure, ja tarkastetaan tulokset.

Tämän tutkielman rakenne noudattaa yllä esitettyä vaihejakoa. Luvussa 2 esitellään sähkömagnetismin perusteoriaa ja johdetaan niin sanotut *Maxwellin yhtälöt*, joihin sähkömagneettisen aallon simulointi tulee perustumaan. Luvussa 3 laaditaan ilmiötä kuvaava matemaattinen malli Maxwellin yhtälöiden pohjalta. Perusajatuksena on, että sirontailmiötä kuvataan sopivalla funktiolla, jonka lauseketta ei kuitenkaan suoraan tunneta. Sen sijaan funktion käyttäytymiselle pystytään asettamaan sellaiset ehdot, että niiden avulla funktion lauseke periaatteessa kyettäisiin yksikäsitteisesti päättämään. Mikäli lausekkeen päättely onnistuu, sanotaan että matemaattiselle mallille löydetään analyttinen ratkaisu.

Sirontailmiötä kuvaava matemaattinen malli tulee olemaan niin monimutkainen, ettei sitä kyetä analyttisesti ratkaisemaan. Silloin joudutaan turvautumaan numeeriseen

simulointiin, joka karkeasti sanoen tarkoittaa arvion laskemista analyttiselle ratkaisulle. Varsinainen lähtölaukaus numeeriseen simulointiin tapahtuukin vasta luvussa 4, kun mallista muodostetaan integraaliyhtälö, joka vastaa heikossa mielessä matemaattista mallia ja joka soveltuu paremmin numeeriseen simulointiin. Luvussa 5 laaditaan heikkoa muotoa vastaava diskreetti tehtävä. Operaatio lienee helpoimmin ymmärrettävissä niin, että siinä asetetaan uusi tehtävä, jonka ratkaisu approksimoi matemaattisen mallin analyttistä ratkaisua. Diskretisointi suoritetaan *äärellisten elementtien menetelmällä* ja tuloksena syntyy lineaarinen yhtälöryhmä.

Tässä vaiheessa sirontailmiön simulointi on kiteytynyt yhtälöryhmän ratkaisemiseen. Itse asiassa samaan tilanteeseen päädytään erittäin monissa numeerisen laskennan tehtävätyypeissä. Esimerkiksi optimointitehtävien kohdalla tilanne on tavallinen. Koska yhtälöryhmän ratkaiseminen on eräs numeerisen laskennan perusongelmia, luvussa 6 tarkastellaan iteratiivisia ratkaisumenetelmiä hieman laajemmin. Erityisesti esitellään *GMRES* menetelmä, jota tullaan soveltamaan simulointitehtävissä. Lopuksi luvussa 7 lasketaan esimerkin vuoksi vaatimattomia simulointitehtäviä ja analysoidaan saatuja tuloksia.

2 Sähkömagnetismi

Siroava sähkömagneettinen aalto kuuluu fysiikassa sähködynamiikan aihepiiriin. Sähködynamiikassa tutkitaan sähkö- ja magneettikenttiä, jotka ovat sekä paikka- että aikariippuvia. Fysiikassa nimitystä *kenttä* käytetään sellaisesta kuvauksesta, joka kuvaa avaruuden pisteen vektoriarvoiseksi suureeksi. Esimerkiksi aikariippuva sähkökenttä kuvaa pisteen vektorisuureeksi, joka ilmaisee kyseisessä pisteessä tietyllä ajanhetkellä vaikuttavien sähköisten voimien suuruuden ja suunnan.

Nykykäsityksen mukaan kaikki sähkö- ja magneettikentät toteuttavat *Maxwellin yhtälöt*. Seuraavaksi johdetaan Maxwellin yhtälöt lähtien staattisista sähkökentistä ja päätyen dynaamisten sähkö- ja magneettikenttien vuorovaikutukseen. Yhtälöiden johdaminen on lainattu pääasiassa kirjasta [33] ja luentomonisteesta [16]. Muihin lähteisiin viitataan erikseen.

2.1 Varaukset

Jo 1700-luvun lopulla opittiin, että on olemassa kahdenlaisia varauksia: positiivisia ja negatiivisia. Kahden liikkumattoman pistevarauksen välillä vallitsee voiman ja vasta-voiman lakia noudattava sähköstaattinen vuorovaikutus:

- Varaukset vaikuttavat toisiinsa voimilla, joiden suunta on niitä yhdistävän suoran suuntainen ja kääntäen verrannollinen varausten välisen etäisyyden neliöön.
- Voimat ovat verrannollisia varausten tuloon siten, että samanmerkkiset varaukset hylkivät toisiaan ja erimerkkiset vetävät toisiaan puoleensa.

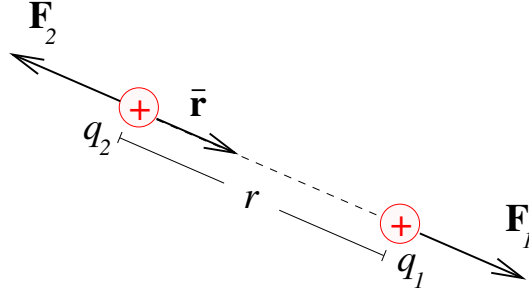
Sähköstaattista vuorovaikutusta kutsutaan *Coulombin* laiksi ja se perustuu kokeellisiin havaintoihin.

2.1.1 Sähköiset voimat

Matemaattisesti ilmaistuna Coulombin laki kertoo, että varaus q_2 vaikuttaa varaukseen q_1 *sähköstaattisella* voimalla

$$\mathbf{F}_1 = k \frac{q_1 q_2}{r^2} \bar{\mathbf{r}}. \quad (2.1)$$

Yksikkövektori $\bar{\mathbf{r}}$ osoittaa varauksesta q_2 varaukseen q_1 kuvan 2.1 mukaisesti, r on varausten välinen etäisyys ja k on verrannollisuuskerroin, joka riippuu käytetystä yksikköjärjestelmästä. Tässä tutkielmassa käytetään SI-järjestelmää, jolloin verrannollisuuskerroin



Kuva 2.1: Kahden positiivisen pistevarauksen välillä vaikuttavat sähköiset voimat.

$$k = \frac{1}{4\pi\epsilon_0}$$

määrittellen tyhjiön permittiivisyyden $\epsilon_0 \approx 8.854 \cdot 10^{-12}$ F/m avulla. Permittiivisyyden yksikkönä on faradi/metri ($\text{F/m} = \text{C}^2/\text{Nm}$) ja varauksen yksikkö coulombi (C).

Useiden varausten muodostamassa systeemissä varaukseen q_i vaikuttava voima lasketaan *superpositioperiaatteella*

$$\mathbf{F}_i = q_i \sum_{j \neq i}^N \frac{q_j}{4\pi\epsilon_0} \frac{\mathbf{r}_{ji}}{r_{ji}^3}, \quad (2.2)$$

joka vastaa kokeellisista mittauksista saatuja tuloksia. Vektori \mathbf{r}_{ji} lähtee varauksesta q_j päättyen varaukseen q_i ja sen pituus on r_{ji} .

2.1.2 Varaustiheys

Kokeellisesti on osoitettu, että yhden elektronin varaus $e \approx 1.6019 \cdot 10^{-19}$ C on pienin luonnossa vapaana esiintyvä varaus ja siitä syystä elektroni toimii yksikkövarauksena. Yhden elektronin varaus on kuitenkin niin pieni, että makroskooppinen varausjakauma muodostuu yleensä suuresta joukosta yksikkövarauksia. Makroskooppisella varausjakaumalla tarkoitetaan varausten sijaintia tilavuudessa, jonka mittakaava on suuri verrattuna varauksien kokoon. Suurten varausjoukkojen käsittelyssä voidaan hyödyntää varaustiheyden käsitettä.

Oletetaan makroskooppinen varausjakauma jatkuvaksi ja sileäksi, jolloin kolmiulotteisessa avaruudessa sijaitsevan tilavuuden varaustiheydeksi määritellään

$$\rho(\mathbf{x}) = \lim_{\Delta V \rightarrow 0} \frac{\Delta q(\mathbf{x})}{\Delta V(\mathbf{x})}$$

ja pintavaraustiheydeksi vastaavasti

$$\sigma(\mathbf{x}) = \lim_{\Delta S \rightarrow 0} \frac{\Delta q(\mathbf{x})}{\Delta S(\mathbf{x})},$$

missä V on tarkasteltava tilavuus ja S tarkasteltava pinta. Varaustiheyden määrittelyssä $\Delta V(\mathbf{x})$ on tilavuus pisteen $\mathbf{x} \in \mathbb{R}^3$ ympäristössä ja $\Delta q(\mathbf{x})$ ilmoittaa kyseisessä tilavuudessa sijaitsevien varauksien määrän. Pintavaraustiheyden merkinnät ovat vastaavia.

Varaustiheyden avulla suuren varausjoukon voimavaikutusta ei tarvitse enää laskea summana (2.2), vaan voidaan siirtyä käyttämään integraalilausekkeita. Olkoon tilavuudessa V varaustiheys ρ ja V :tä rajoittavalla pinnalla S pintavaraustiheys σ . Tällöin pisteessä \mathbf{r} olevaan varaukseen q vaikuttaa voima

$$\mathbf{F}_q = \frac{q}{4\pi\epsilon_0} \int_V \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \rho(\mathbf{r}') dV + \frac{q}{4\pi\epsilon_0} \int_S \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \sigma(\mathbf{r}') dS, \quad (2.3)$$

missä \mathbf{r}' on niin sanottu lähdepiste äärettömän pienestä tarkastelualueesta dV tai dS .

2.2 Staattinen sähkökenttä

Tässä luvussa tutustutaan sähkövarausten aiheuttamaan staattiseen sähkökenttään. Sähköstaattinen vuorovaikutus ajatellaan kaksivaiheiseksi. Yhden tai useamman paikallaan olevan varauksen muodostama staattinen systeemi aiheuttaa *sähkökentän* $\mathbf{E}(\mathbf{x})$. Olkoon pisteessä \mathbf{r} varaus q , joka ei kuulu sähkökentän muodostavaan staattiseen systeemiin. Sähkökenttä \mathbf{E} vaikuttaa varaukseen q voimalla

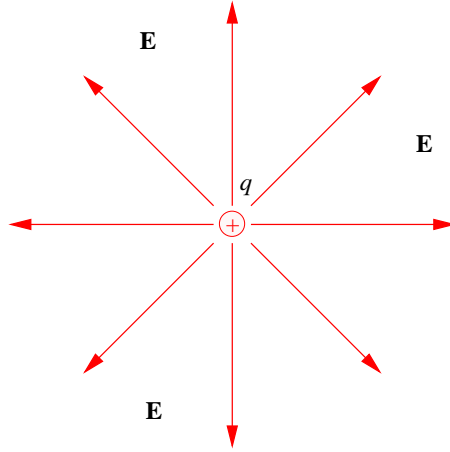
$$\mathbf{F}(\mathbf{r}) = q\mathbf{E}(\mathbf{r}). \quad (2.4)$$

Koska usean varauksen muodostamalle kokonaisvoimalle \mathbf{F} pätee superpositioperiaate (2.2), niin vastaavanlainen periaate pätee myös sähkökentälle. Usean varauksen muodostama kokonaissähkökenttä \mathbf{E} lasketaan yksittäisten varausten muodostamien sähkökenttien summana. Kuvassa 2.2 on esitetty positiivisen pistevarauksen q ympäristöönsä aiheuttama sähkökenttä \mathbf{E} kenttäviivojen avulla. Kenttäviiva on apuväline, jolla havainnollistetaan todellista sähkökenttää ja joka on jokaisessa pisteessään varsinaisen sähkökentän suuntainen.

Yksittäisten varausten ja varausjakautumien muodostama kokonaissähkökenttä pisteessä \mathbf{r} on

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^N q_i \frac{\mathbf{r} - \mathbf{r}_i}{\|\mathbf{r} - \mathbf{r}_i\|^3} + \frac{1}{4\pi\epsilon_0} \int_V \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \rho(\mathbf{r}') dV' + \frac{1}{4\pi\epsilon_0} \int_S \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|^3} \sigma(\mathbf{r}') dS'.$$

Periaatteessa sähkökenttä voidaan siis määrittää laskemalla kaikkien varausjakautumien ja yksittäisten hiukkasten aiheuttamat kentät. Käytännössä tämä on usein täysin ylivoimainen tehtävä.



Kuva 2.2: Positiivisen pistevarauksen ympäristöönä muodostama sähkökenttä.

2.2.1 Sähkökentän vuo

Tarkastellaan origossa sijaitsevan pistevarauksen q muodostamaa sähkökenttää pisteessä \mathbf{r}

$$\mathbf{E}(\mathbf{r}) = \frac{q}{4\pi\epsilon_0} \frac{\bar{\mathbf{r}}}{r^2},$$

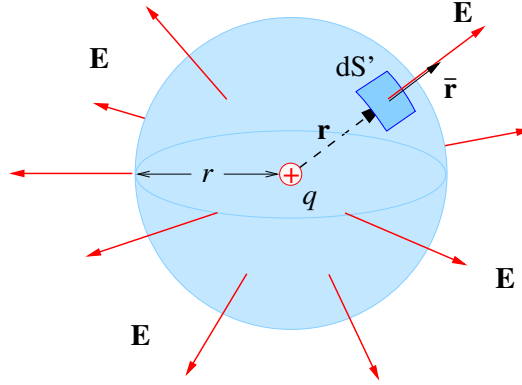
missä $\bar{\mathbf{r}}$ on origosta pisteeseen \mathbf{r} osoittava yksikkövektori. Ympäröidään varaus q pallolla V' , jonka keskipiste on origossa ja jonka säde on r . Tällöin tarkasteltava piste \mathbf{r} sijaitsee pallon pinnalla S' ja $\bar{\mathbf{r}}$ ajatellaan pallon pinnan ulkonormaalivektoriksi. Kuva 2.3 havainnollistaa tilannetta. Tarkoituksena on laskea sähkökentän \mathbf{E} vuo pallon pinnalla S' . Kuvainnollisesti sanottuna sähkökentän vuolla tarkoitetaan sähkökentän "virtausta" suljetun pinnan läpi. Tosiasiassa sähkökenttää ei virtaa samalla tavalla kuin esimerkiksi neste, mutta mielikuva virtauksesta auttaa hahmottamaan vuon abstraktia käsitettä.

Vuo lasketaan summaamalla sähkökentän voimakkuus tarkasteltavan pinnan ulkonormaalivektorin suuntaan jokaisessa pinnan pisteessä. Symbolilla E_{\perp} tarkoitetaan sen sähkökentän komponentin suuruutta, joka on kohtisuorassa tarkasteltavaa pintaa vastaan. Pallon V' pinnalla origossa sijaitsevan varauksen q muodostaman sähkökentän suunta on kohtisuorassa pintaa vastaan, jolloin $\mathbf{E} \cdot \bar{\mathbf{r}} = E_{\perp} = E$. Sähkökentän \mathbf{E} voimakkuus pinnalla S' on

$$E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2}.$$

Jaetaan pallon V' pinta S' äärettömän pieniin osa-alueisiin dS' . Pallon pinta-ala saadaan laskemalla integraali

$$A = \oint_{S'} dS' = 4\pi r^2.$$



Kuva 2.3: Varauksen ympärillä sijaitsevan pallon pinnan läpäisevä sähkökenttä.

Sähkökentän vuo lasketaan integraalina suljetun pinnan yli. Origossa sijaitsevan pistevarauksen q aiheuttaman sähkökentän \mathbf{E} vuo pallon V' pinnalla on

$$\Phi_E = \oint_{S'} \mathbf{E} \cdot \bar{\mathbf{r}} \, dS' = E \oint_{S'} dS' = EA = \frac{q}{\epsilon_0}. \quad (2.5)$$

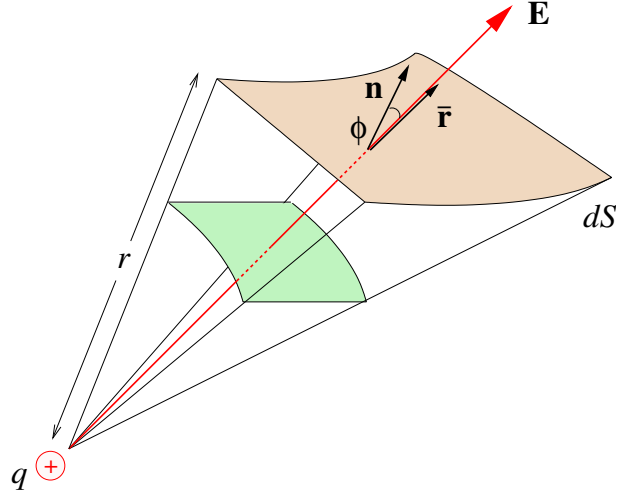
Havaitaan, että laskettu vuo Φ_E on riippuvainen ainoastaan varauksesta q . Erityisesti vuo ei ole riippuvainen pallon säteestä r . Jos esimerkiksi pallon sädettä r kasvatetaan, niin sähkökentän voimakkuus E pienenee samassa suhteessa kuin pinta-ala A kasvaa, jolloin säteen r vaikutus kumoutuu. Seuraavaksi lasketaan sähkökentän vuo varausta q ympäröivän mielivaltaisen kappaleen V pinnalla S . Laskuissa käytetään hyväksi tietoa, ettei pallon pinnalla laskettu vuo ole riippuvainen pallon säteestä.

Kappaleen V pinta jaetaan äärettömän pieniin osa-alueisiin dS . Ajatellaan, että mielivaltainen kappale V pitää sisällään pallon V' ja jokainen osa-alue projisoidaan pallon pinnalle kuvan 2.4 hahmottelemalla tavalla. Osa-alueen dS projektio pallon pinnalle on $dS' = \bar{\mathbf{r}} \cdot \mathbf{n} \, dS$, missä \mathbf{n} on mielivaltaisen pinnan S ulkonormaalivektori [18, sivu 50]. Lasketaan sähkökentän vuo erikseen jokaisessa osa-alueessa dS projektion dS' avulla ja summataan tulokset integroimalla. Yhtälöä (2.5) käyttämällä saadaan lopputuloksena kokonaisvuo pallon pinnalla S' , joka on myös sähkökentän \mathbf{E} vuo mielivaltaisella pinnalla S

$$\Phi_E = \oint_S \mathbf{E} \cdot \mathbf{n} \, dS = E \oint_S \bar{\mathbf{r}} \cdot \mathbf{n} \, dS = E \oint_{S'} dS' = \frac{q}{\epsilon_0}. \quad (2.6)$$

2.2.2 Gaussin laki

Edellä on käsitelty ainoastaan yksittäisen pistevarauksen aiheuttaman sähkökentän vuota suljetulla mielivaltaisella pinnalla. Yleistetään yhtälö (2.6) myös tapauksille, joissa kappaleen V sisällä on määritelty varaustiheys ρ . Jatkossa varaustiheyttä kutsutaan



Kuva 2.4: Hahmotelma mielivaltaisen pinnan osa-alueen projektiosta pallon pinnalle.

varausjakaumaksi. Kappaleen V sisältämä kokonaisvaraus lasketaan varausjakauman avulla

$$Q = \int_V \rho \, dV.$$

Kokonaisvarauksen määritelmää käyttäen yhtälö (2.6) kirjoitetaan yleiseen muotoon

$$\oint_S \mathbf{E} \cdot \mathbf{n} \, dS = \frac{1}{\epsilon_0} \int_V \rho \, dV, \quad (2.7)$$

joka on *Gaussin laki* integraalimuodossa. Laki muotoillaan differentiaalimuotoon käyttämällä seuraavaksi määriteltävää *divergenssiteoreemaa*, josta käytetään myös nimitystä *Gaussin lause*. Riittävän siistille vektorikentälle \mathbf{u} pätee

$$\oint_S \mathbf{u} \cdot \mathbf{n} \, dS = \int_V \nabla \cdot \mathbf{u} \, dV, \quad (2.8)$$

missä \mathbf{n} on tilavuutta V ympäröivän pinnan S ulkonormaalivektori.

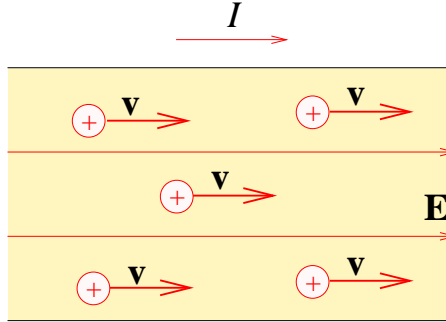
Sovelletaan yhtälöä (2.8) Gaussin lain (2.7) vasempaan puoleen, jolloin on voimassa

$$\int_V \nabla \cdot \mathbf{E} \, dV = \frac{1}{\epsilon_0} \int_V \rho \, dV.$$

Koska yhtälön täytyy olla riippumaton tilavuuden V valinnasta, supistuu se muotoon

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (2.9)$$

joka on Gaussin laki *differentiaalimuodossa*. Sanoin ilmaistuna Gaussin laki määrittelee, että sähkökentän vuo suljetun pinnan yli on yhtä kuin pinnan rajaaman tilavuuden kokonaisvaraus jaettuna vakiolla ϵ_0 . Toisin sanoen laki määrittelee yhteyden pinnan sähkökentälle ja pinnan rajaaman alueen varausjakaumalle.



Kuva 2.5: Sähkökentän liikuttamien varausten johteeseen aiheuttama sähkövirta.

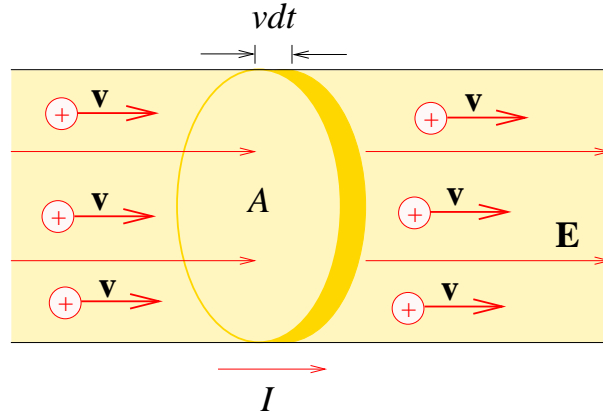
Gaussin laki on *Maxwellin* yhtälöistä ensimmäinen. Jos Maxwellin yhtälöt kirjoitetaan integraalimuodossa, käytetään yhtälöä (2.7) ja differentiaalimuotoon kirjoitettaessa käytetään yhtälöä (2.9).

2.3 Sähkövirta

Edellisessä luvussa tarkasteltiin paikallaan olevien varausten aiheuttamaa sähkökenttää. Tässä luvussa siirrytään tutkimaan liikkuvia varauksia. Johteeseen muodostuu sähkövirta, kun johteen sisällä varauksia siirtyy alueesta toiseen. Kappaletta tai tilavuutta, joka koostuu sellaisesta materiaalista, että varaukset pystyvät siirtymään suhteellisen helposti alueesta toiseen, sanotaan *johteeksi*. Esimerkiksi kuparilanka on johde. Seuraavaksi tarkastellaan, miten sähkökenttää saa johteessa olevat varaukset liikkumaan, ja siten aiheuttaa johteeseen sähkövirran.

Tutkitaan tilannetta, jossa tyhjiössä sijaitsevaan varaukseen q kohdistuu vakiosähkökentän \mathbf{E} aiheuttama vakiovoima $\mathbf{F} = q\mathbf{E}$. Voima saa varauksen kiihtyvään liikkeeseen, jolloin sen nopeus kasvaa ajan edetessä. Johteissa tilanne on toisenlainen, koska liikkuvien varausten suunta muuttuu niiden törmäillessä materiaalihiukkasiin. Tästä seuraa, että johteissa vakiosähkökentän \mathbf{E} aiheuttama voima ei saa varauksia kiihtyvään liikkeeseen, vaan se saa varaukset liikkumaan vakionopeudella \mathbf{v} . Positiiviset varaukset liikkuvat sähkökentän suuntaan, negatiivisten varausten suunta on sähkökenttää vastaan.

Määritellään sähkövirran I suunnaksi positiivisten varausten liikesuunta. Kun positiiviset varaukset virtaavat kuvan 2.5 esittämällä tavalla, johteen oikeaan reunaan ajatellaan kertyvän positiivinen kokonaisvaraus. Kun varaukset ovat negatiivisia, niiden liikesuunta on sähkökenttää vastaan ja jälleen johteen oikeaan reunaan kertyy positiivinen kokonaisvaraus. Havaitaan, että positiivisten ja negatiivisten varausten kokonaisvaikutus on sama, jolloin ne aiheuttavat myös saman suuntaisen sähkövirran. Liikkuvien varausten merkillä ei siis ole vaikutusta virran suuntaan.



Kuva 2.6: Johteen poikkileikkauksen läpi kulkeva sähkövirta.

Jos ajan dt kuluessa johteessa sijaitsevan mielivaltaisen alueen läpi virtaa kokonaisvaraus dQ , joka muodostuu ainoastaan positiivisten varausten liikkeestä, niin sähkövirta I kyseisen alueen läpi on

$$I = \frac{dQ}{dt}. \quad (2.10)$$

Virran SI-yksikkö on ampeeri ($A = C/s$). On oleellista huomata, että vaikka sähkövirralle määriteltiinkin suunta, niin se ei kuitenkaan ole vektorisuure. Sähkövirran suunta kertoo ainoastaan sen, mitä tarkoitetaan positiivisella ja negatiivisella sähkövirralla.

2.3.1 Virrantiheys

Seuraavaksi määritellään virrantiheys sähkövirran I avulla. Tarkastellaan johteen poikkileikkausta A ja oletetaan, että kaikki johteessa liikkuvat varaukset ovat positiivisia. Kuvaan 2.6 on piirretty tarkasteltava poikkileikkaus ja virtaavat varaukset. Olkoon johteen yksikkötilavuudessa n hiukkasta, joiden kaikkien varaus on q . Muuttujaa n kutsutaan lukumäärätiheydeksi ja sen yksikkö on m^{-3} . Oletetaan vielä, että kaikki varaukset liikkuvat samalla nopeudella v . Varauksien nopeuden suuruus on v .

Ajassa dt jokainen varaus etenee matkan vdt . Kuvitellaan, että johteessa sijaitsee sylinteri, jonka pohjan pinta-ala on A ja korkeus vdt kuvan 2.6 osoittamalla tavalla. Vain ne varaukset, jotka jo ajanjakson dt alussa sijaitsevat sylinterissä, voivat ajan dt kuluessa virrata ulos sylinteristä. Lukumäärätiheyden avulla ilmoitettuna, sylinterin tilavuus sisältää $nAvdt$ varausta. Ajassa dt alueen A läpi virtaava kokonaisvaraus on

$$dQ = qnAvdt,$$

jolloin sähkövirraksi saadaan

$$I = \frac{dQ}{dt} = n|q|vA. \quad (2.11)$$

Varauksesta q voidaan ottaa itseisarvo, koska jo aikaisemmin todettiin, että varausten merkki ei vaikuta virran suuntaan.

Sähkövirtaa I yksikköpinta-alan läpi sanotaan virrantiheydeksi J , joka edellä käytetyn poikkileikkauksen A ja yhtälön (2.11) avulla kirjoitettuna on

$$J = \frac{I}{A} = n|q|v. \quad (2.12)$$

Virrantiheyden yksikkö on A/m². Vektorisuureena virrantiheys määritellään yhtälöllä

$$\mathbf{J} = nq\mathbf{v}, \quad (2.13)$$

jolloin yhtälö (2.12) ilmaisee vektoriarvoisen virrantiheyden suuruuden.

Sähkövirta I mielivaltaisen pinnan S läpi voidaan kirjoittaa virrantiheyden \mathbf{J} avulla [11, sivu 113]

$$I = \int_S \mathbf{J} \cdot \mathbf{n} \, dS. \quad (2.14)$$

2.3.2 Jatkuvuusyhtälö

Virrantiheys ja sähkövaraus liittyvät läheisesti toisiinsa. Sovelletaan yhtälöön (2.14) divergenssiteoreemaa (2.8). Tällöin suljetun pinnan S läpi tilavuuteen V tuleva virta on

$$I = - \oint_S \mathbf{J} \cdot \mathbf{n} \, dS = - \int_V \nabla \cdot \mathbf{J} \, dV, \quad (2.15)$$

missä \mathbf{n} on tilavuuden V ulkonormaalivektori. Sähkövirta I voidaan ilmoittaa myös kokonaisvarauksen Q muutoksena ajan suhteen. Kun kokonaisvaraus määritellään varausjakauman ρ avulla

$$Q = \int_V \rho \, dV,$$

niin tilavuuteen V tulevan sähkövirran tulee toteuttaa myös yhtälö

$$I = \frac{dQ}{dt} = \frac{d}{dt} \int_V \rho \, dV. \quad (2.16)$$

Oletetaan tilavuus V kiinteäksi, jolloin yhtälön (2.16) aikaderivaatta voidaan viedä integraalin sisään. Koska ρ on sekä ajan että paikan funktio, siihen kohdistuva kokonaisderivaatta muuttuu osittaisderivaataksi

$$I = \int_V \frac{\partial \rho}{\partial t} \, dV. \quad (2.17)$$

Tilavuuteen V tulevan sähkövirran tulee toteuttaa sekä yhtälö (2.15) että (2.17), jolloin

$$\int_V \frac{\partial \rho}{\partial t} \, dV = - \int_V \nabla \cdot \mathbf{J} \, dV.$$

Koska yhtäsuuruuden täytyy päteä kaikille tilavuuksille, saadaan virralle *jatkuvuusyhtälö*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (2.18)$$

Jos varaustiheys on ajasta riippumaton $\partial \rho / \partial t = 0$, niin siitä seuraa, että sähkövirta on lähteetön $\nabla \cdot \mathbf{J} = 0$. Tällaista virtausta kutsutaan *stationaariseksi*.

2.4 Staattinen magneettikenttä

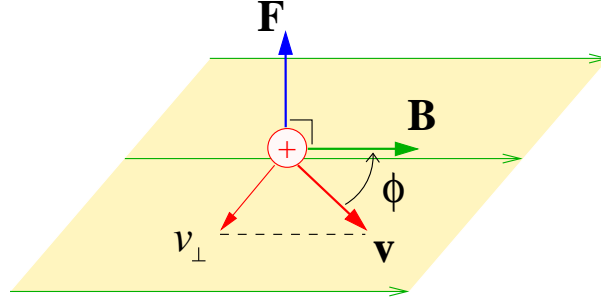
Aikaisemmin on todettu, että paikallaan olevien varausten välillä vaikuttaa sähköisiä voimia, jotka aiheutuvat varausten muodostamista sähkökentistä. Sama pätee myös liikkuville varauksille. Sähköisten voimien lisäksi liikkuvien varausten välillä on myös magneettista vuorovaikutusta, joka havaitaan varauksiin vaikuttavina magneettisina voimina. Ensimmäiset todisteet magnetismin liittymisestä liikkuviin varauksiin löysi *Ørsted* vuonna 1819. Sähköiset ja magneettiset vuorovaikutukset liittyvät tiiviisti toisiinsa ja yhdessä ne muodostavat sähkömagnetismin käsitteen.

Magneettinen vuorovaikutus ajatellaan samalla tavalla kaksivaiheiseksi kuin sähköinen vuorovaikutus. Liikkuva varaus tai sähkövirta, joka on siis varausten liikettä johteessa, aiheuttaa ympäröivään avaruuteen sähkökentän lisäksi myös *magneettikentän* \mathbf{B} . Muodostunut magneettikenttä vaikuttaa voimalla \mathbf{F} muihin kentässä liikkuviin varauksiin ja kentässä sijaitseviin sähkövirtoihin. Edellä kuvattu vuorovaikutus selittää myös magnetismin ehkä tunnetuimmat ilmentymät, magneettiset kappaleet. Myös niissä on pohjimmiltaan kyse liikkuvien varauksien muodostamasta magneettikentästä. Magneettisen kappaleen materiaalirakenne mahdollistaa varausten eli vapaiden elektronien järjestelmällisen liikkumisen kappaleen sisällä niin, että ympäröivään avaruuteen muodostuu magneettikenttä.

2.4.1 Magneettiset voimat

Tutkitaan aluksi, minkälainen voimavaikutus magneettikentällä \mathbf{B} on varauksiin ja virtauksiin. Magneettinen voima \mathbf{F} , joka kohdistuu liikkuvaan varaukseen q , on verrannollinen varauksen nopeuteen \mathbf{v} . Magneettiset voimat eivät vaikuta paikallaan oleviin varauksiin. Sähkökentän aiheuttamasta voimasta poiketen, magneettinen voima \mathbf{F} ei ole magneettikentän \mathbf{B} suuntainen, vaan se on aina kohtisuorassa kenttää \mathbf{B} ja varauksen nopeutta \mathbf{v} vastaan. Voiman suuruus F on verrannollinen nopeuden \mathbf{v} kenttää \mathbf{B} vastaan kohtisuorassa olevaan komponenttiin v_{\perp} . Kun varauksen nopeus on magneettikentän suuntainen, niin $v_{\perp} = 0$ eikä liikkuvaan varaukseen silloin kohdistu magneettisia voimia.

Magneettikentän aiheuttama voima \mathbf{F} on siis aina kohtisuorassa sitä pintaa vastaan,



Kuva 2.7: Magneettikentässä liikkuvaan varaukseen vaikuttava magneettinen voima.

joka sisältää sekä kentän \mathbf{B} että nopeuden \mathbf{v} . Voiman suuruus on

$$F = |q|v_{\perp}B = |q|vB\sin\phi,$$

missä nopeuden \mathbf{v} ja kentän \mathbf{B} välinen kulma ϕ määräytyy kuvan 2.7 esittämällä tavalla. Magneettikentän B yksikkö SI-järjestelmässä on tesla ($T = \text{Ns}/\text{Cm}$). Koska yksi ampeeri A on 1 C/s , niin yksi tesla ampeerin avulla kirjoitettuna on $1 \text{ T} = 1 \text{ N/Am}$.

Voimavektori \mathbf{F} ei ole vielä yksikäsitteinen, koska aina on olemassa kaksi vastakkaista suuntaa, jotka ovat kohtisuorassa pintaa vastaan. Esimerkiksi kuvassa 2.7 myös vektori $-\mathbf{F}$ täyttää kohtisuoruusehdon. Voiman positiivinen suunta määritellään oikean käden sääntöä käyttäen, jolloin voima \mathbf{F} on

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}. \quad (2.19)$$

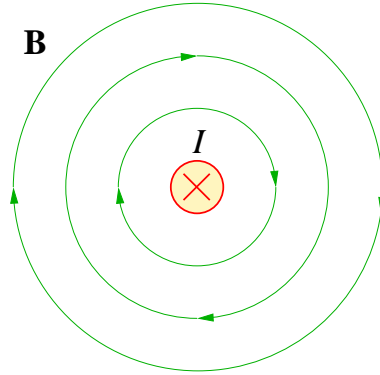
Kun varaus q liikkuu alueessa, jossa vaikuttaa sekä sähkö- \mathbf{E} että magneettikenttä \mathbf{B} , niin varaukseen kohdistuva kokonaisvoima \mathbf{F} lasketaan yhtälöiden (2.4) ja (2.19) summana

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (2.20)$$

Kokonaisvoimaa kutsutaan *Lorentzin voimaksi*.

2.4.2 Magneettikentän vuo

Luvussa 2.2.1 määriteltiin sähkökentän vuo suljetun pinnan yli. Nyt määritellään magneettikentän vuo Φ_B samalla tavalla. Aivan kuten sähkökenttää, magneettikenttää \mathbf{B} voidaan kuvata kenttäviivoilla, jotka ovat jokaisessa pisteessään kentän suuntaisia. Sähkökenttää kuvattaessa varaukset toimivat ikään kuin kentän lähteinä ja nieluina, jolloin varauksista tulee niiden merkistä riippuen kenttäviivojen lähtö- tai päätepisteitä. Kun magneettikenttää kuvataan kenttäviivoilla, viivat eivät lähde eivätkä pääty mihinkään, vaan ne muodostavat aina suljettuja silmukoita. Tämä on seurausta siitä, ettei erillistä magneettista varausta ole olemassa. Kuva 2.8 havainnollistaa, miten magneettikenttää kuvaavat kenttäviivat muodostavat suljettuja silmukoita.



Kuva 2.8: Johteessa kulkevan sähkövirran ympäristöön aiheuttama magneettikenttä.

Lasketaan magneettikentän vuo Φ_B tilavuutta V rajoittavalla suljetulla pinnalla S . Vuo saadaan laskemalla yhteen magneettikentän voimakkuus B tarkasteltavan pinnan ulkoisen normaalivektorin \mathbf{n} suuntaan jokaisessa pinnan pisteessä. Jaetaan pinta pieniin osa-alueisiin dS , jolloin yhden osa-alueen yli laskettu magneettikentän vuo on

$$d\Phi_B = B_{\perp} dS = B \cos\phi dS = \mathbf{B} \cdot \mathbf{n} dS.$$

Magneettivuon yksikkö SI-järjestelmässä on weber ($\text{Wb} = \text{Tm}^2$).

Täsmälleen samalla tavalla kuin sähkökentän vuota laskettaessa, summaamalla yhteen kaikkien osa-alueiden dS vuot, saadaan magneettikentän \mathbf{B} kokonaisvuo suljetun pinnan S yli

$$\Phi_B = \int_S B_{\perp} dS = \int_S \mathbf{B} \cdot \mathbf{n} dS. \quad (2.21)$$

Aikaisemmin todettiin, että magneettikentät ovat lähteettömiä. Johtopäätös on, että jokaisessa avaruuden pisteessä magneettikentän kokonaisvuon täytyy olla nolla. Tilanne on helppo kuvitella ajattelemalla magneettikenttää kuvaavien kenttäviivojen muodostamia suljettuja silmukoita: jokaiseen pisteeseen saapuu ja lähtee kenttäviiva. Seurausena on, että magneettivuo minkä hyvänsä pinnan yli on nolla

$$\oint_S \mathbf{B} \cdot \mathbf{n} dS = 0. \quad (2.22)$$

Yhtälö (2.22) on *Gaussin laki magnetismille* integraalimuodossa. Kun sen vasemman puoleen sovelletaan divergenssiteoreemaa (2.8) ja todetaan, että uuden yhtälön täytyy olla tilavuudesta V riippumaton, niin laki saadaan differentiaalimuotoon

$$\nabla \cdot \mathbf{B} = 0. \quad (2.23)$$

Gaussin laki magnetismille on Maxwellin yhtälöistä toinen ja se määrittelee, että magneettikentät ovat lähteettömiä.

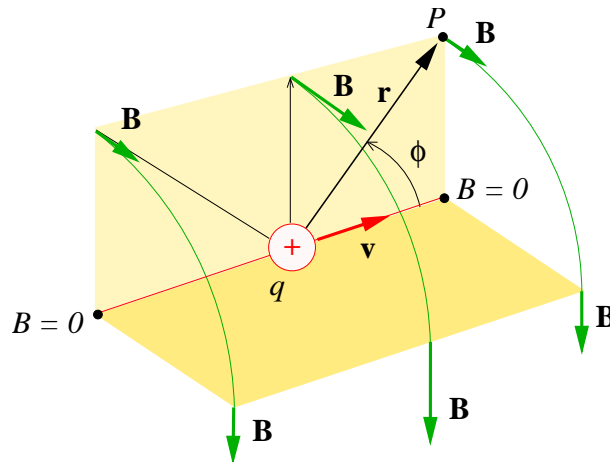
2.5 Magneettikentän alkuperä

Aikaisemmissa luvuissa on tarkasteltu, miten magneettikenttä vaikuttaa liikkuviin varauksiin ja sähkövirtoihin. Nyt tutkitaan, miten liikkuvat varaukset muodostavat magneettikentän. Olkoon avaruuden pisteessä varaus q , joka liikkuessaan vakionopeudella \mathbf{v} aiheuttaa magneettikentän \mathbf{B} . Tarkastellaan muodostuvaa magneettikenttää pisteessä P . Määritellään vektori \mathbf{r} , joka osoittaa varauksesta q tarkastelupisteeseen P ja jonka pituus on pisteiden välinen etäisyys r .

Toisin kuin sähkökenttä, magneettikenttä ei ole vektorin \mathbf{r} suuntainen, vaan se on kuvan 2.9 esittämällä tavalla kohtisuorassa vektorit \mathbf{r} ja \mathbf{v} sisältävää pintaa vastaan. On havaittu, että magneettikentän suuruus B on käänteisesti verrannollinen etäisyyden r neliöön. Lisäksi on havaittu, että magneettikentän suuruus on verrannollinen varauksen q nopeuteen ja vektoreiden \mathbf{v} ja \mathbf{r} väliseen kulmaan ϕ seuraavalla tavalla

$$B = \frac{\mu_0}{4\pi} \frac{|q|v \sin\phi}{r^2}. \quad (2.24)$$

Vakio μ_0 on tyhjiön permeabiliteetti ja sen numeerinen arvo SI-järjestelmän yksiköissä on $4\pi \cdot 10^{-7} \text{Tm/A}$.



Kuva 2.9: Liikkuvan varauksen ympäristöönsä aiheuttama magneettikenttä.

Magneettikentän \mathbf{B} positiivinen suunta määräytyy vektoreiden \mathbf{v} ja \mathbf{r} ristitulosta, joka toteuttaa oikean käden säännön. Magneettikenttä vektorisuurena on yksikkövektorin $\bar{\mathbf{r}} = \mathbf{r}/r$ avulla kirjoitettuna

$$\mathbf{B} = \frac{\mu_0}{4\pi} \frac{q\mathbf{v} \times \bar{\mathbf{r}}}{r^2}. \quad (2.25)$$

Aivan kuten sähkökentille, niin myös magneettikentille on olemassa superpositioperiaate. Useiden liikkuvien varausten muodostama kokonaismagneettikenttä on summa yksittäisten varausten aiheuttamista kentistä.

2.5.1 Sähkövirran aiheuttama magneettikenttä

Johteessa kulkevan sähkövirran ympäristöönsä aiheuttama magneettikenttä on sekä teorian että käytännön kannalta tärkeä erikoistapaus. Tarkastellaan aluksi, minkälaisen magneettikentän johteen lyhyt elementti $d\mathbf{l}$ aiheuttaa ympäristöönsä. Elementtiä kuvaava vektori $d\mathbf{l} = I d\mathbf{l}$ kertoo elementin suunnan ja pituuden. Olkoon elementin poikkileikkauspinta-ala A , jolloin elementin tilavuus on Adl . Oletetaan, että johteen yksikkötilavuudessa on n liikkuvaa hiukkasta ja että jokaisen hiukkasen varaus on q . Tällöin elementin $d\mathbf{l}$ sisältämien liikkuvien hiukkasten kokonaisvaraus on

$$Q = nqAdl.$$

Ajatellaan kokonaisvaraus Q yhtenä yksittäisenä varauksena, joka etenee johteessa nopeudella \mathbf{v} . Nyt voidaan suoraan soveltaa edellisen luvun päätelmiä ja tuloksia. Muistetaan vielä yhtälön (2.11) määritelmä sähkövirralle $I = n|q|vA$. Elementin $d\mathbf{l}$ aiheuttaman magneettikentän voimakkuus tarkastelupisteessä P on

$$dB = \frac{\mu_0}{4\pi} \frac{|Q|v \sin\phi}{r^2} = \frac{\mu_0}{4\pi} \frac{Idl \sin\phi}{r^2}. \quad (2.26)$$

Yksikkövektorin $\bar{\mathbf{r}}$ avulla kirjoitettuna elementin $d\mathbf{l}$ aiheuttama magneettikenttä on

$$d\mathbf{B} = \frac{\mu_0}{4\pi} \frac{Id\mathbf{l} \times \bar{\mathbf{r}}}{r^2}. \quad (2.27)$$

Johteen muodostama kokonaismagneettikenttä \mathbf{B} koostuu yksittäisten elementtien muodostamista magneettikentistä $d\mathbf{B}$. Integroimalla yhtälö (2.27) kaikkien elementtien $d\mathbf{l}$ yli, johteen sähkövirta aiheuttaa ympäristöönsä kokonaismagneettikentän

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int \frac{Id\mathbf{l} \times \bar{\mathbf{r}}}{r^2}. \quad (2.28)$$

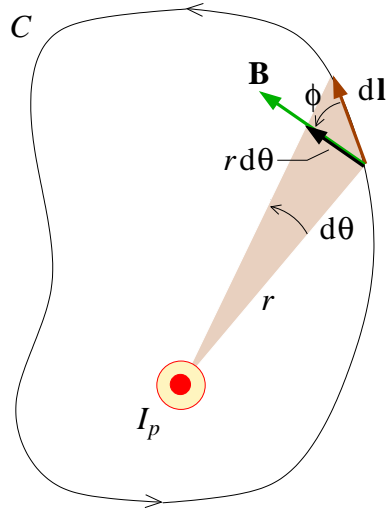
Saadusta yhtälöstä käytetään nimeä *Biot'n ja Savartin laki*.

2.5.2 Ampéren laki

Tässä luvussa tullaan käyttämään esimerkkitapauksena suorassa ja erittäin pitkässä johteessa kulkevan sähkövirran I_p ympäristöönsä aiheuttamaa magneettikenttää. Tässä luvussa sähkövirran alaindeksillä p korostetaan sitä, että kyseessä on yksittäisen johteen kuljettama sähkövirta. Soveltamalla Biot'n ja Savartin lakia, sekä käyttämällä hyväksi tilanteen geometrista symmetrisyyttä, pitkässä ja suorassa johteessa kulkevan sähkövirran I_p aiheuttaman magneettikentän voimakkuus on

$$B = \frac{\mu_0 I_p}{2\pi r}, \quad (2.29)$$

kun r on tarkastelupisteen etäisyys johteesta.



Kuva 2.10: Eräs integrointipolku magneettikentän viivaintegraalin laskemiseen.

Seuraavaksi on tarkoituksena laskea magneettikentän \mathbf{B} viivaintegraali mielivaltaista suljettua polkua pitkin. Viivaintegraalilla tarkoitetaan, että magneettikentän voimakkuus polun suuntaan lasketaan jokaisessa polun pisteessä ja lopuksi saadut tulokset summataan. Olkoon C suljettu mielivaltainen polku. Määritellään vektori $d\mathbf{l}$ jokaisessa polun pisteessä siten, että se on polun suuntainen ja pituudeltaan dl . Tällöin magneettikentän \mathbf{B} viivaintegraali polun C yli on

$$\oint_C \mathbf{B} \cdot d\mathbf{l}.$$

Kuvassa 2.10 mielivaltaisen suljetun polun C sisäpuolisen alueen läpi kulkee erittäin pitkä johde, jota on merkitty pisteen sisältävällä ympyrällä. Johde kuljettaa sähkövirtaa I_p kohtisuoraan katselijan suuntaan. Vektoreiden \mathbf{B} ja $d\mathbf{l}$ välinen kulma on ϕ , jolloin vektoreiden väliselle pistetulolle saadaan yhtäsuuruus $\mathbf{B} \cdot d\mathbf{l} = B dl \cos\phi$. Lisäksi kuvasta 2.10 on pääteltävissä, että $dl \cos\phi = r d\theta$. Todetaan vielä, että suljettu polku voidaan integroida myös kulman θ suhteen ja että seuraava integraali antaa tuloksena täyskulman

$$\oint d\theta = 2\pi.$$

Yhtälön (2.29) avulla kirjoitettuna magneettikentän \mathbf{B} viivaintegraali pitkin mielivaltaista suljettua polkua C on

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \oint_C \frac{\mu_0 I_p}{2\pi r} (r d\theta) = \frac{\mu_0 I_p}{2\pi} \oint_C d\theta = \mu_0 I_p. \quad (2.30)$$

Yhtälö (2.30) pätee myös yleisemmille tapauksille. Mikäli silmukan C rajoittaman alueen läpäisee useita sähkövirtaa kuljettavia johteita, niin magneettikentän \mathbf{B} viivaintegraali polkua C pitkin laskettuna on sähkövirtojen algebrallinen summa I kerrottuna

vakiolla μ_0 . On myös mahdollista osoittaa, että silmukan läpäisevät johteet voivat olla sekä mielivaltaisen muotoisia että suuntaisia. Kirjoitetaan vielä yleistetty yhtälö

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 I, \quad (2.31)$$

jota kutsutaan *Ampèren laiksi*. Laki on Gaussin lain (2.7) vastine magnetismille ja se määrittelee, miten magneettikenttä \mathbf{B} voidaan laskea tunnetun virtajakauman I avulla. Laki pätee ainoastaan staattisille sähkövirroille.

Ampèren laki saadaan differentiaalimuotoon soveltamalla *Stokesin lausetta*: riittävän siistille vektorikentälle \mathbf{u} pätee

$$\oint_C \mathbf{u} \cdot d\mathbf{l} = \int_S \nabla \times \mathbf{u} \cdot \mathbf{n} \, dS, \quad (2.32)$$

missä \mathbf{n} on silmukan C rajoittaman pinnan S normaalivektori. Soveltamalla Stokesin lausetta yhtälön (2.31) vasempaan puoleen ja yhtälöä (2.14) oikeaan puoleen, saadaan

$$\int_S \nabla \times \mathbf{B} \cdot \mathbf{n} \, dS = \mu_0 \int_S \mathbf{J} \cdot \mathbf{n} \, dS. \quad (2.33)$$

Koska yhtälön (2.33) täytyy päteä kaikille pinnoille S , yhtälö supistuu muotoon

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (2.34)$$

joka on Ampèren laki *differentiaalimuodossa*.

2.6 Dynaaminen sähkö- ja magneettikenttä

Tähän saakka on tarkasteltu ainoastaan ajasta riippumattomia kenttiä. Tässä luvussa siirrytään tutkimaan dynaamisten eli ajasta riippuvien sähkö- ja magneettikenttien vuorovaikutusta. Ensin perehdytään niin sanottuun *sähkömagneettiseen induktioon*, jonka jälkeen muokataan Ampèren lakia niin, että se pätee myös ajasta riippuville sähkövirtauksille.

2.6.1 Faradayn laki

Luvussa 2.5.1 todettiin, että sähkövirrat aiheuttavat magneettikenttiä. Ilmiö herätti 1800-luvun alussa seuraavaan kysymyksen: jos kerran sähkövirrat aiheuttavat magneettikenttiä, niin eivätkö magneettikentät puolestaan voisi aiheuttaa sähkövirtoja? *Michael Faradayn* 1830-luvulla suorittamat kokeet johtivat kuitenkin siihen tulokseen, että stationaariset magneettikentät eivät sinänsä aiheuta virtoja, vaan siihen tarvitaan magneettivuo, joka muuttuu ajallisesti.

Sähkövirtojen olemassaolo edellyttää niin sanottua *sähkömotorista voimaa* ε , joka saa varaukset liikkumaan. Olkoon C mielivaltainen polku, joka muodostaa suljetun silmukan. On osoitettu, että staattiselle sähkökentälle pätee

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0.$$

Mikäli kenttä ei ole staattinen, yllä esitetty integraali ei ole nolla ja silmukkaan C sanotaan indusoituvan sähkömotorisen voiman

$$\varepsilon = \oint_C \mathbf{E} \cdot d\mathbf{l}. \quad (2.35)$$

Faradayn tekemien havaintojen mukaan, silmukkaan C indusoituu sähkömotorinen voima silloin, kun silmukan rajoittamalla pinnalla S magneettivuon Φ_B muuttuu ajallisesti

$$\varepsilon = -\frac{d\Phi_B}{dt}. \quad (2.36)$$

Yhtälöistä (2.35) ja (2.36) saadaan *Faradayn induktiolaki*

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}, \quad (2.37)$$

joka on Maxwellin kolmas yhtälö integraalimuodossa. Kyseessä on kokeellinen luonnonlaki, joka ei seuraa mistään muista luonnonlaeista. Lain olemassaolo ei riipu fyysisen silmukan olemassaolosta, vaan se pätee annettua reittiä C pitkin lasketulle integraalille.

Seuraavaksi johdetaan Faradayn induktiolaki differentiaalimuodossa. Soveltamalla magneettivuon määritelmää (2.21) yhtälöön (2.37), saadaan

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{n} \, dS. \quad (2.38)$$

Ajatellaan silmukka C kiinteäksi, jolloin aikaderivaatta voidaan viedä integraalin sisälle. Magneettikenttään \mathbf{B} kohdistuva aikaderivaatta muuttuu osittaisderivaataksi, koska kenttä on sekä paikan että ajan funktio. Kun sovelletaan Stokesin lausetta (2.32) yhtälön (2.38) vasempaan puoleen, on voimassa

$$\int_S \nabla \times \mathbf{E} \cdot \mathbf{n} \, d\mathbf{l} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} \, dS. \quad (2.39)$$

Koska yhtälön (2.39) täytyy päteä kaikille kiinteille pinnoille, Faradayn laki differentiaalimuodossa on

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (2.40)$$

Faradayn laissa oleva miinusmerkki ilmaisee *Lenzin lain*: "Induktiovirta vastustaa muutosta, joka sen aiheuttaa." Yhtälö (2.40) ei sisällä enää mitään viittausta silmukkaan C , vaan sitä on pidettävä yleispätevästä relaationa kenttien \mathbf{E} ja \mathbf{B} välillä. Yhtälö määrittelee, että muuttuva magneettikenttä muodostaa sähkökentän.

2.6.2 Kenttävirta

Tässä luvussa Ampèren lakia laajennetaan niin, että se pätee myös ajan suhteen muuttuville sähkövirroille. Aikoinaan Maxwell huomasi sähkömagneettisia ilmiöitä kuvaavissa yhtälöissä teoreettisen ongelman. Oletetaan, että varaustiheys ja siten sähkökenttä muuttuvat ajallisesti. Otetaan yhtälöstä (2.34) divergenssi puolittain $\nabla \cdot (\nabla \times \mathbf{B}) = \mu_0 \nabla \cdot \mathbf{J}$ ja todetaan, että magneettikentän \mathbf{B} roottorin divergenssi on aina nolla, joten $\nabla \cdot \mathbf{J} = 0$. Virrantiheyden tulee toteuttaa myös jatkuvuusyhtälö (2.18). Kirjoittamalla molemmat vaatimukset samaan yhtälöön huomataan, että jatkuvuusyhtälö ja Ampèren laki ovat ristiriitaisia

$$0 = \nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \neq 0.$$

Tarkastellaan ongelmaa lähtien varauksen jatkuvuusyhtälöstä (2.18), joka Gaussin lain (2.9) avulla kirjoitettuna on

$$\nabla \cdot \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = 0. \quad (2.41)$$

Maxwellin oivallus oli korvata virrantiheys Ampèren laissa (2.34) edellä esitetyllä sulkausekkeella. Tuloksena oli neljäs Maxwellin yhtälöistä differentiaalimuodossa

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \quad (2.42)$$

jota voi hyvällä syyllä nimittää *Ampèren ja Maxwellin laiksi*. Termiä $\epsilon_0(\partial \mathbf{E}/\partial t)$ kutsutaan *kenttävirrantiheydeksi* ja siitä käytetään merkintää \mathbf{J}_D .

Laki saadaan integraalimuotoon, kun yhtälön (2.42) molempiin puoliin sovelletaan sopivaa pintaintegraalia

$$\int_S \nabla \times \mathbf{B} \cdot \mathbf{n} \, dS = \mu_0 \int_S \mathbf{J} \cdot \mathbf{n} \, dS + \mu_0 \epsilon_0 \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot \mathbf{n} \, dS.$$

Magneettikentän integraalilausekkeeseen sovelletaan Stokesin lausetta ja virrantiheyden integraaliin yhtälöä (2.14). Sähkökentän integraalilausekkeeseen sovelletaan ensin kenttävirrantiheyden merkintää ja sitten yhtälöä (2.14), jolloin saadaan neljäs Maxwellin yhtälö integraalimuodossa

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0(i + i_D). \quad (2.43)$$

Merkintä i_D tarkoittaa *kenttävirtaa*. Sähkövirtojen merkitseminen pienillä kirjaimilla korostaa, että virtaukset ovat ajasta riippuvia. Maxwellin neljäs yhtälö kertoo, että ajan suhteen muuttuva sähkökenttä \mathbf{E} aiheuttaa kuvitteellisen sähkövirran, jota nimitetään kenttävirraksi. Tällöin sekä todellinen sähkövirta i että kuvitteellinen kenttävirta i_D aiheuttavat ympäröivään avaruuteen magneettikentän \mathbf{B} .

2.7 Maxwellin yhtälöt

Edellisissä luvuissa on johdettu kaikki neljä Maxwellin yhtälöä, jotka jokaisen sähkö- ja magneettikentän tulee toteuttaa. Jatkuvuusyhtälö (2.18) seuraa Ampéren ja Maxwellin laista yhdessä Gaussin lain kanssa, joten sitä ei tarvitse ottaa mukaan erillisenä yhtälönä. Maxwellin neljä yhtälöä integraalimuodossa ovat

$$\left\{ \begin{array}{ll} \oint_S \mathbf{E} \cdot \mathbf{n} \, dS = \frac{1}{\epsilon_0} \int_V \rho \, dV & \text{Gaussin laki (2.7)} \\ \oint_S \mathbf{B} \cdot \mathbf{n} \, dS = 0 & \text{Gaussin laki magnetismille (2.22)} \\ \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt} & \text{Faradayn induktiolaki (2.37)} \\ \oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0(i + i_D) & \text{Ampéren ja Maxwellin laki (2.43).} \end{array} \right.$$

Yhtälöt on määritelty niin sanotussa *tyhjiömuodossa*. Se tarkoittaa, että varaukset ja sähkövirrat sijaitsevat tyhjiössä, joka ei sisällä muita kappaleita. Vaikka Maxwell ei omakätisesti laatinutkaan kaikkia neljää yhtälöä, ne kantavat hänen nimeään, koska Maxwell kokosi ne yhteen ja niiden avulla ennusti sähkömagneettisten aaltojen olemassaolon.

3 Matemaattinen malli

Maxwellin yhtälöt määrittelevät, että ajan suhteen muuttuva magneettikenttä \mathbf{B} aiheuttaa sähkökentän ja ajan suhteen muuttuva sähkökenttä \mathbf{E} aiheuttaa magneettikentän. Muuttuvat kentät \mathbf{E} ja \mathbf{B} pystyvät siis ikään kuin ylläpitämään toisiaan muodostaen *sähkömagneettisen aallon*, joka kykenee etenemään avaruuden osa-alueesta toiseen ilman minkäänlaista väliainetta. Kun sähkömagneettinen aalto etenee avaruudessa, se aiheuttaa kulkureitilleen muuttuvan sähkö- ja magneettikentän.

Tässä luvussa rakennetaan siroavalle sähkömagneettiselle aallolle matemaattinen malli Maxwellin yhtälöiden avulla. Ensin johdetaan aallon etenemistä kuvaava aaltoyhtälö, sitten määritellään aallon heijastuminen kappaleen pinnalta, jonka jälkeen asetetaan absorboiva reunaehto ja lopuksi kootaan malli. Tässä luvussa esitetyt asiat perustuvat kirjoihin [7], [11], [18] ja [30].

3.1 Aallon eteneminen

Kaikkien sähkö- ja magneettikenttien tulee siis toteuttaa Maxwellin yhtälöt. Aloitetaan aaltoyhtälön johtaminen kirjoittamalla Maxwellin yhtälöt differentiaalimuodossa [18, sivu 492]

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (3.1)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (3.2)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (3.3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (3.4)$$

Yhtälöt on esitetty tyhjiömuodossa. Kertauksen vuoksi mainitaan, että ρ kuvaa sähkövarauksia ja \mathbf{J} virrantiheyksiä, ϵ_0 on tyhjiön sähköinen permittiivisyys, μ_0 tyhjiön magneettinen permeabiliteetti ja c edustaa valonnopeutta. Permittiivisyyden, permeabiliteetin ja valonnopeuden välillä on yhteys $c^{-2} = \epsilon_0 \mu_0$.

Sähkö- ja magneettikenttiä tutkitaan tyhjiössä, joka ei sisällä sähkövarauksia tai sähkövirtoja. Oletuksista seuraa, että $\rho = 0$ ja $\mathbf{J} = 0$. Tällöin *Maxwellin* yhtälöt (3.1) ja (3.4) yksinkertaistuvat muotoon

$$\nabla \cdot \mathbf{E} = 0 \quad (3.5)$$

$$\nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (3.6)$$

Sovelletaan roottori-operaattoria $\nabla \times$ puolittain yhtälöön (3.3). Lisäksi tehdään oletus, että magneettikenttä \mathbf{B} on riittävän säännöllinen sekä paikan että ajan suhteen. Tällöin magneettikenttään \mathbf{B} kohdistuvan aikaderivaatan $\partial/\partial t$ ja roottorin $\nabla \times$ järjestystä voidaan vaihtaa [11, sivu 365]. Kun vielä käytetään yhtälöä (3.6), saadaan magneettikenttä eliminoitua. Tuloksena syntyy yhtälö, joka kuvaa sähkömagneettisen aallon sähkökenttää

$$\begin{aligned}\nabla \times (\nabla \times \mathbf{E}) &= -\nabla \times \left(\frac{\partial \mathbf{B}}{\partial t} \right) = -\frac{\partial (\nabla \times \mathbf{B})}{\partial t} = -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \\ \Leftrightarrow \nabla \times (\nabla \times \mathbf{E}) + \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} &= 0.\end{aligned}\quad (3.7)$$

Yllä esitettyä yhtälöä nimitetään jatkossa aaltoyhtälöksi. Todetaan, että vastaavanlainen yhtälö on mahdollista muodostaa myös magneettikentälle \mathbf{B} . Tällöin roottoria sovellettaisiin yhtälöön (3.6) ja sähkökenttä eliminoitaisiin käyttäen yhtälöä (3.3).

Sähkökenttä $\mathbf{E}(\mathbf{x}, t) : \mathbb{R}^3 \times \mathbb{R}^+ \rightarrow \mathbb{R}^3$ on riippuvainen ajasta ja paikasta. Vain yhdestä taajuudesta koostuvan aikaharmonisen aallon sähkökenttä voidaan esittää muodossa [30, sivu 27]

$$\mathbf{E}_c(\mathbf{x}, t) = \hat{\mathbf{E}}(\mathbf{x})e^{-i\omega t}, \quad (3.8)$$

missä i on imaginaariyksikkö, ω aallon kulmataajuus ja e niin sanottu *Neperin luku*. Sama pätee myös magneettikentälle \mathbf{B} . Yhtälöstä (3.8) seuraa, että sähkökentän aikaderivaatta $\partial/\partial t$ muuttuu kertomiseksi kompleksiluvulla $-i\omega$, koska funktio $\hat{\mathbf{E}}(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{C}^3$ ei ole aikariippuvainen.

Yhtälössä (3.8) kenttä \mathbf{E}_c on kompleksiarvoinen, mutta fysikaaliset kentät ovat reaaliarvoisia. Sähkökenttä saadaan kompleksiarvoisen kentän reaali-osana; kaikki laskutoimitukset voidaan suorittaa kompleksiarvoisina ja sähkökenttä saadaan ottamalla reaali-osa lopputuloksesta [7, sivu 216]

$$\mathbf{E}(\mathbf{x}, t) = \text{Re } \mathbf{E}_c(\mathbf{x}, t). \quad (3.9)$$

Jatkossa sähkömagneettista aaltoa tarkastellaan pelkästään sähkökentän osalta. Näin voidaan tehdä, koska yksitaajuisen aikaharmonisen aallon magneettikenttä voidaan ratkaista sähkökentän avulla. Todetaan asia sijoittamalla yhtälön (3.8) muotoa oleva magneettikenttä ja sähkökenttä yhtälöön (3.3), jolloin saadaan ratkaisuksi

$$\hat{\mathbf{B}} = \frac{1}{i\omega} \nabla \times \hat{\mathbf{E}}. \quad (3.10)$$

Lopullinen aaltoyhtälö, jolla kuvataan sähkömagneettisen aallon etenemistä, saadaan sijoittamalla yhtälön (3.8) oikea puoli yhtälöön (3.7). Saadusta yhtälöstä supistetaan termi $e^{-i\omega t}$. Tulevien laskutoimitusten helpottamiseksi huomioidaan vielä, että kulmataajuus määritellään yhtälöllä $\omega = 2\pi c/\lambda$, missä λ on aallonpituus [11, sivu 370]. Otetaan käyttöön uusi muuttuja α , jota kutsutaan *aaltoluvuksi* ja asetetaan

$\alpha := \omega/c = 2\pi/\lambda$. Edellä mainittujen toimenpiteiden ja sijoitusten jälkeen aaltoyhtälö on

$$\nabla \times (\nabla \times \hat{\mathbf{E}}) - \alpha^2 \hat{\mathbf{E}} = 0. \quad (3.11)$$

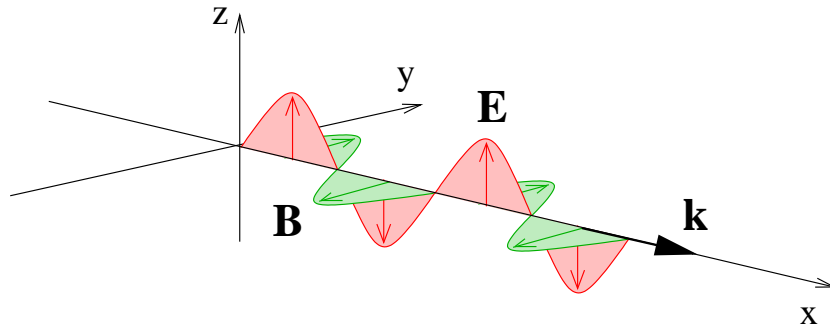
Osoitetaan vielä, että saadun aaltoyhtälön (3.11) ratkaisu $\hat{\mathbf{E}}$ toteuttaa yhtälön (3.5). Tiedetään, että riittävän sileän vektorikentän roottorin divergenssi on aina nolla [3, sivu 7]. Hyödynnetään edellä mainittua ominaisuutta siten, että sovelletaan divergenssioperaattoria $\nabla \cdot$ puolittain saatuun aaltoyhtälöön ja supistetaan termi α^2 . Lopputuloksena saadaan, että

$$\begin{aligned} \nabla \cdot \hat{\mathbf{E}} &= 0 \\ \Rightarrow \nabla \cdot \mathbf{E} &= 0, \end{aligned} \quad (3.12)$$

eli lopullisen aaltoyhtälön ratkaisulle $\hat{\mathbf{E}}$ pätee yhtälö (3.5) automaattisesti.

3.2 Aallon heijastuminen

Seuraavaksi määritellään reunaehto, joka mallintaa etenevän aallon heijastumista kappaleen pinnalta. Käytetään merkintää Ω_s osoittamaan heijastavaa kappaletta. Oletetaan kappaleen pinta Γ_s täydelliseksi johteeksi, jolloin kappaleen sisällä ei voi olla minäänlaista sähkökenttää [11, sivu 394].

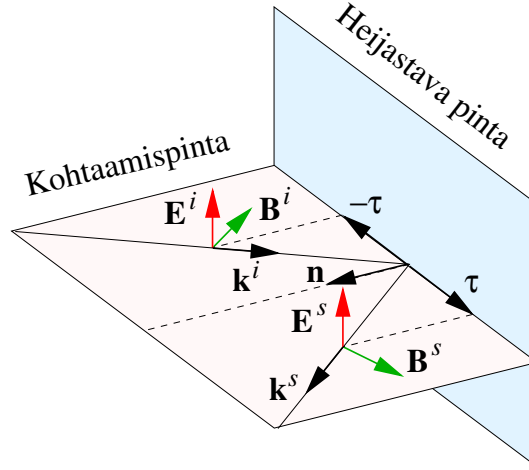


Kuva 3.1: Lineaarisesti polarisoitunut tasoaalto etenee x-akselin suuntaan.

Heijastavaa kappaletta lähestyvä aalto oletetaan lineaarisesti polarisoituneeksi tasoalloksi, jonka sähkökenttä pystytään kirjoittamaan matemaattisesti seuraavalla tavalla [30, sivu 31]

$$\mathbf{E}^i(\mathbf{x}, t) = \mathbf{u} e^{-i(\alpha \mathbf{k} \cdot \mathbf{x} + \omega t)}, \quad (3.13)$$

missä \mathbf{k} on tasoallon etenemissuuntaa kuvaava suuntavektori ja \mathbf{u} on tasoallon sähkökentän suuntavektori. Magneettikenttä \mathbf{B}_i voidaan ilmaista samalla tavalla, kunhan vektori \mathbf{u} korvataan tasoallon magneettikentän suuntavektorilla \mathbf{p} . Kaikki kolme suuntavektoria ovat yksikkövektoreita. Vektori \mathbf{u} on kohtisuorassa aallon etenemissuuntaa



Kuva 3.2: Lineaarisesti polarisoitunut tasoaalto heijastuu suoralta pinnalta.

\mathbf{k} vastaan. Lisäksi vektori \mathbf{p} on kohtisuorassa vektoreita \mathbf{u} ja \mathbf{k} vastaan siten, että $\mathbf{p} = \mathbf{u} \times \mathbf{k}$. Tällöin suuntavektorit muodostavat suorakulmaiset oikeakätiset koordinaattisuunnat [7, sivut 217-218]. Kuvassa 3.1 on esimerkki lineaarisesti polarisoituneesta tasoaallost.

Kappaleen pinnalla osa lähestyvistä tasoallost heijastuu takaisin ja osa läpäisee pinnan tunkeutuen kappaleeseen [7, sivu 223]. Käytetään heijastuvasta aallosta merkintää \mathbf{E}^s ja läpäisevästä \mathbf{E}^t . Kun osa aallosta tunkeutuu kappaleeseen, syntyy kappaleen sisälle sähkökenttä. Kuten aikaisemmin todettiin, tämä ei ole mahdollista täydellisen johteen tapauksessa ja siksi $\mathbf{E}^t \equiv 0$.

Kappaleen pinnalla sähkökentän tangentialikomponentin tulee olla jatkuva [11, sivut 391-404]. Jatkuvuusvaatimus määrittelee aallon heijastumisen kappaleen pinnalta. Jatkuvuus voidaan kirjoittaa pinnan pisteessä $\mathbf{x} \in \Gamma_s$ sijaitsevan mielivaltaisen yksikkötangenttivektorin \mathbf{t} avulla. Lisäksi tarvitaan superpositioperiaatetta, jonka mukaan pisteeseen muodostuva kokonaissähkökenttä voidaan laskea erillisten sähkökenttien summana. Kun vielä supistetaan termi $e^{-i\omega t}$, niin reunaehdoksi tulee

$$\begin{aligned} \mathbf{t} \cdot (\mathbf{E}^s(\mathbf{x}, t) + \mathbf{E}^i(\mathbf{x}, t)) &= 0 \\ \Rightarrow \mathbf{t} \cdot \hat{\mathbf{E}}^s(\mathbf{x}) &= -\mathbf{t} \cdot \mathbf{u} e^{-i\alpha \mathbf{k} \cdot \mathbf{x}}. \end{aligned} \quad (3.14)$$

Kuvassa 3.2 on yksinkertainen esimerkki siitä, miten lineaarisesti polarisoitunut tasoaalto heijastuu suoralta pinnalta.

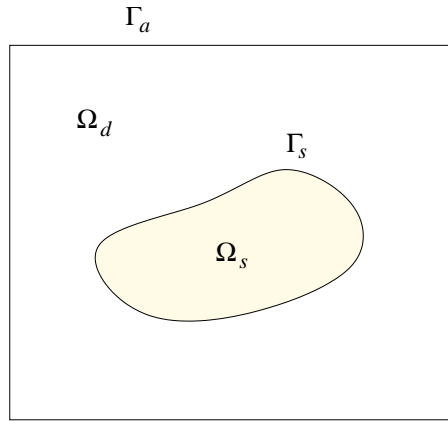
Kun tehtävää ratkaistaan, lähestyvä tasoaalto oletetaan tunnetuksi, eli suuntavektorien \mathbf{k} ja \mathbf{u} sekä aaltoluvun α arvot tiedetään. Edelleen heijastava kappale Ω_s tunnetaan ja silloin myös heijastavan pinnan yksikkötangenttivektorien \mathbf{t} virittämä tangenttitaso tiedetään. Näin ollen reunaehdossa (3.14) ainoa tuntematon on heijastuva aalto $\hat{\mathbf{E}}^s$. Tässä tutkielmassa simuloidaankin ainoastaan heijastunutta aaltoa, koska kokonais-

sähkökenttä \mathbf{E} voidaan superpositioperiaatteen mukaan laskea tunnetun lähestyvän taso-aallon ja ratkaistun heijastuneen aallon summana.

3.3 Absorboiva reunaehto

Jo aikaisemmin tehtiin oletukset, että heijastuva sähkömagneettinen aalto etenee tyhjiössä, joka ei sisällä sähkövarauksia tai sähkövirtoja. Tyhjiö ei sisällä myöskään muita esteitä heijastavan kappaleen Ω_s lisäksi, joka oletettiin täydelliseksi johteeksi. Kun aalto etenee avaruudessa edellä kuvatuissa olosuhteissa, niin energian säilymlakien nojalla aallon kokonaisenergia säilyy, mutta sen amplitudi pienenee [11, sivut 383-392].

Kun numeerinen simulointi suoritetaan tietokoneen avulla ja ratkaisun approksimointimenetelmänä on FEM (*Finite Element Method*), niin tietokoneen vaatimat resurssit ovat suoraan verrannollisia simulointialueen kokoon. Tietokoneen resurssien ollessa rajalliset, aallon koko etenemismatkaa ei voida simuloida, vaan simulointialuetta täytyy rajoittaa. Koska heijastavan kappaleen sisällä ei tarvitse laskea sähkökenttää, voidaan rajoitetuksi simulointialueeksi määritellä $\Omega = \Omega_d \setminus \Omega_s$, missä $\Omega_d \supset \Omega_s$ siis sisältää heijastavan kappaleen. Alueen Ω ulkoreunasta käytetään merkitään $\Gamma_a = \partial\Omega \setminus \partial\Gamma_s$, jota usein kutsutaan keinotekoiseksi reunaksi. Kuvassa 3.3 on esimerkki rajoitetusta simulointialueesta.



Kuva 3.3: Esimerkki rajoitetusta simulointialueesta kaksiulotteisessa tapauksessa.

Koska keinotekoisien reunan ei haluta vaikuttavan aallon etenemissuuntaan millään tavalla, niin reunalla Γ_a täytyy määritellä sellainen reunaehto, että heijastunut aalto pääsee ikään kuin vapaasti poistumaan alueesta. Artikkelissa [1] esitetty absorboiva reunaehto keinotekoisella reunalla on

$$(\mathbf{E} - c\mathbf{B}) \times \mathbf{n} \times \mathbf{n} = \mathbf{G} \times \mathbf{n}, \quad (3.15)$$

missä \mathbf{n} on reunan ulkoinen yksikkönormaalivektori, c on edelleen valonnopeus ja \mathbf{G}

edustaa simulointialueeseen saapuvaa aaltoa. Koska halutaan simuloida ainoastaan heijastuvaa aaltoa, voidaan asettaa $\mathbf{G} \equiv 0$. Yhtälö (3.15) on ensimmäisen kertaluokan absorboiva reunaehto, jota kutsutaan usein *Silver-Müller* -ehdoksi.

Muokataan reunaehto niin, että siirretään magneettikenttä \mathbf{B} yhtälön oikealle puolelle ja sovelletaan roottori-operaattoria $\nabla \times$ puolittain. Kun vielä eliminoidaan magneettikenttä käyttämällä yhtälöä (3.6), saadaan

$$(\nabla \times \mathbf{E}) \times \mathbf{n} = \frac{1}{c} \left(\frac{\partial \mathbf{E}}{\partial t} \right) \times \mathbf{n} \times \mathbf{n}. \quad (3.16)$$

Käytetään vielä yhtälöä (3.8), joka määrittelee yksitaajuisen aikaharmonisen aallon. Supistetaan jälleen termi $e^{-i\omega t}$ ja muistetaan, että simuloitavana on pelkästään heijastunut aalto. Tällöin lopullinen absorboiva reunaehto, jota tullaan jatkossa käyttämään keinotekoisella reunalla Γ_a , on

$$(\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} = -i\alpha \hat{\mathbf{E}}^s \times \mathbf{n} \times \mathbf{n}, \quad (3.17)$$

missä on käytetty myös aaltoluvun α määritelmää.

3.4 Mallin kokoaminen

Edellisissä luvuissa on määritelty yhtälöt aallon etenemiselle (3.11), heijastumiselle (3.14) ja absorboitumiselle (3.17). Lisäksi on voimassa vielä yhtälö (3.12). Koska halutaan simuloida pelkästään heijastunutta aaltoa, käytetään jälleen superpositioperiaatetta ja kirjoitetaan matemaattinen malli vielä yhtälöryhmänä

$$\nabla \times \nabla \times \hat{\mathbf{E}}^s(\mathbf{x}) - \alpha^2 \hat{\mathbf{E}}^s(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega \quad (3.18a)$$

$$\mathbf{t} \cdot \hat{\mathbf{E}}^s(\mathbf{x}) = -\mathbf{t} \cdot \mathbf{u} e^{-i\alpha \mathbf{k} \cdot \mathbf{x}} \quad \mathbf{x} \in \Gamma_s \quad (3.18b)$$

$$(\nabla \times \hat{\mathbf{E}}^s(\mathbf{x})) \times \mathbf{n} = -i\alpha \hat{\mathbf{E}}^s(\mathbf{x}) \times \mathbf{n} \times \mathbf{n} \quad \mathbf{x} \in \Gamma_a, \quad (3.18c)$$

joka heijastuneen sähkömagneettisen aallon tulee toteuttaa simulointialueessa.

Sähkömagneettiset aallot on oletettu yksitaajuisiksi ja aikaharmonisiksi. Simulointialue on oletettu tyhjiöksi, joka ei sisällä sähkövarauksia eikä sähkövirtoja. Lisäksi heijastavan kappaleen pinta on oletettu täydelliseksi johteeksi. Kappaleen pinnan kohtaava lähestyvä aalto on oletettu lineaarisesti polarisoituneeksi tasoaalloksi.

4 Heikko muoto

Matemaattinen malli määrittelee ilmiötä kuvaavan funktion epäsuorasti asettamalla sille joukon rajoituksia, jotka muodostavat yhtälöryhmän. Usein tilanne on se, että sopiva funktio on tuntematon ja se halutaan ratkaista. Näin syntyy tehtävä, jossa täytyy löytää sellainen ilmiötä kuvaava tuntematon funktio, että se toteuttaa yhtälöryhmän sisältämät rajoitukset. Jotta tehtävä olisi mielekäs, tulee rajoitukset toteuttavan funktion olla olemassa ja lisäksi sen tulee olla yksikäsitteinen.

Numeerisella simuloinnilla approksimoidaan matemaattisen mallin analyttistä ratkaisua, siis ilmiötä kuvaavaa funktiota. Sanotaan, että funktiolle laskettu approksimaatio on alkuperäisen tehtävän numeerinen ratkaisu. Tässä tutkielmassa numeerinen ratkaisu muodostetaan käyttämällä approksimointimenetelmää FEM. Menetelmän käyttö vaatii matemaattisen mallin muuntamista erityisellä tavalla; mallista täytyy laatia niin sanottu *heikko muoto* tai toiselta nimeltään *variaatioformulointi*.

Ennen heikon muodon laatimista esitellään joukko funktioavaruuksia, joita tullaan jatkossa tarvitsemaan. Varsinainen heikko muotoilu suoritetaan kaksiosaisena. Ensimmäin laaditaan matemaattista mallia vastaava integraaliyhtälö ja lopuksi integraaliyhtälöä käyttäen määritellään simulointitehtävälle heikko muotoilu homogeenisen Dirichlet'n reunaehdon avulla.

4.1 Funktioavaruudet

Tässä luvussa esitellään kootusti ne funktioavaruudet, joita tullaan tarvitsemaan heikon muodon laatimisessa. Luvun lopussa esitellään niin sanottu *Greenin osittaisintegroitikaava*, joka ei suoranaisesti liity funktioavaruuksiin, mutta joka on seuraavan luvun kannalta tärkeä matemaattinen työkalu. Heikossa muotoilussa tarvitaan funktioavaruuksia, jotka edellyttävät funktioiden määrittelyalueiden reunoilta riittävää säännöllisyyttä. *Lipschitz*-säännöllinen reuna on yleensä riittävä ominaisuus ja sen määrittely on [29, sivu 24]:

Määritelmä 4.1.1 *Olkoon $D \subset \mathbb{R}^d$ avoin ja rajoitettu joukko. Joukon D reuna ∂D on Lipschitz-reuna, jos on olemassa reunan ∂D avoin peite $\{U_i\}_{i \in I}$, $\epsilon > 0$, $N \in \mathbb{N}$ ja $C > 0$ siten, että*

- 1. Kaikilla $\mathbf{x} \in \partial D$ on olemassa $i_x \in I$ siten, että $B(\mathbf{x}, \epsilon) \subset U_{i_x}$.*
- 2. Kaikilla $i \in I$ leikkaus $U_i \cap U_j \neq \emptyset$ korkeintaan N kappaleella indeksejä $j \in I$.*

3. Kaikilla $i \in I$ pätee $U_i \cap D = U_i \cap D_i$, missä

$$D_i = \{(\mathbf{x}', x_d) \in \mathbb{R}^d \mid \mathbf{x}' \in \mathbb{R}^{d-1} \text{ ja } x_d < f_i(\mathbf{x}')\}.$$

Funktio $f_i : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$ on Lipschitz-jatkuva kuvaus, jonka Lipschitz-normille pätee

$$\|f_i\|_{Lip} \leq C.$$

Esitellään seuraavaksi ne funktioavaruudet, jotka määrittelevät tässä tutkielmassa tarvittavat funktiot. Olkoon $D \subset \mathbb{R}^d$ nyt sellainen avoin ja rajoitettu joukko, jonka reuna ∂D on Lipschitz-reuna. Joukosta D kuvautuvat jatkuvat, differentioituvat ja integroituvat funktiot määritellään funktioavaruuksilla [29, sivut 18-25], [8, sivu 499]

$$\begin{aligned} C(D) &= \{f : D \rightarrow \mathbb{K} \mid f \text{ on jatkuva}\}, \\ C^k(D) &= \{f : D \rightarrow \mathbb{K} \mid D^\beta f \in C(D), \forall |\beta| \leq k\}, \\ L^p(D) &= \{f : D \rightarrow \mathbb{K} \mid \int_D |f|^p d\mathbf{x} < \infty\}, \quad 1 \leq p < \infty, \\ L^p_{loc}(D) &= \{f : D \rightarrow \mathbb{K} \mid \text{funktioille } f \text{ integroituvuus} \\ &\quad \int_D |f|^p d\mathbf{x} < \infty \text{ pätee lokaalisti, kun } 1 \leq p < \infty\}, \\ H^k(D) &= \{f \in L^2(D) \mid D^\beta f \in L^2(D), \forall |\beta| \leq k\}, \end{aligned} \tag{4.1}$$

missä $D^\beta f$ on funktion f yleistetty derivaatta. Merkinnän vaihtoehtoinen nimitys on heikko derivaatta. Avaruuksien $L^p(D)$ ja $L^p_{loc}(D)$ määrittelyssä on käytetty *Lebesgue*-integraalia ja avaruudet $H^k(D)$ ovat *Sobolev*-avaruuksien osajoukko.

Jatkossa tarvitaan myös funktioita, jotka ovat neliöintegroituvia ja joiden roottori on neliöintegroituva. Lisäksi tarvitaan sellaisia funktioita, joiden avulla matemaattinen malli epähomogeenisella *Dirichlet*'n reunaehdolla (3.18b) voidaan ratkaista homogeenisen *Dirichlet*'n reunaehdon avulla. Edellä mainitut ominaisuudet löytyvät funktioista, jotka sisältyvät avaruuksiin [8, sivu 204]

$$\begin{aligned} H(\text{curl}, D) &= \{\mathbf{f} \in [L^2(D)]^3 \mid (\nabla \times \mathbf{f}) \in [L^2(D)]^3\}, \\ H_0(\text{curl}, D, \Gamma) &= \{\mathbf{f} \in H(\text{curl}, D) \mid \mathbf{n} \times \mathbf{f} = 0 \text{ reunalla } \Gamma \subseteq \partial D\}, \end{aligned} \tag{4.2}$$

missä \mathbf{n} on reunan pisteessä $\mathbf{x} \in \Gamma$ sijaitseva ulkoinen yksikkönormaalivektori. Molemmat kohdan (4.2) avaruuksista on määritelty, kun $D \subseteq \mathbb{R}^3$. Rajoitteen $\mathbf{n} \times \mathbf{f} = 0$ toteuttavat ne funktiot, joiden tangentialinen komponentti reunalla Γ on nolla. Täsmälleen yhtenevä funktiojoukko saadaan rajoitteella $\mathbf{t} \cdot \mathbf{f} = 0$, missä \mathbf{t} on reunan mielivaltainen yksikkötangenttivektori pisteessä $\mathbf{x} \in \Gamma$.

Tarkasti ottaen funktioavaruuden $H_0(\text{curl}, D, \Gamma)$ formaali määrittely edellyttäisi, että funktioille $\mathbf{v} \in H(\text{curl}, D)$ etsitään niin sanottu *jälki* reunalla Γ . Jälki määrittellään erityisen teoreeman (*Trace theorem*) avulla, jonka todistus löytyy esimerkiksi kirjasta [8, sivut 204-208]. Kuvainnollisesti sanottuna jälki ikään kuin jatkaa funktioiden $\mathbf{v} \in H(\text{curl}, D)$ määrittelyaluetta reunalle Γ . Avaruuden $H_0(\text{curl}, D, \Gamma)$ formaaliin määrittelyyn ei perehdytä tarkemmin tässä tutkielmassa.

Seuraavassa luvussa muokataan integraaliyhtälöitä Greenin osittaisintegrintikaavaa käyttäen. Greenin kaava määrittelee, miten integraalilausekkeen sisällä siirretään osittaisderivaatta funktiolta toiselle. Olkoot $w(\mathbf{x}) : D \rightarrow \mathbb{K}$ ja $q(\mathbf{x}) : D \rightarrow \mathbb{K}$ jatkuvasti differentioituvia funktioita avoimessa ja rajoitetussa alueessa $D \subset \mathbb{R}^d$. Funktioiden w ja q avulla kirjoitettuna Greenin kaava on [35, sivu 66]

$$\int_D \frac{\partial w}{\partial x_i} q \, d\mathbf{x} = - \int_D w \frac{\partial q}{\partial x_i} \, d\mathbf{x} + \int_{\partial D} w q n_i \, ds, \quad (4.3)$$

kun $1 \leq i \leq d$ ja $\mathbf{n} = (n_1, \dots, n_d)$ on reunan ∂D ulkoinen yksikkönormaalivektori. Lisäksi funktioiden w ja q oletetaan kuuluvan avaruuteen $H^1(D)$.

4.2 Integraaliyhtälö

Aloitetaan heikon muodon laatiminen asettamalla joukko oletuksia. Olkoon $\Omega \subset \mathbb{R}^3$ sellainen avoin ja rajoitettu joukko, jonka reuna $\partial\Omega$ on Lipschitz-reuna. Oletetaan, että niin sanottu *testifunktio* \mathbf{v} on äärettömän monta kertaa jatkuvasti differentioituva eli se kuuluu avaruuteen $[C^\infty(\Omega)]^3$. Oletetaan myös, että $\hat{\mathbf{E}}^s \in [C^2(\Omega)]^3$ eli sähkökenttä on kaksi kertaa jatkuvasti differentioituva. Lopuksi vaaditaan, että sähkökenttä $\hat{\mathbf{E}}^s$ ja sen derivaatat ovat äärellisiä, jolloin ne kuuluvat avaruuteen $[L^1(\Omega)]^3$. Määritellään seuraavaksi matemaattisen mallin *residuaali*

$$\mathbf{R}_\Omega = \nabla \times \nabla \times \hat{\mathbf{E}}^s(\mathbf{x}) - \alpha^2 \hat{\mathbf{E}}^s(\mathbf{x}). \quad (4.4)$$

Nyt on ilmeistä, että myös residuaali \mathbf{R}_Ω on äärellinen ja kuuluu avaruuteen $[L^1(\Omega)]^3$.

Seuraavaksi esitellään yksi variaatiolaskennan peruslauseista. Lokaalisti integroituvalla funktiolle $w \in L^1_{loc}(\Omega)$ pätee

$$\begin{aligned} \int_\Omega w v \, d\mathbf{x} &= 0 \quad \forall v \in C_0^\infty(\Omega) \\ \Rightarrow w &= 0 \quad m.k. \mathbf{x} \in \Omega, \end{aligned} \quad (4.5)$$

missä funktiota v sanotaan testifunktioksi ja avaruus $C_0^\infty(\Omega)$ on pienin joukko, jolle lause (4.5) on voimassa. Jatkossa lausetta sovelletaan testifunktioavaruudelle $C^\infty(\Omega)$, joka sisältää joukon $C_0^\infty(\Omega)$. Lisäksi todetaan, että lausetta voidaan soveltaa avaruuden $L^1_{loc}(\Omega)$ osajoukkoon $L^1(\Omega)$ kuuluville funktioille.

Sovelletaan lausetta (4.5) residuaaliin \mathbf{R}_Ω

$$\begin{aligned} \int_{\Omega} \mathbf{R}_\Omega \cdot \mathbf{v} \, d\mathbf{x} &= 0 \quad \forall \mathbf{v} \in \mathbf{C}^\infty(\Omega)^3 \\ \Leftrightarrow \int_{\Omega} \nabla \times \nabla \times \hat{\mathbf{E}}^s \cdot \mathbf{v} \, d\mathbf{x} - \alpha^2 \int_{\Omega} \hat{\mathbf{E}}^s \cdot \mathbf{v} \, d\mathbf{x} &= 0 \quad \forall \mathbf{v} \in \mathbf{C}^\infty(\Omega)^3, \end{aligned} \quad (4.6)$$

jolloin saatu integraaliyhtälö vastaa *heikossa mielessä* matemaattisen mallin yhtälöä (3.18a). Matemaattisen mallin ratkaisu $\hat{\mathbf{E}}^s$ on aina myös heikon muotoilun ratkaisu, mutta päinvastainen väittäminen ei välttämättä aina päde.

Heikon muotoilun (4.6) ensimmäinen integraalilauseke sisältää sähkökentän $\hat{\mathbf{E}}^s$, johon on sovellettu roottori-operaattoria kaksi kertaa. Käyttämällä Greenin kaavaa toinen roottoreista voidaan siirtää testifunktiolle \mathbf{v} . Siirron tarkoituksena on, että funktioiden $\hat{\mathbf{E}}^s$ ja \mathbf{v} jatkuvuusvaatimukset saadaan yhteneviksi. Siirron seurauksena funktion $\hat{\mathbf{E}}^s$ jatkuvuusvaatimukset integraaliyhtälössä ovat lievemmat kuin alkuperäisessä differentiaaliyhtälössä (3.18a) [23, sivu 25]. Roottorin siirtäminen on mielekästä myös siksi, että numeerisen ratkaisun laskeminen helpottuu [17, sivu 13], koska toisen asteen osittaisderivaattoja ei tarvitse enää laskea. Käytetään roottorin määritelmää ja kirjataan sähkökentän tuplaroottorille identiteetti

$$\begin{aligned} \nabla \times \nabla \times \hat{\mathbf{E}}^s &= \left(\frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_3}{\partial x_2} - \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_2}{\partial x_3}, \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_1}{\partial x_3} - \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_3}{\partial x_1}, \right. \\ &\quad \left. \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_2}{\partial x_1} - \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_1}{\partial x_2} \right). \end{aligned} \quad (4.7)$$

Sijoitetaan identiteetti yhtälön (4.6) ensimmäiseen integraalilausekkeeseen ja lasketaan sähkökenttään sovelletun tuplaroottorin ja testifunktion \mathbf{v} sisätulo

$$\begin{aligned} \int_{\Omega} \nabla \times \nabla \times \hat{\mathbf{E}}^s \cdot \mathbf{v} \, d\mathbf{x} &= \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_3}{\partial x_2} v_1 \, d\mathbf{x} - \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_2}{\partial x_3} v_1 \, d\mathbf{x} \\ &\quad + \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_1}{\partial x_3} v_2 \, d\mathbf{x} - \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_3}{\partial x_1} v_2 \, d\mathbf{x} \\ &\quad + \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_2}{\partial x_1} v_3 \, d\mathbf{x} - \int_{\Omega} \frac{\partial(\nabla \times \hat{\mathbf{E}}^s)_1}{\partial x_2} v_3 \, d\mathbf{x}. \end{aligned} \quad (4.8)$$

Saadussa yhtälössä on kuusi osittaiderivaatan sisältävää integraalilauseketta. Sovelletaan erikseen jokaiseen lausekkeeseen Greenin kaavaa (4.3). Kun vielä sopivasti ryhmitellään lausekkeitä, niin yhtälön (4.8) oikeaksi puoleksi tulee

$$\begin{aligned} \int_{\Omega} \left[(\nabla \times \hat{\mathbf{E}}^s)_1 \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \right) + (\nabla \times \hat{\mathbf{E}}^s)_2 \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \right) + (\nabla \times \hat{\mathbf{E}}^s)_3 \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) \right] d\mathbf{x} \\ + \int_{\partial\Omega} \left[((\nabla \times \hat{\mathbf{E}}^s)_3 n_2 - (\nabla \times \hat{\mathbf{E}}^s)_2 n_3) v_1 + ((\nabla \times \hat{\mathbf{E}}^s)_1 n_3 - (\nabla \times \hat{\mathbf{E}}^s)_3 n_1) v_2 + \right. \\ \left. + ((\nabla \times \hat{\mathbf{E}}^s)_2 n_1 - (\nabla \times \hat{\mathbf{E}}^s)_1 n_2) v_3 \right] ds. \end{aligned} \quad (4.9)$$

Havaitaan, että yhtälön alueintegraali pitää sisällään vektoreiden $\nabla \times \hat{\mathbf{E}}^s$ ja $\nabla \times \mathbf{v}$ sisätulon. Toinen alun alkaen sähkökenttään $\hat{\mathbf{E}}^s$ sovelletuista roottoreista on siis siirtynyt testifunktiolle \mathbf{v} . Kun kirjoitetaan sähkökentän roottorin $\nabla \times \hat{\mathbf{E}}^s$ ja ulkoisen yksikkönormaalivektorin \mathbf{n} ristitulo

$$\begin{aligned} (\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} &= ((\nabla \times \hat{\mathbf{E}}^s)_2 n_3 - (\nabla \times \hat{\mathbf{E}}^s)_3 n_2), (\nabla \times \hat{\mathbf{E}}^s)_3 n_1 - (\nabla \times \hat{\mathbf{E}}^s)_1 n_3, \\ &(\nabla \times \hat{\mathbf{E}}^s)_1 n_2 - (\nabla \times \hat{\mathbf{E}}^s)_2 n_1), \end{aligned}$$

niin on tehtävissä toinenkin havainto: yhtälön (4.9) reunaintegraali sisältää lausekkeen $-(\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n}$ ja testifunktion \mathbf{v} välisen sisätulon. Kokoamalla yhtälöiden (4.8) ja (4.9) tulokset sekä tehdyt havainnot, saadaan

$$\int_{\Omega} \nabla \times \nabla \times \hat{\mathbf{E}}^s \cdot \mathbf{v} \, dx = \int_{\Omega} \nabla \times \hat{\mathbf{E}}^s \cdot \nabla \times \mathbf{v} \, dx - \int_{\partial\Omega} (\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} \cdot \mathbf{v} \, ds. \quad (4.10)$$

Nyt on mahdollista lieventää sähkökenttään $\hat{\mathbf{E}}^s$ ja testifunktioon \mathbf{v} liittyviä oletuksia. Riittää, että sähkökenttä ja testifunktio sekä molempien roottorit ovat neliöintegroituja, toisin sanoen $\hat{\mathbf{E}}^s$ ja $\mathbf{v} \in H(\text{curl}, \Omega)$ [14, sivu 1327]. Valittu funktioavaruus takaa, että tehtävän toteuttavalla sähkökentällä on äärellinen energia. Sijoittamalla yhtäsuuruus (4.10) heikkoon muotoiluun (4.6) saadaan uusi integraaliyhtälö, joka funktion $\hat{\mathbf{E}}^s$ tulee toteuttaa heikossa mielessä

$$\begin{aligned} \int_{\Omega} \nabla \times \hat{\mathbf{E}}^s \cdot \nabla \times \mathbf{v} - \alpha^2 \hat{\mathbf{E}}^s \cdot \mathbf{v} \, dx - \int_{\partial\Omega} (\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} \cdot \mathbf{v} \, ds = 0 \\ \forall \mathbf{v} \in H(\text{curl}, \Omega). \end{aligned} \quad (4.11)$$

4.3 Simulointitehtävän heikko muotoilu

Heikko muotoilu (4.11) ei ole vielä käytännöllinen, koska reunaehtoien (3.18b) ja (3.18c) toteutuminen pitää käsitellä erikseen. Seuraavaksi muokataan heikkoa muotoilua niin, että reunaehdot sisältyvät yhtälöön (4.11). Tulevien merkintöjen selkeyttämiseksi esitellään apufunktio $\mathbf{g}(\mathbf{x}) : \Gamma_s \rightarrow \mathbb{C}^3$, joka vastaa reunaehto heijastavan kappaleen pinnalla

$$\mathbf{g}(\mathbf{x}) = -\mathbf{u}e^{-i\alpha\mathbf{k}\cdot\mathbf{x}}.$$

Valitaan sopiva funktio $G \in H(\text{curl}, \Omega)$, jolle pätee $\mathbf{t} \cdot G|_{\Gamma_s} = -\mathbf{t} \cdot \mathbf{g}$. Funktion G rajoittuma reunalle Γ_s toteuttaa siis täsmälleen sen epähomogeenisen reunaehdon (3.18b), jonka ratkaistavan funktion $\hat{\mathbf{E}}^s$ tulee toteuttaa. Määritellään funktio $\hat{\mathbf{E}}^s$ summana $\hat{\mathbf{E}}^s = \hat{\mathbf{E}}_0^s + G$, missä $\hat{\mathbf{E}}_0^s \in H_0(\text{curl}, \Omega, \Gamma_s)$. Funktio $\hat{\mathbf{E}}_0^s$ toteuttaa homogeenisen reunaehdon osareunalla Γ_s , koska se on valittu avaruudesta $H_0(\text{curl}, \Omega, \Gamma_s)$.

Summamääritelmän tarkoituksena on, että funktion $\hat{\mathbf{E}}^s$ sijasta voidaan ratkaista funktio $\hat{\mathbf{E}}_0^s$. Ratkaistavaksi funktioksi halutaan nimenomaan $\hat{\mathbf{E}}_0^s$, koska silloin myös testifunktiot voidaan valita avaruudesta $H_0(\text{curl}, \Omega, \Gamma_s)$. Valinnasta seuraa, että yhtälön

(4.11) reunaintegraali saadaan yksinkertaisempaan muotoon, kuten tullaan myöhemmin näkemään. Virhe, joka tehdään ratkaistaessa funktio ikään kuin väärillä reunaehdoilla, eliminoidaan käyttämällä sopivasti valittua tunnettua funktiota G .

Seuraavaksi kiinnitetään huomio yhtälön (4.11) reunaintegraaliin. Sähkökenttä halutaan ratkaista rajoitetussa alueessa Ω , jonka reuna $\partial\Omega = \Gamma_s \cup \Gamma_a$ koostuu kahdesta osasta. Osareunalla Γ_s tullaan hyödyntämään vektoreiden $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{K}^3$ skalaarikolmitulosääntöä $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ [13, Liite A]. Valitsemalla testifunktio \mathbf{v} avaruudesta $H_0(\text{curl}, \Omega, \Gamma_s)$ ja käyttämällä hyväksi juuri esitettyä skalaarikolmitulosääntöä, saadaan yhtälön (4.11) reunaintegraalilauseke nolaksi reunalla Γ_s

$$\begin{aligned} (\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} \cdot \mathbf{v} &= (\nabla \times \hat{\mathbf{E}}^s) \cdot (\mathbf{n} \times \mathbf{v}) = 0 \quad \mathbf{v} \in H_0(\text{curl}, \Omega, \Gamma_s) \\ \Rightarrow \int_{\Gamma_s} (\nabla \times \hat{\mathbf{E}}^s) \times \mathbf{n} \cdot \mathbf{v} \, ds &= 0 \quad \mathbf{v} \in H_0(\text{curl}, \Omega, \Gamma_s). \end{aligned} \quad (4.12)$$

Koska reunaintegraalilauseke on nolasta eroava ainoastaan osareunalla Γ_a , niin siihen voidaan soveltaa absorboivaa reunaehto (3.18c). Toistuvien merkintöjen vähentämiseksi esitellään niin sanottu *bilineaarimuoto*

$$\begin{aligned} a(\mathbf{w}, \mathbf{q}) &= \int_{\Omega} \nabla \times \mathbf{w} \cdot \nabla \times \mathbf{q} - \alpha^2 \mathbf{w} \cdot \mathbf{q} \, d\mathbf{x} + i\alpha \int_{\Gamma_a} \mathbf{w} \times \mathbf{n} \times \mathbf{n} \cdot \mathbf{q} \, ds \\ \mathbf{w} &\in H(\text{curl}, \Omega), \mathbf{q} \in H_0(\text{curl}, \Omega, \Gamma_s), \end{aligned} \quad (4.13)$$

joka on täsmälleen samaa muotoa kuin yhtälön (4.11) vasen puoli edellä esitettyjen reunaintegraalitulosten jälkeen. Koska bilineaarimuoto (4.13) on nimensä mukaisesti lineaarinen operaattori molempien komponenttiensa suhteen, saadaan funktion $\hat{\mathbf{E}}^s$ summamäärittelyä käyttämällä

$$a(\hat{\mathbf{E}}^s, \mathbf{v}) = a(\hat{\mathbf{E}}_0^s, \mathbf{v}) + a(\mathbf{G}, \mathbf{v}) \quad \mathbf{v} \in H_0(\text{curl}, \Omega, \Gamma_s). \quad (4.14)$$

Yhtälöiden (4.11), (4.13) ja (4.14) avulla pystytään viimein määrittelemään simulointitehtävän heikko muotoilu: etsi funktio $\hat{\mathbf{E}}_0^s \in H_0(\text{curl}, \Omega, \Gamma_s)$ siten, että

$$a(\hat{\mathbf{E}}_0^s, \mathbf{v}) = -a(\mathbf{G}, \mathbf{v}) \quad \forall \mathbf{v} \in H_0(\text{curl}, \Omega, \Gamma_s). \quad (4.15)$$

Lopuksi vielä todetaan, että tehtävä on hyvin asetettu, koska sille on olemassa yksikäsitteinen ratkaisu. Väite on todistettu ainakin artikkeleissa [14, sivut 1338-1339] ja [15, sivut 512-514]. Tehtävä on myös stabiili sähkökentän energiamäärän perusteella [5, sivut 22-24]. Heikon muotoilun yksikäsitteinen ratkaisu vastaa matemaattisen mallin ratkaisua.

5 Simulointitehtävän diskretisointi

Sähkömagneettisen aallon etenemistä ja heijastumista kuvaavan yhtälön (4.15) analyttinen ratkaiseminen on erittäin vaikea tehtävä. Kun tyydytään siihen, että tarkalle ratkaisulle $\hat{\mathbf{E}}_0^s$ lasketaankin vain riittävän hyvä arvio, niin tehtävä helpottuu huomattavasti ja laskennassa pystytään hyödyntämään tietokonetta. Arvion laskeminen edellyttää tehtävän (4.15) diskretisoimista, joka tässä tutkielmassa suoritetaan approksimointimenetelmän FEM avulla.

Äärellisten elementtien menetelmän taustaa ja historiaa esitellään kirjojen [17, sivu 1] ja [28, sivut 1-9] sisältämiin tietoihin nojautuen. Simulointitehtävän diskretisointi aloitetaan esittelemällä elementtimenetelmän peruskäsitteitä oppaaseen [12, sivut 242-272] ja kirjaan [34, sivut 178-191] perustuen. Simulointialueen diskretisointiin käytettävä elementtityyppi esitellään luvussa 5.3, esitys perustuu raporttiin [27]. Lopuksi muodostetaan diskretisoitua simulointitehtävää vastaava lineaarinen yhtälöryhmä. Yhtälöryhmää muodostettaessa lähteinä käytetään kirjoja [5], [17] ja [34].

5.1 Äärellisten elementtien menetelmä

Fyysikot ja insinöörit käyttävät yleisesti differentiaaliyhtälöitä tutkimiensa ilmiöiden mallintamiseen. Usein käytännön ilmiön mallintaminen vaatii niin monimutkaisia differentiaaliyhtälöitä, että niiden analyttinen ratkaiseminen on erittäin vaikeaa ja joskus jopa mahdotonta. Tällöin on tarpeen soveltaa jotakin numeerista ratkaisutapaa riittävän tarkan arvion laskemiseksi analyttiselle ratkaisulle. Yksi käytetyimmistä ratkaisutavoista on *Äärellisten elementtien menetelmä*, joka tunnetaan myös *Elementtimenetelmän* nimellä ja jonka englannin kielestä peritty lyhenne on *FEM* (*Finite Element Method*).

Elementtimenetelmän syntyhetkeä on vaikea määritellä täsmällisesti, mutta yksi ensimmäisistä elementtimenetelmän esittelijöistä oli *Courant* vuonna 1943. Tuolloin menetelmä sai vain vähän huomiota ja se ikään kuin keksittiin uudelleen 1950-luvun lentokoneiteollisuudessa. Henkilöiden *Turner*, *Clough*, *Martin* ja *Topp* vuonna 1956 esittelemä julkaisu innoitti muitakin tiedemiehiä tutkimaan menetelmän mahdollisuuksia. Aluksi menetelmää käytettiin lähinnä rakennesuunnittelussa. Elementtimenetelmän teoreettista taustaa ryhdyttiin tutkimaan toden teolla vasta 1960-luvulla. Tärkeä teoreettinen tulos saavutettiin vuonna 1963, kun *Melosh* osoitti, että elementtimenetelmä onkin itse asiassa muunnos hyvin tunnetusta *Raleigh-Ritz* -menetelmästä. Tulos auttoi tutkijoita tekemään merkittävän huomion: rakennetehtävissä FEM tuottaa joukon lineaarisia tasapainoyhtälöitä minimoimalla järjestelmän potentiaalienergian.

Kun elementtimenetelmän yhteys minimointiin huomattiin, sen käyttö levisi nopeasti eri insinöörialoille. Menetelmää ryhdyttiin soveltamaan erityisesti sellaisiin tehtäviin, joita kuvataan *Laplacen* tai *Poissonin* yhtälöillä, koska kyseisillä yhtälöillä on läheinen yhteys funktionaalien minimointiin. Ensimmäisiä sovelluskohteita olivat lämmönjohtumis- ja virtaustehtävät. Sovelluskohteiden määrä lisääntyi entisestään, kun 1960-luvun lopulla osoitettiin, että edellä mainittujen tehtävien ratkaisemiseen tarvittavat yhtälöt voidaan johtaa käyttämällä *painotettujen residuaalien* menetelmiä, joista esimerkkeinä mainittakoon *Galerkinin* ja *pienimmän neliösumman* menetelmät. Kyseessä oli jälleen erittäin merkittävä teoreettinen tulos, koska sen avulla elementtimenetelmää pystytään soveltamaan hyvin monenlaisiin differentiaaliyhtälöihin.

Nykypäivänä elementtimenetelmän käyttö on hyvin yleistä ja sen sovelluskohteisiin kuuluvat kaikki fysikaaliset ilmiöt, joita voidaan mallintaa differentiaaliyhtälöillä. Menetelmän käyttöä ovat edistäneet sen useat hyvät ominaisuudet:

1. Mallinnettavan kohteen ominaisuudet voivat vaihdella mallinnusalueen eri osissa. Tämä mahdollistaa esimerkiksi monesta eri materiaalista koostuvien kappaleiden mallintamisen.
2. Menetelmää voidaan käyttää myös silloin, kun mallinnusalueen geometria on monimutkainen.
3. Laskentatarkkuutta mallinnusalueen eri osissa voidaan säädellä, jolloin laskentaresursseja voidaan kohdistaa niihin osa-alueisiin, joissa ratkaisulle erityisesti halutaan mahdollisimman suuri tarkkuus.
4. Esimerkiksi epäjatkuvan pintakuormituksen tapaiset reunaehdot eivät aiheuta ongelmia ja menetelmää voidaankin soveltaa kirjavalle joukolle reunaehtoja.
5. Kaikki edellä mainitut ominaisuudet voidaan sisällyttää yleispätevään tietokoneohjelmaan, jonka avulla pystytään ratkaisemaan suuri joukko tietyn tyyppisiä tehtäviä, kuten esimerkiksi lämmönjohtumistehtäviä.

Elementtimenetelmän merkittävimpänä heikkoutena on pidetty sitä, että menetelmän soveltamiseen tarvitaan aina tietokonetta. Jopa yksinkertaisten tehtävien ratkaiseminen elementtimenetelmällä on lähes aina liian työlästä käsin suoritettavaksi. Nykypäivänä tietokoneet ovat niin yleisiä ja tehokkaita, että elementtimenetelmän riippuvuus tietokoneista ei ole enää merkittävä heikkous.

5.2 Elementtimenetelmän peruskäsitteitä

Perusajatus elementtimenetelmän taustalla on, että mitä hyvänsä jatkuvaa suuretta, kuten esimerkiksi lämpötilaa, painetta, siirtymää tai sähkökenttää, voidaan arvioida

diskreetin mallin avulla. Malli koostuu paloittain jatkuvista funktioista, jotka on määritelty äärellisessä määrässä mallinnusalueen osa-alueita. Toisin sanoen jatkuvaa suuretta pyritään arvioimaan yksinkertaisella funktiolla, joka koostetaan joukosta paloittain jatkuvia funktioita ja jonka esittämiseen tarvitaan vain äärellinen määrä parametreja. Diskreetissä simulointitehtävässä parametreille etsitään sellaiset arvot, että yksinkertainen funktio kuvaa mahdollisimman hyvin ratkaistavaa suuretta.

Äärellisten elementtien menetelmässä tuntematonta funktiota \mathbf{u} approksimoidaan yleisessä tapauksessa summalla

$$\mathbf{u}(\mathbf{x}) \approx \tilde{\mathbf{u}}(\mathbf{x}) = q_1 \vec{\Psi}_1(\mathbf{x}) + \dots + q_n \vec{\Psi}_n(\mathbf{x}) = \sum_{i=1}^n q_i \vec{\Psi}_i(\mathbf{x}), \quad (5.1)$$

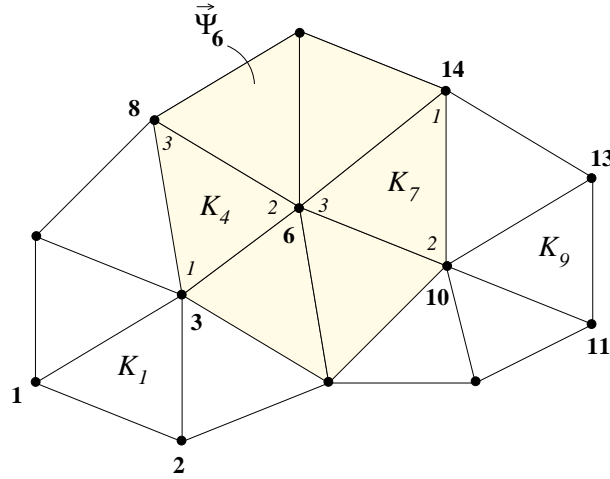
jossa on n kappaletta ratkaistavia kertoimia q_i . Tuntematonta kerrointa q_i kutsutaan funktion $\vec{\Psi}_i$ vapausasteeksi. Funktiot $\vec{\Psi}_i$ tunnetaan ja ne valitaan sopivasti kulloisenkin tehtävän perusteella. Tuntemattomien kertoimien ratkaisemiseksi tarvitaan joukko yhtälöitä. Yhtälöiden muodostaminen on osa simulointitehtävän diskretisointia.

Elementtimenetelmässä simulointialue jaetaan pieniin osa-alueisiin eli *elementteihin*. Elementtien tulee täyttää koko simulointialue siten, että ne eivät ole missään kohdin päällekkäin. Tasoalueiden ja pintojen jakoon käytetään tavallisimmin kolmioita ja nelikulmioita. Kolmiulotteisten alueiden jakaminen tapahtuu yleisimmin nelitahokkaiden ja suorakulmaisilla särmiöillä. Myös käyräviivaisia elementtejä käytetään yleisesti. Huomautettakoon, että alueiden jakaminen elementteihin voidaan suorittaa äärettömän monella tavalla. Kun simulointialue on ositettu elementteihin, siitä käytetään usein nimitystä *verkko*.

Elementtien geometrian määrääviä pisteitä, kuten esimerkiksi kolmion kärkipisteitä, kutsutaan solmuiksi ja elementtien reunoilla olevat solmut ovat vierekkäisille elementeille yhteisiä. Kuvassa 5.1 on esimerkin vuoksi esitetty yksi keino jakaa kaksiulotteinen monikulmio kolmioelementteihin. Simulointialueen jakaminen pieniin osa-alueisiin tilanteeseen sopivilla elementtityypeillä mahdollistaa monimutkaisten geometrioiden mallintamisen.

5.2.1 Kantafunktiot

Yhtälössä (5.1) esiintyviä funktioita $\vec{\Psi}_i$ kutsutaan *kantafunktioiksi*. Useimmiten kantafunktioiksi valitaan matala-asteisia paloittain määriteltyjä polynomeja. Kantafunktiot määritellään yleensä siten, että ne poikkeavat nolasta ainoastaan niiden elementtien alueella, joihin kantafunktioon liittyvä vapausaste q_i kuuluu. Mikäli vapausasteet yhdistetään esimerkiksi solmupisteisiin, niin vapausaste kuuluu kaikkiin niihin elementteihin, joihin vapausasteeseen yhdistetty solmupiste kuuluu. Esimerkiksi kuvassa 5.1 on tilanne, missä vapausasteet on yhdistetty solmupisteisiin ja tällöin kantafunktio $\vec{\Psi}_6$ poikkeaa nolasta ainoastaan tummennettujen elementtien alueella.



Kuva 5.1: Esimerkki kaksiulotteisin monikulmion jakamisesta kolmioelementteihin.

Kantafunktion paloittainen määrittely toteutetaan siten, että kantafunktio määritellään erikseen kaikissa niissä elementeissä, joissa se poikkeaa nolasta. Globaalin kantafunktion $\vec{\Psi}_i$ määritelmää elementissä K_k sanotaan *lokaaliksi kantafunktioksi* ja siitä käytetään merkintää $\vec{\psi}_j^k$, missä j on lokaalin kantafunktion elementtikohtainen indeksi. Globaali kantafunktio koostuu siis joukosta lokaaleja kantafunktioita. Esimerkiksi kuvan 5.1 tapauksessa kantafunktion $\vec{\Psi}_6$ epätäydellinen määrittelmä on

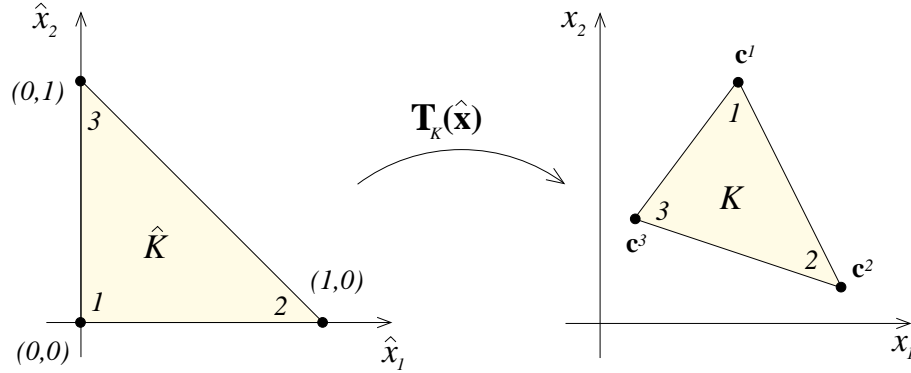
$$\vec{\Psi}_6(\mathbf{x}) = \begin{cases} \vec{\psi}_2^4(\mathbf{x}), & \text{kun } \mathbf{x} \text{ on elementin } K_4 \text{ alueella,} \\ \vdots & \\ \vec{\psi}_3^7(\mathbf{x}), & \text{kun } \mathbf{x} \text{ on elementin } K_7 \text{ alueella,} \\ 0, & \text{muualla.} \end{cases}$$

Lokaalien kantafunktioiden avulla tuntemattomien kertoimien q_i ratkaisemiseen tarvittavat yhtälöt voidaan muodostaa elementteittäin. Lopulliset yhtälöt kertoimien ratkaisemiseen saadaan, kun summataan vierekkäisten elementtien yhteiseen tuntemattomaan kertoimeen kohdistuvat yhtälöt. Yhtälöiden muodostamista kutsutaan *tehtävän kokoamiseksi*.

5.2.2 Viite-elementit

Elementtimenetelmälle muodostettava matemaattisen mallin heikko muotoilu, kuten esimerkiksi (4.15), sisältää aina integraalilausekkeitä. Diskreetin tehtävän kokoaminen edellyttää kyseisten integraalilausekkeiden laskemista. Tällöin ongelmaksi voi muodostua se, että funktioiden integroiminen mielivaltaisten elementtien yli voi olla hankalaa.

Yleisesti käytetty keino ongelman kiertämiseen on määritellä eri elementtityypeille *viite-elementti* \hat{K} *lokaalissa koordinaatistossa* $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$. Lokaali koordinaatisto vali-



Kuva 5.2: Esimerkki viite-elementin kuvautumisesta alkuperäiseen koordinaatistoon.

taan siten, että integrointi viite-elementin yli on mahdollisimman yksinkertaista. Ajatuksena on, että funktioiden integrointi suoritetaan mielivaltaisen elementin sijasta viite-elementin yli. Integrointien laskemiseen viite-elementin avulla tarvitaan muunnoskuvaus $\mathbf{T}_K : \hat{K} \rightarrow K$, joka kuvaa viite-elementin \hat{K} lokaalin koordinaattipisteen alkuperäisessä koordinaatistossa sijaitsevan elementin K pisteeksi. Merkinnällä $\hat{\mathbf{x}}$ tarkoitetaan lokaalin koordinaatiston vektoria.

Kolmiulotteisissa tapauksissa muunnoskuvauksen yleinen komponenttimuoto on

$$\begin{aligned} x_1 &= (T_K)_1(\hat{\mathbf{x}}) = \eta_1(\hat{\mathbf{x}})c_1^1 + \eta_2(\hat{\mathbf{x}})c_1^2 + \dots + \eta_m(\hat{\mathbf{x}})c_1^m, \\ x_2 &= (T_K)_2(\hat{\mathbf{x}}) = \eta_1(\hat{\mathbf{x}})c_2^1 + \eta_2(\hat{\mathbf{x}})c_2^2 + \dots + \eta_m(\hat{\mathbf{x}})c_2^m, \\ x_3 &= (T_K)_3(\hat{\mathbf{x}}) = \eta_1(\hat{\mathbf{x}})c_3^1 + \eta_2(\hat{\mathbf{x}})c_3^2 + \dots + \eta_m(\hat{\mathbf{x}})c_3^m \end{aligned} \quad (5.2)$$

ja vastaava vektorimuoto

$$\mathbf{x} = \mathbf{T}_K(\hat{\mathbf{x}}) = \sum_{i=1}^m \eta_i(\hat{\mathbf{x}})\mathbf{c}^i, \quad (5.3)$$

missä muunnoskuvauksen kantafunktiot η_i on liitetty elementin K geometrisen muodon määritteleviin pisteisiin \mathbf{c}^i . Merkinnällä $(T_K)_i$ tarkoitetaan vektoriarvoisen kuvauksen \mathbf{T}_K i :nnettä komponenttia. Kuvassa 5.2 on esimerkki kaksiulotteisen lineaarisen kolmioelementtityypin viite-elementistä ja sen kuvautumisesta yleiseksi kolmioelementiksi.

Kun funktioiden integroinnissa tehdään koordinaattimuunnos, niin yleisessä tapauksessa tarvitaan vielä muunnoskuvauksen *Jacobin matriisia* ja erityisesti sen determinanttia. Yhtälöiden (5.2) ja (5.3) määrittelemän muunnoskuvauksen \mathbf{T}_K Jacobin

matriisi on [5, sivu 216]

$$D\mathbf{T}_K(\hat{\mathbf{x}}) = \mathbf{J}_K(\hat{\mathbf{x}}) = \begin{bmatrix} \frac{\partial(T_K)_1(\hat{\mathbf{x}})}{\partial\hat{x}_1} & \frac{\partial(T_K)_1(\hat{\mathbf{x}})}{\partial\hat{x}_2} & \frac{\partial(T_K)_1(\hat{\mathbf{x}})}{\partial\hat{x}_3} \\ \frac{\partial(T_K)_2(\hat{\mathbf{x}})}{\partial\hat{x}_1} & \frac{\partial(T_K)_2(\hat{\mathbf{x}})}{\partial\hat{x}_2} & \frac{\partial(T_K)_2(\hat{\mathbf{x}})}{\partial\hat{x}_3} \\ \frac{\partial(T_K)_3(\hat{\mathbf{x}})}{\partial\hat{x}_1} & \frac{\partial(T_K)_3(\hat{\mathbf{x}})}{\partial\hat{x}_2} & \frac{\partial(T_K)_3(\hat{\mathbf{x}})}{\partial\hat{x}_3} \end{bmatrix}. \quad (5.4)$$

Jacobin matriisin determinanttia $\det(D\mathbf{T}_K) = \det(\mathbf{J}_K)$ kutsutaan *Jacobiaaniksi*.

5.2.3 Äärellinen elementti

Edellisissä luvuissa on esitelty elementtimenetelmän käsitteitä lähinnä esimerkkien avulla. Seuraavaksi määritellään täsmällisemmin mitä tarkoitetaan *äärellisellä elementillä*. Tässä luvussa esitety asiat perustuvat kirjoihin [4, sivut 93-96] ja [17, sivut 24-29]. Elementtimenetelmässä rajoitettu simulointialue Ω jaetaan äärelliseen määrään osa-alueita, jolloin osa-alueiden joukko \mathcal{T}_h on simulointialueen diskreetti esitys. Yksittäistä osa-aluetta sanotaan elementiksi ja siitä käytetään tässä luvussa merkintää K . Osa-alueisiin jakaminen eli simulointialueen diskretisointi suoritetaan sillä tavalla, että seuraavat ominaisuudet ovat voimassa:

1. $\bar{\Omega} = \bigcup_{i=1}^m K_i$ $K_i \in \mathcal{T}_h$, missä m on elementtien lukumäärä,
2. kaikki elementit $K_i \in \mathcal{T}_h$ ovat suljettuja ja niiden sisusta K_i^0 ei ole nollamittainen,
3. kaikille erillisille elementeille $K_1, K_2 \in \mathcal{T}_h$ pätee $K_1^0 \cap K_2^0 = \emptyset$,
4. kaikkien elementtien $K_i \in \mathcal{T}_h$ reuna ∂K_i on Lipschitz-reuna.

Äärellisen elementin yleinen määritelmä on:

Määritelmä 5.2.1 *Äärellinen elementti on kolmikko (K, P, Σ) , missä*

- (I) K on avaruuden \mathbb{R}^d suljettu osa-alue, jonka sisusta ei ole nollamittainen ja jonka reuna on Lipschitz-reuna,
- (II) P on alueessa K määritelty äärellisulotteinen funktioavaruus; käytetään merkintää $n = \dim P$,
- (III) Σ on äärellinen joukko lineaarisesti riippumattomia lineaarisia funktioaaleja $\Phi_i, 1 \leq i \leq n$, jotka kuvautuvat avaruudesta P (tai avaruudesta, joka sisältää avaruuden P).

Funktionaalit Φ_i ovat niin sanotusti *P-unisolventteja*, jos mielivaltaiselle joukolle reaalityyppisiä lukuja a_i , $1 \leq i \leq n$ on olemassa yksikäsitteinen funktio $\vec{\varphi} \in P$ siten, että

$$\Phi_i(\vec{\varphi}) = a_i, \quad 1 \leq i \leq n. \quad (5.5)$$

Jos funktionaalit toteuttavat edellä mainitun ehdon, niin ne ovat lineaarisesti riippumattomia. Ehdon toteutumisesta seuraa myös, että on olemassa n kappaletta funktioita $\vec{\varphi}_i \in P$, $1 \leq i \leq n$, joille pätee

$$\Phi_j(\vec{\varphi}_i) = \delta_{ij}, \quad \delta_{ij} = \begin{cases} 1, & \text{kun } i = j, \\ 0, & \text{kun } i \neq j, \end{cases} \quad (5.6)$$

missä $1 \leq j \leq n$. Merkintää δ_{ij} kutsutaan *Kroneckerin symboliksi*.

Kun oletetaan, että funktionaalit Φ_i toteuttavat ehdon (5.5), niin soveltamalla ehdon suoraa seurausta (5.6) saadaan tuloksena identiteetti

$$\vec{\varphi} = \sum_{i=1}^n \Phi_i(\vec{\varphi}) \vec{\varphi}_i, \quad \forall \vec{\varphi} \in P, \quad (5.7)$$

joka tarkoittaa, että kaikki avaruuden P funktiot voidaan esittää funktionaalien Φ_i ja funktioiden $\vec{\varphi}_i$ avulla. Lineaarisia funktionaaleja Φ_i sanotaan äärellisen elementin *vapausasteiksi* ja funktiota $\vec{\varphi}_i$ äärellisen elementin *kantafunktioiksi*.

5.3 Nédélec-elementti

Yksinkertaisimmat elementtityypit ja niissä määritellyt kantafunktiot eivät sovellu sähkökentän $\hat{\mathbf{E}}_0^s$ approksimointiin. Syy on siinä, että käyttämällä yksinkertaisten elementtityyppien kantafunktiota, sähkömagneettisen simulointitehtävän ratkaisulta vaaditaan liian vahvoja jatkuvuusominaisuuksia. Avaruuteen $H(\text{curl}, \Omega)$ kuuluvia funktioita approksimoidaan globaaleilla kantafunktioilla, joiden tangentialisten komponenttien täytyy olla jatkuvia eri elementtien välillä. Tangentiaalinen jatkuvuus kahden elementin välillä ilmaistaan matemaattisesti seuraavalla tavalla:

Lause 5.3.1 *Olkoot K_- ja K_+ kaksi Lipschitz-reunaista monikulmiota (vast. monitahokasta) avaruudesta \mathbb{R}^d , joilla on yhteinen reuna (vast. reuna tai pinta) $e = \partial K_- \cap \partial K_+ \neq \emptyset$ ja joiden yhdisteestä käytetään merkintää $\Omega = \partial K_- \cup \partial K_+$. Vektoreilla \mathbf{n}_- ja \mathbf{n}_+ tarkoitetaan monikulmioiden tai -tahokkaiden ulkonormaalivektoreita. Funktio \mathbf{u} kuuluu avaruuteen $H(\text{curl}, \Omega)$ jos ja vain jos rajoittuma $\mathbf{u}|_{K_-} = \mathbf{u}_-$ kuuluu avaruuteen $H(\text{curl}, K_-)$, rajoittuma $\mathbf{u}|_{K_+} = \mathbf{u}_+$ kuuluu avaruuteen $H(\text{curl}, K_+)$ ja tangentiaalinen hyppy leikkauksen e yli häviää: $(\mathbf{u}_- \times \mathbf{n}_-) + (\mathbf{u}_+ \times \mathbf{n}_+) = 0 \quad \mathbf{x} \in e$.*

Tangentiaalisen jatkuvuuden toteuttavat esimerkiksi kantafunktiot, jotka määritellään niin sanotuilla *Nédélec-elementeillä*. Tässä tutkielmassa käytetään sähkökentän $\hat{\mathbf{E}}_0^s$ approksimointiin ensimmäisen asteen Nédélec-kuutioelementtejä, jotka $J. C.$

Nédélec esitteli artikkelissaan [22] vuonna 1980. Varsinaiset kuutioelementit, joilla simulointialue jaetaan osa-alueisiin, kuvataan viite-elementin avulla. Aivan kuten yleinen äärellinen elementti, niin myös Nédélec-kuutioviiite-elementti määritellään elementin geometrian, kantafunktioiden ja vapausasteiden valinnoilla. Luonnollinen valinta kuutioviiite-elementin geometriaksi on kuutio $\hat{C} = [0, 1]^3$ koordinaatistossa $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$. Koordinaatistossa (x_1, x_2, x_3) sijaitsevat varsinaiset kuutioelementit C kuvataan affiinilla muunnoskuvaksella \mathbf{T}_C siten, että

$$C = \mathbf{T}_C(\hat{C}), \quad C \ni \mathbf{x} = \mathbf{T}_C(\hat{\mathbf{x}}) = \mathbf{B}_C \hat{\mathbf{x}} + \hat{\mathbf{b}}_C \quad \hat{\mathbf{x}} \in \hat{C}, \quad (5.8)$$

missä matriisi \mathbf{B}_C on neliömatriisi.

Viite-elementin kantafunktiot valitaan vektoriarvoisten polynomifunktioiden joukosta. Seuraavaksi esitellään joukko polynomifunktioavaruuksia, joiden avulla määritellään elementin kantafunktioiden komponenttifunktiot:

Määritelmä 5.3.1 *Funktioavaruudet $\mathcal{Q}_{l,m,n}$ sisältävät sellaiset polynomit, jotka kuvautuvat viite-elementistä \hat{C} ja joiden aste muuttujien \hat{x}_1, \hat{x}_2 ja \hat{x}_3 suhteen on korkeintaan l, m ja n vastaavassa järjestyksessä.*

Edellistä määritelmää käyttäen elementin \hat{C} kantafunktioavaruudeksi valitaan

$$\mathcal{P}^k = \left\{ \vec{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} : \varphi_1 \in \mathcal{Q}_{k-1,k,k}, \varphi_2 \in \mathcal{Q}_{k,k-1,k}, \varphi_3 \in \mathcal{Q}_{k,k,k-1} \right\}. \quad (5.9)$$

Koska tässä tutkielmassa käytetään ensimmäisen asteen elementtejä, niin kaikki kantafunktiot ovat silloin muotoa

$$\vec{\varphi}(\hat{\mathbf{x}}) = \begin{pmatrix} a_1 \hat{x}_2 + b_1 \hat{x}_3 + c_1 \hat{x}_2 \hat{x}_3 + d_1 \\ a_2 \hat{x}_1 + b_2 \hat{x}_3 + c_2 \hat{x}_1 \hat{x}_3 + d_2 \\ a_3 \hat{x}_1 + b_3 \hat{x}_2 + c_3 \hat{x}_1 \hat{x}_2 + d_3 \end{pmatrix} \in \mathcal{P}^1. \quad (5.10)$$

Viite-elementin \hat{C} kantafunktioiden määrittelemiseksi täytyy kertoimet a_i, b_i, c_i ja $d_i \in \mathbb{R}$ ratkaista. Kertoimet saadaan ratkaistua vapausasteiden Φ avulla. Esitellään seuraavaksi Nédélec-kuutioelementtiin liittyvät vapausasteet:

Määritelmä 5.3.2 *Olkoon \hat{C} kuutioviiite-elementti, $\hat{\mathbf{t}}$ kuution särmen yksikkötangenttivektori ja $\hat{\mathbf{n}}$ kuution tahkon ulkoinen yksikkönormaalivektori. Kantafunktioavaruudesta \mathcal{P}^k kuvautuvien vapausasteiden joukko Σ koostuu lineaarisista funktionaaleista, jotka on yhdistetty kuution eri osa-alueisiin:*

vapausasteet särmillä

$$\Phi(\vec{\varphi}) := \int_{\hat{e}} (\hat{\mathbf{t}} \cdot \vec{\varphi}) v \, d\hat{\mathbf{s}}, \quad \forall v \in \mathbb{P}_{k-1}(\hat{e}),$$

jokaisella kuution \hat{C} särmillä \hat{e} . Kuution särmillä on yhteensä $12k$ vapausastetta.

vapausasteet tahkoilla

$$\Phi(\vec{\varphi}) := \int_{\hat{f}} (\vec{\varphi} \times \hat{\mathbf{n}}) \cdot \mathbf{v} \, d\hat{\mathbf{a}}, \quad \forall \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad v_1 \in \mathcal{Q}_{k-2,k-1}(\hat{f}), \quad v_2 \in \mathcal{Q}_{k-1,k-2}(\hat{f}),$$

jokaisella kuution \hat{C} tahkolla \hat{f} . Kuution tahkoilla on yhteensä $6 \cdot 2k(k-1)$ vapausastetta.

vapausasteet kuution sisällä

$$\Phi(\vec{\varphi}) := \int_{\hat{C}} \vec{\varphi} \cdot \mathbf{v} \, d\hat{\mathbf{x}}, \quad \forall \mathbf{v} = \begin{pmatrix} v_1 \in \mathcal{Q}_{k-1,k-2,k-2} \\ v_2 \in \mathcal{Q}_{k-2,k-1,k-2} \\ v_3 \in \mathcal{Q}_{k-2,k-2,k-1} \end{pmatrix}.$$

Vapausasteita kuution sisällä on yhteensä $3k(k-1)^2$.

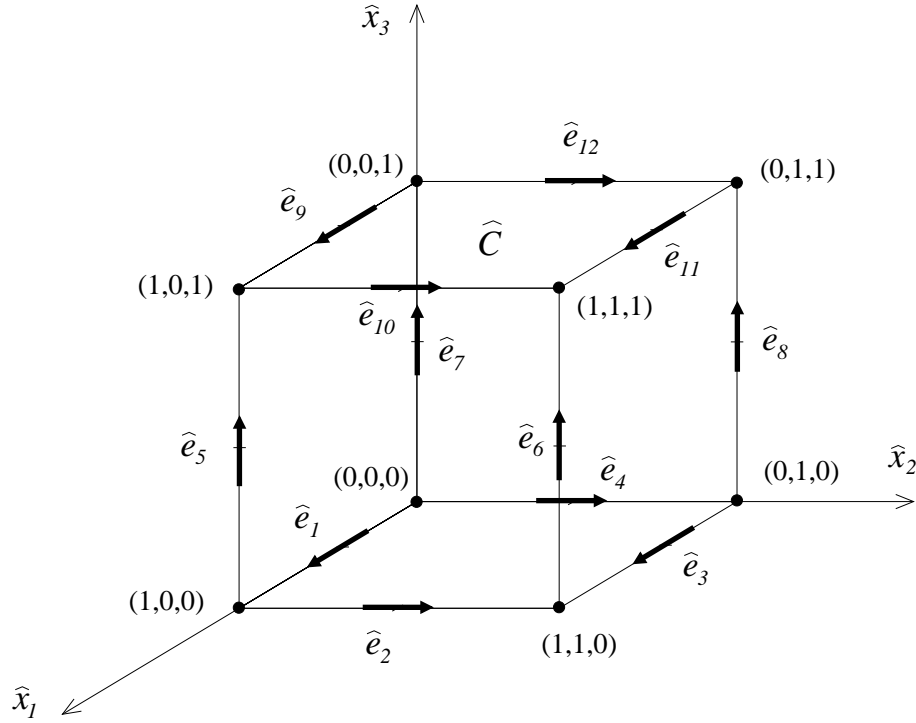
Funktioavaruus $\mathbb{P}_{k-1}(\hat{e})$ sisältää kaikki sellaiset polynomit, jotka on määritelty särmillä \hat{e} ja ovat korkeintaan astetta $k-1$. Vapausasteiden määritelmän (5.3.2) perusteella ensimmäisen asteen kuutioelementillä on kaksitoista vapausastetta, yksi jokaisella särmillä. Seuraavaksi selvitetään kyseisiin vapausasteisiin liittyvien kantafunktioiden kertoimet a_i, b_i, c_i ja $d_i \in \mathbb{R}$, kantafunktioiden muoto esiteltiin yhtälössä (5.10). Koska kuution särmiin yhdistettyjen vapausasteiden määrittelyssä esiintyvien integraaliyhtälöiden tulee olla voimassa kaikilla testifunktioilla $v \in \mathbb{P}_0$, niin erityisesti niiden tulee olla voimassa, kun $v \equiv 1$. Lisäksi tiedetään, että vapausasteiden tulee olla lineaarisesti riippumattomia. Vaaditaan, että vapausasteet toteuttavat ehdon

$$\Phi_i(\vec{\varphi}_j) = \delta_{ij}, \quad i, j = 1, \dots, 12, \quad (5.11)$$

jolloin vapausasteet ovat lineaarisesti riippumattomia.

Vapausasteille asetettu ehto (5.11) tuottaa kaksitoista yhtälöä jokaista kantafunktiota kohden. Koska kaikilla kantafunktioilla on kaksitoista tuntematonta kerrointa, niin jokaisen kantafunktion kertoimet voidaan ratkaista yksikäsitteisesti yhtälöryhmästä, jossa on sama määrä yhtälöitä kuin tuntemattomia. Kun numeroidaan kuution särmit kuvan 5.3 esittämällä tavalla ja ratkaistaan kaikkiin kantafunktioihin liittyvät yhtälöryhmät, niin kuutioviite-elementin kantafunktioiksi saadaan

$$\begin{aligned} \vec{\varphi}_1(\hat{\mathbf{x}}) &= (1 - \hat{x}_2 - \hat{x}_3 + \hat{x}_2\hat{x}_3, 0, 0), & \vec{\varphi}_2(\hat{\mathbf{x}}) &= (0, \hat{x}_1 - \hat{x}_1\hat{x}_3, 0), \\ \vec{\varphi}_3(\hat{\mathbf{x}}) &= (\hat{x}_2 - \hat{x}_2\hat{x}_3, 0, 0), & \vec{\varphi}_4(\hat{\mathbf{x}}) &= (0, 1 - \hat{x}_1 - \hat{x}_3 + \hat{x}_1\hat{x}_3, 0), \\ \vec{\varphi}_5(\hat{\mathbf{x}}) &= (0, 0, \hat{x}_1 - \hat{x}_1\hat{x}_2), & \vec{\varphi}_6(\hat{\mathbf{x}}) &= (0, 0, \hat{x}_1\hat{x}_2), \\ \vec{\varphi}_7(\hat{\mathbf{x}}) &= (0, 0, 1 - \hat{x}_1 - \hat{x}_2 + \hat{x}_1\hat{x}_2), & \vec{\varphi}_8(\hat{\mathbf{x}}) &= (0, 0, \hat{x}_2 - \hat{x}_1\hat{x}_2), \\ \vec{\varphi}_9(\hat{\mathbf{x}}) &= (\hat{x}_3 - \hat{x}_2\hat{x}_3, 0, 0), & \vec{\varphi}_{10}(\hat{\mathbf{x}}) &= (0, \hat{x}_1\hat{x}_3, 0), \\ \vec{\varphi}_{11}(\hat{\mathbf{x}}) &= (\hat{x}_2\hat{x}_3, 0, 0), & \vec{\varphi}_{12}(\hat{\mathbf{x}}) &= (0, \hat{x}_3 - \hat{x}_1\hat{x}_3, 0). \end{aligned} \quad (5.12)$$



Kuva 5.3: Nédélec-kuutioviite-elementti ja sen särmien numerointi.

Kuvassa 5.3 on nuolilla merkitty viite-elementin särmien tangenttivektoreiden suunta. Edelle esitetyt määritelmät yhdessä muodostavat äärellisen elementin $(\hat{C}, \mathcal{P}^1, \Sigma)$, joka on *unisolventti* ja *konformi* avaruudessa $H(\text{curl}, \hat{C})$ [22, sivu 331].

Koska varsinaiset kuutioelementit C kuvataan viite-elementin \hat{C} avulla, niin myös elementtien C kantafunktiot $\vec{\psi}_i$ täytyy kuvata viite-elementin kantafunktioiden $\vec{\varphi}_i$ avulla. Ensimmäisen asteen Nédélec-elementin $C = \mathbf{T}_C(\hat{C})$ kantafunktiot saadaan kuvauksella

$$\vec{\psi}_i(\mathbf{x}) = \mathcal{M}_C(\vec{\varphi}_i(\hat{\mathbf{x}})) = (D\mathbf{T}_C^{-T} \vec{\varphi}_i) \circ \mathbf{T}_C^{-1}(\mathbf{x}), \quad (5.13)$$

missä $\mathbf{T}_C^{-1}(\mathbf{x})$ on muunnoskuvauksen $\mathbf{T}_C(\hat{\mathbf{x}})$ käänteiskuvaus ja $D\mathbf{T}_C^{-T}$ on lyhennysmerkintä Jacobin käänteismatriisin transpoosille $((D\mathbf{T}_C)^{-1})^T$. Koska kuvaus \mathbf{T}_C on määritelty affiiniksi (5.8), sen Jacobin matriisi $D\mathbf{T}_C$ on yksinkertaisesti vakiomatriisi \mathbf{B}_C ja Jacobiaani $\det(D\mathbf{T}_C)$ on vakioarvoinen reaaliluku $\det(\mathbf{B}_C)$. Jatkossa Jacobiaani oletetaan positiiviseksi $\det(\mathbf{B}_C) > 0$. Tällöin yhtälö (5.13) sievenee muotoon

$$\vec{\psi}_i(\mathbf{x}) = \mathcal{M}_C(\vec{\varphi}_i(\hat{\mathbf{x}})) = \mathbf{B}_C^{-T}(\vec{\varphi}_i \circ \mathbf{T}_C^{-1})(\mathbf{x}). \quad (5.14)$$

Koordinaatistossa (x_1, x_2, x_3) määritelty roottori $\nabla \times$ täytyy myös laskea muunnoskuvauksen avulla. Olkoon \mathbf{u} kuutiosta $C = \mathbf{T}_C(\hat{C})$ kuvautuva vektoriarvoinen funktio, joka on määritelty muunnoksella (5.14) vektoriarvoisesta funktiosta $\hat{\mathbf{u}}$. Tällöin funktion

\mathbf{u} roottori on laskettavissa kuvauksella

$$\nabla \times \mathbf{u}(\mathbf{x}) = \frac{1}{\det(\mathbf{B}_C)} \mathbf{B}_C(\overline{\nabla \times \hat{\mathbf{u}}} \circ \mathbf{T}_C^{-1})(\mathbf{x}), \quad (5.15)$$

missä merkinnällä $\overline{\nabla \times}$ tarkoitetaan koordinaatistossa $(\hat{x}_1, \hat{x}_2, \hat{x}_3)$ määriteltyä roottoria.

Edellä on määritelty kuutioviite-elementti, jonka avulla kuvataan kaikki simulointialueen varsinaiset kuutioelementit. Jokaisen elementin C alueella tutkittavaa suuretta, kuten esimerkiksi sähkökenttää $\hat{\mathbf{E}}_0^s$, voidaan approksimoida kyseisen elementin kantafunktioiden lineaarikombinaatiolla. Suuretta halutaan kuitenkin approksimoida koko simulointialueessa Ω globaaleilla kantafunktioilla, jotka ovat paloittain määriteltyjä usean elementin alueella. Jotta globaalit kantafunktiot kuuluisivat funktioavaruuteen $H(\text{curl}, \Omega)$, täytyy elementtien C kantafunktiot kuvata viite-elementin \hat{C} kantafunktioista siten, että vierekkäisten elementtien välillä globaalit kantafunktiot ovat tangentialisesti jatkuvia.

Oletetaan, että simulointialueen diskretisointi on suoritettu ja näin ollen tiedetään todellisen verkon solmupisteet. Erityisesti oletetaan, että diskretisoinnin jälkeen tunnetaan särmien tangenttivektorit. Valitaan tangenttivektoreiden joukosta mielivaltaisesti vektori \mathbf{t}_j , missä alaindeksi j on sen särmän globaali tunnistenumero, johon vektori liittyy. Nyt vaaditaan, että kaikki särmään j liittyvät lokaalit kantafunktiot $\vec{\psi}_i$ ovat särmällä j tangenttivektorin \mathbf{t}_j suuntaisia. Tällöin kaikki globaalit kantafunktiot olisivat tangentialisesti jatkuvia särmän j yli. Vaatimus täyttyy, jos ensin muodostetaan kaikki särmään liittyvät lokaalit kantafunktiot $\vec{\psi}_i$ normaalisti muunnoksella (5.14). Tämän jälkeen kantafunktioiden suuntaa verrataan sisätulon avulla tangenttivektoriin \mathbf{t}_j . Viite-elementin särmään i liittyvän kantafunktion $\vec{\varphi}_i$ suunnaksi särmällä i valitaan esimerkiksi särmän tangenttivektorin suunta.

Jos $(\vec{\psi}_i, \mathbf{t}_j) \geq 0$, niin ei ryhdytä minkäänlaisiin toimenpiteisiin. Mutta jos sisätulo onkin negatiivinen $(\vec{\psi}_i, \mathbf{t}_j) < 0$, niin suoritetaan päivitys $\vec{\psi}_i \leftarrow (-\vec{\psi}_i)$. Edellisellä merkinnällä tarkoitetaan, että lokaalin kantafunktion $\vec{\psi}_i$ suunta vaihdetaan vastakkaisuuntaiseksi alkuperäiseen suuntaan verrattuna. Koska lokaalit kantafunktiot $\vec{\psi}_i$ voivat muunnoksessa (5.14) kuvautua ainoastaan saman suuntaiseksi tai vastakkaisuuntaiseksi tangenttivektorin \mathbf{t}_j kanssa, niin päivitys takaa, että lokaalit kantafunktiot $\vec{\psi}_i$ ovat operaation jälkeen särmällä j tangenttivektorin \mathbf{t}_j suuntaisia. Todetaan, että suoritettamalla päivitysoperaatio kaikille verkon särmille saadaan myös kaikki verkon lokaalit kantafunktiot särmillä tangenttivektoreiden suuntaisiksi. Näin ollen kaikki globaalit kantafunktiot ovat tangentialisesti jatkuvia vierekkäisten elementtien välillä.

Koska kaikki globaalit kantafunktiot $\vec{\Psi}$ ovat tangentialisesti jatkuvia vierekkäisten elementtien välillä, niin ne kuuluvat avaruuteen $H(\text{curl}, \Omega)$. Tällöin tutkittavaa suuretta voidaan approksimoida globaalien kantafunktioiden lineaarikombinaationa. Kan-

tafunktioit valitaan avaruudesta V_h , jonka määritelmä on [21, sivu 11]

$$\begin{aligned} U_h &= \{\mathbf{u}_h \in H(\text{curl}, \Omega) \mid \mathbf{u}_h(\mathbf{x}) \in \mathcal{P}^1, \mathbf{x} \in C \quad \forall C \in \mathcal{T}_h\}, \\ V_h &= \{\mathbf{v}_h \in U_h \mid \mathbf{v}_h \times \mathbf{n} = 0, \mathbf{x} \in \Gamma_s\} \subset H_0(\text{curl}, \Omega, \Gamma_s). \end{aligned} \quad (5.16)$$

Kun tutkittavaa suuretta approksimoidaan käyttämällä avaruuden V_h kantafunktioita, aiheutuu virhe, joka on $H(\text{curl})$ -normilla mitattuna kertaluokkaa $\mathcal{O}(h)$. Muuttuja h edustaa verkon \mathcal{T}_h suurimman kuutioelementin särmän pituutta.

5.4 Yhtälöryhmän kokoaminen

Viimeinen vaihe simulointitehtävän diskretisoinnissa on yhtälöryhmän muodostaminen. Jaetaan luvussa 3.3 määritelty simulointialue Ω äärelliseen määrään osa-alueita luvun 5.2.3 esittämällä tavalla. Käytetään osa-alueiden joukosta merkintää \mathcal{T}_h . Lisäksi määritellään joukko \mathcal{F}_h , jonka alkioita ovat kaikkien kuutioelementtien $K \in \mathcal{T}_h$ ne tahkot, jotka kuuluvat reunalle Γ_a .

Tehtävän (4.15) diskretisoitu muoto saadaan käyttämällä kohdan (5.16) äärellisluotteista aliavaruutta V_h . Sähkökenttää $\hat{\mathbf{E}}_0^s$ approksimoidaan funktiolla \mathbf{E}_h . Simulointitehtävän diskreetti muoto on: etsi funktio $\mathbf{E}_h \in V_h$ siten, että

$$a(\mathbf{E}_h, \mathbf{v}_h) = -a(\mathbf{G}, \mathbf{v}_h) \quad \forall \mathbf{v}_h \in V_h. \quad (5.17)$$

Bilineaarimuoto $a(\cdot, \cdot)$ on määritelty yhtälössä (4.13) ja funktio $G \in H(\text{curl}, \Omega)$ voidaan valita sopivasti, kunhan se toteuttaa luvussa 4.3 esitetyt ehdot.

Olkoon $\{\vec{\Psi}_i\}_{i=1}^n$ joukko globaaleja kantafunktioita ja n verkkoon \mathcal{T}_h sisältyvien särmien lukumäärä. Määritellään funktio \mathbf{E}_h lineaarikombinaationa globaaleista kantafunktioista $\vec{\Psi}_i$

$$\hat{\mathbf{E}}_0^s(\mathbf{x}) \approx \mathbf{E}_h(\mathbf{x}) = \sum_{i=1}^n q_i \vec{\Psi}_i(\mathbf{x}) \quad \mathbf{x} \in \Omega, \quad (5.18)$$

missä kertoimet $q_i \in \mathbb{K}$ ovat joko reaali- tai kompleksilukuja. Tehtävän diskretisointi suoritetaan Galerkinin menetelmällä, joka on eräs elementtimenetelmän variaatio. Tällöin testifunktioina \mathbf{v}_h käytetään kantafunktioita $\vec{\Psi}_i$. Sijoittamalla summalauseke (5.18) tehtävään (5.17) saadaan

$$\begin{aligned} a\left(\sum_{j=1}^n q_j \vec{\Psi}_j, \vec{\Psi}_i\right) &= -a(\mathbf{G}, \vec{\Psi}_i) \quad i = 1, \dots, n \\ \Leftrightarrow \sum_{j=1}^n a(\vec{\Psi}_j, \vec{\Psi}_i) q_j &= -a(\mathbf{G}, \vec{\Psi}_i) \quad i = 1, \dots, n. \end{aligned} \quad (5.19)$$

Kirjoitetaan yhtälö (5.19) matriisimuodossa

$$\mathbf{A}\mathbf{q} = \mathbf{f}, \quad (5.20)$$

jolloin saadaan lineaarinen yhtälöryhmä, josta ratkaistaan vektori \mathbf{q} . Matriisi $\mathbf{A} = \{a_{i,j}\}_{i,j=1}^n$ muodostuu alkioista $a_{i,j} := a(\vec{\Psi}_i, \vec{\Psi}_j)$ ja vektori $\mathbf{f} = \{f_i\}_{i=1}^n$ alkioista $f_i := -a(\mathbf{G}, \vec{\Psi}_i)$. Yhtälöryhmästä (5.20) ratkaistava vektori $\mathbf{q} = \{q_i\}_{i=1}^n$ koostuu yksinkertaisesti yhtälön (5.18) tuntemattomista kertoimista q_i . Tarkastellaan seuraavaksi miten matriisin \mathbf{A} ja vektorin \mathbf{f} alkiot lasketaan.

Kirjoitetaan matriisin \mathbf{A} alkioiden $a_{i,j}$ lausekkeet kokonaisuudessaan bilineaarimuodon avulla

$$a_{i,j} = a(\vec{\Psi}_i, \vec{\Psi}_j) = \int_{\Omega} \nabla \times \vec{\Psi}_i \cdot \nabla \times \vec{\Psi}_j - \alpha^2 \vec{\Psi}_i \cdot \vec{\Psi}_j \, d\mathbf{x} + \mathbf{i}\alpha \int_{\Gamma_a} \vec{\Psi}_i \times \mathbf{n} \times \mathbf{n} \cdot \vec{\Psi}_j \, ds.$$

Koska simulointialue Ω on ositettu luvun 5.2.3 esittämällä tavalla, tällöin pätee erityisesti $\Omega = \mathcal{T}_h$ ja $\Gamma_a = \mathcal{F}_h$, niin edellisen lausekkeen integraalit voidaan laskea summana integraaleista elementtien ja tahkojen yli

$$a_{i,j} = \sum_{k=1}^n \int_{C_k} \nabla \times \vec{\Psi}_i \cdot \nabla \times \vec{\Psi}_j - \alpha^2 \vec{\Psi}_i \cdot \vec{\Psi}_j \, d\mathbf{x} + \mathbf{i}\alpha \sum_{l=1}^m \int_{S_l} \vec{\Psi}_i \times \mathbf{n} \times \mathbf{n} \cdot \vec{\Psi}_j \, ds$$

$$C_k \in \mathcal{T}_h, S_l \in \mathcal{F}_h, \quad (5.21)$$

missä m on joukkoon \mathcal{F}_h sisältyvien tahkojen lukumäärä.

Globaalit kantafunktiot $\vec{\Psi}$ on määritelty elementeissä C_k ja tahkoilla S_l lokaalien elementtikohtaisten kantafunktioiden $\vec{\psi}$ avulla. Tällöin yhtälössä (5.21) globaalit kantafunktiot voidaan korvata lokaaleilla kantafunktioilla

$$a_{i,j} = \sum_{k=1}^n \int_{C_k} \nabla \times \vec{\psi}_i^k \cdot \nabla \times \vec{\psi}_j^k - \alpha^2 \vec{\psi}_i^k \cdot \vec{\psi}_j^k \, d\mathbf{x} + \mathbf{i}\alpha \sum_{l=1}^m \int_{S_l} \vec{\psi}_i^l \times \mathbf{n} \times \mathbf{n} \cdot \vec{\psi}_j^l \, ds$$

$$C_k \in \mathcal{T}_h, S_l \in \mathcal{F}_h. \quad (5.22)$$

On oleellista huomata, että vaikka kantafunktio $\vec{\psi}_i^k$ on globaalin kantafunktion $\vec{\Psi}_i$ määritelmä elementissä C_k , niin merkinnän $\vec{\psi}_i^k$ alaindeksi i ei viittaa globaaliin kantafunktioon. Funktion $\vec{\psi}_i^k$ alaindeksi on elementtikohtainen tunniste, joka voi tässä tutkielmassa saada arvot $1 \dots 12$, kun taas funktion $\vec{\Psi}_i$ alaindeksi voi saada arvot $1 \dots n$. Tässä erityistapauksessa globaalin kantafunktion alaindeksi on tunniste verkon \mathcal{T}_h särmälle ja lokaalin kantafunktion alaindeksi on tunniste elementin C_k särmälle.

Tutkitaan seuraavaksi, miten yhtälön (5.22) integraalit elementtien ja tahkojen yli lasketaan. Integraali yksittäisen elementin C_k yli halutaan laskea viite-elementin \hat{C} avulla, jolloin täytyy suorittaa koordinaattimuunnos

$$\int_{C_k} \nabla \times \vec{\psi}_i^k(\mathbf{x}) \cdot \nabla \times \vec{\psi}_j^k(\mathbf{x}) - \alpha^2 \vec{\psi}_i^k(\mathbf{x}) \cdot \vec{\psi}_j^k(\mathbf{x}) \, d\mathbf{x}$$

$$= \int_{\hat{C}} (\nabla \times \vec{\psi}_i^k \circ \mathbf{T}_k(\hat{\mathbf{x}}) \cdot \nabla \times \vec{\psi}_j^k \circ \mathbf{T}_k(\hat{\mathbf{x}}) - \alpha^2 \vec{\psi}_i^k \circ \mathbf{T}_k(\hat{\mathbf{x}}) \cdot \vec{\psi}_j^k \circ \mathbf{T}_k(\hat{\mathbf{x}})) |\det(D\mathbf{T}_k)| \, d\hat{\mathbf{x}}.$$

Kun edellisen lausekkeen kantafunktioihin sovelletaan muunnoskuvausta (5.14) ja kantafunktioiden roottoreihin kuvausta (5.15), saadaan

$$\int_{\hat{C}} \left(\frac{\mathbf{B}_k}{\det(\mathbf{B}_k)} \overline{\nabla \times \vec{\varphi}_i(\hat{\mathbf{x}})} \cdot \frac{\mathbf{B}_k}{\det(\mathbf{B}_k)} \overline{\nabla \times \vec{\varphi}_j(\hat{\mathbf{x}})} - \alpha^2 \mathbf{B}_k^{-T} \vec{\varphi}_i(\hat{\mathbf{x}}) \cdot \mathbf{B}_k^{-T} \vec{\varphi}_j(\hat{\mathbf{x}}) \right) \det(\mathbf{B}_k) d\hat{\mathbf{x}}. \quad (5.23)$$

Jakolasku $\mathbf{B}_k/\det(\mathbf{B}_k)$ on hyvin määritelty, koska luvussa 5.3 tehtiin oletetus, että $\det(\mathbf{B}_k) > 0$. Sovelletaan integraalilausekkeen (5.23) sisätuloihin laskusääntöä $\mathbf{B}\mathbf{u} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{B}^T \mathbf{w}$ [24, sivu 7]. Tällöin lopullinen lauseke yksittäisen elementin C_k integraalille viite-elementin \hat{C} avulla kirjoitettuna on

$$I_k^C(i, j) = \int_{\hat{C}} \frac{\mathbf{B}_k^T \mathbf{B}_k}{\det(\mathbf{B}_k)} \overline{\nabla \times \vec{\varphi}_i(\hat{\mathbf{x}})} \cdot \overline{\nabla \times \vec{\varphi}_j(\hat{\mathbf{x}})} - \alpha^2 \det(\mathbf{B}_k) (\mathbf{B}_k^T \mathbf{B}_k)^{-1} \vec{\varphi}_i(\hat{\mathbf{x}}) \cdot \vec{\varphi}_j(\hat{\mathbf{x}}) d\hat{\mathbf{x}}. \quad (5.24)$$

Yleinen käytäntö simulointitehtävissä on, että integraalilausekkeiden laskemiseen käytetään numeerista integrointia. Se tarkoittaa, että integraalilausekkeen tarkalle arvolle lasketaan arvio. Numeerisen integrointiyhtälön yleinen muoto on [17, sivu 52]

$$\int_K \mathbf{u}(\mathbf{x}) d\mathbf{x} \approx \sum_{p=1}^r w_p^K \mathbf{u}(\mathbf{x}_p^K), \quad (5.25)$$

missä painokertoimet $w_p^K \in \mathbb{R}$, integrointipisteet $\mathbf{x}_p^K \in K \subset \mathbb{R}^d$ ja integrointipisteiden lukumäärä $0 < r \in \mathbb{N}$ määräytyvät käytettävän numeerisen integrointimenetelmän mukaan. Lasketaan yhtälön (5.24) integraalilausekkeelle arvio soveltamalla numeerista integrointiyhtälöä (5.25)

$$I_k^C(i, j) \approx \tilde{I}_k^C(i, j) = \sum_{p=1}^r \hat{w}_p \left(\frac{\mathbf{B}_k^T \mathbf{B}_k}{\det(\mathbf{B}_k)} \overline{\nabla \times \vec{\varphi}_i(\hat{\mathbf{x}}_p)} \cdot \overline{\nabla \times \vec{\varphi}_j(\hat{\mathbf{x}}_p)} - \alpha^2 \det(\mathbf{B}_k) (\mathbf{B}_k^T \mathbf{B}_k)^{-1} \vec{\varphi}_i(\hat{\mathbf{x}}_p) \cdot \vec{\varphi}_j(\hat{\mathbf{x}}_p) \right). \quad (5.26)$$

Seuraavaksi määritellään miten yhtälössä (5.22) lasketaan integraali yksittäisen tahkon S_l yli. Integrointiyhtälön määrittäminen tahkolle suoritetaan samalla tavalla kuin edellä suoritettiin vastaavan elementtikohtaisen yhtälön määrittäminen. Tahkon integraaliyhtälö saadaan soveltamalla yhtälön (5.22) oikeanpuoleiseen integraalilausekkeeseen koordinaattimuunnosta ja kuvausta (5.14)

$$\begin{aligned} \int_{S_l} \vec{\psi}_i^l \times \mathbf{n} \times \mathbf{n} \cdot \vec{\psi}_j^l ds &= \int_{\hat{S}} (\mathbf{B}_l^{-T} \vec{\varphi}_i(\hat{\mathbf{x}}) \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \cdot \mathbf{B}_l^{-T} \vec{\varphi}_j(\hat{\mathbf{x}})) \det(\mathbf{B}_l) d\hat{\mathbf{s}} \\ &= \int_{\hat{S}} \det(\mathbf{B}_l) (\mathbf{B}_l^T \mathbf{B}_l)^{-1} \vec{\varphi}_i(\hat{\mathbf{x}}) \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \cdot \vec{\varphi}_j(\hat{\mathbf{x}}) d\hat{\mathbf{s}} = I_l^S(i, j). \end{aligned}$$

Saadulle integraalille lasketaan arvio numeerisen integrointiyhtälön (5.25) avulla

$$I_l^S(i, j) \approx \tilde{I}_l^S(i, j) = \sum_{p=1}^t \hat{w}_p \det(\mathbf{B}_l) (\mathbf{B}_l^T \mathbf{B}_l)^{-1} \vec{\varphi}_i(\hat{\mathbf{x}}_p) \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \times \mathbf{B}_l^{-T} \hat{\mathbf{n}} \cdot \vec{\varphi}_j(\hat{\mathbf{x}}_p). \quad (5.27)$$

Lopuksi määritellään lausekkeet vektorin \mathbf{f} alkioden laskemiseen. Funktio \mathbf{G} valitaan siten, että se poikkeaa nolasta ainoastaan sellaisten elementtien alueella, joiden tahko kuuluu heijastavalle reunalle Γ_s . Erityisesti oletetaan, että $\mathbf{G}(\mathbf{x}) \equiv 0$, $\mathbf{x} \in \Gamma_a$. Alkion f_i lauseke on oletuksen nojalla

$$f_i = \sum_{k=1}^n \int_{C_k} \nabla \times \mathbf{G} \cdot \nabla \times \vec{\psi}_i^k - \alpha^2 \mathbf{G} \cdot \vec{\psi}_i^k \, d\mathbf{x} \quad C_k \in \mathcal{T}_h. \quad (5.28)$$

Integraalilauseke yksittäisen elementin yli on

$$I_k^F(i) = \int_{\hat{C}} \mathbf{B}_k^T \nabla \times \mathbf{G} \circ \mathbf{T}_k(\hat{\mathbf{x}}) \cdot \overline{\nabla \times \vec{\varphi}_i(\hat{\mathbf{x}})} - \alpha^2 \det(\mathbf{B}_k) \mathbf{B}_k^{-1} \mathbf{G} \circ \mathbf{T}_k(\hat{\mathbf{x}}) \cdot \vec{\varphi}_i(\hat{\mathbf{x}}) \, d\hat{\mathbf{x}}$$

ja sitä voidaan arvioida soveltamalla numeerista integrointiyhtälöä

$$I_k^F(i) \approx \tilde{I}_k^F(i) = \sum_{p=1}^r \hat{w}_p (\mathbf{B}_k^T \nabla \times \mathbf{G} \circ \mathbf{T}_k(\hat{\mathbf{x}}_p) \cdot \overline{\nabla \times \vec{\varphi}_i(\hat{\mathbf{x}}_p)} - \alpha^2 \det(\mathbf{B}_k) \mathbf{B}_k^{-1} \mathbf{G} \circ \mathbf{T}_k(\hat{\mathbf{x}}_p) \cdot \vec{\varphi}_i(\hat{\mathbf{x}}_p)). \quad (5.29)$$

Nyt on määritelty kaikki lineaarisen yhtälöryhmän (5.20) kokoamiseen tarvittavat lausekkeet. Matriisin \mathbf{A} alkiot $a_{i,j}$ lasketaan yhtälöiden (5.22), (5.26) ja (5.27) avulla. Vektorin \mathbf{f} alkiot f_i lasketaan käyttäen yhtälöitä (5.28) ja (5.29). Näin muodostettu matriisi \mathbf{A} ei yleensä ole symmetrinen eikä positiivisesti definiitti.

Kertoimet $\{a_{i,j}\}_{i,j=1}^n$ ja $\{f_i\}_{i=1}^n$ tunnetaan. Yhtälöryhmässä on siis saman verran yhtälöitä kuin tuntemattomia kertoimia $\{q_i\}_{i=1}^n$. Yllä esitetty yhtälöryhmä voidaan kirjoittaa kompaktisti matriisimuodossa. Olkoon $A \in \mathbb{K}^{n \times n}$ neliömatriisi, jossa on n kappaletta rivejä ja sarakkeita. Matriisi A muodostuu kertoimista $\{a_{i,j}\}_{i,j=1}^n$. Vektorit $\mathbf{q}, \mathbf{f} \in \mathbb{K}^n$ koostuvat puolestaan kertoimista $\{q_i\}_{i=1}^n$ ja $\{f_i\}_{i=1}^n$. Vektori \mathbf{q} ratkaistaan nyt lineaarisesta yhtälöryhmästä, joka matriisimuotoon kirjoitettuna on

$$A\mathbf{q} = \mathbf{f}. \quad (6.1)$$

Vektori \mathbf{q} voidaan periaatteessa ratkaista lausekkeesta $\mathbf{q} = A^{-1}\mathbf{f}$ käänteismatriisin A^{-1} avulla. Käytännössä tilanne on kuitenkin se, että käänteismatriisin A^{-1} laskeminen suurten yhtälöryhmien kerroinmatriisille A on joitakin poikkeuksia lukuun ottamatta liian työlästä. Tällöin yhtälöryhmän (6.1) ratkaisemisessa joudutaan turvautumaan esimerkiksi iteratiivisiin ratkaisumenetelmiin. Seuraavaksi määritellään mitä iteratiivisilla menetelmillä tarkoitetaan ja esitellään menetelmiin liittyvät peruskäsitteet. Lähtien jostakin alkuarvauksesta \mathbf{q}_0 , iteratiivinen menetelmä muodostaa ratkaisuvektorille \mathbf{q} jonon approksimaatioita \mathbf{q}_j siten, että

$$\lim_{j \rightarrow \infty} \mathbf{q}_j = \mathbf{q}.$$

Toisin sanoen approksimaatiojono suppenee kohti ratkaisuvektoria.

Uuden approksimaation \mathbf{q}_j laskemista sanotaan iteraatioaskeleeksi ja sen yleinen lauseke on

$$\mathbf{q}_j = \mathbf{h}_j(\mathbf{q}_0, \dots, \mathbf{q}_{j-1}, A, \mathbf{f}). \quad (6.2)$$

Jatkossa vektoreiden alaindeksillä $j \in \mathbb{N}$ tarkoitetaan, että kyseinen vektori liittyy alaindeksin osoittamaan iteraatioaskeleeseen. Edellä iteraatioaskel määriteltiin funktiolla $\mathbf{h}_j : \mathbb{K}^{n \times (j+n+1)} \rightarrow \mathbb{K}^n$, joka voi olla hyvinkin monimutkainen ja jopa epälineaarinen. Jokaiseen iteraatioaskeleeseen liittyy virhe ja niin sanottu residuaali. Virhe on yksinkertaisesti kulloisenkin approksimaation poikkeama ratkaisuvektorista

$$\mathbf{e}_j = \mathbf{q}_j - \mathbf{q}$$

ja residuaaliksi määritellään

$$\mathbf{r}_j = A\mathbf{q}_j - \mathbf{f}.$$

Virheen ja residuaalin välillä on seuraavanlainen yhteys

$$A\mathbf{e}_j = \mathbf{r}_j. \quad (6.3)$$

Iteraatioaskeleelle esiteltiin yhtälössä (6.2) kokolailla abstrakti määritelmä. Luonteva tulkinta iteraatioaskeleelle saadaan käyttämällä *korjausvektoria* $\mathbf{d}_j \in \mathbb{K}^n$. Ajatuksena on, että jokaisella iteraatioaskeleella vektori \mathbf{d}_j määrittelee mihin suuntaan vanhaa

approksimaatiota \mathbf{q}_{j-1} korjataan. Tietysti halutaan, että korjauksen tuloksena syntyy uusi parempi approksimaatio ratkaisuvektorille \mathbf{q} . Edellä esitetty ajatus saadaan yhtälömuotoon kirjoittamalla

$$\mathbf{q}_j = \mathbf{q}_{j-1} + \mathbf{d}_j. \quad (6.4)$$

Uuteen approksimaation \mathbf{q}_j liittyvä virhe ja residuaali saadaan päivittämällä vanhaan approksimaation \mathbf{q}_{j-1} liittyvää virhettä ja residuaalia vektorin \mathbf{d}_j avulla

$$\begin{aligned} \mathbf{e}_j &= \mathbf{e}_{j-1} + \mathbf{d}_j \\ \mathbf{r}_j &= \mathbf{r}_{j-1} + \mathbf{A}\mathbf{d}_j. \end{aligned}$$

Askeleella j optimaalinen korjausvektori on

$$\mathbf{d}_j = -\mathbf{e}_{j-1},$$

koska silloin $\mathbf{e}_j = 0$ ja approksimaatio \mathbf{q}_j onkin itse asiassa yhtälöryhmän tarkka ratkaisu. Soveltamalla optimaalista korjausta yhtälöön (6.3) saadaan uusi lineaarinen yhtälöryhmä

$$\mathbf{A}\mathbf{d}_j = -\mathbf{r}_{j-1}, \quad (6.5)$$

jonka ratkaisuvektori on siis optimaalinen korjausvektori. On kuitenkin syytä olla tietoinen siitä, että vektorin \mathbf{d}_j laskeminen on yhtä vaikeaa kuin vektorin \mathbf{q} ratkaiseminen alkuperäisestä yhtälöryhmästä (6.1).

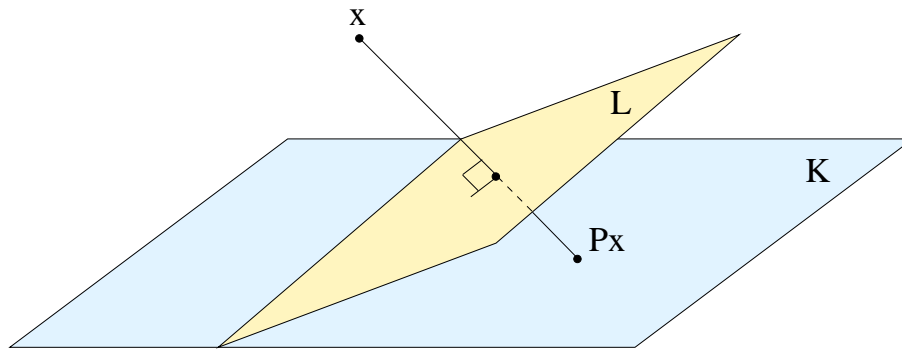
Lopuksi tehdään vielä yksi huomio. Kun asetetaan

$$\mathbf{d}_j = \mathbf{h}_j(\mathbf{q}_0, \dots, \mathbf{q}_{j-1}, \mathbf{A}, \mathbf{f}) - \mathbf{q}_{j-1},$$

on mahdollista havaita, että iteraatioaskeleen määritelmät (6.2) ja (6.4) ovat ilmaisuvoimaltaan saman veroisia ja molempien avulla voidaan kuvata kaikki iteratiiviset menetelmät. Artikkelin [26] luo katsauksen iteratiivisten menetelmien historiaan ja tarjoaa samalla monipuolisen lähdeluettelon iteratiivisista menetelmistä kiinnostuneille.

6.2 Projektiomenetelmät

Merkittävä osa nykyisistä suurten lineaaristen yhtälöryhmien ratkaisemiseen tarkoitettuista iteratiivisista menetelmistä, joilla on käytännön merkitystä, hyödyntävät tavalla tai toisella niin sanottua *projektiotekniikkaa*. Olkoon lineaarisen yhtälöryhmän (6.1) tarkka ratkaisu $\mathbf{q} \in \mathbb{K}^n$. Projektiotekniikan idea on, että ratkaisuvektorille \mathbf{q} muodostetaan approksimaatio $\bar{\mathbf{q}}$, joka kuuluu avaruuden \mathbb{K}^n johonkin toistaiseksi määrittelemättömään aliavaruuteen. Käytetään merkintää \mathcal{K} siitä aliavaruudesta, joka pitää sisällään kaikki approksimaatiokandidaatit. Avaruutta \mathcal{K} nimitetään joskus *hakuavaruudeksi* (*search supspace*).



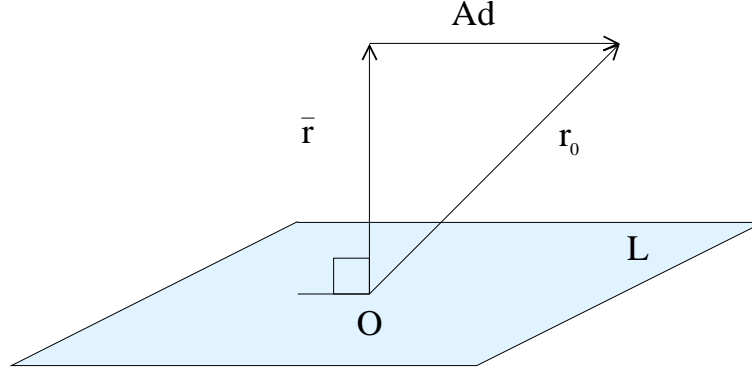
Kuva 6.1: Vektorin vinoprojektio aliavaruuteen \mathcal{K} .

Olkoon hakuavaruuden \mathcal{K} dimensio m . Sekaannusten välttämiseksi todettakoon, että avaruuden dimensiolla tarkoitetaan avaruuden kantavektoreiden lukumäärää. Nyt approksimaatiovektori on lineaarikombinaatio aliavaruuden \mathcal{K} kantavektoreista. Approksimaatiovektori on siis summa aliavaruuden kantavektoreista, missä jokainen kantavektori on kerrottu jollakin kertoimella. Jotta approksimaatiovektori voidaan yksikäsitteisesti valita approksimaatiokandidaattien joukosta, täytyy kantavektoreiden kertoimet pystyä määrittämään yksikäsitteisesti. Määrittely onnistuu, kunhan jokaista tuntematonta vastaa yksi rajoite. Näin ollen tarvitaan m kappaletta rajoitteita yksikäsitteisen approksimaatiovektorin määräämiseen.

Tyypillinen keino rajoitteiden muodostamiseen on asettaa m keskenään riippumatonta kohtisuoruusehtoa (*orthogonality conditions*). Erityisesti residuaalivektoria $A\mathbf{q} - \mathbf{f}$ voidaan rajoittaa, kun vaaditaan siltä kohtisuoruus m lineaarisesti riippumattomaa vektoria vastaan. Lineaarisesti riippumattomat vektorit muodostavat aliavaruuden $\mathcal{L} \subset \mathbb{K}^n$, jota kutsutaan usein *rajoiteavaruudeksi* (*subspace of constraints*). Residuaalivektorin rajoittaminen kohtisuoruusehdoilla on monille erilaisille matemaattisille menetelmille yhteinen tekniikka ja se tunnetaankin *Petrov-Galerkin* ehtona.

Projektiomenetelmille on olemassa kaksi tärkeää alaluokkaa: kohtisuorat ja vinot menetelmät. Kohtisuorissa menetelmissä rajoiteavaruus \mathcal{L} on täsmälleen hakuavaruus \mathcal{K} . Vinomenetelmissä \mathcal{L} ja \mathcal{K} ovat eri avaruuksia, eikä niillä välttämättä ole mitään tekemistä toistensa kanssa. Projektiomenetelmien jaottelu kahteen alaluokkaan, kohtisuoriin ja vinoihin menetelmiin, on perusteltua, koska alaluokat tuottavat erityyppisiä iterointialgoritmeja. Kuva 6.1 saattaa auttaa ymmärtämään, mistä nimi projektimenetelmä juontaa juurensa. Kuvassa vektori \mathbf{x} on projisoitu aliavaruuteen \mathcal{K} siten, että vektori $\mathbf{x} - P\mathbf{x}$ on kohtisuorassa avaruutta \mathcal{L} vastaan.

Seuraavaksi tutkitaan yleisellä tasolla miten projektioaskel suoritetaan. Esitetyt asiat pätevät sekä kohtisuorille että vinoille menetelmille. Projektimenetelmissä etsitään siis lineaarisen yhtälöryhmän tarkalle ratkaisulle approksimaatiota $\bar{\mathbf{q}}$, joka kuuluu hakuavaruuteen \mathcal{K} . Lisäksi approksimaatioon liittyvän residuaalin $\bar{\mathbf{r}}$ tulee olla kohti-



Kuva 6.2: Residuaalivektorin kohtisuoruus avaruutta \mathcal{L} vastaan.

suorassa rajoiteavaruutta \mathcal{L} vastaan. Matemaattisin merkinnöin projektiioaskeleen formulointi on:

$$\text{Etsi } \bar{\mathbf{q}} \in \mathcal{K} \text{ siten, että } \bar{\mathbf{r}} = A\bar{\mathbf{q}} - \mathbf{f} \perp \mathcal{L}.$$

Mikäli projektiomenetelmissä halutaan hyödyntää alkuarvausta \mathbf{q}_0 , täytyy approksimaatiovektoria etsiä affiinista aliavaruudesta $\mathbf{q}_0 + \mathcal{K}$. Tällöin projektiioaskeleen uusi formulointi on:

$$\text{Etsi } \bar{\mathbf{q}} \in \mathbf{q}_0 + \mathcal{K} \text{ siten, että } \bar{\mathbf{r}} = A\bar{\mathbf{q}} - \mathbf{f} \perp \mathcal{L}. \quad (6.6)$$

Oletetaan, että ratkaisuvektorin approksimaatio saadaan alkuarvauksesta $\bar{\mathbf{q}} = \mathbf{q}_0 + \mathbf{d}$, kun $\mathbf{d} \in \mathcal{K}$. Lisäksi oletetaan, että alkuperäiseksi residuaaliksi on määritetty $\mathbf{r}_0 = A\mathbf{q}_0 - \mathbf{f}$. Projektiioaskeleen formulointiin liittyvä kohtisuoruusehto on nyt

$$A(\mathbf{q}_0 + \mathbf{d}) - \mathbf{f} \perp \mathcal{L} \quad \Leftrightarrow \quad A\mathbf{d} + \mathbf{r}_0 \perp \mathcal{L}. \quad (6.7)$$

Toisin sanoen approksimaatiovektoriksi voidaan määritellä

$$\begin{aligned} \bar{\mathbf{q}} &= \mathbf{q}_0 + \mathbf{d}, \quad \mathbf{d} \in \mathcal{K}, \\ (A\mathbf{d} + \mathbf{r}_0, \mathbf{w}) &= 0, \quad \forall \mathbf{w} \in \mathcal{L}. \end{aligned}$$

Approksimaatiovektoriin liittyvän residuaalin $\bar{\mathbf{r}} = A\mathbf{d} + \mathbf{r}_0$ kohtisuoruusehtoa on havainnollistettu kuvassa 6.2.

Yllä on esitelty projektiioaskeleen perusteet yleisessä muodossaan. Suurin osa tavallisimmin käytetyistä menetelmistä suorittaa projektiioaskeleita perätysten useita kertoja. Tyypillisesti uusi projektiioaskel käyttää uutta aliavaruusparia \mathcal{K} ja \mathcal{L} , alkuarvauksena \mathbf{q}_0 käytetään edellisellä askeleella laskettua approksimaatiota. Projektiomenetelmät muodostavat yhtenäisen teorian monille numeerisen laskennan menetelmille. Aina kun approksimaatio määritellään käyttäen m kappaletta vapausasteita (aliavaruus \mathcal{K}) ja m kappaletta rajoitteita (aliavaruus \mathcal{L}), kyseessä on projektiotekniikka.

6.2.1 Projektiomenetelmien luokittelu

Tässä luvussa tutkitaan yksityiskohtaisemmin projektiomenetelmiä ja esitellään määritelmä, jonka avulla projektiomenetelmiä voidaan luokitella. Raportissa [31] ja artikkelissa [32], joihin tämä luku perustuu, puhutaan ortogonalisointimenetelmistä (*orthogonalization methods*). Kyseiset menetelmät vastaavat kuitenkin projektiomenetelmiä ja tässä yhteydessä käytetäänkin jälkimmäistä nimitystä.

Iteraatioaskeleeseen j liittyvä hakuavaruus olkoon $\mathcal{K}_j = \text{span}(\mathbf{v}_{j-\sigma_j, j}, \dots, \mathbf{v}_{j-1, j})$, eli kantavektorit $\mathbf{v}_{j-i, j}$ virittävät aliavaruuden \mathcal{K}_j , kun $i = 1, \dots, \sigma_j$. Seuraavaksi esitellään eräs projektiomenetelmien määritelmä, joka kattaa suuren joukon iteratiivisia menetelmiä:

Määritelmä 6.2.1 *Olkoon \mathbf{q}_0 alkuarvaus ja $\mathbf{r}_0 = \mathbf{A}\mathbf{q}_0 - \mathbf{f}$ alkuarvausta vastaava residuaali. Mikäli iteraatioaskel suoritetaan niin, että alla luetellavat ominaisuudet pätevät, iteratiivinen menetelmä kuuluu projektiomenetelmien joukkoon. Askeleella $j \geq 1$ lasketaan ratkaisuvektorin \mathbf{q} approksimaatio \mathbf{q}_j siten, että*

$$\mathbf{q}_j \in \tilde{\mathbf{q}}_j + \text{span}(\mathbf{v}_{j-\sigma_j, j}, \dots, \mathbf{v}_{j-1, j}), \quad (6.8)$$

missä $\mathbf{v}_{j-i, j} \in \mathbb{K}^n$ ja $\tilde{\mathbf{q}}_j \in \text{span}(\mathbf{q}_{j-\sigma_j}, \dots, \mathbf{q}_{j-1})$. Lisäksi approksimaatioon liittyvälle residuaalille \mathbf{r}_j tulee olla voimassa kohtisuoruusehto

$$\mathbf{r}_j^T \mathbf{Z}_j \mathbf{v}_{j-i, j} = 0, \quad (6.9)$$

kun $i = 1, \dots, \sigma_j$. Matriisit \mathbf{Z}_j eivät ole singulaarisia ja niitä kutsutaan apu- tai tukimatriiseiksi. Projektiomenetelmää sanotaan

- tarkaksi, jos $\sigma_j = j$, eli uuden iterantin laskemiseen käytetään j kappaletta vektoreita $\mathbf{v}_{j-i, j}$.
- uudelleen käynnistyväksi, jos $\sigma_j = (j - 1) \bmod \sigma_{res}$, eli uudelleen käynnistys suoritetaan aina, kun on edetty σ_{res} askelta edellisestä käynnistyksestä.
- typistetyksi, jos $\sigma_j = \min(j, \sigma_{max})$, missä σ_{max} on kiinnitetty. Tällöin käytetään korkeintaan σ_{max} kappaletta vektoreita $\mathbf{v}_{j-i, j}$ uuden iterantin laskemiseen.
- yhdistetyksi, jos typistetty menetelmä käynnistetään jaksoittain uudelleen. Toisin sanoen $\sigma_j = \min((j - 1) \bmod \sigma_{res} + 1, \sigma_{max})$.

Koska määritelmä (6.2.1) sisältää runsaasti parametreja, on se luonteeltaan tekninen. Parametrisoinnilla on pyritty siihen, että määritelmän avulla voidaan kattavasti luokitella projektiomenetelmiä. Määritelmän teknisen ulkokuoren alta paljastuu kuitenkin jo entuudestaan tuttuja asioita. Esimerkiksi jos matriisi $\mathbf{Z}_j \in \mathcal{K}^{n \times n}$ ajatellaan

lineaarikuvaukseksi hakuavaruudelta \mathcal{K}_j rajoiteavaruudelle \mathcal{L}_j , eli Z_j kuvaa esimerkiksi kantavektorit $\mathbf{v}_{j-i,j}$ aliavaruuteen \mathcal{L}_j , niin kohtisuoruusehto (6.9) vastaa tismalleen ehtoa (6.7).

Todetaan, että yleisessä tapauksessa affiini siirtovektori on joko $\tilde{\mathbf{q}}_j = \mathbf{q}_{j-\sigma_j}$ tai $\tilde{\mathbf{q}}_j = \mathbf{q}_{j-1}$. Otetaan käyttöön lyhennemerkinnyt

$$\begin{aligned}\tilde{\mathbf{e}}_j &= \tilde{\mathbf{q}}_j - \mathbf{q}, \\ \tilde{\mathbf{r}}_j &= A\tilde{\mathbf{q}}_j - \mathbf{f}.\end{aligned}$$

Koska aliavaruuden \mathcal{K}_j virittävät vektorit tunnetaan, niin approksimaatiolle \mathbf{q}_j saadaan lauseke

$$\mathbf{q}_j = \sum_{i=1}^{\sigma_j} y_{i,j} \mathbf{v}_{j-i,j} + \tilde{\mathbf{q}}_j. \quad (6.10)$$

Merkinnällä $y_{i,j}$ ei tarkoiteta matriisin alkioita, vaan iteraatioaskeleella j muodostettavan lineaarikombinaation kerrointa i .

Edelleen käyttämällä yhtälöä (6.10), juuri esitettyjä lyhennemerkinntöjä sekä residuaalin ja virheen alkuperäisiä määritelmiä, saadaan tuloksena

$$\mathbf{e}_j = \sum_{i=1}^{\sigma_j} y_{i,j} \mathbf{v}_{j-i,j} + \tilde{\mathbf{e}}_j, \quad (6.11)$$

$$\mathbf{r}_j = \sum_{i=1}^{\sigma_j} y_{i,j} A \mathbf{v}_{j-i,j} + \tilde{\mathbf{r}}_j. \quad (6.12)$$

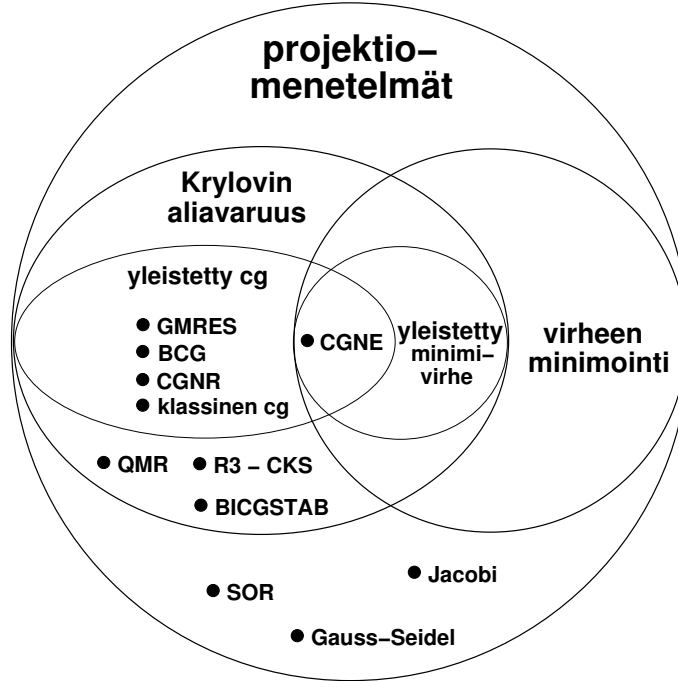
Kun sijoitetaan residuaalilauseke (6.12) ehtoon (6.9), syntyy joukko yhtälöitä

$$\sum_{i=1}^{\sigma_k} y_{i,j} \mathbf{v}_{j-i,j}^T A^T Z_j \mathbf{v}_{j-k,j} = -\tilde{\mathbf{r}}_j^T Z_j \mathbf{v}_{j-k,j}, \quad (6.13)$$

missä $k = 1, \dots, \sigma_j$. Kyseessä on itse asiassa lineaarinen yhtälöryhmä, josta voidaan ratkaista kertoimet $y_{i,j}$. Erilaisia projektiomenetelmiä muodostetaan määrittelemällä parametrit $\mathbf{v}_{j-i,j}$, $\tilde{\mathbf{q}}_j$, Z_j ja σ_j . Raportissa [31, sivu 7] on esitelty taulukko, josta löytyvät monien tunnettujen iteratiivisten menetelmien käyttämät parametrit. Kuvassa 6.3 on osittainen luokittelu projektiomenetelmille.

Tarkat ortogonalisointimenetelmät voidaan esittää yksinkertaisemmin matriisimuodossa. Muistetaan, että tarkoissa menetelmissä $\sigma_j = j$ ja että jokaisella iteraatiokierroksella approksimaatiovektori on lineaarikombinaatio aliavaruuden \mathcal{K}_j kantavektoreista. Lineaarikombinaatioon liittyvistä kantavektoreiden kertoimista käytetään merkitään $y_{1,j}, \dots, y_{j,j}$. Olkoon

$$\begin{aligned}\mathbf{V}_j &= (\mathbf{v}_{0,j}, \dots, \mathbf{v}_{j-1,j}) \in \mathbb{K}^{n \times j}, \\ \mathbf{y}_j &= (y_{j,j}, \dots, y_{1,j})^T \in \mathbb{K}^n,\end{aligned}$$



Kuva 6.3: Projektimenetelmien osittainen luokittelukaavio.

eli matriisiin V_j sarakkeet ovat kulloisenkin aliavaruuden kantavektoreita ja vektorin \mathbf{y}_j alkioina ovat kantavektoreiden lineaarikombinaatioon liittyvät kertoimet.

Nyt yhtälöt (6.10), (6.11) ja (6.12) ovat ekvivalentteja seuraavien yhtälöiden kanssa

$$\begin{aligned}\mathbf{q}_j &= V_j \mathbf{y}_j + \tilde{\mathbf{q}}_j, \\ \mathbf{e}_j &= V_j \mathbf{y}_j + \tilde{\mathbf{e}}_j, \\ \mathbf{r}_j &= AV_j \mathbf{y}_j + \tilde{\mathbf{r}}_j.\end{aligned}$$

Edelleen kohtisuoruusehto (6.9) on matriisimuodossa

$$\mathbf{r}_j^T Z_j V_j = 0$$

ja lineaarinen yhtälöryhmä (6.13) saadaan muotoon

$$V_j^T Z_j^T AV_j \mathbf{y}_j = -V_j^T Z_j^T \tilde{\mathbf{r}}_j. \quad (6.14)$$

6.2.2 Krylovin aliavaruusmenetelmät

Projektimenetelmien ehkä tärkein osajoukko muodostuu Krylovin aliavaruusmenetelmistä (*Krylov subspace methods*), joita jatkossa nimitetään lyhyemmin Krylovin menetelmiksi. Nykypäivänä yleisesti käytettyjen Krylovin menetelmien suosio on lähtöisin siitä, että menetelmät perustuvat yksinkertaisiin matriisi-vektori kertolaskuihin. Aliavaruutta sanotaan Krylovin aliavaruudeksi, jos se voidaan parametrisoida seuraavalla

tavalla

$$\mathcal{K}_j(\mathbf{B}, \mathbf{v}_1) = \text{span}(\mathbf{v}_1, \mathbf{B}\mathbf{v}_1, \mathbf{B}^2\mathbf{v}_1, \dots, \mathbf{B}^{j-1}\mathbf{v}_1),$$

missä $\mathbf{B} \in \mathbb{K}^{n \times n}$ ja $\mathbf{v}_1 \in \mathbb{K}^n$. Krylovin menetelmissä hakuavaruuden kantavektoreille ja affinille siirtovektorille pätee

$$\mathbf{v}_{j-i,j} \in \mathcal{K}_j(\mathbf{B}, \mathbf{v}_1), \quad \tilde{\mathbf{q}}_j = \mathbf{q}_{j-\sigma_j},$$

missä $i = 1, \dots, \sigma_j$. Mikäli $\mathbf{B} \neq \mathbf{A}$, joudutaan jokaisella iteraatioaskeleella suorittamaan kaksi matriisi-vektori kertolaskua: yksi matriisilla \mathbf{B} kantavektorin $\mathbf{v}_{j-i,j}$ selvittämiseksi ja yksi matriisilla \mathbf{A} residuaalin \mathbf{r}_j laskemiseksi. Jos $\mathbf{B} = \mathbf{A}$, niin yleisessä tapauksessa tarvitaan vain yksi matriisi-vektori kertolasku per iteraatioaskel.

Krylovin liittomenetelmissä (*Conjugate Krylov supspace methods*) asetetaan $\mathbf{B} = \mathbf{A}$ ja $\mathbf{v}_1 = \mathbf{r}_0$, jolloin

$$\mathbf{v}_{j-i,j} \in \mathcal{K}_j(\mathbf{A}, \mathbf{r}_0).$$

Nyt approksimaatiovektorille sekä siihen liittyvälle virheelle ja residuaalille pätee

$$\mathbf{q}_j \in \mathbf{q}_0 + \mathcal{K}_j(\mathbf{A}, \mathbf{r}_0), \quad (6.15)$$

$$\mathbf{e}_j \in \mathbf{e}_0 + \mathbf{A}\mathcal{K}_j(\mathbf{A}, \mathbf{e}_0), \quad (6.16)$$

$$\mathbf{r}_j \in \mathbf{r}_0 + \mathbf{A}\mathcal{K}_j(\mathbf{A}, \mathbf{r}_0). \quad (6.17)$$

Tarkoille Krylovin liittomenetelmille asetetaan $\mathbf{v}_{j-i,j} = \mathbf{A}^{j-i}\mathbf{r}_0$, jolloin kohtisuoruusehtoa (6.9) vastaa yhtälö

$$\mathbf{r}_j^T \mathbf{Z}_j \mathbf{A}^{j-i} \mathbf{r}_0 = 0,$$

missä $i = 1, \dots, j$. Asettamalla $\mathbf{Z}_j = \mathbf{Z} = \text{vakio}$ saadaan Krylovin liittomenetelmien aliluokka, *yleistetyt gradienttimenetelmät* (*Generalized conjugate gradient methods*).

Krylovin menetelmiä voidaan tarkastella myös toisesta näkökulmasta. Olkoon polynomifunktioiden avaruus

$$\mathcal{P}_n = \{\phi(\lambda) = \sigma_0 + \sigma_1\lambda + \dots + \sigma_n\lambda^n \mid \sigma_0, \sigma_1, \dots, \sigma_n \in \mathbb{K}\}.$$

Residuaalin \mathbf{r}_0 ja matriisin \mathbf{A} virittämä Krylovin aliavaruus on polynomifunktioiden avulla kirjoitettuna

$$\mathcal{K}_j(\mathbf{A}, \mathbf{r}_0) = \{\phi(\mathbf{A})\mathbf{r}_0 \mid \phi \in \mathcal{P}_{j-1}\}$$

ja tästä syystä menetelmiä, jonka approksimaatiovektorit ovat muotoa (6.15), sanotaankin usein polynomimenetelmiksi. Polynomimenetelmien näkökulmasta iteraatioaskeleella j approksimaatiovektori on $\mathbf{q}_j = \mathbf{q}_0 + \phi_{j-1}(\mathbf{A})\mathbf{r}_0$, $\phi_{j-1} \in \mathcal{P}_{j-1}$. Erityisesti approksimaatioon liittyvälle residuaalille on olemassa polynomimääritelmä

$$\mathbf{r}_j = \mathbf{A}\mathbf{q}_j - \mathbf{f} = \psi_j(\mathbf{A})\mathbf{r}_0,$$

missä $\psi_j \in \mathcal{P}_j$ ja $\psi_j(0) = 1$.

Krylovin menetelmissä on ajatuksena, että jokaisella iteraatioaskeleella minimoidaan approksimaatioon liittyvä residuaali. Polynomimenetelmien näkökulmasta ongelmanna on löytää se polynomi, joka minimoi kulloisenkin residuaalin. Tässä kappaleessa seuraavaksi esitettävät asiat perustuvat julkaisuun [9, sivu 77]. Oletetaan, että matriisilla A on ortonormaali joukko ominaisvektoreita $\{\mathbf{v}_i\}_{i=1}^n$. Ominaisvektoreita vastaa joukko ominaisarvoja $\{\lambda_1, \dots, \lambda_n\} = \sigma(A)$. Jos alkuarvausta vastaava residuaali on $\mathbf{r}_0 = \sum_{i=1}^n \beta_i \mathbf{v}_i$, niin askelta j vastaava residuaali on $\mathbf{r}_j = \sum_{i=1}^n \beta_i \psi_j(\lambda_i) \mathbf{v}_i$. Tällöin residuaalin normille pätee

$$\|\mathbf{r}_j\| = \left(\sum_{i=1}^n \beta_i^2 \psi_j(\lambda_i)^2 \right)^{1/2} \leq \max_{\lambda \in \sigma(A)} |\psi_j(\lambda)| \|\mathbf{r}_0\|.$$

Eli residuaalin \mathbf{r}_j normi on pieni, mikäli residuaaliin liittyvän polynomien itseisarvo $|\psi_j|$ on pieni kaikilla matriisin A ominaisarvoilla.

6.3 GMRES

Vuonna 1986 julkaistu *Saadin* ja *Schultzin* kirjoittama artikkeli [25] toi GMRES menetelmän alan harrastajien ja tutkijoiden tietoisuuteen. GMRES on iteratiivinen menetelmä lineaaristen yhtälöryhmien ratkaisemiseen ja se kuuluu yleistettyjen gradienttimenetelmien joukkoon; tällöin se on myös projektiomenetelmä ja Krylovin aliavaruusmenetelmä. GMRES menetelmässä asetetaan $\mathcal{K} = \mathcal{K}_m$ ja $\mathcal{L} = A\mathcal{K}_m$, missä \mathcal{K}_m on järjestyksessään m . Krylovin aliavaruus valinnoin $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|$ ja $B = A$. GMRES on nykypäivänä yleisesti käytetty menetelmä, koska se soveltuu myös sellaisten yhtälöryhmien ratkaisemiseen, joissa kerroinmatriisi ei välttämättä ole symmetrinen tai hermiittinen eikä positiivisesti definiitti.

6.3.1 Arnoldin menetelmä

GMRES pohjautuu *Arnoldin menetelmään*, joka esiteltiin jo vuonna 1951. Arnoldin menetelmä perustuu edelleen *Gram-Schmidt* menetelmään, jolla muodostetaan lineaarisesti riippumattomasta vektorijoukosta ortonormaali vektorijoukko. Vektorijoukko $\{\mathbf{v}_i\}_{i=1}^n$ on ortonormaali, jos sille pätee

$$(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij},$$

missä δ_{ij} on Kroneckerin symboli. Arnoldin menetelmä muodostaa Krylovin aliavaruudelle $\mathcal{K}_m(A, \mathbf{v}_1) = \text{span}(\mathbf{v}_1, A\mathbf{v}_1, A^2\mathbf{v}_1, \dots, A^{m-1}\mathbf{v}_1)$ ortonormaalin kannan. Perusalgoritmi Arnoldin menetelmälle on

Algoritmi 6.3.1 : Arnoldi.

1. Valitse aloitusvektori $\mathbf{v}_1 \in \mathbb{K}^n$ siten, että $\|\mathbf{v}_1\| = 1$.

2. Suorita iteroimalla $j = 1, \dots, m$:
3. Laske jokaiselle $i = 1, \dots, j$: $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$.
4. $\tilde{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^j h_{i,j}\mathbf{v}_i$,
5. Jos $\|\tilde{\mathbf{v}}_{j+1}\| = 0$, lopeta.
6. $h_{j+1,j} = \|\tilde{\mathbf{v}}_{j+1}\|$.
7. $\mathbf{v}_{j+1} = \tilde{\mathbf{v}}_{j+1}/h_{j+1,j}$,
8. lopeta iterointi.

Jokaisella iteraatiokierroksella algoritmi kertoo edellisen Arnoldin vektorin \mathbf{v}_j matriisilla \mathbf{A} ja ortonormalisoi tuloksena saadun vektorin kaikkia aikaisempia \mathbf{v}_i vektoreita vastaan. Käytetään merkintää $\mathbf{V}_m \in \mathbb{K}^{n \times m}$ matriisista, jonka sarakkeina ovat algoritmin tuottamat kantavektorit, eli $\mathbf{V}_m = (\mathbf{v}_1, \dots, \mathbf{v}_m)$. Matriisi $\mathbf{H}_m \in \mathbb{K}^{m \times m}$ olkoon puolestaan sellainen, että sen ainoat nollasta eroavat alkiot ovat $h_{i,j}$. Matriisista \mathbf{H}_m muodostuu niin sanottu *Hessenberg*-matriisi, koska sille pätee $h_{i,j} = 0$, kun $i > j + 1$. Kun Arnoldin algoritmi on suoritettu onnistuneesti, pätee yhtäsuuruus

$$\mathbf{H}_m = \mathbf{V}_m^T \mathbf{A} \mathbf{V}_m.$$

6.3.2 Algoritmi GMRES menetelmälle

Seuraavaksi esitellään algoritmi GMRES menetelmälle. Algoritmi voidaan johtaa kahdella tapaa. Toinen keinoista on käyttää kohtisuoruusehtoa (6.14), missä $\mathbf{Z} = \mathbf{A}$. Näin ei kuitenkaan tehdä tässä tutkielmassa, vaan algoritmi johdetaan suoraan residuaalivektorin minimointitehtävästä. GMRES hyödyntää Arnoldin menetelmää muodostaessaan ortonormaalia kantaa aliavaruudelle $\mathcal{K}_m(\mathbf{A}, \mathbf{r}_0)$. Aluksi tehdään huomio, että suorittamalla Arnoldin algoritmia j iteraatioaskelta, saadaan muodostettua ortonormaali kanta \mathbf{V}_{j+1} . Samalla on muodostettu matriisi $\tilde{\mathbf{H}}_j \in \mathbb{K}^{(j+1) \times j}$, jonka ainoat nollasta eroavat alkiot ovat $h_{i,j}$. Matriisi $\tilde{\mathbf{H}}_j$ vastaa aikaisemmin määriteltyä matriisia \mathbf{H}_j sillä lisäyksellä, että $\tilde{\mathbf{H}}_j$ sisältää yhden uuden rivin, jonka ainoa nollasta eroava alkio on $h_{j+1,j}$. Muodostetut matriisit toteuttavat yhtälön

$$\mathbf{A}\mathbf{V}_j = \mathbf{V}_{j+1}\tilde{\mathbf{H}}_j. \quad (6.18)$$

GMRES menetelmän ajatuksena on, että jokaisella iteraatioaskeleella approksimaatiovektori valitaan siten, että siihen liittyvä residuaali on minimaalinen. Aikaisempien lukujen perusteella tiedetään, että Krylovin liittomenetelmissä approksimaatiovektori, joka liittyy iteraatioaskeleeseen j , on muotoa $\mathbf{q}_j = \mathbf{q}_0 + \mathbf{d}$, $\mathbf{d} \in \mathcal{K}_j$. Pienimmän mahdollisen residuaalin saavuttamiseksi täytyy jokaisella iteraatioaskeleella ratkaista minimointitehtävä

$$\min_{\mathbf{d} \in \mathcal{K}_j} \|\mathbf{r}_j\| = \min_{\mathbf{d} \in \mathcal{K}_j} \|\mathbf{A}(\mathbf{q}_0 + \mathbf{d}) - \mathbf{f}\| = \min_{\mathbf{d} \in \mathcal{K}_j} \|\mathbf{A}\mathbf{d} - \mathbf{r}_0\|.$$

Optimaalinen korjausvektori \mathbf{d} lasketaan käyttämällä Arnoldin menetelmällä muodostettua aliavaruuden \mathcal{K}_j kantaa V_j siten, että $\mathbf{d} = V_j \mathbf{y}$. Tällöin residuaalin miniminormi saadaan jokaisella iteraatioaskeleella minimoimalla seuraava vektorista \mathbf{y} riippuva funktio:

$$J(\mathbf{y}) = \|\mathbf{A}V_j \mathbf{y} - \beta \mathbf{v}_1\|. \quad (6.19)$$

Yhtälössä (6.19) on käytetty merkintää $\beta = \|\mathbf{r}_0\|$. On oleellista huomata, että optimaalisen vektorin $\mathbf{y} \in \mathbb{K}^j$ ratkaiseminen on pienimmän neliösumman tehtävä, koska tuntemattomia on vähemmän kuin yhtälöitä. Seuraavaksi sovelletaan yhtälöä (6.18) funktioon (6.19)

$$J(\mathbf{y}) = \|\mathbf{V}_{j+1}(\tilde{\mathbf{H}}_j \mathbf{y} - \beta \mathbf{e}_1)\|. \quad (6.20)$$

Vektorilla \mathbf{e}_1 tarkoitetaan $(j+1) \times (j+1)$ dimensioisen identtisen matriisin ensimmäistä saraketta, jolloin $V_{j+1} \mathbf{e}_1 = \mathbf{v}_1$. Yhtälöä (6.20) voidaan vieläkin sieventää. Todetaan, että matriisi V_{j+1} on ortogonaalinen ja että vektorin pituus säilyy ennallaan, kun sitä kerrotaan ortogonaalisella matriisilla. Juuri esitetyillä perusteluilla yhtälö (6.20) sievenee muotoon

$$J(\mathbf{y}) = \|\tilde{\mathbf{H}}_j \mathbf{y} - \beta \mathbf{e}_1\|. \quad (6.21)$$

Approksimaatiovektori \mathbf{q}_j , joka minimoi residuaalin \mathbf{r}_j iteraatioaskeleella j , on

$$\mathbf{q}_j = \mathbf{q}_0 + V_j \mathbf{y}_j,$$

missä \mathbf{y}_j minimoi funktion (6.21) yli kaikkien $\mathbf{y} \in \mathbb{K}^j$. Ensimmäinen versio GMRES menetelmän algoritmista on

Algoritmi 6.3.2 : GMRES.

1. Valitse alkuarvaus \mathbf{q}_0 . Laske $\mathbf{r}_0 = \mathbf{A}\mathbf{q}_0 - \mathbf{f}$ ja $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|$.
2. Suorita iteroimalla $j = 1, \dots, m$:
 Laske jokaiselle $i = 1, \dots, j$: $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$.
 $\tilde{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^j h_{i,j} \mathbf{v}_i$,
 Jos $\|\tilde{\mathbf{v}}_{j+1}\| = 0$, lopeta.
 $h_{j+1,j} = \|\tilde{\mathbf{v}}_{j+1}\|$.
 $\mathbf{v}_{j+1} = \tilde{\mathbf{v}}_{j+1}/h_{j+1,j}$.
3. Muodosta approksimaatiovektori:
 Laske $\mathbf{q}_m = \mathbf{q}_0 + V_m \mathbf{y}_m$, missä \mathbf{y}_m minimoi funktion $\|\tilde{\mathbf{H}}_m \mathbf{y} - \beta \mathbf{e}_1\|$, $\mathbf{y} \in \mathbb{K}^m$.

Minimoivan vektorin \mathbf{y}_m laskeminen on suhteellisen helppoa, kun m on pieni. Vektorin laskeminen edellyttää $(m+1) \times m$ dimensioisen pienimmän neliösummatehtävän

$$\tilde{\mathbf{H}}_j \mathbf{y}_m = \beta \mathbf{e}_1 \quad (6.22)$$

ratkaisemista. Algoritmin akilleen kantapää on se, että jokaisella iteraatioaskeleella algoritmin resurssivaatimukset kasvavat merkittävästi. Talletettavien vektorien lukumäärä on suorassa suhteessa iteraatioaskeleeseen j ja algoritmissa suoritettavien laskutoimitusten lukumäärä kasvaa neliöllisesti iteraatioaskeleiden edetessä $\frac{1}{2}j^2n$. Eräs keino ongelman kiertämiseen on uudelleen käynnistää algoritmi aina, kun m kappaletta iteraatioaskeleita on suoritettu. Kokonaislukuparametri m kiinnitetään ennen algoritmin aloitusta. Uudelleen käynnistettävä versio GMRES menetelmän algoritmista on

Algoritmi 6.3.3 : GMRES(m).

1. Valitse alkuarvaus \mathbf{q}_0 . Laske $\mathbf{r}_0 = \mathbf{A}\mathbf{q}_0 - \mathbf{f}$ ja $\mathbf{v}_1 = \mathbf{r}_0/\|\mathbf{r}_0\|$.
2. Suorita iteroimalla $j = 1, \dots, m$:
 Laske jokaiselle $i = 1, \dots, j$: $h_{i,j} = (\mathbf{A}\mathbf{v}_j, \mathbf{v}_i)$.
 $\tilde{\mathbf{v}}_{j+1} = \mathbf{A}\mathbf{v}_j - \sum_{i=1}^j h_{i,j}\mathbf{v}_i$,
 Jos $\|\tilde{\mathbf{v}}_{j+1}\| = 0$, lopeta.
 $h_{j+1,j} = \|\tilde{\mathbf{v}}_{j+1}\|$.
 $\mathbf{v}_{j+1} = \tilde{\mathbf{v}}_{j+1}/h_{j+1,j}$.
3. Muodosta approksimaatiovektori:
 Laske $\mathbf{q}_m = \mathbf{q}_0 + \mathbf{V}_m\mathbf{y}_m$, missä \mathbf{y}_m minimoi funktion $\|\tilde{\mathbf{H}}_m\mathbf{y} - \beta\mathbf{e}_1\|$, $\mathbf{y} \in \mathbb{K}^m$.
4. Uudelleen käynnistys:
 Laske $\mathbf{r}_m = \mathbf{A}\mathbf{q}_m - \mathbf{f}$; jos residuaali on kelvollinen, niin lopeta.
 Muutoin aseta $\mathbf{q}_0 := \mathbf{q}_m$, $\mathbf{v}_1 = \mathbf{r}_m/\|\mathbf{r}_m\|$ ja jatka kohdasta 2.

Uudelleen käynnistettävän GMRES algoritmin heikkous on puolestaan se, että sille ei voida taata approksimaatioiden suppenemista kohti ratkaisuvektoria. Algoritmi voi jäädä ikään kuin kävelemään paikalleen.

6.3.3 Algoritmin käytännöllinen toteutus

Seuraavaksi käsitellään joitakin tärkeitä yksityiskohtia liittyen edellisessä luvussa esitettyjen algoritmien toteutukseen. Tarkastellaan tilannetta, jossa halutaan ratkaista pienimmän neliösumman tehtävä

$$\min_{\mathbf{y}} \|\tilde{\mathbf{H}}_m\mathbf{y} - \beta\mathbf{e}_1\|.$$

Klassinen keino tehtävän ratkaisemiseen on muodostaa matriisille $\tilde{\mathbf{H}}_m$ niin sanottu QR-hajotelma $\mathbf{Q}_m\mathbf{R}_m$ käyttäen *Givens*-muunnoksia. Ajatus on, että Givens-muunnosten avulla Hessenberg-matriisi $\tilde{\mathbf{H}}_m$ saadaan yläkolmiomuotoon, jolloin pienimmän neliösummatehtävän ratkaiseminen on suoraviivaista.

muokataan edelleen matriisi $\tilde{\mathbf{R}}_{j+1}$. Muokkaus aloitetaan soveltamalla matriisiin $\tilde{\mathbf{H}}_{j+1}$ uuteen sarakkeeseen $j+1$ aikaisempia kiertomatriiseja $\mathbf{F}_i, i = 1, \dots, j$. Lopputuloksena saadaan $(j+2) \times (j+1)$ matriisi

$$\left[\begin{array}{cccc|cc} x & x & x & x & \cdots & x & | & x \\ 0 & x & x & x & \cdots & x & | & x \\ & 0 & x & x & \cdots & x & | & x \\ & & 0 & & & \vdots & | & x \\ \vdots & \vdots & & & x & x & | & x \\ & & \cdots & 0 & x & | & x \\ 0 & & \cdots & & 0 & | & r \\ 0 & & \cdots & & 0 & | & h \end{array} \right]. \quad (6.23)$$

Yllä olevassa matriisissa on pystypalkein eroteltu uuden sarakkeen alkiot. Lisäksi on voimassa yhtäsuuruus $h = h_{j+2,j+1}$, koska aikaisemmat kiertomatriisit eivät vaikuta alkioon paikassa $(j+2) \times (j+1)$. Kohdan (6.23) matriisista saadaan yläkolmiomatriisi $\tilde{\mathbf{R}}_{j+1}$ nollaamalla alkio h . Se onnistuu rotaatiomatriisilla \mathbf{F}_{j+1} , missä

$$c_{j+1} = r/(r^2 + h^2)^{1/2}, \quad s_{j+1} = -h/(r^2 + h^2)^{1/2}.$$

Edellä on havainnollistettu, miten algoritmin jokaisella iteraatioaskeella muodostetaan yläkolmiomatriisia. Nyt voidaan palata askel taaksepäin ja ottaa käyttöön merkintä $\mathbf{Q}_j = \mathbf{F}_j \mathbf{F}_{j-1} \cdots \mathbf{F}_2 \mathbf{F}_1$. Tällöin matriisi $\mathbf{Q}_j \in \mathbb{K}^{(j+1) \times (j+1)}$ on kiertomatriisien tulo eli kumuloitunut kiertomatriisi. Askeleen j jälkeen on saatu aikaan hajotelma

$$\mathbf{Q}_j \tilde{\mathbf{H}}_j = \tilde{\mathbf{R}}_j.$$

Toisin sanoen kertomalla matriisia $\tilde{\mathbf{H}}_k$ vasemmalta matriisilla \mathbf{Q}_j on askeleen j jälkeen saatu muodostettua yläkolmiomatriisi $\tilde{\mathbf{R}}_j$, jonka viimeinen rivi on nollarivi. On oleellista huomata, että myös niin sanottu oikean käden vektori $\beta \mathbf{e}_1$ täytyy kertoa kumuloituneella kiertomatriisilla

$$\mathbf{g}_j = \mathbf{Q}_j \beta \mathbf{e}_1.$$

Vektorilla $\mathbf{g}_j \in \mathbb{K}^{(j+1)}$ on $(j+1)$ komponenttia. Määritellään vektori $\mathbf{c} \in \mathbb{K}^j$, joka sisältää vektorin \mathbf{g}_j ensimmäiset j komponenttia. Vektorin \mathbf{g}_j viimeistä komponenttia vastatkoon muuttuja $d = (\mathbf{g}_j)_{j+1}$. Lisäksi käytetään merkintää $\mathbf{R}_j \in \mathbb{K}^{j \times j}$ matriisista, joka saadaan poistamalla matriisin $\tilde{\mathbf{R}}_j$ viimeinen nollarivi. Todetaan vielä, että matriisi \mathbf{Q}_j on unitaarinen, eikä näin ollen muuta vektorien pituutta matriisi-vektori kertolaskuissa. Funktiolle (6.21) saadaan nyt uusi lauseke

$$\begin{aligned} J(\mathbf{y}) &= \|\tilde{\mathbf{H}}_j \mathbf{y} - \beta \mathbf{e}_1\| = \|\mathbf{Q}_j (\tilde{\mathbf{H}}_j \mathbf{y} - \beta \mathbf{e}_1)\| = \|\tilde{\mathbf{R}}_j \mathbf{y} - \mathbf{g}_j\| \\ &= \left\| \begin{pmatrix} \mathbf{R}_j \mathbf{y} - \mathbf{c} \\ d \end{pmatrix} \right\| = \|\mathbf{R}_j \mathbf{y} - \mathbf{c}\| + \|d\|. \end{aligned} \quad (6.24)$$

Funktio (6.24) saa minimiarvonsa $\|d\|$, kun

$$R_j \mathbf{y} = \mathbf{c}. \quad (6.25)$$

GMRES algoritmi takaa, että $(R_j)_{i,i} \neq 0$ kaikilla $i = 1, \dots, j$. Koska R_j on yläkolmionmatriisi, niin vektori \mathbf{y} on helppo ratkaista takenevilla sijoituksilla $j \times j$ dimensioisesta yhtälöryhmästä (6.25). Nyt on nähtävissä, että jokaisella iteraatioaskeleella residuaalin \mathbf{r}_j normi on $\|d\|$, eli vektorin \mathbf{g}_j viimeisen komponentin itseisarvo. Näin ollen päätös algoritmin jatkamisesta tai lopettamisesta voidaan tehdä laskematta lainkaan vektoria \mathbf{y}_j . Vektori \mathbf{y}_j ratkaistaan lineaarisesta yhtälöryhmästä vain silloin, kun approksimaatiovektori $\mathbf{q}_j = \mathbf{q}_0 + V\mathbf{y}_j$ täytyy muodostaa. Tällaisia tilanteita ovat algoritmin uudelleenkäynnistys ja lopetusvaihe.

Lopuksi mainitaan muutama sana siitä, milloin algoritmien (6.3.2) ja (6.3.3) iterointi lopetetaan. Ilmeinen lopetusehto on se, että residuaalin normi on riittävän pieni. Algoritmit on määritelty kuitenkin siten, että iterointi lopetetaan myös, kun $\|\tilde{\mathbf{v}}_j\| = h_{j+1,j} = 0$. Tällöin iteraatioaskeleeseen liittyvän residuaalin normi on nolla ja algoritmi antaa tuloksena tarkan ratkaisun. Myös käänteinen pätee: jos algoritmi loppuu askeleella j , niin $h_{j+1,j} = 0$.

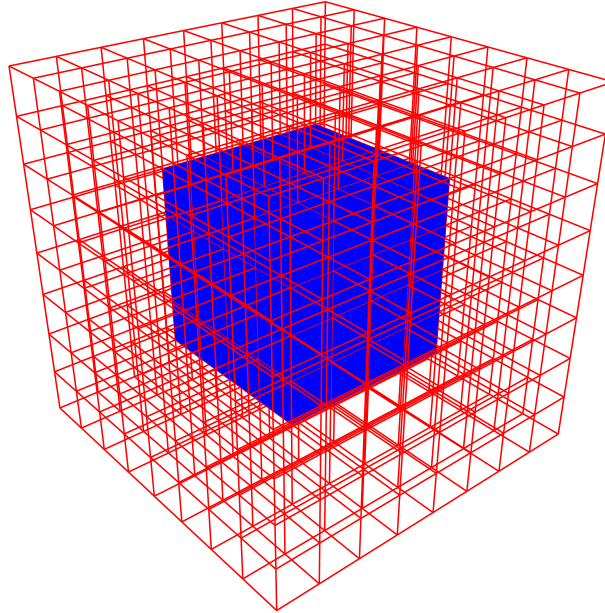
7 Laskentaesimerkkejä

Tässä luvussa simuloidaan sähkömagneettisen aallon heijastumista kolme erilaisen kappaleen pinnalta. Simuloinneissa sovelletaan aikaisemmissa luvuissa esitettyjä menetelmiä. Simuloitavat tilanteet on valittu yksinkertaisiksi, eivätkä ne edusta sen paremmin mitään tiettyä luonnossa esiintyvää tilannetta kuin teknistä sovellusta. Simulointien päätarkoitus on havainnollistaa, miten aiemmissa luvuissa teorioiltaan tutuiksi tulleita menetelmiä sovelletaan käytäntöön. Heijastavina kappaleina tullaan käyttämään kuutiota, porrasta ja onkaloa. Kaikille kolmelle kappaleelle lasketaan heijastunut sähkömagneettinen aalto kolmella erilaisella verkolla. Lisäksi simulointeja suoritetaan kahdella eri aallonpituudella. Simuloinnin tietokonetoteutuksessa käytetään apuna verkko-generaattoria, matemaattista aliohjelmakirjastoa sekä visualisointiohjelmia, jotka kaikki ovat *Janne Martikaisen* ohjelmoimia. Simuloinneista saatuja tuloksia tarkastellaan ja joitakin tuloksiin perustuvia havaintoja annetaan. Lopuksi esitetään sähkömagneettisesta simuloinnista tehtyjä päätelmiä.

7.1 Esimerkkitapausten simulointi

Simulointialueeksi asetetaan yksikkökuutio $[0, 1] \times [0, 1] \times [0, 1]$. Heijastavaa kappaletta lähestyvän lineaarisesti polarisoituneen tasoallon oletetaan etenevän x-akselin suuntaan $\mathbf{k} = (1, 0, 0)^T$ ja lähestyvän aallon sähkökenttä oletetaan z-akselin suuntaiseksi $\mathbf{u} = (0, 0, 1)^T$. Niin sanottuna aaltolukuna käytetään sekä $\alpha_1 = 5$ että $\alpha_2 = 10$, jolloin simuloitavien aaltojen aallonpituudet ovat $\lambda_1 = 2\pi/\alpha_1 \approx 1.257$ ja $\lambda_2 \approx 0.628$. Eli mitä suurempi aaltoluku on, sitä lyhyempi on aallonpituus. Koska aaltoluku vaikuttaa suoraan aallonpituuteen, niin se määrittelee samalla sähkömagneettisen aallon taajuuden. Tässä luvussa kaikki pituudet ja etäisyydet ilmoitetaan metreissä. Heijastavat kappaleet kuutio, porras ja onkalo edellä mainitussa järjestyksessä olkoot $\Omega^k = [\frac{1}{4}, \frac{3}{4}] \times [\frac{1}{4}, \frac{3}{4}] \times [\frac{1}{4}, \frac{3}{4}]$, $\Omega^p = \Omega^k \setminus [\frac{1}{4}, \frac{1}{2}[\times[\frac{1}{4}, \frac{3}{4}]\times]\frac{1}{2}, \frac{3}{4}]$, $\Omega^o = \Omega^k \setminus [\frac{1}{4}, \frac{5}{8}[\times]\frac{3}{8}, \frac{5}{8}[\times]\frac{3}{8}, \frac{5}{8}[\times]\frac{3}{8}]$.

Simulointialueeseen muodostetaan harva verkko siten, että yksikkökuutio ositetaan jokaiseen koordinaattisuuntaan yhdeksällä jakopisteellä. Jako suoritetaan vieläpä niin, että yksikkökuutio jakaantuu jokaisessa koordinaattisuunnassa kahdeksaan tasamittaiseen osaväliin. Tällöin harvan verkon jokainen elementti on kuutio, jonka särmän pituus on $1/8$. Samalla tavalla muodostetaan kaksi tiheämpää verkkoa, ainoastaan jakopisteiden määrää lisätään. Tiheämpien verkkojen muodostamiseen käytetään 17 ja 25 jakopistettä, jolloin kuutioelementtien särmien pituudet ovat $1/16$ ja $1/24$. Sähkömagneettisissa simuloinneissa on oleellista, kuinka monta elementtiä mahtuu yhdellä aallonpituudella. Kun aaltoluku on 5, niin kyseiset lukumäärät edellä määritellyille



Kuva 7.1: Simulointialueeseen muodostettu harva verkko ja heijastava kuutio.

verkoille ovat 10, 20 ja 30. Aaltoluvulla 10 simulointi suoritetaan ainoastaan tiheimmällä verkolla. Tällöin elementtejä mahtuu 15 aallonpituuden matkalle. Kuvassa 7.1 on simulointialueeseen muodostettu harva verkko, jonka keskellä sijaitsee heijastavana kappaleena toimiva kuutio.

Matemaattinen malli diskretisoidaan äärellisten elementtien menetelmällä. Muodostunut diskreetti tehtävä kootaan lineaariseksi yhtälöryhmäksi, josta eliminoidaan kaikki ne vapausasteet, jotka liittyvät heijastavan kappaleen pinnalla sijaitseviin särmiin. Eliminointi on mahdollista, koska tehtävän tarkka ratkaisu tunnetaan heijastavan kappaleen pinnalla alkutietojen perusteella. Eliminoinnilla saadaan ratkaistavan yhtälöryhmän kokoa pienennettyä. Eliminoinnin jälkeen yhtälöryhmä ratkaistaan GMRES menetelmällä, joka käynnistetään uudelleen aina 50 iteraatioaskeleen jälkeen. Iteroinnissa käytetään suhteellista lopetuskriteeriä. Se tarkoittaa, että iterointi lopetetaan, kun residuaalivektorin normin ja oikean puolen vektorin normin välinen suhdeluku on alle valitun rajan. Tässä tutkielmassa suhdeluvun rajaksi on valittu 10^{-6} .

Simulointien yhteydessä mitattiin iteraatioaskelten määrä ja laskenta-aika, jonka GMRES menetelmä tarvitsi kuhunkin tehtävään liittyvän lineaarisen yhtälöryhmän ratkaisemiseen. Mitattu aika ei siis sisällä yhtälöryhmän kokoamista eikä varsinaisen ratkaisun laskemista. Varsinainen simulointitehtävän ratkaisu, eli simulointialueeseen muodostuva sähkökenttä, lasketaan yhtälöryhmästä ratkaistujen tuntemattomien avulla. Taulukko 7.1 sisältää simulointitehtävät sekä niille mitatut tulokset. Taulukon neljäs sarake kertoo, kuinka paljon tehtäviin liittyvät lineaariset yhtälöryhmät sisältävät tuntemattomia kertoimia. Jokainen simulointitehtävä on laskettu kolme kertaa ja tau-

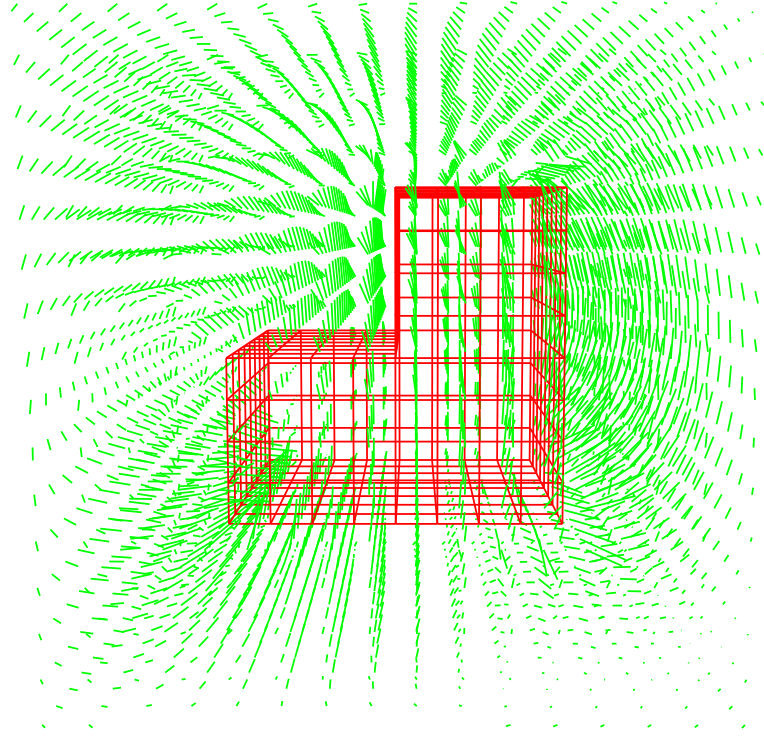
lukossa ilmoitettu aika on kolmen tehtävän keskiarvo. Simuloinnit suoritettiin *Fujitsu Siemensin* laitesarjaan *Amilo-A* kuuluvalla kannettavalla tietokoneella *A1400+*. Simulointihetkellä tietokoneessa oli *Mobile AMD XP 1400+* prosessori, *256 Mt* keskusmuistia ja *Windows XP Home Edition* käyttöjärjestelmä.

kappale	verkko	aaltoluku	tuntem.	elementit	iteraatiot	aika(s)
kuutio	8×8×8	5	1644	448	209	2.6
onkalo	8×8×8	5	1659	460	216	2.6
porras	8×8×8	5	1700	464	260	3.4
kuutio	16×16×16	5	11928	3584	589	64.4
onkalo	16×16×16	5	12126	3680	617	70.0
porras	16×16×16	5	12344	3712	647	73.1
kuutio	24×24×24	5	38916	12096	1008	381.4
onkalo	24×24×24	5	39681	12420	1081	439.6
porras	24×24×24	5	40284	12528	1064	426.9
kuutio	24×24×24	10	38916	12096	790	327.3
onkalo	24×24×24	10	39681	12420	874	385.0
porras	24×24×24	10	40284	12528	717	310.6

7.2 Tulosten tarkastelu

Kun GMRES menetelmän iterointinopeutta tarkasteltiin simuloinnin aikana, oli selkeästi havaittavissa, että nopeus kiihtyi algoritmin uudelleen käynnistyksen seurauksena. Edelleen oli havaittavissa, että iterointinopeus pieneni jokaisella iteraatioaskeleella, joka otettiin käynnistyksen jälkeen ja nopeus oli hitaimmillaan juuri ennen seuraavaa käynnistystä. Havainnot tukevat GMRES menetelmän esittelyn yhteydessä ilmi tulleita teoreettisia tuloksia. Lähes itsestään selvä havainto taulukon 7.1 tuloksista on, että mitä enemmän tehtävä sisältää tuntemattomia sitä enemmän iteraatioita ja aikaa tehtävän ratkaiseminen vaatii.

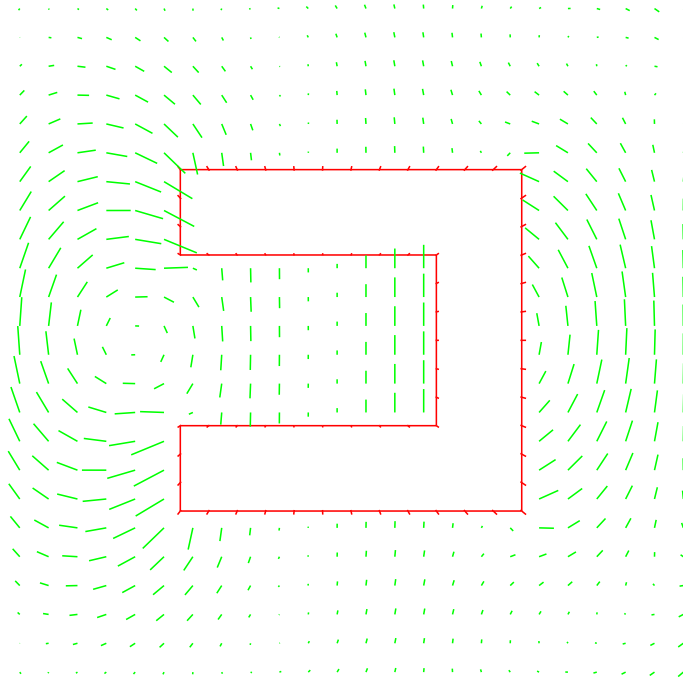
Yllättävää kuitenkin oli, että tiheimmälle verkolle laskenta-aika lyheni, kun aallon taajuutta nostettiin. Esimerkiksi artikkelissa [20] esitettyjen tulosten perusteella, täysin päinvastainen seuraus olisi ollut odotettavissa. Taulukossa 7.1 esitettyjä tuloksia tarkastelemalla havaitaan, että laskenta-ajat ovat kasvaneet jyrkimmin onkalo-tehtäville. Erityisesti näyttäisi siltä, että tiheämmällä verkolla ja korkeammalla taajuudella sähkökentän ratkaiseminen onkalosta heijastuneelle aallolle on laskennallisesti vaikeampaa kuin kuutiolle ja portaalle. Varovainen arvio onkin, että heijastavien kappaleiden geometriset ominaisuudet vaikuttavat tehtävän haasteellisuuteen. Simulointituloksia visuaalisesti tarkastelemalla osoittautui, että harvalla verkolla lasketuista ratkaisuista oli vaikea hahmottaa sähkömagneettisen aallon muotoa. Syynä on se, että harvalla verkolla aallonpituuden matkalle mahtuu vain vähän elementtejä.



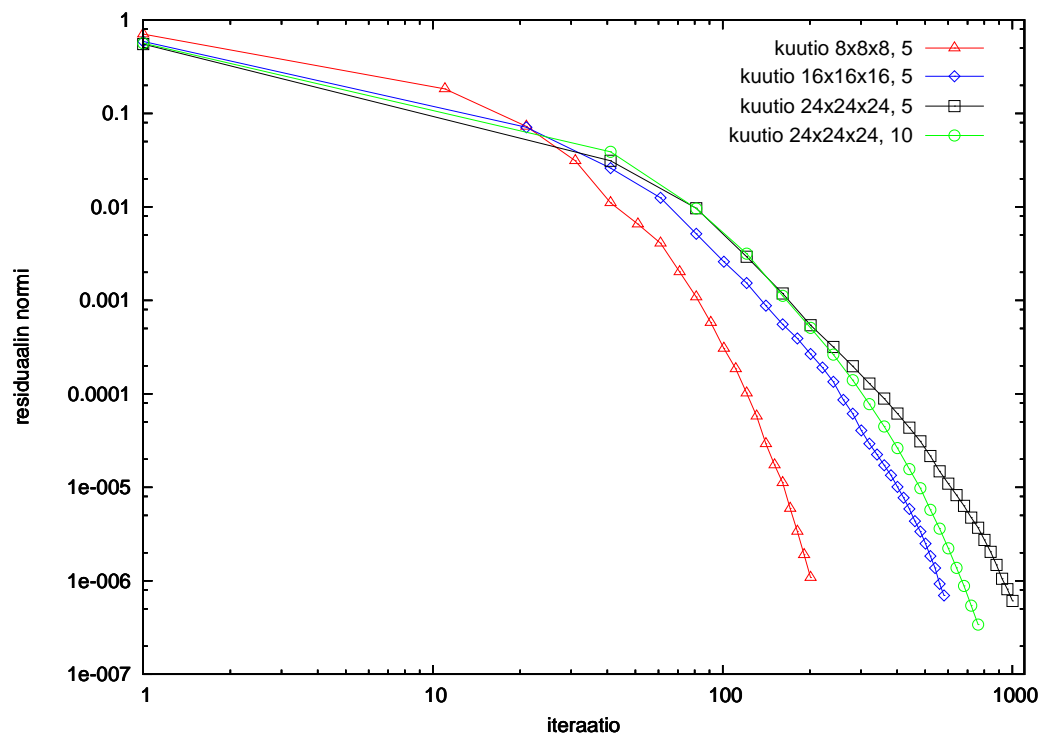
Kuva 7.2: Portaasta heijastuneen aallon muodostama sähkökenttä.

Kuvassa 7.2 on portaasta heijastuneen aallon muodostama sähkökenttä $\hat{\mathbf{E}}^s$ yksikkökuutiossa. Simulointi on suoritettu verkolla, joka muodostettiin 17 jakopisteellä. Aaltoluku on 5. Kuvassa porrasta lähestyvä aalto etenee vasemmalta oikealle ja lähestyvän aallon sähkökentän suunta on kuvan pystysuunta. Kuva 7.3 on poikkileikkaus onkalosta heijastuneen aallon sähkökentästä. Simulointi on suoritettu tiheimmällä verkolla ja aaltoluku on 10. Kuvassa 7.3 positiivisen x-akselin suunta on vasemmalta oikealle ja positiivisen z-akselin suunta ylhäältä alaspäin. Molemmissa kuvissa sähkökenttä on laskettu jokaisen elementin keskipisteessä. Kuvaan 7.4 piirretyt käyrät osoittavat, miten residuaalivektorien normit suppenevat eri simulointitehtävissä. Yhteistä käyriin liittyvillä tehtävillä on, että kaikissa heijastavana kappaleena toimii kuutio. Käyristä kolme edustaa tehtäviä, joissa aaltoluku on 5; ainoastaan verkon tiheys vaihtelee. Neljännen käyrän tehtävässä simulointi suoritettiin tiheimmällä verkolla aaltoluvun ollessa 10. Kuvan 7.4 vaaka- ja pystyakseleille on valittu logaritminen asteikko.

Lopuksi mainitaan muutama sana laskentaesimerkeissä syntyneiden tulosten tarkkuudesta. Ideaalitulanteessa tuloksia voitaisiin verrata esimerkiksi keinotekoisella reunaanalyttisesti laskettuihin lausekkeisiin. Valitettavasti matemaattinen malli on niin monimutkainen, ettei analyttisen lausekeen laskeminen onnistu edes tässä luvussa esitettyjen yksinkertaisten laskentaesimerkkien tapauksissa. Vaihtoehtoinen keino tuloksien tarkasteluun on käyttää esimerkiksi *a posteriori* -virhearviota. Tällöin si-



Kuva 7.3: Poikkileikkauskuva tiheimmällä verkolla lasketusta sähkökentästä.



Kuva 7.4: Residuaalivektorien normin suppeneminen iteraatioaskelten edetessä.

muloinnista saatavat tulokset sijoitetaan takaisin matemaattiseen malliin. Sijoituksen jälkeen mallin residuaali lasketaan numeerisesti, jolloin tuloksille saadaan paikasta riippuva virhearvio. Tässä tutkielmassa varsinaista virheen arviointia ei suoriteta. Saatuja tuloksia on kuitenkin verrattu viitetuloksiin ja todettu, että vertailtavat vastaavat toisiaan.

7.3 Johtopäätökset

Simulointitehtävien vaatima laskenta-aika riippuu oleellisesti siitä, kuinka suuri tarkkuus tuloksilta vaaditaan. Esimerkiksi tässä tutkielmassa keinotekoiselle reunalle asetettu Silver-Müller -ehto ja diskreetin tehtävän muodostamisessa käytetyt ensimmäisen asteen Nédélec -elementit ovat tarkkuudeltaan kertaluokkaa $\mathcal{O}(h)$, missä h edustaa kuutioelementtien tapauksessa särmän pituutta. Tällöin simulointituloksien tarkkuus on suoraan verrannollinen elementtien kokoon. Jotta tulosten tarkkuuden kertaluokkaa saataisiin nostettua, tulisi tehtävän diskretisoinnissa käyttää korkeamman asteen Nédélec -elementtejä. Lisäksi keinotekoisella reunalla täytyisi soveltaa esimerkiksi *PML* (*Perfectly Matched Layer*) -tekniikkaa [2]. Mikäli keinotekoista reunaa ei huomioida, niin tulosten tarkkuutta voidaan parantaa karkeasti sanoen kahdella tavalla: tiheämmällä verkolla tai korkeamman asteen elementeillä. Molemmat johtavat entistä suurempiin yhtälöryhmiin ja usein onkin tarpeen selvittää, kummalla tavalla haluttu tarkkuus saavutetaan edullisemmin. Sekä PML että korkeamman asteen Nédélec -elementit olisivat hyviä jatkotutkimuskohteita.

GMRES menetelmän tehokas käyttö haastavien simulointitehtävien ratkaisemiseen edellyttää niin sanottua *pohjustinta* (*preconditioner*); pohjustimilla kiihdytetään iteratiivisten menetelmien konvergenssinopeutta. Lyhyesti sanottuna pohjustin on matriisi, jota sovelletaan alkuperäiseen lineaariseen yhtälöryhmään jollakin pohjustustekniikasta riippuvalla tavalla. Operaation jälkeen ratkaistaankin muunnettu yhtälöryhmä alkuperäisen yhtälöryhmän sijasta. Tehokkaita pohjustimia saadaan muodostettua yhdistämällä *FDM* (*Fictitious Domain Method*) [19] ja jokin nopea suora ratkaisija (*Fast Direct Solver*), joka perustuu *FFT*:hen (*Fast Fourier Transform*) [6]. Artikkelissa [20] on esitelty eräs nopea suora ratkaisija, joka perustuu nimenomaan *FFT*:hen ja joka soveltuu heijastuneen sähkömagneettisen aallon simulointiin. Kyseisessä artikkelissa julkaistut konvergenssitulokset ovat merkittävästi parempia kuin esimerkiksi tässä luvussa esitetyt tulokset. Tehokkaat pohjustimet ja nopeat suorat ratkaisijat sähkömagneettisille simulointitehtäville olisivatkin haastava ja mielenkiintoinen jatkotutkimuskohde.

On osoittautunut, että kolmiulotteiset sähkömagneettiset simulointitehtävät ovat yksi haastavimmista tehtäväjoukoista. Syitä on useita. Ilmiöitä kuvaavat matemaattiset lausekkeet ovat monimutkaisia ja niiden numeerinen laskeminen vaatii paljon laskutoimituksia. Laskuissa joudutaan käyttämään kompleksilukuaritmetiikkaa, joka on laskennallisesti raskaampaa kuin reaalityyppinen laskenta. Tärkein syy on kuitenkin se, että

sähkömagneettisen aallon aallonpituuden matkalla tulee olla vähintään 8 - 12 elementtiä, jotta saatu ratkaisu kuvaisi edes tyydyttävästi aallon käyttäytymistä. Seurauksena on, että vaatimattomissakin simulointitehtävissä tuntemattomien määrä kasvaa kymmeniin tuhansiin. Käytännön merkitystä omaavissa simulointitehtävissä tuntemattomia on helposti miljoonia.

8 Yhteenveto

Tässä tutkielmassa tarkasteltiin sähkömagneettisen sironnan numeerista simulointia. Tutkielman alkuosa keskittyi sironnan fysikaalisiin taustoihin ja matemaattisiin perusteisiin. Aivan ensimmäiseksi tutkittiin sähkömagnetismia fysikaalisena ilmiönä. Merkittävimpänä yksittäisenä asiana esiteltiin Maxwellin yhtälöt. Yhtälöiden pohjalta johdettiin matemaattinen malli kappaleen pinnalta heijastuneen sähkömagneettisen aallon sähkökentälle. Malli koostui aaltoyhtälöstä, heijastavan kappaleen pinnalle määrittelystä reunaehdosta ja keinotekoiselle reunalle asetetusta Silver-Müller -ehdosta. Matemaattista mallia muodostettaessa tehtiin joukko oletuksia. Tärkeimpinä mainittakoon, että ympäröivä avaruus oletettiin tyhjiöksi, joka ei sisällä sähköisiä varauksia eikä virtoja, sähkömagneettinen aalto oletettiin yhdestä taajuudesta koostuvaksi aikaharmoniseksi aalloksi ja heijastavan kappaleen pinta oletettiin täydelliseksi johteeksi.

Tutkielman seuraava looginen kokonaisuus oli matemaattista mallia vastaavan diskreetin tehtävän muodostaminen. Ensitöiksi mallista laadittiin heikko muotoilu, joka on integraalivastine matemaattiseksi malliksi kutsutulle osittaisdifferentiaaliyhtälöryhmälle. Seuraavaksi heikko muotoilu diskretisoitiin äärellisten menetelmien menetelmällä. Diskretisoinnissa käytettiin ensimmäisen asteen Nédélec -kuutioelementtejä. Kuutioelementtien käyttö oli mielekästä, koska heijastavat kappaleet oletettiin yksinkertaisiksi. Monimutkaisempien kappaleiden mallintamiseen tarvittaisiin elementtejä, jotka ovat esimerkiksi nelitahokkaita tai käyräreunaisia. Lopuksi diskreetti tehtävä koottiin lineaariseksi yhtälöryhmäksi. Diskretisointi olisi voitu tehdä vaihtoehtoisesti vaikkapa reunaelementtimenetelmällä (*Boundary Element Method*). Kiintoisa kohde jatkotutkimukselle olisikin tutkia reunaelementtimenetelmän soveltuvuutta sähkömagneettisiin simulointitehtäviin.

Viimeinen osio tutkielmassa piti sisällään lineaarisen yhtälöryhmän ratkaisemisen ja simulointitulosten tarkastelua. Yhtälöryhmien ratkaisemiseen liittyen tarkasteltiin iteratiivisia menetelmiä ja niistä erityisesti projektio- ja Krylovin aliavaruusmenetelmiä. GMRES menetelmälle esiteltiin konkreettinen algoritmi ja perehdyttiin algoritmin käyttökelpoiseen toteutukseen. Vihoviimeiseksi ratkaistiin joukko laskentaesimerkkejä GMRES menetelmän avulla. Laskentaesimerkit olivat yksinkertaisia simulointitehtäviä, joiden tarkoituksena oli tuoda ilmi konkreettiseen simulointiin liittyviä seikkoja. Tärkein havainto oli, että kolmiulotteiset sähkömagneettiset simulointitehtävät ovat laskennallisesti erittäin haasteellisia. Huolimatta siitä, että niiden ratkaisemista on tutkittu jo vuosikymmeniä, ne muodostavat edelleenkin aktiivisen tutkimusalueen. Tekniikan kehityksen myötä tutkimuksissa syntyville menetelmille ja keksinnöille löytyy päivä päivältä enemmän sovelluskohteita.

9 Kirjallisuutta

- [1] F. Assous, P. Ciarlet, Jr., and J. Segré. Numerical solution to the time-dependent Maxwell equations in two-dimensional singular domains: the singular complement method. *J. Comput. Phys.*, 161(1):218–249, 2000.
- [2] J.-P. Berenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 127(2):363–379, 1996.
- [3] A. Bespalov. Numerical simulation of 3D electromagnetic scattering by algebraic fictitious domain method. Technical Report RR-2729, INRIA, 1995.
- [4] P. G. Ciarlet and J.-L. Lions, editors. *Handbook of numerical analysis. Vol. II.* Handbook of Numerical Analysis, II. North-Holland, Amsterdam, 1991. Finite element methods. Part 1.
- [5] G. C. Cohen. *Higher-order numerical methods for transient wave equations.* Scientific Computation. Springer-Verlag, Berlin, 2002. With a foreword by R. Glowinski.
- [6] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19:297–301, 1965.
- [7] C. Cronström and P. Lipas. *Johdatus elektrodynamiikkaan ja suhteellisuusteoriaan.* Limes ry, Helsinki, 1986.
- [8] R. Dautray and J.-L. Lions. *Mathematical analysis and numerical methods for science and technology. Vol. 3.* Springer-Verlag, Berlin, 1990. Spectral theory and applications, With the collaboration of Michel Artola and Michel Cessenat, Translated from the French by John C. Amson.
- [9] H. C. Elman. Iterative methods for linear systems. In *Large-scale matrix problems and the numerical solution of partial differential equations (Lancaster, 1992)*, Adv. Numer. Anal., III, pages 69–118. Oxford Univ. Press, New York, 1994.
- [10] R. Freund, G. Golub, and N. Nachtigal. Iterative solution of linear systems. In *Acta Numerica 1992*, pages 57–100. Cambridge University Press, 1992.
- [11] I. S. Grant and W. R. Phillips. *Electromagnetism.* John Wiley & Sons, Chichester, 2. edition, 1990.

- [12] J. Haataja, J. Heikonen, Y. Leino, J. Rahola, J. Ruokolainen, and V. Savolainen. *Numeeriset menetelmät käytännössä*. CSC - Tieteellinen laskenta Oy. Saatavilla WWW-muodossa <URL:<http://www.csc.fi/oppaat/num.kayt/num.kayt.pdf>>, 2.9.2002.
- [13] J. Jin. *The Finite Element Method in Electromagnetics*. Wiley-Interscience [John Wiley & Sons], New York, 1993.
- [14] A. Kirsch and P. Monk. A finite element/spectral method for approximating the time-harmonic Maxwell system in \mathbb{R}^3 . *SIAM J. Appl. Math.*, 55(5):1324–1344, 1995.
- [15] A. Kirsch and P. Monk. A finite element method for approximating electromagnetic scattering from a conducting object. *Numer. Math.*, 92(3):501–534, 2002.
- [16] H. Koskinen and A. Viljanen. *Elektrodynamiikka*. Helsingin yliopisto, fysikaalisten tieteiden laitos. Saatavilla WWW-muodossa <URL:http://www.physics.helsinki.fi/~tfo_www/lectures/ed/luento/ed2003.pdf>, 28.4.2003.
- [17] M. Křížek and P. Neittaanmäki. *Finite element approximation of variational problems and applications*, volume 50 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow, 1990.
- [18] P. Lorrain, D. P. Corson, and F. Lorrain. *Electromagnetic Fields and Waves*. W. H. Freeman and Company, New York, 3. edition, 1988.
- [19] G. I. Marchuk, Y. A. Kuznetsov, and A. M. Matsokin. Fictitious domain and domain decomposition methods. *Soviet J. Numer. Anal. Math. Modelling*, 1(1):3–35, 1986.
- [20] J. Martikainen, T. Rossi, K. Rogovin, and J. Toivanen. Fast direct solver for an electromagnetic problem with an application. In *Proceedings of Waves 2003 - the Sixth International Conference on Mathematical and Numerical Aspects of Wave Propagation*, pages 252–256. Springer-Verlag, 2003.
- [21] P. Monk and L. Demkowicz. Discrete compactness and the approximation of Maxwell's equations in \mathbb{R}^3 . *Math. Comp.*, 70(234):507–523, 2001.
- [22] J.-C. Nédélec. Mixed finite elements in \mathbb{R}^3 . *Numer. Math.*, 35(3):315–341, 1980.
- [23] J. N. Reddy. *An Introduction to the Finite Element Method*. McGraw-Hill, New York, 1984.

- [24] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS publishing, 1. edition, 1996. Saatavilla WWW-muodossa <URL:ftp://ftp.cs.umn.edu/dept/users/saad/PS/all_pdf.zip>, 22.01.2002.
- [25] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.
- [26] Y. Saad and H. A. van der Vorst. Iterative solution of linear systems in the 20th century. *J. Comput. Appl. Math.*, 123(1-2):1–33, 2000. Numerical analysis 2000, Vol. III. Linear algebra.
- [27] A. Schneebeli. *An $H(\text{curl};\Omega)$ -conforming FEM: Nédélec elements of first type*. Saatavilla WWW-muodossa <URL:<http://www.dealii.unihd.de/developer/reports/nedelec/nedelec.pdf>>, 2.5.2003.
- [28] L. S. Segerlind. *Applied Finite Element Analysis*. John Wiley & Sons, Inc., New York, 1976.
- [29] J. Tervo and M. Vauhkonen. *Elementtimenetelmän perusteita*. Kuopion yliopisto, sovelletun fysiikan laitos. Saatavilla WWW-muodossa <URL:<http://venda.uku.fi/studies/kurssit/FEM/fem99.pdf>>, 22.6.2000.
- [30] B. Thidé. *Electromagnetic Field Theory*. Uppsalan yliopisto, tähtitieteen ja avaruusfysiikan laitos. Saatavilla WWW-muodossa <URL:http://www.plasma.uu.se/CED/Book/EMFT_book.pdf>, 19.5.2003.
- [31] R. Weiss. Orthogonalization methods. Raportti 52/94, Karlsruhe yliopisto, laskentakeskus, 1994.
- [32] R. Weiss. A theoretical overview of Krylov subspace methods. *Applied Numerical Mathematics: Transactions of IMACS*, 19(3):207–233, 1996.
- [33] H. D. Young and R. A. Freedman. *University Physics with Modern Physics*. Addison-Wesley, San Francisco, 10. edition, 2000.
- [34] O. C. Zienkiewicz. *The Finite Element Method*. McGraw-Hill, London, 1978. The third, expanded and revised edition of *The Finite Element Method in Engineering Science*.
- [35] O. C. Zienkiewicz and K. Morgan. *Finite elements and approximation*. A Wiley-Interscience Publication. John Wiley & Sons Inc., New York, 1983.