

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II
DIPARTIMENTO DI INGEGNERIA ELETTRICA

Tesi per il conseguimento del titolo di
Dottore di Ricerca in
Ingegneria Elettrotecnica

**Analisi Numerica 3D di Problemi Elettromagnetici e MHD
tramite la Minimizzazione dell'Errore Costitutivo.**

IL TUTORE
(PROF. RAFFAELE MARTONE)

IL DOTTORANDO
(RAFFAELE FRESA)

IL COORDINATORE DEL CORSO DI DOTTORATO
(PROF. LUCIANO DE MENNA)

Ringraziamenti

La lista delle persone cui sento di dovere dei ringraziamenti è estremamente lunga, ed inizia con mio padre sempre presente con le giuste parole ogni volta che ho attraversato momenti difficili. Ringrazio ancora mia madre e le mie sorelle, tutti i miei amici e colleghi che mi sono stati vicini nel corso di questi anni.

Un ringraziamento particolare va a tutti i professori con i quali ho collaborato in questi ultimi anni. In particolare, al prof. R.Martone, mio tutore, e ai professori R.Albanese e G. Rubinacci, senza i consigli e le ipercritiche osservazioni dei quali questo lavoro non sarebbe stato possibile. Desidero ringraziare, infine, il prof.Coccorese, il prof. Bobbio e il prof. De Menna, coordinatore del corso di dottorato, per la simpatia e la disponibilità che hanno sempre mostrato nei miei confronti.

R.Fresa

**Analisi numerica 3D di problemi elettromagnetici e MHD
tramite la minimizzazione dell'errore costitutivo.**

Indice

Introduzione

**Capitolo I: Stato dell'arte sui metodi di analisi dei campi
Elettromagnetici.**

**1.1) *Generalità sui metodi analitici e numerici per la soluzione
dei problemi Elettromagnetici.***

1.1.1) Metodi analitici

- a) Funzione di Green*
- b) Metodo di separazione delle variabili*
- c) Metodo delle trasformazioni conformi*

1.1.2) Metodi numerici

- a) Metodo delle differenze finite (FDM)*
- b) Metodo dei boundary elements (BEM)*
- c) Metodo degli elementi finiti (FEM).*

1.2) *Modelli E.M. classici per formulazioni F.E.M.*

1.2.1) Magnetostatica

1.2.2) Elettrocinetica quasi-stazionaria

1.2.3) Propagazione

1.3) *Rassegna dei metodi basati su principi variazionali complementari. Il metodo della minimizzazione dell'errore costitutivo.*

1.3.1) Campi statici

1.3.2) Elettrocinetica quasi-stazionaria

1.3.3) Propagazione in regime sinusoidale

Capitolo 2: L'approccio basato sulla minimizzazione dell'errore costitutivo nell' Elettromagnetismo.

2.1) *Generalità*

2.2) *Formulazione per problemi di campo 3D nel dominio del tempo*

2.2.1) Modello matematico

2.2.2) Definizione del funzionale di errore.

2.2.3) Espressioni duali delle grandezze integrali

2.3) *Formulazione per problemi circuitali.*

2.4) *Confronto tra le formulazioni basate sull'errore costitutivo nei problemi di campo e nei circuiti a parametri concentrati.*

2.5) *Formulazione per le equazioni di Maxwell in regime sinusoidale.*

Capitolo 3: Aspetti computazionali .

3.1) Soluzione numerica dei problemi a valore iniziale

3.1.1) Generalità

3.1.2) Soluzione numerica dei problemi a valore iniziale o di Cauchy

- *Il metodo implicito*
- *Il metodo Crank-Nicolson*
- *Il metodo esplicito*
- *Il metodo θ .*

3.2) *Metodi di soluzione di sistemi di equazioni algebriche lineari e non lineari*

3.2.1) Risoluzione di sistemi di equazioni lineari.

- *Introduzione*
- *Generalità intorno ai sistemi lineari*
- *Tecniche di immagazzinamento delle matrici*
- *Metodi di risoluzione dei sistemi lineari*

3.2.2) Risoluzione di sistemi di equazioni non lineari

- *Generalità*
- *Problemi non-lineari mono-dimensionali*
- *Problemi non-lineari multi-dimensionali*

3.3) *Metodi agli elementi finiti: gli edge elements*

3.4) *Formulazione numerica del metodo della minimizz. dell'errore costitutivo.*

- *Discretizzazione numerica dei potenziali vettori incogniti*

- Espressione del funzionale di errore nelle incognite discretizzate
- Formulazione discreta del problema di minimizzazione del funzionale di errore
- Imposizione delle condizioni al contorno
- Risoluzione numerica del sistema non lineare associato alla minimizzazione.

3.5) Rappresentazione grafica dei campi di errore.

3.6) Implementazione numerica del metodo error-based per una rete del secondo ordine.

Capitolo 4: Casi test e risultati

Introduzione

4.1) Circuiti a parametri concentrati: Rete RLC lineare.

4.2) Magnetostatica: Cubo attraversato da corrente uniforme

4.3) Correnti parassite: il Felix-brick.

4.4) Propagazione.

-Incidenza di un'onda piana su uno slab di materiale dielettrico lineare.

-Incidenza di un'onda piana su uno slab di materiale dielettrico non lineare.

Capitolo 5: Minimizzazione dell'errore costitutivo nella MHD e applicazioni ai modelli dinamici di plasma.

5.1) Generalita' sulla fusione nucleare

5.2) Il modello MHD.

-Modello matematico del plasma: le equazioni MHD

-Il modello del plasma all'equilibrio

-Il modello dinamico del plasma

5.3) Il problema MHD in assenza di forze di pressione: un approccio basato sulla minimizzazione dell'errore costitutivo.

Conclusioni

Appendice A: **Edge functions**

Appendice B: **Scelta dei pesi sulle relazioni costitutive per il disaccoppiamento delle variabili**

Appendice C: **Imposizione delle condizioni di assorbimento in problemi monodimensionali**

INTRODUZIONE

Le leggi dell'Elettromagnetismo hanno raggiunto l'attuale forma compatta ed elegante, intitolata al nome del loro artefice finale, Maxwell, solo a valle di una lunga fase di rielaborazione ed integrazione di relazioni di carattere particolare; esse si pongono come obiettivo finale l'inquadramento in un contesto unitario dell'intero universo dei fenomeni fisici di natura elettromagnetica.

In maniera affatto opposta al processo storico, appare oggi naturale postulare le equazioni di Maxwell come le leggi dell'Elettromagnetismo, dalle quali le più semplici leggi di Coulomb, Ampere, Faraday possono essere derivate come casi particolari. Questo viene fatto semplicemente rimuovendo tra i termini di sorgente quelli che, nello specifico problema in esame, apportano ai campi un contributo ritenuto trascurabile. Aggiungendo alle equazioni di Maxwell, così particolarizzate, opportune condizioni iniziali e al contorno e le relazioni costitutive, ovvero i modelli matematici delle proprietà fisiche dei materiali che occupano la regione in oggetto, sotto opportune ipotesi limitative sul carattere di queste ultime, si può dimostrare che il problema Elettromagnetico resta completamente definito in modo da ammettere soluzione unica.

Il generico modello matematico del problema fisico è costituito da un sistema di equazioni integro-differenziali le cui incognite sono campi elettromagnetici o grandezze a questi correlate. Poiché solo in rarissimi casi (geometrie particolari, caratteristiche del mezzo lineari ed omogenee) il

sistema matematico ammette una soluzione in forma chiusa, si ricorre all'ausilio di metodi numerici che sono capaci di fornire soluzioni soddisfacenti da un punto di vista ingegneristico per classi di problemi, limitate da vincoli molto meno rigidi rispetto a quelle imposti dai metodi analitici. Pertanto è pratica comune per gli ingegneri risolvere in forma numerica la maggior parte dei complessi problemi che si manifestano in fase di analisi e di progetto delle moderne sofisticate apparecchiature elettromagnetiche adoperate nei più disparati settori: dalla medicina ai processi industriali, dalla fusione nucleare agli acceleratori di particelle. Questo enorme successo delle tecniche risolutive di tipo numerico è largamente dovuto agli enormi progressi che le tecnologie informatiche hanno registrato negli ultimi anni in termini sia di potenza di calcolo, che di capacità di memoria.

A fronte dei suoi enormi vantaggi, l'analisi numerica, cioè la tecnica di risolvere problemi matematici attraverso approssimazione numerica, comporta però una serie di problemi che, allo stato attuale, sono ancora oggetto di ricerca e di sviluppo.

Le soluzioni trovate sono affette da errori di discretizzazione e di arrotondamento, tanto più elevati quanto più è rozza la schematizzazione adoperata per descrivere il problema fisico. La accuratezza della discretizzazione è d'altronde limitata dalla disponibilità di memoria e dalla velocità del computer (anche se con il passare del tempo l'avanzare della tecnologia rende sempre meno stretti questi vincoli). Se si considera un metodo basato sugli elementi finiti, per poter ottenere una soluzione numerica soddisfacente appare dunque chiara la necessità di operare una modellizzazione giudiziosa del problema in esame che spinga maggiormente il raffinamento della discretizzazione nelle zone ove l'errore numerico può concentrarsi. Naturalmente questa strategia presuppone che sia possibile, attraverso una fase di post-processing della soluzione, ricavare informazioni circa la distribuzione dell'errore, compito tutt'altro che facile con la maggior parte dei metodi numerici tradizionali.

Un ulteriore aspetto dell'analisi numerica Elettromagnetica non appare ancora pienamente soddisfacente. Si è osservato come l'intero universo dei fenomeni elettromagnetici sia rappresentato nelle equazioni di Maxwell; e quanto sia semplice passare da queste equazioni di carattere generale a quelle che descrivono i problemi dell'elettrostatica, della magnetostatica, della propagazione, dell'elettrocinemica, etc. Ad una siffatta semplicità nella scrittura dei modelli, purtroppo, non ha fatto sinora riscontro un metodo risolutivo numerico di carattere generale. Sono state così proposte nel passato una moltitudine di metodi ciascuno specializzato nel trattamento di uno solo o, comunque, pochi tra i particolari sottoproblemi di carattere elettromagnetico sopra citati. Complicazioni ulteriori si hanno, come appare naturale, nella trattazione di problemi accoppiati, ed anzi, in diversi casi, esistono problematiche che, ancora oggi, non si è in grado di trattare se non al prezzo di semplificazioni tanto notevoli da inficiare la validità dei risultati ottenibili. E' questo ad esempio il caso del sistema delle equazioni MHD, che interessa fortemente il settore della fusione termonucleare controllata, la cui risoluzione operativa, salvo casi particolari, è disponibile solo nel caso di geometria assisimmetrica o quasi-assisimmetrica.

In questo lavoro il metodo solutivo basato sulla minimizzazione dell'errore sulle relazioni costitutive, proposto in passato per la risoluzione di problemi stazionari e quasi-stazionari, viene considerato e riproposto in una formulazione abbastanza generale da permettere lo studio non solo di una estesa classe di problemi di elettromagnetismo puro (la sua applicazione nella presente versione esclude la sola presenza di dispersività nello spazio o nel tempo dei mezzi materiali), ma anche di lasciare aperta una finestra verso il mondo esterno, proponendosi così come una valida alternativa nella soluzione di problemi accoppiati di tipo termico-elettromagnetico, meccanico-elettromagnetico, elettro-fluido-dinamico. Al vantaggio di un trattamento unificato dei problemi elettromagnetici, il presente metodo aggiunge il pregio di fornire informazioni immediate e precise circa la distribuzione spaziale e

temporale dell'errore numerico associato alla soluzione, in modo da suggerire un raffinamento opportunamente differenziato della discretizzazione ove questo appaia necessario.

Nelle presente tesi si applicherà innanzitutto il metodo della minimizzazione dell'errore costitutivo alle equazioni di Maxwell. Si mostrerà successivamente come il suddetto metodo sia estensibile al trattamento di problemi accoppiati, con particolare riferimento a quelli descritti dal modello MHD.

CAPITOLO I

Stato dell'arte sui metodi di analisi dei campi Elettromagnetici.

Introduzione

Le equazioni di Maxwell sono universalmente riconosciute come un esempio mirabile ed elegante di modello fisico-matematico per il carattere estremamente compatto col quale inquadrano ogni tipo di problematica si possa presentare nel settore dell'elettromagnetismo. Trascurando, quando possibile, in maniera opportuna qualcuno tra i termini sorgente ed aggiungendo relazioni costitutive, condizioni al contorno, ed, eventualmente, condizioni iniziali, si ricavano in maniera automatica i corretti modelli per qualunque problema di magnetostatica, di elettrostatica, di elettrocinematica, di propagazione. Nell'ipotesi di mezzi non dispersivi nel tempo o nello spazio, le equazioni per ciascuno dei particolari problemi elettromagnetici citati sono riportate in tabella I.

Ciascuno dei problemi che compaiono nella tabella è stato già soddisfacentemente risolto attraverso metodologie mirate di carattere analitico o numerico. Si sono cioè costruite delle formulazioni in termini di campo o di potenziale -differenziando maggiormente questi ultimi in ridotti o totali, scalari o vettori, elettrici o magnetici-, ciascuna adatto in particolare a trattare problemi che rientrino o in una sola delle classi citate, oppure, addirittura, in un particolare sottinsieme di una singola classe. E' il caso dei metodi analitici il cui impiego è tipicamente possibile solo per problemi il cui dominio di definizione sia particolarmente semplice, ad esempio riconducibile a un rettangolo, in geometria bidimensionale, o ad un parallelepipedo, se si opera in 3D, attraverso opportune trasformazioni di coordinate.

In questo capitolo sarà, innanzitutto, presentata una breve panoramica sulle diverse modalità operative disponibili per la soluzione di un problema E.M: metodi analitici, metodi numerici alle differenze finite, agli elementi finiti, agli elementi di frontiera.

In seguito, limitando l'attenzione alle sole formulazioni pensate per una implementazione numerica agli elementi finiti, saranno rivisitate, nella vasta letteratura disponibile, evidenziandone limiti e vantaggi, quelle che, per ciascuna delle principali classi di problemi elettromagnetici, maggiore successo hanno incontrato fino al momento attuale.

Il paragrafo conclusivo sarà, invece, dedicato alla presentazione del metodo basato sulla minimizzazione dell'errore costitutivo, soffermandosi sui principali risultati raggiunti nel recente passato nel trattamento di problemi E.M. stazionari e quasi-stazionari.

1.1) Generalità sui metodi analitici e numerici impiegati nella soluzione dei problemi elettromagnetici.

1.1.1) Metodi analitici

a) Funzione di Green

b) Metodo di separazione delle variabili

c) Metodo delle immagini

d) Metodo delle trasformazioni conformi

1.1.2) Metodi numerici

Nei vari settori dell'ingegneria e della fisica ci si trova spesso nella situazione di voler descrivere o predire la risposta di un sistema soggetto a influenze esterne che variano il suo stato.

Dal punto di vista pratico si procede dapprima a costruire un modello matematico approssimato del sistema, che includa tutte le informazioni essenziali del problema fisico, e, successivamente si procede alla sua inversione alla ricerca della soluzione. All'approssimazione associata alla modellizzazione matematica del sistema reale, è quasi sempre inevitabile aggiungere quella ulteriore inerente alla soluzione numerica del modello stesso. Sono proposte in letteratura varie possibili strategie per la risoluzione numerica di un modello matematico di un problema fisico. Tra tutte indugeremo su quelle che maggiore diffusione e successo hanno ottenuto negli ultimi anni: il *Metodo delle Differenze Finite (FDM)*, il *Metodo degli Elementi di Frontiera (BEM)*, ed il *Metodo degli Elementi Finiti (FEM)*. Tale descrizione appare opportuna oltre che per doveroso richiamo storico e per necessità di completezza, anche quale utile premessa all'argomento oggetto della tesi.

A scopo di chiarezza definiamo in modo descrittivo, piuttosto che matematico, quattro concetti che useremo ricorrentemente nel prosieguo del discorso e che risultano utili nell'impostazione del modello di qualunque problema di natura fisica: *sistema*, *dominio*, *equazioni governanti* e *condizioni di carico*.

Il *sistema* è tipicamente, ma non necessariamente, un oggetto fisico composto di vari materiali: solidi, liquidi, gas, plasmi, o combinazione di questi. Un sistema elettromagnetico può includere materiali e/o lo spazio vuoto.

Il *dominio* del problema è tipicamente la regione spazio/temporale nella quale interessa considerare il sistema. In termini matematici esso rappresenta l'insieme di definizione delle variabili indipendenti spazio e tempo. Talvolta tali

variabili indipendenti possono anche essere grandezze di natura diversa come frequenze o coefficienti spettrali.

Le *equazioni governanti* sono le relazioni che vincolano le variabili descrittive del sistema sulla base delle leggi fisiche del problema in esame. Tipicamente si tratta di equazioni integro-differenziali che esprimono la conservazione o il bilancio di qualche proprietà fisica, come la massa, la carica, il momento o l'energia. Queste possono essere anche delle equazioni che esprimono un principio variazionale, ed allora si presentano come un vincolo di minimo di un'opportuno funzionale come un'energia potenziale. Sono incluse tra le equazioni governanti anche le relazioni costitutive, che descrivono le particolare proprietà fisiche dei materiali costituenti il sistema.

Le *condizioni di carico* descrivono le sollecitazioni di origine esterna che interagiscono col sistema: forze, temperature, correnti e campi impressi, etc. I carichi agenti internamente al dominio, vengono chiamati carichi interni e compaiono nelle equazioni governanti. I carichi che intervengono sulla frontiera del dominio, sono invece chiamati carichi di frontiera, e compaiono in equazioni separate dette condizioni al contorno.

a) Metodo delle differenze finite (FDM)

b) Metodo degli elementi di frontiera(BEM)

Un problema di campo elettromagnetico è in generale definito da un'equazione differenziale e appropriate condizioni iniziali e al contorno. Avendo definito una funzione di Green che soddisfi l'equazione differenziale e le condizioni al contorno per una sorgente puntuale, la soluzione del problema di campo può essere espressa come l'integrale della funzione sorgente moltiplicata per la funzione di Green. Con questo sistema l'equazione

differenziale viene ricondotta ad un'equazione integrale che a volte può essere risolta più facilmente.

Durante gli anni '60 R.Harrington[1] sviluppò il metodo dei Momenti per la soluzione numerica delle equazioni integrali. Con questo metodo le equazioni integrali sono trasformate in equazioni algebriche attraverso l'introduzione di funzioni test e peso per il campo incognito.

Alla fine degli anni '70 C.Brebbia[2] ha proposto il cosiddetto metodo degli elementi di frontiera (BEM) dove le equazioni integrali sono risolte discretizzando la frontiera in Elementi Finiti, detti appunto gli Elementi di Frontiera. Al contrario dei metodi FEM che richiedono il trattamento di grosse matrici di tipo sparso, il BEM è caratterizzato da basso sforzo di modellizzazione e da matrici relativamente compatte, ma non simmetriche e densamente popolate. I metodi BEM sono stati applicati a problemi bidimensionali e tridimensionali di magnetostatica, elettrostatica, elettrocinematica, propagazione. I maggiori inconvenienti ad essi collegati possono essere riassunti nei seguenti punti:

- anche nel caso in cui l'operatore di partenza sia autoaggiunto le matrici associate al sistema da risolvere cui si perviene non sono nè simmetriche nè tantomeno definite positive, e non è sempre possibile avere sufficienti garanzie di convergenza.

- e' richiesta la soluzione numerica di un'equazione integrale sulla frontiera, operazione che può produrre forti malcondizionamenti (i cosiddetti "incubi algoritmici") a causa dei nuclei singolari presenti nelle funzioni di Green sotto il segno di integrale[3].

- Il vantaggio di ridurre fortemente lo sforzo nella generazione della mesh rispetto ai metodi FEM può non essere significativo da un punto di vista operativo. Difatti, generalmente nello studio di un problema pratico richiedono spesso richieste analisi ripetitive, o per la stima dell'errore o, abbastanza frequentemente, per iterazioni progettuali o per studi parametrici. Il costo connesso con la preparazione di una complicata mesh FE è pertanto solo una

piccola frazione del costo globale dell'analisi, ed anzi, grazie all'introduzione di sempre migliori preprocessor automatici tende a diventare sempre meno significativo. Un ulteriore punto a favore dei metodi FEM è costituito dalle proprietà locali di superconvergenza che non sono dimostrate per i metodi BEM

[1] **R.F.Harrington**. *Field computation by moment methods*. The Macmillan Company, New York, (1968)

[2] **C.A.Brebbia, S.Walker**, *Boundary element techniques in Engineering*, Newnes-Butterworths,London (1980)

[3] **R.Albanese** - *Tesi PHD*, p.41.

c) Metodo degli elementi finiti (FEM).

Il Metodo degli Elementi Finiti deve la sua attuale formalizzazione al contributo fornito negli ultimi decenni da vari ricercatori appartenenti ai diversi settori dell'ingegneria. Tra tutti il pioniere può essere considerato Courant con un lavoro che risale al 1943 [1]. Nuovi sostanziali apporti al metodo si ebbero tuttavia solo a partire dalla metà degli anni'50, grazie soprattutto alle favorevoli prospettive aperte dall'introduzione di calcolatori digitali dotati di memoria per l'immagazzinamento dei programmi e capaci di svolgere computazioni ad "alta velocità". L'attuale denominazione del metodo si deve attribuire a Clough[2]. Successivi apporti determinanti sono stati da una parte quelli, di Melosh, Besseling, Jones e de Veubeke, tutti comparsi tra il 1963 e il 1964, nei quali viene mostrato come il metodo FEM possa essere identificato come una forma del metodo variazionale di Ritz in cui vengono adoperate delle funzioni polinomiali come funzioni di base (Trial functions); e dall'altra quelli del 1965 di Zienkiewicz e Cheung che hanno allargato enormemente il dominio di applicazione del FEM a tutti i problemi di campo che possono essere posti in forma variazionale. A partire da quella data i metodi FEM hanno conosciuto vastissima diffusione ed innumerevoli sono i loro campi di applicazione: meccanica dei solidi, elasticità,

elettromagnetismo, termodinamica, acustica, fenomeni accoppiati, e praticamente ogni problema matematico o fisico descritto da equazioni differenziali, integrali, integro-differenziali o variazionali.

Diamo ora una sintetica descrizione del modo in cui si opera nei metodi FEM.

Il dominio del problema è discretizzato in regioni più piccole chiamate elementi. Elementi adiacenti si toccano senza sovrapporsi, e non esiste alcun spazio del dominio non coperto da elementi. Le forme degli elementi sono intenzionalmente costruite nella maniera più semplice possibile: generalmente triangoli o quadrilaterali in 2D, e tetraedri, pentaedri o esaedri in geometria 3D. L'intero mosaico ottenuto prende il nome di griglia ("mesh"). Il processo di generazione della mesh, un tempo laborioso perchè effettuato in forma manuale, si è ora velocizzato notevolmente grazie all'introduzione di preprocessori automatici che accettano quali dati in ingresso delle macro-descrizioni del dominio da discretizzare.

Per ogni elemento le equazioni governanti, usualmente in forma differenziale o variazionale (integrale), sono trasformate in equazioni algebriche, dette equazioni di elemento, tramite un processo di approssimazione. Questo passo viene eseguito attraverso la proiezione del campo incognito in uno spazio funzionale la cui base è costituita da funzioni di tipo polinomiale ciascuna delle quali assume valore diverso da zero solo in una limitata regione della griglia, costituita da uno o al più pochi elementi tra loro adiacenti. I gradi di libertà della base dello spazio funzionale nel quale viene ricercata la soluzione, vengono associati a parti geometriche, o *simplices*, della mesh - e cioè, nodi, lati, facce, volumi-. A seconda della scelta operata si parla di *node-elements*, *edges-elements*, *facet-elements*, *volume-elements*.

Invece di calcolare direttamente le funzioni di forma associate ai simplices degli elementi reali della mesh si preferisce adoperare una via indiretta. Si supponga di rappresentare la mesh del dominio in un sistema di riferimento

cartesiano e sia \mathbf{x} il vettore delle coordinate del generico punto della mesh in tale sistema (\mathbf{x} è definito *vettore delle coordinate reali o globali*). Si introduce un elemento *master* (o *parent*) di caratteristiche geometriche regolari (ad esempio per trattare elementi esaedrali, l'elemento master è un cubo di spigolo di lunghezza unitaria) in un sistema di coordinate locali ξ . Si definiscono per l'elemento master le funzioni di forma $\phi_i(\xi)$ (dove, ad esempio, per i nodes-elements, i va da 1 al numero dei nodi dell'elemento), associate al particolare tipo di simplices che si intende utilizzare (sono queste le cosiddette *parent shape functions*). A questo punto per ciascuno degli elementi reali, (e), della mesh si definisce una trasformazione invertibile $\mathbf{x}=\chi^{(e)}(\xi)$ che mappi l'elemento master sull'elemento reale. Le funzioni di forma associate a ciascun elemento reale della mesh possono così essere calcolate attraverso le parent shape functions e la funzione inversa di mapping dell'elemento $\xi=\xi^{(e)}(\mathbf{x})$ come:

$$\phi_i^{(e)}(\mathbf{x})=\phi_i^{(e)}(\xi^{(e)}(\mathbf{x}))$$

Tra i vari possibili modi di stabilire il mapping, viene frequentemente adoperato il cosiddetto *approccio isoparametrico*:

$$\mathbf{x} = \sum_{k=1}^l \mathbf{x}_k^{(e)} \phi_k(\xi)$$

dove $\mathbf{x}_k^{(e)}$, $k=1,2,\dots,l$ sono le coordinate degli l nodi nell' (e)^{mo} elemento reale della mesh, mentre le $\phi_k(\xi)$ sono le funzioni di forma nodali. Il nome isoparametrico deriva dal fatto che coi nodes-elements si adoperano le stesse funzioni $\phi_k(\xi)$ per approssimare il campo incognito $\mathbf{U}(\mathbf{x})$.

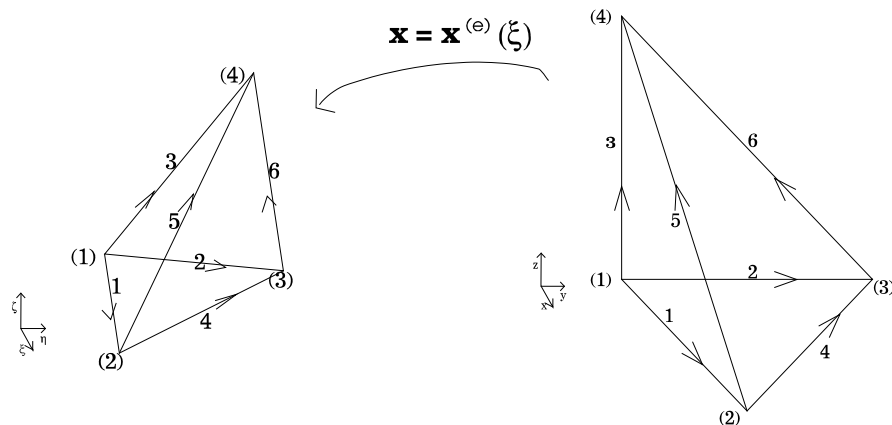


Fig.[] Mapping dell'elemento master tetraedrale sull'elemento reale (e) attraverso la trasformazione $\mathbf{x} = \chi^{(e)}(\xi)$

L'opportunità di trasformare le equazioni di partenza in equazioni algebriche consente una radicale semplificazione delle complessità di calcolo e permette l'utilizzo di una delle numerose procedure risolutive ciascuna delle quali specializzata nell'affrontare una particolare classe di sistemi di equazioni algebriche.

Due aspetti importanti vanno rimarcati. Primo, le equazioni di elemento sono dal punto di vista della struttura algebrica identiche per tutti gli elementi che sono dotati delle stesse caratteristiche geometriche. Pertanto lo sforzo analitico di formulare le equazioni di elemento va eseguito una sola volta per ciascuno degli elementi tipici utilizzati (solitamente uno o due) e non per tutti gli elementi della mesh. Secondo, poichè tanto la geometria degli elementi, quanto la struttura dello spazio funzionale approssimante vengono scelti di struttura particolarmente semplice, la scrittura del sistema algebrico delle equazioni di elemento è molto agevole e non comprende più di pochi tipi di equazioni algebriche (tipicamente da 2 a 20).

I coefficienti delle equazioni di elemento sono valutati numericamente, tramite l'ausilio di un calcolatore, per ciascun elemento della mesh e successivamente assemblati in una matrice complessiva di dimensioni solitamente molto più grandi detta delle *equazioni di sistema*. Tale sistema di

equazioni può tuttavia essere risolto in tempi ragionevoli sfruttando il carattere sparso della matrice dei coefficienti, risultato della scelta della base di proiezione.

Solo a questo livello vengono imposte le condizioni al contorno ed, eventualmente, quelle di interfaccia tra sottoinsiemi del dominio attraverso l'opportuna modifica delle equazioni di sistema o l'aggiunta di ulteriori equazioni di vincolo.

Resta ora da chiarire cosa siano le equazioni governanti per il metodo degli elementi finiti. A questo scopo definiamo una quantità chiamata residuo. Si supponga, per semplicità, che l'intero modello del problema sia rappresentato da una singola equazione differenziale. Si trasferiscano tutti i termini al lato sinistro (LHS), di modo che il lato destro (RHS) risulti nullo. Se ora si sostituisce alla funzione incognita, l'esatta soluzione $\mathbf{U}(\mathbf{x})$ dell'equazione differenziale, il LHS dell'equazione si annullerebbe per ogni \mathbf{x} eguagliando pertanto il RHS. Se, invece dell'esatta soluzione, ne usassimo nella sostituzione una approssimata $\mathbf{U}_i(\mathbf{x}, \mathbf{a})$ -dove \mathbf{a} rappresenta il vettore dei parametri liberi associati alle funzioni approssimanti $\phi_i(\mathbf{x})$ - , la differenza sarebbe una funzione totalmente diversa da zero detta residuo dell'equazione, e denotata da $\mathbf{R}(\mathbf{x}, \mathbf{a})$. Il principio ispiratore di tutti i metodi basati sui residui pesati è che la soluzione numerica di un sistema di equazioni può essere affidata alla ricerca di quell'insieme di parametri capace di rendere in tutto il dominio di definizione la funzione residuo prossima per quanto è possibile alla funzione nulla.

Ciò che, all'interno di questa classe di metodi, distingue uno schema da un altro è il criterio col quale si definisce la vicinanza tra la funzione nulla e la funzione residuo. Nascono così i metodi del collocamento, del sottodominio, dei minimi quadrati, di Galerkin. Quest'ultimo, in particolare, impone che, per ogni parametro \mathbf{a}_i , una media pesata di $\mathbf{R}(\mathbf{x}, \mathbf{a})$ sull'intero dominio sia zero. Le funzioni peso coincidono con le funzioni di base $\phi_i(\mathbf{x})$ associate a ciascun \mathbf{a}_i . Si

ottiene in questo modo un sistema in cui le equazioni -dette governanti- sono in numero pari al numero di funzioni di base e, dunque, di incognite.

Il metodo dei residui di Galerkin è strettamente legato ad un quinto schema solutivo che trova vasta rilevanza nelle formulazioni FEM e che va sotto il nome di metodo Variazionale di Ritz. Mentre la parola variazionale riporta al vasto e complesso ambito del calcolo delle variazioni, il nome Ritz richiama una specifica, semplice e ben definita tecnica per ottenere soluzioni numeriche approssimate a problemi descritti in forma variazionale. Con un'espressione di carattere generale, si può affermare che vari problemi retti da sistemi di equazioni differenziali, ammettono una formulazione equivalente di carattere variazionale. In questo caso, se denotiamo con $I(\mathbf{U})$ un possibile funzionale integrale (che a volte può essere interpretato da un punto di vista fisico come un'energia associata al sistema), a valore scalare, equivalente al modello differenziale del problema in oggetto, il metodo variazionale di Ritz consiste nel ricercare all'interno dello spazio delle funzioni base la funzione $\mathbf{U}_i(\mathbf{x}, \mathbf{a})$ che trova un estremo del funzionale $I(\mathbf{U})$. Si può dimostrare che le equazioni di elemento alle quali si perviene sono identiche a quelle ottenute attraverso il metodo di Galerkin. Va tuttavia osservato che il metodo di Galerkin ha un maggiore spettro di applicazioni rispetto a quello di Ritz, non richiedendo necessariamente che il problema ammetta una formulazione variazionale. D'altro canto la formulazione di Ritz, ove applicabile, risulta concettualmente superiore a quella di Galerkin. Si consideri ad esempio un sistema fisico in equilibrio stabile: ad esso corrisponde un minimo della sua energia potenziale; in questo caso il principio variazionale associato alle equazioni è un principio di minimo. In problemi siffatti la sequenza di valori dell'integrale $I(\mathbf{U}_i)$, corrispondente ad una sequenza di soluzioni approssimate, risulta convergere monotonicamente a patto che la sequenza degli spazi delle soluzioni sia costruita in modo che ogni termine della sequenza includa come suo sottinsieme tutte le soluzioni rappresentabili attraverso lo spazio delle soluzioni dei termini precedenti

(il che può essere ottenuto aggiungendo nuove funzioni di base senza modificare o rimuovere quelle precedenti, ovvero partizionando alcuni degli elementi della mesh in elementi più piccoli). Grazie alla monotonia della convergenza può in ogni momento essere stabilito un limite superiore al valore del funzionale $I(\mathbf{U})$. Inoltre, se esiste anche un principio variazionale complementare, ovvero se esiste la possibilità di trovare una espressione di I in un ulteriore campo incognito \mathbf{W} , per il quale il principio variazionale governante sia un principio di massimo, allora è possibile stabilire anche un limite superiore per il valore del funzionale. Trovare limiti per $I(\mathbf{U})$ può essere importante quando il funzionale rappresenta una quantità fisica di interesse; è questo il caso ad esempio dell'energia potenziale di una configurazione di equilibrio o della frequenza di risonanza di un sistema vibrante.

[1] **R.Courant**, *Variational methods for the solutions of Problems of Equilibrium and Vibrations*. Bull Am.Math.Soc., 49:1-23, (1943)

[2] **R.W.Clough**, *The Finite Element Method in Plane Stress Analysis*, Proc.2d ASCE Conf. Electronic Computation, Pittsburgh, PA., **Sept.1960**.

1.2) Modelli E.M. classici per formulazioni F.E.M.

1.2.1) Magnetostatica

Il calcolo del campo magnetostatico in geometria tridimensionale avviene usualmente attraverso la formulazione basata sul potenziale scalare.

Secondo l'approccio basato sul potenziale scalare ridotto il campo magnetico \mathbf{H} è diviso nelle sue componenti irrotazionale \mathbf{H}_{irr} e solenoidale \mathbf{H}_{sol} . La parte solenoidale \mathbf{H}_{sol} può essere calcolata facilmente attraverso la legge dell'azione elementare di Laplace:

$$\mathbf{H}_{\text{sol}}(\mathbf{r}) = \frac{1}{4\pi} \int_{V_j} \frac{\mathbf{J}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} dV'$$

dove V_j è la regione in cui la densità di corrente \mathbf{J} è diversa da zero.

La parte irrotazionale di \mathbf{H}_{irr} , viene invece rappresentata come il gradiente di un potenziale scalare ϕ :

$$\mathbf{H}_{\text{irr}} = -\nabla\phi,$$

per cui risulta:

$$\mathbf{H} = \mathbf{H}_{\text{sol}} - \nabla\phi$$

Il potenziale scalare ridotto è poi ottenuto risolvendo la seguente equazione:

$$\nabla \cdot \mu(\mathbf{H}_{\text{sol}} - \nabla\phi) = 0$$

che impone la solenoidalità del campo di induzione magnetica \mathbf{B} .

Il principale svantaggio di quest'approccio è la scarsa accuratezza dovuta ai fenomeni di cancellazione tra campi calcolati in modo indipendente, specialmente nelle regioni in cui la permeabilità magnetica è elevata. In aggiunta, uno sforzo computazionale notevole deve essere eseguito durante il calcolo dell'integrale in ogni punto del dominio ove \mathbf{H} è richiesto.

Per superare questi problemi, una formulazione alternativa combina l'uso del potenziale scalare ridotto con quello del potenziale scalare totale. In questo caso il campo magnetico è descritto da

$$\mathbf{H} = \mathbf{H}_{\text{sol}} - \nabla\phi \quad \text{in } V_j$$

$$\mathbf{H} = -\nabla\psi \quad \text{nella regione esterna } V - V_j$$

I potenziali scalari sono così determinati dalle seguenti equazioni:

$$\nabla \cdot \mu(\mathbf{H}_{\text{sol}} - \nabla\phi) = 0 \quad \text{in } V_j$$

$$\nabla \cdot \mu \nabla\psi = 0 \quad \text{nella regione esterna } V - V_j$$

che vengono accoppiate attraverso le condizioni di interfaccia sulla frontiera ∂V_j (cioè la continuità di $\mathbf{H} \times \mathbf{n}$ e $\mathbf{B} \cdot \mathbf{n}$ dove \mathbf{n} è il versore normale a ∂V_j).

Questa impostazione, tuttavia, riduce il problema della cancellazione solo se la permeabilità è costante nella regione V_j . Inoltre, a causa della polidromia della funzione potenziale scalare per regioni molteplicemente connesse devono essere introdotti opportuni tagli.

Questi problemi scompaiono laddove siano adottate formulazioni basate su potenziali vettori. Il prezzo da pagare è in questo caso costituito dall'accresciuto numero di incognite scalari e dalla necessità, in alcuni approcci, di imporre condizioni addizionali sui potenziali vettori allo scopo di garantire l'unicità della soluzione.

Due classiche formulazioni di quest'ultimo tipo sono basate l'una sul potenziale vettore magnetico \mathbf{A} (definito da $\mathbf{B} = \nabla \times \mathbf{A}$) e l'altra sul potenziale vettore elettrico \mathbf{T} (definito da $\mathbf{J} = \nabla \times \mathbf{T}$). Anche queste formulazioni possono essere usate in congiunzione coll'approccio del potenziale scalare totale.

Come detto, vincoli addizionali, detti Gauges, devono essere imposti allo scopo di avere una corrispondenza biunivoca tra i potenziali vettori e i loro rotori. Una tra le Gauge più comuni nel problema magnetostatico è quella di Coulomb:

$$\nabla \cdot \mathbf{A} = 0 \quad \nabla \cdot \mathbf{T} = 0$$

Un'altra possibile scelta è :

$$\mathbf{w} \cdot \mathbf{A} = 0 \quad \mathbf{w} \cdot \mathbf{T} = 0$$

dove \mathbf{w} è un arbitrario campo vettoriale non dotato di linee di campo chiuse.

Adottando quest'ultima gauge e usando congiuntamente i potenziali vettoriali e i potenziali scalari totali, si arriva alle due formulazioni complementari magnetica (\mathbf{A} - ψ) ed elettrica (\mathbf{T} - Ω).

Nel caso di mezzi lineari, la formulazione magnetica conduce al seguente sistema:

$$\begin{aligned} \nabla \times (\mu^{-1} \nabla \times \mathbf{A}) &= \mathbf{J} && \text{in } V_j \\ \mathbf{w} \cdot \mathbf{A} &= 0 && \text{in } V_j \\ \nabla \cdot \mu \nabla \psi &= 0 && \text{nella regione esterna } V - V_j \end{aligned}$$

Per quanto riguarda l'imposizione delle condizioni di interfaccia e l'eventuale introduzione di tagli, la formulazione presenta caratteristiche simili a quelle che combinano l'uso dei potenziali scalari ridotto e totale, col vantaggio, però, di eliminare i problemi legati alla cancellazione. Con questo approccio, nel caso di domini a connessione multipla, è, tuttavia, possibile, estendere la regione V_j , in modo da rendere semplicemente connessa la regione esterna alle correnti ed avere di conseguenza una funzione potenziale scalare ψ monodroma.

La formulazione basata sul potenziale vettore elettrico \mathbf{T} è nota come il metodo \mathbf{T} - Ω . In questo metodo \mathbf{T} differisce da \mathbf{H} per il gradiente di un potenziale scalare Ω :

$$\mathbf{H} = \nabla \mathbf{T} - \Omega$$

Tenendo conto che il potenziale vettore elettrico può essere definito solo nella regione ove la densità di corrente è diversa da zero, utilizzando per il resto del dominio il potenziale scalare totale, si perviene al sistema :

$$\begin{array}{ll}
 \nabla \times \mathbf{T} = \mathbf{J} & \text{in } V_j \\
 \mathbf{w} \cdot \mathbf{T} = 0 & \text{in } V_j \\
 \nabla \cdot \mu (\mathbf{T} - \nabla \Omega) = 0 & \text{in } V_j \\
 \nabla \cdot \mu \nabla \Omega = 0 & \text{nella regione esterna } V - V_j
 \end{array}$$

Per quel che concerne tagli e condizioni di interfaccia valgono considerazioni analoghe a quelle relative alla formulazione $\mathbf{A}-\psi$.

Potrebbe sembrare che l'approccio $\mathbf{T}-\Omega$ sia simile a quello basato sui due potenziali scalari (\mathbf{H}_{sol} potrebbe, infatti, essere interpretato come il particolare potenziale vettore elettrico \mathbf{T} avente divergenza nulla). La differenza tra le due formulazioni è che nell'approccio coi potenziali scalari \mathbf{T} , o se si preferisce \mathbf{H}_{sol} , viene calcolato in maniera analitica, mentre in quello basato sul potenziale vettore elettrico la sua valutazione è ottenuta per via numerica proiettandolo sullo stesso spazio funzionale adoperato per approssimare il gradiente di Ω , la qualcosa limita notevolmente il problema degli errori di cancellazione.

[*R.Albanese, G.Rubinacci, Magnetostatic field computations in terms of two-component vector potentials, Int. j. num.meth. eng.,29,515-532,1990*

1.2.2) Elettrostatica

1.2.3) Elettrocinemica quasi-stazionaria

1.2.4) Propagazione

1.3) Rassegna sui metodi basati sui principi variazionali complementari. Il metodo della minimizzazione dell'errore costitutivo.

1.3.1) CAMPI STATICI

La formulazione basata sulla minimizzazione dell'errore costitutivo è nata nell'ambito della teoria dei principi variazionali complementari (**CVP**). I metodi CVP, già brevemente introdotti nelle pagine precedenti, sono stati sviluppati originariamente per applicazioni nei settori della meccanica analitica e dell'elasticità e, successivamente adattati per essere applicati in problemi di elettromagnetismo[1,2]. Questi metodi, essenzialmente equivalenti per problemi statici, sono simili nel fatto che essi pongono prima una serie di ipotesi sulle specifiche del problema, ad esempio escludono domini molteplici e connessi e sorgenti superficiali oppure pongono altre limitazioni restringendo la classe dei problemi trattabili, e poi derivano in vario modo le espressioni dei funzionali complementari.

In particolare, Penman e Fraser[2] identificano due insiemi di equazioni canoniche che descrivono la struttura matematica del problema di campo stazionario, e le integrano alternativamente nei funzionali di Hu-Washizu, funzionali usati originariamente nella teoria dell'elasticità. Giungono così a due coppie di funzionali complementari. La loro formulazione, pur rimarchevole per la concisione con la quale viene evidenziata la struttura del problema di campo statico, fallisce in presenza di mezzi non-lineari, di regioni molteplici e connesse, di sorgenti superficiali ed, inoltre, in presenza di condizioni al contorno che, anche se fisicamente ammissibili, non riescono sempre ad essere convenientemente inquadrare nel modello da loro proposto.

Hammond[1] parte, invece, dalle equazioni della meccanica analitica formulate attraverso il principio dei lavori virtuali. Quindi, con appropriate sostituzioni, ed attraverso trasformazioni degli integrali di volume risultanti e integrazioni rispetto ai parametri variati, perviene a due coppie di funzionali

complementari. Rispetto alla formulazione precedente questa se ne avvantaggia in due punti: consente il trattamento di non-linearità e propone una interpretazione fisica a processi variazionali di carattere essenzialmente matematico (il funzionale da lui proposto si ricollega direttamente alla teoria Hamiltoniana)..

Rikabi si mette sulla scia di questa attività, giungendo ad una formulazione che ha valenza molto più generale. Nel suo lavoro [3], si preoccupa innanzitutto di definire da una parte la struttura formale di un problema stazionario ben posto e dall'altra il range di applicabilità della teoria CVP.

Partendo dall'ipotesi di relazioni costitutive monotone, ed usando congiuntamente simboli magnetostatici -per i campi e le distribuzioni di corrente- ed elettrostatici -per le distribuzioni di carica- sviluppa una formulazione valida per ogni problema elettromagnetico di natura stazionaria. Con riferimento alle due coppie di campi $(\mathbf{H}_1(\mathbf{r}), \mathbf{B}_1(\mathbf{r}))$ e $(\mathbf{H}_2(\mathbf{r}), \mathbf{B}_2(\mathbf{r}))$ nelle ipotesi che:

- I) i campi $\mathbf{H}(\mathbf{r})$ rispettino la legge di Ampere;
- II) i campi $\mathbf{B}(\mathbf{r})$ rispettino la legge di Gauss;
- III) per tutti i campi valgano le rispettive condizioni di interfaccia associate alle leggi;
- IV) che ciascuna coppia (\mathbf{H}, \mathbf{B}) soddisfi la relazione costitutiva;

dimostra che, se le condizioni al contorno sono imposte in termini di \mathbf{H}_t , \mathbf{B}_n e di regolarità dei campi all'infinito, l'unicità della soluzione è garantita.

Partendo da queste premesse, e conservando l'ipotesi di monotonia per le relazioni costitutive, propone di usare un modello composto da due sistemi complementari, comunicanti solo attraverso la relazione costitutiva e basati rispettivamente sui campi incogniti \mathbf{H} e \mathbf{B} , in modo che il sistema \mathbf{H} rispetti la legge di Ampere e la relativa condizione d'interfaccia, e il sistema \mathbf{B} soddisfi la legge di Gauss con la condizione d'interfaccia associata.

Attraverso le espressioni dell'energia e della coenergia associate ai campi, costruisce, quindi, un funzionale scalare $\Lambda(\mathbf{H}, \mathbf{B})$ definito positivo, il cui valore è zero solo nell'ipotesi che i sistemi complementari forniscano campi che soddisfano la relazione costitutiva.

Introducendo un potenziale scalare Ω nel sistema \mathbf{H} e un potenziale vettore \mathbf{A} nel sistema \mathbf{B} in modo che i campi siano espressi attraverso le relazioni

$$\mathbf{H} = \mathbf{G} - \nabla\Omega \quad (1)$$

$$\mathbf{B} = \mathbf{C} + \nabla \times \mathbf{A} \quad (2)$$

dove \mathbf{G} e \mathbf{C} sono campi predefiniti per soddisfare le relazioni

$$\nabla \times \mathbf{G} = \mathbf{J} \quad (3)$$

$$\nabla \cdot \mathbf{C} = \rho_0 \quad (4)$$

mostra come usando alternativamente le relazioni (1) e (2) e le condizioni al contorno, purchè queste rientrino tra quelle dichiarate ammissibili nella dimostrazione di unicità, sia possibile ottenere la separazione del funzionale $\Lambda(\mathbf{H}, \mathbf{B})$ in due coppie duali di funzionali complementari. Si possono così ottenere in principio ben quattro funzionali, ognuno dipendente dalle variabili appartenenti ad uno solo tra i due sistemi \mathbf{H} e \mathbf{B} , e tutti capaci di fornire, attraverso una procedura di estremizzazione, una soluzione valida.

$$\begin{aligned} \Lambda(\mathbf{H}, \mathbf{B}) &= \Theta_p(\mathbf{H}) + \Xi_p(\mathbf{B}) \\ &= \Theta_d(\mathbf{H}) + \Xi_d(\mathbf{B}). \end{aligned} \quad (5)$$

Nella pratica, tuttavia, Rikabi consiglia di scegliere tra le decomposizioni duali dei funzionali quella che risulta più conveniente nel particolare problema in esame.

Rikabi mostra, infine, come, ai funzionali Θ e Ξ sia possibile imporre dei limiti inferiori e superiori, e ottenere una utile stima dell'errore commesso sulla relazione costitutiva. Nell'impostazione di Rikabi rimane il limite di non poter trattare problemi per i quali le condizioni al contorno sono specificate attraverso relazioni tra le variabili del sistema \mathbf{H} e del sistema \mathbf{B} . In questi

casi, difatti, non è più possibile lo splitting del funzionale $\Lambda(\mathbf{H},\mathbf{B})$. Resta naturalmente aperta la possibilità di arrivare ad una soluzione attraverso la diretta minimizzazione di $\Lambda(\mathbf{H},\mathbf{B})$, anche se, così facendo, nel processo solutivo dovranno essere coinvolte contemporaneamente le variabili di entrambi i sistemi

[1]**P.Hammond**, *Energy Methods in Electromagnetism*, Clarendon Press, Oxford,(1981).

[2]**J.Penman,J.R.Fraser**, 'Dual and complementary methods in electromagnetism', IEEE Trans., MAG-19, 2311-2316,(1983).

[3]**J.Rikabi,C.F.Bryant,E.M.Freeman**, 'An error-based approach to complementary formulations of static field solutions', Int.j.numer.methods eng., 26, (1988)

1.3.2) Elettrocinesica quasi-stazionaria

(Ref. Albanese, Rikabi)

1.3.3) Propagazione in regime sinusoidale

(Ref. L.Pichon, A.Bossavit)

CAPITOLO II

L'approccio basato sulla minimizzazione dell'errore costitutivo nell'Elettromagnetismo

2.1) Illustrazione del metodo proposto per la trattazione di problemi E.M. in dominio 3D al finito costituito da mezzi non-lineari, anisotropi, tempo-varianti, non dispersivi nel tempo e nello spazio. Rimozione della

cura nella ricerca dello splitting. Vantaggi di questa impostazione rispetto all'approccio Liguriano: maggiore generalita'. Svantaggi: le incognite elettriche e magnetiche non sono piu' disaccoppiate: maggiore carico computazionale. Scelta dei pesi per le incognite elettriche e magnetiche.

La soluzione delle equazioni di Maxwell nella loro forma completa nel dominio del tempo e in geometria tridimensionale è a tutt'oggi un problema di fondamentale importanza per il quale si è ben lungi dall'aver fornito risposte pienamente soddisfacenti. Basti dire a tal proposito che le prime metodologie proposte, tutte naturalmente di carattere numerico e, in generale, basate sui metodi FEM, risalgono solo al recente passato[1,2].

Una impostazione corretta del problema elettromagnetico include le due fondamentali equazioni di campo (vale a dire la legge di Faraday e la legge di Ampère), le relazioni costitutive e naturalmente le condizioni iniziali, al contorno e sulle interfacce.

I metodi tradizionali forzano le relazioni costitutive ed una sola delle due equazioni di campo, accontentandosi di soddisfare approssimativamente l'altra.

L'approccio basato sulla minimizzazione dell'errore costitutivo, qui trattato, adopera una formulazione globale in termini di due potenziali vettori incogniti, l'uno elettrico, l'altro magnetico. La soluzione viene quindi calcolata imponendo la minimizzazione di un funzionale correlato all'errore sulle relazioni costitutive.

Questo schema provvede due soluzioni complementari, e proprio in questa sua caratteristica, si trovano contemporaneamente il suo più grande vantaggio e il maggiore inconveniente. Da un lato, infatti, le soluzioni complementari forniscono la stima degli errori locali e globali associati alla soluzione numerica trovata, informazioni queste che possono essere vantaggiosamente utilizzate nella fase di raffinamento della discretizzazione del problema sotto studio. D'altra parte l'introduzione di due potenziali vettori comporta un forte aumento

del carico computazionale (sei incognite scalari rispetto alle tre delle formulazioni tradizionali).

E' stato già osservato che il metodo è stato introdotto per trattare casi statici e quasi-statici. La formulazione qui presentata generalizza il metodo di modo da consentirne l'utilizzazione per una più vasta classe di problemi elettromagnetici che comprende quelli nei quali le correnti di spostamento non possono essere trascurate. Più precisamente, almeno dal punto di vista concettuale, il modello sembra adeguato per il trattamento di ogni problema tridimensionale elettromagnetico che sia caratterizzato da relazioni costitutive eventualmente non-lineari, anisotrope, non-omogenee e tempo-varianti, purchè monotone e non dispersive nel tempo e nello spazio.

[1] **C.W.Trowbridge**, "Overview of the 'State of the Art' in Electromagnetic analysis and design", in "Industrial Application of Electromagnetic Computer Codes", (Crutzen,Molinari,Rubinacci), p.4, Kluwer Academic Publishers, **1990**.

[2] Computer Physics Communications, Vol.68 (special issue), **1991**

2.2) Sviluppo analitico del metodo proposto. Modello matematico. Condizioni iniziali . Condizioni al contorno . (Condizioni all'infinito). Funzionale d'errore e sue derivate. Espressioni complementari per l'energia elettrostatica e magnetostatica, per la potenza ohmica, per il vettore di Poynting.

Un qualunque problema di elettromagnetismo definito nel dominio spaziale V e nell'intervallo temporale $[0, T]$ può essere formulato attraverso le equazioni di Maxwell che vengono qui presentate in forma differenziale.

$$\nabla \cdot \mathbf{D}(\mathbf{x},t) = \rho(\mathbf{x},t) \quad (1a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{x},t) = 0 \quad (1b)$$

$$\nabla \times \mathbf{H}(\mathbf{x},t) = \mathbf{J}(\mathbf{x},t) + \partial \mathbf{D}(\mathbf{x},t) / \partial t \quad (1c)$$

$$\nabla \times \mathbf{E}(\mathbf{x},t) = -\partial \mathbf{B}(\mathbf{x},t)/\partial t \quad (1d)$$

$$\nabla \cdot \mathbf{J}(\mathbf{x},t) = -\partial \rho(\mathbf{x},t)/\partial t \quad (1e)$$

valide $\forall \mathbf{x} \in V$ e $\forall t \in [0, T]$.

Alle espressioni (1) si accompagnano le relazioni di interfaccia valide sulle eventuali superfici di separazione Σ tra materiali diversi

$$[\mathbf{D}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = \rho_\Sigma(\mathbf{x},t) \quad (2a)$$

$$[\mathbf{B}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (2b)$$

$$[\mathbf{E}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (2c)$$

$$[\mathbf{H}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma(\mathbf{x},t) \quad (2d)$$

$$[\mathbf{J}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = -\partial \rho_\Sigma(\mathbf{x},t)/\partial t \quad (2e)$$

che devono essere soddisfatte $\forall \mathbf{x} \in \Sigma$ e $\forall t \in [0, T]$

Perchè la soluzione del problema elettromagnetico sia unica è necessario aggiungere $\forall \mathbf{x} \in V$ alle equazioni di Maxwell le condizioni iniziali:

$$\mathbf{D}(\mathbf{x},0) = \mathbf{D}_o(\mathbf{x}), \quad (3a)$$

$$\mathbf{B}(\mathbf{x},0) = \mathbf{B}_o(\mathbf{x}) \quad (3b)$$

le relazioni costitutive:

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t), \quad (4a)$$

$$\mathbf{D}(\mathbf{x},t) = \mathbf{f}_D(\mathbf{E}, \mathbf{x}, t), \quad (4b)$$

$$\mathbf{J}(\mathbf{x},t) = \mathbf{f}_J(\mathbf{E}, \mathbf{x}, t) \quad (4c)$$

e, indicata con ∂V la frontiera esterna di V , $\forall \mathbf{x} \in \partial V$ e $\forall t \in [0, T]$ opportune condizioni al contorno.

Ricordiamo, per inciso, che la relazione (1.e) si può ricavare dalle relazioni (1c) e (1a).

Assumiamo ora le seguenti ipotesi:

I) Le relazioni costitutive (4), anche se eventualmente non-lineari, anisotrope, non-omogenee e tempo-varianti, siano comunque non dispersive nel tempo e nello spazio. In tal caso le (4a,b,c) diventano:

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t), \mathbf{x}, t), \quad (4a)$$

$$\mathbf{D}(\mathbf{x},t) = \mathbf{f}_D(\mathbf{E}(\mathbf{x},t), \mathbf{x}, t), \quad (4b)$$

$$\mathbf{J}(\mathbf{x},t) = \mathbf{f}_J(\mathbf{E}(\mathbf{x},t), \mathbf{x}, t) \quad (4c)$$

II) Le relazioni costitutive (4) siano monotone, il che si può porre in forma matematica imponendo che per ogni trasformazione elementare siano valide le seguenti disequaglianze:

$$\Delta \mathbf{B} \cdot \Delta \mathbf{H} \geq 0 \quad (4d)$$

$$\Delta \mathbf{D} \cdot \Delta \mathbf{E} \geq 0 \quad (4e)$$

$$\Delta \mathbf{J} \cdot \Delta \mathbf{E} \geq 0 \quad (4f)$$

Si ponga:

$$\mathbf{D}_T(\mathbf{x},t) = \mathbf{D}_{T_0}(\mathbf{x}) + \int_0^t (\mathbf{J}(\mathbf{x},t) + \partial \mathbf{D}(\mathbf{x},t) / \partial t) dt \quad (5)$$

Dove per il teorema di Helmholtz si può scrivere

$$\mathbf{D}_T(\mathbf{x},t) = \mathbf{D}_T^{(irr)}(\mathbf{x},t) + \mathbf{D}_T^{(sol)}(\mathbf{x},t)$$

con la posizione:

$$\nabla \cdot \mathbf{D}_T^{(irr)}(\mathbf{x}) = \nabla \cdot \mathbf{D}_{T_0}^{(irr)}(\mathbf{x}) = \rho_{T_0}(\mathbf{x}) \quad (6)$$

ove al termine \mathbf{D}_T si è assegnato il nome di "Vettore di spostamento totale" per sottolinearne il significato di generalizzazione di \mathbf{D} , "Vettore di spostamento". La densità di carica $\rho_{T_0}(\mathbf{x})$ rappresenta la sorgente del campo di background $\mathbf{D}_T^{(irr)}$, un campo cioè che non modifica nè è modificato da alcuno dei fenomeni elettromagnetici che si verificano dopo l'istante $t=0$. Per i nostri scopi è dunque lecito assumere per $\rho_{T_0}(\mathbf{x})$ un valore qualsiasi e in particolare

$\rho_{T_0}(\mathbf{x})=0$. Giacchè anche il valore della componente solenoidale \mathbf{D}_{T_0} non modifica l'evoluzione dei successivi fenomeni elettromagnetici si può porre senz'altro $\mathbf{D}_{T_0}(\mathbf{x})=0$.

Le equazioni di Maxwell si possono allora riscrivere nella forma seguente:

$$\nabla \cdot \mathbf{D}_T = 0 \quad (7a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (7b)$$

$$\nabla \times \mathbf{H} = \partial \mathbf{D}_T / \partial t \quad (7c)$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t \quad (7d)$$

e le condizioni di continuità sulle superfici di interfaccia comportano:

$$[\mathbf{D}_T(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (8a)$$

$$[\mathbf{B}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (8b)$$

$$[\mathbf{E}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (8c)$$

$$[\mathbf{H}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (8d)$$

Le relazioni costitutive (4a,b,c) si possono condensare nelle seguenti espressioni:

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t), \mathbf{x}, t), \quad (9a)$$

$$\mathbf{D}_T(\mathbf{x},t) = \mathbf{D}_{T_0}(\mathbf{x}) - \mathbf{D}_o(\mathbf{x}) + \mathbf{f}_D(\mathbf{E}(\mathbf{x},t), \mathbf{x}, t) + \int_0^t \mathbf{f}_J(\mathbf{E}(\mathbf{x},t'), \mathbf{x}, t') dt' \quad (9b)$$

Alle condizioni iniziali (3) bisogna, infine aggiungere quella su \mathbf{D}_T :

$$\mathbf{D}(\mathbf{x},0) = \mathbf{D}_o(\mathbf{x}), \quad (3a)$$

$$\mathbf{B}(\mathbf{x},0) = \mathbf{B}_o(\mathbf{x}) \quad (3b)$$

$$\mathbf{D}_T(\mathbf{x},0) = \mathbf{D}_{T_0}(\mathbf{x}), \quad (3c)$$

Come detto, il campo vettoriale $\mathbf{D}_{T_0}(\mathbf{x})$ può essere scelto a piacere e, in particolare, può essere il campo identicamente nullo.

Notiamo, per inciso, che nelle comuni applicazioni le grandezze dei due termini \mathbf{J} e $\partial\mathbf{D}/\partial t$ in un generico punto del dominio raramente si trovano ad avere valori comparabili.

Se ora definiamo due potenziali vettori nel seguente modo:

$$\mathbf{A} = - \int_0^t \mathbf{E} dt \quad (10a)$$

$$\mathbf{W} = \int_0^t \mathbf{H} dt \quad (10b)CC$$

il che equivale:

$$\mathbf{E} = -\partial\mathbf{A}/\partial t \quad (11a)$$

$$\mathbf{H} = \partial\mathbf{W}/\partial t \quad (11b)$$

e poniamo:

$$\mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A} \quad (11c)$$

$$\mathbf{D}_{T_0} = \mathbf{D}_{T_0} + \nabla \times \mathbf{W} \quad (11d)$$

si perviene alle formulazioni duali:

$$\begin{aligned} \nabla \times \mathbf{H} &= \partial \mathbf{D}_T / \partial t \\ \Downarrow \\ \mathbf{H} &= \partial \mathbf{W} / \partial t, \quad \mathbf{D}_T = \mathbf{D}_{T_0} + \nabla \times \mathbf{W} \end{aligned}$$

$$\begin{aligned} \nabla \times \mathbf{E} &= -\partial \mathbf{B} / \partial t \\ \Downarrow \\ \mathbf{E} &= -\partial \mathbf{A} / \partial t, \quad \mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A} \end{aligned}$$

$$\nabla \times \mathbf{H} - \partial \mathbf{D}_T / \partial t = 0, \quad \forall \mathbf{W}$$

$$\nabla \times \mathbf{E} + \partial \mathbf{B} / \partial t = 0, \quad \forall \mathbf{A}$$

nelle quali, per ogni possibile scelta dei potenziali \mathbf{A} e \mathbf{W} , le equazioni di Maxwell (7) risultano automaticamente soddisfatte.

Procediamo ora alla definizione di un funzionale di errore locale $\lambda(\mathbf{x},t) \geq 0$ per descrivere la misura con la quale le relazioni costitutive sono soddisfatte; esso dovrà essere nullo solo nei punti spazio-temporali dove le relazioni costitutive (9) sono rigorosamente rispettate.

Una definizione ammissibile potrebbe ad esempio essere:

$$\lambda(\mathbf{x},t) = \alpha_H [\mathbf{B}(\mathbf{x},t) - \mathbf{f}_B(\mathbf{H}(\mathbf{x},t),t)]^2 + \alpha_E [\mathbf{D}_T(\mathbf{x},t) - \mathbf{D}_{T_0}(\mathbf{x}) - \mathbf{f}_D(\mathbf{E}(\mathbf{x},t),\mathbf{x},t) + \mathbf{D}_0(\mathbf{x}) - \int_0^t \mathbf{f}_J(\mathbf{E}(\mathbf{x},t),t) dt]^2 \quad (12)$$

dove α_H e α_E sono opportuni coefficienti di peso sulle relazioni costitutive magnetiche ed elettriche, che sono, in generale, funzione della dinamica delle sorgenti e che possono assumere eventualmente nello stesso problema valori diversi per differenti materiali.

Una volta definito un valore conveniente per i coefficienti di peso, il problema si riduce alla ricerca dei due potenziali vettori $\mathbf{A}(\mathbf{x},t)$ e $\mathbf{W}(\mathbf{x},t)$ capaci di minimizzare il funzionale di errore globale Λ definito come

$$\Lambda = \int_0^T \int_V \lambda(\mathbf{x},t) dV dt \quad (13)$$

dove V è la regione spaziale di integrazione e $(0,T)$ l'intervallo temporale di interesse.

Tra i maggiori vantaggi che la formulazione basata sull'errore costitutivo offre spicca senz'altro quello di avere a disposizione diversi modi per calcolare le più rilevanti grandezze integrali di interesse elettromagnetico: da una parte l'energia elettrica e l'energia magnetica che alla fine di ogni time-step sono immagazzinate nel sistema, e dall'altra la potenza ohmica dissipata. Inoltre, attraverso una riformulazione del vettore di Poynting nelle variabili di campo introdotte, possono essere stimati, alla fine di ciascun time-step, le variazioni di energia del sistema dovute a scambi energetici col mondo esterno:oppure, come si vedrà in seguito, a dissipazione numerica.

Ogni variabile di campo può essere rappresentata in due maniere distinte ognuna delle quali basata su uno solo dei potenziali vettori (Vedi tab.[])

Campo	Espressione in \mathbf{A}	Espressione in \mathbf{W}
$\mathbf{B}(\mathbf{x},t)$	$\mathbf{B}_o + \nabla \times \mathbf{A}$	$\mathbf{f}_B(\partial \mathbf{W} / \partial t, \mathbf{x}, t)$
$\mathbf{D}_T(\mathbf{x},t)$	$\mathbf{f}_{D_T}(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{D}_{T_o} + \nabla \times \mathbf{W}$
$\mathbf{H}(\mathbf{x},t)$	$\mathbf{f}_H(\mathbf{B}_o + \nabla \times \mathbf{A}, \mathbf{x}, t)$	$\partial \mathbf{W} / \partial t$
$\mathbf{E}(\mathbf{x},t)$	$-\partial \mathbf{A} / \partial t$	$\mathbf{f}_{E-D_T}(\mathbf{D}_{T_o} + \nabla \times \mathbf{W}, \mathbf{x}, t)$
$\mathbf{D}(\mathbf{x},t)$	$\mathbf{f}_D(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{f}_D[\mathbf{f}_{E-D_T}(\mathbf{D}_{T_o} + \nabla \times \mathbf{W}, \mathbf{x}, t), \mathbf{x}, t]$
$\mathbf{J}(\mathbf{x},t)$	$\mathbf{f}_J(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{f}_J[\mathbf{f}_{E-D_T}(\mathbf{D}_{T_o} + \nabla \times \mathbf{W}, \mathbf{x}, t), \mathbf{x}, t]$

Sfruttando le varie espressioni in tabella è semplice dedurre le espressioni per le varie grandezze integrali nei casi lineari e non lineari.

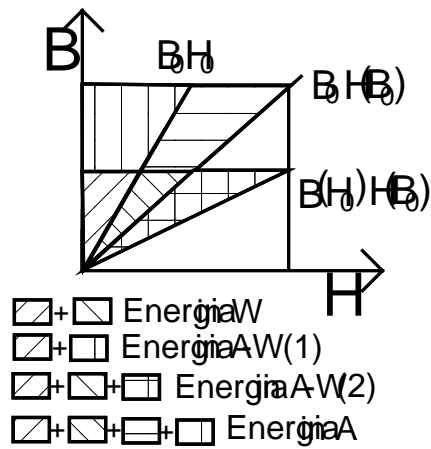


Fig.): Significato delle diverse espressioni per l'energia magnetica. (Caso lineare)

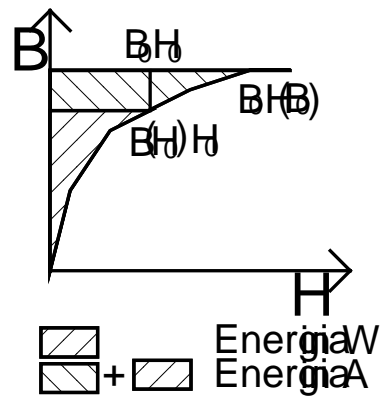


Fig.): Significato delle diverse espressioni per l'energia magnetica. (Caso non lineare)

ENERGIA ELETTRICA

Si assumo che per il mezzo valga la relazione costitutiva lineare

$$\mathbf{D}(\mathbf{x},t) = \mathbf{f}_D(\mathbf{E}(\mathbf{x},t), \mathbf{x},t)$$

l'energia elettrica immagazzinata all'istante t nel volume V può essere stimata come::

$$\begin{aligned}
 W_{El,A} &= \frac{1}{2} \int_V \mathbf{f}_D \left(\int \dot{\mathbf{A}} \right) \cdot \left(\int -\dot{\mathbf{A}} \right) dV \\
 W_{El,W} &= \frac{1}{2} \int_V \mathbf{f}_D \left[\int_{E-D_T} (\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] \cdot \mathbf{f}_{E-D_T} (\mathbf{D}_{T0} + \nabla \times \mathbf{W}) dV \\
 W'_{El,A-W} &= \frac{1}{2} \int_V \mathbf{f}_D \left[\int_{E-D_T} (\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] \cdot \left(-\dot{\mathbf{A}} \right) dV \\
 W''_{El,A-W} &= \frac{1}{2} \int_V \mathbf{f}_D \left(\int \dot{\mathbf{A}} \right) \cdot \left[\int_{E-D_T} (\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] dV
 \end{aligned}$$

Se in V la conducibilità elettrica è nulla, $\mathbf{D}(\mathbf{x},t) = \mathbf{D}_T(\mathbf{x},t)$ e vale la seguente espressione:

$$W'''_{El,A-W} = \frac{1}{2} \int_V (\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \cdot (-\dot{\mathbf{A}}) dV$$

Se per il mezzo vale la relazione costitutiva non lineare

$$\mathbf{D}(\mathbf{x},t) = \mathbf{f}_D(\mathbf{E}(\mathbf{x},t),\mathbf{x},t)$$

l'energia magnetica immagazzinata all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni::

$$W_{El,A} = \frac{1}{2} \int_V \left[\int_0^{\mathbf{f}_D[-\mathbf{A}(t_k+dtk)]} \int_0^{\mathbf{f}_E(\mathbf{D}')} \mathbf{f}_E(\mathbf{D}') \cdot d\mathbf{D}' \right] dV$$

$$W_{El,W} = \frac{1}{2} \int_V \left[\int_0^{\mathbf{f}_D[\mathbf{f}_{E-DT}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}(t_{k+1}))]} \int_0^{\mathbf{f}_E(\mathbf{D}')} \mathbf{f}_E(\mathbf{D}') \cdot d\mathbf{D}' \right] dV$$

ENERGIA MAGNETICA

Si assuma che per il mezzo valga la relazione costitutiva lineare

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t),\mathbf{x},t)$$

l'energia magnetica immagazzinata all'istante t nel volume V può essere stimata in uno dei modi seguenti :

$$W_{Mag,A} = \frac{1}{2} \int_V \mathbf{f}_H(\nabla \times \mathbf{A}) \int \nabla \times \mathbf{A} dV$$

$$W_{Mag,W} = \frac{1}{2} \int_V \mathbf{f}_B(\nabla \times \mathbf{W}) \cdot \nabla \times \mathbf{W} dV$$

$$W'_{Mag,A-W} = \frac{1}{2} \int_V \nabla \times \mathbf{A} \cdot \nabla \times \mathbf{W} dV$$

$$W'''_{MagA-W} = \frac{1}{2} \int_V \mathbf{f}_B(\nabla \times \mathbf{W}) \cdot \mathbf{f}_H(\mathbf{B}_0 + \nabla \times \mathbf{A}) dV$$

Se per il mezzo vale la relazione costitutiva non lineare

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t),\mathbf{x},t)$$

l'energia magnetica immagazzinata all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni::

$$W_{\text{Mag,A}} = \frac{1}{2} \int_V \left[\int_0^{\mathbf{B}_0 + \nabla \times \mathbf{A}(t_{k+1})} \mathbf{f}_H(\mathbf{B}') \cdot d\mathbf{B}' \right] dV$$

$$W_{\text{Mag,W}} = \frac{1}{2} \int_V \left[\int_0^{f_B \left[\dot{\mathbf{W}}(tk+dtk) \right]} \mathbf{f}_H(\mathbf{B}') \cdot d\mathbf{B}' \right] dV$$

POTENZA OHMICA-POTENZA GENERATA

Si assuma che per il mezzo valga la relazione costitutiva, eventualmente non lineare:

$$\mathbf{J}(\mathbf{x},t) = \mathbf{f}_J(\mathbf{E}(\mathbf{x},t), \mathbf{x},t)$$

La differenza tra la potenza dissipata per effetto ohmico e quella generata all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni:

$$P_{\text{Oh,A}}(t) - P_{\text{Gen,A}}(t) = \frac{1}{2} \int_V \mathbf{f}_J \left(\int \dot{\mathbf{A}} \right) \cdot \left(\int \dot{\mathbf{A}} \right) dV$$

$$P_{\text{Oh,W}}(t) - P_{\text{Gen,W}}(t) = \frac{1}{2} \int_V \mathbf{f}_J \left[\mathbf{f}_{\mathbf{E-D}_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] \cdot \mathbf{f}_{\mathbf{E-D}_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}) dV$$

$$(P_{\text{EOhA-W}}(t) - P_{\text{Gen,A-W}}(t))' = \frac{1}{2} \int_V \mathbf{f}_J \left[\mathbf{f}_{\mathbf{E-D}_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] \cdot \left(-\dot{\mathbf{A}} \right) dV$$

$$(P_{\text{Oh,A-W}}(t) - P_{\text{Gen,A-W}}(t))'' = \frac{1}{2} \int_V \mathbf{f}_J \left(\int \dot{\mathbf{A}} \right) \cdot \left[\mathbf{f}_{\mathbf{E-D}_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}) \right] dV$$

Se in V la corrente di spostamento è trascurabile rispetto alla corrente di conduzione, allora $\mathbf{J} = \partial \mathbf{D}_T / \partial t$ e vale la seguente espressione:

$$(P_{\text{Oh,A-W}}(t) - P_{\text{Gen,A-W}}(t))''' = \frac{1}{2} \int_V \left(-\dot{\mathbf{A}} \right) \cdot (\nabla \int \dot{\mathbf{W}}) dV$$

VEVTORE DI POYNTING

Si ricorda l'espressione del bilancio elettromagnetico di potenza per il volume V .

$$P_{\text{Gen}} = P_{\text{Ohm}} + \frac{d}{dt}(W_{\text{Mag}} + W_{\text{El}}) + \oint_{\partial V} \mathbf{S} \cdot \mathbf{n} d\Sigma \quad ()$$

dove: \mathbf{S} è il vettore di Poynting ed è stata scelta l'orientazione uscente dal volume per la normale \mathbf{n} alla frontiera ∂V . Integrando la () nell'intervallo $(t_k, t_k + \Delta t_k)$ l'energia uscente dal volume nel time-step è data dalla seguente espressione.

$$\int_{t_k}^{t_k + \Delta t_k} \oint_{\partial V} \mathbf{S} \cdot \mathbf{n} d\Sigma = \left[(W_{\text{El}}^{(k)} - W_{\text{El}}^{(k+1)}) + \int_{t_k}^{t_k + \Delta t_k} (P_{\text{Gen}} - P_{\text{Ohm}}) dt \right] + (W_{\text{Mag}}^{(k)} - W_{\text{Mag}}^{(k+1)})$$

Sostituendo le diverse espressioni di energia elettrica, energia magnetica e potenza, è possibile ottenere il vettore di Poynting in termini della sola A , di W oppure attraverso l'uso combinato dei due potenziali.

Si consideri ora un mezzo per il quale la relazione costitutiva $\mathbf{f}_D(\mathbf{E}, \mathbf{x}, t)$ è lineare::

$$\mathbf{D}(\mathbf{x}, t) = \mathbf{f}_D(\mathbf{E}, \mathbf{x}, t) = \varepsilon(\mathbf{x}, t) \mathbf{E}(\mathbf{x}, t) \quad ()$$

In questo caso si può trovare una formulazione del teorema di Poynting in cui la potenza ohmica e quella generata vengono sostituiti da un' espressione integrale in cui compare il vettore \mathbf{D}_T .

Si ha in ogni caso:

$$P_{\text{Gen}} - P_{\text{Ohm}} = - \int_V \mathbf{J} \cdot \mathbf{E} dV \quad ()$$

Per la linearità ipotizzata si può inoltre scrivere:

$$W_{El} = \frac{1}{2} \int_V \mathbf{D} \cdot \mathbf{E} dV \quad (1)$$

Se si assume che $\mathbf{E}(\mathbf{x}, t) = \mathbf{E}(\mathbf{x}, t_{k+1}) \quad \forall t \in (t_k, t_k + dt_k)$ (il che è comunque vero per $\mathbf{E}(\mathbf{A}) = -\partial \mathbf{A} / \partial t$, poichè per \mathbf{A} si è impiegata un'interpolazione nel tempo lineare a tratti), e si usano le relazioni (1) e (2) è possibile portare \mathbf{E} fuori dall'integrale nel tempo della (1), che diventa:

$$\int_V -\mathbf{E}(t_{K+1}^-) \cdot \left[\int_{t_k}^{t_k + \Delta t_k} \mathbf{J} dt + \frac{1}{2} \mathbf{D}(t_{K+1}) \right] + \frac{1}{2} \mathbf{E}(t_K^-) \mathbf{D}(t_K) + (W_{Mag}^{(k)} - W_{Mag}^{(k+1)}) - \int_{t_k}^{t_k + \Delta t_k} \oint_{\partial V} \mathbf{S} \cdot \mathbf{n} d\Sigma = 0$$

e successivamente

$$\int_V -\mathbf{E}(t_{K+1}^-) \cdot \left[\int_{t_k}^{t_k + \Delta t_k} \mathbf{J} dt + \mathbf{D}(t_{K+1}) \right] + \frac{1}{2} \mathbf{E}(t_{K+1}^-) \cdot \mathbf{D}(t_{K+1}) + \frac{1}{2} \mathbf{E}(t_K^-) \mathbf{D}(t_K) + (W_{Mag}^{(k)} - W_{Mag}^{(k+1)}) - \int_{t_k}^{t_k + \Delta t_k} \oint_{\partial V} \mathbf{S} \cdot \mathbf{n} d\Sigma = 0$$

$$\int_{t_k}^{t_k + \Delta t_k} \oint_{\partial V} \mathbf{S} \cdot \mathbf{n} d\Sigma = \int_V -\mathbf{E}(t_{K+1}^-) \cdot [\mathbf{D}_T(t_{K+1}) - \mathbf{Hist}(t_K)] + (W_{El}^{(k+1)} + W_{El}^{(k)}) + (W_{Mag}^{(k)} - W_{Mag}^{(k+1)})$$

dove con $\mathbf{Hist}(t_k)$ si è indicata l'espressione :

$$\mathbf{Hist}(\mathbf{x}, t_k) = \mathbf{D}_{T0}(\mathbf{x}) - \mathbf{D}_0(\mathbf{x}) + \int_0^{t_k} \mathbf{f}_J(\mathbf{A}(\mathbf{x}, t), \mathbf{x}, t) dt$$

che, nel caso in cui le caratteristiche di conducibilità del mezzo siano lineari e tempo invariane: $\mathbf{J}(\mathbf{x}, t) = \mathbf{f}_J(\mathbf{E}, \mathbf{x}, t) = \sigma(\mathbf{x})[\mathbf{E}(\mathbf{x}, t) + \mathbf{E}^*(\mathbf{x}, t)]$; diventa:

$$\mathbf{Hist}(\mathbf{x}, t_k) = \mathbf{D}_{T0}(\mathbf{x}) - \mathbf{D}_0(\mathbf{x}) + \sigma(\mathbf{x}) \mathbf{A}(\mathbf{x}, t_k) + \int_0^{t_k} \sigma(\mathbf{x}) \mathbf{E}^*(\mathbf{x}, t) dt$$

2.3) Applicazione dell'approccio basato sulla minimizzazione dell'errore costitutivo ad un circuito RLC. Introduzione delle "Cariche di maglie" e dei "Flussi di nodo" quali variabili di stato. Equazioni risolutive. Funzionale d'errore e sue derivate.

L'approccio basato sulla minimizzazione dell'errore costitutivo può essere adoperato anche per il trattamento dei circuiti a parametri concentrati. La cosa appare d'altronde ovvia se si osserva che la teoria delle reti consegue in modo naturale dalle eq. di Maxwell, allorchè siano verificate le ipotesi di funzionamento quasi-stazionario. In tal caso è possibile riscrivere in forma integrale le caratteristiche dei materiali ed associarle a zone di singolarità spaziale: i componenti elettrici. Se si ridefiniscono anche i campi in forma integrale -evitando attentamente di includere nel dominio di integrazione le zone di singolarità spaziali- è possibile descrivere una estesa classe di fenomeni elettromagnetici attraverso modelli *zero-dimensional*, ovvero attraverso modelli nei quali sono totalmente assenti le variabili spaziali. Le incognite integrali, per le quali il problema così riformulato deve essere invertito, sono le tensioni e le correnti.

In questo paragrafo, viene definito un funzionale di errore associato alle caratteristiche dei componenti di una generica rete lineare priva di generatori dipendenti e tale da risultare strettamente positivo se le caratteristiche non sono esattamente soddisfatte e nullo in caso contrario.

Si consideri una generica rete lineare composta eventualmente da generatori indipendenti di corrente e di tensione, resistori, condensatori ed induttori (eventualmente accoppiati), ma priva di generatori dipendenti.

La rete sia composta di l lati e n nodi. Si scelgano i riferimenti per le tensioni e le correnti associate ai lati della rete e si costruiscano la matrice di incidenza \mathbf{A} e la matrice degli anelli $\mathbf{M}[\]$. Siano \mathbf{v}^l ed \mathbf{i}^l funzioni vettoriali i cui elementi sono gli andamenti temporali rispettivamente delle tensioni e delle correnti di ciascun lato della rete.

Assunto un qualunque nodo come riferimento, si associ la funzione $e_j(t)$ a ciascuno dei nodi rimanenti (potenziali di nodo). Sia \mathbf{e} il vettore dei potenziali di nodo composto dagli $n-1$ elementi $\{e_j(t)\}$. Per ogni possibile valore di \mathbf{e} l'equazione di Kirchoff alle tensioni risulta automaticamente soddisfatta. Il

vettore delle tensioni di lato può essere espressa attraverso la matrice di incidenza e il vettore dei potenziali di nodo come:

$$\mathbf{v}^l = \mathbf{A}^T \mathbf{e} \quad (14)$$

In maniera duale, individuato un albero della rete in cui non compaiano lati su cui si trovano generatori indipendenti di corrente, si costruisca il vettore delle correnti di maglia \mathbf{i} associando la funzione $i_j(t)$ [$j= 1, l-(n-1)$] all' i -mo lato del coalbero. In questo caso per ogni possibile scelta di \mathbf{i} le equazioni di Kirchoff alle correnti sono soddisfatte. Il vettore delle correnti di lato può essere espressa attraverso la matrice degli anelli e il vettore delle correnti di maglia come:

$$\mathbf{i}^l = \mathbf{M}^T \mathbf{i} \quad (15)$$

Utilizzando le variabili \mathbf{e} ed \mathbf{i} come incognite, attraverso l'integrazione temporale dell'errore sull'equazione caratteristica è possibile definire per ogni lato del circuito un funzionale locale di errore. Sommando i contributi dei funzionali locali ed integrando nel tempo si giunge alla definizione del funzionale globale di errore. La soluzione della rete può essere trovata attraverso la minimizzazione del funzionale globale di errore.

Nella definizione delle espressioni dei funzionali locali di errore si supponrà che, per ciascun lato k della rete, l'equazione caratteristica sia compresa nel seguente elenco:

$$v_k^l(t) = R_k i_k^l(t) \quad (\text{resistenza}) \quad (16)$$

$$v_k^l(t) = \sum_{j=1}^l L_{kj} \frac{di_j^l(t)}{dt}; \quad i_k^l(0)=i_{k0} \quad (\text{induttore accoppiato}) \quad (17)$$

$$i_k^l(t) = C_k \frac{dv_k^l(t)}{dt}; \quad v_k^l(0)=v_{k0} \quad (\text{condensatore}) \quad (18)$$

$$i_k^l(t) = i_k^*(t) \quad (\text{generatore ind. di corrente}) \quad (19)$$

$$e_k(t) = e_k^*(t) \quad (\text{generatore ind. di tensione}) \quad (20)$$

Le relazioni 19 e 20 possono essere imposte direttamente come vincoli sulle variabili. Per il generico lato k soggetto rispettivamente alle condizioni (19) o (20) si ha pertanto: :

$$(\underline{\mathbf{M}}^T \mathbf{i})_k = i_k^*(t) \quad (21)$$

$$(\underline{\mathbf{A}}^T \mathbf{e})_k = e_k^*(t) \quad (22)$$

Si noti che, così facendo, le condizioni (21) e (22) sono analoghe alle condizioni al contorno che devono essere imposte nel problema di campo e, dunque, i lati su cui si trovano i generatori indipendenti potrebbero essere visti come i "*lati di frontiera*" della rete!!!

Per il generico lato k, una possibile definizione dei funzionali locali di errore associati alle relazioni (16), (17) e (18) che soddisfa tutte le condizioni richieste è la seguente:

$$\begin{aligned} \lambda_{R,k}(t) &= \frac{1}{2} [v_k^l(t) - R_k i_k^l(t)]^T \cdot [(R_k)^{-1} v_k^l(t) - i_k^l(t)] = \\ &= \frac{1}{2} [(\underline{\mathbf{A}}^T \mathbf{e})_k - R_k (\underline{\mathbf{M}}^T \mathbf{i})_k]^T \cdot [(R_k)^{-1} (\underline{\mathbf{A}}^T \mathbf{e})_k - (\underline{\mathbf{M}}^T \mathbf{i})_k] ; \end{aligned} \quad (23)$$

$$\begin{aligned} \lambda_{L,k}(t) &= \frac{1}{2} \left[v_k^l(t) - \sum_{j=1}^1 L_{kj} \frac{di_j^l(t)}{dt} \right]^T \cdot \left[\sum_{j=1}^1 \Gamma_{kj} \int_0^t v_j^l(t') dt' - i_k^l(t) - i_{k0}^l \right] = \\ &= \frac{1}{2} \left[(\underline{\mathbf{A}}^T \mathbf{e})_k - \sum_{j=1}^1 L_{kj} \frac{d(\underline{\mathbf{M}}^T \mathbf{i})_j}{dt} \right]^T \cdot \left[\sum_{j=1}^1 \Gamma_{kj} \int_0^t (\underline{\mathbf{A}}^T \mathbf{e})_j dt' - (\underline{\mathbf{M}}^T \mathbf{i})_k - i_{k0}^l \right] ; \end{aligned} \quad (24)$$

$$\begin{aligned} \lambda_{C,k}(t) &= \frac{1}{2} \left[v_k^l(t) - v_{k0}^l - (C_k)^{-1} \int_0^t i_k^l(t') dt' \right]^T \cdot \left[C_k \frac{dv_k^l(t)}{dt} - i_k^l(t) \right] = \\ &= \frac{1}{2} \left[(\underline{\mathbf{A}}^T \mathbf{e})_k - v_{k0}^l - (C_k)^{-1} \int_0^t (\underline{\mathbf{M}}^T \mathbf{i})_k dt' \right]^T \cdot \left[C_k \frac{d(\underline{\mathbf{A}}^T \mathbf{e})_k}{dt} - (\underline{\mathbf{M}}^T \mathbf{i})_k \right] ; \end{aligned} \quad (25)$$

dove si è indicato con L_{kj} e Γ_{kj} l'elemento k,j rispettivamente delle matrici di induttanza e di induttanza reciproca associate alla rete e si è usata la simbologia classica della teoria dei circuiti per tutti gli altri termini..

Per evitare problemi numerici dovuti al fatto che le grandezze incognite possono differire per diversi ordini di grandezza può risultare vantaggioso moltiplicare, nella costruzione del funzionale globale di errore, ciascuno dei funzionali locali per un peso opportuno α_k .

Se gli I lati della rete contengono N_E generatori indipendenti di tensione, N_J generatori indipendenti di corrente, N_R resistori, N_L induttori e N_C condensatori, il funzionale globale di errore si scrive come:

$$\Lambda(\mathbf{e}, \mathbf{i}, \boldsymbol{\alpha}) = \int_0^T \sum_{k=1}^{N_R} \alpha_k \lambda_{R,k}(\mathbf{e}, \mathbf{i}, t) + \sum_{k=1}^{N_L} \alpha_k \lambda_{L,k}(\mathbf{e}, \mathbf{i}, t) + \sum_{k=1}^{N_C} \alpha_k \lambda_{C,k}(\mathbf{e}, \mathbf{i}, t) dt \quad (26)$$

[C.A.Desoer-E.S.Kuh, *Fondamenti di teoria dei circuiti*, Franco Angeli, (1989)

(Ref. Rikabi)

2.4) Analogia tra il modello associato alle equazioni di Maxwell e quello associato alla rete RLC nella formulazione dell'errore costitutivo. Comparazione tra le incognite, le equazioni, le condizioni iniziali, i funzionali d'errore.

Uno dei maggiori problemi che il metodo basato sulla minimizzazione dell'errore costitutivo presenta dal punto di vista operativo è la scelta di valori adeguati per i pesi. Questi, come detto, variano in funzione delle caratteristiche del problema (caratteristiche del materiale, dinamica delle sorgenti). Si vedrà, inoltre, nel capitolo seguente che altri parametri influenzano pesantemente la qualità delle soluzioni numeriche e sono il metodo di integrazione e l'ampiezza

del time-step per l'avanzamento nel tempo del sistema di equazioni differenziali. La scelta ottimale di questi valori può essere condotta attraverso un'analisi parametrica. D'altra parte per problemi di campo tridimensionali un'analisi parametrica è un compito tutt'altro che semplice da eseguire a causa del notevole onere computazionale associato alla soluzione del singolo problema. Si può allora pensare di costruire un rozzo modello circuitale del problema di campo e ricavare da questo, con peso computazionale modesto, tutte le informazioni che occorrono per scegliere in modo soddisfacente i parametri per il problema di campo.

Allo scopo di meglio evidenziare le analogie che esistono tra il problema tridimensionale e il corrispondente modello zero-dimensionale, può risultare conveniente assumere come incognite per il circuito a parametri concentrati anzichè i potenziali di nodo e le correnti di maglia, gli integrali temporali di queste grandezze:

$$q_m(t) = \int_0^t i_m(t') dt' \quad m=1, l - (n-1) \quad (27)$$

$$\Phi_j(t) = \int_0^t e_j(t') dt' \quad j=1, n-1 \quad (28)$$

In tabella [] vengono comparati il modello tri-dimensionale e quello zero-dimensionale:

:

Problemi di campo

Circuiti a parametri concentrati

Equazioni di Maxwell

$$\nabla \times \mathbf{H} = \partial \mathbf{D}_T / \partial t$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$$

$$\nabla \cdot \mathbf{D}_T = 0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\forall \mathbf{x} \in V \text{ e } \forall t \in [0, T]$$

Incognite duali

potenziale vettore elettrico \mathbf{W}

potenziale vettore magnetico \mathbf{A}

Campi

$$\mathbf{E} = -\partial \mathbf{A} / \partial t; \quad \mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A}$$

$$\mathbf{H} = \partial \mathbf{W} / \partial t; \quad \mathbf{D}_T = \mathbf{D}_{T0} + \nabla \times \mathbf{W}$$

Equazioni di Kirchoff

(Circuito con l lati ed n nodi)

$$\sum_{k=1}^{N_j} i_k^j = 0; \quad j=1, \dots, n-1$$

$$\sum_{k=1}^{N_m} v_k^m = 0; \quad m=1, \dots, l - (n-1)$$

Incognite duali

Cariche di maglia q_m ; $m=1, l - (n-1)$

Flussi di nodo Φ_j ; $j=1, n-1$

Correnti di maglia, Potenziali di nodo

$$e_j = d(\Phi_j) / dt; \quad j=1, n-1$$

$$i_m = d(q_m) / dt; \quad m=1, l - (n-1)$$

Condizioni iniziali

$$\mathbf{D}(\mathbf{x},0)=\mathbf{D}_0(\mathbf{x}); \quad \mathbf{B}(\mathbf{x},0)=\mathbf{B}_0(\mathbf{x});$$

$$\mathbf{D}_T(\mathbf{x},0)=0$$

$$\forall \mathbf{x} \in V$$

Condizioni iniziali

$$q_{Cm}(0)=q_{C0m}; \quad m=1, N_C$$

$$\varphi_{Lj}(0) = \varphi_{L0j}; \quad j=1, N_L$$

$$q_k(0)=0 \quad ; \quad k=1, l - (n-1)$$

Condizioni al contorno

$$\mathbf{A}(\mathbf{x},t)=\mathbf{A}^*(\mathbf{x},t);$$

$$\mathbf{W}(\mathbf{x},t)=\mathbf{W}^*(\mathbf{x},t);$$

$$\forall \mathbf{x} \in \partial V \text{ e } \forall t \in [0, T]$$

Condizioni al contorno

$$(\underline{\mathbf{A}}^T \mathbf{e})_k = \int_0^t e^*_k(t') dt'; \quad k=1, N_E$$

$$(\underline{\mathbf{M}}^T \mathbf{q})_k = \int_0^t i^*_k(t') dt'; \quad k=1, N_J$$

A scopo di esempio si consideri il problema di campo di fig [.a] modellabile attraverso la rete RLC di fig [.b]:

Le condizioni iniziali del problema di campo

$$\mathbf{D}(\mathbf{x},0)=\mathbf{D}_0(\mathbf{x}); \quad \text{e} \quad \mathbf{B}(\mathbf{x},0)=\mathbf{B}_0(\mathbf{x});$$

corrispondono nella rete elettrica alle condizioni iniziali sulla carica del condensatore e sul flusso dell'induttore:

$$q_C(0)=q_{C0};$$

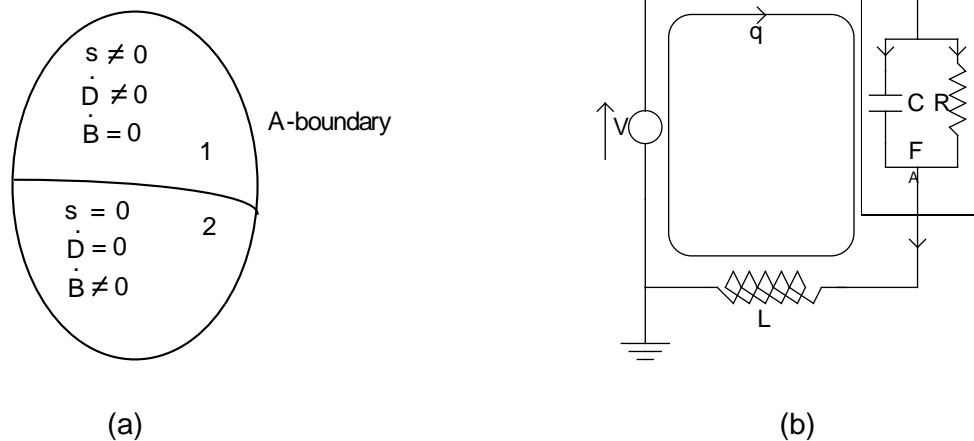
$$\varphi_L(0) = \varphi_{L0};$$

Le condizioni al contorno sul potenziale vettore **A** sono rappresentate attraverso il generatore di tensione V .

Le caratteristiche dei bipoli passivi sono:

$$i_R = i_R(v_C); \quad (29.a) \qquad q_C = q_C(v_C); \quad (29.b) \qquad i_L = i_L(\Phi_L); \quad (29.c)$$

Se il coppia RC viene trattato come un singolo bipolo, come variabili nei termini delle quali risolvere la rete si possono assumere la carica di maglia, q , e il flusso di nodo $\Phi = \Phi_L$. Si può porre arbitrariamente $q(0) = 0$.



Fig[1] (a) Problema di campo; (b) Modello circuitale del problema (a)

Adottando per tutti i componenti passivi la convenzione dell'utilizzatore, le leggi di Kirchoff alle tensioni e alle correnti sono:

$$V = v_C + \dot{\Phi}; \qquad (30.a)$$

$$i_L = i_R + i_C; \quad \rightarrow \quad \int_0^t i_L dt' = \int_0^t (i_R + i_C) dt'; \quad \rightarrow \quad q(t) = q_C(t) - q_{C0} + \int_0^t i_R dt'; \qquad (30.b)$$

Dalle relazioni (29.a), (29.b) e (30.b) si ricava l'equazione caratteristica per il cappio RC:

$$q = q_C(V - \dot{\Phi}) - q_{C0} + \int_0^t i_R(V - \dot{\Phi}) dt'; \quad (29.d)$$

Nel caso lineare le relazioni costitutive (29 a,b,c) diventano:

$$i_R = \frac{v_C}{R} = \frac{V - \dot{\Phi}}{R}; \quad (31.a) \quad q_C = C v_C = C(V - \dot{\Phi}); \quad (31.b) \quad \dot{q} = \frac{\Phi}{L}; \quad (31.c)$$

e la (29.d) diventa: $q = C(V - \dot{\Phi}) - q_{C0} + \int_0^t \frac{V - \dot{\Phi}}{R} dt'; \quad (31.d)$

Se per q e Φ si adotta una approssimazione lineare nel generico intervallo $[t_k, t_k + \Delta t_k]$ si ha:

$$q(t) = \frac{q_{k+1} - q_k}{\Delta t_k} (t - t_k) + q_k; \quad (32.a) \quad \dot{q}(t) = \frac{q_{k+1} - q_k}{\Delta t_k} \quad (32.b)$$

$$\Phi(t) = \frac{\Phi_{k+1} - \Phi_k}{\Delta t_k} (t - t_k) + \Phi_k; \quad (33.a) \quad \dot{\Phi}(t) = \frac{\Phi_{k+1} - \Phi_k}{\Delta t_k} \quad (33.b)$$

dove: $q(t_k) = q_k$; e $\Phi(t_k) = \Phi_k$; sono assunte note dal passo precedente mentre $q(t_k + \Delta t_k) = q_{k+1}$; e $\Phi(t_k + \Delta t_k) = \Phi_{k+1}$; sono le incognite da determinare. Si ponga:

$$k = q_0 - q_{C0} + \int_0^{t_k} \frac{V - \dot{\Phi}}{R} dt = q_0 - q_{C0} + \sum_{m=0}^{K-1} \left[V(t_m^+) - \Phi(t_m^+) \right] \Delta t_m$$

Le caratteristiche (31.c) e (31.d) si scrivono come:

$$\dot{q} = \frac{\dot{\Phi}}{L}; \quad \left(V - \dot{\Phi} \right) = (q - k) \frac{R}{(t - t_k) + RC} = (q - k)A(t);$$

dove si è posto
$$A(t) = \frac{R}{(t - t_k) + RC};$$

I funzionali locali di errore, tenendo conto delle (32) e (33) e del fatto che col metodo del time-stepping, che sarà illustrato nel prossimo capitolo, sono:

$$\lambda_L(\Phi_{k+1}, q_{k+1}, t) = \frac{1}{2} \left[\dot{q} - \frac{\dot{\Phi}}{L} \right] \cdot \left[L \dot{q} - \dot{\Phi} \right]; \quad (34.a)$$

$$\lambda_{RC}(\Phi_{k+1}, q_{k+1}, t) = \frac{1}{2} \left[\left(V - \dot{\Phi} \right) - (q - k)A(t) \right] \cdot \left[\frac{V - \dot{\Phi}}{A(t)} - (q - k) \right]; \quad (34.b)$$

$$\Lambda_k(\Phi_{k+1}, q_{k+1}, \alpha) = \int_{t_k}^{t_k + \Delta t_k} \alpha_L \lambda_L(\Phi, q, t) + \alpha_{RC} \lambda_{RC}(\Phi, q, t) dt; \quad t \in [t_k, t_k + \Delta t_k] \quad (35)$$

2.5) Formulazione nei termini dell'errore costitutivo per le equazioni di Maxwell in regime sinusoidale.

Con un procedimento formalmente simile a quello adoperato nel dominio del tempo è possibile ricavare un'espressione del funzionale di errore adeguata a trattare problemi nei quali i campi si trovano in condizioni di regime sinusoidale. Le equazioni di Maxwell (1) nel dominio della frequenza assumono il seguente aspetto:

$$\nabla \cdot \hat{\mathbf{B}}(\mathbf{x}, \omega) = 0 \quad (1b)$$

$$\nabla \times \hat{\mathbf{H}}(\mathbf{x}, \omega) = \hat{\mathbf{J}}(\mathbf{x}, \omega) + j \omega \hat{\mathbf{D}}(\mathbf{x}, \omega) \quad (1c)$$

$$\nabla \times \hat{\mathbf{E}}(\mathbf{x}, \omega) = -j\omega \hat{\mathbf{B}}(\mathbf{x}, \omega) \quad (1d)$$

$$\nabla \cdot \hat{\mathbf{J}}(\mathbf{x}, \omega) = -j\omega \hat{\rho}(\mathbf{x}, \omega) \quad (1e)$$

Mentre le relazioni di interfaccia valide sulle eventuali superfici di separazione Σ tra materiali diversi diventano

$$[\hat{\mathbf{D}}(\mathbf{x}, \omega)] \cdot \hat{\mathbf{n}}_{\Sigma} = \hat{\rho}_{\Sigma}(\mathbf{x}, \omega) \quad (2a)$$

$$[\hat{\mathbf{B}}(\mathbf{x}, \omega)] \cdot \hat{\mathbf{n}}_{\Sigma} = 0 \quad (2b)$$

$$[\hat{\mathbf{E}}(\mathbf{x}, \omega)] \times \hat{\mathbf{n}}_{\Sigma} = 0 \quad (2c)$$

$$[\hat{\mathbf{H}}(\mathbf{x}, \omega)] \times \hat{\mathbf{n}}_{\Sigma} = \hat{\mathbf{J}}_{\Sigma}(\mathbf{x}, \omega) \quad (2d)$$

$$[\hat{\mathbf{J}}(\mathbf{x}, \omega)] \cdot \hat{\mathbf{n}}_{\Sigma} = -j\omega \hat{\rho}_{\Sigma}(\mathbf{x}, \omega) \quad (2e)$$

che devono essere soddisfatte $\forall \mathbf{x} \in \Sigma$.

Per definire in modo completo il problema resta da aggiungere $\forall \mathbf{x} \in V$ alle equazioni di Maxwell le relazioni costitutive:

$$\hat{\mathbf{B}}(\mathbf{x}, \omega) = \mathbf{f}_{\mathbf{B}}(\hat{\mathbf{H}}, \mathbf{x}, \omega), \quad (4a)$$

$$\hat{\mathbf{D}}(\mathbf{x}, \omega) = \mathbf{f}_{\mathbf{D}}(\hat{\mathbf{E}}, \mathbf{x}, \omega), \quad (4b)$$

$$\hat{\mathbf{J}}(\mathbf{x}, \omega) = \mathbf{f}_{\mathbf{J}}(\hat{\mathbf{E}}, \mathbf{x}, \omega) \quad (4c)$$

e, indicata con ∂V la frontiera esterna di V , $\forall \mathbf{x} \in \partial V$ opportune condizioni al contorno.

Naturalmente la relazione (1.e) si può ricavare dalle relazioni (1c) e (1a).

Si assumano le seguenti ipotesi:

- l) Le relazioni costitutive (4) possono eventualmente essere anisotrope e non-omogenee ma devono essere lineari (per evitare dispersione in frequenza), tempo-invarianti (per evitare in frequenza la presenza di

integrali di convoluzione), e non dispersive nello spazio, nel tempo e in frequenza. In tal caso le (4a,b,c) diventano:

$$\hat{\mathbf{B}}(\mathbf{x},\omega) = \mathbf{f}_B(\hat{\mathbf{H}}(\mathbf{x},\omega),\mathbf{x}) = \boldsymbol{\mu}(\mathbf{x})\hat{\mathbf{H}}(\mathbf{x},\omega) \quad (4a)$$

$$\hat{\mathbf{D}}(\mathbf{x},\omega) = \mathbf{f}_D(\hat{\mathbf{E}}(\mathbf{x},\omega),\mathbf{x}) = \boldsymbol{\varepsilon}(\mathbf{x})\hat{\mathbf{E}}(\mathbf{x},\omega) \quad (4b)$$

$$\hat{\mathbf{J}}(\mathbf{x},\omega) = \mathbf{f}_J(\hat{\mathbf{E}}(\mathbf{x},\omega),\mathbf{x}) = \boldsymbol{\sigma}(\mathbf{x})(\hat{\mathbf{E}}(\mathbf{x},\omega) + \hat{\mathbf{E}}_{(s)}(\mathbf{x},\omega)) \quad (4c)$$

dove $\boldsymbol{\mu}$ ed $\boldsymbol{\varepsilon}$ possono eventualmente essere quantità complesse con parte immaginaria non positiva:

$$\boldsymbol{\mu}(\mathbf{x}) = \boldsymbol{\mu}'(\mathbf{x}) - j\boldsymbol{\mu}''(\mathbf{x}); \quad \boldsymbol{\varepsilon}(\mathbf{x}) = \boldsymbol{\varepsilon}'(\mathbf{x}) - j\boldsymbol{\varepsilon}''(\mathbf{x})$$

II) Le relazioni costitutive (4) siano monotone, il che si può porre in forma matematica imponendo che per ogni trasformazione elementare siano valide le seguenti disequaglianze:

$$\Delta\hat{\mathbf{B}} \cdot \Delta\hat{\mathbf{H}} \geq 0 \quad (4d)$$

$$\Delta\hat{\mathbf{D}} \cdot \Delta\hat{\mathbf{E}} \geq 0 \quad (4e)$$

$$\Delta\hat{\mathbf{J}} \cdot \Delta\hat{\mathbf{E}} \geq 0 \quad (4f)$$

Si ponga:

$$\hat{\mathbf{D}}_T(\mathbf{x},\omega) = (j\omega)^{-1} \hat{\mathbf{J}}(\mathbf{x},\omega) + \hat{\mathbf{D}}(\mathbf{x},\omega) \quad (5)$$

Le equazioni di Maxwell si possono allora riscrivere nella forma seguente:

$$\nabla \cdot \hat{\mathbf{D}}_T = 0 \quad (7a)$$

$$\nabla \cdot \hat{\mathbf{B}} = 0 \quad (7b)$$

$$\nabla \times \hat{\mathbf{H}} = j\omega \hat{\mathbf{D}}_T \quad (7c)$$

$$\nabla \times \hat{\mathbf{E}} = -j\omega \hat{\mathbf{B}} \quad (7d)$$

e le condizioni di continuità sulle superfici di interfaccia comportano:

$$[\hat{\mathbf{D}}_T(\mathbf{x},\omega)] \cdot \mathbf{n}_\Sigma = 0 \quad (8a)$$

$$[\hat{\mathbf{B}}(\mathbf{x},\omega)] \cdot \mathbf{n}_\Sigma = 0 \quad (8b)$$

$$[\hat{\mathbf{E}}(\mathbf{x},\omega)] \times \mathbf{n}_\Sigma = 0 \quad (8c)$$

$$[\hat{\mathbf{H}}(\mathbf{x},\omega)] \times \mathbf{n}_\Sigma = 0 \quad (8d)$$

Le relazioni costitutive (4a,b,c) si possono condensare nelle seguenti espressioni:

$$\hat{\mathbf{B}}(\mathbf{x},t) = \mathbf{f}_B(\hat{\mathbf{H}}(\mathbf{x},\omega), \mathbf{x}) = \mu(\mathbf{x}) \hat{\mathbf{H}}(\mathbf{x},\omega) \quad (9a)$$

$$\begin{aligned} \hat{\mathbf{D}}_T(\mathbf{x},t) &= \mathbf{f}_D(\hat{\mathbf{E}}(\mathbf{x},\omega), \mathbf{x}) + (j\omega)^{-1} \mathbf{f}_J(\hat{\mathbf{E}}(\mathbf{x},\omega), \mathbf{x}) = \\ &= [\boldsymbol{\varepsilon}(\mathbf{x}) + (j\omega)^{-1} \boldsymbol{\sigma}(\mathbf{x})] \hat{\mathbf{E}}(\mathbf{x},\omega) + (j\omega)^{-1} \boldsymbol{\sigma}(\mathbf{x}) \hat{\mathbf{E}}_{(s)}(\mathbf{x},\omega) \end{aligned} \quad (9b)$$

Si definiscano i due potenziali vettori nel seguente modo:

$$\hat{\mathbf{A}} = -(j\omega)^{-1} \hat{\mathbf{E}} \quad (10a)$$

$$\hat{\mathbf{W}} = (j\omega)^{-1} \hat{\mathbf{H}} \quad (10b)_{CC}$$

e si ponga

$$\hat{\mathbf{B}} = \nabla \times \hat{\mathbf{A}} \quad (11c)$$

$$\hat{\mathbf{D}}_T = \nabla \times \hat{\mathbf{W}} \quad (11d)$$

si perviene alle formulazioni duali:

$$\begin{aligned} \nabla \times \hat{\mathbf{H}} &= j\omega \hat{\mathbf{D}}_T \\ \Downarrow \end{aligned}$$

$$\begin{aligned} \nabla \times \hat{\mathbf{E}} &= -j\omega \hat{\mathbf{B}} \\ \Downarrow \end{aligned}$$

$$\hat{\mathbf{H}} = j\omega \hat{\mathbf{W}}, \quad \hat{\mathbf{D}}_T = \nabla \times \hat{\mathbf{W}}$$

$$\hat{\mathbf{E}} = j\omega \hat{\mathbf{A}}, \quad \hat{\mathbf{B}} = \nabla \times \hat{\mathbf{A}}$$

$$\nabla \times \hat{\mathbf{H}} - j\omega \hat{\mathbf{D}}_T = \mathbf{0}, \quad \forall \hat{\mathbf{W}}$$

$$\nabla \times \hat{\mathbf{E}} + j\omega \hat{\mathbf{B}} = \mathbf{0}, \quad \forall \hat{\mathbf{A}}$$

nelle quali, per ogni possibile scelta dei potenziali $\hat{\mathbf{A}}$ e $\hat{\mathbf{W}}$, le equazioni di Maxwell (7) risultano automaticamente soddisfatte.

Si noti che le equazioni cui si è pervenuto sono molto simili a quelle già proposte in letteratura da Pichon e Bossavit [1] con la sola differenza che qui si è usato $\hat{\mathbf{D}}_T$ al posto di $\hat{\mathbf{D}}$

Occorre adesso definire un funzionale di errore locale $\lambda(\mathbf{x}, \omega) \geq 0$ per descrivere la misura con la quale le relazioni costitutive sono soddisfatte; esso dovrà essere nullo solo nei punti spazio-frequenza dove le relazioni costitutive (9) sono rigorosamente rispettate,

A questo scopo si tenga in conto il fatto che il quadrato di un vettore complesso $\hat{\mathbf{x}} = \mathbf{x}_r + j\mathbf{x}_i$, dove \mathbf{x}_r e \mathbf{x}_i sono vettori reali, è la quantità complessa: $(\hat{\mathbf{x}})^2 = (\mathbf{x}_r \cdot \mathbf{x}_r)^2 - (\mathbf{x}_i \cdot \mathbf{x}_i)^2 + j(\mathbf{x}_r \cdot \mathbf{x}_i)$.

Si propone di usare il funzionale scalare e di tipo reale funzione di variabili vettoriali complesse, i potenziali vettori :

$$\lambda(\mathbf{x}, \omega) = \alpha_H (\mu)^{-1} [\hat{\mathbf{B}} - \mu \hat{\mathbf{H}}] \cdot [\hat{\mathbf{B}} - \mu \hat{\mathbf{H}}]^* + \alpha_E [\epsilon^2 + (\sigma/\omega)^2]^{-1/2} \\ \cdot [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}] - [\epsilon + (j\omega)^{-1} \sigma] \hat{\mathbf{E}}] \cdot [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}] - [\epsilon + (j\omega)^{-1} \sigma] \hat{\mathbf{E}}]^* \quad (12)$$

Appare evidente che se le relazioni costitutive sono soddisfatte il funzionale di errore si annulla.

Supposto per semplicità ϵ e μ reali, ipotesi assolutamente non necessaria ai fini del discorso, si può dimostrare che il funzionale sopra definito è positivo e di tipo reale scrivendo in modo esplicito ogni campo nella forma di parte reale e immaginaria. Con qualche manipolazione il funzionale (12) assume la forma:

$$[\hat{\mathbf{B}} - \mu \hat{\mathbf{H}}] \cdot [\hat{\mathbf{B}}^* - \mu \hat{\mathbf{H}}^*] = (\mathbf{B}^{(r)} - \mu \mathbf{H}^{(r)})^2 + (\mathbf{B}^{(i)} - \mu \mathbf{H}^{(i)})^2$$

$$\begin{aligned} & [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}] - [\epsilon + (j\omega)^{-1} \sigma] \hat{\mathbf{E}}] \cdot [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}]^* - [\epsilon - (j\omega)^{-1} \sigma] \hat{\mathbf{E}}^*] = \\ & = (\mathbf{D}_T^{(r)} - (\sigma/\omega) \mathbf{E}_{(s)}^{(i)} - \epsilon \mathbf{E}^{(r)} - (\sigma/\omega) \mathbf{E}^{(i)})^2 + (\mathbf{D}_T^{(i)} + (\sigma/\omega) \mathbf{E}_{(s)}^{(r)} - \epsilon \mathbf{E}^{(i)} + (\sigma/\omega) \mathbf{E}^{(r)})^2 \end{aligned}$$

$$\begin{aligned} \lambda(\mathbf{x}, \omega) = & \alpha_H (\mu)^{-1} (\mathbf{B}^{(r)} - \mu \mathbf{H}^{(r)})^2 + (\mathbf{B}^{(i)} - \mu \mathbf{H}^{(i)})^2 + \alpha_E [\epsilon^2 + (\sigma/\omega)^2]^{-1/2} \\ & (\mathbf{D}_T^{(r)} - (\sigma/\omega) \mathbf{E}_{(s)}^{(i)} - \epsilon \mathbf{E}^{(r)} - (\sigma/\omega) \mathbf{E}^{(i)})^2 + (\mathbf{D}_T^{(i)} + (\sigma/\omega) \mathbf{E}_{(s)}^{(r)} - \epsilon \mathbf{E}^{(i)} + (\sigma/\omega) \mathbf{E}^{(r)})^2 \end{aligned}$$

nella quale appare chiaramente come esso sia reale e definito positivo.

Si arriva così alla definizione del funzionale globale di errore:

$$\Lambda(\omega) = \int_V \lambda(\mathbf{x}, \omega) dV \quad (13)$$

Il problema di minimizzazione ha soluzione unica a meno che $\omega > 0$ non corrisponda ad una frequenza di risonanza (il che, d'altra parte non può accadere se nel mezzo ci sono perdite (ovvero σ e/o μ'' e/o ϵ'' sono diversi da zero)).

Come sottolineato nel lavoro sopra citato, questa formulazione, se accoppiata con l'uso degli edge elements, permette di evitare la comparsa di modi spuri. Con questo nome si indicano le soluzioni numeriche del problema elettromagnetico che non verificano la condizione di solenoidalità per il vettore di induzione magnetico $\hat{\mathbf{B}}$ e, ove questa sia richiesta, per il vettore spostamento elettrico $\hat{\mathbf{D}}$, e che pertanto sono prive di significato fisico.

[L.Pichon,A.Bossavit, *A New Variational Formulation Free of Spurious Modes, for the Problem of Loaded Cavities*, IEEE Trans.on Magn.Vol.29, N.2, (marzo 1993)

INTRODUZIONE

Le leggi dell'Elettromagnetismo hanno raggiunto l'attuale forma compatta ed elegante, intitolata al nome del loro artefice finale, Maxwell, solo a valle di una lunga fase di rielaborazione ed integrazione; esse si pongono, come obiettivo finale, l'inquadramento in un contesto unitario dell'intero universo dei fenomeni fisici di natura elettromagnetica.

In maniera affatto opposta al processo storico, appare oggi naturale postulare le equazioni di Maxwell come le leggi dell'Elettromagnetismo, e far discendere i modelli di una serie di problemi di notevole interesse applicativo come casi particolari ed approssimati. Questo viene fatto o aggiungendo opportune ipotesi semplificative (quali ad esempio la stazionarietà dei campi nei problemi statici) o rimuovendo termini trascurabili all'interno delle equazioni. Aggiungendo alle equazioni di Maxwell, così particolarizzate, opportune condizioni iniziali e al contorno e le relazioni costitutive, ovvero i modelli matematici delle proprietà fisiche dei materiali che occupano la regione in oggetto, sotto opportune ipotesi limitative sul carattere di queste ultime, si può dimostrare che il problema Elettromagnetico resta completamente definito in modo da ammettere una soluzione unica.

Il generico modello matematico del problema fisico è costituito da un sistema di equazioni integro-differenziali le cui incognite sono campi elettromagnetici o grandezze a questi correlate. Poiché solo in rarissimi casi (geometrie particolari, caratteristiche del mezzo lineari ed omogenee) il sistema matematico ammette una soluzione in forma chiusa, è necessario

ricorrere all'ausilio di metodi numerici, che sono capaci di fornire soluzioni soddisfacenti da un punto di vista ingegneristico per classi di problemi limitate da vincoli molto meno rigidi rispetto a quelli imposti dai metodi analitici. Pertanto è pratica comune per gli ingegneri risolvere in forma numerica la maggior parte dei complessi problemi che si manifestano in fase di analisi e di progetto delle moderne sofisticate apparecchiature elettromagnetiche adoperate nei più disparati settori: dalla medicina ai processi industriali, dalla fusione nucleare agli acceleratori di particelle. Questo grande successo delle tecniche solutive di tipo numerico è largamente dovuto agli enormi progressi che le tecnologie informatiche hanno registrato negli ultimi anni in termini sia di potenza di calcolo, sia di capacità di memoria.

A fronte dei suoi enormi vantaggi, l'analisi numerica, cioè la tecnica di risolvere problemi matematici attraverso approssimazione numerica, comporta però una serie di problemi che, allo stato attuale, sono ancora oggetto di ricerca e di sviluppo.

Le soluzioni trovate sono affette da errori di discretizzazione e di arrotondamento, tanto più elevati quanto più è rozza la schematizzazione adoperata per descrivere il problema fisico. L'accuratezza della discretizzazione è, d'altronde, limitata dalla disponibilità di memoria e dalla velocità del computer (anche se con il passare del tempo l'avanzare della tecnologia rende sempre meno stretti questi vincoli). Se si considera un metodo basato sugli elementi finiti, per poter ottenere una soluzione numerica soddisfacente appare, dunque, chiara la necessità di operare una modellizzazione giudiziosa del problema in esame che spinga maggiormente il raffinamento della discretizzazione nelle zone ove l'errore numerico può concentrarsi. Naturalmente questa strategia presuppone che sia possibile, attraverso una fase di post-processing della soluzione, ricavare informazioni sulla distribuzione dell'errore, compito tutt'altro che facile con la maggior parte dei metodi numerici tradizionali.

Un ulteriore aspetto dell'analisi numerica Elettromagnetica non appare ancora pienamente soddisfacente. Si è osservato come l'intero universo dei fenomeni elettromagnetici sia rappresentato nelle equazioni di Maxwell; e quanto sia semplice passare da queste equazioni di carattere generale a quelle che descrivono i problemi dell'elettrostatica, della magnetostatica, della propagazione, dell'elettrocinetica, etc. Ad una siffatta semplicità nella formulazione dei modelli, purtroppo, non fa riscontro la disponibilità di metodi risolutivi numerici di carattere sufficientemente generale. E' stata così proposta nel passato una moltitudine di metodi ciascuno specializzato nel trattamento di uno solo o, comunque, pochi tra i particolari sottoproblemi di carattere elettromagnetico sopra citati. Complicazioni ulteriori si hanno, come appare naturale, nella trattazione di problemi accoppiati e, anzi, in diversi casi, esistono problematiche che, ancora oggi, non è possibile trattare se non al prezzo di semplificazioni tanto notevoli da inficiare la validità dei risultati ottenibili. E' questo, ad esempio, il caso del sistema delle equazioni MHD, che interessa fortemente il settore della fusione termonucleare controllata, la cui risoluzione operativa, salvo casi particolari, è disponibile solo nel caso di geometria assisimmetrica o quasi-assisimmetrica.

In questo lavoro il metodo solutivo basato sulla minimizzazione dell'errore sulle relazioni costitutive, proposto in passato per la risoluzione di problemi stazionari e quasi-stazionari, viene considerato e riproposto in una formulazione abbastanza generale da permettere lo studio non solo di una estesa classe di problemi di elettromagnetismo, ma anche di candidarsi come una valida alternativa nella soluzione di problemi accoppiati di tipo termico-elettromagnetico, meccanico-elettromagnetico, elettro-fluido-dinamico. Al vantaggio di un trattamento unificato dei problemi elettromagnetici, il presente metodo aggiunge il pregio di fornire informazioni immediate e precise circa la distribuzione spaziale e temporale dell'errore numerico associato alla soluzione. Queste informazioni possono essere utilizzate per ampliare in maniera opportuna lo spazio funzionale in cui si cerca la soluzione (ad

esempio, attraverso un raffinamento differenziato della discretizzazione del dominio del problema).

In queste pagine si illustrerà, innanzitutto, il metodo della minimizzazione dell'errore costitutivo applicato alle equazioni di Maxwell. Si mostrerà successivamente come il suddetto metodo sia estensibile al trattamento di problemi accoppiati, con particolare riferimento a quelli descritti dal modello MHD.

CAPITOLO I

Stato dell'arte sui metodi di analisi dei campi Elettromagnetici.

Introduzione

Le equazioni di Maxwell sono universalmente riconosciute come un esempio mirabile ed elegante di modello fisico-matematico per il carattere estremamente compatto col quale inquadrano ogni tipo di problematica si possa presentare nel settore dell'elettromagnetismo. Trascurando, quando possibile, in maniera opportuna qualcuno tra i termini sorgente ed aggiungendo relazioni costitutive, condizioni al contorno, ed, eventualmente, condizioni iniziali, si ricavano in maniera automatica i corretti modelli per qualunque problema di magnetostatica, di elettrostatica, di elettrocinematica, di propagazione. Nell'ipotesi di mezzi non dispersivi nel tempo o nello spazio, le equazioni per ciascuno dei particolari problemi elettromagnetici citati sono riportate in tabella 1 (e riferiti alla generica geometria di fig.1)

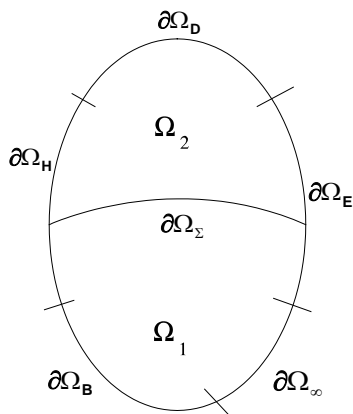


Fig.1 Dominio geometrico per i problemi di tabella 1

Problema	Eq.di Maxwell	Eq.costitutive	Eq.su interfacce tra mezzi diversi	Condizioni iniziali	Condizioni al contorno
Full Maxwell $\Omega = \Omega_1 \cup \Omega_2$ $\partial\Omega_\Sigma = \Omega_1 \cap \Omega_2$ $\partial\Omega = \partial\Omega_E \cup \partial\Omega_H \cup \partial\Omega_\infty$ $t \in [0, T]$	$\nabla \cdot \mathbf{D} = \rho$ $\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{H} = \mathbf{J} + \partial\mathbf{D}/\partial t$ $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\nabla \cdot \mathbf{J} = -\partial\rho/\partial t$ $\forall t, \forall \mathbf{x} \in \Omega$	$\mathbf{B}(\mathbf{x}, t) = \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t)$ $\mathbf{D}(\mathbf{x}, t) = \mathbf{f}_D(\mathbf{E}, \mathbf{x}, t)$ $\mathbf{J}(\mathbf{x}, t) = \mathbf{f}_J(\mathbf{E}, \mathbf{x}, t)$ $\forall t, \forall \mathbf{x} \in \Omega$	$[\mathbf{D}] \cdot \mathbf{n}_\Sigma = \rho_\Sigma$ $[\mathbf{B}] \cdot \mathbf{n}_\Sigma = 0$ $[\mathbf{E}] \times \mathbf{n}_\Sigma = 0$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma$ $[\mathbf{J}] \cdot \mathbf{n}_\Sigma = -\partial\rho_\Sigma/\partial t$ $\forall t, \forall \mathbf{x} \in \partial\Omega_\Sigma$	$\mathbf{D} _{t=0} = \mathbf{D}_0$ $\mathbf{B} _{t=0} = \mathbf{B}_0$ $\forall \mathbf{x} \in \Omega$	$[\mathbf{E}] \times \mathbf{n}_\Sigma = \mathbf{E}^*$ su $\partial\Omega_E$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{H}^*$ su $\partial\Omega_H$ regolarità su $\partial\Omega_\infty$ $\forall t, \forall \mathbf{x} \in \partial\Omega$
Magnetostatica $\Omega = \Omega_1 \cup \Omega_2$ $\partial\Omega_\Sigma = \Omega_1 \cap \Omega_2$ $\partial\Omega = \partial\Omega_B \cup \partial\Omega_H \cup \partial\Omega_\infty$	$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{H} = \mathbf{J}_S$ $\forall \mathbf{x} \in \Omega$	$\mathbf{B}(\mathbf{x}) = \mathbf{f}_B(\mathbf{H}, \mathbf{x})$ $\forall \mathbf{x} \in \Omega$	$[\mathbf{B}] \cdot \mathbf{n}_\Sigma = 0$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma$ $\forall \mathbf{x} \in \partial\Omega_\Sigma$		$[\mathbf{B}] \cdot \mathbf{n}_\Sigma = \mathbf{B}^*$ su $\partial\Omega_B$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{H}^*$ su $\partial\Omega_H$ regolarità su $\partial\Omega_\infty$ $\forall \mathbf{x} \in \partial\Omega$
Elettrostatica $\Omega = \Omega_1 \cup \Omega_2$ $\partial\Omega_\Sigma = \Omega_1 \cap \Omega_2$ $\partial\Omega = \partial\Omega_D \cup \partial\Omega_E \cup \partial\Omega_\infty$	$\nabla \cdot \mathbf{D} = \rho$ $\nabla \times \mathbf{E} = \mathbf{0}$ $\forall \mathbf{x} \in \Omega$	$\mathbf{D}(\mathbf{x}, t) = \mathbf{f}_D(\mathbf{E}, \mathbf{x})$ $\forall \mathbf{x} \in \Omega$	$[\mathbf{D}] \cdot \mathbf{n}_\Sigma = \rho_\Sigma$ $[\mathbf{E}] \times \mathbf{n}_\Sigma = 0$ $\forall \mathbf{x} \in \partial\Omega_\Sigma$		$[\mathbf{E}] \times \mathbf{n}_\Sigma = \mathbf{E}^*$ su $\partial\Omega_E$ $[\mathbf{D}] \cdot \mathbf{n}_\Sigma = \mathbf{D}^*$ su $\partial\Omega_D$ regolarità su $\partial\Omega_\infty$ $\forall \mathbf{x} \in \partial\Omega$
Eddy Currents (quasi-stazionario magnetico) $\Omega = \Omega_1 \cup \Omega_2$ $\partial\Omega_\Sigma = \Omega_1 \cap \Omega_2$ $\partial\Omega = \partial\Omega_E \cup \partial\Omega_H \cup \partial\Omega_\infty$ $t \in [0, T]$	$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{H} = \mathbf{J}$ $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\nabla \cdot \mathbf{J} = 0$ $\forall t, \forall \mathbf{x} \in \Omega$	$\mathbf{B}(\mathbf{x}, t) = \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t)$ $\mathbf{J}(\mathbf{x}, t) = \mathbf{f}_J(\mathbf{E}, \mathbf{x}, t)$ $\forall t, \forall \mathbf{x} \in \Omega$	$[\mathbf{B}] \cdot \mathbf{n}_\Sigma = 0$ $[\mathbf{E}] \times \mathbf{n}_\Sigma = 0$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma$ $[\mathbf{J}] \cdot \mathbf{n}_\Sigma = 0$ $\forall t, \forall \mathbf{x} \in \partial\Omega_\Sigma$	$\mathbf{B} _{t=0} = \mathbf{B}_0$ $\forall \mathbf{x} \in \Omega$	$[\mathbf{E}] \times \mathbf{n}_\Sigma = \mathbf{E}^*$ su $\partial\Omega_E$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{H}^*$ su $\partial\Omega_H$ regolarità su $\partial\Omega_\infty$ $\forall t, \forall \mathbf{x} \in \partial\Omega$
Propagazione in mezzi non conduttori (assenza di sorgenti) $\Omega = \Omega_1 \cup \Omega_2$ $\partial\Omega_\Sigma = \Omega_1 \cap \Omega_2$ $\partial\Omega = \partial\Omega_E \cup \partial\Omega_H \cup \partial\Omega_\infty$ $t \in [0, T]$	$\nabla \cdot \mathbf{D} = 0$ $\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{H} = \partial\mathbf{D}/\partial t$ $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ $\forall t, \forall \mathbf{x} \in \Omega$	$\mathbf{B}(\mathbf{x}, t) = \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t)$ $\mathbf{D}(\mathbf{x}, t) = \mathbf{f}_D(\mathbf{E}, \mathbf{x}, t)$ $\forall t, \forall \mathbf{x} \in \Omega$	$[\mathbf{D}] \cdot \mathbf{n}_\Sigma = \rho_\Sigma$ $[\mathbf{B}] \cdot \mathbf{n}_\Sigma = 0$ $[\mathbf{E}] \times \mathbf{n}_\Sigma = 0$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma$ $\forall t, \forall \mathbf{x} \in \partial\Omega_\Sigma$	$\mathbf{D} _{t=0} = \mathbf{D}_0$ $\mathbf{B} _{t=0} = \mathbf{B}_0$ $\forall \mathbf{x} \in \Omega$	$[\mathbf{E}] \times \mathbf{n}_\Sigma = \mathbf{E}^*$ su $\partial\Omega_E$ $[\mathbf{H}] \times \mathbf{n}_\Sigma = \mathbf{H}^*$ su $\partial\Omega_H$ regolarità su $\partial\Omega_\infty$ $\forall t, \forall \mathbf{x} \in \partial\Omega$

Tabella 1: Modelli matematici dei più comuni problemi elettromagnetici

Ciascuno dei problemi che compaiono nella tabella è stato già soddisfacentemente risolto utilizzando specifiche metodologie di carattere analitico o numerico. Si sono cioè costruite delle formulazioni in termini di campo o di potenziale -differenziando maggiormente questi ultimi in ridotti o totali, scalari o vettori, elettrici o magnetici-, ciascuna adatta in particolare a trattare problemi che rientrino o in una sola delle classi citate, oppure,

addirittura, in un particolare sottinsieme di una singola classe. E' il caso dei metodi analitici il cui impiego è tipicamente possibile solo per problemi il cui dominio di definizione sia particolarmente semplice, ad esempio riconducibile a un rettangolo, in geometria bidimensionale, o ad un parallelepipedo, se si opera in 3D, attraverso opportune trasformazioni di coordinate.

In questo capitolo sarà, innanzitutto, presentata una breve panoramica sulle diverse modalità operative disponibili per la soluzione di un problema E.M: metodi analitici, metodi numerici alle differenze finite, agli elementi finiti, agli elementi di frontiera.

In seguito, limitando l'attenzione alle sole formulazioni immaginate per una implementazione attraverso gli elementi finiti, saranno rivisitate, nel quadro della vasta letteratura disponibile, evidenziandone limiti e vantaggi, quelle che, per ciascuna delle principali classi di problemi elettromagnetici, maggiore successo hanno incontrato fino al momento attuale.

Il paragrafo conclusivo sarà, invece, dedicato alla presentazione del metodo basato sulla minimizzazione dell'errore costitutivo, soffermandosi sui principali risultati raggiunti nel recente passato nel trattamento di problemi E.M. stazionari e quasi-stazionari.

1.1) Generalità sui metodi analitici e numerici impiegati nella soluzione dei problemi elettromagnetici.

1.1.1) Metodi analitici

Fra i vari metodi analitici che si possono utilizzare per risolvere problemi di campo elettromagnetico, qui ci limiteremo a citare la funzione di Green, la separazione delle variabili e le trasformazioni conformi.

a) Funzione di Green

Si consideri per semplicità il problema di Poisson entro un volume finito V con condizioni di Dirichlet sul contorno Σ :

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0} \quad \text{in } V \quad (1a)$$

$$\phi = \phi_0 \quad \text{su } \Sigma \quad (1b)$$

dove ρ rappresenta una distribuzione nota di termini sorgente, e ϕ è il campo scalare incognito vincolato ad assumere sulla frontiera Σ il valore assegnato ϕ_0 .

La soluzione del problema può essere ottenuta per mezzo del teorema di Green:

$$\int_V (\phi \nabla^2 G - G \nabla^2 \phi) dV = \oint_{\Sigma} \left(\phi \frac{\partial G}{\partial n} - G \frac{\partial \phi}{\partial n} \right) d\Sigma \quad (2)$$

applicato a ϕ e alla *funzione di Green* $G(\mathbf{x}, \mathbf{x}')$ definita da:

$$\nabla^2 G(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}') \quad \text{in } V \quad (3a)$$

$$G = G_0 \quad \text{su } \Sigma \quad (3b)$$

Dal teorema di Green, dalla (3a) e dalla (1) si ottiene:

$$\phi(\mathbf{x}) = \frac{1}{\epsilon_0} \int_V \rho(\mathbf{x}') G(\mathbf{x}, \mathbf{x}') dV' + \oint_{\Sigma} \left[G(\mathbf{x}, \mathbf{x}') \frac{\partial \phi}{\partial n} - \phi(\mathbf{x}') \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n} \right] d\Sigma' \quad (4)$$

Se si sa risolvere il problema (3) per $G_0=0$, allora è possibile, attraverso la (4), risolvere il problema di Poisson con condizioni al contorno di Dirichlet in ogni punto del dominio V noti che siano i soli valori assunti da ϕ sulla frontiera Σ .

Attraverso un procedimento analogo si può risolvere il problema di Poisson con condizioni al contorno di Neumann.[1]

b) Metodo di separazione delle variabili

Si supponga di modellare un problema elettromagnetico definito in un dominio V attraverso una equazione alle derivate parziali ed opportune

condizioni imposte sulla frontiera ∂V del dominio. Se l'incognita è il campo scalare Φ , funzione delle n variabili indipendenti $\mathbf{x}=(x_1,\dots,x_n)$, l'equazione differenziale, supposta lineare e di ordine M , si può scrivere nella seguente forma di carattere generale [2]:

$$F\left(x_1, x_2, \dots, x_n, \frac{\partial \Phi}{\partial x_1}, \frac{\partial \Phi}{\partial x_2}, \dots, \frac{\partial \Phi}{\partial x_n}, \frac{\partial^2 \Phi}{\partial x_1^2}, \frac{\partial^2 \Phi}{\partial x_1 \partial x_2}, \dots, \frac{\partial^2 \Phi}{\partial x_n^2}, \dots\right) = 0 \quad \text{con } \mathbf{x} \in V$$

nella quale compare almeno una derivata parziale di ordine M .

In molti casi cercare una soluzione del tipo:

$$\Phi(x_1, x_2, \dots, x_n) = \varphi_1(x_1) \varphi_0(x_2, \dots, x_n)$$

permette di riscrivere l'equazione differenziale nella "forma separata":

$$F_1\left(x_1, \varphi_1, \frac{\partial \varphi_1}{\partial x_1}, \frac{\partial^2 \varphi_1}{\partial x_1^2}, \dots\right) = F_0\left(x_2, \dots, x_n, \varphi_0, \frac{\partial \varphi_0}{\partial x_2}, \dots, \frac{\partial \varphi_0}{\partial x_n}, \frac{\partial^2 \varphi_0}{\partial x_2^2}, \dots, \frac{\partial^2 \varphi_0}{\partial x_n^2}, \dots\right)$$

In questo caso la soluzione si può ottenere sovrapponendo le soluzioni di equazioni differenziali della forma:

$$\begin{cases} F_1\left(x_1, \varphi_1, \frac{\partial \varphi_1}{\partial x_1}, \frac{\partial^2 \varphi_1}{\partial x_1^2}, \dots\right) = C \\ F_0\left(x_2, \dots, x_n, \varphi_0, \frac{\partial \varphi_0}{\partial x_2}, \dots, \frac{\partial \varphi_0}{\partial x_n}, \frac{\partial^2 \varphi_0}{\partial x_2^2}, \dots, \frac{\partial^2 \varphi_0}{\partial x_n^2}, \dots\right) = C \end{cases}$$

dove C è una costante.

c) Metodo delle trasformazioni conformi

Un approccio matematico assai generale a problemi, in geometria bidimensionale, nei quali viene richiesto di determinare la distribuzione di un campo utilizza la teoria delle funzioni di variabile complessa. Sia $u(x,y)$ un campo vettoriale piano armonico (cioè irrotazionale e solenoidale) definito in una regione a connessione semplice. Sotto queste ipotesi sono definibili due funzioni scalari armoniche, $v(x,y)$ e $g(x,y)$ -denominate rispettivamente funzione potenziale e funzione flusso-, ciascuna associata in maniera biunivoca, a meno di due costanti arbitrarie, al campo $u(x,y)$. Si riconosce facilmente che le linee equipotenziali e quelle a flusso costante fanno parte di due famiglie di curve reciprocamente ortogonali. Si introducano la variabile complessa $z=x+jy$ e la funzione *potenziale complesso per il campo piano armonico* $F(z)=v(x,y)+jg(x,y)$. Si riconosce facilmente che $F(z)$ è una funzione analitica e che, dalla conoscenza di questa, si risale al campo u .

La possibilità di rappresentare un campo piano armonico tramite il potenziale complesso viene sfruttata per trasformare un problema di assegnate condizioni al contorno, e definito in un dominio Ω complesso, in un altro problema, ad esso correlato, ma definito in un dominio Ω^* per il quale la soluzione può essere calcolata più agevolmente. Le funzioni complesse analitiche $z^*=W(z)$ e $z=W^{-1}(z^*)$, che mettono in corrispondenza biunivoca i domini Ω e Ω^* , trasformano anche le linee equipotenziali e le linee di campo di un dominio nelle corrispondenti linee equipotenziali e di campo dell'altro. Viene pertanto conservata nella trasformazione da un dominio all'altro la proprietà di ortogonalità delle due famiglie di curve. Il metodo delle trasformazioni conformi trae il suo nome proprio da questa ultima circostanza.

Tra le varie trasformazioni conformi citiamo:

-la trasformazione bilineare che trasforma le circonferenze e le rette del piano Ω nelle rette e circonferenze del piano Ω^* : $z^* = \frac{\alpha z + \beta}{\gamma z + \delta}$ dove $\alpha, \beta, \gamma, \delta$,

sono costanti reali o complesse tali che $\alpha\delta - \beta\gamma \neq 0$.

-l'inversione per raggi reciproci che -indicato con \tilde{z} il numero complesso coniugato di z - associa a $z=r e^{j\varphi}$ il valore complesso

$$z^* = r^* e^{j\varphi^*} = \frac{r_0^2}{z} = \frac{r_0^2}{r} e^{j\varphi}$$

Questa trasformazione associa, in altre parole, ad ogni punto $P(r,\varphi)$ del piano Ω la sua immagine $P(r^*,\varphi^*=\varphi)$ rispetto alla circonferenza γ di raggio r_0 centrata nell'origine. Con tale trasformazione ogni problema esterno al cerchio centrato nell'origine si tramuta in un problema interno e viceversa.

1.1.2) Metodi numerici

L'analisi o il progetto di un sistema di interesse applicativo, richiede tipicamente anzitutto la definizione di un modello matematico approssimato del sistema che includa tutte le informazioni essenziali del problema fisico, e, successivamente, lo studio del modello alla ricerca della soluzione. All'approssimazione associata alla modellizzazione matematica del sistema reale è quasi sempre inevitabile aggiungere quella ulteriore corrispondente alla soluzione numerica del modello stesso. Sono proposte in letteratura varie possibili strategie per la risoluzione numerica dei principali modelli spaziotemporali. Tra tutte ci soffermeremo su quelle che maggiore diffusione e successo hanno ottenuto negli ultimi anni e, segnatamente, il *Metodo delle Differenze Finite (FDM)*, il *Metodo degli Elementi di Frontiera (BEM)*, ed il *Metodo degli Elementi Finiti (FEM)*. Tale descrizione appare opportuna oltre che per doveroso richiamo storico e per necessità di completezza, anche quale utile premessa all'argomento oggetto della tesi.

A scopo di chiarezza definiamo in modo descrittivo, piuttosto che matematico, quattro concetti che useremo ricorrentemente nel prosieguo del discorso e che risultano utili nell'impostazione del modello di qualunque problema di natura fisica: *sistema*, *dominio*, *equazioni governanti* e *termini di sorgente*.

Il *sistema* è tipicamente, ma non necessariamente, un oggetto fisico composto di vari materiali: solidi, liquidi, gas, plasmi, o combinazione

di questi. Un sistema elettromagnetico può includere materiali e/o lo spazio vuoto.

Il *dominio* del problema è tipicamente la regione spazio/temporale nella quale interessa studiare il sistema. In termini matematici esso rappresenta l'insieme di definizione delle variabili indipendenti spazio e tempo. Talvolta, in dipendenza dalle formulazioni, le variabili indipendenti possono anche essere grandezze di natura diversa come frequenze o coefficienti spettrali.

Le *equazioni governanti* sono le relazioni che vincolano le variabili descrittive del sistema sulla base delle leggi fisiche del problema in esame. Tipicamente si tratta di equazioni integro-differenziali che esprimono la conservazione o il bilancio di qualche proprietà fisica, come la massa, la carica, il momento o l'energia. Queste possono essere anche delle equazioni che esprimono un principio variazionale, ed allora si presentano come un vincolo di minimo di un'opportuno funzionale come un'energia potenziale. Sono incluse tra le equazioni governanti anche le relazioni costitutive, che descrivono le particolari proprietà fisiche dei materiali costituenti il sistema.

I *termini sorgente o carichi* descrivono le sollecitazioni di origine esterna che interagiscono col sistema: forze, correnti e campi impressi, etc. Le sorgenti agenti internamente al dominio vengono chiamate sorgenti interne e compaiono nelle equazioni governanti. Le sorgenti che intervengono sulla frontiera del dominio sono invece chiamate sorgenti di frontiera, e compaiono in equazioni separate dette condizioni al contorno.

a) Metodi alle differenze finite (FDM)

Tra i metodi numerici adoperati per risolvere equazioni differenziali alle derivate parziali il metodo alle differenze finite si distingue per l'ampiezza della classe dei problemi alla quale è applicabile. Sebbene la paternità del metodo sia attribuita a C.F.Gauss (1777-1855), e Boltzmann abbia dimostrato nel 1892 l'applicabilità delle equazioni alle differenze per risolvere l'equazione di

Laplace, è solo a partire dagli anni '40 che le differenze finite si sono vastamente diffuse. Rinviano a testi specializzati[2] per quanto riguarda le possibilità di applicazioni dei FDM alla soluzione di equazioni differenziali alle derivate parziali di forma generale, si illustrerà qui il metodo attraverso un semplice problema di elettrostatica definito all'interno del dominio bidimensionale Σ e modellabile attraverso l'equazione di Poisson ed opportune condizioni al contorno sulla frontiera $\partial\Sigma$ del dominio di definizione. In un sistema di riferimento cartesiano si può scrivere:

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = -\frac{\rho}{\varepsilon_0} \quad \text{su } \Sigma \quad (1)$$

$$\Phi(x,y) = \Phi_0(x,y) \quad \text{su } \partial\Sigma$$

dove $\Phi(x,y)$ è il potenziale scalare incognito, $\Phi_0(x,y)$ è il valore noto del potenziale sulla frontiera e ρ è la distribuzione di carica all'interno del dominio Σ .

Per risolvere il problema (1) attraverso i metodi FD bisogna suddividere la regione Σ in una rete di linee perpendicolari fra loro con un numero finito di intersezioni. L'equazione di Poisson viene quindi condotta in forma algebrica introducendo delle espressioni approssimate per le derivate seconde parziali rispetto a x e a y . In particolare, espandendo in serie di Taylor il potenziale scalare Φ attorno al punto (x,y) si ottiene:

$$\Phi(x+h, y) \cong \Phi(x, y) + h \frac{\partial\Phi(x, y)}{\partial x} + \frac{h^2}{2} \frac{\partial^2\Phi(x, y)}{\partial x^2} \quad (2)$$

$$\Phi(x-h, y) \cong \Phi(x, y) - h \frac{\partial\Phi(x, y)}{\partial x} + \frac{h^2}{2} \frac{\partial^2\Phi(x, y)}{\partial x^2} \quad (3)$$

Sommando la (2) e la (3) ed arrangiando i termini, la derivata parziale seconda rispetto a x si può scrivere in funzione del solo potenziale $\Phi(x,y)$. Attraverso operazioni analoghe si arriva ad un'espressione della derivata

parziale seconda rispetto a y in funzione del solo potenziale $\Phi(x,y)$. Sostituendo queste espressioni nell'equazione differenziale (1) si arriva alla seguente forma approssimata dell'equazione di Poisson:

$$\Phi(x+h, y) + \Phi(x-h, y) + \Phi(x, y+h) + \Phi(x, y-h) - 4\Phi(x, y) = -\frac{\rho h^2}{\varepsilon_0} \quad (4)$$

dove per semplicità si è preso lo stesso incremento di distanza h nelle due direzioni.

Si può scrivere, per ogni intersezione della rete, una equazione del tipo (4): si avrà, infine, un sistema di n equazioni nelle n incognite rappresentate dai potenziali dei nodi interni della rete. I nodi che si trovano sulla frontiera della rete non sono incognite in quanto vincolati ad assumere i valori imposti dalle condizioni al contorno. In principio il sistema algebrico cui si è giunti può essere invertito. Tuttavia, se n è un numero molto grande, può essere conveniente trovare la soluzione mediante aggiustamenti successivi dei valori dei potenziali fino ad arrivare a soddisfare le equazioni con un certo grado di approssimazione.

Il metodo descritto si complica quando la frontiera del dominio non è regolare, quando le condizioni al contorno sono di tipo Neumann o di altro tipo, e quando l'operatore differenziale è di natura diversa da quella dell'operatore di Laplace. In questi casi, può risultare conveniente utilizzare metodi basati sugli elementi finiti.

b) Metodo degli elementi di frontiera(BEM)

Un problema elettromagnetico può essere descritto da un sistema di equazioni differenziali associato con condizioni iniziali e sulla frontiera del dominio di definizione. Avendo definito una funzione di Green che soddisfi l'equazione differenziale e le condizioni al contorno omogenee per una

sorgente concentrata, la soluzione del problema di campo può essere espressa attraverso la somma di due termini, il primo dei quali è l'integrale esteso su tutto il dominio di definizione del problema della funzione sorgente moltiplicata per la funzione di Green, mentre il secondo termine è un integrale di superficie che tiene conto delle condizioni imposte sulla frontiera. Con questo sistema l'equazione differenziale viene ricondotta ad un'equazione integrale tipo la (1.1.1-4) che a volte può essere risolta più facilmente[3].

Durante gli anni '60 R.Harrington[4] sviluppò il metodo dei Momenti per la soluzione numerica delle equazioni integrali. In questo metodo le equazioni integrali sono trasformate in equazioni algebriche attraverso l'introduzione di funzioni test e peso per il campo incognito.

Alla fine degli anni '70 C.Brebbia[5] ha proposto il cosiddetto metodo degli elementi di frontiera (BEM) nel quale le equazioni integrali sono risolte discretizzando la frontiera in Elementi Finiti, detti appunto gli Elementi di Frontiera.

Per illustrare lo spirito del metodo si supponga di voler calcolare il valore del potenziale $\phi(\mathbf{x})$ che soddisfa il problema di Poisson nel dominio V con condizioni al contorno di tipo Neumann sulla frontiera ∂V :

$$\nabla^2\phi = -\frac{\rho}{\epsilon_0} \quad \text{in } V \quad (5a)$$

$$\frac{\partial\phi}{\partial n} = f_0(\mathbf{x}) \quad \text{su } \partial V \quad (5b)$$

Applicando il teorema di Green si è visto che attraverso la (1.1.1-4) è possibile calcolare il valore di $\phi(\mathbf{x})$ per ogni punto appartenente al dominio V a partire dalla funzione di Green che soddisfa l'equazione differenziale (5a) e dai valori assunti dal potenziale $\phi(\mathbf{x})$ e dalla sua derivata normale sulla frontiera ∂V . Di norma è dato conoscere solo una delle due ultime funzioni richieste (ad esempio, nel caso in esame la (5b) fornisce il valore della derivata del potenziale). Si è visto d'altronde che, se si è in grado di calcolare la funzione di Green che soddisfi alla (5a) e che inoltre abbia derivata normale nulla su ∂V ,

allora la sola condizione (5b) basta per determinare in maniera univoca il valore di $\phi(\mathbf{x})$ in ogni punto del dominio.

Va subito aggiunto che, in generale, è impresa tutt'altro che agevole determinare la funzione di Green che soddisfi alle condizioni richieste.

$$\phi(\mathbf{x}) = \frac{1}{\varepsilon_0} \int_V \rho(\mathbf{x}') G(\mathbf{x}, \mathbf{x}') dV' + \oint_{\Sigma} \left[G(\mathbf{x}, \mathbf{x}') \frac{\partial \phi}{\partial n} - \phi(\mathbf{x}') \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n} \right] d\Sigma'$$

Nei metodi BEM si calcola, allora, la particolare funzione di Green che soddisfa la condizione (5a) e le condizioni di regolarità all'infinito e si determina il valore assunto da $\phi(\mathbf{x})$ sulla sola frontiera ∂V risolvendo ivi numericamente l'equazione integrale (1.1.1-4) che nei punti regolari della frontiera assume la forma:

$$\frac{1}{2} \phi(\mathbf{x}) = \frac{1}{\varepsilon_0} \int_V \rho(\mathbf{x}') G(\mathbf{x}, \mathbf{x}') dV' + \oint_{\Sigma} \left[G(\mathbf{x}, \mathbf{x}') \frac{\partial \phi}{\partial n} - \phi(\mathbf{x}') \frac{\partial G(\mathbf{x}, \mathbf{x}')}{\partial n} \right] d\Sigma'$$

Calcolato in questo modo il valore del potenziale sulla frontiera, dalla (1.1.1-4) si può determinare il valore di $\phi(\mathbf{x})$ in ogni punto del dominio V .

Al contrario dei metodi FEM, che richiedono grossi sforzi di discretizzazione e utilizzano grosse matrici di tipo sparso, il BEM è caratterizzato da basso sforzo di modellizzazione e fa uso di matrici relativamente compatte anche se densamente popolate e non simmetriche. Nel quadro dell'elettromagnetismo i metodi BEM sono stati applicati a problemi bidimensionali e tridimensionali di magnetostatica, elettrostatica, elettrocinematica, propagazione. I maggiori inconvenienti ad essi collegati possono essere riassunti nei seguenti punti:

- anche nel caso in cui l'operatore di partenza sia autoaggiunto le matrici associate al sistema da risolvere cui si perviene non sono nè simmetriche nè tantomeno definite positive, e pertanto non è sempre possibile avere sufficienti garanzie di convergenza. Un ulteriore punto a favore dei metodi FEM è

costituito dal fatto che la soluzione numerica ha in alcuni punti un tasso di convergenza notevolmente più alto rispetto a quello medio sull'intero dominio (*superconvergenza*), proprietà questa che non è dimostrata per i metodi BEM.

- e' richiesta la soluzione numerica di un'equazione integrale sulla frontiera, operazione che può produrre forti malcondizionamenti (i cosiddetti "incubi algoritmici") a causa dei nuclei singolari presenti nelle funzioni di Green sotto il segno di integrale[6].

- Il minore sforzo di discretizzazione non è sempre un vantaggio significativo da un punto di vista operativo. Difatti, generalmente, nello studio di un problema pratico sono spesso richieste analisi ripetitive, sia al fine di stimare l'errore numerico associato alla soluzione sia, situazione questa abbastanza frequente, per iterazioni progettuali o, ancora, per studi parametrici. Il costo connesso con la preparazione di una complicata mesh FE è pertanto solo una piccola frazione del costo globale dell'analisi, ed anzi, grazie all'introduzione di sempre migliori preprocessori automatici, tende a diventare sempre meno significativo.

c) Metodo degli elementi finiti (FEM).

Il Metodo degli Elementi Finiti deve la sua attuale formalizzazione al contributo fornito, negli ultimi decenni, da vari ricercatori appartenenti ai diversi settori dell'ingegneria. Tra tutti, il pioniere può essere considerato Courant con un lavoro che risale al 1943 [7]. Nuovi sostanziali apporti al metodo si ebbero a partire dalla metà degli anni'50, grazie soprattutto alle favorevoli prospettive aperte dall'introduzione di calcolatori digitali dotati di memoria per l'immagazzinamento dei dati e dei programmi e capaci di svolgere computazioni ad "alta velocità". L'attuale denominazione del metodo si deve attribuire a Clough[8]. Successivi apporti determinanti sono stati da una parte quelli di Melosh, Besseling, Jones e de Veubeke, tutti compresi tra il 1963 e il 1964, che mostrarono come il metodo FEM potesse essere inquadrato nel metodo variazionale di Ritz adoperando delle funzioni polinomiali come

funzioni di base (Trial functions); altri contributi fondamentali sono quelli di Zienkiewicz e Cheung (1965) che hanno allargato enormemente il dominio di applicazione del metodo a tutti i problemi di campo che possono essere posti in forma variazionale. A partire da quegli anni i metodi FEM hanno conosciuto vastissima diffusione, ed innumerevoli sono i loro campi di applicazione: meccanica dei solidi, elasticità, elettromagnetismo, termodinamica, acustica, fenomeni accoppiati, e praticamente ogni problema matematico o fisico descritto da equazioni differenziali, integrali, integro-differenziali o variazionali.

Diamo ora una sintetica descrizione del modo in cui si opera nei metodi FEM.

Il dominio del problema è discretizzato in regioni più piccole chiamate elementi. Elementi adiacenti si toccano senza sovrapporsi, e non esiste alcun spazio del dominio non coperto da elementi. Le forme degli elementi sono tipicamente molto semplici: generalmente triangoli o quadrilaterali in 2D, e tetraedri, pentaedri o esaedri in geometria 3D. L'intero mosaico ottenuto prende il nome di griglia ("mesh"). Il processo di generazione della mesh, un tempo laborioso perchè effettuato in forma manuale, è ora più semplice e veloce grazie all'introduzione di preprocessori automatici che accettano quali dati in ingresso delle macro-descrizioni del dominio da discretizzare.

Per ogni elemento le equazioni governanti, usualmente in forma differenziale o variazionale (integrale), sono trasformate in equazioni algebriche, dette equazioni di elemento, tramite un processo di approssimazione. Questo passo viene eseguito attraverso la proiezione del campo incognito in uno spazio funzionale la cui base è costituita da funzioni di tipo polinomiale a tratti, ciascuna delle quali assume valore diverso da zero solo in una limitata regione della griglia, costituita da uno o, al più, pochi elementi tra loro adiacenti. I gradi di libertà della base dello spazio funzionale nel quale viene ricercata la soluzione vengono associati a parti geometriche, o *simplices*, della mesh - e cioè, nodi, lati, facce, volumi-. A seconda della scelta

operata si parla di *node-elements*, *edge-elements*, *facet-elements*, *volume-elements*.

Invece di calcolare direttamente le funzioni di forma associate ai simplices degli elementi reali della mesh si preferisce operare in maniera alternativa. Si supponga di rappresentare la mesh del dominio in un sistema di riferimento cartesiano, e sia \mathbf{x} il vettore delle coordinate del generico punto della mesh in tale sistema (\mathbf{x} è definito *vettore delle coordinate reali o globali*). Si introduce un elemento *master* (o *parent*) di caratteristiche geometriche regolari (ad esempio per trattare elementi esaedrali, l'elemento master è un cubo di spigolo di lunghezza unitaria) in un sistema di coordinate locali ξ . Si definiscono per l'elemento master le funzioni di forma $\phi_i(\xi)$ (dove, ad esempio, per i *node-elements*, i va da 1 al numero dei nodi dell'elemento), associate al particolare tipo di simplices che si intende utilizzare (sono queste le cosiddette *parent shape functions*). A questo punto, per ciascuno degli elementi reali (e) della mesh, si definisce un'applicazione invertibile $\mathbf{x}=\chi^{(e)}(\xi)$ che trasforma l'elemento master nell'elemento reale (*mapping*). Le funzioni di forma associate a ciascun elemento reale della mesh possono così essere calcolate attraverso le *parent shape functions* e la funzione inversa di *mapping* dell'elemento $\xi=\xi^{(e)}(\mathbf{x})$ come:

$$\phi_i^{(e)}(\mathbf{x})=\phi_i(\xi^{(e)}(\mathbf{x}))$$

Tra i vari possibili modi di stabilire il *mapping*, viene frequentemente adoperato il cosiddetto *approccio isoparametrico*:

$$\mathbf{x} = \sum_{k=1}^l \mathbf{x}_k^{(e)} \phi_k(\xi)$$

dove $\mathbf{x}_k^{(e)}$, $k=1,2,\dots,l$ sono le coordinate degli l nodi nell' $(e)^{\text{mo}}$ elemento reale della mesh, mentre le $\phi_k(\xi)$ sono le funzioni di forma nodali. Il nome *isoparametrico* deriva dal fatto che coi *node-elements* si adoperano le stesse funzioni $\phi_k(\xi)$ per approssimare il campo incognito $\mathbf{U}(\mathbf{x})$.

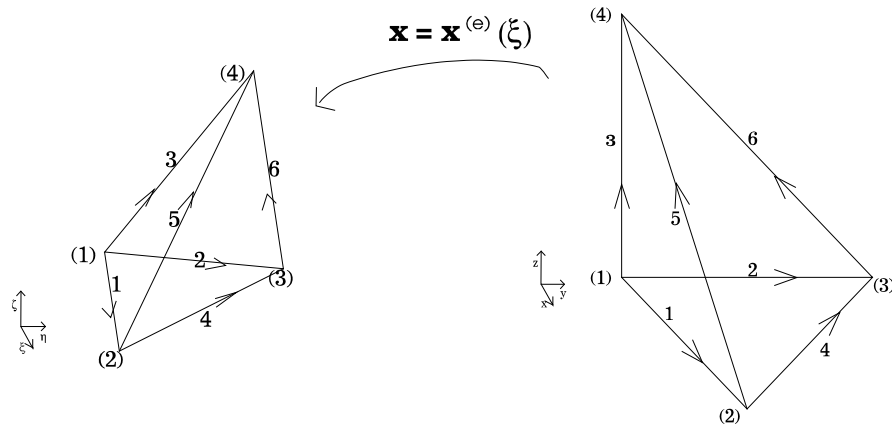


Fig.[2] Mapping dell'elemento master sull'elemento reale (e) attraverso la trasformazione $\mathbf{x}=\chi^{(e)}(\xi)$

L'opportunità di trasformare le equazioni di partenza in equazioni algebriche consente una radicale semplificazione delle complessità di calcolo e permette l'utilizzo di una delle numerose procedure solutive ciascuna delle quali specializzata nell'affrontare una particolare classe di sistemi di equazioni algebriche.

Due aspetti importanti vanno rimarcati. Primo: le equazioni di elemento sono dal punto di vista della struttura algebrica identiche per tutti gli elementi che sono dotati delle stesse caratteristiche geometriche. Pertanto lo sforzo analitico di formulare le equazioni di elemento va eseguito una sola volta per ciascuno degli elementi tipici utilizzati (solitamente uno o due) e non per tutti gli elementi della mesh. Secondo: poichè tanto la geometria degli elementi, quanto la struttura dello spazio funzionale approssimante vengono scelti di forma particolarmente semplice, la scrittura del sistema delle equazioni di elemento è molto agevole e non comprende più che pochi tipi di equazioni algebriche (tipicamente da 2 a 20).

I coefficienti delle equazioni di elemento sono valutati automaticamente, per ciascun elemento della mesh e successivamente assemblati in una matrice complessiva detta delle *equazioni di sistema*. Tale sistema di equazioni

può tuttavia essere risolto in tempi ragionevoli sfruttando il carattere di sparsità della matrice dei coefficienti, risultato questo del criterio adottato per la scelta della base di proiezione.

Solo a questo livello vengono imposte le condizioni al contorno ed, eventualmente, quelle di interfaccia tra sottoinsiemi del dominio, attraverso l'opportuna modifica delle equazioni di sistema o l'aggiunta di ulteriori equazioni di vincolo.

Resta ora da chiarire cosa siano le equazioni governanti per il metodo degli elementi finiti. A questo scopo definiamo una quantità chiamata residuo. Si supponga, per semplicità, che l'intero modello del problema sia rappresentato da una singola equazione differenziale. Si trasferiscano tutti i termini al lato sinistro (LHS), di modo che il lato destro (RHS) risulti nullo. Se ora si sostituisce alla funzione incognita, l'esatta soluzione $\mathbf{U}(\mathbf{x})$ dell'equazione differenziale, il LHS dell'equazione si annullerebbe per ogni \mathbf{x} eguagliando pertanto il RHS. Se, invece dell'esatta soluzione, ne usassimo per la sostituzione una approssimante $\mathbf{U}_t(\mathbf{x}, \mathbf{a})$ -dove \mathbf{a} rappresenta il vettore dei parametri liberi associati alle funzioni approssimanti $\phi_i(\mathbf{x})$ -, la differenza sarebbe una funzione $\mathbf{R}(\mathbf{x}, \mathbf{a})$ diversa da zero che viene detta residuo dell'equazione. Il principio ispiratore di tutti i metodi basati sui residui pesati è che la soluzione numerica di un sistema di equazioni può essere trasformata nella ricerca di quell'insieme di parametri capace di rendere minimo, nel dominio di definizione, la funzione residuo.

Ciò che, all'interno di questa classe di metodi, distingue uno schema da un altro è il criterio col quale si definisce la vicinanza tra la funzione nulla e la funzione residuo. Nascono, così, i metodi del collocamento, del sottodominio, dei minimi quadrati, di Galerkin. Quest'ultimo, in particolare, impone che, per ogni parametro \mathbf{a}_i , una media pesata di $\mathbf{R}(\mathbf{x}, \mathbf{a})$ sull'intero dominio sia zero. Le funzioni peso coincidono con le funzioni di base $\phi_i(\mathbf{x})$ associate a ciascun \mathbf{a}_i . Si ottiene in questo modo un sistema in cui le equazioni -dette governanti- sono in numero pari al numero di funzioni di base e, dunque, di incognite.

Il metodo dei residui di Galerkin è strettamente legato ad uno schema solutivo che trova vasta rilevanza nelle formulazioni FEM e che va sotto il nome di metodo Variazionale di Ritz. Mentre la parola variazionale riporta al vasto e complesso ambito del calcolo delle variazioni, il nome Ritz richiama una specifica, semplice e ben definita tecnica per ottenere soluzioni numeriche approssimate a problemi descritti in forma variazionale. Con un'espressione di carattere generale, si può affermare che vari problemi, retti da sistemi di equazioni differenziali, ammettono una formulazione equivalente di carattere variazionale. In questo caso, se denotiamo con $I(\mathbf{U})$ un possibile funzionale scalare di tipo integrale (che a volte può essere interpretato da un punto di vista fisico come un'energia associata al sistema), corrispondente al modello differenziale del problema in oggetto, il metodo variazionale di Ritz consiste nel ricercare, all'interno dello spazio delle funzioni base, la funzione $\mathbf{U}_t(\mathbf{x}, \mathbf{a})$ che rende estrema il funzionale $I(\mathbf{U})$. Si può dimostrare che le equazioni di elemento alle quali si perviene sono identiche a quelle ottenute attraverso il metodo di Galerkin. Va tuttavia osservato che il metodo di Galerkin ha un maggiore spettro di applicazioni rispetto a quello di Ritz, non richiedendo necessariamente che il problema ammetta una formulazione variazionale. D'altro canto la formulazione di Ritz, ove applicabile, risulta concettualmente superiore a quella di Galerkin. Si consideri, ad esempio, un sistema fisico in equilibrio stabile: ad esso corrisponde un minimo della sua energia potenziale; in questo caso il principio variazionale associato alle equazioni è un principio di minimo. In problemi siffatti la sequenza di valori dell'integrale $I(\mathbf{U}_t)$, corrispondente ad una sequenza di soluzioni approssimate, risulta convergere monotonicamente a patto che la sequenza degli spazi delle soluzioni sia costruita in modo che ogni termine della sequenza includa come suo sottoinsieme tutte le soluzioni rappresentabili attraverso lo spazio delle soluzioni dei termini precedenti (il che può essere ottenuto aggiungendo nuove funzioni di base senza modificare o rimuovere quelle precedenti, ovvero partizionando alcuni degli elementi della mesh in elementi più piccoli). Grazie

alla monotonia della convergenza può, in ogni momento, essere stabilito un limite superiore al valore del funzionale $I(\mathbf{U}_t)$. Inoltre, se esiste anche un principio variazionale complementare, ovvero se esiste la possibilità di trovare una espressione di I in un ulteriore campo incognito \mathbf{W} , per il quale il principio variazionale governante sia un principio di massimo, allora è possibile stabilire anche un limite inferiore per il valore del funzionale. Trovare limiti per $I(\mathbf{U}_t)$ può essere importante quando il funzionale rappresenta una quantità fisica di interesse; è questo il caso, ad esempio, dell'energia potenziale di una configurazione di equilibrio o della frequenza di risonanza di un sistema vibrante.

1.2) Modelli E.M. classici per formulazioni F.E.M.

1.2.1) Magnetostatica

Il calcolo del campo magnetostatico in geometria tridimensionale può essere effettuato attraverso una formulazione in termini di potenziale scalare.

Secondo l'approccio detto del "potenziale scalare ridotto", il campo magnetico \mathbf{H} è diviso nelle sue componenti irrotazionale \mathbf{H}_{irr} e solenoidale \mathbf{H}_{sol} . La parte solenoidale \mathbf{H}_{sol} può essere calcolata facilmente attraverso la legge dell'azione elementare di Laplace:

$$\mathbf{H}_{\text{sol}}(\mathbf{r}) = \frac{1}{4\pi} \int_{V_j} \frac{\mathbf{J}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} dV'$$

dove V_j è la regione in cui la densità di corrente \mathbf{J} è diversa da zero.

La parte irrotazionale di \mathbf{H}_{irr} , viene invece rappresentata come il gradiente di un potenziale scalare ϕ :

$$\mathbf{H}_{\text{irr}} = -\nabla\phi ,$$

per cui risulta:

$$\mathbf{H} = \mathbf{H}_{\text{sol}} - \nabla\phi$$

Il potenziale scalare ridotto è poi ottenuto risolvendo la seguente equazione:

$$\nabla \cdot \mu(\mathbf{H}_{\text{sol}} - \nabla\phi) = 0$$

che impone la solenoidalità del campo di induzione magnetica \mathbf{B} .

Il principale svantaggio di quest'approccio è la scarsa accuratezza dovuta ai fenomeni di cancellazione tra le due componenti del campo calcolate in modo indipendente, specialmente nelle regioni in cui la permeabilità magnetica è elevata. In aggiunta, uno sforzo computazionale notevole è richiesto per il calcolo dell'integrale in ogni punto del dominio ove \mathbf{H} è richiesto.

Per superare questi problemi, una formulazione alternativa combina l'uso del potenziale scalare ridotto con quello del potenziale scalare totale[9]. In questo caso il campo magnetico è descritto da

$$\begin{aligned} \mathbf{H} &= \mathbf{H}_{\text{sol}} - \nabla \phi && \text{in } V_j \\ \mathbf{H} &= -\nabla \psi && \text{nella regione } V - V_j \end{aligned}$$

dove V_j è la regione sede delle correnti.

I potenziali scalari sono così determinati dalle seguenti equazioni:

$$\begin{aligned} \nabla \cdot \mu(\mathbf{H}_{\text{sol}} - \nabla \phi) &= 0 && \text{in } V_j \\ \nabla \cdot \mu \nabla \psi &= 0 && \text{nella regione esterna } V - V_j \end{aligned}$$

che vengono accoppiate attraverso le condizioni di interfaccia sulla frontiera ∂V_j (cioè la continuità di $\mathbf{H} \times \mathbf{n}$ e $\mathbf{B} \cdot \mathbf{n}$ dove \mathbf{n} è il versore normale a ∂V_j).

Questa impostazione, tuttavia, attenua il problema della cancellazione solo se la permeabilità è costante nella regione V_j . Inoltre, a causa della polidromia della funzione potenziale scalare, per regioni molteplicemente connesse devono essere introdotti opportuni tagli.

Questi problemi scompaiono utilizzando formulazioni basate su potenziali vettori. Il prezzo da pagare è, in questo caso, costituito dall'accresciuto numero di incognite scalari e dalla necessità, in alcuni approcci, di imporre condizioni aggiuntive sui potenziali vettori allo scopo di garantire l'unicità della soluzione.

Appartengono a questa famiglia due classiche formulazioni basate l'una sul potenziale vettore magnetico \mathbf{A} (definito da $\mathbf{B}=\nabla\times\mathbf{A}$) e l'altra sul potenziale vettore elettrico \mathbf{T} (definito da $\mathbf{J}=\nabla\times\mathbf{T}$). Anche queste formulazioni possono essere usate in congiunzione coll'approccio del potenziale scalare totale.

Come già osservato, vincoli aggiuntivi, detti Gauges, devono essere imposti allo scopo di avere una corrispondenza biunivoca tra i potenziali vettori e i loro rotori. Una tra le Gauges più comuni nel problema magnetostatico è quella di Coulomb:

$$\nabla\cdot\mathbf{A}=0 \quad \nabla\cdot\mathbf{T}=0$$

Un'altra possibile scelta è :

$$\mathbf{w}\cdot\mathbf{A}=0 \quad \mathbf{w}\cdot\mathbf{T}=0$$

dove \mathbf{w} è un arbitrario campo vettoriale non dotato di linee di campo chiuse.

Adottando quest'ultima gauge e usando congiuntamente i potenziali vettori e i potenziali scalari totali, si arriva alle due formulazioni complementari magnetica ($\mathbf{A}-\psi$) ed elettrica ($\mathbf{T}-\Omega$).

Nel caso di mezzi lineari, la formulazione magnetica conduce al seguente sistema:

$$\begin{array}{ll} \nabla\times(\mu^{-1}\nabla\times\mathbf{A})=\mathbf{J} & \text{in } V_j \\ \mathbf{w}\cdot\mathbf{A}=0 & \text{in } V_j \\ \nabla\cdot\mu\nabla\psi=0 & \text{nella regione esterna } V-V_j \end{array}$$

Per quanto riguarda l'imposizione delle condizioni di interfaccia e l'eventuale introduzione di tagli, la formulazione presenta caratteristiche simili a quelle che combinano l'uso dei potenziali scalari ridotto e totale, col vantaggio, però, di eliminare i problemi legati alla cancellazione. Con questo approccio, nel caso di domini a connessione multipla, è, tuttavia, possibile estendere la regione V_j , in modo da rendere semplicemente connessa la regione esterna alle correnti ed avere di conseguenza una funzione potenziale scalare ψ monodroma.

La formulazione basata sul potenziale vettore elettrico \mathbf{T} è nota come il metodo $\mathbf{T}-\Omega$. In questo metodo \mathbf{T} differisce da \mathbf{H} per il gradiente di un potenziale scalare Ω :

$$\mathbf{H} = \mathbf{T} - \nabla\Omega$$

Tenendo conto che il potenziale vettore elettrico può essere definito solo nella regione ove la densità di corrente è diversa da zero, utilizzando per il resto del dominio il potenziale scalare totale, si perviene al sistema :

$$\begin{array}{ll} \nabla \times \mathbf{T} = \mathbf{J} & \text{in } V_j \\ \mathbf{w} \cdot \mathbf{T} = 0 & \text{in } V_j \\ \nabla \cdot \mu (\mathbf{T} - \nabla\Omega) = 0 & \text{in } V_j \\ \nabla \cdot \mu \nabla\Omega = 0 & \text{nella regione esterna } V - V_j \end{array}$$

Per quel che concerne tagli e condizioni di interfaccia valgono considerazioni analoghe a quelle relative alla formulazione $\mathbf{A}-\psi$.

Potrebbe sembrare che l'approccio $\mathbf{T}-\Omega$ sia simile a quello basato sui due potenziali scalari (\mathbf{H}_{sol} potrebbe, infatti, essere interpretato come il particolare potenziale vettore elettrico \mathbf{T} avente divergenza nulla). La differenza tra le due formulazioni è che nell'approccio coi potenziali scalari \mathbf{T} , o se si preferisce \mathbf{H}_{sol} , viene calcolato in maniera analitica, mentre in quello basato sul potenziale vettore elettrico la sua valutazione è ottenuta per via numerica proiettandolo sullo stesso spazio funzionale adoperato per approssimare il gradiente di Ω , la qual cosa limita notevolmente il problema degli errori di cancellazione.

1.2.2) Elettrocinematica quasi-stazionaria

La determinazione delle correnti parassite (Eddy currents) indotte nei conduttori da un campo magnetico tempo-variante è basata sulla soluzione delle equazioni di Maxwell scritte sotto l'approssimazione quasi-stazionaria:

$$\nabla \cdot \mathbf{D} = \rho \quad (1)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (3)$$

$$\nabla \times \mathbf{H} = \mathbf{J} \quad (4)$$

$$\nabla \cdot \mathbf{J} = 0 \quad (5)$$

dove $\mathbf{D}, \mathbf{B}, \mathbf{E}, \mathbf{H}$ sono gli usuali campi vettoriali, mentre ρ e \mathbf{J} sono rispettivamente la densità di carica libera e la densità di corrente. Si noti che la (5) non è un'equazione indipendente dalle altre 4, ma è direttamente implicata dalla (4).

Il modello si completa aggiungendo le relazioni costitutive, le eventuali relazioni di interfaccia, le condizioni iniziali e al contorno.

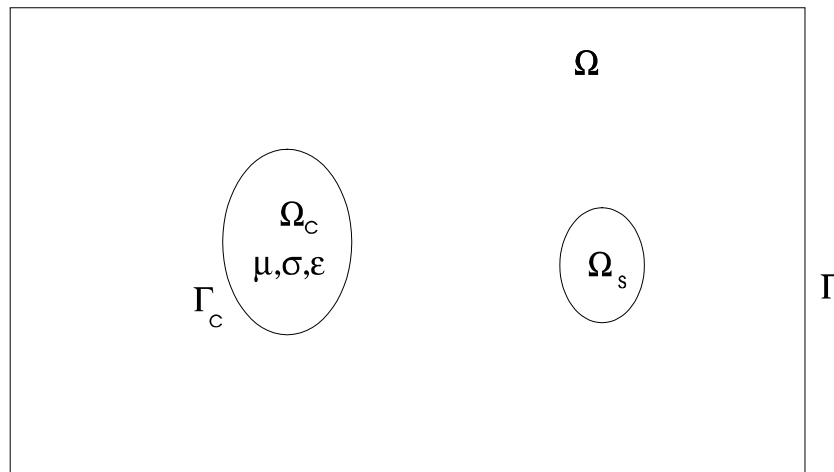


Fig.[3] Semplice configurazione per il problema delle correnti parassite

Per la risoluzione del problema sono stati proposti diverse formulazioni. Nel seguito vengono richiamate alcune tra le più comuni, adeguate, tra l'altro, per il trattamento di domini molteplici e connessi.

Metodo A-V-φ

Una prima possibilità è quella di fare uso del potenziale vettore magnetico \mathbf{A} , ponendo [10-11]

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (6)$$

L'equazione (3) diventa:

$$\nabla \times \left(\mathbf{E} - \frac{\partial \mathbf{A}}{\partial t} \right) = \mathbf{0} \quad (7)$$

da cui si ricava:

$$\mathbf{E} = - \left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) \quad (8)$$

dove V è un potenziale scalare. Per avere unicità della soluzione sarà necessario imporre una condizione sulla divergenza di \mathbf{A} (Gauge) e sul valore assunto da V sul contorno del dominio.

Le Gauges più diffuse nelle applicazioni sono quelle di Coulomb e di Lorentz.

Per mezzi materiali lineari, nella regione dei conduttori, le equazioni di campo (4) e (5) diventano:

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{A} + \sigma \left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) = \mathbf{J}_s \quad (9)$$

$$\nabla \cdot \sigma \left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) = 0 \quad (10)$$

dove l'equazione (10) non è indipendente dalla (9) potendosi ricavare da questa attraverso l'operatore divergenza.

Nella regione non conduttrice dove $\sigma=0$ e $\nabla \times \mathbf{H} = \mathbf{J}_s$ si può adoperare il potenziale scalare magnetico ridotto:

$$\nabla \cdot \mu \nabla \varphi = 0 \quad (11)$$

$$\mathbf{H} = \mathbf{H}_S - \nabla \varphi \quad (12)$$

oppure una delle altre formulazioni illustrate nel precedente paragrafo sulla magnetostatica.

Sulla frontiera dei conduttori la condizione di continuità sulla componente normale di \mathbf{J} implica che:

$$\mathbf{J}_n = -\sigma \left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) \cdot \mathbf{n} = 0$$

$$\frac{\partial \mathbf{A}_n}{\partial t} + \frac{\partial V}{\partial n} = 0$$

mentre all'interfaccia tra mezzi conduttori di diversa conducibilità risulta:

$$\sigma_1 \left(\frac{\partial \mathbf{A}_1}{\partial t} + \nabla V_1 \right) \cdot \mathbf{n} = \sigma_2 \left(\frac{\partial \mathbf{A}_2}{\partial t} + \nabla V_2 \right) \cdot \mathbf{n}$$

Se si impone la gauge di Coulomb, $\nabla \cdot \mathbf{A} = 0$, l'equazione (9) diventa:

$$\nabla^2 \left(\frac{1}{\mu} \nabla \cdot \mathbf{A} \right) = 0 \quad (13)$$

che è indipendente dalla (10).

Dalla (13) si può vedere che se si impone $\nabla \cdot \mathbf{A} = 0$ si forza in modo automatico la gauge di Coulomb.

Se, viceversa, si adopera la gauge di Lorentz, che per i problemi di correnti parassite assume la forma $\nabla \cdot \mathbf{A} = -\mu \sigma V$ si perviene, nel caso di mezzi lineari ed omogenei, alle equazioni:

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{A} = -\sigma \frac{\partial \mathbf{A}}{\partial t} + \sigma \nabla \left(\frac{1}{\mu \sigma} \nabla \cdot \mathbf{A} \right) \quad (14)$$

$$-\mu \sigma^2 \frac{\partial V}{\partial t} + \nabla \cdot \sigma (\nabla V) = 0 \quad (15)$$

Il maggior vantaggio di questa gauge è il disaccoppiamento che determina tra le due equazioni. Le equazioni (13) e (14) ammettono soluzione unica se viene assegnato V oppure $\mathbf{A} \cdot \mathbf{n}$ sulla superficie del conduttore. L'unicità è garantita anche nel dominio della frequenza purchè sia $\omega \neq 0$. A frequenza zero, invece, l'unicità della soluzione è garantita solo se la condizione al contorno viene imposta attraverso $\mathbf{A} \cdot \mathbf{n}$.

Così come la Gauge di Coulomb, la Gauge di Lorentz può essere applicata anche in presenza di mezzi disomogenei o non lineari. Tuttavia, in questi casi, il disaccoppiamento tra \mathbf{A} e V non è garantito.

Metodo del potenziale vettore modificato \mathbf{A}^* e metodo \mathbf{E}

Se, con riferimento al metodo precedente poniamo:

$$\mathbf{E} = -\frac{\partial \mathbf{A}^*}{\partial t} = -\left(\frac{\partial \mathbf{A}}{\partial t} + \nabla V \right) \quad (16)$$

è possibile eliminare il potenziale scalare V . Le equazioni di campo in termini di \mathbf{A}^* -chiamato *potenziale vettore modificato* [12]- si riducono alla sola

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{A}^* + \sigma \frac{\partial \mathbf{A}^*}{\partial t} = \mathbf{J}_s$$

La posizione (16) equivale ad imporre la condizione di Gauge $\nabla \cdot (\sigma \mathbf{A}) = 0$, che per una singola regione conduttrice con σ costante equivale alla Gauge di Coulomb.

I maggiori limiti di questa formulazione sono essenzialmente due:

-la formulazione diventa estremamente pesante qualora si intenda trattare situazioni in cui nel dominio del problema siano presenti regioni adiacenti di differente conducibilità a causa delle difficoltà operative che si incontrano

nell'imporre la continuità della componente normale di \mathbf{J} sulla superficie di interfaccia;

-la presenza dell'operatore rot-rot può produrre matrici singolari o molto mal-condizionate quando la conducibilità tende a zero oppure nel caso di problemi in bassa frequenza.

Si noti che impiegare il potenziale vettore modificato equivale a risolvere il problema direttamente in termini del campo elettrico \mathbf{E} .

Vettore campo magnetico \mathbf{H} - φ

Se, al posto di introdurre potenziali vettori, si risolve direttamente nell'incognita \mathbf{H} , nelle regioni a conducibilità costante si giunge all'equazione[13]:

$$\nabla \times \nabla \times \mathbf{H} + \sigma \frac{\partial (\mu \mathbf{H})}{\partial t} = 0 \quad (17)$$

Nelle regioni in cui σ è uguale a zero e non vi sono correnti impresse, si può porre $\mathbf{H} = \nabla \varphi$ dove φ è una funzione armonica che può essere facilmente accoppiata ad \mathbf{H} sulle frontiere di interfaccia tra mezzi conduttori e non-conduttori imponendo che la componente tangente del campo \mathbf{H} coincida con la derivata direzionale di φ lungo la tangente.

Nelle implementazioni FEM, per le difficoltà associate all'imposizione delle condizioni fisiche di interfaccia sulla componente tangente di \mathbf{H} attraverso funzioni di base nodali, questa formulazione viene comunemente associata all'uso delle edge functions.

Potenziale vettore elettrico (metodo \mathbf{T} - Ω)

La condizione di solenoidalità sulla densità di corrente \mathbf{J} , può essere imposta attraverso la posizione[14]:

$$\underline{\mathbf{J} = \nabla \times \mathbf{T}}$$

Il campo magnetico \mathbf{H} , si può invece esprimere come:

$$\mathbf{H} = (\mathbf{T} - \nabla\Omega)$$

dove Ω è il potenziale scalare magnetico.

Le equazioni (2) e (3) diventano in questa formulazione:

$$\begin{aligned}\nabla \cdot \mu (\mathbf{T} - \nabla\Omega) &= 0 \\ \nabla \times \frac{1}{\sigma} \nabla \times \mathbf{T} + \mu \frac{\partial (\mathbf{T} - \nabla\Omega)}{\partial t} &= 0\end{aligned}$$

Come per la formulazione \mathbf{H} , è possibile adoperare un potenziale scalare armonico nelle regioni in cui σ è uguale a zero e non vi sono correnti impresse. Anche per il potenziale vettore elettrico \mathbf{T} è necessario imporre una gauge per avere soluzione unica. Una possibile scelta è:

$$\mathbf{T} \cdot \mathbf{u} = 0$$

dove \mathbf{u} è un arbitrario campo vettoriale che non possiede linee di campo chiuse. In questo modo, oltre ad assicurare l'unicità della soluzione, si ha l'ulteriore vantaggio di ridurre il numero di gradi di libertà di \mathbf{T} da tre a due.

Il principale limite della formulazione è la scarsa accuratezza dei risultati in presenza di regioni ferromagnetiche, specie durante le fasi transitorie[15].

1.2.3) Propagazione

I fenomeni di tipo propagativo costituiscono una famiglia di problemi estremamente vasta e diversificata. Appare opportuno introdurre una classificazione più dettagliata:

1) *Problemi propagativi transitori*. Vengono generalmente affrontati nel dominio del tempo. Vengono adoperati metodi alle differenze finite (time-

stepping) per l'avanzamento nel tempo delle equazioni differenziali associate al problema.

2) *Problemi propagativi in regime armonico nello spazio libero o in mezzi lineari*: L'approccio classico è quello di ridurre le equazioni differenziali in equazioni algebriche complesse attraverso la trasformata di Fourier. La soluzione viene quindi determinata nel dominio della frequenza.

Data la vastità del problema, risulta impossibile anche solo menzionare tutte le formulazioni proposte e i problemi di interesse del settore. Ci si limiterà, pertanto, ad accennare solo ad alcuni argomenti, e in particolare:

- la determinazione degli autovalori nelle cavità risonanti;
- la propagazione nelle guide d'onda in condizioni di regime sinusoidale;
- la questione dei modi spuri;

Riguardo ai problemi propagativi a carattere transitorio si rimanda invece al capitolo IV.

Determinazione degli autovalori nelle cavità risonanti

L'energia elettromagnetica può essere trasmessa attraverso conduttori cavi dette guide d'onda. Nel corso della propagazione, i campi possono assumere diverse configurazioni dette modi la cui forma dipende dalle caratteristiche geometriche ed elettriche della struttura guidante. Una cavità risonante può essere considerata una guida d'onda cortocircuitata ad una estremità. Se la guida viene alimentata all'altra estremità si stabilisce un'onda stazionaria in analogia a quanto succede ad una linea in corto-circuito; in particolare, indicata con λ_g la lunghezza d'onda del campo elettromagnetico nella guida (il cui valore è legato alla geometria della guida e alla lunghezza d'onda del segnale), a distanza $\lambda_g/2$ dalla terminazione il campo è nullo. Si può pertanto chiudere la guida con un piano posto a distanza $\lambda_g/2$ dalla terminazione senza disturbare la distribuzione del campo: si realizza in questa maniera una cavità risonante. Il risonatore può avere più modi di oscillazione e quindi più frequenze di risonanza: il modo fondamentale è quello che ha la frequenza più bassa.

La soluzione del problema della cavità risonante implica la soluzione di un problema agli autovalori. Le equazioni governanti in una formulazione basata sul potenziale vettore magnetico e sul potenziale scalare sono[16-17]:

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{A} = (\varepsilon \omega^2 - j \omega \sigma) \mathbf{A} - (\sigma + j \omega \varepsilon) \nabla V \quad \text{in } \Omega,$$

dove per le cavità risonanti si assume che la variazione temporale dei campi è armonica e che la frequenza sia abbastanza alta da escludere la penetrazione del campo nei conduttori (questo implica che i conduttori possano essere rappresentati in termini di condizioni sulla frontiera, Σ , mentre all'interno del dominio, Ω , si possa assumere $\sigma=0$). Per assicurare l'unicità della soluzione, può essere adottata la gauge di Lorentz nel limite ad alta frequenza:

$$\begin{aligned} \nabla \cdot \mathbf{A} &= -j \omega \mu \varepsilon V & \text{in } \Omega \\ \mathbf{A} \cdot \mathbf{n} &= -j \omega \varepsilon \beta V & \text{su } \Sigma \end{aligned}$$

dove β è una costante positiva arbitraria, mentre \mathbf{n} è il versore uscente normale a Σ .

Problemi in guide d'onda

Se il problema consiste nel determinare il comportamento di un dato modo (noto) propagantesi in una guida d'onda nella quale vari ostacoli possono essere incontrati (ad esempio conduttori non-ideali, permittività complesse) bisogna adottare la gauge di Lorentz nella sua forma completa[18]:

$$\nabla \cdot \mathbf{A} = -\mu \sigma V - \mu \varepsilon \frac{\partial V}{\partial t} \quad \text{in } \Omega$$

in tal caso l'equazione governante assume la forma:

$$\nabla \times \frac{1}{\mu} \nabla \times \mathbf{A} + \sigma \frac{\partial \mathbf{A}}{\partial t} + \varepsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \frac{1}{\mu} \nabla \cdot \mathbf{A} = \mathbf{0} \quad \text{in } \Omega$$

In questo modo si determina il potenziale vettore magnetico e quindi il campo magnetico. Per la determinazione del campo elettrico è necessario calcolare V . Esso si ottiene risolvendo l'equazione:

$$\nabla \cdot \sigma \nabla V + \nabla \cdot \varepsilon \frac{\partial \nabla V}{\partial t} = -\nabla \cdot \sigma \frac{\partial \mathbf{A}}{\partial t} - \nabla \cdot \varepsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} \quad \text{in } \Omega$$

dove il potenziale \mathbf{A} è una quantità nota dalla soluzione dell'equazione precedente.

Modi spuri

Tra gli argomenti che, nel settore della propagazione, attualmente maggiore interesse destano nel mondo scientifico, spicca senz'altro quello dell'identificazione e dell'eliminazione dei cosiddetti modi spuri dalle soluzioni numeriche dei problemi di cavità risonanti. Una possibile origine delle soluzioni spurie si individua facilmente esaminando il semplice modello del problema propagativo quando ci si trova in condizioni di regime armonico, in presenza di mezzi omogenei non-dissipativi e in assenza di termini sorgente:

$$\begin{array}{llll} \nabla \times \hat{\mathbf{E}} = -j\omega\mu_0\mu_r \hat{\mathbf{H}} & \text{in } \Omega & \mathbf{(1a)} & \nabla \cdot \hat{\mathbf{H}} = 0 & \text{in } \Omega & \mathbf{(2a)} \\ \nabla \times \hat{\mathbf{H}} = j\omega\varepsilon_0\varepsilon_r \hat{\mathbf{E}} & \text{in } \Omega & \mathbf{(1b)} & \nabla \cdot \hat{\mathbf{E}} = 0 & \text{in } \Omega & \mathbf{(2b)} \\ \hat{\mathbf{E}} \times \mathbf{n}_\Sigma = 0 & \text{in } \Sigma_E & \mathbf{(1c)} & \hat{\mathbf{H}} \cdot \mathbf{n}_\Sigma = 0 & \text{in } \Sigma_E & \mathbf{(2c)} \\ \hat{\mathbf{H}} \times \mathbf{n}_\Sigma = 0 & \text{in } \Sigma_H & \mathbf{(1d)} & \hat{\mathbf{E}} \cdot \mathbf{n}_\Sigma = 0 & \text{in } \Sigma_H & \mathbf{(2d)} \end{array}$$

dove la frontiera del dominio di definizione del problema Ω è stata divisa in due parti per tenere conto delle due possibili condizioni al contorno: Σ_H (piano di simmetria) e Σ_E (conduttore perfetto o piano di anti-simmetria).

Per $\omega=0$ tale problema si scinde in due problemi distinti (uno per $\hat{\mathbf{E}}$ e uno per $\hat{\mathbf{H}}$). Ricordando il teorema di Helmholtz è facile vedere che l'unica soluzione possibile per i campi è quella banale $\hat{\mathbf{E}}=\hat{\mathbf{H}}=\mathbf{0}$.

Per $\omega \neq 0$ le equazioni relative alla divergenza e alle componenti normali dei campi sulla frontiera sono automaticamente dedotte dalle equazioni immediatamente precedenti e, pertanto, esse spesso non vengono incluse esplicitamente nelle formulazioni. Dato, però, che l'implicazione non sussiste per $\omega=0$ e, d'altra parte, le equazioni (2) sono essenziali perchè la soluzione sia $\hat{\mathbf{E}}=\hat{\mathbf{H}}=\mathbf{0}$, la loro omissione genera soluzioni non banali a frequenza nulla che il problema originale non ammette: le soluzioni spurie.

Il problema modellato attraverso le relazioni (1-2) può essere riformulato in forma equivalente attraverso uno solo dei campi ad esempio il campo elettrico $\hat{\mathbf{E}}$ come:

$$\begin{aligned}
 (\epsilon\mu)^{-1}\nabla\times\nabla\times\hat{\mathbf{E}} &= k^2\hat{\mathbf{E}} & \text{in } \Omega & \quad \text{(3a)} & \quad \nabla\cdot\hat{\mathbf{E}} &= 0 & \text{in } \Omega & \quad \text{(3b)} & \quad \hat{\mathbf{E}}\times\mathbf{n}_\Sigma &= \\
 0 & & \text{in } \Sigma_E & \quad \text{(3c)} & \quad \nabla\times\hat{\mathbf{E}}\times\mathbf{n}_\Sigma &= 0 & \text{in } \Sigma_H & \quad \text{(3d)} \\
 \hat{\mathbf{E}}\cdot\mathbf{n}_\Sigma &= 0 & & & & & \text{in } \Sigma_H & \quad \text{(3e)}
 \end{aligned}$$

dove $k^2 = \omega^2\mu_0\epsilon_0$. Si può facilmente dimostrare che la formulazione(3) esclude a priori la presenza di modi spurie.

Va, tuttavia, osservato che la (3) impone quattro equazioni scalari, lasciando a disposizione tre soli gradi di libertà. Nel caso delle guide d'onda e delle cavità a simmetria assiale il problema si risolve esplicitando la (3b) in termini di una delle componenti di $\hat{\mathbf{E}}$, sostituendo nella (3a) ed usando poi due sole delle tre equazioni che la (3a) mette a disposizione[19]. Nel caso di formulazioni 3D, tuttavia il metodo non è applicabile e pertanto vanno usate formulazioni differenziali che non sono a priori esenti dalla presenza di modi spurie.

Appare opportuno, a questo proposito, distinguere le formulazioni proposte in due gruppi: da una parte si possono collocare i metodi caratterizzati

dall'impiego di funzioni di forma nodali, dall'altra quelli basati sull'impiego degli edge elements.

Nel caso di elementi finiti nodali, le approssimazioni inaccurate del numero d'onda zero non possono essere distinte dagli autovalori fisicamente diversi da zero, che anzi sovente risultano contaminati da parte dei modi spuri. D'altra parte il fatto che la divergenza delle soluzioni spurie è diversa da zero può essere adoperata per filtrare le soluzioni reali da quelle dovute all'approssimazione numerica.

Uno dei primi lavori in cui si tratta intorno alle soluzioni spurie, è quello di Davies[20], il quale, pur non adottando alcun accorgimento per forzare la condizione di solenoidalità dei campi, propone un semplice criterio per identificare le soluzioni spurie: integrare il quadrato della divergenza.

Successivamente venne proposto di eliminare la divergenza dei campi attraverso l'introduzione di un termine di penalità[21]. Questo approccio, tuttavia, anzichè eliminare le soluzioni spurie, finisce col creare una nuova sorgente di modi spuri[22]. Webb risolve il problema attraverso l'introduzione di un algoritmo iterativo che aggiusta in maniera conveniente per ciascun modo i coefficienti di penalità[23].

Nelle formulazioni basate sugli edge elements poichè i modi spuri compaiono a frequenza esattamente nulla ed anche in numero prevedibile a priori, si può evitare di calcolarli. Per questo in letteratura le formulazioni basate sugli edge elements sono indicate con una terminologia un po' impropria come esenti da soluzioni spurie[24-25].

1.3) Rassegna sui metodi basati sui principi variazionali complementari. Il metodo della minimizzazione dell'errore costitutivo.

1.3.1) CAMPI STATICI

La formulazione basata sulla minimizzazione dell'errore costitutivo è nata nell'ambito della teoria dei principi variazionali complementari (**CVP**). I

metodi CVP, già brevemente introdotti nelle pagine precedenti, sono stati sviluppati originariamente per applicazioni nei settori della meccanica analitica e dell'elasticità e, successivamente adattati per essere applicati in problemi di elettromagnetismo[26,27]. Questi metodi, formulano una serie di ipotesi sulle specifiche del problema (ad esempio, escludendo domini molteplici e connessi e sorgenti superficiali o, in generale, restringendo la classe dei problemi trattabili) e, poi, derivano in vario modo le espressioni dei funzionali complementari[28].

In particolare, Penman e Fraser[27] identificano due insiemi di equazioni canoniche che descrivono la struttura matematica del problema di campo stazionario, e le integrano alternativamente nei funzionali di Hu-Washizu, usati originariamente nella teoria dell'elasticità. Giungono, così, a due coppie di funzionali complementari. La loro formulazione, pur rimarchevole per la concisione con la quale viene evidenziata la struttura del problema di campo statico, fallisce in presenza di mezzi non-lineari, di regioni molteplici e connesse, di sorgenti superficiali ed, inoltre, in presenza di condizioni al contorno che, anche se fisicamente ammissibili, non riescono sempre ad essere convenientemente inquadrare nel modello da loro proposto.

Hammond[26] parte, invece, dalle equazioni della meccanica analitica formulate attraverso il principio dei lavori virtuali. Quindi, con appropriate sostituzioni, ed attraverso trasformazioni degli integrali di volume coinvolti e integrazioni rispetto ai parametri variati, perviene a due coppie di funzionali complementari. Rispetto alla formulazione precedente essa rappresenta un significativo progresso in due punti in particolare: consente il trattamento di non-linearità e propone una interpretazione fisica a processi variazionali di carattere essenzialmente matematico (il funzionale da lui proposto si ricollega direttamente alla teoria Hamiltoniana).

Rikabi si mette sulla scia di questa attività, giungendo ad una formulazione che ha valenza molto più generale. Egli difatti [28], si

preoccupa innanzitutto di definire da una parte la struttura formale di un problema stazionario ben posto e dall'altra il range di applicabilità della teoria CVP.

Partendo dall'ipotesi di relazioni costitutive monotone, ed usando congiuntamente simboli magnetostatici -per i campi e le distribuzioni di corrente- ed elettrostatici -per le distribuzioni di carica- sviluppa una formulazione valida per ogni problema elettromagnetico di natura stazionaria. Con riferimento alle due coppie di campi $(\mathbf{H}_1(\mathbf{r}), \mathbf{B}_1(\mathbf{r}))$ e $(\mathbf{H}_2(\mathbf{r}), \mathbf{B}_2(\mathbf{r}))$ nelle ipotesi che:

- I) i campi $\mathbf{H}(\mathbf{r})$ rispettino la legge di Ampère;
- II) i campi $\mathbf{B}(\mathbf{r})$ rispettino la legge di Gauss;
- III) per tutti i campi valgano le rispettive condizioni di interfaccia associate alle leggi;
- IV) che ciascuna coppia (\mathbf{H}, \mathbf{B}) soddisfi la relazione costitutiva;

dimostra che, se le condizioni al contorno sono imposte in termini di \mathbf{H}_t , \mathbf{B}_n e di regolarità dei campi all'infinito, l'unicità della soluzione è garantita.

Partendo da queste premesse, e conservando l'ipotesi di monotonia per le relazioni costitutive, propone di usare un modello composto da due sistemi complementari, comunicanti solo attraverso la relazione costitutiva e basati rispettivamente sui campi incogniti \mathbf{H} e \mathbf{B} , in modo che il sistema \mathbf{H} rispetti la legge di Ampère e la relativa condizione d'interfaccia, e il sistema \mathbf{B} soddisfi la legge di Gauss con la condizione d'interfaccia associata.

Attraverso le espressioni dell'energia e della coenergia associate ai campi, costruisce, quindi, un funzionale scalare $\Lambda(\mathbf{H}, \mathbf{B})$ definito positivo, il cui valore è zero solo nell'ipotesi che i sistemi complementari forniscano campi che soddisfano la relazione costitutiva.

Introduce quindi un potenziale scalare Ω nel sistema \mathbf{H} e un potenziale vettore \mathbf{A} nel sistema \mathbf{B} in modo che i campi siano espressi attraverso le relazioni

$$\mathbf{H} = \mathbf{G} - \nabla\Omega \quad (1)$$

$$\mathbf{B} = \mathbf{C} + \nabla \times \mathbf{A} \quad (2)$$

dove \mathbf{G} e \mathbf{C} sono campi predefiniti per soddisfare le relazioni

$$\nabla \times \mathbf{G} = \mathbf{J} \quad (3)$$

$$\nabla \cdot \mathbf{C} = \rho_0 \quad (4)$$

con $\rho_0=0$ nel problema magnetostatico, e $\mathbf{J}=\mathbf{0}$ nel problema elettrostatico.

Mostra, poi, come usando alternativamente le relazioni (1) e (2) e le condizioni al contorno, purchè queste rientrino tra quelle dichiarate ammissibili nella dimostrazione di unicità, sia possibile ottenere la separazione del funzionale $\Lambda(\mathbf{H}, \mathbf{B})$ in due coppie duali di funzionali complementari. Si possono così ottenere in principio ben quattro funzionali, ognuno dipendente dalle variabili appartenenti ad uno solo tra i due sistemi \mathbf{H} e \mathbf{B} , e tutti capaci di fornire, attraverso una procedura di estremizzazione, una soluzione valida.

$$\begin{aligned} \Lambda(\mathbf{H}, \mathbf{B}) &= \Theta_p(\mathbf{H}) + \Xi_p(\mathbf{B}) \\ &= \Theta_d(\mathbf{H}) + \Xi_d(\mathbf{B}). \end{aligned} \quad (5)$$

Nella pratica, tuttavia, Rikabi consiglia di scegliere tra le decomposizioni duali dei funzionali quella che risulta più conveniente nel particolare problema in esame.

Rikabi mostra, infine, come, tenendo conto che il minimo di $\Lambda(\mathbf{H}, \mathbf{B})$ è zero, ai funzionali Θ e Ξ sia possibile imporre dei limiti inferiori e superiori, e ottenere una utile stima dell'errore commesso sulla relazione costitutiva: detta, infatti, $(\mathbf{H}^*, \mathbf{B}^*)$ la soluzione esatta, risulta

$$\Lambda(\mathbf{H}^*, \mathbf{B}^*) = \Theta(\mathbf{H}^*) + \Xi(\mathbf{B}^*) = 0 \quad \rightarrow \quad -\Theta(\mathbf{H}^*) = \Xi(\mathbf{B}^*)$$

Se ora si considerano le coppie di campi $(\mathbf{H}^*, \mathbf{B})$ e $(\mathbf{H}, \mathbf{B}^*)$ risulta:

$$\Theta(\mathbf{H}^*) + \Xi(\mathbf{B}) = \Lambda(\mathbf{H}^*, \mathbf{B}) \geq \Lambda(\mathbf{H}^*, \mathbf{B}^*) = \Theta(\mathbf{H}^*) + \Xi(\mathbf{B}^*) \rightarrow \Xi(\mathbf{B}) \geq \Xi(\mathbf{B}^*)$$

$$\Theta(\mathbf{H}) + \Xi(\mathbf{B}^*) = \Lambda(\mathbf{H}, \mathbf{B}^*) \geq \Lambda(\mathbf{H}^*, \mathbf{B}^*) = \Theta(\mathbf{H}^*) + \Xi(\mathbf{B}^*) \rightarrow \Theta(\mathbf{H}) \geq \Theta(\mathbf{H}^*)$$

e pertanto: $-\Theta(\mathbf{H}) \leq -\Theta(\mathbf{H}^*) = \Xi(\mathbf{B}^*) \leq \Xi(\mathbf{B})$

Nell'impostazione di Rikabi rimane il limite di non poter trattare problemi per i quali le condizioni al contorno sono specificate attraverso relazioni tra le variabili del sistema \mathbf{H} e del sistema \mathbf{B} . In questi casi, difatti, non è più possibile la separazione (splitting) del funzionale $\Lambda(\mathbf{H}, \mathbf{B})$. Resta, naturalmente, aperta la possibilità di arrivare ad una soluzione attraverso la diretta minimizzazione di $\Lambda(\mathbf{H}, \mathbf{B})$, anche se, così facendo, nel processoolutivo dovranno purtroppo essere coinvolte contemporaneamente le variabili di entrambi i sistemi

1.3.2) Elettrocinemica quasi-stazionaria

Il modello matematico del problema delle correnti parassite in presenza di mezzi lineari viene definito dalle seguenti equazioni di campo:

$$\begin{array}{ll} \nabla \times \mathbf{H} = \mathbf{J} & \text{in } V \\ \nabla \cdot \mathbf{J} = 0 & \text{in } V \\ \nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t & \text{in } V \\ \nabla \cdot \mathbf{B} = 0 & \text{in } V \end{array}$$

dalle proprietà dei materiali:

$$\begin{array}{l} \mathbf{B} = \mu \mathbf{H} \\ \mathbf{J} = \sigma \mathbf{E} + \mathbf{J}_s \end{array}$$

e da opportune condizioni iniziali, al contorno e sulle interfacce tra materiali differenti.

Anche per questo tipo di problema sono stati introdotti in passato approcciolutivi basati sul metodo di minimizzazione dell'errore costitutivo.

E' Rikabi stesso il primo [29] a tentare di derivare formulazioni complementari per il problema delle correnti parassite; egli, tuttavia, percorrendo una strada parallela a quella seguita per il problema magnetostatico giunge alla conclusione che in questo caso non esiste un principio di minimo. Le specifiche fisiche del problema delle eddy-currents non permettono infatti generalmente di dividere il funzionale di errore (che nel lavoro citato viene soprannominato *Liguriano*) in due funzionali indipendenti funzioni rispettivamente delle sole incognite elettriche e magnetiche. In condizioni di regime sinusoidale di pulsazione ω il liguriano assume la forma:

$$\Lambda(t) = \Lambda^0 + \operatorname{Re}\{\Lambda^c e^{2j\omega t}\} = \Lambda^0 + \Lambda^r \cos 2\omega t - \Lambda^i \sin 2\omega t \quad (1)$$

dove Λ^0 è la parte reale tempo-invariante del Liguriano, che deve soddisfare la diseuguaglianza:

$$\Lambda^0 \geq |\Lambda^c| \geq 0 \quad (2)$$

affinché sia soddisfatta la condizione $\Lambda \geq 0, \forall t$.

Rikabi trova che per la componente complessa a frequenza doppia del liguriano, Λ^c , è possibile arrivare alla separazione delle incognite elettriche e magnetiche:

$$\Lambda^c = \Theta^c(\mathbf{H}) + \Xi^c(\mathbf{E})$$

e pertanto la minimizzazione di Λ^c conduce a principi variazionali complementari

$$\delta \Theta^c(\mathbf{H}) = 0 \quad \text{e} \quad \delta \Xi^c(\mathbf{E}) = 0$$

Tuttavia, essendo i funzionali concavi-convessi, le soluzioni ottenute non costituiscono dei limiti nè per l'energia nè per la potenza. Naturalmente, se la

soluzione ottenuta fosse quella corretta, si avrebbe l'annullamento della parte complessa a frequenza doppia del funzionale di errore:

$$\Lambda^c=0 \quad \Rightarrow \quad \Theta^c(\mathbf{H})=-\Xi^c(\mathbf{E})$$

Viceversa la parte reale, tempo-invariante, del Liguriano, sebbene associata per la (2) con un principio di minimo, non si divide in due funzionali complementari e pertanto non dà origine a principi variazionali complementari.

Nonostante nel caso delle correnti parassite non si riesca ad arrivare allo splitting delle incognite in due sistemi indipendenti e non si possano stabilire limiti di energia, il funzionale liguriano -conclude Rikabi- può ancora essere valutato e attraverso di esso è possibile stabilire facilmente una misura dell'errore commesso nella soluzione numerica.

Albanese e Rubinacci [30] riprendono il lavoro di Rikabi e riescono, attraverso un uso particolareggiato della trasformata di Laplace (questa viene calcolata unicamente sulla parte positiva dell'asse reale) ad ottenere lo splitting delle variabili e a stabilire limiti sui singoli funzionali di errore.

Specificamente assumendo:

$$\begin{aligned} \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} - \nabla \frac{\partial \varphi}{\partial t} & \rightarrow & \quad \mathbf{B} = \nabla \times \mathbf{A} \\ \mathbf{H} &= \mathbf{T} - \nabla \Omega & \rightarrow & \quad \mathbf{J} = \nabla \times \mathbf{T} \end{aligned} \quad (3)$$

(dove \mathbf{A} e \mathbf{T} sono potenziali vettori a due componenti che vengono approssimati attraverso funzioni di forma basate sugli elementi di lato) pervengono alle seguenti espressioni dei funzionali locali di errore associati alle relazioni costitutive elettrica e magnetica:

$$\lambda_m(\hat{\mathbf{B}}, \hat{\mathbf{H}}, p) = \frac{(\hat{\mathbf{B}} - \mu \hat{\mathbf{H}})^2}{2\mu} \quad (4)$$

$$\lambda_e(\hat{\mathbf{J}}, \hat{\mathbf{E}}, p) = \frac{(\hat{\mathbf{J}} - \sigma \hat{\mathbf{E}} - \hat{\mathbf{J}}_s)^2}{2p\sigma}$$

dove si sono indicate col soprassetto le trasformate di Laplace sull'asse reale positivo dei campi e dove p è una quantità reale e positiva.

Integrando nel dominio la somma delle relazioni (4) e sostituendo i campi attraverso le (3) si perviene al funzionale globale di errore Λ a valori reali, che attraverso l'introduzione delle condizioni al contorno si divide nei due funzionali complementari Θ e Ξ :

$$\Lambda \left[\hat{\mathbf{A}} + \nabla \hat{\phi}, \hat{\mathbf{T}} - \nabla \hat{\Omega}, p \right] = \Xi \left[\hat{\mathbf{A}} - \nabla \hat{\phi}, p \right] + \Theta \left[\hat{\mathbf{T}} - \nabla \hat{\Omega}, p \right] \geq 0$$

In questo caso valgono le seguenti disequaglianze attraverso le quali è possibile stabilire dei limiti inferiori e superiori per il valore assunto dai funzionali complementari:

$$\Xi \left[\hat{\mathbf{A}} - \nabla \hat{\phi}, p \right] \geq \Xi \left[-\hat{\mathbf{E}}_0, p \right] \geq -\Theta \left[\hat{\mathbf{H}}_0, p \right] \geq -\Theta \left[\hat{\mathbf{T}} - \nabla \hat{\Omega}, p \right]$$

Nello stesso lavoro viene, tra l'altro, evidenziato che tra le due formulazioni complementari, elettrica e magnetica, alle quali questo approccio conduce, è impossibile individuare quale sia in generale più efficiente: la scelta del metodo solutivo deve cioè essere effettuata portando in conto le particolari specifiche del problema. Il consiglio degli autori è tuttavia quello di risolvere in ogni caso il problema con entrambe le formulazioni, allo scopo da un lato di valutare in maniera affidabile la bontà della soluzione e di avere utili indicazioni circa un

eventuale raffinamento locale della discretizzazione, e dall'altro di disporre di uno strumento affidabile per identificare eventuali errori tra i dati di ingresso del problema.

1.3.3) Propagazione in regime sinusoidale

L'approccio basato sulla minimizzazione dell'errore costitutivo è stato utilizzato anche per problemi propagativi in mezzi lineari e in condizioni di regime sinusoidale[31]. Facendo uso della trasformata di Fourier si perviene al seguente modello del problema:

$$\begin{aligned} \nabla \times \hat{\mathbf{H}}(\mathbf{x}, \omega) &= j \omega \hat{\mathbf{D}}(\mathbf{x}, \omega) & (1a) & \quad \nabla \times \hat{\mathbf{E}}(\mathbf{x}, \omega) = -j \omega \hat{\mathbf{B}}(\mathbf{x}, \omega) & (2a) \\ \mathbf{n}_{\Sigma} \times \hat{\mathbf{H}}(\mathbf{x}, \omega) &= \mathbf{n}_{\Sigma} \times \hat{\mathbf{H}}_S(\mathbf{x}, \omega) \text{ su } \Sigma_H & (1b) & \quad \mathbf{n}_{\Sigma} \times \hat{\mathbf{E}}(\mathbf{x}, \omega) = \mathbf{n}_{\Sigma} \times \hat{\mathbf{E}}_S(\mathbf{x}, \omega) \text{ su } \Sigma_E & (2b) \\ \hat{\mathbf{D}} &= \varepsilon \hat{\mathbf{E}}; & (1c) & \quad \hat{\mathbf{B}} = \mu \hat{\mathbf{H}}; & (2c) \end{aligned}$$

dove $\hat{\mathbf{H}}_S$ e $\hat{\mathbf{E}}_S$ sono campi assegnati rispettivamente sulle frontiere Σ_H e Σ_E , ed ε e μ possono assumere valori complessi con parte immaginaria non-positiva:

$$\mu = \mu' - j\mu''; \quad (3a) \quad \varepsilon = \varepsilon' - j\varepsilon'' \quad (3b)$$

Adoperando come incognite direttamente i campi $\hat{\mathbf{E}}$ e $\hat{\mathbf{H}}$, ed esprimendo i campi $\hat{\mathbf{D}}$ e $\hat{\mathbf{B}}$ attraverso le (1a) e (2a), ci si riduce a determinare i campi $\hat{\mathbf{E}}$ e $\hat{\mathbf{H}}$ che soddisfano le condizioni (1b) e (2b) e rendono stazionario il seguente funzionale di errore a valori complessi:

$$\mathfrak{I}(\hat{\mathbf{E}}, \hat{\mathbf{H}}) = j\omega \left[\int_D \varepsilon^{-1} \left(\hat{\mathbf{D}} - \varepsilon \hat{\mathbf{E}} \right)^2 dV + \int_D \mu^{-1} \left(\hat{\mathbf{B}} - \mu \hat{\mathbf{H}} \right)^2 dV \right] \quad (4)$$

Il funzionale (4) si può decomporre in due funzionali complementari a valori complessi, ed il problema si riduce a trovare separatamente i valori di $\hat{\mathbf{E}}$ e $\hat{\mathbf{H}}$ che li rendono stazionari:

$$\mathfrak{I}^e(\hat{\mathbf{E}}) = \int_D j\omega \varepsilon \hat{\mathbf{E}}^2 dV + \int_D (j\omega \mu)^{-1} \left(\nabla \times \hat{\mathbf{E}} \right)^2 dV - 2 \int_{\partial D} \mathbf{n} \times \hat{\mathbf{H}}_S \cdot \hat{\mathbf{E}} dS \quad (5)$$

$$\mathfrak{J}^h(\hat{\mathbf{H}}) = \int_D j\omega\mu \hat{\mathbf{H}}^2 dV + \int_D (j\omega\varepsilon)^{-1} \left(\nabla \times \hat{\mathbf{H}} \right)^2 dV + 2 \int_{\partial D} \mathbf{n} \times \hat{\mathbf{E}}_s \cdot \hat{\mathbf{H}} dS \quad (6)$$

Il problema così impostato ammette soluzione unica, a meno che $\omega > 0$ non corrisponda ad una frequenza di risonanza. Ciò è d'altra parte escluso se nel mezzo vi sono perdite, ovvero se μ'' e/o ε'' sono maggiori di zero in qualche parte del dominio.

I funzionali (5) e (6) rappresentano a meno di un fattore moltiplicativo, le energie medie attive e reattive immagazzinate nel dominio di integrazione.

Gli autori mostrano come questo approccio accoppiato all'impiego degli edge-elements permetta di evitare la comparsa di modi spuri (con questo termine si indicano i modi associati a soluzioni numeriche per i valori dei campi che non soddisfano a tutti i vincoli imposti dalle equazioni di Maxwell e che risultano pertanto prive di significato fisico). Sottolineano infine la vantaggiosa possibilità di ottenere in modo semplice, attraverso questo approccio, una stima dell'errore numerico locale utile per eventuali successivi raffinamenti della mesh.

CAPITOLO II

L'approccio basato sulla minimizzazione dell'errore costitutivo nell'Elettromagnetismo

2.1) Generalità

La soluzione delle equazioni di Maxwell nella loro forma completa nel dominio del tempo e in geometria tridimensionale è un problema di fondamentale importanza per il quale a tutt'oggi si è lontani dalla disponibilità di risposte pienamente generali e soddisfacenti. Basti dire a tal proposito che le prime metodologie proposte, tutte naturalmente di carattere numerico e, in generale, basate sui metodi FEM, risalgono solo al recente passato[32,33].

Una impostazione corretta del problema elettromagnetico include le due fondamentali equazioni di campo (vale a dire la legge di Faraday e la legge di Ampère), le relazioni costitutive e naturalmente le condizioni iniziali, al contorno e sulle interfacce.

I metodi tradizionali impongono in maniera esatta le relazioni costitutive ed una sola delle due equazioni di campo, accontentandosi di soddisfare approssimativamente l'altra.

L'approccio basato sulla minimizzazione dell'errore costitutivo, qui trattato, adopera una formulazione globale in termini di due potenziali vettori incogniti, l'uno elettrico, l'altro magnetico. La soluzione soddisfa esattamente ambedue le equazioni di campo e viene quindi calcolata imponendo la minimizzazione di un funzionale correlato all'errore sulle relazioni costitutive.

Questo schema provvede due soluzioni complementari, e proprio in questa sua caratteristica, si trovano contemporaneamente il suo più grande vantaggio

e il maggiore inconveniente. Da un lato, infatti, le soluzioni complementari forniscono la stima degli errori locali e globali associati alla soluzione numerica trovata, informazioni queste che possono essere vantaggiosamente utilizzate non solo ai fini della verifica di validità dei risultati conseguiti, ma anche nell'eventuale fase di ampliamento dello spazio funzionale in cui si cerca la soluzione (*raffinamento della discretizzazione*). D'altra parte l'introduzione di due potenziali vettori comporta un forte aumento del carico computazionale (sei incognite scalari rispetto alle tre delle formulazioni tradizionali).

E' stato già osservato che il metodo è stato introdotto per trattare casi statici e quasi-statici. La formulazione presentata in questa tesi generalizza il metodo di modo da consentirne l'utilizzazione per una più vasta classe di problemi elettromagnetici che comprende quelli nei quali le correnti di spostamento non possono essere trascurate. Più precisamente, almeno dal punto di vista concettuale, il modello sembra adeguato per il trattamento di ogni problema tridimensionale elettromagnetico che sia caratterizzato da relazioni costitutive eventualmente non-lineari, anisotrope, non-omogenee e tempo-varianti.

2.2) Sviluppo analitico dell'approccio basato sulla minimizzazione dell'errore costitutivo per problemi di campo 3D nel dominio del tempo

2.2.1) Modello matematico

Un qualsiasi problema di elettromagnetismo definito nel dominio spaziale V e nell'intervallo temporale $[0, T]$ può essere formulato utilizzando le equazioni di Maxwell che con le usuali notazioni, si presentano come:

$$\nabla \cdot \mathbf{D}(\mathbf{x},t) = \rho(\mathbf{x},t) \quad (1a)$$

$$\nabla \cdot \mathbf{B}(\mathbf{x},t) = 0 \quad (1b)$$

$$\nabla \times \mathbf{H}(\mathbf{x},t) = \mathbf{J}(\mathbf{x},t) + \partial \mathbf{D}(\mathbf{x},t) / \partial t \quad (1c)$$

$$\nabla \times \mathbf{E}(\mathbf{x},t) = -\partial \mathbf{B}(\mathbf{x},t) / \partial t \quad (1d)$$

$$\nabla \cdot \mathbf{J}(\mathbf{x},t) = -\partial \rho(\mathbf{x},t) / \partial t \quad (1e)$$

valide $\forall \mathbf{x} \in V$ e $\forall t \in [0, T]$.

Alle espressioni (1) si accompagnano le relazioni di interfaccia sulle eventuali superfici di separazione Σ tra materiali diversi

$$[\mathbf{D}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = \rho_\Sigma(\mathbf{x},t) \quad (2a)$$

$$[\mathbf{B}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (2b)$$

$$[\mathbf{E}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (2c)$$

$$[\mathbf{H}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = \mathbf{J}_\Sigma(\mathbf{x},t) \quad (2d)$$

$$[\mathbf{J}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = -\partial \rho_\Sigma(\mathbf{x},t) / \partial t \quad (2e)$$

che devono essere soddisfatte $\forall \mathbf{x} \in \Sigma$ e $\forall t \in [0, T]$ e dove con i termini ρ_Σ e \mathbf{J}_Σ si sono indicati i termini sorgente che sono rispettivamente densità superficiali di carica e di corrente.

Perchè la soluzione del problema elettromagnetico sia unica è necessario aggiungere $\forall \mathbf{x} \in V$ alle equazioni di Maxwell le condizioni iniziali:

$$\mathbf{D}(\mathbf{x},0)=\mathbf{D}_0(\mathbf{x}), \quad (3a)$$

$$\mathbf{B}(\mathbf{x},0)=\mathbf{B}_0(\mathbf{x}) \quad (3b)$$

le relazioni costitutive:

$$\mathbf{B}(\mathbf{x},t)=\mathbf{f}_B(\mathbf{H},\mathbf{x},t), \quad (4a)$$

$$\mathbf{D}(\mathbf{x},t)=\mathbf{f}_D(\mathbf{E},\mathbf{x},t), \quad (4b)$$

$$\mathbf{J}(\mathbf{x},t)=\mathbf{f}_J(\mathbf{E},\mathbf{x},t) \quad (4c)$$

e, indicata con ∂V la frontiera di V , $\forall \mathbf{x} \in \partial V$ e $\forall t \in [0, T]$ opportune condizioni al contorno.

Ricordiamo, per inciso, che le relazioni (1a), (1c) ed (1e) non sono indipendenti potendosi ad esempio ricavare la (1e) dalle altre.

Assumiamo ora le seguenti ipotesi:

- I) Le relazioni costitutive (4), anche se eventualmente non-lineari, anisotrope, non-omogenee e tempo-varianti, siano comunque non dispersive nel tempo e nello spazio. In tal caso le (4a,b,c) diventano:

$$\mathbf{B}(\mathbf{x},t)=\mathbf{f}_B(\mathbf{H}(\mathbf{x},t),\mathbf{x},t), \quad (4a)$$

$$\mathbf{D}(\mathbf{x},t)=\mathbf{f}_D(\mathbf{E}(\mathbf{x},t),\mathbf{x},t), \quad (4b)$$

$$\mathbf{J}(\mathbf{x},t)=\mathbf{f}_J(\mathbf{E}(\mathbf{x},t),\mathbf{x},t) \quad (4c)$$

- II) Le relazioni costitutive (4) siano monotone. Con riferimento alla (4a) questo corrisponde a imporre che prese due coppie di campi $(\mathbf{H}_1, \mathbf{B}_1)$ e $(\mathbf{H}_2, \mathbf{B}_2)$ che soddisfano la relazione costitutiva magnetica sia valida la disuguaglianza:

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot (\mathbf{H}_2 - \mathbf{H}_1) \geq 0 \quad (4d)$$

analogamente per le coppie di campi $(\mathbf{E}_1, \mathbf{D}_1)$ e $(\mathbf{E}_2, \mathbf{D}_2)$ che soddisfano la (4b) e per le coppie $(\mathbf{E}_1, \mathbf{J}_1)$ e $(\mathbf{E}_2, \mathbf{J}_2)$ che soddisfano la (4c) devono valere le relazioni:

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot (\mathbf{E}_2 - \mathbf{E}_1) \geq 0 \quad (4e)$$

$$(\mathbf{J}_2 - \mathbf{J}_1) \cdot (\mathbf{E}_2 - \mathbf{E}_1) \geq 0 \quad (4f)$$

Si ponga:

$$\mathbf{D}_T(\mathbf{x}, t) = \mathbf{D}_{T_0}(\mathbf{x}) + \int_0^t (\mathbf{J}(\mathbf{x}, t) + \partial \mathbf{D}(\mathbf{x}, t) / \partial t) dt \quad (5)$$

Per il teorema di Helmholtz si può scrivere

$$\mathbf{D}_T(\mathbf{x}, t) = \mathbf{D}_T^{(irr)}(\mathbf{x}, t) + \mathbf{D}_T^{(sol)}(\mathbf{x}, t)$$

con la posizione:

$$\nabla \cdot \mathbf{D}_T^{(irr)}(\mathbf{x}) = \nabla \cdot \mathbf{D}_{T_0}^{(irr)}(\mathbf{x}) = \rho_{T_0}(\mathbf{x}) \quad (6)$$

ove al termine \mathbf{D}_T si può assegnare il nome di "*Vettore di spostamento totale*" per sottolinearne la sua capacità di generalizzare l'usuale "*Vettore di spostamento*" \mathbf{D} . La densità di carica $\rho_{T_0}(\mathbf{x})$ rappresenta la sorgente del campo di background $\mathbf{D}_T^{(irr)}$, un campo cioè che non modifica nè è modificato da alcuno dei fenomeni elettromagnetici che si verificano dopo l'istante $t=0$. E' dunque lecito assumere quale condizione iniziale per la densità di carica per $\rho_{T_0}(\mathbf{x})$ una funzione arbitraria e in particolare $\rho_{T_0}(\mathbf{x})=0$. Giacchè anche l'eventuale componente solenoidale di \mathbf{D}_{T_0} non condiziona l'evoluzione dei successivi fenomeni elettromagnetici si può porre senz'altro analogamente $\mathbf{D}_{T_0}(\mathbf{x})=0$.

Le equazioni di Maxwell si possono allora riscrivere nella forma seguente:

$$\nabla \cdot \mathbf{D}_T = 0 \quad (7a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (7b)$$

$$\nabla \times \mathbf{H} = \partial \mathbf{D}_T / \partial t \quad (7c)$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t \quad (7d)$$

e le condizioni di continuità sulle superfici di interfaccia diventano:

$$[\mathbf{D}_T(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (8a)$$

$$[\mathbf{B}(\mathbf{x},t)] \cdot \mathbf{n}_\Sigma = 0 \quad (8b)$$

$$[\mathbf{E}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (8c)$$

$$[\mathbf{H}(\mathbf{x},t)] \times \mathbf{n}_\Sigma = 0 \quad (8d)$$

Le relazioni costitutive (4a,b,c) si possono condensare nelle seguenti espressioni:

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t), \mathbf{x}, t), \quad (9a)$$

$$\mathbf{D}_T(\mathbf{x},t) = \mathbf{D}_{T_0}(\mathbf{x}) - \mathbf{D}_0(\mathbf{x}) + \mathbf{f}_D(\mathbf{E}(\mathbf{x},t), \mathbf{x}, t) + \int_0^t \mathbf{f}_J(\mathbf{E}(\mathbf{x},t'), \mathbf{x}, t') dt' \quad (9b)$$

Alle condizioni iniziali (3) bisogna, infine aggiungere quella su \mathbf{D}_T :

$$\mathbf{D}(\mathbf{x},0) = \mathbf{D}_0(\mathbf{x}), \quad (3a)$$

$$\mathbf{B}(\mathbf{x},0) = \mathbf{B}_0(\mathbf{x}) \quad (3b)$$

$$\mathbf{D}_T(\mathbf{x},0) = \mathbf{D}_{T_0}(\mathbf{x}), \quad (3c)$$

dove, come già osservato, $\mathbf{D}_{T_0}(\mathbf{x})$ può arbitrariamente essere scelto nullo.

Se ora definiamo due potenziali vettori nel seguente modo:

$$\mathbf{A} = - \int_0^t \mathbf{E} dt \quad (10a)$$

$$\mathbf{W} = \int_0^t \mathbf{H} dt \quad (10b)$$

il che comporta:

$$\mathbf{E} = -\partial\mathbf{A}/\partial t \quad (11a)$$

$$\mathbf{H} = \partial\mathbf{W}/\partial t \quad (11b)$$

e poniamo:

$$\mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A} \quad (11c)$$

$$\mathbf{D}_T = \mathbf{D}_{T0} + \nabla \times \mathbf{W} \quad (11d)$$

si perviene alle formulazioni duali:

$$\begin{aligned} \nabla \times \mathbf{H} &= \partial \mathbf{D}_T / \partial t \\ \Downarrow \\ \mathbf{H} &= \partial \mathbf{W} / \partial t, \quad \mathbf{D}_T = \mathbf{D}_{T0} + \nabla \times \mathbf{W} \end{aligned}$$

$$\begin{aligned} \nabla \times \mathbf{E} &= -\partial \mathbf{B} / \partial t \\ \Downarrow \\ \mathbf{E} &= -\partial \mathbf{A} / \partial t, \quad \mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A} \end{aligned}$$

$$\nabla \times \mathbf{H} - \partial \mathbf{D}_T / \partial t = 0, \quad \forall \mathbf{W}$$

$$\nabla \times \mathbf{E} + \partial \mathbf{B} / \partial t = 0, \quad \forall \mathbf{A}$$

nelle quali, per ogni arbitraria scelta dei potenziali \mathbf{A} e \mathbf{W} , le equazioni di Maxwell (7) risultano automaticamente soddisfatte.

2.2.2) Definizione del funzionale di errore.

Si procede a questo punto alla definizione di un funzionale di errore locale $\lambda(\mathbf{x},t) \geq 0$ che descriva la misura con la quale le relazioni costitutive sono soddisfatte; esso dovrà essere nullo solo nei punti spazio-temporali dove le relazioni costitutive (2.2.1-9) sono rigorosamente rispettate.

Una tra le definizioni ammissibili potrebbe essere:

$$\lambda(\mathbf{x},t) = \alpha_H [\mathbf{B}(\mathbf{x},t) - \mathbf{f}_B(\mathbf{H}(\mathbf{x},t),t)]^2 + \alpha_E [\mathbf{D}_T(\mathbf{x},t) - \mathbf{D}_{T0}(\mathbf{x}) - \mathbf{f}_D(\mathbf{E}(\mathbf{x},t),\mathbf{x},t) + \mathbf{D}_0(\mathbf{x}) - \int_0^t \mathbf{f}_J(\mathbf{E}(\mathbf{x},t),t) dt]^2 \quad (1)$$

dove α_H e α_E sono opportuni coefficienti di peso sulle relazioni costitutive magnetiche ed elettriche, che possono essere, in generale, funzione del tempo e che possono assumere eventualmente nello stesso problema valori diversi per differenti materiali.

In questa tesi tuttavia si è preferito usare la seguente forma:

$$\lambda(\mathbf{x},t) = \alpha_H \left[\int_{\mathbf{H}^*(\mathbf{x},t)}^{\mathbf{H}(\mathbf{x},t)} \mathbf{f}_B(\mathbf{h},\mathbf{x},t) \cdot d\mathbf{h} + \int_{\mathbf{B}^*(\mathbf{x},t)}^{\mathbf{B}(\mathbf{x},t)} \mathbf{f}_H(\mathbf{b},\mathbf{x},t) \cdot d\mathbf{b} + \mathbf{H}^*(\mathbf{x},t) \cdot \mathbf{B}^*(\mathbf{x},t) - \mathbf{H}(\mathbf{x},t) \cdot \mathbf{B}(\mathbf{x},t) \right] + \alpha_E \left[\int_{\mathbf{D}_T^*(\mathbf{x},t)}^{\mathbf{D}_T(\mathbf{x},t)} \mathbf{f}_E(\mathbf{d}_T,\mathbf{x},t) \cdot d\mathbf{d}_T + \int_{\mathbf{E}^*(\mathbf{x},t)}^{\mathbf{E}(\mathbf{x},t)} \mathbf{f}_{D_T}(\mathbf{e},\mathbf{x},t) \cdot d\mathbf{e} + \mathbf{E}^*(\mathbf{x},t) \cdot \mathbf{D}_T^*(\mathbf{x},t) - \mathbf{E}(\mathbf{x},t) \cdot \mathbf{D}_T(\mathbf{x},t) \right] \quad \forall t \in (t_k, t_{k+1}) \quad (2)$$

dove $\mathbf{f}_H(\mathbf{B},\mathbf{x},t)$ è la caratteristica magnetica inversa che fornisce \mathbf{H} rispetto a \mathbf{B} , per \mathbf{x} e t fissati; $(\mathbf{H}^*, \mathbf{B}^*)$ è una coppia che soddisfa la relazione costitutiva magnetica (2.2.1-9a); $\mathbf{f}_{D_T}(\mathbf{E},\mathbf{x},t)$ è il mapping non-lineare di $\mathbf{E}(\mathbf{x},t)$ su $\mathbf{D}_T(\mathbf{x},t)$

definito dalla relazione (2.2.1-9b); $f_E(\mathbf{D}_T, \mathbf{x}, t)$ è la sua inversa rispetto a $\mathbf{E}(\mathbf{x}, t)$ ed, infine, $(\mathbf{E}^*, \mathbf{D}_T^*)$ è una coppia che soddisfa la relazione costitutiva elettrica (2.2.1-9b). Si noti che la relazione (2.2.1-9b) è dispersiva nel dominio del tempo, per cui un legame di tipo diretto tra i valori istantanei di $\mathbf{E}(\mathbf{x}, t)$ e $\mathbf{D}_T(\mathbf{x}, t)$ è possibile solo per casi particolari quali ad esempio quello in cui i potenziali \mathbf{A} e \mathbf{W} sono scelti a dinamica lineare in (t_k, t_{k+1}) . Ad esempio nel caso particolare di condizioni iniziali nulle, di assenza di sorgenti nel dominio e di linearità, isotropia, tempo-invarianza dei mezzi si ha:

$$\mathbf{B} = \mu \mathbf{H} \quad (3a)$$

$$\mathbf{J} = \sigma \mathbf{E} \quad (3b)$$

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (3c)$$

da cui, assunto $\mathbf{D}_{T_0}(\mathbf{x}) = 0$, si ricava:

$$f_H(\mathbf{B}, \mathbf{x}, t) = \mathbf{B} / \mu \quad (4a)$$

$$f_B(\mathbf{H}, \mathbf{x}, t) = \mu \mathbf{H} \quad (4b)$$

$$f_{D_T}(\mathbf{E}, \mathbf{x}, t) = [(t - t_K)\sigma + \varepsilon] \mathbf{E} - \sigma \mathbf{A}(t_K) \quad (4c)$$

$$f_E(\mathbf{D}_T, \mathbf{x}, t) = [\mathbf{D}_T + \sigma \mathbf{A}(t_K)] / [(t - t_K)\sigma + \varepsilon] \quad (4d)$$

Come coppie di campi che soddisfano rispettivamente le relazioni costitutive elettrica e magnetica si possono scegliere:

$$\mathbf{E}^* = 0$$

$$\mathbf{D}_T^* = -\sigma \mathbf{A}(t_K)$$

$$\mathbf{H}^* = 0$$

$$\mathbf{B}^* = 0$$

In questo modo l'errore locale diventa:

$$\lambda(\mathbf{x}, t) = \alpha_H \{ [\mu \mathbf{H}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, t)] \cdot [\mathbf{H}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, t) / \mu] \} / 2 +$$

$$\alpha_E \{ [\mathbf{D}_T(\mathbf{x}, t) - \mathbf{f}_{D_T}(\mathbf{E}(\mathbf{x}, t), \mathbf{x}, t)] \cdot [\mathbf{E}(\mathbf{x}, t) - \mathbf{f}_E(\mathbf{D}_T(\mathbf{x}, t), \mathbf{x}, t)] \} \quad (5)$$

$$\forall t \in (t_k, t_{k+1})$$

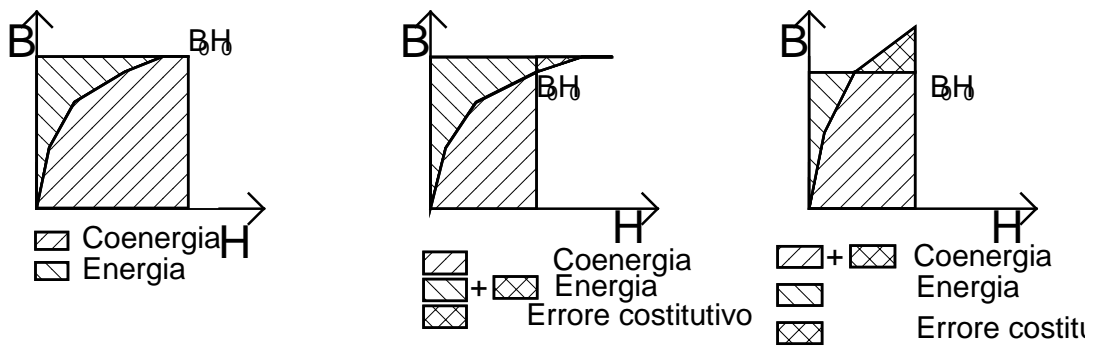


Fig.[1] Il funzionale di errore locale definito dalla (2) per relazioni costitutive crescenti è sempre positivo e vale 0 solo se la relazione costitutiva è verificata esattamente.

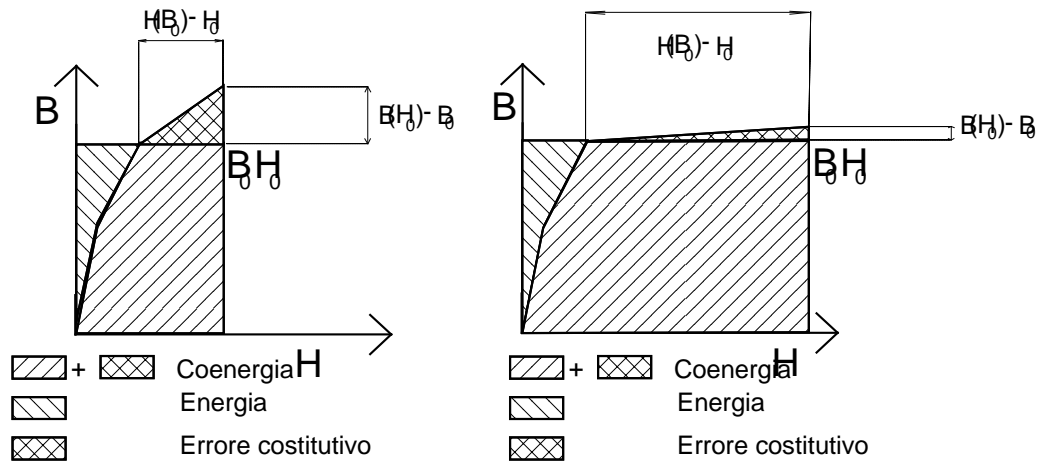


Fig.[2] In questo caso le relazioni (1)

Fig.[3] Situazione patologica per la

e (2) sono ambedue accettabili.

relazione (1), ma non per la (2)

Il motivo della scelta sulla definizione dell'errore locale appare evidente se si fa riferimento al diagramma della relazione costitutiva B-H riportato in figura (1). Sulla stessa figura è stato evidenziato il significato geometrico degli errori associati alle relazioni (1) e (2). Si osservi che l'errore associato alla (1) è una distanza, mentre quello associato alla relazione (2) rappresenta fisicamente un'area. Questa scelta permette da un lato di evitare situazioni patologiche (vedi figura (3)), e dall'altro di bilanciare equamente il peso dell'errore sulle incognite. Si può notare a questo proposito l'inerente simmetria della (2) rispetto alle soluzioni nelle variabili \mathbf{A} e \mathbf{W} .

Una volta definito un valore conveniente per i coefficienti di peso, il problema si riduce alla ricerca dei due potenziali vettori $\mathbf{A}(\mathbf{x},t)$ e $\mathbf{W}(\mathbf{x},t)$ capaci di minimizzare il funzionale di errore globale Λ definito come

$$\Lambda = \int_0^T \int_V \lambda(\mathbf{x},t) dVdt \quad (6)$$

dove V è la regione spaziale di integrazione e $(0,T)$ l'intervallo temporale di interesse.

2.2.3) Espressioni duali dei campi e delle grandezze integrali

La formulazione basata sull'errore costitutivo mette a disposizione due diversi modi per calcolare i campi e le più rilevanti grandezze integrali di interesse elettromagnetico: da una parte l'energia elettrica e l'energia magnetica che istante per istante sono immagazzinate nel sistema, e dall'altra la potenza ohmica dissipata. Inoltre, attraverso una riformulazione del vettore di Poynting nelle variabili di campo introdotte, possono essere stimate le variazioni di energia del sistema dovute a scambi energetici col mondo esterno: oppure, come si vedrà in seguito, a dissipazione numerica.

Ogni variabile di campo può essere rappresentata in due maniere distinte ognuna delle quali basata su uno solo dei potenziali vettori (Vedi tab.[1])

Campo	Espressione in \mathbf{A}	Espressione in \mathbf{W}
$\mathbf{B}(\mathbf{x},t)$	$\mathbf{B}_o + \nabla \times \mathbf{A}$	$\mathbf{f}_B(\partial \mathbf{W} / \partial t, \mathbf{x}, t)$
$\mathbf{D}_T(\mathbf{x},t)$	$\mathbf{f}_{D_T}(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{D}_{T_0} + \nabla \times \mathbf{W}$
$\mathbf{H}(\mathbf{x},t)$	$\mathbf{f}_H(\mathbf{B}_o + \nabla \times \mathbf{A}, \mathbf{x}, t)$	$\partial \mathbf{W} / \partial t$
$\mathbf{E}(\mathbf{x},t)$	$-\partial \mathbf{A} / \partial t$	$\mathbf{f}_{E-D_T}(\mathbf{D}_{T_0} + \nabla \times \mathbf{W}, \mathbf{x}, t)$
$\mathbf{D}(\mathbf{x},t)$	$\mathbf{f}_D(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{f}_D[\mathbf{f}_{E-D_T}(\mathbf{D}_{T_0} + \nabla \times \mathbf{W}, \mathbf{x}, t), \mathbf{x}, t]$
$\mathbf{J}(\mathbf{x},t)$	$\mathbf{f}_J(-\partial \mathbf{A} / \partial t, \mathbf{x}, t)$	$\mathbf{f}_J[\mathbf{f}_{E-D_T}(\mathbf{D}_{T_0} + \nabla \times \mathbf{W}, \mathbf{x}, t), \mathbf{x}, t]$

Tabella[1] Espressioni duali dei campi

dove con \mathbf{f}_{D_T} si è indicato la relazione (2.2.1-9b) e con \mathbf{f}_{E-D_T} si è indicato la sua inversa.

Sfruttando le varie espressioni in tabella è semplice dedurre le espressioni per le varie grandezze integrali.

ENERGIA ELETTRICA

Se per il mezzo vale la relazione costitutiva non lineare

$$\mathbf{D}(\mathbf{x},t) = \mathbf{f}_D(\mathbf{E}(\mathbf{x},t), \mathbf{x}, t)$$

l'energia elettrica immagazzinata all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni::

$$W_{El,A} = \frac{1}{2} \iiint_V \left[\mathbf{f}_D \left[-\dot{\mathbf{A}}(t_k + dt_k) \right] \int_0^{\mathbf{D}'} \mathbf{f}_E(\mathbf{D}') \cdot d\mathbf{D}' \right] dV$$

$$W_{El,W} = \frac{1}{2} \iiint_V \left[\mathbf{f}_D \left[\mathbf{f}_{E-D_T}(\mathbf{D}_{T_0} + \nabla \times \mathbf{W}(t_{k+1})) \right] \int_0^{\mathbf{D}'} \mathbf{f}_E(\mathbf{D}') \cdot d\mathbf{D}' \right] dV$$

ENERGIA MAGNETICA

Se per il mezzo vale la relazione costitutiva non lineare

$$\mathbf{B}(\mathbf{x},t) = \mathbf{f}_B(\mathbf{H}(\mathbf{x},t),\mathbf{x},t)$$

l'energia magnetica immagazzinata all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni (Fig.[4]):

$$W_{\text{Mag,A}} = \frac{1}{2} \iiint_V \left[\int_0^{\mathbf{B}_0 + \nabla \times \mathbf{A}(t_{k+1})} \mathbf{f}_H(\mathbf{B}') \cdot d\mathbf{B}' \right] dV$$

$$W_{\text{Mag,W}} = \frac{1}{2} \iiint_V \left[\int_0^{f_B[\dot{\mathbf{W}}(tk+dtk)]} \mathbf{f}_H(\mathbf{B}') \cdot d\mathbf{B}' \right] dV$$

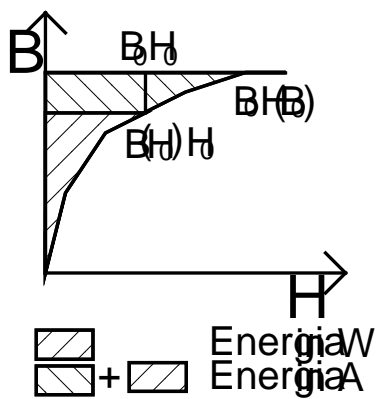


Fig.[4]: Le due diverse espressioni della energia magnetica specifica.

POTENZA ELETTRICA

Si assuma che per il mezzo valga la relazione costitutiva, eventualmente non lineare:

$$\mathbf{J}(\mathbf{x},t) = \mathbf{f}_J(\mathbf{E}(\mathbf{x},t),\mathbf{x},t)$$

La potenza elettrica assorbita all'istante t nel volume V può essere stimata attraverso una delle seguenti espressioni:

$$P_{Ei,A}(t) = \frac{1}{2} \iiint_V \mathbf{f}_J(-\dot{\mathbf{A}}) \cdot (-\dot{\mathbf{A}}) dV$$

$$P_{Ei,W}(t) = \frac{1}{2} \iiint_V \mathbf{f}_J[\mathbf{f}_{E-D_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W})] \cdot \mathbf{f}_{E-D_T}(\mathbf{D}_{T0} + \nabla \times \mathbf{W}) dV$$

Se in V la corrente di spostamento è trascurabile rispetto alla corrente di conduzione, allora $\mathbf{J} = \partial \mathbf{D}_T / \partial t$ e vale la seguente espressione:

$$P_{Ei,A-W}(t) = \frac{1}{2} \iiint_V (-\dot{\mathbf{A}}) \cdot (\nabla \times \dot{\mathbf{W}}) dV$$

VETTORE DI POYNTING

Attraverso il teorema di Poynting si perviene all'espressione del bilancio elettromagnetico di potenza per il volume V :

$$\oint_{\partial V} \mathbf{S} \cdot n d\Sigma = -P_{Ei} - \frac{d}{dt} (W_{Mag} + W_{Ei}) \quad (1)$$

dove $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ è il vettore di Poynting ed è stata scelta l'orientazione uscente dal volume per la normale n alla frontiera ∂V .

Può essere interessante ripercorrere i passaggi che conducono alla (1) partendo dalle equazioni di Maxwell scritte attraverso la (7) allo scopo di appurare se il teorema di Poynting debba continuare a valere anche per le soluzioni approssimate per via numerica attraverso il metodo basato sulla minimizzazione dell'errore.

Si moltiplichino attraverso il prodotto scalare entrambi i membri della (7d) per \mathbf{H} e della (7c) per \mathbf{E} ; sottraendo la seconda espressione dalla prima si ricava

$$\mathbf{H} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{H} = -\mathbf{E} \cdot \partial \mathbf{D}_T / \partial t - \mathbf{H} \cdot \partial \mathbf{B} / \partial t \quad (2)$$

Applicando una ben nota identità vettoriale, integrando nel volume V ed applicando il teorema della divergenza al primo membro la (2) diventa:

$$\int_{\partial V} (\mathbf{E} \times \mathbf{H}) \cdot \hat{\mathbf{n}} d\Sigma = - \int_V (\mathbf{E} \cdot \dot{\mathbf{D}}_T + \mathbf{H} \cdot \dot{\mathbf{B}}) dV \quad (3)$$

Si noti ora che poichè si è giunti a questa espressione adoperando esclusivamente le equazioni di Maxwell e queste sono automaticamente verificate quando si adoperi il metodo basato sulla minimizzazione dell'errore costitutivo, il teorema di Poynting (3) è ancora vero per le soluzioni numeriche. Questo è vero, beninteso, solo se ci si serve delle espressioni (11a) e (11c) in termini del potenziale vettore \mathbf{A} per rappresentare i campi \mathbf{E} e \mathbf{B} e delle espressioni (11b) e (11d) in termini del potenziale vettore \mathbf{W} per rappresentare i campi \mathbf{H} e \mathbf{D}_T . Si noti difatti che le relazioni costitutive (9a) e (9b) non sono soddisfatte esattamente e perciò in generale:

$$\int_V \mathbf{H} \cdot \dot{\mathbf{B}} dV \neq \int_V \mathbf{f}_H(\mathbf{B}) \cdot \dot{\mathbf{B}} dV = \dot{W}_{Mag} \quad (9a)$$

$$\int_V \mathbf{E} \cdot \dot{\mathbf{D}}_T dV \neq \int_V \mathbf{f}_{E-D_T}(\mathbf{D}_T) \cdot \dot{\mathbf{D}}_T dV = \dot{W}_{El} + P_{El}$$

dove con $\mathbf{f}_H(\mathbf{B})$ e $\mathbf{f}_{E-D_T}(\mathbf{D}_T)$ si sono indicate rispettivamente le relazioni inverse della (9a) e della (9b).

Appare dunque chiaro che se si sostituiscono nella (1) le espressioni dell'energia elettrica, dell'energia magnetica e della potenza elettrica ricavate nel paragrafo precedente il secondo membro risulterà in generale differente dal primo e quindi il teorema di Poynting non è soddisfatto. Questo equivale a dire che per le soluzioni numeriche ottenute attraverso l'approccio qui discusso non vale il principio di conservazione dell'energia. Si può giustificare in questo modo la presenza di fenomeni di dissipazione o amplificazione numerica nella simulazione numerica di problemi elettromagnetici di tipo dinamico.

2.3) L'approccio basato sulla minimizzazione dell'errore costitutivo per i circuiti lineari a parametri concentrati.

L'approccio basato sulla minimizzazione dell'errore costitutivo può essere adoperato anche per il trattamento dei circuiti a parametri concentrati. La cosa appare d'altronde ovvia se si osserva che la teoria delle reti consegue in modo naturale dalle eq. di Maxwell, allorchè siano verificate le ipotesi di funzionamento quasi-stazionario. In tal caso è possibile riscrivere in forma integrale le caratteristiche dei materiali ed associarle a zone di singolarità spaziale: i componenti elettrici. Se si ridefiniscono anche i campi in forma integrale -evitando attentamente di includere nel dominio di integrazione le zone di singolarità spaziali- è possibile descrivere una estesa classe di fenomeni elettromagnetici attraverso modelli *zero-dimensionali*, ovvero attraverso modelli nei quali sono totalmente assenti le variabili spaziali. Le incognite integrali, per le quali il problema così riformulato deve essere invertito, sono le tensioni e le correnti.

In questo paragrafo, viene definito un funzionale di errore associato alle caratteristiche dei componenti di una generica rete lineare priva di generatori dipendenti e tale da risultare strettamente positivo se le caratteristiche non sono esattamente soddisfatte e nullo in caso contrario.

Si consideri una generica rete lineare composta eventualmente da generatori indipendenti di corrente e di tensione, resistori, condensatori ed induttori (eventualmente accoppiati), ma priva di generatori dipendenti.

La rete sia composta di l lati e n nodi. Si scelgano i riferimenti per le tensioni e le correnti associate ai lati della rete e si costruiscano la matrice di incidenza \mathbf{A} e la matrice degli anelli \mathbf{M} . [34] Siano \mathbf{v}^l ed \mathbf{i}^l funzioni vettoriali i cui elementi sono gli andamenti temporali rispettivamente delle tensioni e delle correnti di ciascun lato della rete.

Assunto un qualunque nodo come riferimento, si associ la funzione $e_j(t)$ a ciascuno dei nodi rimanenti (potenziali di nodo). Sia \mathbf{e} il vettore dei potenziali di nodo composto dagli $n-1$ elementi $\{e_j(t)\}$. Per ogni possibile valore di \mathbf{e} l'equazione di Kirchhoff alle tensioni risulta automaticamente soddisfatta. Il vettore delle tensioni di lato può essere espressa attraverso la matrice di incidenza e il vettore dei potenziali di nodo come:

$$\mathbf{v}^l = \mathbf{A}^T \mathbf{e} \quad (1)$$

In maniera duale, individuato un albero della rete in cui non compaiano lati su cui si trovano generatori indipendenti di corrente, si costruisca il vettore delle correnti di maglia \mathbf{i} associando la funzione $i_j(t)$ [$j= 1, l-(n-1)$] all' i -mo lato del coalbero. In questo caso per ogni possibile scelta di \mathbf{i} le equazioni di Kirchhoff alle correnti sono soddisfatte. Il vettore delle correnti di lato può essere espressa attraverso la matrice degli anelli e il vettore delle correnti di maglia come:

$$\mathbf{i}^l = \mathbf{M}^T \mathbf{i} \quad (2)$$

Utilizzando le variabili \mathbf{e} ed \mathbf{i} come incognite, attraverso l'integrazione temporale dell'errore sull'equazione caratteristica è possibile definire per ogni lato del circuito un funzionale locale di errore. Sommando i contributi dei funzionali locali ed integrando nel tempo si giunge alla definizione del funzionale globale di errore. La soluzione della rete può essere trovata attraverso la minimizzazione del funzionale globale di errore.

Nella definizione delle espressioni dei funzionali locali di errore si supporrà che, per ciascun lato k della rete, l'equazione caratteristica sia compresa nel seguente elenco:

$$v_k^l(t) = R_k i_k^l(t) \quad (\text{resistenza}) \quad (3)$$

$$v_k^l(t) = \sum_{j=1}^l L_{kj} \frac{di_j^l(t)}{dt}; \quad i_k^l(0)=i_{k0} \quad (\text{induttore accoppiato}) \quad (4)$$

$$i_k^l(t) = C_k \frac{dv_k^l(t)}{dt}; \quad v_k^l(0) = v_{k0}^l \quad (\text{condensatore}) \quad (5)$$

$$i_k^l(t) = i_k^*(t) \quad (\text{generatore ind. di corrente}) \quad (6)$$

$$e_k^l(t) = e_k^*(t) \quad (\text{generatore ind. di tensione}) \quad (7)$$

Le relazioni (6) e (7) possono essere imposte direttamente come vincoli sulle variabili. Per il generico lato k soggetto rispettivamente alle condizioni (6) o (7) si ha pertanto: :

$$(\mathbf{M}^T \mathbf{i})_k = i_k^*(t) \quad (8)$$

$$(\mathbf{A}^T \mathbf{e})_k = e_k^*(t) \quad (9)$$

Si noti che, così facendo, le condizioni (8) e (9) sono analoghe alle condizioni al contorno che devono essere imposte nel problema di campo e, dunque, i lati su cui si trovano i generatori indipendenti potrebbero essere visti come i "*lati di frontiera*" della rete.

Per il generico lato k, una possibile definizione dei funzionali locali di errore associati alle relazioni (3), (4) e (5) che soddisfa tutte le condizioni richieste è la seguente:

$$\begin{aligned} \lambda_{R,k}(t) &= \frac{1}{2} [v_k^l(t) - R_k i_k^l(t)]^T \cdot [(R_k)^{-1} v_k^l(t) - i_k^l(t)] = \\ &= \frac{1}{2} [(\mathbf{A}^T \mathbf{e})_k - R_k (\mathbf{M}^T \mathbf{i})_k]^T \cdot [(R_k)^{-1} (\mathbf{A}^T \mathbf{e})_k - (\mathbf{M}^T \mathbf{i})_k]; \quad (10) \end{aligned}$$

$$\begin{aligned} \lambda_{L,k}(t) &= \frac{1}{2} \left[v_k^l(t) - \sum_{j=1}^1 L_{kj} \frac{di_j^l(t)}{dt} \right]^T \cdot \left[\sum_{j=1}^1 \Gamma_{kj} \int_0^t v_j^l(t') dt' - i_k^l(t) + i_{k0}^l \right] = \\ &= \frac{1}{2} \left[(\mathbf{A}^T \mathbf{e})_k - \sum_{j=1}^1 L_{kj} \frac{d(\mathbf{M}^T \mathbf{i})_j}{dt} \right]^T \cdot \left[\sum_{j=1}^1 \Gamma_{kj} \int_0^t (\mathbf{A}^T \mathbf{e})_j dt' - (\mathbf{M}^T \mathbf{i})_k + i_{k0}^l \right]; \quad (11) \end{aligned}$$

$$\lambda_{C,k}(t) = \frac{1}{2} \left[v_k^l(t) - v_{k0}^l - (C_k)^{-1} \int_0^t i_k^l(t) dt' \right]^T \cdot \left[C_k \frac{dv_k^l(t)}{dt} - i_k^l(t) \right] =$$

$$\frac{1}{2} \left[(\underline{\mathbf{A}}^T \mathbf{e})_k - v_{k0} - (\mathbf{C}_k)^{-1} \int_0^t (\underline{\mathbf{M}}^T \mathbf{i})_k dt' \right]^T \cdot \left[\mathbf{C}_k \frac{d(\underline{\mathbf{A}}^T \mathbf{e})_k}{dt} - (\underline{\mathbf{M}}^T \mathbf{i})_k \right]; \quad (12)$$

dove si è indicato con L_{kj} e Γ_{kj} l'elemento k,j rispettivamente delle matrici di induttanza e di induttanza reciproca associate alla rete e si è usata la simbologia classica della teoria dei circuiti per tutti gli altri termini.

Per evitare problemi numerici dovuti al fatto che le grandezze incognite possono differire per diversi ordini di grandezza può risultare vantaggioso moltiplicare, nella costruzione del funzionale globale di errore, ciascuno dei funzionali locali per un peso opportuno α_k .

Se gli l lati della rete contengono N_E generatori indipendenti di tensione, N_J generatori indipendenti di corrente, N_R resistori, N_L induttori e N_C condensatori, il funzionale globale di errore si scrive come:

$$\Lambda(\mathbf{e}, \mathbf{i}, \boldsymbol{\alpha}) = \int_0^T \left[\sum_{k=1}^{N_R} \alpha_k \lambda_{R,k}(\mathbf{e}, \mathbf{i}, t) + \sum_{k=1}^{N_L} \alpha_k \lambda_{L,k}(\mathbf{e}, \mathbf{i}, t) + \sum_{k=1}^{N_C} \alpha_k \lambda_{C,k}(\mathbf{e}, \mathbf{i}, t) \right] dt \quad (13)$$

2.4) Confronto tra le formulazioni basate sull'errore costitutivo nei problemi di campo e nei circuiti a parametri concentrati.

Uno dei maggiori problemi che il metodo basato sulla minimizzazione dell'errore costitutivo presenta dal punto di vista operativo è la scelta di valori adeguati per i pesi. Questi, come detto, variano in funzione delle caratteristiche del problema (caratteristiche del materiale, dinamica delle sorgenti). Si vedrà, inoltre, nel capitolo seguente, che altri parametri influenzano notevolmente la qualità delle soluzioni numeriche; e sono il metodo di integrazione e l'ampiezza del time-step per l'avanzamento nel tempo del sistema di equazioni differenziali. La scelta ottimale di questi valori può essere condotta attraverso un'analisi parametrica. D'altra parte, per problemi di campo tridimensionali,

un'analisi parametrica è un compito tutt'altro che semplice da eseguire a causa del notevole onere computazionale associato alla soluzione del singolo problema. Si può, allora, pensare di costruire un rozzo modello circuitale del problema di campo e ricavare da questo, con peso computazionale modesto, tutte le informazioni che occorrono per scegliere in modo soddisfacente i parametri per il problema di campo.

Allo scopo di meglio evidenziare le analogie che esistono tra il problema tri-dimensionale e il corrispondente modello zero-dimensionale, può risultare conveniente assumere come incognite per il circuito a parametri concentrati anzichè i potenziali di nodo e le correnti di maglia, gli integrali temporali di queste grandezze:

$$q_m(t) = \int_0^t i_m(t') dt' \quad m=1, l - (n-1) \quad (1)$$

$$\Phi_j(t) = \int_0^t e_j(t') dt' \quad j=1, n-1 \quad (2)$$

In tabella [2] vengono confrontati il modello tri-dimensionale e quello zero-dimensionale:

Problemi di campo

Circuiti a parametri concentrati

Equazioni di Maxwell

$$\nabla \times \mathbf{H} = \partial \mathbf{D}_T / \partial t$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$$

$$\nabla \cdot \mathbf{D}_T = 0$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\forall \mathbf{x} \in V \text{ e } \forall t \in [0, T]$$

Incognite duali

potenziale e vettore elettrico \mathbf{W}
potenziale e vettore magnetico \mathbf{A}

Campi

$$\mathbf{E} = -\partial \mathbf{A} / \partial t; \quad \mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A}$$

$$\mathbf{H} = \partial \mathbf{W} / \partial t; \quad \mathbf{D}_T = \mathbf{D}_{T0} + \nabla \times \mathbf{W}$$

Condizioni iniziali

$$\mathbf{D}(\mathbf{x}, 0) = \mathbf{D}_0(\mathbf{x}); \quad \mathbf{B}(\mathbf{x}, 0) = \mathbf{B}_0(\mathbf{x});$$

$$\mathbf{D}_T(\mathbf{x}, 0) = 0$$

$$\forall \mathbf{x} \in V$$

Condizioni al contorno

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}^*(\mathbf{x}, t);$$

$$\mathbf{W}(\mathbf{x}, t) = \mathbf{W}^*(\mathbf{x}, t);$$

$$\forall \mathbf{x} \in \partial V \text{ e } \forall t \in [0, T]$$

Equazioni di Kirchhoff

(Circuito con l lati ed n nodi)

$$\sum_{k=1}^{N_j} i_k = 0; \quad j=1, \dots, n-1$$

$$\sum_{k=1}^{N_m} v_k = 0; \quad m=1, \dots, l - (n-1)$$

Incognite duali

Cariche di maglia q_m ; $m=1, l - (n-1)$
Flussi di nodo Φ_j ; $j=1, n-1$

Correnti di maglia, Potenziali di nodo

$$e_j = d(\Phi_j) / dt; \quad j=1, n-1$$

$$i_m = d(q_m) / dt; \quad m=1, l - (n-1)$$

Condizioni iniziali

$$q_{Cm}(0) = q_{C0m}; \quad m=1, N_C$$

$$\varphi_{Lj}(0) = \varphi_{L0j}; \quad j=1, N_L$$

$$q_k(0) = 0; \quad k=1, l - (n-1)$$

Condizioni al contorno

$$(\underline{\mathbf{A}}^T \mathbf{F})_k = \int_0^t e_k^*(t') dt'; \quad k=1, N_E$$

$$(\underline{\mathbf{M}}^T \mathbf{q})_k = \int_0^t i_k^*(t') dt'; \quad k=1, N_J$$

Tabella [2]: Confronto tra le variabili e le equazioni del problema di campo e quelle del circuito a parametri concentrati scritte attraverso il metodo basato sull'errore costitutivo

A scopo di esempio si consideri il problema di campo di fig.[5.a] modellabile attraverso la rete RLC di fig. [5.b]:

Le condizioni iniziali del problema di campo

$$\mathbf{D}(\mathbf{x},0)=\mathbf{D}_0(\mathbf{x}); \quad \text{e} \quad \mathbf{B}(\mathbf{x},0)=\mathbf{B}_0(\mathbf{x});$$

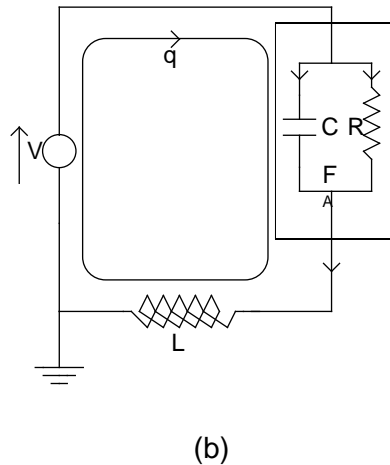
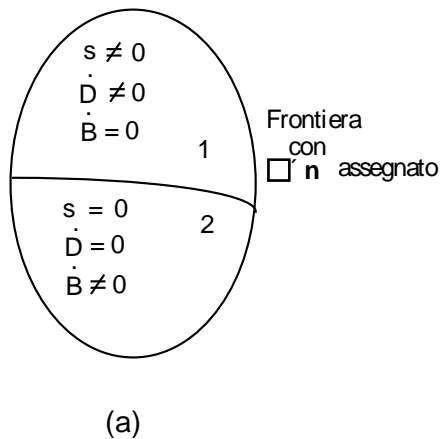
corrispondono nella rete elettrica alle condizioni iniziali sulla carica del condensatore e sul flusso dell'induttore:

$$q_C(0) = q_{C0}; \quad \phi_L(0) = \phi_{L0};$$

Le condizioni al contorno sul potenziale vettore \mathbf{A} sono rappresentate attraverso il generatore di tensione V . Le caratteristiche dei bipoli passivi sono:

$$i_R = i_R(v_C); \quad \mathbf{(3.a)} \quad q_C = q_C(v_C); \quad \mathbf{(3.b)} \quad i_L = i_L(\Phi_L); \quad \mathbf{(3.c)}$$

Se il coppia RC viene trattato come un singolo bipolo, come variabili nei termini delle quali risolvere la rete si possono assumere la carica di maglia, q , e il flusso di nodo $\Phi = \Phi_L$. Si può porre arbitrariamente $q(0) = 0$.



Fig[5] (a) Problema di campo; (b) Modello circuitale del problema (a)

Adottando per tutti i componenti passivi la convenzione dell'utilizzatore, le leggi di Kirchhoff alle tensioni e alle correnti sono:

$$V = v_C + \dot{\Phi}; \quad \mathbf{(4.a)}$$

$$i_L = i_R + i_C; \rightarrow \int_0^t i_L dt' = \int_0^t (i_R + i_C) dt'; \rightarrow q(t) = q_C(t) - q_{C0} + \int_0^t i_R dt';$$

(4.b)

Dalle relazioni (3.a), (3.b) e (4.b) si ricava l'equazione caratteristica per il cappio RC:

$$q = q_C(V - \dot{\Phi}) - q_{C0} + \int_0^t i_R(V - \dot{\Phi}) dt';$$

(3.d)

Nel caso lineare le relazioni costitutive (3 a,b,c) diventano:

$$i_R = \frac{v_C}{R} = \frac{V - \dot{\Phi}}{R}; \quad \mathbf{(5.a)} \quad q_C = C v_C = C(V - \dot{\Phi}); \quad \mathbf{(5.b)} \quad \dot{q} = \frac{\Phi}{L}; \quad \mathbf{(5.c)}$$

e la (3.d) diventa: $q = C(V - \dot{\Phi}) - q_{C0} + \int_0^t \frac{V - \dot{\Phi}}{R} dt';$ **(5.d)**

Se per q e Φ si adotta una approssimazione lineare nel generico intervallo $[t_k, t_k + \Delta t_k]$ si ha:

$$q(t) = \frac{q_{k+1} - q_k}{\Delta t_k} (t - t_k) + q_k; \quad \mathbf{(6.a)} \quad \dot{q}(t) = \frac{q_{k+1} - q_k}{\Delta t_k} \quad \mathbf{(6.b)}$$

$$\Phi(t) = \frac{\Phi_{k+1} - \Phi_k}{\Delta t_k} (t - t_k) + \Phi_k; \quad \mathbf{(7.a)} \quad \dot{\Phi}(t) = \frac{\Phi_{k+1} - \Phi_k}{\Delta t_k} \quad \mathbf{(7.b)}$$

dove: $q(t_k) = q_k$; e $\Phi(t_k) = \Phi_k$; sono assunte note dal passo precedente mentre $q(t_k + \Delta t_k) = q_{k+1}$; e $\Phi(t_k + \Delta t_k) = \Phi_{k+1}$; sono le incognite da determinare. Si ponga:

$$Cost = q_0 - q_{C0} + \int_0^{t_k} \frac{V - \dot{\Phi}}{R} dt = q_0 - q_{C0} + \sum_{m=0}^{K-1} \left[V(t_m^+) - \dot{\Phi}(t_m^+) \right] \Delta t_m$$

Le caratteristiche (5.c) e (5.d) si scrivono come:

$$\dot{q} = \frac{\Phi}{L}; \quad \left(V - \dot{\Phi} \right) = (q - Cost) \frac{R}{(t - t_k) + RC} = (q - Cost) A(t);$$

dove si è posto
$$A(t) = \frac{R}{(t - t_k) + RC};$$

I funzionali locali di errore, tenendo conto delle (6) e (7), e adoperando il metodo del time-stepping che sarà illustrato nel prossimo capitolo, sono:

$$\lambda_L(\Phi_{k+1}, q_{k+1}, t) = \frac{1}{2} \left[\dot{q} - \frac{\Phi}{L} \right] \cdot \left[L \dot{q} - \Phi \right]; \quad (8.a)$$

$$\lambda_{RC}(\Phi_{k+1}, q_{k+1}, t) = \frac{1}{2} \left[\left(V - \dot{\Phi} \right) - (q - Cost) A(t) \right] \cdot \left[\frac{V - \dot{\Phi}}{A(t)} - (q - Cost) \right]; \quad (8.b)$$

$$\Lambda_k(\Phi_{k+1}, q_{k+1}, \alpha) = \int_{t_k}^{t_k + \Delta t_k} \alpha_L \lambda_L(\Phi, q, t) + \alpha_{RC} \lambda_{RC}(\Phi, q, t) dt; \quad t \in [t_k, t_k + \Delta t_k] \quad (9)$$

Si riportano infine le costanti di tempo caratteristiche della rete:

$$\tau_1 = \frac{1}{\lambda_1} = \frac{1}{\left[-\frac{1}{2RC} - \sqrt{\left(\frac{1}{2RC} \right)^2 - \frac{1}{LC}} \right]};$$

$$\tau_2 = \frac{1}{\lambda_2} = \frac{1}{\left[-\frac{1}{2RC} + \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}} \right]};$$

2.5) Formulazione nei termini dell'errore costitutivo per le equazioni di Maxwell in regime sinusoidale.

Con un procedimento formalmente simile a quello adoperato nel dominio del tempo, sotto l'ipotesi di linearità delle relazioni costitutive, è possibile ricavare un'espressione del funzionale di errore adeguata a trattare problemi nei quali i campi si trovano in condizioni di regime sinusoidale. Le equazioni di Maxwell (2.2.1-1a-e) nel dominio della frequenza assumono il seguente aspetto:

$$\nabla \cdot \hat{\mathbf{D}}(\mathbf{x}, \omega) = \hat{\rho}(\mathbf{x}, \omega) \quad (1a)$$

$$\nabla \cdot \hat{\mathbf{B}}(\mathbf{x}, \omega) = 0 \quad (1b)$$

$$\nabla \times \hat{\mathbf{H}}(\mathbf{x}, \omega) = \hat{\mathbf{J}}(\mathbf{x}, \omega) + j \omega \hat{\mathbf{D}}(\mathbf{x}, \omega) \quad (1c)$$

$$\nabla \times \hat{\mathbf{E}}(\mathbf{x}, \omega) = -j \omega \hat{\mathbf{B}}(\mathbf{x}, \omega) \quad (1d)$$

$$\nabla \cdot \hat{\mathbf{J}}(\mathbf{x}, \omega) = -j \omega \hat{\rho}(\mathbf{x}, \omega) \quad (1e)$$

Mentre le relazioni di interfaccia valide sulle eventuali superfici di separazione Σ tra materiali diversi diventano

$$[\hat{\mathbf{D}}(\mathbf{x}, \omega)] \cdot \mathbf{n}_\Sigma = \hat{\rho}_\Sigma(\mathbf{x}, \omega) \quad (2a)$$

$$[\hat{\mathbf{B}}(\mathbf{x}, \omega)] \cdot \mathbf{n}_\Sigma = 0 \quad (2b)$$

$$[\hat{\mathbf{E}}(\mathbf{x}, \omega)] \times \mathbf{n}_\Sigma = 0 \quad (2c)$$

$$[\hat{\mathbf{H}}(\mathbf{x}, \omega)] \times \mathbf{n}_\Sigma = \hat{\mathbf{J}}_\Sigma(\mathbf{x}, \omega) \quad (2d)$$

$$[\hat{\mathbf{J}}(\mathbf{x},\omega)] \cdot \mathbf{n}_\Sigma = -j\omega \hat{\rho}_\Sigma(\mathbf{x},\omega) \quad (2e)$$

che devono essere soddisfatte $\forall \mathbf{x} \in \Sigma$.

Per definire in modo completo il problema resta da aggiungere $\forall \mathbf{x} \in V$ alle equazioni di Maxwell le relazioni costitutive:

$$\hat{\mathbf{B}}(\mathbf{x},\omega) = \mathbf{f}_B(\hat{\mathbf{H}}, \mathbf{x}, \omega), \quad (3a)$$

$$\hat{\mathbf{D}}(\mathbf{x},\omega) = \mathbf{f}_D(\hat{\mathbf{E}}, \mathbf{x}, \omega), \quad (3b)$$

$$\hat{\mathbf{J}}(\mathbf{x},\omega) = \mathbf{f}_J(\hat{\mathbf{E}}, \mathbf{x}, \omega) \quad (3c)$$

e, indicata con ∂V la frontiera esterna di V , $\forall \mathbf{x} \in \partial V$ opportune condizioni al contorno.

Naturalmente la relazione (1.e) si può ricavare dalle relazioni (1c) e (1a).

Si assumano le seguenti ipotesi:

Le relazioni costitutive (3) possono eventualmente essere anisotrope e non-omogenee ma devono essere non dispersive nello spazio e nel tempo, ed in più, per evitare dispersione in frequenza, lineari e tempo-invarianti (si rammenta a tal proposito che, se $x(t)$ è la rappresentazione di un campo scalare nel dominio nel tempo e $\hat{X}(\omega)$ è la corrispondente trasformata di Fourier, al quadrato del campo nel dominio del tempo $[x(t)]^2$ corrisponde in frequenza l'integrale di convoluzione $\hat{X}(\omega) * \hat{X}(\omega)$). Sotto tali ipotesi le relazioni (3) diventano:

$$\hat{\mathbf{B}}(\mathbf{x},\omega) = \mathbf{f}_B(\hat{\mathbf{H}}(\mathbf{x},\omega), \mathbf{x}) = \boldsymbol{\mu}(\mathbf{x}) \hat{\mathbf{H}}(\mathbf{x},\omega) \quad (3a)$$

$$\hat{\mathbf{D}}(\mathbf{x},\omega) = \mathbf{f}_D(\hat{\mathbf{E}}(\mathbf{x},\omega), \mathbf{x}) = \boldsymbol{\varepsilon}(\mathbf{x}) \hat{\mathbf{E}}(\mathbf{x},\omega) \quad (3b)$$

$$\hat{\mathbf{J}}(\mathbf{x},\omega) = \mathbf{f}_J(\hat{\mathbf{E}}(\mathbf{x},\omega), \mathbf{x}) = \boldsymbol{\sigma}(\mathbf{x}) \hat{\mathbf{E}}(\mathbf{x},\omega) + \hat{\mathbf{E}}_{(s)}(\mathbf{x},\omega) \quad (3c)$$

dove μ ed ε possono eventualmente essere quantità complesse:

$$\mu(\mathbf{x}) = \mu'(\mathbf{x}) - j\mu''(\mathbf{x}); \quad \varepsilon(\mathbf{x}) = \varepsilon'(\mathbf{x}) - j\varepsilon''(\mathbf{x})$$

Si ponga:

$$\hat{\mathbf{D}}_T(\mathbf{x}, \omega) = (j\omega)^{-1} \hat{\mathbf{J}}(\mathbf{x}, \omega) + \hat{\mathbf{D}}(\mathbf{x}, \omega) \quad (4)$$

Le equazioni di Maxwell si possono allora riscrivere nella forma seguente:

$$\nabla \cdot \hat{\mathbf{D}}_T = 0 \quad (5a)$$

$$\nabla \cdot \hat{\mathbf{B}} = 0 \quad (5b)$$

$$\nabla \times \hat{\mathbf{H}} = j\omega \hat{\mathbf{D}}_T \quad (5c)$$

$$\nabla \times \hat{\mathbf{E}} = -j\omega \hat{\mathbf{B}} \quad (5d)$$

e le condizioni di continuità sulle superfici di interfaccia comportano:

$$[\hat{\mathbf{D}}_T(\mathbf{x}, \omega)] \cdot \mathbf{n}_\Sigma = 0 \quad (6a)$$

$$[\hat{\mathbf{B}}(\mathbf{x}, \omega)] \cdot \mathbf{n}_\Sigma = 0 \quad (6b)$$

$$[\hat{\mathbf{E}}(\mathbf{x}, \omega)] \times \mathbf{n}_\Sigma = 0 \quad (6c)$$

$$[\hat{\mathbf{H}}(\mathbf{x}, \omega)] \times \mathbf{n}_\Sigma = 0 \quad (6d)$$

Le relazioni costitutive (3a,b,c) si possono condensare nelle seguenti espressioni:

$$\hat{\mathbf{B}}(\mathbf{x}, t) = \mathbf{f}_B(\hat{\mathbf{H}}(\mathbf{x}, \omega), \mathbf{x}) = \mu(\mathbf{x}) \hat{\mathbf{H}}(\mathbf{x}, \omega) \quad (7a)$$

$$\begin{aligned} \hat{\mathbf{D}}_T(\mathbf{x}, t) &= \mathbf{f}_D(\hat{\mathbf{E}}(\mathbf{x}, \omega), \mathbf{x}) + (j\omega)^{-1} \mathbf{f}_J(\hat{\mathbf{E}}(\mathbf{x}, \omega), \mathbf{x}) = \\ &= [\varepsilon(\mathbf{x}) + (j\omega)^{-1} \sigma(\mathbf{x})] \hat{\mathbf{E}}(\mathbf{x}, \omega) + (j\omega)^{-1} \sigma(\mathbf{x}) \hat{\mathbf{E}}_{(s)}(\mathbf{x}, \omega) \end{aligned} \quad (7b)$$

Si definiscano i due potenziali vettori nel seguente modo:

$$\hat{\mathbf{A}} = -(\mathrm{j}\omega)^{-1} \hat{\mathbf{E}} \quad (8a)$$

$$\hat{\mathbf{W}} = (\mathrm{j}\omega)^{-1} \hat{\mathbf{H}} \quad (8b)$$

e si ponga

$$\hat{\mathbf{B}} = \nabla \times \hat{\mathbf{A}} \quad (9a)$$

$$\hat{\mathbf{D}}_T = \nabla \times \hat{\mathbf{W}} \quad (9b)$$

si perviene alle formulazioni duali:

$$\begin{aligned} \nabla \times \hat{\mathbf{H}} &= \mathrm{j}\omega \hat{\mathbf{D}}_T \\ \Downarrow \\ \hat{\mathbf{H}} &= \mathrm{j}\omega \hat{\mathbf{W}}, \quad \hat{\mathbf{D}}_T = \nabla \times \hat{\mathbf{W}} \end{aligned}$$

$$\begin{aligned} \nabla \times \hat{\mathbf{E}} &= -\mathrm{j}\omega \hat{\mathbf{B}} \\ \Downarrow \\ \hat{\mathbf{E}} &= \mathrm{j}\omega \hat{\mathbf{A}}, \quad \hat{\mathbf{B}} = \nabla \times \hat{\mathbf{A}} \end{aligned}$$

$$\nabla \times \hat{\mathbf{H}} - \mathrm{j}\omega \hat{\mathbf{D}}_T = \mathbf{0}, \quad \forall \hat{\mathbf{W}}$$

$$\nabla \times \hat{\mathbf{E}} + \mathrm{j}\omega \hat{\mathbf{B}} = \mathbf{0}, \quad \forall \hat{\mathbf{A}}$$

nelle quali, per ogni possibile scelta dei potenziali $\hat{\mathbf{A}}$ e $\hat{\mathbf{W}}$, le equazioni di Maxwell (5) risultano automaticamente soddisfatte.

Si noti che le equazioni cui si è pervenuto sono molto simili a quelle già proposte in letteratura da Pichon e Bossavit [31] con la sola differenza che qui si è usato $\hat{\mathbf{D}}_T$ al posto di $\hat{\mathbf{D}}$

Occorre, adesso, definire un funzionale di errore locale $\lambda(\mathbf{x}, \omega) \geq 0$ per descrivere la misura con la quale le relazioni costitutive sono soddisfatte; esso

dovrà essere nullo solo nei punti spazio-frequenza dove le relazioni costitutive (7) sono rigorosamente rispettate.

A questo scopo si tenga in conto il fatto che il quadrato di un vettore complesso $\hat{\mathbf{x}} = \mathbf{x}_r + j\mathbf{x}_i$, dove \mathbf{x}_r e \mathbf{x}_i sono vettori reali, è la quantità complessa: $(\hat{\mathbf{x}})^2 = (\mathbf{x}_r \cdot \mathbf{x}_r)^2 - (\mathbf{x}_i \cdot \mathbf{x}_i)^2 + j(\mathbf{x}_r \cdot \mathbf{x}_i)$.

Supposti ϵ e μ reali, si propone di usare il funzionale scalare e di tipo reale funzione di variabili vettoriali complesse, i potenziali vettori :

$$\lambda(\mathbf{x}, \omega) = \alpha_H (\mu)^{-1} [\hat{\mathbf{B}} - \mu \hat{\mathbf{H}}] \cdot [\hat{\mathbf{B}} - \mu \hat{\mathbf{H}}]^* + \alpha_E [\epsilon^2 + (\sigma/\omega)^2]^{-1/2} \cdot [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}] - [\epsilon + (j\omega)^{-1} \sigma] \hat{\mathbf{E}}] \cdot [[\hat{\mathbf{D}}_T - (j\omega)^{-1} \sigma \hat{\mathbf{E}}_{(s)}] - [\epsilon + (j\omega)^{-1} \sigma] \hat{\mathbf{E}}]^* \quad (10)$$

Appare evidente che se le relazioni costitutive sono soddisfatte il funzionale di errore si annulla.

Si può dimostrare che il funzionale sopra definito è positivo e di tipo reale scrivendo in modo esplicito ogni campo nella forma di parte reale e immaginaria.

Si arriva così alla definizione del funzionale globale di errore:

$$\Lambda(\omega) = \int_V \lambda(\mathbf{x}, \omega) dV \quad (11)$$

Il problema di minimizzazione ha soluzione unica a meno che $\omega > 0$ non corrisponda ad una frequenza di risonanza (il che, d'altra parte non può accadere se nel mezzo ci sono perdite (ovvero σ e/o μ'' e/o ϵ'' sono diversi da zero)[31].

Nelle formulazioni basate sugli edge elements poichè i modi spuri compaiono a frequenza esattamente nulla ed anche in numero prevedibile a priori, si può evitare di calcolarli.

Come sottolineato nel lavoro sopra citato e ribadito nel paragrafo (1.2.3), se si accoppia questa formulazione all'uso degli edge elements, si può evitare il problema dei modi spuri. Con questo nome si indicano le soluzioni numeriche

del problema elettromagnetico che non verificano la condizione di solenoidalità per il vettore di induzione magnetico $\hat{\mathbf{B}}$ e, ove questa sia richiesta, per il vettore spostamento elettrico $\hat{\mathbf{D}}$, e che pertanto sono prive di significato fisico (vedi paragrafo 1.2.3).

CAPITOLO III

Aspetti computazionali

3.1) Soluzione numerica dei problemi a valore iniziale

3.1.1) Generalità

Le equazioni differenziali alle derivate parziali sono usualmente classificati nelle tre categorie di equazioni ellittiche, paraboliche ed iperboliche in base alle loro *curve caratteristiche*, ovvero alle curve che determinano il modo in cui l'informazione si propaga.

Tipici esempi di equazioni iperboliche sono le equazioni di convezione e le equazioni delle onde. Quest'ultima, per il caso monodimensionale si scrive come:

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

dove $v = \text{cost.}$ è la velocità di propagazione dell'onda.

Un tipico esempio di equazione parabolica è l'equazione di diffusione:

$$\frac{\partial u}{\partial t} = - \frac{\partial}{\partial x} \left(D \frac{\partial u}{\partial x} \right)$$

(2)

dove D è il coefficiente di diffusione.

Infine il prototipo di un'equazione ellittica è l'equazione di Poisson:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \rho(x,y) \quad (3)$$

dove \square è il termine sorgente. Se quest'ultimo è identicamente nullo l'equazione risultante è la ben nota equazione di Laplace.

Da un punto di vista computazionale la distinzione che assume maggiore importanza è quella tra equazioni di tipo ellittico ed equazioni di altra categoria. Difatti, mentre la soluzione dell'equazione (3) è una funzione statica, per trovare la quale è necessario specificare i suoi valori, oppure quelli della derivata normale, sulla frontiera del dominio di definizione (problemi siffatti sono detti, per questo motivo, *di valore al contorno*), i problemi del tipo (1) e (2), partendo dalla conoscenza di u all'istante iniziale, descrivono come la funzione incognita evolva nel tempo. Questi ultimi problemi sono generalmente definiti *a valore iniziale* o *di Cauchy*.

La sottoclassificazione dei problemi di Cauchy in iperbolici e parabolici è poco significativa dal punto di vista pratico in quanto mentre da una parte molti dei problemi di interesse pratico sono naturalmente di tipo misto, dall'altra gli schemi computazionali usualmente adoperati fanno sì che la maggior parte dei problemi iperbolici assuma forma mista all'atto della computazione.

3.1.2) Soluzione numerica dei problemi a valore iniziale o di Cauchy

Tra le varie proposte per la soluzione numerica dei problemi a valore iniziale, sarà presentata quella che maggiore fortuna ha incontrato in letteratura: questa prevede l'uso congiunto degli Elementi finiti, per il trattamento delle variabili spaziali, e del cosiddetto time-stepping per la rappresentazione dell'evoluzione temporale.

Attraverso un metodo basato sugli elementi finiti si perviene, al termine della fase di assemblaggio delle matrici, ad un sistema di N equazione differenziali ordinarie accoppiate che per i problemi di tipo parabolico ha la forma seguente:

$$[\mathbf{C}] \left(\frac{\partial \mathbf{a}(t)}{\partial t} \right) + [\mathbf{K}](\mathbf{a}(t)) = (\mathbf{f}(t)); \quad t \in [t_0, t_0+T] \quad \text{con } \mathbf{a}(t_0) = \mathbf{a}_0 \quad (4)$$

dove $[C]$ è detta *matrice di capacità*, $[K]$ è la *matrice di rigidità* ed f è il *vettore di carico*, e a_0 è il vettore delle condizioni iniziali assegnate.

Nei problemi misti il sistema differenziale si presenta come:

$$[M] \left(\frac{\partial^2 \mathbf{a}(t)}{\partial t^2} \right) + [C] \left(\frac{\partial \mathbf{a}(t)}{\partial t} \right) + [K] \mathbf{a}(t) = \mathbf{f}(t) \quad t \in [t_0, t_0+T]$$

$$\text{con } \mathbf{a}(t_0) = \mathbf{a}_0 \quad (5)$$

dove $[M]$ è la *matrice di massa*, mentre $[C]$ viene ora chiamata *matrice di smorzamento*, anzichè *di capacità*.

Nei problemi iperbolici vale ancora la (5), a patto che $[C]$ coincida con la matrice nulla di ordine N.

In tutti i metodi basati sul time-stepping il dominio temporale $[t_0, t_0+T]$ è diviso in una successione di intervalli Δt_i , $i=1, N_{\text{step}}$ a partire dall'istante t_0 . In alcuni metodi il passo degli intervalli deve essere mantenuto costante, in altri può avere lunghezza diversa. A questo punto, anzichè cercare la soluzione del sistema differenziale come un vettore $\mathbf{a}(t)$ le cui componenti $a_i(t)$ sono funzione continua, si assumono come incognite i valori discreti $\mathbf{a} = \mathbf{a}(t_i)$ che le funzioni assumono alla fine di ciascun time-step. Gli $\{\mathbf{a}_i\}$, $i=1, N_{\text{step}}$, vengono calcolati attraverso una relazione di ricorrenza, ovvero attraverso una equazione algebrica che approssima l'equazione differenziale e che correla il valore di due (metodi a singolo step) o più (metodi multi-step) termini consecutivi della successione \mathbf{a}_i . Nei metodi mono-step, ad esempio, vengono correlati i valori discreti degli estremi di un singolo step: in tal caso si usano i valori iniziali noti \mathbf{a}_0 per determinare i valori \mathbf{a}_1 assunti dal vettore soluzione alla fine del primo time-step; per il secondo time step si procede in maniera analoga usando per la determinazione del vettore \mathbf{a}_2 al posto di \mathbf{a}_0 il vettore \mathbf{a}_1 determinato al passo precedente, e così via per tutti i time-steps successivi.

Per problemi di diffusione sono quasi universalmente adoperati metodi ad uno o due step, mentre per problemi misti sono soventemente impiegati

metodi multi-step di ordine superiore. I metodi mono-step includono tre classiche formule delle differenze finite:

- *Il metodo implicito* noto anche come la *regola di Eulero a ritroso*.
- *Il metodo delle differenze centrali*, detto *metodo Crank-Nicolson* o *della regola trapezoidale*.
- *Il metodo esplicito* noto come *regola di Eulero*.

I tre metodi precedenti sono dei casi speciali di una formula più generale indicata come il *metodo θ* .

Da un punto di vista pratico nella scelta del metodo da adottare bisogna tenere conto di due indici fondamentali, che dipendono dal solo schema di integrazione e non dal particolare sistema differenziale: il tasso di convergenza asintotico e la stabilità.

Il tasso di convergenza asintotico misura quanto rapidamente decada, in un dato istante t , l'errore sulla soluzione al diminuire del time-step, Δt , per $\Delta t \rightarrow 0$.

La stabilità dello schema di integrazione fornisce informazioni sul comportamento della soluzione quando $t \rightarrow \infty$, l'ampiezza del time-step è mantenuta costante. Il comportamento stabile o instabile di un metodo può essere determinato osservando l'evoluzione libera di un sistema del primo ordine (che, come è noto, è una funzione che decade con andamento esponenziale). Un metodo instabile produrrà una soluzione che oscilla con ampiezza sempre crescente attorno alla curva di decadimento esponenziale. In un metodo stabile, invece, la soluzione o conterrà oscillazioni che vanno decedendo con $t \rightarrow \infty$, oppure sarà del tutto priva di oscillazioni.

Premesso che, pur se sviluppate con maggiore complessità, le idee di fondo di tutti i metodi multi-step ricalcano quelle già presenti negli schemi a singolo step, sono illustrate brevemente di seguito le caratteristiche peculiari del metodo mono-step applicato al problema di diffusione il cui modello è rappresentato dal sistema differenziale (4).

Metodo implicito

Per tutti gli schemi del primo ordine il valore della derivata prima per il k-mo time-step, può essere approssimato attraverso la relazione:

$$\left(\frac{\partial \mathbf{a}(t)}{\partial t} \right) = \frac{\{\mathbf{a}\}_k - \{\mathbf{a}\}_{k-1}}{\Delta t_k} \quad \forall t \in [t_k, t_k + \Delta t_k] \quad (6)$$

Nello schema implicito il sistema (4) si può scrivere come:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_k + [\mathbf{K}] \{\mathbf{a}\}_k = \{\mathbf{f}\}_k \quad (7)$$

dove si è indicato con il pedice k il valore assunto dalla grandezza all'istante t_k . Sostituendo la (6) nella (7) si perviene alla relazione di ricorrenza:

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + [\mathbf{K}] \right) \frac{1}{\Delta t_k} [\mathbf{C}] \{\mathbf{a}\}_k = \{\mathbf{f}\}_k + \frac{1}{\Delta t_k} [\mathbf{C}] \{\mathbf{a}\}_{k-1} \quad (8)$$

Il tasso di convergenza asintotico del metodo implicito è del primo ordine, cioè l'errore sulla soluzione in un dato istante t è proporzionale a Δt quando $\Delta t \rightarrow 0$.

La maggiore attrattiva di questo metodo è di essere del tutto privo di oscillazioni e incondizionatamente stabile.

Metodo di Crank-Nicolson

Nello schema alla Crank-Nicolson si adopera per approssimare il sistema differenziale (4) l'espressione:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{k-1/2} + [\mathbf{K}] \{\mathbf{a}\}_{k-1/2} = \{\mathbf{f}\}_{k-1/2} \quad (9)$$

che determina il seguente sistema algebrico:

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + \frac{1}{2} [\mathbf{K}] \right) \{\mathbf{a}\}_k = \{\mathbf{f}\}_{k-1/2} + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - \frac{1}{2} [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (10)$$

Lo schema delle differenze centrali converge in maniera proporzionale a Δt^2 , è stabile ma presenta oscillazioni nella soluzione.

Metodo di Eulero

Il metodo di Eulero usa come schematizzazione della (4) l'espressione:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{k-1} + [\mathbf{K}] \{\mathbf{a}\}_{k-1} = \{\mathbf{f}\}_{k-1} \quad (11)$$

che comporta:

$$\frac{1}{\Delta t_k} [\mathbf{C}] \{\mathbf{a}\}_k = \{\mathbf{f}\}_{k-1} + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (12)$$

A differenza dei metodi precedenti, nello schema di Eulero la matrice di rigidità, $[\mathbf{K}]$, non compare più al primo membro della relazione di ricorrenza, mentre resta presente la matrice di capacità. Poichè quest'ultima normalmente non è diagonale, lo schema (12) è ancora implicito. Tuttavia, in questo caso, la matrice $[\mathbf{C}]$ può essere resa diagonale, la qual cosa disaccoppia le incognite e rende questo metodo estremamente veloce. La diagonalizzazione della matrice può essere ottenuta rapidamente attraverso delle tecniche che vanno sotto il generico nome di *lumping*. L'efficacia computazionale del *lumping* si può misurare osservando quanto vari la velocità di convergenza e la precisione della soluzione se si risolve lo schema modificato al posto dello schema (12). In generale si potrà osservare una diminuzione del tasso di

convergenza e, cosa interessante, un miglioramento a volte significativo nella precisione.

Tra le varie tecniche di lumping, la più semplice, applicata prima della fase di assemblaggio alle matrici di capacità dei singoli elementi, consiste nel ricavare la matrice concentrata $[\mathbf{CL}]^{(\theta)}$ a partire dalla matrice di capacità $[\mathbf{C}]$ col seguente algoritmo:

$$CL_{ii}^{(e)} = \sum_{j=1}^n C_{ij}^{(e)}; \text{ per } i=1, 2, \dots, n; CL_{ij}^{(e)} = 0; \text{ per } i \neq j; \quad (13)$$

A fronte del pregio di essere più veloce dal punto di vista computazionale rispetto ai metodi precedenti, il metodo esplicito presenta il grave inconveniente di essere potenzialmente instabile.

Metodo θ .

Questo metodo è la naturale generalizzazione dei precedenti. In questo caso la (4) viene valutata nella generica locazione dell'intervallo individuata dal parametro adimensionale θ :

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{\Theta} + [\mathbf{K}] \{\mathbf{a}\}_{\Theta} = \{\mathbf{f}\}_{\Theta} \quad (14)$$

dove :

$$\Theta = \frac{t - t_{k-1}}{\Delta t_k} \quad (15)$$

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + \Theta [\mathbf{K}] \right) \{\mathbf{a}\}_k = (1 - \Theta) \{\mathbf{f}\}_{k-1} + \Theta \{\mathbf{f}\}_k + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - (1 - \Theta) [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (16)$$

Per quanto riguarda la stabilità, si può dire in generale che:

- Per $1 \geq \theta \geq 0.5$ lo schema di integrazione è incondizionatamente stabile.

- Per $0 \leq \theta < 0.5$ lo schema è stabile solo se, indicato λ_{Max} con il maggior autovalore della matrice $[C^{-1}K]$ viene soddisfatta la seguente disuguaglianza, che pone un limite sulla massima ampiezza del time-step:

$$\lambda_{Max} \Delta t < \frac{2}{1 - 2\theta} \quad (17)$$

Esistono delle formule che permettono di stimare λ_{Max} senza passare per il calcolo degli autovalori della matrice []

Oltre a quei valori (0, .5, 1) per i quali si ritorna agli schemi già esaminati, altri valori di θ che spesso vengono adoperati sono $\theta=2/3$ e $\theta=0.878$ (che rappresenta il valore minore per il quale la soluzione non presenta oscillazioni).

[D.Burnett, *Finite Element Analysis*, Addison Wesley 1987

Ref. Bossavit, Kameari, Albanese-Rubinacci, van Welij, Comincioli, Num. Recipes, Bulirsch..)

3.2) Metodi di soluzione di sistemi di eq. lineari e non lineari

3.2.1) Risoluzione di sistemi di equazioni lineari.

Introduzione

La risoluzione dei sistemi lineari riveste un ruolo basilare nella matematica applicata. Mentre, infatti, diversi modelli matematici significativi nascono direttamente lineari, si è già mostrato come, attraverso i metodi numerici, anche problemi di dimensione infinita (ad esempio quelli modellati attraverso equazioni integro-differenziali non lineari), siano riconducibili a sistemi non

lineari di dimensione finita che è spesso possibile linearizzare (ovvero risolvere per via iterativa attraverso problemi lineari).

Alla varietà di contesti da cui un sistema lineare può scaturire, corrisponde una uguale differenziazione nella struttura della matrice associata al sistema. Questa, infatti, può essere di volta in volta piena, bandata o sparsa, simmetrica oppure definita positiva. E' importante, dunque, avere a disposizione una varietà di algoritmi efficienti tra i quali scegliere il più adeguato per stabilità, occupazione di memoria, velocità di esecuzione a risolvere su un determinato calcolatore un problema particolare.

In questo paragrafo si definiranno alcuni termini adoperati usualmente nella classificazione delle matrici. Si daranno quindi dei brevi cenni circa i metodi di immagazzinamento delle matrici sparse. La parte conclusiva è dedicata, seguendo l'usuale ripartizione in metodi diretti ed iterativi, all'illustrazione di alcuni delle tecniche adoperate nella risoluzione di sistemi lineari $N \times N$ non-singolari.

Generalità intorno alla risoluzione di sistemi lineari

Si riportano alcune definizioni adoperate usualmente in Algebra Lineare nella classificazione delle matrici :

Una matrice è detta reale se tutti i suoi elementi appartengono al dominio dei numeri reali.

Una matrice \mathbf{A} quadrata di ordine N ad elementi reali tale che $\mathbf{A}=\mathbf{A}^t$ è detta *simmetrica*.

Una matrice \mathbf{A} quadrata di ordine N ad elementi complessi tale che $\mathbf{A}=\mathbf{A}^*$ è detta *hermitiana*.

Una matrice reale \mathbf{A} di ordine N tale che per ogni $\mathbf{x} \in \mathbf{R}^n$ diverso dall'elemento nullo si abbia $\mathbf{x}^t \mathbf{A} \mathbf{x} > 0$ viene detta *definita positiva*.

Una matrice reale \mathbf{A} quadrata di ordine N e di elementi $a_{i,j}$ tale che si abbia:

$$a_{i,i} > a_{i,j} \quad \forall i,j \in [1,N] \quad \text{con } i \neq j$$

è detta *a dominanza diagonale*

Una matrice quadrata **D** è detta diagonale se tutti i suoi elementi che non si trovano sulla diagonale principale sono nulli.

Una matrice quadrata **U** è detta triangolare superiore se tutti i suoi elementi che si trovano sotto la diagonale principale sono nulli.

Una matrice quadrata **L** è detta triangolare inferiore se tutti i suoi elementi che giacciono sopra la diagonale principale sono nulli.

Si dimostra che se **A** è una matrice quadrata simmetrica è sempre possibile trovare una matrice diagonale **D** e una matrice triangolare superiore **U** tale che $\mathbf{A}=\mathbf{U}^T\mathbf{D}\mathbf{U}$ (Fattorizzazione di **A**).

Si è già evidenziata la stretta correlazione tra la struttura di una matrice e il tipo di metodo numerico che si sta adoperando. Una matrice in cui prevalgono termini diversi da zero viene chiamata "*piena*", mentre una in cui dominano i termini nulli si classifica come *sparsa*.

I metodi numerici di tipo integrale (ad esempio il metodo degli elementi di frontiera) sono caratterizzati da matrici associate piene di dimensione relativamente piccola (100-1000 incognite), viceversa i metodi basati sugli elementi finiti generano sistemi con matrici associate sparse e di ordine elevato (1000-30000 incognite).

Esistono diverse tecniche di immagazzinamento dei coefficienti delle matrici di tipo sparso, e la scelta della più adatta al problema in esame viene determinata dalle caratteristiche strutturali di sparsità della matrice. I differenti approcci vengono usualmente classificati in letteratura sotto le etichette di *metodi a bande*, *metodi d'inviluppo* e *metodi di sparsità generale*.

Da un punto di vista pratico capita spesso di incontrare problemi basati su metodi FEM che richiedono la soluzione di diversi sistemi lineari di tipo sparso per i quali la struttura della matrice -legata alla mesh, alle funzioni di base ed alle funzioni test-, non subisce modifiche, mentre possono variare sia i coefficienti del sistema che i termini noti. In questi casi l'approccio risolutivo generalmente adoperato passa attraverso tre stadi che vanno sotto i nomi di

fattorizzazione simbolica , *fattorizzazione numerica* della matrice e *soluzione del sistema* .

La fattorizzazione simbolica viene eseguito una volta sola, all'inizio del processo solutivo e il suo scopo è quello di determinare i legami tra le caratteristiche di sparsità delle matrici **A** e **U**. Il processo di fattorizzazione numerica va invece ripetuto ogni volta che un nuovo sistema deve essere risolto. Tuttavia le informazioni acquisite durante la fase di fattorizzazione simbolica fanno sì che quest'ultima operazione risulti estremamente veloce

Per i metodi a banda o di involuppo, caratterizzati da sparsità strutturali, i blocchi contenenti elementi non nulli sono semplicemente identici e pertanto il processo di fattorizzazione simbolica può essere evitato.

Nel caso dei metodi di sparsità generali, invece i legami tra **A** e **U** sono molto meno intuitivi e dipendono oltre che dallo schema di fattorizzazione (che determina il cosiddetto *fill in* o riempimento della matrice) anche dalla numerazione dei nodi della mesh. Per questo motivo i metodi di sparsità generale vengono usualmente preceduti da algoritmi -il più famoso dei quali è quello di Cuthill-McKee- capaci di minimizzare il fill in attraverso il riordinamento della numerazione dei nodi. In questo caso il costo dell'operazione. Questo insieme di operazioni può essere molto costoso dal punto di vista computazionale e, dunque, la fase preliminare di fattorizzazione simbolica produce un consistente vantaggio.

Un elemento che bisogna portare in conto nella scelta dello schema risolutivo è l'onere computazionale che questo impone. Questo parametro viene comunemente valutato attraverso il numero di moltiplicazioni e divisioni - le operazioni elementari più costose per la macchina calcolatrice- che l'algoritmo richiede per la risoluzione di un sistema di ordine N .

Metodi di immagazzinamento delle matrici

Metodi di banda

I metodi di banda sono adeguati per il trattamento di matrici, caratterizzate dalla concentrazione dei termini non nulli entro uno stretto numero di diagonali adiacenti la diagonale principale. Il parametro caratteristico, m_A , associato alla matrice *bandata* \mathbf{A} viene chiamato *ampiezza di banda*, ed è definito come il più piccolo intero tale che $a_{i,j}=0$ per $|i-j|>m_A$. Una importante proprietà delle matrici a bande è che la fattorizzazione $\mathbf{A}=\mathbf{U}^T\mathbf{D}\mathbf{U}$, non comporta allargamento della banda per la matrice triangolare superiore: per questo particolare processo di fattorizzazione risulta cioè $m_A=m_U$.

Sistemi di matrici bandate provengono, ad esempio, da problemi agli elementi finiti monodimensionali, o anche bidimensionali ma definiti in domini geometricamente molto regolare.

Per tutti i tipi di matrici sparse sono state elaborate efficaci strategie per ridurre al minimo la quantità di memoria necessaria per l'immagazzinamento dei coefficienti. In particolare nel caso di matrici bandate simmetriche si usa immagazzinare le diagonali che contengono termini non nulli in un array bidimensionale. La diagonale principale riempie la prima riga dell'array, quella immediatamente sopra viene immagazzinata nella seconda riga, e così di seguito finché l'intera banda collocata nel triangolo superiore delle matrici di partenza non ha trovato posto nell'array. Non è necessario per la simmetria della matrice stipare anche i termini del triangolo inferiore. Un esempio è mostrato in figura [].

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 \\ a_{12} & a_{22} & a_{23} & a_{24} & 0 & 0 \\ a_{13} & a_{23} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & a_{35} & a_{45} & a_{55} & a_{56} \\ 0 & 0 & a_{36} & a_{46} & a_{56} & a_{66} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{44} & a_{55} & a_{66} \\ a_{12} & a_{23} & a_{34} & a_{45} & a_{56} & * \\ a_{13} & a_{24} & a_{35} & a_{46} & * & * \\ a_{14} & 0 & a_{36} & * & * & * \end{bmatrix}$$

(a) (b)

Figura []

(a) Matrice bandata \mathbf{A} con $m_A=3$;

(b) Schema di immagazzinamento per \mathbf{A} con un array bidimensionale.

(* = Elementi dell'array di immagazzinamento non adoperati)

Adoperando il criterio sopra illustrato lo spazio di immagazzinamento di una matrice bandata di ordine N e di ampiezza di banda m_A passa da N^2 a $N \times (1+m_A)$ locazioni di memoria che includono anche $m_A(1+m_A)/2$ posizioni dell'array che non vengono adoperate. Queste locazioni vengono tuttavia mantenute allo scopo di facilitare l'accesso all'intera struttura, e, in ogni caso non occupano porzioni di memoria significative, essendo $m_A \ll N$. Si noti infine che le righe del triangolo superiore di \mathbf{A} coincidono con le colonne dell'array di immagazzinamento. Questo torna molto utile nei linguaggi come il Fortran che immagazzinano gli array per colonne, in quanto la corrispondenza dello schema di immagazzinamento al linguaggio di programmazione riduce al minimo la difficoltà di accedere alle righe di \mathbf{A} con un complessivo aumento delle prestazioni dell'algoritmo in termini di velocità.

Metodi di involuppo

Un problema FEM bidimensionale, con geometria del dominio di definizione abbastanza complessa, conduce ad un moderato grado di variazione locale dell'ampiezza di banda della matrice. Attraverso una opportuna rinumerazione degli elementi geometrici (nodi o lati) cui sono associate le incognite del problema è possibile ottenere una matrice nella quale l'involuppo degli elementi non nulli ha una forma lenticolare. In questo caso, sebbene l'ampiezza di banda della matrice resti uguale a m_A , fatta eccezione per la riga centrale $N/2$, gli elementi non nulli si trovano generalmente ben all'interno della banda di ampiezza m_A . Detta \mathbf{A} la matrice, che supponiamo simmetrica, ed \mathbf{U} la matrice triangolare superiore ottenuta dalla fattorizzazione di \mathbf{A} , l'analisi del processo di fattorizzazione mostra che $u_{ij}=0$ se $i < f_j \equiv \min(\{k: a_{kj} \neq 0\})$; cioè, per ciascuna colonna gli elementi di \mathbf{U} diversi da zero sono localizzati tra il primo elemento non nullo di \mathbf{A} sulla stessa colonna e la diagonale principale. Gli elementi di \mathbf{A} appartenenti all'insieme $\text{Env}(\mathbf{A}) = \{a_{ij}: f_j < i < j\}$ costituiscono l'involuppo di \mathbf{A} .

La struttura di immagazzinamento dell'involuppo di \mathbf{A} è più complessa, ma, in certi casi, più efficiente rispetto a quella esaminata precedentemente che conservava l'intera banda della matrice. Gli elementi vengono immagazzinati per colonne in un'array monodimensionale che denoteremo con A . In questo schema, risulta necessario definire un ulteriore array I_A , i cui elementi sono dei puntatori ai termini che formano la diagonale principale di \mathbf{A} . Un esempio della struttura è mostrato in figura [1].

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & a_{17} \\ a_{12} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ 0 & a_{23} & a_{33} & a_{34} & 0 & 0 & 0 \\ 0 & 0 & a_{34} & a_{44} & a_{45} & 0 & 0 \\ 0 & 0 & 0 & a_{45} & a_{55} & a_{56} & 0 \\ 0 & 0 & 0 & 0 & a_{56} & a_{66} & a_{67} \\ a_{17} & 0 & 0 & 0 & 0 & a_{67} & a_{77} \end{bmatrix} \quad (\text{a})$$

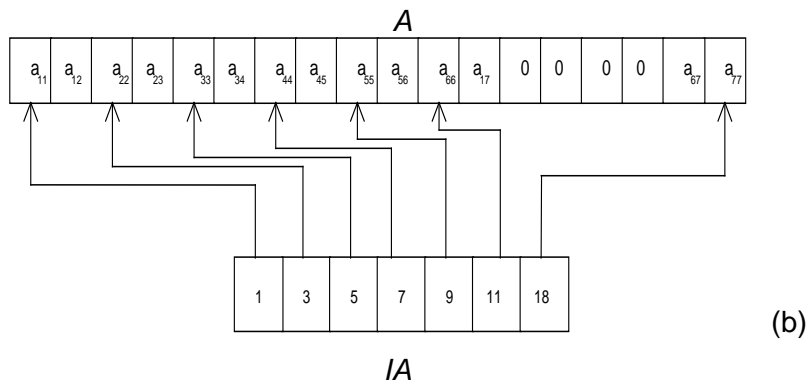


Figura [] : (a) Matrice **A**
 (b) Struttura di immagazzinamento di involuppo per **A**

Con questo schema di immagazzinamento è semplice calcolare f_j come:

$$f_j = \begin{cases} 1 & \text{se } j = 1 \\ j + 1 - (IA(j) - IA(j - 1)) & \text{se } j \neq 1 \end{cases}$$

e da qui accedere agli elementi di ogni colonna della matrice.

La struttura di involuppo risulta più efficiente rispetto a quella di banda in vari casi. Un esempio è costituito dai problemi bidimensionali in cui si sia provveduto a numerare i nodi della mesh lungo le diagonali (Vedi Fig. []). In tal caso si può osservare come lo schema di immagazzinamento di involuppo consenta un risparmio del 33% in termini di locazioni di memoria e del 50% nel tempo di fattorizzazione.

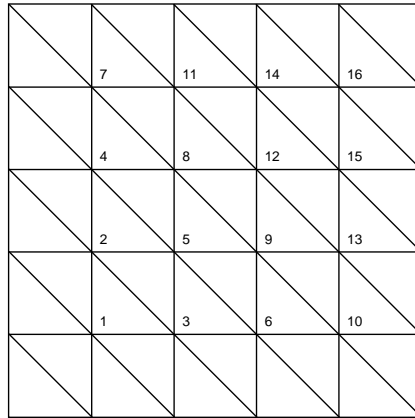


Figura []: Ordinamento diagonale dei nodi

Metodi di sparsità generale

I metodi a banda e d'inviluppo, mentre risultano vantaggiosi nel trattamento di matrici che presentano blocchi strutturali compatti (sia cioè possibile partizionare la matrice in zone in cui si trovano soltanto elementi nulli e in zone che presentano un'alta densità di elementi diversi da zero), si dimostrano totalmente insoddisfacenti in tutti i problemi che danno origine a matrici la cui sparsità è di tipo non strutturale: ciò è vero, in particolare, per una vasta classe di problemi FEM bi- e tri-dimensionali. Si preferiscono utilizzare in questi casi dei metodi che a prezzo di una maggiore complessità della struttura di immagazzinamento dei coefficienti garantiscono una significativa riduzione della quantità di dati da conservare.

Sia \mathbf{A} una matrice simmetrica di ordine N che presenta caratteristiche di sparsità non strutturali. Per la simmetria sarà possibile immagazzinare solo i coefficienti che si trovano sulla diagonale \mathbf{D} oppure nel blocco triangolare inferiore \mathbf{E} di \mathbf{A} . Per l'immagazzinamento dei coefficienti si utilizzano tre vettori monodimensionali \mathbf{A} , \mathbf{IA} e \mathbf{JA} .

\mathbf{A} è un vettore di reali che contiene, ordinati per righe, tutti gli elementi di \mathbf{E} e \mathbf{D} diversi da zero.

JA è un vettore di interi che contiene l'indice di colonna degli elementi corrispondenti immagazzinati in A

IA contiene, infine, i puntatori agli elementi diagonali immagazzinati in A .

Se il numero di elementi non nulli di $E + D$ è $nsize$, allora la dimensione di A e JA sarà $nsize$, mentre quella di IA sarà N .

Gli elementi a_{ij} con $j \leq i$ dell' i -ma riga di A sono immagazzinati nelle posizioni di A che vanno dall'indice $IA(i-1)+1$ all'indice $IA(i)$

In figura [] è mostrato come possa essere immagazzinata attraverso questo schema la matrice di figura [.a]:

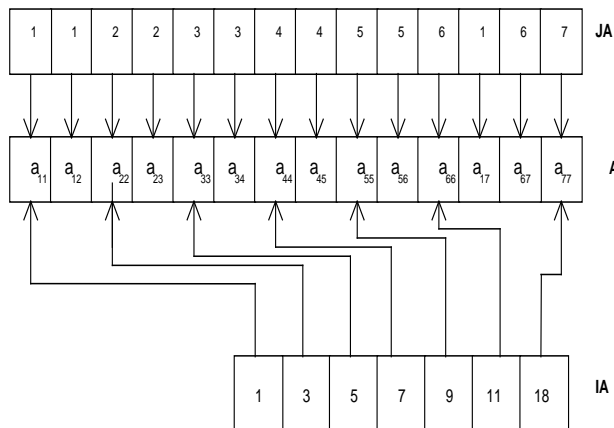


Fig.[] Immagazzinamento di una matrice con il metodo di sparsità generale

Si nota immediatamente che questo schema non permette né un accesso random, né un accesso per colonne agli elementi di A .

Va infine notato che se -come spesso succede- bisogna risolvere diversi sistemi lineari ai quali sono associate matrici sparse della medesima struttura ed eventualmente con diversi coefficienti, allora le sole informazioni contenute in A debbono essere aggiornate mentre i vettori JA e IA si mantengono costanti e possono pertanto essere calcolati preventivamente al processo solutivo. Ovviamente se lo schema risolutivo che si intende adottare prevede la fattorizzazione della matrice A , non si può prescindere nella generazione di JA

e IA dal portare in conto anche gli elementi che pur nulli per A non lo saranno a causa del fill in U nella matrice U che viene generata dalla fattorizzazione: come si era precedentemente anticipato, in questi casi la fattorizzazione simbolica diventa indispensabile.

Metodi di risoluzione dei sistemi lineari

a) Metodi diretti [1]

I metodi diretti si basano essenzialmente sull'idea dell'eliminazione di Gauss, che consiste nel manipolare una equazione del sistema in modo che una delle incognite venga espressa esplicitamente in funzione delle altre; sostituendo l'espressione ricavata in tutte le equazioni rimanenti, si riduce di uno l'ordine del sistema. Iterando il procedimento, il problema originario viene ricondotto ad un sistema di una equazione in una incognita, banale da invertire. Le altre componenti della soluzione possono successivamente essere calcolate attraverso una procedura di sostituzione a ritroso.

Da un punto di vista numerico esistono due concetti essenziali che possono determinare il successo o il fallimento di un processo risolutivo: il condizionamento della matrice e la stabilità del metodo.

Dalle nozioni di algebra lineare è noto che per un sistema lineare

$$\mathbf{Ax}=\mathbf{b} \quad (1)$$

di n equazioni indipendenti in n incognite la soluzione esiste ed è univocamente determinata. Da un punto di vista numerico le cose possono andare in maniera diversa. E' difatti possibile che, causa gli arrotondamenti di macchina, il sistema, originariamente non singolare, possa diventarlo. Questa situazione va sotto il nome di malcondizionamento della matrice. Una strategia per ovviare può essere quella di trovare un'espressione equivalente del

sistema di partenza, ma più lontana dalla condizione di singolarità, attraverso la moltiplicazioni con matrici opportune. Questa trasformazione corrisponde fisicamente a cambiare le unità di misura delle grandezze in gioco, e matematicamente a riscalarare tanto i coefficienti della matrice quanto i termini noti e le variabili incognite, in modo tale che il sistema risultante sia meglio condizionato. E' importante rimarcare il fatto che, benchè matematicamente equivalenti, differenti formulazioni possono dare risultati numerici differenti.

Un algoritmo basato sul metodo di sostituzione, che elimini in ordine la prima equazione e successivamente le altre, si arresterà nel caso in cui uno dei coefficienti della diagonale principale sia nullo. Sono comuni inoltre situazioni in cui l'algoritmo, a causa della precisione finita della macchina, produce soluzioni completamente errate anche se è chiamato ad operare su sistemi ben condizionati. Per ovviare a questo problema si usa una tecnica che a sotto il nome di pivoting. L'idea di base è quella di scegliere a ciascun passo k in maniera oculata la variabile da eliminare e l'equazione in cui eliminare la variabile scelta. Tra i vari criteri che possono guidare in questa scelta si cita quello adoperato nel cosiddetto pivoting scalato. Definito col termine di grandezza della riga, d_i , l'elemento in valore assoluto maggiore che si trova sulla i -ma riga, al generico passo k dell'eliminazione si assume come equazione pivotale quella tra le $n-k$ rimanenti che ha il coefficiente dell'incognita x_k , che si vuole eliminare, più grande in modulo quando rapportato alla grandezza della riga. In altre parole si cerca l'indice p compreso tra k e n per il quale:

$$\frac{a_{pk}}{d_p} \geq \frac{a_{ik}}{d_i} \quad \forall i = k, \dots, n$$

Annotiamo che su matrici simmetriche definite positive non è necessario, almeno teoricamente, usare la strategia del pivoting in quanto i minori principali sono tutti diversi da zero.

Durante la soluzione di sistemi lineari ben difficilmente si opera direttamente sulla matrice di partenza. Molto più frequentemente, infatti, si preferisce decomporla in matrici di struttura più semplice -diagonale o triangolare-, per le quali esistono implementazioni del procedimento di eliminazione di Gauss particolarmente efficienti.

Tra le varie possibili strategie di fattorizzare una generica matrice **A** citiamo:

- 1) La decomposizione **LU** dove **L** è una matrice triangolare inferiore, e **U** una matrice triangolare superiore.
- 2) La decomposizione **LDM^T**, dove **L** ed **M** sono matrici triangolari inferiori con elementi unitari sulla diagonale principale e **D** è una matrice diagonale.
- 3) Le decomposizioni di Doolittle e di Crout, analoghe alla decomposizione **LU**, con la differenza che **L** nel metodo di Doolittle ed **U** nel metodo di Crout presentano elementi tutti unitari sulla diagonale principale.
- 4) La decomposizione, possibile solo se **A** è simmetrica, **LDL^T**, analoga alla 2) con **M=L**.
- 5) La decomposizione **RR^T**, detta di Cholesky, possibile se **A** è simmetrica e definita positiva, con **R** matrice triangolare inferiore con elementi tutti positivi sulla diagonale.

Esistono due grossi limiti all'utilizzo dei metodi diretti:

1) Le decomposizioni e il pivoting in generale distruggono la struttura originaria della matrice, e ciò risulta particolarmente fastidioso per matrici di tipo sparso. Questa osservazione non è valida per le matrici bandate a dominanza diagonale e per le matrici bandate simmetriche e definite positive che non necessitano dell'operazione di pivoting e possono essere decomposte con algoritmi che conservano la struttura bandata anche nelle matrici fattori,

2) Un limite ulteriore all'impiego dei metodi diretti è l'ingente costo computazionale associato al processo di fattorizzazione della matrice. Per matrici piene esso risulta proporzionale a N^3 (in particolare, se si adotta la fattorizzazione di Cholesky vale $N^3/3$), dove N è il numero delle incognite, che,

nei problemi FEM di interesse pratico, risulta tipicamente dell'ordine di $1E3$ - $1E5$. Una stima del tempo occorrente per risolvere un sistema lineare si ottiene facilmente tenendo conto che la velocità dei minicomputers, le macchine comunemente adoperate nei laboratori di calcolo scientifico è attualmente dell'ordine delle decine di Megaflops (1 Megaflop corrisponde ad un milione di moltiplicazioni per secondo). Il tempo di elaborazione richiesto per l'inversione di un problema di interesse pratico può quindi essere dell'ordine dei mesi, il che è, ovviamente, assolutamente inaccettabile. Se le matrici sono di tipo sparso, l'impiego di metodi di banda rende il costo del processo di fattorizzazione proporzionale a $m_A N^2$, dove m_A è l'ampiezza di banda della matrice che tipicamente è di ordine 100; seppure l'impiego di metodi di inviluppo renda il costo computazionale della fattorizzazione ancora minore, questa tuttavia conserva un peso notevole nell'economia complessiva del processo solutivo. Risulta per questo motivo conveniente adoperare i metodi diretti nei casi in cui l'ordine del sistema non sia particolarmente elevato e/o quando il problema richieda di risolvere un grande numero di sistemi di equazione nei quali il termine noto cambi mentre la matrice dei coefficienti si conservi costante; in casi diversi possono essere preferiti schemi iterativi di soluzione.

[V.Comincioli, *Analisi Numerica*, Mc Graw-Hill,1990

[V.Comincioli, *Fortran 77.Introduzione e applicazioni numeriche*,Mc Graw-Hill,1990

[G.F.Carey-J.T.Oden, *Finite elements vol.III*, Prentice-Hall 1984

[W.H.Press et al., *Numerical Recipes*, Cambridge University Press,1986

b) Metodi iterativi

Per matrici dotate di sparsità non strutturale e di dimensione elevata, un'alternativa ai metodi diretti è costituita dagli schemi di soluzione iterativi che lasciano inalterata la struttura originale della matrice e non richiedono il processo di fattorizzazione. Essendo, tuttavia, in questo caso la soluzione

ottenuta come limite di una successione, per essere un'alternativa valida ai metodi diretti, essi hanno bisogno di opportune tecniche di accelerazione.

L'idea comune a tutti i metodi iterativi è la seguente.

1) Si dà una stima del vettore soluzione $x^{(0)}$ del sistema $\mathbf{Ax}=\mathbf{b}$, dove \mathbf{A} è una matrice sparsa di ordine N non singolare.

2) Si costruisce una successione di soluzioni $x^{(k)}$ attraverso la successiva risoluzione di sistemi lineari più semplici.

In forma generale consideriamo una decomposizione della matrice \mathbf{A} della forma: $\mathbf{A}=\mathbf{M}-\mathbf{N}$ con \mathbf{M} non singolare.

Si ha allora:
$$\mathbf{Ax}=\mathbf{b} \Leftrightarrow \mathbf{Mx}=\mathbf{Nx} + \mathbf{b} \quad (2)$$

da cui il procedimento iterativo:
$$\mathbf{Mx}^{(k+1)}=\mathbf{Nx}^{(k)} + \mathbf{b} \quad (3)$$

La matrice
$$\mathbf{B} = \mathbf{M}^{-1}\mathbf{N} = \mathbf{M}^{-1}(\mathbf{M}-\mathbf{A}) = \mathbf{I} - \mathbf{M}^{-1}\mathbf{A} \quad (4)$$

è detta *matrice di iterazione*; essa individua un particolare metodo e le sue proprietà sono fondamentali per la qualità e la rapidità della convergenza del processo iterativo.

Facendo riferimento alla decomposizione $\mathbf{A}=\mathbf{D}-\mathbf{E}-\mathbf{F}$ (vedi fig.[]):

$$\mathbf{A} = \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & \mathbf{D} & & & \\ & & & & & \\ & & -\mathbf{E} & & & \\ & & & \ddots & & \\ & & & & & \ddots \end{bmatrix} \quad (5)$$

Fig.[]: Decomposizione adottata per la matrice \mathbf{A} negli schemi di Jacobi e Gauss-Seidel

citiamo alcuni tra gli schemi iterativi più comunemente impiegati:

-Il metodo di Jacobi, o *delle sostituzioni simultanee* ove si assume

$$\mathbf{M}=\mathbf{D}; \quad \mathbf{N}=\mathbf{E}+\mathbf{F}; \quad (6)$$

e quindi:
$$\mathbf{B}_j = \mathbf{D}^{-1}(\mathbf{E}+\mathbf{F}) = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A} \quad (7)$$

-Il metodo di Gauss Seidel, o *delle sostituzioni successive* ove si assume

$$\mathbf{M}=\mathbf{D}-\mathbf{E}; \quad \mathbf{N}=\mathbf{F}; \quad (8)$$

$$\text{e quindi: } \mathbf{B}_j = (\mathbf{D}-\mathbf{E})^{-1}\mathbf{F} = (\mathbf{I} - \mathbf{D}^{-1}\mathbf{E})^{-1}\mathbf{D}^{-1}\mathbf{F} \quad (9)$$

Con l'obiettivo di accelerare il processo di convergenza i metodi precedenti possono essere modificati introducendo un parametro reale ω e definendo la componente i -ma della soluzione al $(k+1)$ -mo passo iterativo come:

$$x_i^{(k+1)} = \omega x_i^{(k+1/2)} + (1-\omega) x_i^{(k)} \quad (10)$$

In questo modo, partendo dal metodo di Jacobi si perviene al seguente schema:

$$\mathbf{x}^{(k+1)} = (\mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}) \mathbf{x}^{(k)} + \omega \mathbf{D}^{-1}\mathbf{b} \quad (11)$$

Mentre partendo dal metodo di Gauss-Seidel si giunge al metodo di rilassamento (detto anche SOR) in cui si assume:

$$\mathbf{M} = \left(\frac{\mathbf{D}}{\omega} - \mathbf{E} \right); \quad \mathbf{N} = \left(\frac{1}{\omega} - 1 \right) \mathbf{D} + \mathbf{F}; \quad (12)$$

a cui corrisponde la seguente matrice di iterazione:

$$\mathbf{B}_\omega = (\mathbf{D}-\omega\mathbf{E})^{-1} [(1-\omega) \mathbf{D} + \omega \mathbf{F}] \quad (13)$$

Circa la convergenza di questi schemi iterativi della successione $\mathbf{x}^{(k)}$ alla soluzione \mathbf{x} valgono i seguenti teoremi :

-Se \mathbf{A} è una matrice di ordine N a stretta predominanza diagonale, allora i metodi di Jacobi e di Gauss-Seidel sono convergenti.

-Se \mathbf{A} è una matrice hermitiana definita positiva, allora il metodo di rilassamento converge se e solo se $0 < \omega < 2$.

Per quanto riguarda la velocità di convergenza, questa dipende in maniera critica, e questo è il più grande inconveniente del metodo SOR, dalla scelta del fattore di rilassamento ω . Young, a questo proposito, ha mostrato [1] come la scelta ottimale sia:

$$\omega_b = \frac{2}{1 + (1 - z^2)^{\frac{1}{2}}} \quad (14)$$

dove z è il più grande autovalore di $\mathbf{D}^{-1}(\mathbf{E}+\mathbf{F})$.

[1] **D.M.Young**, *Iterative Solution of Large Linear Systems*, Academic Press, New York, **1971**

Metodo del gradiente coniugato

Un altro schema iterativo che può essere utilizzato vantaggiosamente nella soluzione di problemi FEM è quello del Gradiente Coniugato. Questo metodo è stato introdotto per la prima volta da Hestenes e Stiefel nel 1952 [2], ed ha il vantaggio (puramente teorico) di risolvere esattamente un sistema di ordine N , in al più N iterazioni (in pratica, causa gli errori di arrotondamento, questa proprietà viene persa, ed inoltre N iterazioni sono un costo computazionale usualmente troppo alto da pagare).

Il metodo del gradiente coniugato sfrutta il fatto che risolvere il sistema (1) è equivalente a minimizzare il funzionale quadratico:

$$\mathbf{Q}(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{b} \quad (15)$$

o in maniera equivalente il funzionale:

$$\mathbf{E}(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{A} (\mathbf{x} - \bar{\mathbf{x}}) = \mathbf{e}^T \mathbf{A} \mathbf{e}$$

dove $\bar{\mathbf{x}}$ è la soluzione esatta del sistema (1), ed $\mathbf{e} = \mathbf{x} - \bar{\mathbf{x}}$.

La minimizzazione di (16) viene condotta per via iterativa scegliendo al passo k una direzione $\mathbf{p}_k \neq 0$ e uno scalare α_k in maniera che, posto

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (16)$$

si abbia $\mathbf{E}(\mathbf{x}_{k+1}) \leq \mathbf{E}(\mathbf{x}_k)$.

Fissata la direzione \mathbf{p}_k e definito il residuo $\mathbf{r}_k(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}_k = \mathbf{A}(\bar{\mathbf{x}} - \mathbf{x}_k)$ si dimostra che la scelta ottimale per α_k è

$$\alpha_k = \frac{(\mathbf{r}_k, \mathbf{p}_k)}{(\mathbf{A}\mathbf{p}_k, \mathbf{p}_k)} \quad (17)$$

Per quel che concerne la scelta delle direzioni \mathbf{p}_k ricordiamo innanzitutto che due vettori \mathbf{x} e \mathbf{y} si dicono coniugati rispetto all'applicazione bilineare \mathbf{A} se vale la relazione $(\mathbf{A}\mathbf{x}, \mathbf{y}) = 0$.

Il nome dell'algoritmo è così legato al criterio di ortogonalità rispetto al prodotto scalare $(\mathbf{A}\mathbf{p}_i, \mathbf{p}_k) \forall i=1, k-1$ col quale, al k -mo passo di iterazione, viene scelta la direzione lungo le quali effettuare la minimizzazione. Da un punto di vista pratico se si adotta per α_k la scelta (17), la direzione \mathbf{p}_k è determinata attraverso le relazioni:

$$\mathbf{p}_k = \mathbf{r}_k + \beta_k \mathbf{p}_{k-1}; \quad (18) \quad \text{con} \quad \beta_k = -\frac{(\mathbf{A}\mathbf{p}_{k-1}, \mathbf{r}_k)}{(\mathbf{A}\mathbf{p}_{k-1}, \mathbf{p}_{k-1})}; \quad (19)$$

Per quanto riguarda il numero di operazioni richieste si vede che il costo essenziale è dovuto al calcolo di $\mathbf{A}\mathbf{p}$, che per matrici piene è pari a N^2 mentre per matrici sparse non strutturate è uguale a $nsize$ (numero degli elementi non nulli della matrice). Se il numero di iterazioni è N , il costo totale per matrici piene risulta pari a N^3 superiore a quello di $N^3/3$ richiesto nel processo di fattorizzazione di Cholesky. Appare chiaro dunque che il metodo diventa interessante quando la matrice è sparsa e/o quando il numero di iterazioni necessarie è decisamente inferiore a N . A questo scopo risulta molto spesso di grande efficacia l'accorgimento di preconditionare la matrice \mathbf{A} attraverso la moltiplicazione con una matrice non-singolare \mathbf{R} . Si può infatti osservare che se la matrice \mathbf{R} viene scelta in maniera opportuna l'algoritmo del gradiente coniugato convergerà molto più velocemente alla soluzione del sistema equivalente

$$\mathbf{R}^{-1}\mathbf{A}(\mathbf{R}^{-1})^T (\mathbf{R}^T\mathbf{x}) = (\mathbf{R}^{-1}\mathbf{b}) \Leftrightarrow \mathbf{R}^{-1}\mathbf{Ax} = (\mathbf{R}^{-1}\mathbf{b}) \Leftrightarrow \mathbf{M}^{-1}\mathbf{Ax} = (\mathbf{M}^{-1}\mathbf{b}) = \mathbf{s} \quad (20)$$

ove, essendo \mathbf{R} non singolare, $\mathbf{M}=\mathbf{RR}^T$ è simmetrica e definita positiva.

Il calcolo di \mathbf{s} deve essere eseguito ad ogni iterazione e non deve naturalmente passare per l'inversione di \mathbf{M} . E' pertanto essenziale scegliere \mathbf{M} in modo tale che il sistema lineare $\mathbf{Ms}=\mathbf{b}$ possa essere risolto in maniera economica. Tra le varie possibili scelte di \mathbf{M} si ricordano (vedi fig []):

- $\mathbf{M}=\mathbf{D}$ che dà origine al *metodo del gradiente coniugato di Jacobi*.
- $\mathbf{M} = \frac{1}{\omega(2-\omega)} (\mathbf{D} - \omega\mathbf{F}^T) \mathbf{D}^{-1} (\mathbf{D} - \omega\mathbf{E}^T)$, usata nel metodo SSOR, ove, a

differenza del metodo SOR la scelta di ω non è più critica.

- Un approccio leggermente differente è quello di scegliere \mathbf{R} in modo che sia un'approssimazione della matrice triangolare inferiore \mathbf{L} che risulta dalla fattorizzazione di Cholesky di \mathbf{A} . Se per l'immagazzinamento di \mathbf{A} si è adoperato un metodo di matrici sparse, allora può essere conveniente calcolare \mathbf{R} attraverso una fattorizzazione parziale di Cholesky di \mathbf{A} con un metodo di eliminazione Gaussiana, in cui il *fill in* è trascurato, oppure portato in conto solo in zone limitate della matrice fattorizzata (ad esempio all'interno di un fissato numero di diagonali localizzate attorno la diagonale principale). Allo scopo di limitare l'insorgere di mal-condizionamenti per effetto degli arrotondamenti numerici e di evitare il fallimento del processo di fattorizzazione, può risultare conveniente scalare la matrice \mathbf{A} in modo da riportare i termini diagonali a valori unitari e, eventualmente, moltiplicare i termini fuori diagonale per il fattore $1/(1+\alpha)$, ove α è un numero reale positivo.

(Ref. Numerical recipes, Rapporto di doc. progr. di ottimizzazione, Dorny, Comincioli, Bulirsch...)

3.2.2) Risoluzione di sistemi di equazioni non lineari

Generalità

Mentre per i sistemi di equazioni lineari la soluzione può essere trovata tanto con metodi diretti quanto per via iterativa, nei sistemi di equazioni non lineari il vettore delle radici viene, in ogni caso, determinato per via iterativa e, dunque, diventano determinanti nella scelta del metodo risolutivo le proprietà di applicabilità e di convergenza.

A differenza dei sistemi lineari, in questo tipo di problema non è garantita né l'esistenza (almeno nel campo reale) né l'unicità della soluzione: infatti, poichè i sistemi non-lineari sono usualmente posti nella forma implicita, $\mathbf{F}(\mathbf{x})=\mathbf{0}$, vale, a proposito dell'esistenza della soluzione, il teorema del Dini che nel caso di funzioni continue dà qualche speranza, ma non la certezza, di poter soddisfare simultaneamente N equazioni in N incognite. Inoltre, nei problemi multi-dimensionali, anche nel caso in cui sia dimostrabile l'esistenza di una soluzione, la convergenza ad essa può dipendere criticamente dalla scelta del punto di partenza. Presentate subito le difficoltà che si possono incontrare nel risolvere un generico sistema non-lineare, va subito aggiunto che le cose si semplificano in maniera significativa se il sistema non-lineare proviene dalla minimizzazione di un funzionale convesso. In questo caso, infatti, le equazioni del sistema non sono funzioni indipendenti, ma le componenti del vettore gradiente del funzionale. Esse devono pertanto obbedire alle condizioni di integrabilità che sono altamente restrittive. In parole semplici è sempre possibile trovare il minimo di un funzionale convesso scendendo lungo la direzione più ripida e questo ci rassicura circa l'esistenza di uno zero del vettore gradiente.

Problemi non-lineari mono-dimensionali

Il metodo più semplice per localizzare le radici di una funzione mono-dimensionale non-lineare e continua, $f(x)$, consiste nel tabellare la funzione con passo opportuno e di individuare due valori di x , che denoteremo con i simboli a e b , per i quali la funzione $f(x)$ assuma valori di segno opposto. Per il

teorema degli zeri delle funzioni continue la radice dell'equazione deve essere compresa all'interno dell'intervallo aperto $]a,b[$. Una volta localizzato l'intervallo si può applicare uno tra i diversi algoritmi monodimensionali di ricerca presenti in letteratura. Di seguito ne vengono presentati alcuni, insieme con le loro caratteristiche di convergenza. Per tutti l'idea di base comune è quella di ridurre l'ampiezza dell'intervallo fino a che non sia soddisfatto un test di tolleranza sulla soluzione ottenuta ($|b_k - a_k| < \epsilon_1$ oppure $f(x_k) < \epsilon_2$) oppure non sia stato raggiunto il numero massimo di iterazioni consentite dall'utente.

- L'*algoritmo di bisezione* genera una successione di intervalli (a_k, b_k) con:

$$]a_k, b_k[\subset]a_{k-1}, b_{k-1}[\quad \text{e} \quad |b_k - a_k| = \frac{1}{2} |b_{k-1} - a_{k-1}|$$

Una volta raggiunto il criterio di arresto, si assume come soluzione la media degli estremi cui si è ristretto l'intervallo durante l'ultima iterazione: $x_{k+1} = \frac{a_k + b_k}{2}$; che comporta un'errore pari a: $e_k = \frac{b_k - a_k}{2}$. Il tasso

di convergenza viene determinato come $\frac{|e_{k+1}|}{|e_k|} \approx \frac{1}{2}$ ed è chiaramente lineare.

- Mentre l'algoritmo di bisezione tiene conto della sola informazione sul segno, l'algoritmo che va sotto il nome di *regula falsi* utilizza anche i valori che la funzione assume negli estremi dell'intervallo. Ad ogni nuova iterazione il valore di x viene determinato attraverso interpolazione lineare:

$$x_{k+1} = \frac{a_k f(b_k) + b_k f(a_k)}{f(b_k) - f(a_k)} \quad (21)$$

Si può dimostrare che questo metodo, seppure più veloce del metodo di bisezione, presenta una convergenza di carattere superlineare, ma la stima del suo esatto ordine è piuttosto complicata.

- Una variante del metodo precedente è rappresentata dal *metodo delle secanti*, nel quale per il calcolo dell'interpolazione lineare si utilizzano gli ultimi due punti calcolati. Questo metodo ha un tasso di convergenza maggiore di entrambi i metodi precedenti e pari al cosiddetto "numero d'oro", 1.618..., ma presenta lo svantaggio di non fornire necessariamente un'intervallo che

contenga la radice. Per funzioni non sufficientemente regolari la convergenza non è pertanto garantita.

Sistemi multi-dimensionali di equazioni non-lineari

I metodi illustrati precedentemente sono tutti concepiti per trattare il solo caso monodimensionale. Nel caso in cui sia richiesto di trovare le radici di un sistema pluri-dimensionale $\mathbf{F}(\mathbf{x})=[f_1(\mathbf{x}),\dots,f_N(\mathbf{x})]$, in cui il vettore delle incognite $\mathbf{x}\in\mathbb{R}^N$, bisogna necessariamente ricorrere all' algoritmo, anch'esso di carattere iterativo, detto di Newton-Raphson. In questo schema è necessario saper valutare nel generico punto \mathbf{x} oltre che le funzioni $f_i(\mathbf{x})$ anche lo Jacobiano $\mathbf{A}(\mathbf{x})$ di $\mathbf{F}(\mathbf{x})$. L'idea di base dell'algoritmo consiste nell'espandere, alla $k+1$ -ma iterazione, $\mathbf{F}(\mathbf{x})$ in serie di Taylor. Indicando con $\mathbf{H}(\mathbf{x})$ la matrice Hessiana di $\mathbf{F}(\mathbf{x})$ si ha pertanto:

$$\mathbf{F}(\mathbf{x}_k+d\mathbf{x})\approx\mathbf{F}(\mathbf{x}_k)+\mathbf{A}(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}_k}d\mathbf{x}+\frac{1}{2}\mathbf{H}(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}_k}d\mathbf{x}^2; \quad \text{con } \mathbf{x}\in[\mathbf{x}_k,\mathbf{x}_k+d\mathbf{x}] \quad (21)$$

Trascurando i termini del secondo ordine la (21) diventa:

$$\mathbf{F}(\mathbf{x}_k+d\mathbf{x})\approx\mathbf{F}(\mathbf{x}_k)+\mathbf{A}(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}_k}d\mathbf{x}=\mathbf{0} \quad (22)$$

$$\mathbf{x}_{k+1}=\mathbf{x}_k+d\mathbf{x} \quad (23)$$

Il sistema lineare (22) può essere risolto utilizzando tra gli algoritmi presentati nel paragrafo precedente quello che risulta più vantaggioso in base alla struttura di $\mathbf{A}(\mathbf{x})$.

La potenza del metodo di Newton giace nel fatto che il suo tasso di convergenza è quadratico, vale a dire che *nelle vicinanze di una radice* il numero di cifre significative esatte approssimativamente raddoppia ad ogni step.

Anche questo metodo presenta tuttavia degli inconvenienti. In primo luogo la sua applicabilità dipende dal fatto che lo Jacobiano sia continuo e diverso da

zero in prossimità delle radici. Inoltre, lontano dalle radici, dove i termini dello sviluppo di ordine superiore acquistano importanza rilevante, il metodo di Newton può dare grosse correzioni del tutto prive di significato. Infine, soprattutto nel caso in cui le funzioni $f_i(\mathbf{x})$ siano costruite per interpolazione non lineare di valori tabellati possono insorgere patologie che determinano il completo fallimento dello schema solutivo.

Un modo per ridurre il rischio di insorgenza di situazioni patologiche e per migliorare ulteriormente il tasso di convergenza è quello di assumere per la soluzione al passo $k+1$, anziché il valore fornito dalla (23), quello determinato dall'espressione:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \mathbf{dx} \quad (23)$$

dove α è l'incognita scalare del problema monodimensionale non-lineare:

$$\mathbf{G}(\mathbf{x}_k + \alpha \mathbf{dx}) = \mathbf{F}(\mathbf{x}_k + \alpha \mathbf{dx}) \times \mathbf{dx} = \mathbf{0} \quad (24)$$

che può essere risolto con il metodo *regula falsi*, oppure con il *metodo delle secanti*.

3.3) Generalità sui metodi FEM. Edges elements e loro proprietà . Elementi di libreria: Tetraedri, prismi, esaedri a 12 e a 36 lati : p-refinement.

3.3) Implementazione numerica del metodo della minimizzazione dell'errore costitutivo. Metodo FEM con edges elementi nello spazio , metodo delle differenze finite nel tempo. Problemi concernenti l'uso delle Penalty functions per imporre le boundary conditions. Trattamento delle simmetrie. Imposizione delle condizioni iniziali direttamente nei punti di gauss spaziali o attraverso una fase di pre-polarizzazione. Storia. Pesi differenti per le relazioni costitutive elettriche e magnetiche. Scelta dei pesi per ottenere lo splitting delle incognite.

Discretizzazione numerica dei potenziali vettori incogniti

Per il metodo basato sulla minimizzazione dell'errore costitutivo viene proposta un'implementazione basata sugli Elementi Finiti nel dominio spaziale e sulle differenze finite nel dominio temporale.

I potenziali vettori a 3 componenti $\mathbf{A}(\mathbf{x},t)$ e $\mathbf{W}(\mathbf{x},t)$ sono approssimati usando le funzioni di forma spaziale associate agli edge-elements, che permettono la gestione automatica di disomogeneità e discontinuità spaziali.

$$\mathbf{A}(\mathbf{x},t) = \sum_{i=1}^{Nedges} A_i(t) \mathbf{N}_i(\mathbf{x}) \quad (1a)$$

$$\mathbf{W}(\mathbf{x},t) = \sum_{i=1}^{Nedges} W_i(t) \mathbf{N}_i(\mathbf{x}) \quad (1b)$$

dove $Nedges$ è il numero dei lati della mesh usata per rappresentare il dominio del problema.

La soluzione è trovata partizionando l'intervallo di tempo di interesse $(0, T)$ in sotto-intervalli nei quali è assunta una interpolazione lineare. Per il generico sotto-intervallo $[t_k, t_k + dt_k]$ si ha pertanto:

$$A_i(t) = \frac{a_{i,k+1} - a_{i,k}}{\Delta t_k} (t - t_k) + a_{i,k} \quad (2a)$$

$$W_i(t) = \frac{w_{i,k+1} - w_{i,k}}{\Delta t_k} (t - t_k) + w_{i,k} \quad (2b)$$

Espressione del funzionale di errore nelle incognite discretizzate

Con questa scelta per la discretizzazione dei potenziali vettoriali incogniti il problema si riduce alla soluzione di una successione di sotto-problemi, ciascuno dei quali è definito all'interno di un intervallo elementare. Le condizioni finali incognite di un time-step sono usate come condizioni iniziali

per il time-step successivo, il che equivale ad imporre la continuità nel tempo di \mathbf{A} e \mathbf{W} .

La soluzione per ogni sotto-problema può così essere ottenuta attraverso la minimizzazione di un errore globale in $V \times [t_k, t_{k+1}]$, definito attraverso la (2-13) e di una conveniente espressione dell'errore locale, come ad esempio la (2-12). In questa tesi tuttavia si è preferito usare la seguente forma:

$$\lambda(\mathbf{x}, t) =$$

$$\alpha_H \left[\int_{\mathbf{H}^*(\mathbf{x}, t)}^{\mathbf{H}(\mathbf{x}, t)} \mathbf{f}_B(\mathbf{h}, \mathbf{x}, t) \cdot d\mathbf{h} + \int_{\mathbf{B}^*(\mathbf{x}, t)}^{\mathbf{B}(\mathbf{x}, t)} \mathbf{f}_H(\mathbf{b}, \mathbf{x}, t) \cdot d\mathbf{b} + \mathbf{H}^*(\mathbf{x}, t) \cdot \mathbf{B}^*(\mathbf{x}, t) - \mathbf{H}(\mathbf{x}, t) \cdot \mathbf{B}(\mathbf{x}, t) \right] +$$

$$\alpha_E \left[\int_{\mathbf{D}_T^*(\mathbf{x}, t)}^{\mathbf{D}_T(\mathbf{x}, t)} \mathbf{f}_E(\mathbf{d}_T, \mathbf{x}, t) \cdot d\mathbf{d}_T + \int_{\mathbf{E}^*(\mathbf{x}, t)}^{\mathbf{E}(\mathbf{x}, t)} \mathbf{f}_{D_T}(\mathbf{e}, \mathbf{x}, t) \cdot d\mathbf{e} + \mathbf{E}^*(\mathbf{x}, t) \cdot \mathbf{D}_T^*(\mathbf{x}, t) - \mathbf{E}(\mathbf{x}, t) \cdot \mathbf{D}_T(\mathbf{x}, t) \right]$$

$$\forall t \in (t_k, t_{k+1}) \quad (3)$$

dove $\mathbf{f}_H(\mathbf{B}, \mathbf{x}, t)$ è l'inverso della caratteristica magnetica rispetto a \mathbf{B} , per \mathbf{x} e t fissati; $(\mathbf{H}^*, \mathbf{B}^*)$ è una coppia che soddisfa la relazione costitutiva magnetica (2-9a); $\mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t)$ è il mapping non-lineare di $\mathbf{E}(\mathbf{x}, t)$ su $\mathbf{D}_T(\mathbf{x}, t)$ definito dalla relazione (2-9b); $\mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t)$ è la sua inversa rispetto a $\mathbf{E}(\mathbf{x}, t)$ ed, infine, $(\mathbf{E}^*, \mathbf{D}_T^*)$ è una coppia che soddisfa la relazione costitutiva elettrica (2-9b).

Si noti che la relazione (2-9b) è dispersiva nel dominio del tempo, per cui un legame di tipo algebrico tra i valori istantanei di $\mathbf{E}(\mathbf{x}, t)$ e $\mathbf{D}_T(\mathbf{x}, t)$ è in generale possibile solo quando i potenziali \mathbf{A} e \mathbf{W} sono scelti lineari in (t_k, t_{k+1}) .

Ad esempio nel caso particolare di condizioni iniziali nulle, di assenza di sorgenti nel dominio e di linearità, isotropia, tempo-invarianza dei mezzi si ha:

$$\mathbf{B} = \mu \mathbf{H} \quad (4a)$$

$$\mathbf{J} = \sigma \mathbf{E} \quad (4b)$$

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (4c)$$

da cui, assunto $\mathbf{D}_{T_0}(\mathbf{x}) = 0$, si ricava:

$$\mathbf{f}_H(\mathbf{B}, \mathbf{x}, t) = \mathbf{B} / \mu \quad (5a)$$

$$\mathbf{f}_B(\mathbf{H}, \mathbf{x}, t) = \mu \mathbf{H} \quad (5b)$$

$$\mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t) = [(t - t_K)\sigma + \varepsilon] \mathbf{E} - \sigma \mathbf{A}(t_K) \quad (5c)$$

$$\mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t) = [\mathbf{D}_T + \sigma \mathbf{A}(t_K)] / [(t - t_K)\sigma + \varepsilon] \quad (5d)$$

Come coppie di campi che soddisfano rispettivamente le relazioni costitutive elettrica e magnetica si possono scegliere:

$$\begin{aligned} \mathbf{E}^* &= 0 \\ \mathbf{D}_T^* &= -\sigma \mathbf{A}(t_K) \\ \mathbf{H}^* &= 0 \\ \mathbf{B}^* &= 0 \end{aligned}$$

In questo modo l'errore locale diventa:

$$\begin{aligned} \lambda(\mathbf{x}, t) = & \alpha_H \{ [\mu \mathbf{H}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, t)] \cdot [\mathbf{H}(\mathbf{x}, t) - \mathbf{B}(\mathbf{x}, t) / \mu] \} / 2 + \\ & \alpha_E \{ [\mathbf{D}_T(\mathbf{x}, t) - \mathbf{f}_{D_T}(\mathbf{E}(\mathbf{x}, t), \mathbf{x}, t)] \cdot [\mathbf{E}(\mathbf{x}, t) - \mathbf{f}_E(\mathbf{D}_T(\mathbf{x}, t), \mathbf{x}, t)] \} \quad (6) \end{aligned}$$

$$\forall t \in (t_k, t_{k+1})$$

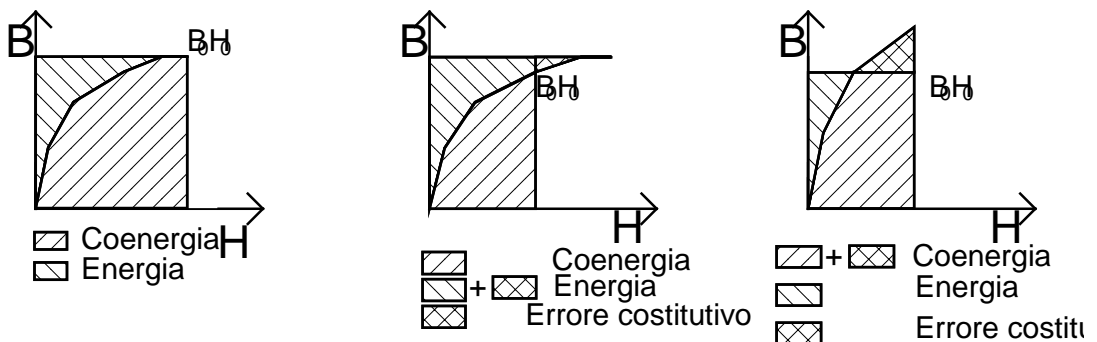


Fig.[] Il funzionale di errore definito dalla (3) per relazioni costitutive crescenti è sempre positivo e vale 0 solo se la relazione costitutiva è verificata esattamente.

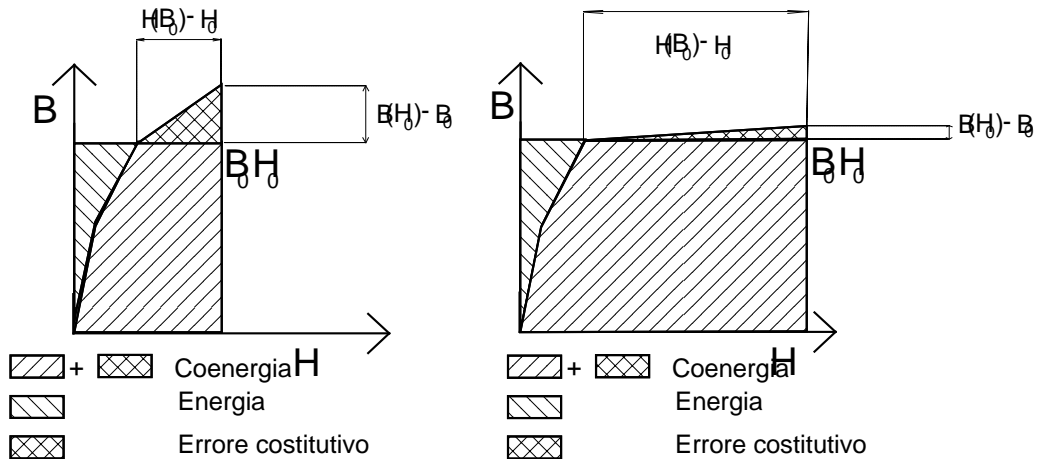


Fig.[] In questo caso le relazioni (2.12) e (3) sono equivalenti.

Fig.[] Situazione patologica per la relazione (2.12), ma non per la (3)

Il motivo della scelta sulla definizione dell'errore locale appare evidente se si fa riferimento al diagramma della relazione costitutiva $B-H$ riportato in figura (). Sulla stessa figura è stato evidenziato il significato geometrico degli errori associati alle relazioni (3) e (2-12). Si osservi che l'errore associato alla (2-12) è una distanza, mentre quello associato alla relazione (3) rappresenta fisicamente un'area. Questa scelta permette da un lato di evitare situazioni patologiche (vedi figura ()), e dall'altro di bilanciare equamente il peso dell'errore sulle incognite. Si può notare a questo proposito l'inerente simmetria della (3) rispetto alle soluzioni nelle variabili A e W .

Formulazione discreta del problema di minimizzazione del funzionale di errore

La composizione delle equazioni governanti inizia attraverso la sostituzione nella (3) dei campi con le espressioni (2-11) dei potenziali vettori. I potenziali vettori sono descritti attraverso le posizioni (1). Il funzionale (3) è trasformato in

questa maniera da integro-differenziale in algebrico nelle incognite vettoriali \mathbf{A} e \mathbf{W} , la cui dimensione risulta pari al numero di lati della mesh.

Il problema si riduce quindi a cercare i valori di \mathbf{A} e \mathbf{W} che rendono minimo la restrizione del funzionale globale di errore Λ nel time-step $[t_k, t_{k+1}]$. Nel seguito indicheremo con \mathbf{a} e \mathbf{w} rispettivamente i vettori $\mathbf{A}(t_{k+1})$ e $\mathbf{W}(t_{k+1})$. e con Λ_k Una formulazione equivalente del problema è quella di cercare i valori di \mathbf{a} e \mathbf{w} che annullano il Jacobiano di $\Lambda_k(\mathbf{a}, \mathbf{w})$:

$$\min \Lambda_k(\mathbf{a}, \mathbf{w}) \quad \rightarrow \quad \begin{bmatrix} \frac{\partial \Lambda_k}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_k}{\partial \mathbf{w}} \end{bmatrix} = 0$$

dove se si adotta per il funzionale locale di errore l'espressione (3), le derivate di Λ_k rispetto alle incognite del lato r della mesh si scrivono :

$$\frac{\partial \Lambda_k}{\partial \mathbf{w}_r^{k+1}} = \alpha_E \iiint_V \nabla \times \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k} [\mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t) - \mathbf{E}] dt +$$

$$+ \alpha_H \iiint_V \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{[\mathbf{f}_B(\mathbf{H}, \mathbf{x}, t) - \mathbf{B}]}{\Delta t_k} dt$$

$$\frac{\partial \Lambda_k}{\partial \mathbf{a}_r^{k+1}} = \alpha_E \iiint_V \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} - \frac{[\mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t) - \mathbf{D}_T]}{\Delta t_k} dt +$$

$$+ \alpha_H \iiint_V \nabla \times \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k} [\mathbf{f}_H(\mathbf{B}, \mathbf{x}, t) - \mathbf{H}] dt$$

mentre i coefficienti della matrice Hessiana di Λ_k valgono:

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \mathbf{a}_s} = \iiint_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \alpha_E \left[\frac{\mathbf{N}_r}{\Delta t_k^2} \cdot \left(\frac{\partial \mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t)}{\partial \mathbf{E}} \mathbf{N}_s \right) \right] + \alpha_H \left[\frac{(t-t_k)^2}{\Delta t_k^2} \nabla \times \mathbf{N}_r \cdot \left(\frac{\partial \mathbf{f}_H(\mathbf{B}, \mathbf{x}, t)}{\partial \mathbf{B}} \nabla \times \mathbf{N}_s \right) \right] dt$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \mathbf{w}_s} = \iiint_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k^2} \{ \alpha_E [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_H [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \} dt$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \mathbf{a}_s} = \iiint_V d\mathbf{x} \int_{t_k}^{t_{k+1}} - \frac{(t-t_k)}{\Delta t_k^2} \{ \alpha_E [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_H [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \} dt$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \mathbf{w}_s} = \iiint_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \alpha_E \left[\frac{(t-t_k)^2}{\Delta t_k^2} \nabla \times \mathbf{N}_r \cdot \left(\frac{\partial \mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t)}{\partial \mathbf{D}_T} \nabla \times \mathbf{N}_s \right) \right] + \alpha_H \left[\frac{\mathbf{N}_r}{\Delta t_k^2} \cdot \left(\frac{\partial \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t)}{\partial \mathbf{H}} \mathbf{N}_s \right) \right] dt$$

Imposizione delle condizioni al contorno

Partizionata la superficie $\partial\Omega$ del dominio di definizione del problema in quattro zone a seconda delle condizioni al contorno imposte su ciascuna:

$$\partial\Omega = \partial\Omega_H \cup \partial\Omega_E \cup \partial\Omega_B \cup \partial\Omega_{D_T}$$

dalle condizioni al contorno imposte rispettivamente per $\partial\Omega_E$ e $\partial\Omega_H$ sulle componenti tangenti di \mathbf{E} ed \mathbf{H} si risale per integrazione temporale alle corrispondenti condizioni per i potenziali vettori.

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}^*(\mathbf{x}, t) = - \int_0^t \mathbf{E}^*(\mathbf{x}, t') dt'$$

$$\mathbf{W}(\mathbf{x}, t) = \mathbf{W}^*(\mathbf{x}, t) = \int_0^t \mathbf{H}^*(\mathbf{x}, t') dt'$$

Invece dalle condizioni al contorno imposte rispettivamente per $\partial\Omega_B$ e $\partial\Omega_{D_T}$ sulle componenti normali di \mathbf{B} e \mathbf{D}_T si determinano per ciascuna faccia di contorno della mesh $\Sigma_l^{(e)} \quad l=1, N_{\text{faces}}$ le circuitazioni di \mathbf{A} e \mathbf{W} su $\partial\Sigma_l^{(e)}$:

$$\oint_{\partial\Sigma_l^{(e)}} \mathbf{A}(\mathbf{x}, t) \cdot d\mathbf{x} = \oint_{\partial\Sigma_l^{(e)}} \mathbf{A}^*(\mathbf{x}, t) \cdot d\mathbf{x} = \iint_{\Sigma_l^{(e)} \times} \mathbf{B}^*(\mathbf{x}, t) \cdot \hat{\mathbf{n}} d\Sigma$$

$$\oint_{\partial\Sigma_l^{(e)}} \mathbf{W}(\mathbf{x}, t) \cdot d\mathbf{x} = \oint_{\partial\Sigma_l^{(e)}} \mathbf{W}^*(\mathbf{x}, t) \cdot d\mathbf{x} = \iint_{\Sigma_l^{(e)} \times} \mathbf{D}_T^*(\mathbf{x}, t) \cdot \hat{\mathbf{n}} d\Sigma$$

Queste condizioni nel problema discretizzato si traducono in equazioni vincolari sulle incognite e possono essere imposte attraverso il metodo delle Penalty Functions o per sostituzione diretta nel sistema (7).

Imposizione delle condizioni al contorno attraverso il metodo delle Penalty Functions

Con il metodo delle Penalty functions il funzionale di errore si modifica nella forma seguente:

$$\Lambda_{Mod,k}(\mathbf{a}, \mathbf{w}) = \Lambda_k(\mathbf{a}, \mathbf{w}) + \frac{C_P}{2} \left(\sum_{r=1}^{N_A} (a_r - a_r^*)^2 + \sum_{s=1}^{N_W} (w_s - w_s^*)^2 \right)$$

dove per i lati appartenenti all'insieme $\partial\Omega_E$ e $\partial\Omega_H$ si ha semplicemente:

$$a_r^* = \int_{l_r} \mathbf{A}^*(\mathbf{x}, t_k + \Delta t_k) \cdot d\mathbf{x}; \quad w_r^* = \int_{l_r} \mathbf{W}^*(\mathbf{x}, t_k + \Delta t_k) \cdot d\mathbf{x};$$

mentre se le condizioni di frontiera sono imposte attraverso le componenti normali di \mathbf{B} o \mathbf{D}_T il valore dei coefficienti può essere determinato attraverso la soluzione dei sistemi di equazioni:

$$\sum_{r=1}^{N_l} a_r^* = \iint_{\Sigma_l^{(e)}} \mathbf{B}^*(\mathbf{x}, t) \cdot \hat{\mathbf{n}} d\Sigma \quad l=1, N_{\text{faces}, B};$$

$$\sum_{r=1}^{N_l} w_r^* = \iint_{\Sigma_l^{(e)}} \mathbf{D}_T^*(\mathbf{x}, t) \cdot \hat{\mathbf{n}} d\Sigma \quad l=1, N_{\text{faces}, D_T}$$

dove N_l è il numero dei lati della faccia di frontiera $\Sigma_l^{(e)}$

$N_{\text{faces}, B}$ è il numero delle facce di frontiera $\Sigma_l^{(e)} \in \partial\Omega_B$

N_{faces, D_T} è il numero delle facce di frontiera $\Sigma_l^{(e)} \in \partial\Omega_{D_T}$

I lati che si trovano sull'intersezione di due diverse zone di frontiera saranno tenuti a rispettare i vincoli legati ad entrambe le frontiere.

Lo jacobiano e la matrice hessiana del funzionale di errore si modificano in maniera consistente con la nuova definizione del funzionale di errore.

Il maggiore vantaggio del metodo delle Penalty è rappresentato dalla facilità con la quale le condizioni al contorno vengono imposte. Operativamente bisogna, infatti, semplicemente aggiornare di volta in volta il termine noto, mentre i coefficienti della matrice non devono essere modificati.

Particolarmente critica nel metodo delle Penalty è la scelta del valore del coefficiente di penalità c_p . Questo deve essere abbastanza grande da dominare in maniera significativa su tutti i coefficienti dell'equazione che si intende vincolare. D'altra parte una scelta di c_p troppo elevata compromette il buon condizionamento della matrice e conduce a sistemi quasi-singolari: in questi casi il metodo fallisce oppure converge verso la soluzione solo dopo un elevatissimo numero di iterazioni di gradiente coniugato.

Imposizione delle condizioni al contorno attraverso il metodo di sostituzione

Nel metodo di sostituzione le condizioni di frontiera vengono imposte operando direttamente sulla matrice hessiana e sullo jacobiano del funzionale di errore.

Ad esempio il vincolo $a_r = a_r^*$ sul lato r viene posto nella seguente maniera:

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \right)_r = \mathbf{a}_r^{(j+1)} - \mathbf{a}_r^* ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{a}_r} = 1$$

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \right)_s = \left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \right)_s + \frac{\partial^2 \Lambda_k^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{a}_s} \mathbf{a}_r^* ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{a}_s} = 0$$

$s=1, \dots, Nedges; s \neq r$

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \right)_s = \left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \right)_s + \frac{\partial^2 \Lambda_k^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{w}_s} \mathbf{a}_r^* ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{w}_s} = 0$$

$s=1, \dots, Nedges ;$

A differenza del metodo di penalità, in questo caso il buon condizionamento della matrice non risulta compromesso. L'inconveniente maggiore è rappresentato dal maggiore lavoro che bisogna compiere ogni volta che si risolve un nuovo sistema: infatti, in questo caso, oltre al termine noto è necessario modificare anche la matrice del sistema.

Condizioni al contorno di tipo miste:

Nei problemi di propagazione si usa imporre sulla frontiera del dominio le cosiddette condizioni di assorbimento che anzichè imporre un particolare valore sulle componenti tangenti del campo elettrico oppure del campo magnetico, stabiliscono che questi campi restino incogniti ma siano vincolati a soddisfare una qualche relazione lineare.

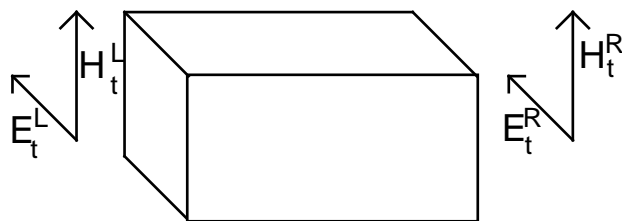


Fig [] Imposizione delle condizioni di assorbimento

Con riferimento alla struttura di figura [] le condizioni da imporre sulle basi siano le seguenti $\forall \mathbf{x} \in \partial V^L$ e $\forall \mathbf{x} \in \partial V^R$

$$\begin{cases} a_{1L} \mathbf{E}_t^L + a_{2L} \mathbf{H}_t^L + a_{3L} = 0 \\ a_{1R} \mathbf{E}_t^R + a_{2R} \mathbf{H}_t^R + a_{3R} = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{H}_t^L = -\frac{a_{3L}}{a_{2L}} - \frac{a_{1L}}{a_{2L}} \mathbf{E}_t^L; & (a) \\ \mathbf{H}_t^R = -\frac{a_{3R}}{a_{2R}} - \frac{a_{1R}}{a_{2R}} \mathbf{E}_t^R; & (b) \end{cases}$$

(2)

dove i campi $(\mathbf{E}_t^L, \mathbf{H}_t^L), (\mathbf{E}_t^R, \mathbf{H}_t^R)$ sono incognite del problema.

Le condizioni di tipo miste non possono essere imposte attraverso il metodo di sostituzione, pena la perdita di simmetria della matrice.

D'altra parte non è neanche conveniente adoperare le funzioni di penalità, in quanto la matrice risultante pur conservando la simmetria, sarebbe dominata da coefficienti fuori diagonale, con un sistema risultante molto mal condizionato; e scarse speranze di convergenza dello schema iterativo verso la soluzione.

Esiste tuttavia la possibilità di imporre condizioni miste con uno schema numerico efficace che conduce alla soluzione del problema attraverso quella di tre problemi simili a quello di interesse, ma con condizioni al contorno miste sostituite da quelle di tipo standard elencate di seguito,

- (I): $\mathbf{E}_{t,1}^{*L}=1; \mathbf{E}_{t,1}^{*R}=0$
- (II) $\mathbf{E}_{t,2}^{*L}=0; \mathbf{E}_{t,2}^{*R}=1 \quad \forall \mathbf{x} \in \partial V^L \quad \text{e} \quad \forall \mathbf{x} \in \partial V^R$
- (III) $\mathbf{E}_{t,3}^{*L}=0; \mathbf{E}_{t,3}^{*R}=0$

Risolvendo i problemi (I), (II), (III) si ottengono sulle frontiere L e R rispettivamente i seguenti valori del campo magnetico \mathbf{H} :

- (I): $\mathbf{H}_{t,1}^{*L}; \mathbf{H}_{t,1}^{*R}$
 - (II) $\mathbf{H}_{t,2}^{*L}; \mathbf{H}_{t,2}^{*R}$
 - (III) $\mathbf{H}_{t,3}^{*L}; \mathbf{H}_{t,3}^{*R}$
- $\forall \mathbf{x} \in \partial V^L \quad \text{e} \quad \forall \mathbf{x} \in \partial V^R$

I campi \mathbf{E} ed \mathbf{H} soluzioni sulle frontiere L ed R del problema originale si possono scrivere come combinazione lineare dei campi \mathbf{E} ed \mathbf{H} soluzioni dei problemi (i),(ii) e(iii) negli stessi punti della frontiera:

$$\mathbf{E}_t^L(\mathbf{x}) = c_1(\mathbf{x})\mathbf{E}_{t,1}^{*L}(\mathbf{x}) + c_2(\mathbf{x})\mathbf{E}_{t,2}^{*L}(\mathbf{x}) + \mathbf{E}_{t,3}^{*L}(\mathbf{x}) = c_1(\mathbf{x}) \quad (3.a)$$

$$\mathbf{E}_t^R(\mathbf{x}) = c_1(\mathbf{x})\mathbf{E}_{t,1}^{*R}(\mathbf{x}) + c_2(\mathbf{x})\mathbf{E}_{t,2}^{*R}(\mathbf{x}) + \mathbf{E}_{t,3}^{*R}(\mathbf{x}) = c_2(\mathbf{x}) \quad (3.b)$$

$$\mathbf{H}_t^L(\mathbf{x}) = c_1(\mathbf{x})\mathbf{H}_{t,1}^{*L}(\mathbf{x}) + c_2(\mathbf{x})\mathbf{H}_{t,2}^{*L}(\mathbf{x}) + \mathbf{H}_{t,3}^{*L}(\mathbf{x}) \quad (4.a)$$

$$\mathbf{H}_t^R(\mathbf{x}) = c_1(\mathbf{x})\mathbf{H}_{t,1}^{*R}(\mathbf{x}) + c_2(\mathbf{x})\mathbf{H}_{t,2}^{*R}(\mathbf{x}) + \mathbf{H}_{t,3}^{*R}(\mathbf{x}) \quad (4.b)$$

$$\forall \mathbf{x} \in \partial V^L \text{ e } \forall \mathbf{x} \in \partial V^R$$

Sostituendo la (2.a) nella (3.b) e la (2.b) nella (4.b) si ottiene:

$$\left\{ \begin{array}{l} \mathbf{E}_t^L(\mathbf{x}) = -\frac{a_{3L}}{a_{1L}} - c_1(\mathbf{x})\frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,1}^{*L}(\mathbf{x}) - c_2(\mathbf{x})\frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,2}^{*L}(\mathbf{x}) - \frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,3}^{*L}(\mathbf{x}); \\ \mathbf{E}_t^R(\mathbf{x}) = -\frac{a_{3R}}{a_{2R}} - c_1(\mathbf{x})\frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,1}^{*R}(\mathbf{x}) - c_2(\mathbf{x})\frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,2}^{*R}(\mathbf{x}) - \frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,3}^{*R}(\mathbf{x}); \end{array} \right. \quad (5.a)$$

$$\left\{ \begin{array}{l} \mathbf{E}_t^L(\mathbf{x}) = -\frac{a_{3L}}{a_{1L}} - c_1(\mathbf{x})\frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,1}^{*L}(\mathbf{x}) - c_2(\mathbf{x})\frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,2}^{*L}(\mathbf{x}) - \frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,3}^{*L}(\mathbf{x}); \\ \mathbf{E}_t^R(\mathbf{x}) = -\frac{a_{3R}}{a_{2R}} - c_1(\mathbf{x})\frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,1}^{*R}(\mathbf{x}) - c_2(\mathbf{x})\frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,2}^{*R}(\mathbf{x}) - \frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,3}^{*R}(\mathbf{x}); \end{array} \right. \quad (5.b)$$

$$\forall \mathbf{x} \in \partial V^L \text{ e } \forall \mathbf{x} \in \partial V^R$$

Sostituendo la (5.a) nella (3.a) e la (5.b) nella (4.a) si ottiene per il punto \mathbf{x} appartenente alle frontiere ∂V^L e ∂V^R il seguente sistema nelle incognite $c_1(\mathbf{x})$ e $c_2(\mathbf{x})$:

$$\begin{bmatrix} 1 + \frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,1}^{*L}(\mathbf{x}) & \frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,2}^{*L}(\mathbf{x}) \\ \frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,1}^{*R}(\mathbf{x}) & 1 + \frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,2}^{*R}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} c_1(\mathbf{x}) \\ c_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} -\frac{a_{3L}}{a_{1L}} - \frac{a_{2L}}{a_{1L}}\mathbf{H}_{t,3}^{*L}(\mathbf{x}) \\ -\frac{a_{3R}}{a_{2R}} - \frac{a_{2R}}{a_{1R}}\mathbf{H}_{t,3}^{*R}(\mathbf{x}) \end{bmatrix}$$

$$\forall \mathbf{x} \in \partial V^L \text{ e } \forall \mathbf{x} \in \partial V^R$$

Calcolati così i coefficienti $c_1(\mathbf{x})$ e $c_2(\mathbf{x})$ è possibile attraverso le (3.a) e (3.b) calcolare i campi $\mathbf{E}_t^L(\mathbf{x})$ ed $\mathbf{E}_t^R(\mathbf{x})$ in tutti i punti delle frontiere ∂V^L e ∂V^R . La

soluzione del problema originale si può infine trovare risolvendo un quarto problema di tipo standard in cui sulla frontiera al posto delle condizioni di assorbimento si impongono i valori dei campi $\mathbf{E}_t^L(\mathbf{x})$ ed $\mathbf{E}_t^R(\mathbf{x})$.

Questo metodo sfrutta il principio di sovrapposizione degli effetti, ma può essere adoperato anche per problemi in cui siano presenti relazioni costitutive non lineari a patto che queste siano linearizzate all'interno del time-step.

Risoluzione numerica del sistema non lineare associato alla minimizzazione.

Il sistema algebrico finale può essere risolto per via numerica attraverso il metodo di Newton. Lo Jacobiano di $\Lambda_{Mod,k}$ viene approssimato in serie di Taylor:

$$\begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j+1)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j+1)}}{\partial \mathbf{w}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{w}} \\ \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{w}} \end{bmatrix} \begin{bmatrix} d\mathbf{a}^{(j+1)} \\ d\mathbf{w}^{(j+1)} \end{bmatrix} = \mathbf{0}$$

Alla j -ma iterazione di Newton bisogna pertanto risolvere il seguente sistema lineare;

$$\begin{bmatrix} \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{w}} \\ \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{w}} \end{bmatrix} \begin{bmatrix} d\mathbf{a}^{(j+1)} \\ d\mathbf{w}^{(j+1)} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \end{bmatrix}$$

Il sistema è sparso, simmetrico e definito positivo. I suoi coefficienti vengono immagazzinati attraverso il metodo di sparsità generale descritto nel capitolo precedente.

Il sistema viene prima preconditionato attraverso il metodo della fattorizzazione incompleta di Cholesky e poi risolto per via iterativa col metodo del gradiente coniugato.. Per ridurre il rischio di patologie ed accelerare il processo di convergenza, si assume alla $(j+1)^{ma}$ iterazione di Newton come soluzione il vettore:

$$\begin{bmatrix} \mathbf{a}^{(j+1)} \\ \mathbf{w}^{(j+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{a}^{(j)} \\ \mathbf{w}^{(j)} \end{bmatrix} + \beta \begin{bmatrix} d\mathbf{a}^{(j+1)} \\ d\mathbf{w}^{(j+1)} \end{bmatrix}$$

dove β è la soluzione del problema di minimizzazione monodimensionale del funzionale;

$$\min \Lambda_{\text{Mod},k}(\mathbf{a}^{(j)} + \beta d\mathbf{a}^{(j+1)}, \mathbf{w}^{(j)} + \beta d\mathbf{w}^{(j+1)}) ;$$

e viene ricavata attraverso il metodo delle secanti.

Attraverso una scelta opportuna dei coefficienti adimensionali α_H e α_E che compaiono nell'espressione del funzionale di errore (e precisamente, $\alpha_H = \alpha_E$), i due campi incogniti $\mathbf{A}(t_{K+1})$ e $\mathbf{W}(t_{K+1})$ possono essere ottenuti come le soluzioni di due sistemi di equazione indipendenti. Vale la pena tuttavia di osservare che questo splitting non consente di separare i due problemi per \mathbf{A} e \mathbf{W} in tutto l'intervallo $(0, T)$ perchè nel nostro caso ogni potenziale a t_{K+1} dipende da entrambi i potenziali a t_K .

3.4) Rappresentazione grafica dei campi di errore.

Il metodo qui proposto provvede due soluzioni complementari che esattamente verificano le equazioni di Maxwell. Precisamente la soluzione in \mathbf{A} fornisce i campi \mathbf{E} e \mathbf{B} che verificano esattamente la legge di Faraday (2-7d), mentre la soluzione in termini di \mathbf{W} produce i campi \mathbf{D}_T e \mathbf{H} che soddisfano rigorosamente la legge di Ampere (2-7c). In questa maniera l'errore numerico

risulta associato esclusivamente alle relazioni costitutive del mezzo (2-9a) e (2-9b).

Una visualizzazione della distribuzione della densità dell'errore locale può essere ottenuta, almeno per materiali isotropi, riportando la nuvola di punti (${}^0\mathbf{B}^0 = {}^0\mathbf{B}_0 + \nabla \times \mathbf{A}^0, {}^0\mathbf{H}^0 = {}^0\partial\mathbf{W}/\partial t^0$) in un piano B-H insieme con la curva rappresentante l'equazione costitutiva (2-9a).

Da un punto di vista pratico, il punto debole di questa rappresentazione di visualizzazione è la mancanza dell'informazione correlata alla distribuzione spaziale della densità di errore; sebbene non esista una dimostrazione formale della diretta correlazione tra errore locale e raffinamento della mesh, l'esperienza pratica indica che una discretizzazione più raffinata localmente è generalmente utile per migliorare l'accuratezza locale. A questo scopo, due possibili rappresentazioni grafiche possono essere di supporto: un primo approccio può essere quello di disegnare nel dominio di interesse i vettori freccia che rappresentano i campi vettoriali $\mathbf{B}-\mathbf{B}(\mathbf{H})$ o alternativamente $\mathbf{D}_T-\mathbf{D}_T(\mathbf{E})$; la seconda possibilità è quella di disegnare la quantità scalare λ attraverso curve di livello in problemi 2D oppure, nel caso 3D, attraverso il modulo di vettori freccia che puntano lungo una direzione preassegnata.

3.5) Implementazione numerica del metodo error-based per la rete RLC.

Posto :

$$A(t_j) = \frac{R}{(t_j - t_k) + RC} ; \quad \Delta t_m = t_{m+1} - t_m ;$$

$$k = q_0 - q_{C0} + \int_0^{t_k} \frac{V - \dot{\Phi}}{R} dt = q_0 - q_{C0} + \sum_{m=0}^{K-1} \left[V(t_m^+) - \Phi(\dot{t}_m^+) \right] \Delta t_m$$

Usando il time-stepping e tenendo conto delle espressioni (2.32) e (2.33) il funzionale di errore (2.35) della rete RLC esaminata nel capitolo 2 può essere posto in forma esplicita come.

$$\Lambda_k(\Phi_{k+1}, q_{k+1}, \alpha) = \frac{1}{2} \sum_{j=1}^{N_G} w_j \left\{ \alpha_L \left[\frac{\Phi^2}{L} + L(\dot{q})^2 - 2\Phi \dot{q} \right] + \alpha_{RC} \left[\frac{(V - \Phi)^2}{A(t)} + A(t)(q - k)^2 - 2(V - \Phi)(q - k) \right] \right\}$$

Dove per l'integrale nell'intervallo di tempo si possono usare le formule di integrazione gaussiana oppure uno dei metodi monostep descritti nel paragrafo (3.1.2). La minimizzazione del funzionale di errore si ottiene uguagliando a zero la derivata di Λ_k rispetto alle incognite Φ_{k+1} e q_{k+1} :

$$\begin{bmatrix} \frac{\partial \Lambda_k}{\partial q_{k+1}} \\ \frac{\partial \Lambda_k}{\partial \Phi_{k+1}} \end{bmatrix} = \mathbf{0}$$

che conduce al sistema :

$$\begin{bmatrix} \sum_{j=1}^{N_G} w_G \left[\alpha_L \frac{L}{\Delta t_K} + \alpha_{RC} \left(\frac{A(t_j)(t_j - t_K)^2}{\Delta t_K} \right) \right] & \sum_{j=1}^{N_G} w_G \left[(\alpha_L - \alpha_{RC}) \left(\frac{t_j - t_K}{\Delta t_K} \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[(\alpha_L - \alpha_{RC}) \left(\frac{t_j - t_K}{\Delta t_K} \right) \right] & \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \frac{1}{A(t_j)\Delta t_K} + \alpha_L \left(\frac{(t_j - t_K)^2}{L\Delta t_K} \right) \right] \end{bmatrix} \begin{bmatrix} q_{k+1} \\ \Phi_{k+1} \end{bmatrix} =$$

$$\begin{aligned}
& \left[\begin{array}{cc} \sum_{j=1}^{N_G} w_G \left[\alpha_L \frac{L}{\Delta t_K} + \alpha_{RC} (t_j - t_K) A(t_j) \left(\frac{t_j - t_K}{\Delta t_K} - 1 \right) \right] & \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \frac{t_j - t_K}{\Delta t_K} + \alpha_L \left(1 - \frac{t_j - t_K}{\Delta t_K} \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left(\frac{t_j - t_K}{\Delta t_K} - 1 \right) - \alpha_L \frac{t_j - t_K}{\Delta t_K} \right] & \sum_{j=1}^{N_G} w_G \left[\frac{\alpha_{RC}}{A(t_j) \Delta t_K} + \alpha_L \frac{t_j - t_K}{L} \left(\frac{t_j - t_K}{L \Delta t_K} - 1 \right) \right] \end{array} \right] \begin{bmatrix} q_k \\ \Phi_k \end{bmatrix} \\
& + \left[\begin{array}{c} \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left((kA(t_j) + V(t_j))(t_j - t_K) \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left(k \frac{V(t_j)}{A(t_j)} \right) \right] \end{array} \right]
\end{aligned}$$

Si noti che la scelta $\alpha_L = \alpha_{RC}$ produce il disaccoppiamento delle equazioni. Ciò ovviamente non significa che i sistemi in q e Φ possono essere avanzati nel tempo indipendentemente. Difatti in ogni caso il termine noto di ciascuna equazione dipende dai valori assunti nel time-step precedente da entrambe le variabili.

CAPITOLO IV

Casi test e risultati

Introduzione

Una delle caratteristiche più interessanti del metodo proposto è il trattamento unificato di problemi stazionari, quasi-stazionari e propagativi. Infatti, la corrente di spostamento è tenuta in conto in maniera automatica, senza alcuna complicazione addizionale, sia nel caso in cui sia importante che in quello in cui sia numericamente trascurabile. Inoltre, i problemi stazionari possono essere trattati come la situazione di regime di un transitorio fittizio.

Si è già visto come i campi possano essere espressi attraverso due espressioni differenti ed indipendenti. Ad esempio, il campo magnetico di induzione \mathbf{B} si può scrivere rispettivamente in funzione di \mathbf{A} e come:

$$\mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A}$$

$$\mathbf{B} = \mathbf{f}_B(\partial \mathbf{W} / \partial t)$$

Inoltre, come visto, anche per parametri integrali, quali ad esempio l'energia magnetica o la potenza ohmica, sono disponibili varie espressioni. Appare tuttavia opportuno ricordare che queste espressioni stabiliscono limiti superiori ed inferiori per alcune grandezze integrali, quali l'energia elettrostatica e magnetostatica, e non per altre come la potenza ohmica. Per questo motivo, il metodo proposto non può essere considerato, a rigore, una vera formulazione CVP.

Il primo paragrafo sarà dedicato alla discussione dei risultati numerici ottenuti applicando la formulazione basata sulla minimizzazione dell'errore costitutivo alla particolare rete RLC introdotta nelle pagine precedenti. Sarà evidenziato, tra l'altro, come il raffinamento temporale induca un miglioramento della convergenza nell'approssimazione della soluzione e nella minimizzazione dell'errore sul funzionale col raffinamento temporale di ordine quadratico.

Nei successivi paragrafi saranno esaminati tre casi test, che si riferiscono rispettivamente a problemi di magnetostatica, correnti parassite e propagazione. Le soluzioni ottenute con l'approccio basato sulla minimizzazione dell'errore costitutivo saranno confrontate con quelle riportate in letteratura.

4.1) Rete RLC : Risultati e discussione. Oscillazioni iniziali. Costanti di tempo lontane tra loro. Analisi delle situazioni di transitorio, regime DC, regime AC. Importanza dei pesi e delle scale di tempo sulle quali ci si colloca. Limiti $R=0$, $L=0$, $C=0$. Confronti in termini di tempo ed accuratezza. Convergenza quadratica nella soluzione e nell'errore del funzionale.

4.2) Magnetostatica. Caso test : Cubo percorso da corrente . Confronto con risultati precedenti.

Il primo esempio che viene riportato si riferisce al calcolo del campo magnetico bidimensionale prodotto da una corrente di densità uniforme parallela ad uno dei lati di un cubo di permeabilità uniforme. La lunghezza del lato sia $l=1$ m, mentre la permeabilità magnetica quella del vuoto; la densità di corrente sia fissata pari a $J=10^7$ A/m².

I risultati sono stati ottenuti usando due mesh: una più rozza costituita da due soli elementi per lato, e l'altra più raffinata con otto elementi per lato. I risultati numerici in termini di energia magnetica sono riportati in tavola 1 insieme con quelli ottenuti in [] dove si sono adoperate le formulazioni in **A** e in **T-Ω**.

Tavola 1

Calcolo del campo magnetostatico in un cubo
($l=1\text{m}$, $J_x=J_y=0$, $J_z=10\text{MA/m}^2$)

<i>Mesh</i>		
elementi per lato	2	8
nodi	27	729
elementi	8	512
lati	54	1944
<i>Gradi di libertà</i>		
formulazione in A []	5	833
formulazione in T-Ω []	54	1944
formulazione in A-W	108	3888
<i>Occupazione di memoria</i>		
form. in A [] (matrice bandata)	14	77914
form. in T-Ω [(matrice bandata)	257	119718
form. in A-W (matrice sparsa)	510	27672

energia magnetica (MJ)

Valore analitico	2.2082	2.2082
form.in A []	1.4726	2.1572
form.in T-Ω []	2.6180	2.2380
form.in A-W (int. volume $\mathbf{B}^2/2\mu$)	1.4726	2.1572
form.in A-W (int. volume $\mu\mathbf{H}^2/2$)	2.6180	2.2381
form.in A-W (int. volume $\mathbf{B}\cdot\mathbf{H}/2$)	1.4726	2.1572

[]**R.Albanese,G.Rubinacci**,*Magnetostatic field computations in terms of two-component vector potentials*, Int. j. num.meth. eng.,29,515-532,**1990**

4.3) Eddy currents. Caso test: Felix-brick. Confronto con risultati precedenti.

Come esempio di applicazione per il problema delle eddy currents è stato scelto il cosiddetto Felix Brick, uno dei problemi usati come test da una associazione internazionale di ricercatori, che va sotto il nome di TEAM Workshop, per misurare le prestazioni di differenti codici numerici[].

Il dominio del problema è costituito da un mattone di alluminio di forma rettangolare con un foro ugualmente rettangolare che lo attraversa completamente. Un campo magnetico uniforme, che decade esponenzialmente col tempo, è applicato perpendicolarmente alle facce forate. Il confronto tra i vari codici viene effettuato comparando la perdita totale di potenza per effetto ohmico, e il campo magnetico in varie posizioni.

Il materiale di cui è composto il mattone è una lega di alluminio con valore di resistività $\rho=3.94 \mu\Omega\text{cm}$. Le dimensioni del mattone sono $0.1524\text{m}\times 0.1016\text{m}\times 0.0508\text{m}$. Il foro rettangolare che attraversa completamente il mattone dalla parte della facce più larghe ha dimensioni $0.0889\text{m}\times 0.0381\text{m}$. Il campo magnetico esterno ha la sola componente nella direzione z ed è uniforme nello spazio. Per $t>0$, decade in maniera esponenziale nel tempo con la seguente legge:

$$B_z = B_0 e^{t/\tau} \quad (1)$$

dove $B_0=0.1$ T e $\tau=0.0119$ s. Questa condizione può essere imposta attraverso la scelta di opportune condizioni al contorno.

Grazie alle condizioni di simmetria, è necessario modellare solo un ottavo del mattone.

Per i calcoli effettuati con la formulazione proposta è stata adoperata la mesh consigliata nella referenza e mostrata nelle figure 1 e 2. Essa è composta da 1344 nodi, 1001 elementi esaedrali, 3656 lati e, pertanto, 7312 incognite per il metodo **A-W**. I valori computati per la potenza ohmica e per il campo indotto nel centro del buco ($x=y=z=0$) sono riportati in tabella 2, e sono in buon accordo con i risultati ottenuti attraverso le altre formulazioni (vedi Tabella 3).

E' interessante riportare il valore incrementale di Λ in un time-step $\Delta t=0.5$ ms (1.66 mJs a $t=10.5$ ms, quando la potenza ohmica è massima). che è sufficientemente piccolo quando comparato al valore della figura di riferimento (2) che è pari a 373.91 mJs:

$$\Lambda_{ref} = \alpha_H \Delta t \int_V \mathbf{B} \cdot \mathbf{H} / 2 dV + \alpha_E \Delta t \int_V (\mathbf{E} \cdot \mathbf{D} / 2 + \mathbf{E} \cdot \mathbf{J} \Delta t) dV \quad (2)$$

Tavola 2

Un problema di eddy currents: il Felix brick
($t=10.5$ ms, $\Delta t=0.5$ ms)

Potenza Ohmica

form.in A-W (int. volume \mathbf{J}^2/σ)	106.98 W
form.in A-W (int. volume $\sigma \mathbf{E}^2$)	106.03 W
form.in A-W (int. volume $\mathbf{E} \cdot \mathbf{J}$)	105.53 W

Campo magnetico indotto in ($x=y=z=0$)

form.in A-W (\mathbf{B}_z)	0.04108 T
form.in A-W ($\mu \mathbf{H}_z$)	0.03907 T

[] **A.Kameari**. *Results for benchmark calculations of problem 4 (the Felix brick)*. COMPEL, vol.7, N.1 & 2, 65-80 (1988)

4.4) Propagazione. Caso test : Incidenza di un'onda su uno slab di materiale dielettrico lineare. Confronto con la soluzione analitica. Incidenza di un'onda su uno slab di materiale dalle caratteristiche non lineari.

Come ultimo esempio si è scelto un semplice problema di propagazione: l'incidenza di un'onda piana su una lastra di lunghezza indefinita nelle direzioni x e y e dello spessore di 20 metri. Si è assunto $\mu_r=1$, $\epsilon_r=10$ e $\sigma=0$). Sebbene il problema sia di natura monodimensionale, è stato studiato come se fosse 3D usando una mesh consistente in una pila di elementi esaedrici lungo la direzione z (direzione di propagazione dell'onda). Sono state usate due discretizzazioni, ciascuna delle quali includeva 200 metri di spazio vuoto prima e 100 metri oltre la lastra. La mesh più rozza consiste di 128 elementi e 2056 incognite, mentre quella più raffinata di 320 elementi e 5128 incognite.

Si è supposto che sulla lastra incidesse un'onda piana il cui andamento temporale fosse rappresentato dalla semplice forma di un gradino. Si è assunto inoltre che il campo magnetico avesse ampiezza unitaria e fosse linearmente polarizzato lungo la direzione y . Questo caso, nonostante la semplicità della sua soluzione analitica è piuttosto critico dal punto di vista numerico a causa della brusca variazione del fronte d'onda. Le soluzioni ottenute attraverso formulazioni FEM di tipo tradizionale possono, infatti, mostrare accentuate oscillazioni che precedono e seguono il fronte d'onda.

Le seguenti condizioni al contorno sono state assunte: $\mathbf{E} \times \mathbf{n} = 0$, che implica $\mathbf{A} \cdot \mathbf{n} = 0$, sui due piani di simmetria a $x = \text{costante}$; $\mathbf{H} \times \mathbf{n} = 0$, che si traduce in $\mathbf{W} \cdot \mathbf{n} = 0$, sui piani di simmetria a $y = \text{costante}$; sulla frontiera $z = \text{cost}$ situata dietro la lastra si è imposto, attraverso l'incognita \mathbf{W} , \mathbf{H}_x uguale a zero e \mathbf{H}_y essere un gradino unitario; ed infine, sul piano di frontiera $z = \text{cost}$ collocato

oltre la lastra si è imposta la condizione $\mathbf{H} \times \mathbf{n} = 0$, ovvero $\mathbf{W} \cdot \mathbf{n} = 0$. In questo modo il problema dell'incidenza dell'onda piana sulla lastra dielettrica nello spazio libero è stato rappresentato correttamente, almeno fino al momento in cui uno dei fronti d'onda uscenti dalla lastra raggiunge le superfici di frontiera $z = \text{cost}$.

In figura, è mostrato il confronto tra i risultati numerici e la soluzione analitica. Al fronte d'onda delle soluzioni numeriche è vietato variare in maniera ripida dalle intrinseche proprietà di continuità sulle componenti tangenti delle edge functions attraverso le quali sono rappresentati gli andamenti spaziali dei potenziali vettori \mathbf{A} e \mathbf{W} , e conseguentemente i campi \mathbf{E} e \mathbf{H} . Tuttavia, come si può osservare il fenomeno di dispersione del fronte d'onda coinvolge un numero molto limitato di elementi (all'incirca 3).

Si sottolinea che la formulazione basata sull'uso congiunto dei potenziali \mathbf{A} e \mathbf{W} e degli edge elements non comporta oscillazioni nei campi in prossimità del fronte d'onda, ed in questo si dimostra superiore rispetto alle tradizionali soluzioni agli elementi finiti. Si può infine notare che il raffinamento della mesh produce una diminuzione del funzionale di errore Λ , che vale $0.22\text{E-}13$ Js per la mesh rozza e $0.19\text{E-}14$ nell'intervallo di tempo $(0, 0.68 \mu\text{s})$. Sfortunatamente, nel caso dell'analisi nel dominio del tempo di un fenomeno transitorio di propagazione, l'informazione sulla localizzazione dell'errore massimo è poco significativa, perchè essa coincide ovviamente col fronte d'onda che si sta muovendo. Nondimeno, l'informazione può ridiventare preziosa se si adottano mesh mobili oppure nel caso dell'analisi nel dominio complesso di fenomeni propagativi pervenuti nella condizione di regime.

CAPITOLO V

Minimizzazione dell'errore costitutivo nella MHD e applicazioni ai modelli dinamici di plasma.

5.1) Generalita' sulla fusione nucleare. Plasma. Equazioni MHD. Problema aperti : equilibrio e evoluzione del confinamento del plasma in strutture tridimensionali quali i Tokamak . Halo currents, diffusione resistiva.

(Ref. Tesi Phd . Albanese,...)

5.2) Review delle metodologie e dei codici attualmente adoperati. Il codice Proteus. Limite della quasi assi-simmetria .

(Ref. Albanese-Rubinacci...)

5.3) Modello matematico per le equazioni MHD col metodo della minimizzazione dell'errore costitutivo.

CONCLUSIONI

L'approccio basato sulla minimizzazione dell'errore costitutivo è stato esteso allo studio di problemi tridimensionali generali di natura elettromagnetica. I principali vantaggi di questa estensione sono i seguenti:

a) la soluzione è ottenuta attraverso la minimizzazione di un funzionale di errore globale e, da questo, la matrice delle equazioni di sistema è simmetrica e definita positiva;

b) è facilmente disponibile una stima dell'errore numerico che si sta commettendo.

c) consente un trattamento unificato di problemi stazionari, quasi-stazionari e propagativi.

D'altra parte, le incognite sono in numero doppio rispetto alle classiche formulazioni in termini di **B** o di **E**. Tuttavia, attraverso una scelta conveniente dei pesi, è possibile ottenere lo splitting del modello discretizzato; in tal modo, ci si riconduce a dover risolvere, ad ogni passo nel tempo, due sistemi indipendenti di equazioni rispettivamente nelle incognite **A** e **W**.

CAPITOLO III

Aspetti computazionali

3.1) Soluzione numerica dei problemi a valore iniziale

3.1.1) Generalità

Le equazioni differenziali alle derivate parziali sono usualmente classificate nelle tre categorie di equazioni ellittiche, paraboliche ed iperboliche in base alle loro *curve caratteristiche*, ovvero alle curve che determinano il modo in cui l'informazione si propaga.

Tipici esempi di equazioni iperboliche sono le equazioni di convezione e le equazioni delle onde. Quest'ultima, per il caso monodimensionale, si scrive come:

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

dove $v = \text{cost.}$ è la velocità di propagazione dell'onda.

Un tipico esempio di equazione parabolica è l'equazione di diffusione:

$$\frac{\partial u}{\partial t} = -\frac{\partial}{\partial x} \left(D \frac{\partial u}{\partial x} \right)$$

(2)

dove D è il coefficiente di diffusione.

Infine il prototipo di un'equazione ellittica è l'equazione di Poisson:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \rho(x,y) \quad (3)$$

dove \square è il termine sorgente. Se quest'ultimo è identicamente nullo l'equazione risultante è la ben nota equazione di Laplace.

Da un punto di vista computazionale la distinzione che assume maggiore importanza è quella tra equazioni di tipo ellittico ed equazioni di altra categoria. Difatti, mentre la soluzione dell'equazione (3) è una funzione statica, per trovare la quale è necessario specificare i suoi valori, oppure quelli della derivata normale, sulla frontiera del dominio di definizione (problemi siffatti sono detti, per questo motivo, *di valore al contorno*), i problemi del tipo (1) e (2), partendo dalla conoscenza di u all'istante iniziale, descrivono come la funzione incognita evolva nel tempo. Questi ultimi problemi sono generalmente definiti *a valore iniziale* o *di Cauchy*.

Per quel che riguarda la sottoclassificazione dei problemi di Cauchy in iperbolici e parabolici si dirà solo che dal punto di vista pratico riveste importanza soprattutto per quel che riguarda il differente tipo di condizioni al contorno da imporre nelle due classi di problemi.

3.1.2) Soluzione numerica dei problemi a valore iniziale o di Cauchy

Tra le varie proposte per la soluzione numerica dei problemi a valore iniziale, sarà presentata quella che maggiore fortuna ha incontrato in letteratura: questa prevede l'uso congiunto degli Elementi finiti, per il trattamento delle variabili spaziali, e del cosiddetto time-stepping per la rappresentazione dell'evoluzione temporale.

Attraverso un metodo basato sugli elementi finiti si perviene, al termine della fase di assemblaggio delle matrici, ad un sistema di N equazioni differenziali ordinarie accoppiate che per i problemi di tipo parabolico ha la forma seguente:

$$[\mathbf{C}] \left(\frac{\partial \mathbf{a}(t)}{\partial t} \right) + [\mathbf{K}] \mathbf{a}(t) = \mathbf{f}(t); \quad t \in [t_0, t_0+T] \quad \text{con } \mathbf{a}(t_0) = \mathbf{a}_0 \quad (1)$$

dove $[\mathbf{C}]$ è detta *matrice di capacità*, $[\mathbf{K}]$ è la *matrice di rigidità* ed \mathbf{f} è il *vettore di carico*, e \mathbf{a}_0 è il vettore delle condizioni iniziali assegnate.

Nei problemi misti il sistema differenziale si presenta come:

$$\begin{cases} [\mathbf{M}]\left(\frac{\partial^2 \mathbf{a}(t)}{\partial t^2}\right) + [\mathbf{C}]\left(\frac{\partial \mathbf{a}(t)}{\partial t}\right) + [\mathbf{K}]\mathbf{a}(t) = \mathbf{f}(t) \\ \mathbf{a}(t_0) = \mathbf{a}_0 \\ \left.\frac{\partial \mathbf{a}(t)}{\partial t}\right|_{t=t_0} = \mathbf{f}(t_0) \end{cases} \quad t \in [t_0, t_0+T] \quad (2)$$

dove $[\mathbf{M}]$ è la *matrice di massa*, mentre $[\mathbf{C}]$ viene ora chiamata *matrice di smorzamento*, anzichè *di capacità*. Nei problemi iperbolici vale ancora la (2), a patto che $[\mathbf{C}]$ coincida con la matrice nulla di ordine N .

In tutti i metodi basati sul time-stepping il dominio temporale $[t_0, t_0+T]$ è diviso in una successione di intervalli Δt_i , $i=1, N_{\text{step}}$ a partire dall'istante t_0 . In alcuni metodi il passo degli intervalli deve essere mantenuto costante, in altri può avere lunghezza diversa. A questo punto, anzichè cercare la soluzione del sistema differenziale come un vettore $\mathbf{a}(t)$ le cui componenti $a_j(t)$ siano funzioni continue, si assumono come incognite i valori discreti $\mathbf{a}_i = \mathbf{a}(t_i)$ che le funzioni assumono alla fine di ciascun time-step. Gli $\{\mathbf{a}_i\}$, $i=1, N_{\text{step}}$, vengono calcolati attraverso una relazione di ricorrenza, ovvero attraverso una equazione algebrica che approssima l'equazione differenziale e che correla il valore di due (metodi a singolo step) o più (metodi multi-step) termini consecutivi della successione \mathbf{a}_i . Nei metodi mono-step, ad esempio, vengono correlati i valori discreti degli estremi di un singolo step: in tal caso si usano i valori iniziali noti \mathbf{a}_0 per determinare i valori \mathbf{a}_1 assunti dal vettore soluzione alla fine del primo time-step; per il secondo time step si procede in maniera analoga usando per la determinazione del vettore \mathbf{a}_2 al posto di \mathbf{a}_0 il vettore \mathbf{a}_1 determinato al passo precedente, e così via per tutti i time-steps successivi.

Per problemi di diffusione sono quasi universalmente adoperati metodi ad uno o due step, mentre per problemi misti sono soventemente impiegati metodi multi-step di ordine superiore. I metodi mono-step includono tre classiche formule delle differenze finite:

- Il metodo implicito noto anche come la *regola di Eulero a ritroso*.

- Il metodo delle differenze centrali, detto metodo Crank-Nicolson o della regola trapezoidale.

- Il metodo esplicito noto come regola di Eulero.

I tre metodi precedenti sono dei casi speciali di una formula più generale indicata come il metodo θ .

Da un punto di vista pratico nella scelta del metodo da adottare bisogna tenere conto di due indici fondamentali, che dipendono dal solo schema di integrazione e non dal particolare sistema differenziale: il tasso di convergenza asintotico e la stabilità.

Il tasso di convergenza asintotico misura quanto rapidamente decada, in un dato istante t , l'errore sulla soluzione al diminuire del time-step, Δt , per $\Delta t \rightarrow 0$.

Il comportamento stabile o instabile di un metodo può essere determinato osservando l'evoluzione libera di un sistema del primo ordine (che, come è noto, è una funzione che decade con andamento esponenziale). Un metodo instabile produrrà una soluzione che oscilla con ampiezza sempre crescente attorno alla curva di decadimento esponenziale. In un metodo stabile, invece, la soluzione o conterrà oscillazioni che vanno decadendo con $t \rightarrow \infty$, oppure sarà del tutto priva di oscillazioni.

Premesso che, pur se sviluppate con maggiore complessità, le idee di fondo di tutti i metodi multi-step ricalcano quelle già presenti negli schemi a singolo step, sono illustrate brevemente di seguito le caratteristiche peculiari del metodo mono-step applicato al problema di diffusione il cui modello è rappresentato dal sistema differenziale (1).

- Metodo implicito

Per tutti gli schemi del primo ordine il valore della derivata prima per il k -mo time-step, può essere approssimato attraverso la relazione:

$$\left(\frac{\partial \mathbf{a}(t)}{\partial t} \right) = \frac{\{\mathbf{a}\}_k - \{\mathbf{a}\}_{k-1}}{\Delta t_k} \quad \forall t \in [t_k, t_k + \Delta t_k] \quad (3)$$

Nello schema implicito il sistema (1) si può scrivere come:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_k + [\mathbf{K}] \{\mathbf{a}\}_k = \{\mathbf{f}\}_k \quad (4)$$

dove si è indicato con il pedice k il valore assunto dalla grandezza all'istante t_k . Sostituendo la (3) nella (4) si perviene alla relazione di ricorrenza:

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + [\mathbf{K}] \right) \{\mathbf{a}\}_k = \{\mathbf{f}\}_k + \frac{1}{\Delta t_k} [\mathbf{C}] \{\mathbf{a}\}_{k-1} \quad (5)$$

Il tasso di convergenza asintotico del metodo implicito è del primo ordine, cioè l'errore sulla soluzione in un dato istante t è proporzionale a Δt quando $\Delta t \rightarrow 0$.

La maggiore attrattiva di questo metodo è di essere del tutto privo di oscillazioni e incondizionatamente stabile.

I maggiori svantaggi possono così riassumersi:

- bisogna invertire $\left(\frac{1}{\Delta t_k} [\mathbf{C}] + [\mathbf{K}] \right)$, il che è generalmente meno agevole rispetto ad invertire $\frac{1}{\Delta t_k} [\mathbf{C}]$

-il metodo può risultare stabile anche per problemi fisicamente instabili;

- la stabilità numerica è dovuta ad un meccanismo di "dissipazione" che può dare fastidio se vogliamo studiare problemi fisicamente non dissipativi (ovviamente si fa riferimento a problemi di tipo iperbolico o simile).

- Metodo di Crank-Nicolson

Nello schema alla Crank-Nicolson si adopera per approssimare il sistema differenziale (1) l'espressione:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{k-1/2} + [\mathbf{K}] \{\mathbf{a}\}_{k-1/2} = \{\mathbf{f}\}_{k-1/2} \quad (6)$$

che determina il seguente sistema algebrico:

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + \frac{1}{2} [\mathbf{K}] \right) \{\mathbf{a}\}_k = \{\mathbf{f}\}_{k-1/2} + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - \frac{1}{2} [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (7)$$

Lo schema delle differenze centrali converge in maniera proporzionale a Δt^2 , è stabile ma presenta oscillazioni nella soluzione.

- Metodo di Eulero

Il metodo di Eulero usa come schematizzazione della (1) l'espressione:

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{k-1} + [\mathbf{K}] \{\mathbf{a}\}_{k-1} = \{\mathbf{f}\}_{k-1} \quad (8)$$

che comporta:

$$\frac{1}{\Delta t_k} [\mathbf{C}] \{\mathbf{a}\}_k = \{\mathbf{f}\}_{k-1} + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (9)$$

A differenza dei metodi precedenti, nello schema di Eulero la matrice di rigidità, $[\mathbf{K}]$, non compare più al primo membro della relazione di ricorrenza, mentre resta presente la matrice di capacità. Poichè quest'ultima normalmente non è diagonale, lo schema (9) è ancora implicito. Tuttavia la matrice $[\mathbf{C}]$ può essere resa diagonale, la qual cosa disaccoppia le incognite e rende questo

metodo estremamente veloce. La diagonalizzazione della matrice può essere ottenuta rapidamente attraverso delle tecniche che vanno sotto il generico nome di *lumping*. L'efficacia computazionale del *lumping* si può misurare osservando quanto vari la velocità di convergenza e la precisione della soluzione se si risolve lo schema modificato al posto dello schema (9). In generale si potrà osservare una diminuzione del tasso di convergenza e, cosa interessante, un miglioramento a volte significativo nella precisione.

Tra le varie tecniche di *lumping*, la più semplice, applicata prima della fase di assemblaggio alle matrici di capacità dei singoli elementi, consiste nel ricavare la matrice concentrata $[\mathbf{CL}]^{(e)}$ a partire dalla matrice di capacità $[\mathbf{C}]^{(e)}$ col seguente algoritmo:

$$CL_{ii}^{(e)} = \sum_{j=1}^n C_{ij}^{(e)}; \text{ per } i=1, 2, \dots, n; CL_{ij}^{(e)} = 0; \text{ per } i \neq j; \quad (10)$$

A fronte del pregio di essere più veloce dal punto di vista computazionale rispetto ai metodi precedenti, il metodo esplicito presenta il grave inconveniente di essere condizionatamente stabile (vedi paragrafo successivo).

- Metodo θ .

Questo metodo è la naturale generalizzazione dei precedenti. In questo caso la (1) viene valutata nella generica locazione dell'intervallo individuata dal parametro adimensionale θ :

$$[\mathbf{C}] \left\{ \frac{\partial \mathbf{a}}{\partial t} \right\}_{\Theta} + [\mathbf{K}] \{\mathbf{a}\}_{\Theta} = \{\mathbf{f}\}_{\Theta} \quad (11)$$

dove :

$$\Theta = \frac{t - t_{k-1}}{\Delta t_k} \quad (12)$$

$$\left(\frac{1}{\Delta t_k} [\mathbf{C}] + \Theta [\mathbf{K}] \right) \{\mathbf{a}\}_k = (1 - \Theta) \{\mathbf{f}\}_{k-1} + \Theta \{\mathbf{f}\}_k + \left(\frac{1}{\Delta t_k} [\mathbf{C}] - (1 - \Theta) [\mathbf{K}] \right) \{\mathbf{a}\}_{k-1} \quad (13)$$

Per quanto riguarda la stabilità, si può dire in generale che:

- Per $1 \geq \theta \geq 0.5$ lo schema di integrazione è incondizionatamente stabile.
- Per $0 \leq \theta < 0.5$ lo schema è stabile solo se, indicato con λ_{Max} il maggior autovalore della matrice $[\mathbf{C}^{-1}\mathbf{K}]$ viene soddisfatta la seguente disuguaglianza, che pone un limite sulla massima ampiezza del time-step:

$$\lambda_{Max} \Delta t < \frac{2}{1 - 2\Theta} \quad (14)$$

Esistono delle formule che permettono di stimare λ_{Max} senza passare per il calcolo degli autovalori della matrice [35]

Oltre a quei valori (0,5,1) per i quali si ritorna agli schemi già esaminati, altri valori di θ che spesso vengono adoperati sono $\theta=2/3$ e $\theta=0.878$ (che rappresenta il valore minore per il quale la soluzione non presenta oscillazioni).

3.2) Metodi di soluzione di sistemi di eq. lineari e non lineari

3.2.1) Risoluzione di sistemi di equazioni lineari.

- Introduzione

La risoluzione dei sistemi lineari riveste un ruolo basilare nella matematica applicata. Mentre, infatti, diversi modelli matematici significativi nascono direttamente lineari, si è già mostrato come, attraverso i metodi numerici,

anche problemi di dimensione infinita (ad esempio quelli modellati attraverso equazioni integro-differenziali non lineari), siano riconducibili a sistemi non lineari di dimensione finita che è spesso possibile linearizzare (ovvero risolvere per via iterativa attraverso problemi lineari).

Alla varietà di contesti da cui un sistema lineare può scaturire, corrisponde una uguale differenziazione nella struttura della matrice associata al sistema. Questa, infatti, può essere di volta in volta piena, bandata o sparsa, simmetrica oppure definita positiva. E' importante, dunque, avere a disposizione una varietà di algoritmi efficienti tra i quali scegliere il più adeguato per stabilità, occupazione di memoria, velocità di esecuzione a risolvere su un determinato calcolatore un problema particolare.

In questo paragrafo si definiranno alcuni termini adoperati usualmente nella classificazione delle matrici. Si daranno quindi dei brevi cenni circa i metodi di immagazzinamento delle matrici sparse. La parte conclusiva è dedicata, seguendo l'usuale ripartizione in metodi diretti ed iterativi, all'illustrazione di alcune delle tecniche adoperate nella risoluzione di sistemi lineari $N \times N$ non-singolari.

- Generalità intorno alla risoluzione di sistemi lineari

Si riportano alcune definizioni adoperate usualmente in Algebra Lineare nella classificazione delle matrici :

Una matrice è detta reale se tutti i suoi elementi appartengono al dominio dei numeri reali.

Una matrice \mathbf{A} quadrata di ordine N ad elementi reali tale che $\mathbf{A}=\mathbf{A}^t$ è detta *simmetrica*.

Una matrice \mathbf{A} quadrata di ordine N ad elementi complessi tale che $\mathbf{A}=\mathbf{A}^*$ è detta *hermitiana*.

Una matrice reale \mathbf{A} di ordine N tale che per ogni $\mathbf{x} \in \mathbb{R}^n$ diverso dall'elemento nullo si abbia $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ viene detta *definita positiva*.

Una matrice reale \mathbf{A} quadrata di ordine N e di elementi a_{ij} tale che si abbia:

$$a_{ii} \geq \sum_{j=1, j \neq i}^N |a_{ij}| \quad \forall i \in [1, N]$$

è detta a *dominanza diagonale*

Una matrice quadrata \mathbf{D} è detta diagonale se tutti i suoi elementi che non si trovano sulla diagonale principale sono nulli.

Una matrice quadrata \mathbf{U} è detta triangolare superiore se tutti i suoi elementi che si trovano sotto la diagonale principale sono nulli.

Una matrice quadrata \mathbf{L} è detta triangolare superiore se tutti i suoi elementi che giacciono sopra la diagonale principale sono nulli.

Si dimostra che se \mathbf{A} è una matrice quadrata simmetrica è sempre possibile trovare una matrice diagonale \mathbf{D} e una matrice triangolare superiore \mathbf{U} tale che $\mathbf{A} = \mathbf{U}^T \mathbf{D} \mathbf{U}$ (Fattorizzazione di \mathbf{A}).

Si è già evidenziata la stretta correlazione tra la struttura di una matrice e il tipo di metodo numerico che si sta adoperando. Una matrice in cui prevalgono termini diversi da zero viene chiamata "*piena*", mentre una in cui dominano i termini nulli si classifica come *sparsa*.

I metodi numerici di tipo integrale (ad esempio il metodo degli elementi di frontiera) sono caratterizzati da matrici associate piene di dimensione relativamente piccola (100-1000 incognite), viceversa i metodi basati sugli elementi finiti generano sistemi con matrici associate sparse e di ordine elevato (1000-30000 incognite).

Esistono diverse tecniche di immagazzinamento dei coefficienti delle matrici di tipo sparso, e la scelta della più adatta al problema in esame viene determinata dalle caratteristiche strutturali di sparsità della matrice. I differenti approcci vengono usualmente classificati in letteratura sotto le etichette di *metodi a bande*, *metodi d'involuppo* e *metodi di sparsità generale*.

Da un punto di vista pratico capita spesso di incontrare problemi basati su metodi FEM che richiedono la soluzione di diversi sistemi lineari di tipo sparso per i quali la struttura della matrice -legata alla mesh, alle funzioni di base ed alle funzioni test-, non subisce modifiche, mentre possono variare sia i coefficienti del sistema che i termini noti. In questi casi l'approccio risolutivo generalmente adoperato passa attraverso tre stadi che vanno sotto i nomi di *fattorizzazione simbolica*, *fattorizzazione numerica* della matrice e *soluzione del sistema*.

La fattorizzazione simbolica viene eseguita una volta sola, all'inizio del processo solutivo e il suo scopo è quello di determinare i legami tra le caratteristiche di sparsità delle matrici **A** e **U**. Il processo di fattorizzazione numerica va invece ripetuto ogni volta che un nuovo sistema deve essere risolto. Tuttavia le informazioni acquisite durante la fase di fattorizzazione simbolica fanno sì che quest'ultima operazione risulti estremamente veloce.

Per i metodi a banda o di involuppo, caratterizzati da sparsità strutturali, i blocchi contenenti elementi non nulli sono semplicemente identici e, pertanto, il processo di fattorizzazione simbolica può essere evitato.

Nel caso dei metodi di sparsità generali, invece i legami tra **A** e **U** sono molto meno intuitivi e dipendono oltre che dallo schema di fattorizzazione (che determina il cosiddetto *fill in* o riempimento della matrice) anche dalla numerazione dei nodi della mesh. Per questo motivo i metodi di sparsità generale vengono usualmente preceduti da algoritmi -il più famoso dei quali è quello di Cuthill-McKee- capaci di minimizzare il fill in attraverso il riordinamento della numerazione dei nodi. Quest'insieme di operazioni può essere molto costoso dal punto di vista computazionale perciò la fase preliminare di fattorizzazione simbolica produce un consistente vantaggio.

Un elemento che bisogna portare in conto nella scelta dello schema risolutivo è l'onere computazionale che esso impone. Questo parametro viene comunemente valutato attraverso il numero di moltiplicazioni e divisioni -le

operazioni elementari più costose per la macchina calcolatrice- che l'algoritmo richiede per la risoluzione di un sistema di ordine N.

- Metodi di immagazzinamento delle matrici

Metodi di banda

I metodi di banda [36] sono adeguati per il trattamento di matrici, caratterizzate dalla concentrazione dei termini non nulli entro uno stretto numero di diagonali adiacenti la diagonale principale. Il parametro caratteristico, m_A , associato alla matrice *bandata* \mathbf{A} viene chiamato *ampiezza di banda*, ed è definito come il più piccolo intero tale che $a_{i,j}=0$ per $|i-j|>m_A$. Una importante proprietà delle matrici a bande è che la fattorizzazione $\mathbf{A}=\mathbf{U}^T\mathbf{D}\mathbf{U}$, non comporta allargamento della banda per la matrice triangolare superiore: per questo particolare processo di fattorizzazione risulta cioè $m_A=m_U$.

Sistemi di matrici bandate provengono, ad esempio, da problemi agli elementi finiti monodimensionali, o anche bidimensionali ma definiti in domini geometricamente molto regolare. Un esempio di matrice bandata è riportato in fig.[1].

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 \\ a_{12} & a_{22} & a_{23} & a_{24} & 0 & 0 \\ a_{13} & a_{23} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & a_{35} & a_{45} & a_{55} & a_{56} \\ 0 & 0 & a_{36} & a_{46} & a_{56} & a_{66} \end{bmatrix}$$

(a)

$$\begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{44} & a_{55} & a_{66} \\ a_{12} & a_{23} & a_{34} & a_{45} & a_{56} & * \\ a_{13} & a_{24} & a_{35} & a_{46} & * & * \\ a_{14} & 0 & a_{36} & * & * & * \end{bmatrix}$$

(b)

Figura [1]

- (a) Matrice bandata \mathbf{A} con $m_A=3$;
 - (b) Schema di immagazzinamento per \mathbf{A} con un array bidimensionale.
- (* = Elementi dell'array di immagazzinamento non adoperati)

Adoperando il criterio sopra illustrato lo spazio di immagazzinamento di una matrice bandata di ordine N e di ampiezza di banda m_A passa da N^2 a $N \times (1+m_A)$ locazioni di memoria che includono anche $m_A(1+m_A)/2$ posizioni dell'array che non vengono adoperate.

Metodi di involuppo

Un problema FEM bidimensionale, con geometria del dominio di definizione abbastanza complessa, conduce ad un moderato grado di variazione locale dell'ampiezza di banda della matrice. Attraverso una opportuna rinumerazione degli elementi geometrici (nodi o lati) cui sono associate le incognite del problema è possibile ottenere una matrice nella quale l'involuppo degli elementi non nulli ha una forma lenticolare. In questo caso, sebbene l'ampiezza di banda della matrice resti uguale a m_A , fatta eccezione per la riga centrale $N/2$, gli elementi non nulli si trovano generalmente ben all'interno della banda di ampiezza m_A . Questa circostanza viene sfruttata nei metodi di involuppo che al prezzo di una maggiore complessità della struttura di immagazzinamento possono in alcuni casi risultare maggiormente efficienti. Un esempio è costituito dai problemi bidimensionali in cui si sia provveduto a numerare i nodi della mesh lungo le diagonali (Vedi Fig. [2]). In tal caso lo schema di immagazzinamento di involuppo consenta un risparmio del 33% in termini di locazioni di memoria e del 50% nel tempo di fattorizzazione.

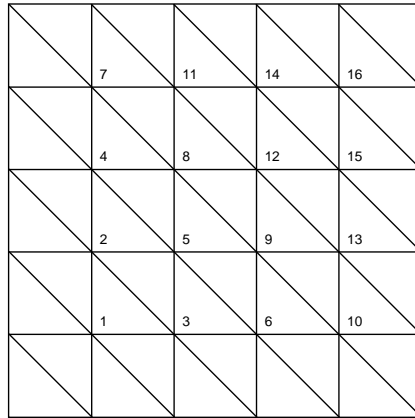


Figura [2]: Ordinamento diagonale dei nodi

Metodi di sparsità generale

I metodi a banda e d'inviluppo, mentre risultano vantaggiosi nel trattamento di matrici che presentano blocchi strutturali compatti (matrici cioè che sia possibile partizionare in zone in cui si trovano soltanto elementi nulli e in zone che presentano un'alta densità di elementi diversi da zero), si dimostrano totalmente insoddisfacenti in tutti i problemi che danno origine a matrici la cui sparsità è di tipo non strutturale: ciò è vero, in particolare, per una vasta classe di problemi FEM bi- e tri-dimensionali. Si preferiscono utilizzare in questi casi dei metodi che a prezzo di una maggiore complessità della struttura di immagazzinamento dei coefficienti garantiscono una significativa riduzione della quantità di dati da conservare.

Sia \mathbf{A} una matrice simmetrica di ordine N che presenta caratteristiche di sparsità non strutturali. Per la simmetria sarà possibile immagazzinare solo i coefficienti che si trovano sulla diagonale \mathbf{D} oppure nel blocco triangolare inferiore \mathbf{E} di \mathbf{A} . Per l'immagazzinamento dei coefficienti si utilizzano tre vettori monodimensionali \mathbf{A} , \mathbf{IA} e \mathbf{JA} .

\mathbf{A} è un vettore di reali che contiene, ordinati per righe, tutti gli elementi di \mathbf{E} e \mathbf{D} diversi da zero.

JA è un vettore di interi che contiene l'indice di colonna degli elementi corrispondenti immagazzinati in A

IA contiene, infine, i puntatori agli elementi diagonali immagazzinati in A .

Se il numero di elementi non nulli di $E + D$ è $nsize$, allora la dimensione di A e JA sarà $nsize$, mentre quella di IA sarà N .

Gli elementi a_{ij} con $j \leq i$ dell' i -ma riga di A sono immagazzinati nelle posizioni di A che vanno dall'indice $IA(i-1)+1$ all'indice $IA(i)$

In figura [4] è mostrato come possa essere immagazzinata attraverso questo schema la matrice di figura [3]:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & a_{17} \\ a_{12} & a_{22} & a_{23} & 0 & 0 & 0 & 0 \\ 0 & a_{23} & a_{33} & a_{34} & 0 & 0 & 0 \\ 0 & 0 & a_{34} & a_{44} & a_{45} & 0 & 0 \\ 0 & 0 & 0 & a_{45} & a_{55} & a_{56} & 0 \\ 0 & 0 & 0 & 0 & a_{56} & a_{66} & a_{67} \\ a_{17} & 0 & 0 & 0 & 0 & a_{67} & a_{77} \end{bmatrix} \quad (a)$$

Fig.[3] Esempio di matrice sparsa

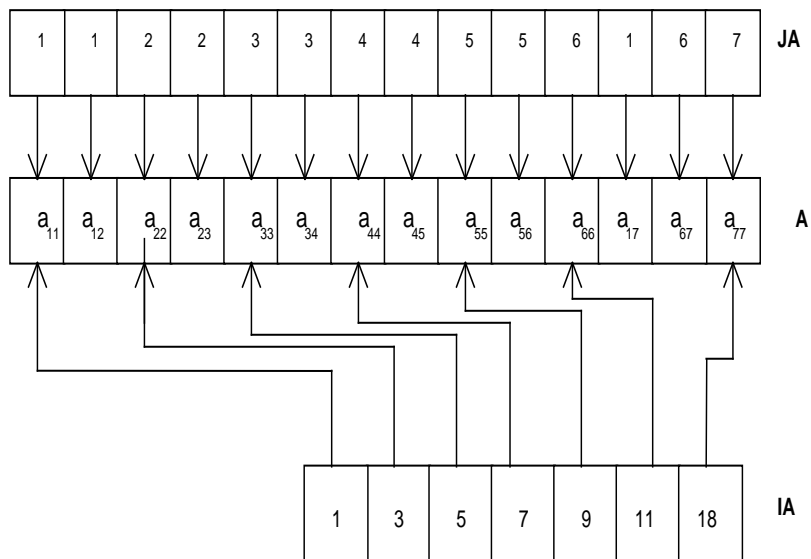


Fig.[4] Immagazzinamento di una matrice con il metodo di sparsità generale

Si nota immediatamente che questo schema non permette né un accesso random, né un accesso per colonne agli elementi di \mathbf{A} .

Va infine notato che se -come spesso succede- bisogna risolvere diversi sistemi lineari ai quali sono associate matrici sparse della medesima struttura ed eventualmente con diversi coefficienti, allora le sole informazioni contenute in \mathbf{A} debbono essere aggiornate mentre i vettori \mathbf{JA} e \mathbf{IA} si mantengono costanti e possono pertanto essere calcolati preventivamente al processo solutivo. Ovviamente se lo schema risolutivo che si intende adottare prevede la fattorizzazione della matrice \mathbf{A} , non si può prescindere, nella generazione di \mathbf{JA} e \mathbf{IA} , dal portare in conto anche gli elementi che pur nulli per \mathbf{A} non lo saranno a causa del fill in nella matrice \mathbf{U} che viene generata dalla fattorizzazione: come si era precedentemente anticipato, in questi casi la fattorizzazione simbolica diventa indispensabile.

- Metodi di risoluzione dei sistemi lineari

a) Metodi diretti

I metodi diretti [36-39] si basano essenzialmente sull'idea dell'eliminazione di Gauss, che consiste nel manipolare un'equazione del sistema in modo che una delle incognite venga espressa esplicitamente in funzione delle altre; sostituendo l'espressione ricavata in tutte le equazioni rimanenti, si riduce di uno l'ordine del sistema. Iterando il procedimento, il problema originario viene ricondotto ad un sistema di una equazione in una incognita, banale da invertire. Le altre componenti della soluzione possono successivamente essere calcolate attraverso una procedura di sostituzione a ritroso.

Da un punto di vista numerico esistono due concetti essenziali che possono determinare il successo o il fallimento di un processo risolutivo: il condizionamento della matrice e la stabilità del metodo.

Dalle nozioni di algebra lineare è noto che per un sistema lineare

$$\mathbf{Ax}=\mathbf{b} \quad (1)$$

di n equazioni indipendenti in n incognite la soluzione esiste ed è univocamente determinata, purchè \mathbf{A} abbia determinante non nullo. Da un punto di vista numerico le cose possono andare in maniera diversa. E' difatti possibile che, causa gli arrotondamenti di macchina, il sistema, originariamente non singolare, possa diventarlo. Questa situazione va sotto il nome di malcondizionamento della matrice. Una strategia per ovviare può essere quella di trovare un'espressione equivalente del sistema di partenza, ma più lontana dalla condizione di singolarità, attraverso la moltiplicazioni con matrici opportune. Questa trasformazione corrisponde fisicamente a cambiare le unità di misura delle grandezze in gioco, e matematicamente a riscaldare tanto i coefficienti della matrice quanto i termini noti e le variabili incognite, in modo tale che il sistema risultante sia meglio condizionato. E' importante rimarcare il

fatto che, benchè matematicamente equivalenti, differenti formulazioni possono dare risultati numerici differenti.

Un algoritmo basato sul metodo di sostituzione, che elimini in ordine la prima equazione e successivamente le altre, si arresterà nel caso in cui uno dei coefficienti della diagonale principale sia nullo. Sono comuni, inoltre, situazioni in cui l'algoritmo, a causa della precisione finita della macchina, produce soluzioni completamente errate anche se è chiamato ad operare su sistemi ben condizionati. Per ovviare a questo problema, si usa una tecnica che va sotto il nome di pivoting. L'idea di base è quella di scegliere, a ciascun passo k , in maniera oculata la variabile da eliminare e l'equazione in cui eliminare la variabile scelta. Tra i vari criteri che possono guidare in questa scelta si cita quello adoperato nel cosiddetto pivoting scalato. Definito col termine di grandezza della riga, d_i , l'elemento in valore assoluto maggiore che si trova sulla i -ma riga, al generico passo k dell'eliminazione si assume come equazione pivotale quella tra le $n-k$ rimanenti che ha il coefficiente dell'incognita x_k , che si vuole eliminare, più grande in modulo quando rapportato alla grandezza della riga. In altre parole si cerca l'indice p compreso tra k e n per il quale:

$$\frac{a_{pk}}{d_p} \geq \frac{a_{ik}}{d_i} \quad \forall i = k, \dots, n$$

Annotiamo che su matrici simmetriche definite positive non è necessario, almeno teoricamente, usare la strategia del pivoting in quanto i minori principali sono tutti diversi da zero.

Durante la soluzione di sistemi lineari ben difficilmente si opera direttamente sulla matrice di partenza. Molto più frequentemente, infatti, si preferisce decomporla in matrici di struttura più semplice -diagonale o triangolare-, per le quali esistono implementazioni del procedimento di eliminazione di Gauss particolarmente efficienti.

Tra le varie possibili strategie di fattorizzare una generica matrice **A** citiamo:

- 1) La decomposizione **LU** dove **L** è una matrice triangolare inferiore, e **U** una matrice triangolare superiore.
- 2) La decomposizione **LDM^T**, dove **L** ed **M** sono matrici triangolari inferiori con elementi unitari sulla diagonale principale e **D** è una matrice diagonale.
- 3) Le decomposizioni di Doolittle e di Crout, analoghe alla decomposizione **LU**, con la differenza che **L** nel metodo di Doolittle ed **U** nel metodo di Crout presentano elementi tutti unitari sulla diagonale principale.
- 4) La decomposizione, possibile solo se **A** è simmetrica, **LDL^T**, analoga alla 2) con **M=L**.
- 5) La decomposizione **RR^T**, detta di Cholesky, possibile se **A** è simmetrica e definita positiva, con **R** matrice triangolare inferiore con elementi tutti positivi sulla diagonale.

Esistono due grossi limiti all'utilizzo dei metodi diretti:

- 1) Le decomposizioni e il pivoting in generale distruggono la struttura originaria della matrice, e ciò risulta particolarmente fastidioso per matrici di tipo sparso. Questa osservazione non è valida per le matrici bandate a dominanza diagonale e per le matrici bandate simmetriche e definite positive che non necessitano dell'operazione di pivoting e possono essere decomposte con algoritmi che conservano la struttura bandata anche nelle matrici fattori,

- 2) Un limite ulteriore all'impiego dei metodi diretti è l'ingente costo computazionale associato al processo di fattorizzazione della matrice. Per matrici piene esso risulta proporzionale a N^3 (in particolare, se si adotta la fattorizzazione di Cholesky vale $N^3/3$), dove N è il numero delle incognite, che, nei problemi FEM di interesse pratico, risulta tipicamente dell'ordine di $1E3-1E5$. Una stima del tempo occorrente per risolvere un sistema lineare siffatto si ottiene facilmente tenendo conto che la velocità dei minicomputers, le macchine comunemente adoperate nei laboratori di calcolo scientifico è attualmente dell'ordine delle decine di Megaflops (1 Megaflop corrisponde ad

un milione di moltiplicazioni per secondo). Il tempo di elaborazione richiesto per l'inversione di un problema di interesse pratico può quindi essere dell'ordine dei mesi, il che è, ovviamente, assolutamente inaccettabile. Se le matrici sono di tipo sparso, l'impiego di metodi di banda rende il costo del processo di fattorizzazione proporzionale a $m_A N^2$, dove m_A è l'ampiezza di banda della matrice che tipicamente è di ordine 100; seppure l'impiego di metodi di inviluppo renda il costo computazionale della fattorizzazione ancora minore, questa tuttavia conserva un peso notevole nell'economia complessiva del processo solutivo. Risulta per questo motivo conveniente adoperare i metodi diretti nei casi in cui l'ordine del sistema non sia particolarmente elevato e/o quando il problema richieda di risolvere un grande numero di sistemi di equazione, nei quali il termine noto cambi mentre la matrice dei coefficienti si conservi costante; in casi diversi, possono essere preferiti schemi iterativi di soluzione.

b) Metodi iterativi

Per matrici dotate di sparsità non strutturale e di dimensione elevata, un'alternativa ai metodi diretti è costituita dagli schemi di soluzione iterativi che lasciano inalterata la struttura originale della matrice e non richiedono il processo di fattorizzazione. Essendo, tuttavia, in questo caso la soluzione ottenuta come limite di una successione, per essere un'alternativa valida ai metodi diretti, essi hanno bisogno di opportune tecniche di accelerazione.

L'idea comune a tutti i metodi iterativi è la seguente.

1) Si dà una stima del vettore soluzione $x^{(0)}$ del sistema $\mathbf{Ax}=\mathbf{b}$, dove \mathbf{A} è una matrice sparsa di ordine N non singolare.

2) Si costruisce una successione di soluzioni $x^{(k)}$ attraverso la successiva risoluzione di sistemi lineari più semplici.

In forma generale consideriamo una decomposizione della matrice \mathbf{A} della forma: $\mathbf{A}=\mathbf{M}-\mathbf{N}$ con \mathbf{M} non singolare.

Si ha allora:

$$\mathbf{Ax}=\mathbf{b} \Leftrightarrow \mathbf{Mx}=\mathbf{Nx} + \mathbf{b} \quad (2)$$

da cui il procedimento iterativo : $\mathbf{M}\mathbf{x}^{(k+1)} = \mathbf{N}\mathbf{x}^{(k)} + \mathbf{b}$ (3)

La matrice $\mathbf{B} = \mathbf{M}^{-1}\mathbf{N} = \mathbf{M}^{-1}(\mathbf{M}-\mathbf{A}) = \mathbf{I} - \mathbf{M}^{-1}\mathbf{A}$ (4)

è detta *matrice di iterazione*; essa individua un particolare metodo e le sue proprietà sono fondamentali per la qualità e la rapidità della convergenza del processo iterativo.

Facendo riferimento alla decomposizione $\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{F}$ (vedi fig.[5]):

$$\mathbf{A} = \begin{bmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & \mathbf{D} & & -\mathbf{F} & \\ & & & & & \ddots \\ & -\mathbf{E} & & \ddots & & \\ & & & & & \ddots \end{bmatrix} \quad (5)$$

Fig.[5]: Decomposizione adottata per la matrice **A** negli schemi di Jacobi e Gauss-Seidel

citiamo alcuni tra gli schemi iterativi più comunemente impiegati:

-Il metodo di Jacobi, o *delle sostituzioni simultanee* ove si assume

$$\mathbf{M} = \mathbf{D}; \quad \mathbf{N} = \mathbf{E} + \mathbf{F}; \quad (6)$$

e quindi: $\mathbf{B}_j = \mathbf{D}^{-1}(\mathbf{E} + \mathbf{F}) = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A}$ (7)

-Il metodo di Gauss Seidel, o *delle sostituzioni successive* ove si assume

$$\mathbf{M} = \mathbf{D} - \mathbf{E}; \quad \mathbf{N} = \mathbf{F}; \quad (8)$$

e quindi: $\mathbf{B}_j = (\mathbf{D} - \mathbf{E})^{-1}\mathbf{F} = (\mathbf{I} - \mathbf{D}^{-1}\mathbf{E})^{-1}\mathbf{D}^{-1}\mathbf{F}$ (9)

Con l'obiettivo di accelerare il processo di convergenza i metodi precedenti possono essere modificati introducendo un parametro reale ω e definendo la componente i -ma della soluzione al $(k+1)$ -mo passo iterativo come:

$$x_i^{(k+1)} = \omega x_i^{(k+1/2)} + (1-\omega) x_i^{(k)} \quad (10)$$

In questo modo, partendo dal metodo di Jacobi si perviene al seguente schema:

$$\mathbf{x}^{(k+1)} = (\mathbf{I} - \omega \mathbf{D}^{-1} \mathbf{A}) \mathbf{x}^{(k)} + \omega \mathbf{D}^{-1} \mathbf{b} \quad (11)$$

Mentre partendo dal metodo di Gauss-Seidel si giunge al metodo di rilassamento (detto anche SOR) in cui si assume:

$$\mathbf{M} = \left(\frac{\mathbf{D}}{\omega} - \mathbf{E} \right); \quad \mathbf{N} = \left(\frac{1}{\omega} - 1 \right) \mathbf{D} + \mathbf{F}; \quad (12)$$

a cui corrisponde la seguente matrice di iterazione:

$$\mathbf{B}_\omega = (\mathbf{D} - \omega \mathbf{E})^{-1} [(1 - \omega) \mathbf{D} + \omega \mathbf{F}] \quad (13)$$

Circa la convergenza di questi schemi iterativi della successione $\mathbf{x}^{(k)}$ alla soluzione \mathbf{x} valgono i seguenti teoremi :

-Se \mathbf{A} è una matrice di ordine N a stretta predominanza diagonale, allora i metodi di Jacobi e di Gauss-Seidel sono convergenti.

-Se \mathbf{A} è una matrice hermitiana definita positiva, allora il metodo di rilassamento converge se e solo se $0 < \omega < 2$.

Per quanto riguarda la velocità di convergenza, essa dipende in maniera critica (e questo è il più grande inconveniente del metodo SOR) dalla scelta del fattore di rilassamento. Young, a questo proposito, ha mostrato [40] come la scelta ottimale sia:

$$\omega_b = \frac{2}{1 + (1 - z^2)^{\frac{1}{2}}} \quad (14)$$

dove z è il più grande autovalore di $\mathbf{D}^{-1}(\mathbf{E} + \mathbf{F})$.

Metodo del gradiente coniugato

Un altro schema iterativo che può essere utilizzato vantaggiosamente nella soluzione di problemi FEM è quello del Gradiente Coniugato. Questo metodo è stato introdotto per la prima volta da Hestenes e Stiefel nel 1952 [41], ed ha il vantaggio (puramente teorico) di risolvere esattamente un sistema di ordine N , in al più N iterazioni (in pratica, causa gli errori di arrotondamento, questa proprietà viene persa, ed inoltre N iterazioni sono un costo computazionale usualmente troppo alto da pagare).

Il metodo del gradiente coniugato sfrutta il fatto che risolvere il sistema (1) è equivalente a minimizzare il funzionale quadratico:

$$\mathbf{Q}(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{x}^T \mathbf{b} \quad (15)$$

o in maniera equivalente il funzionale:

$$\mathbf{E}(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{A} (\mathbf{x} - \bar{\mathbf{x}}) = \mathbf{e}^T \mathbf{A} \mathbf{e}$$

dove $\bar{\mathbf{x}}$ è la soluzione esatta del sistema (1), ed $\mathbf{e} = \mathbf{x} - \bar{\mathbf{x}}$.

Qui si vede che \mathbf{A} deve essere necessariamente definita positiva altrimenti $\mathbf{E}(\mathbf{x})$ non è minimo per $\mathbf{e} = \mathbf{0}$.

La minimizzazione di (16) viene condotta per via iterativa scegliendo al passo k una direzione $\mathbf{p}_k \neq \mathbf{0}$ e uno scalare α_k in maniera che, posto

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (16)$$

si abbia $\mathbf{E}(\mathbf{x}_{k+1}) \leq \mathbf{E}(\mathbf{x}_k)$.

Fissata la direzione \mathbf{p}_k e definito il residuo $\mathbf{r}_k(\mathbf{x}) = \mathbf{b} - \mathbf{A} \mathbf{x}_k = \mathbf{A} (\bar{\mathbf{x}} - \mathbf{x}_k)$ si dimostra che la scelta ottimale per α_k è

$$\alpha_k = \frac{(\mathbf{r}_k, \mathbf{p}_k)}{(\mathbf{A} \mathbf{p}_k, \mathbf{p}_k)} \quad (17)$$

Per quel che concerne la scelta delle direzioni \mathbf{p}_k ricordiamo, innanzitutto, che due vettori \mathbf{x} e \mathbf{y} si dicono coniugati rispetto all'applicazione bilineare \mathbf{A} se vale la relazione $(\mathbf{A} \mathbf{x}, \mathbf{y}) = 0$.

Il nome dell'algorithm è così legato al criterio di ortogonalità rispetto al prodotto scalare $(\mathbf{A} \mathbf{p}_i, \mathbf{p}_k) \forall i=1, k-1$ col quale, al k -mo passo di iterazione, viene scelta la direzione lungo le quali effettuare la minimizzazione. Da un punto di

vista pratico, se si adotta per α_k la scelta (17), la direzione \mathbf{p}_k è determinata attraverso le relazioni:

$$\mathbf{p}_k = \mathbf{r}_k + \beta_k \mathbf{p}_{k-1}; \quad (18) \quad \text{con} \quad \beta_k = -\frac{(\mathbf{A}\mathbf{p}_{k-1}, \mathbf{r}_k)}{(\mathbf{A}\mathbf{p}_{k-1}, \mathbf{p}_{k-1})}; \quad (19)$$

Per quanto riguarda il numero di operazioni richieste si vede che il costo essenziale è dovuto al calcolo di $\mathbf{A}\mathbf{p}$, il quale per matrici piene è pari a N^2 mentre per matrici sparse non strutturate è uguale a $nsize$ (numero degli elementi non nulli della matrice). Se il numero di iterazioni è N , il costo totale per matrici piene risulta pari a N^3 superiore a quello di $N^3/3$ richiesto nel processo di fattorizzazione di Cholesky. Appare chiaro, dunque, che il metodo diventa interessante quando la matrice è sparsa e/o quando il numero di iterazioni necessarie è decisamente inferiore a N . A questo scopo risulta molto spesso di grande efficacia l'accorgimento di preconditionare la matrice \mathbf{A} attraverso la moltiplicazione con una matrice non-singolare \mathbf{R} . Si può, infatti, osservare che se la matrice \mathbf{R} viene scelta in maniera opportuna (\mathbf{R} quanto più possibile vicina ad \mathbf{A}^{-1}), l'algoritmo del gradiente coniugato convergerà molto più velocemente alla soluzione del sistema equivalente

$$\mathbf{R}^{-1}\mathbf{A}(\mathbf{R}^{-1})^T (\mathbf{R}^T\mathbf{x}) = (\mathbf{R}^{-1}\mathbf{b}) \Leftrightarrow \mathbf{R}^{-1}\mathbf{A}\mathbf{x} = (\mathbf{R}^{-1}\mathbf{b}) \Leftrightarrow \mathbf{M}^{-1}\mathbf{A}\mathbf{x} = (\mathbf{M}^{-1}\mathbf{b}) = \mathbf{s} \quad (20)$$

ove, essendo \mathbf{R} non singolare, $\mathbf{M} = \mathbf{R}\mathbf{R}^T$ è simmetrica e definita positiva.

Il calcolo di \mathbf{s} deve essere eseguito ad ogni iterazione e non deve, naturalmente, passare per l'inversione di \mathbf{M} . E' pertanto essenziale scegliere \mathbf{M} in modo tale che il sistema lineare $\mathbf{M}\mathbf{s} = \mathbf{b}$ possa essere risolto in maniera economica. Tra le varie possibili scelte di \mathbf{M} si ricordano (vedi fig [5]):

- $\mathbf{M} = \mathbf{D}$ che dà origine al *metodo del gradiente coniugato di Jacobi*.
- $\mathbf{M} = \frac{1}{\omega(2-\omega)} (\mathbf{D} - \omega\mathbf{F}^T) \mathbf{D}^{-1} (\mathbf{D} - \omega\mathbf{E}^T)$, usata nel metodo SSOR, ove, a differenza del metodo SOR, la scelta di ω non è più critica.

- Un approccio leggermente differente è quello di scegliere \mathbf{R} in modo che sia un'approssimazione della matrice triangolare inferiore \mathbf{L} che risulta dalla fattorizzazione di Cholesky di \mathbf{A} . Se per l'immagazzinamento di \mathbf{A} si è adoperato un metodo di matrici sparse, allora può essere conveniente calcolare \mathbf{R} attraverso una fattorizzazione parziale di Cholesky di \mathbf{A} con un metodo di eliminazione Gaussiana, in cui il *fill in* è trascurato, oppure portato in conto solo in zone limitate della matrice fattorizzata (ad esempio all'interno di un fissato numero di diagonali localizzate attorno la diagonale principale). Allo scopo di limitare l'insorgere di mal-condizionamenti per effetto degli arrotondamenti numerici e di evitare il fallimento del processo di fattorizzazione, può risultare conveniente scalare la matrice \mathbf{A} in modo da riportare i termini diagonali a valori unitari e, eventualmente, moltiplicare i termini fuori diagonale per il fattore $1/(1 + \alpha)$, ove α è un numero reale positivo.

3.2.2) Risoluzione di sistemi di equazioni non lineari

- Generalità

Mentre per i sistemi di equazioni lineari la soluzione può essere trovata tanto con metodi diretti quanto per via iterativa, nei sistemi di equazioni non lineari il vettore delle radici viene, in ogni caso, determinato per via iterativa; e, dunque, diventano determinanti nella scelta del metodo risolutivo le proprietà di applicabilità e di convergenza.

A differenza dei sistemi lineari, in questo tipo di problema non è garantita né l'esistenza (almeno nel campo reale) né l'unicità della soluzione: infatti, poichè i sistemi non-lineari sono usualmente posti nella forma implicita, $\mathbf{F}(\mathbf{x})=\mathbf{0}$, vale, a proposito dell'esistenza della soluzione, il teorema del Dini che, nel caso di funzioni continue, dà qualche speranza, ma non la certezza, di poter soddisfare simultaneamente N equazioni in N incognite[39]. Inoltre, nei problemi multi-dimensionali, anche nel caso in cui sia dimostrabile l'esistenza di una soluzione, la convergenza ad essa può dipendere criticamente dalla

scelta del punto di partenza. Presentate subito le difficoltà che si possono incontrare nel risolvere un generico sistema non-lineare, va subito aggiunto che le cose si semplificano in maniera significativa se il sistema non-lineare proviene dalla minimizzazione di un funzionale convesso. In questo caso, infatti, le equazioni del sistema non sono funzioni indipendenti, ma le componenti del vettore gradiente del funzionale. Esse devono pertanto obbedire alle condizioni di integrabilità che sono altamente restrittive. In parole semplici, è sempre possibile trovare il minimo di un funzionale convesso scendendo lungo la direzione più ripida, il che ci rassicura circa l'esistenza di uno zero del vettore gradiente.

- Problemi non-lineari mono-dimensionali

Il metodo più semplice per localizzare le radici di una funzione mono-dimensionale non-lineare e continua, $f(x)$, consiste nel tabellare la funzione con passo opportuno e di individuare due valori di x , che denoteremo con i simboli a e b , per i quali la funzione $f(x)$ assuma valori di segno opposto. Per il *teorema degli zeri delle funzioni continue* la radice dell'equazione deve essere compresa all'interno dell'intervallo aperto $]a,b[$. Una volta localizzato l'intervallo, si può applicare uno tra i diversi algoritmi monodimensionali di ricerca presenti in letteratura. Di seguito ne vengono presentati alcuni, insieme con le loro caratteristiche di convergenza. Per tutti l'idea di base comune è quella di ridurre l'ampiezza dell'intervallo fino a che non sia soddisfatto un test di tolleranza sulla soluzione ottenuta ($|b_k - a_k| < \epsilon_1$ o $f(x_k) < \epsilon_2$) oppure non sia stato raggiunto il numero massimo di iterazioni consentite dall'utente.

- L'*algoritmo di bisezione* genera una successione di intervalli (a_k, b_k) con:

$$]a_k, b_k[\subset]a_{k-1}, b_{k-1}[\quad \text{e} \quad |b_k - a_k| = \frac{1}{2} |b_{k-1} - a_{k-1}|$$

Una volta raggiunto il criterio di arresto, si assume come soluzione la media degli estremi cui si è ristretto l'intervallo durante l'ultima

iterazione: $x_{k+1} = \frac{a_k + b_k}{2}$; che comporta un'errore pari a: $e_k = \frac{b_k - a_k}{2}$. Il tasso di convergenza viene determinato come $\frac{|e_{k+1}|}{|e_k|} \approx \frac{1}{2}$ ed è chiaramente lineare.

- Mentre l'algoritmo di bisezione tiene conto della sola informazione sul segno, l'algoritmo che va sotto il nome di *regula falsi*, utilizza anche i valori che la funzione assume negli estremi dell'intervallo. Ad ogni nuova iterazione il valore di x viene determinato attraverso interpolazione lineare:

$$x_{k+1} = \frac{a_k f(b_k) + b_k f(a_k)}{f(b_k) - f(a_k)} \quad (1)$$

Si può dimostrare che questo metodo presenta una convergenza di carattere superlineare, ma la stima del suo esatto ordine è piuttosto complicata.

- Una variante del metodo precedente è rappresentata dal *metodo delle secanti*, nel quale per il calcolo dell'interpolazione lineare si utilizzano gli ultimi due punti calcolati. Questo metodo ha un tasso di convergenza maggiore di entrambi i metodi precedenti e pari al cosiddetto "numero d'oro", 1.618..., ma presenta lo svantaggio di non fornire necessariamente un'intervallo che contenga la radice. Per funzioni non sufficientemente regolari la convergenza non è pertanto garantita.

- Problemi non-lineari multi-dimensionali

I metodi illustrati precedentemente sono tutti concepiti per trattare il solo caso monodimensionale. Nel caso in cui sia richiesto di trovare le radici di un sistema pluri-dimensionale $\mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_N(\mathbf{x})]$, in cui il vettore delle incognite $\mathbf{x} \in \mathbb{R}^N$, bisogna necessariamente ricorrere all'algoritmo, anch'esso di carattere iterativo, detto di Newton-Raphson. In questo schema è necessario saper valutare nel generico punto \mathbf{x} oltre che le funzioni $f_i(\mathbf{x})$ anche lo Jacobiano $\mathbf{A}(\mathbf{x})$ di $\mathbf{F}(\mathbf{x})$. L'idea di base dell'algoritmo consiste nell'espandere, alla $k+1$ -ma iterazione, $\mathbf{F}(\mathbf{x})$ in serie di Taylor. Indicando con $\mathbf{H}(\mathbf{x})$ la matrice Hessiana di $\mathbf{F}(\mathbf{x})$ si ha pertanto:

$$\mathbf{F}(\mathbf{x}_k + d\mathbf{x}) = \mathbf{F}(\mathbf{x}_k) + \mathbf{A}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_k} d\mathbf{x} + \frac{1}{2} \mathbf{H}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_k} d\mathbf{x}^2; \quad \text{con } \mathbf{x} \in [\mathbf{x}_k, \mathbf{x}_k + d\mathbf{x}] \quad (2)$$

Trascurando i termini del secondo ordine la (2) diventa:

$$\mathbf{F}(\mathbf{x}_k + d\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_k) + \mathbf{A}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_k} d\mathbf{x} = \mathbf{0} \quad (3)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + d\mathbf{x} \quad (4)$$

Il sistema lineare (3) può essere risolto utilizzando tra gli algoritmi presentati nel paragrafo precedente quello che risulta più vantaggioso in base alla struttura di $\mathbf{A}(\mathbf{x})$.

La potenza del metodo di Newton giace nel fatto che il suo tasso di convergenza è quadratico, vale a dire che *nelle vicinanze di una radice* il numero di cifre significative esatte approssimativamente raddoppia ad ogni step.

Anche questo metodo presenta, tuttavia, degli inconvenienti. In primo luogo la sua applicabilità dipende dal fatto che lo Jacobiano sia continuo e diverso da zero in prossimità delle radici. Inoltre, lontano dalle radici, dove i termini dello sviluppo di ordine superiore acquistano importanza rilevante, il metodo di Newton può dare grosse correzioni del tutto prive di significato. Infine, soprattutto nel caso in cui le funzioni $f_i(\mathbf{x})$ siano costruite per interpolazione non lineare di valori tabellati, possono insorgere patologie che determinano il completo fallimento dello schema solutivo.

Un modo per ridurre il rischio di insorgenza di situazioni patologiche e per migliorare ulteriormente il tasso di convergenza, è quello di assumere per la soluzione al passo $k+1$, anziché il valore fornito dalla (4), quello determinato dall'espressione:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + a d\mathbf{x} \quad (5)$$

dove a è l'incognita scalare del problema monodimensionale non-lineare:

$$\mathbf{G}(\mathbf{x}_k + a d\mathbf{x}) = \mathbf{F}(\mathbf{x}_k + a d\mathbf{x}) \times d\mathbf{x} = \mathbf{0} \quad (6)$$

che può essere risolto con il metodo *regula falsi*, oppure con il *metodo delle secanti*.

3.3) Metodi agli elementi finiti: gli edge elements.

Si è già detto, introducendo le formulazioni agli elementi finiti, che i gradi di libertà, nei termini dei quali il problema di interesse deve essere invertito, sono associati ai cosiddetti *simplices* (nodi, lati, facce, volumi) della mesh.

Le formulazioni tradizionali basate sui metodi agli elementi finiti associano i gradi di libertà ai nodi della mesh. Se il dominio di definizione del problema è costituito da materiali, le proprietà costitutive dei quali variano bruscamente, e l'incognita è un campo vettoriale che deve soddisfare il vincolo di continuità sulla componente tangente nelle zone di interfaccia tra i materiali, le soluzioni ottenute attraverso approcci che adottano i nodi come incognite possono rivelarsi poco precise.

Risultati decisamente migliori possono essere ottenuti adottando formulazioni basate sui lati della mesh, e per questo comunemente etichettate col nome di "Edge elements". In questo caso i gradi di libertà degli edge elements sono associati alle componenti del campo vettoriale lungo gli spigoli del generico elemento della mesh[42]. In tal modo la continuità delle componenti tangenti risulta automaticamente verificata, mentre nessun vincolo è imposto sulle componenti normali dei campi che possono perciò essere discontinue.

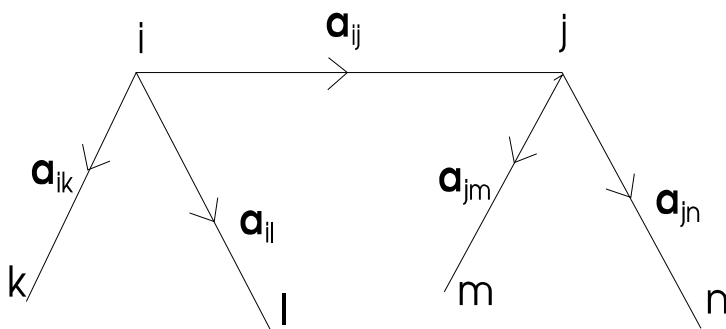


Fig.[6] Tre lati per nodo

Si consideri il caso di un elemento del primo ordine per il quale ogni nodo appartiene a tre lati. Con riferimento alla fig.[6] la funzione di forma di un lato per un elemento siffatto può essere determinata nel modo seguente [43,44]:

Per il generico lato a_{ij} dell'elemento master si definiscano due vettori, che sono rispettivamente legati ai nodi i e j :

$$\mathbf{s}_{ij} = \psi_i \nabla \Pi(\mathbf{a}_{ik}, \mathbf{a}_{il}) \quad (1a)$$

$$\mathbf{s}_{ji} = \psi_j \nabla \Pi(\mathbf{a}_{jm}, \mathbf{a}_{jn}) \quad (1b)$$

dove ψ_i è la funzione di forma scalare isoparametrica associata al nodo i (che vale 1 nel nodo i e 0 in tutti gli altri nodi dell'elemento), mentre $\Pi(\mathbf{a}, \mathbf{b})=0$ è l'equazione del piano definito dai vettori \mathbf{a} e \mathbf{b} . Il vettore $\nabla \Pi$ risulta, evidentemente, perpendicolare ai vettori \mathbf{a} e \mathbf{b} . Per la dipendenza da ψ_i , \mathbf{s}_{ij} è un vettore nullo su tutti i lati dell'elemento eccetto su quelli che partono dal nodo i . Inoltre, grazie all'operatore gradiente, \mathbf{s}_{ij} risulta perpendicolare ai lati \mathbf{a}_{ik} e \mathbf{a}_{il} . Per questi motivi, l'integrale di linea della componente tangenziale \mathbf{s}_{ij} lungo tutti i lati, fatta eccezione per il lato \mathbf{a}_{ij} vale 0. Questa proprietà è ugualmente vera per \mathbf{s}_{ji} . Pertanto,

$$\mathbf{N}_{ij} = c \mathbf{s}_{ij} - d \mathbf{s}_{ji}; \quad (2)$$

è la funzione di forma cercata per il lato \mathbf{a}_{ij} . E' possibile determinare i coefficienti scalari c e d imponendo:

$$|c|=|d| \quad (3a) \quad \int_{lato \mathbf{a}_{ij}} \mathbf{N}_{ij} \cdot d\mathbf{l} = 1 \quad (3b)$$

Inoltre tenendo conto che, se f e g sono dei generici campi scalari, valgono le identità:

$$\nabla \times (\nabla f) = 0; \quad (4a) \quad \nabla \times (f \nabla g) = f \nabla \times (\nabla g) + \nabla g \times \nabla f = \nabla g \times \nabla f; \quad (4b)$$

si ha per il rotore di \mathbf{N}_{ij} :

$$\begin{aligned}\nabla \times \mathbf{N}_{IJ} &= \nabla \times (\mathbf{c} \mathbf{s}_{ij} - \mathbf{d} \mathbf{s}_{ji}) = \mathbf{c} \nabla \times (\psi_i \nabla \Pi(\mathbf{a}_{ik}, \mathbf{a}_{il})) - \mathbf{d} \nabla \times (\psi_j \nabla \Pi(\mathbf{a}_{jm}, \mathbf{a}_{jn})) = \\ &= \mathbf{c} \nabla \psi_i \times \nabla \Pi(\mathbf{a}_{ik}, \mathbf{a}_{il}) - \mathbf{d} \nabla \psi_j \times \nabla \Pi(\mathbf{a}_{jm}, \mathbf{a}_{jn});\end{aligned}\quad (5)$$

Ad esempio, per un elemento tetraedrico, la funzione di forma associata allo spigolo avente per estremi i nodi i e j e il suo rotore si esprimono come:

$$\mathbf{N}_{IJ} = \psi_i \nabla \psi_j - \psi_j \nabla \psi_i; \quad (6a) \quad \nabla \times \mathbf{N}_{IJ} = 2 \nabla \psi_i \times \nabla \psi_j \quad (6b)$$

Per quanto riguarda il passaggio dalle funzioni di forma locali a quelle dell'elemento reale, si tenga conto del fatto che, se $\mathbf{s}(\xi)$ è un generico campo scalare, funzione delle coordinate locali ξ , il suo gradiente, nel sistema di riferimento globale, può essere espresso come:

$$[\nabla_{\mathbf{x}} (s(\xi))]_{\mathbf{x}} = \left[\frac{\partial s}{\partial \xi} \frac{\partial \xi}{\partial x} + \frac{\partial s}{\partial \eta} \frac{\partial \eta}{\partial x} + \frac{\partial s}{\partial \zeta} \frac{\partial \zeta}{\partial x} \right] \hat{\mathbf{i}}_x = \left[\frac{\partial s}{\partial \xi} \mathbf{J}_{11}^{-1} + \frac{\partial s}{\partial \eta} \mathbf{J}_{21}^{-1} + \frac{\partial s}{\partial \zeta} \mathbf{J}_{31}^{-1} \right] \hat{\mathbf{i}}_x$$

$$[\nabla_{\mathbf{x}} (s(\xi))]_{\mathbf{y}} = \left[\frac{\partial s}{\partial \xi} \frac{\partial \xi}{\partial y} + \frac{\partial s}{\partial \eta} \frac{\partial \eta}{\partial y} + \frac{\partial s}{\partial \zeta} \frac{\partial \zeta}{\partial y} \right] \hat{\mathbf{i}}_y = \left[\frac{\partial s}{\partial \xi} \mathbf{J}_{12}^{-1} + \frac{\partial s}{\partial \eta} \mathbf{J}_{22}^{-1} + \frac{\partial s}{\partial \zeta} \mathbf{J}_{32}^{-1} \right] \hat{\mathbf{i}}_y$$

$$[\nabla_{\mathbf{x}} (s(\xi))]_{\mathbf{z}} = \left[\frac{\partial s}{\partial \xi} \frac{\partial \xi}{\partial z} + \frac{\partial s}{\partial \eta} \frac{\partial \eta}{\partial z} + \frac{\partial s}{\partial \zeta} \frac{\partial \zeta}{\partial z} \right] \hat{\mathbf{i}}_z = \left[\frac{\partial s}{\partial \xi} \mathbf{J}_{13}^{-1} + \frac{\partial s}{\partial \eta} \mathbf{J}_{23}^{-1} + \frac{\partial s}{\partial \zeta} \mathbf{J}_{33}^{-1} \right] \hat{\mathbf{i}}_z$$

e cioè, in forma compatta:

$$[\nabla_{\mathbf{x}} (s(\xi)))] = \mathbf{J}^{-T} [\nabla_{\mathbf{x}} (s(\xi)))] \quad (7)$$

dove $\mathbf{J}(\xi)$ è la matrice jacobiana dell'applicazione $\mathbf{x}=\chi^{(e)}(\xi)$ che proietta l'elemento master sull'elemento reale (e):

$$\mathbf{J}(\xi, \eta, \zeta) = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{vmatrix} \quad (8)$$

La funzione globale di edge $\mathbf{N}_i^{(e)}$, relativa al lato i per l'elemento reale (e) può essere ricavata attraverso l'uso combinato della (5) e della (7).

Il rotore della funzione globale di lato $\mathbf{N}_i^{(e)}$ relativa al lato i dell'elemento reale (e) e calcolato rispetto al sistema di riferimento globale \mathbf{x} si ricava dal rotore della corrispondente funzione locale di lato $\mathbf{N}_i^{(l)}$ valutato nel sistema di riferimento locale ξ attraverso la seguente formula:

$$\nabla_{\mathbf{x}} \times \mathbf{N}_i^{(e)} = \text{Det}[\mathbf{J}(\xi)]^{-1} (\mathbf{J}(\xi) \nabla_{\xi} \times \mathbf{N}_i^{(l)})$$

Nell'analisi dei problemi elettromagnetici capita spesso di dover modellare forme geometriche estremamente complesse. Il compito può risultare estremamente complicato, o ad ogni modo terminare in una cattiva rappresentazione del dominio originale, se non si dispone di elementi di libreria di diverse caratteristiche geometriche. Mentre, in generale, l'uso degli esaedri produce risultati qualitativamente migliori[45-46], in alcuni casi (ad esempio per problemi a simmetria

rotazionale) l'impiego di elementi prismatici o tetraedrale può diventare essenziale[44].

Se si dispone, inoltre, di elementi lineari e quadratici della medesima forma geometrica, è possibile migliorare la qualità dei risultati lasciando inalterato il numero degli elementi e sostituendo gli elementi di ordine superiore a quelli lineari. E' questa la logica del cosiddetto p-refinement[36].

Le espressioni delle funzioni di forma nodale, delle funzioni di lato e dei rotori delle funzioni di lato per il tetraedro, il pentaedro e l'esaedro di ordine 1 e per l'esaedro di ordine 2 sono riportati nell'appendice A.

3.4) Formulazione numerica del metodo della minimizzazione dell'errore costitutivo.

- Discretizzazione numerica dei potenziali vettori incogniti

Per il metodo basato sulla minimizzazione dell'errore costitutivo viene proposta un'implementazione basata sugli Elementi Finiti nel dominio spaziale e sulle differenze finite nel dominio temporale.

I potenziali vettori a 3 componenti $\mathbf{A}(\mathbf{x},t)$ e $\mathbf{W}(\mathbf{x},t)$ sono approssimati usando le funzioni di forma spaziale associate agli edge-elements, che permettono la gestione automatica di disomogeneità e discontinuità spaziali.

$$\mathbf{A}(\mathbf{x},t) = \sum_{i=1}^{Nedges} A_i(t) \mathbf{N}_i(\mathbf{x}) \quad (1a)$$

$$\mathbf{W}(\mathbf{x},t) = \sum_{i=1}^{Nedges} W_i(t) \mathbf{N}_i(\mathbf{x}) \quad (1b)$$

dove $Nedges$ è il numero dei lati della mesh usata per rappresentare il dominio del problema.

La soluzione è trovata partizionando l'intervallo di tempo di interesse $(0, T)$ in sotto-intervalli nei quali è assunta una interpolazione lineare. Per il generico sotto-intervallo $[t_k, t_{k+1}=t_k+\Delta t_k]$ si ha pertanto:

$$A_i(t) = \frac{a_{i,k+1} - a_{i,k}}{\Delta t_k} (t - t_k) + a_{i,k} \quad (2a)$$

$$W_i(t) = \frac{w_{i,k+1} - w_{i,k}}{\Delta t_k} (t - t_k) + w_{i,k} \quad (2b)$$

dove $a_{i,k}=A_i(t_k)$ e $w_{i,k}=W_i(t_k)$.

- Espressione del funzionale di errore nelle incognite discretizzate

Con questa scelta per la discretizzazione dei potenziali vettori incogniti il problema si riduce alla soluzione di una successione di sotto-problemi, ciascuno dei quali è definito all'interno di un intervallo elementare. Le condizioni finali incognite di un time-step sono usate come condizioni iniziali per il time-step successivo, il che equivale ad imporre la continuità nel tempo di \mathbf{A} e \mathbf{W} (che è una proprietà intrinseca risultante dalla loro definizione, eq.(2a-b)).

La soluzione per ogni sotto-problema può così essere ottenuta attraverso la minimizzazione di un errore globale in $V \times [t_k, t_{k+1}]$, definito attraverso la (2.2.2-2), e di una conveniente espressione dell'errore locale, come ad esempio la (2.2.2-1).

- Formulazione discreta del problema di minimizzazione del funzionale di errore

La composizione delle equazioni governanti inizia attraverso la sostituzione nella (2.2.2-2) dei campi con le espressioni (2.2.1-11) dei potenziali vettori. I potenziali vettori sono descritti attraverso le posizioni (1). Il funzionale (2.2.2-2) è trasformato in questa maniera da integro-differenziale in algebrico nelle incognite potenziali vettori, \mathbf{A} e \mathbf{W} , la dimensione dei quali risulta pari al numero di lati della mesh.

Il problema si riduce quindi a cercare i valori di \mathbf{A} e \mathbf{W} all'istante t_{k+1} che rendono minima la restrizione del funzionale globale di errore Λ nel time-step $[t_k, t_{k+1}]$. Nel seguito indicheremo con \mathbf{a} e \mathbf{w} rispettivamente i vettori $\mathbf{A}(t_{k+1})$ e $\mathbf{W}(t_{k+1})$ e con Λ_k la restrizione del funzionale globale di errore all'intervallo $[t_k, t_{k+1}]$. Una formulazione equivalente del problema è quella di cercare i valori di \mathbf{a} e \mathbf{w} che annullano lo Jacobiano di $\Lambda_k(\mathbf{a}, \mathbf{w})$:

$$\min \Lambda_k(\mathbf{a}, \mathbf{w}) \quad \rightarrow \quad \begin{bmatrix} \frac{\partial \Lambda_k}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_k}{\partial \mathbf{w}} \end{bmatrix} = 0 \quad (3)$$

dove, se si adotta per il funzionale locale di errore l'espressione (2.2.2-2), le derivate di Λ_k rispetto alle incognite del lato r della mesh si scrivono:

$$\begin{aligned} \frac{\partial \Lambda_k}{\partial \mathbf{w}_r^{k+1}} = & \alpha_E \iiint_V \nabla \times \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k} [\mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t) - \mathbf{E}] dt + \\ & + \alpha_H \iiint_V \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{[\mathbf{f}_B(\mathbf{H}, \mathbf{x}, t) - \mathbf{B}]}{\Delta t_k} dt \end{aligned} \quad (4a)$$

$$\begin{aligned} \frac{\partial \Lambda_k}{\partial \mathbf{a}_r^{k+1}} = & \alpha_E \iiint_V \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{[\mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t) - \mathbf{D}_T]}{\Delta t_k} dt + \\ & + \alpha_H \iiint_V \nabla \times \mathbf{N}_r d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k} [\mathbf{f}_H(\mathbf{B}, \mathbf{x}, t) - \mathbf{H}] dt \end{aligned} \quad (4b)$$

mentre i coefficienti della matrice Hessiana di Λ_k valgono:

$$\begin{aligned} \frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \partial \mathbf{a}_s} = & \iiint_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \alpha_E \left[\frac{\mathbf{N}_r}{\Delta t_k^2} \cdot \left(\frac{\partial \mathbf{f}_{D_T}(\mathbf{E}, \mathbf{x}, t)}{\partial \mathbf{E}} \mathbf{N}_s \right) \right] + \alpha_H \left[\frac{(t-t_k)^2}{\Delta t_k^2} \nabla \times \mathbf{N}_r \cdot \left(\frac{\partial \mathbf{f}_H(\mathbf{B}, \mathbf{x}, t)}{\partial \mathbf{B}} \nabla \times \mathbf{N}_s \right) \right] dt \end{aligned} \quad (5a)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \partial \mathbf{w}_s} = \int_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k^2} \left\{ \alpha_E [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_H [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \right\} dt \quad (5b)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \mathbf{a}_s} = \int_V d\mathbf{x} \int_{t_k}^{t_{k+1}} -\frac{(t-t_k)}{\Delta t_k} \left\{ \alpha_H [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_E [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \right\} dt \quad (5c)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \mathbf{w}_s} = \int_V d\mathbf{x} \int_{t_k}^{t_{k+1}} \alpha_E \left[\frac{(t-t_k)^2}{\Delta t_k^2} \nabla \times \mathbf{N}_r \cdot \left(\frac{\partial \mathbf{f}_E(\mathbf{D}_T, \mathbf{x}, t)}{\partial \mathbf{D}_T} \nabla \times \mathbf{N}_s \right) \right] + \alpha_H \left[\frac{\mathbf{N}_r}{\Delta t_k^2} \cdot \left(\frac{\partial \mathbf{f}_B(\mathbf{H}, \mathbf{x}, t)}{\partial \mathbf{H}} \mathbf{N}_s \right) \right] dt \quad (5d)$$

La matrice Hessiana ci tornerà utile quando, attraverso il metodo di Newton, si andrà a determinare la soluzione del sistema di equazioni.

- Imposizione delle condizioni al contorno

Partizionata la superficie $\partial\Omega$ del dominio di definizione del problema in quattro zone a seconda delle condizioni al contorno imposte su ciascuna:

$$\partial\Omega = \partial\Omega_H \cup \partial\Omega_E \cup \partial\Omega_B \cup \partial\Omega_{D_T}$$

dalle condizioni al contorno imposte rispettivamente per $\partial\Omega_E$ e $\partial\Omega_H$ sulle componenti tangenti di \mathbf{E} ed \mathbf{H} si risale per integrazione temporale alle corrispondenti condizioni per i potenziali vettori.

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}^*(\mathbf{x}, t) = -\int_0^t \mathbf{E}^*(\mathbf{x}, t') dt'$$

$$\mathbf{W}(\mathbf{x}, t) = \mathbf{W}^*(\mathbf{x}, t) = \int_0^t \mathbf{H}^*(\mathbf{x}, t') dt'$$

Nei problemi di magnetostatica e di elettrostatica le condizioni al contorno sono, talvolta, assegnate attraverso le componenti normali di \mathbf{B} e \mathbf{D}_T .

Ovviamente nei problemi statici \mathbf{A} e \mathbf{W} non sono univocamente determinati (si perde, infatti, il riferimento all'integrale temporale di \mathbf{E}). Pertanto, facendo riferimento al problema magnetostatico, pur potendosi operare in vari modi all'interno del dominio (ad esempio considerando la situazione stazionaria come lo stato di regime di un problema transitorio), sul contorno basterà

assegnare opportunamente \mathbf{A} tangente in modo che la sua circuitazione sia uguale al flusso di \mathbf{B}_n . [47]

Condizioni al contorno di simmetria

Simmetria di riflessione

Se all'interno del dominio V , i campi sono vincolati ad essere simmetrici rispetto ad un piano Σ (*piano di riflessione*), ci si può limitare a discretizzare soltanto una metà del dominio. Sul piano Σ vanno imposte le cosiddette condizioni di simmetria. Queste possono essere di due tipi:

- 1) $\mathbf{E}_t(\mathbf{x})=0$ \rightarrow $\mathbf{B}_n(\mathbf{x})=0$ $\forall \mathbf{x} \in \Sigma$
- 2) $\mathbf{H}_t(\mathbf{x})=0$ \rightarrow $\mathbf{D}_n(\mathbf{x})=0, \mathbf{J}_n(\mathbf{x})=0$ $\forall \mathbf{x} \in \Sigma$

Dalle condizioni di simmetria sui campi si ricavano le corrispondenti condizioni da imporre sui potenziali vettori:

1) $\mathbf{A}_t(\mathbf{x})=0$

2) $\mathbf{W}_t(\mathbf{x})=0$

Simmetria di rotazione

PERIODIC BOUNDARY

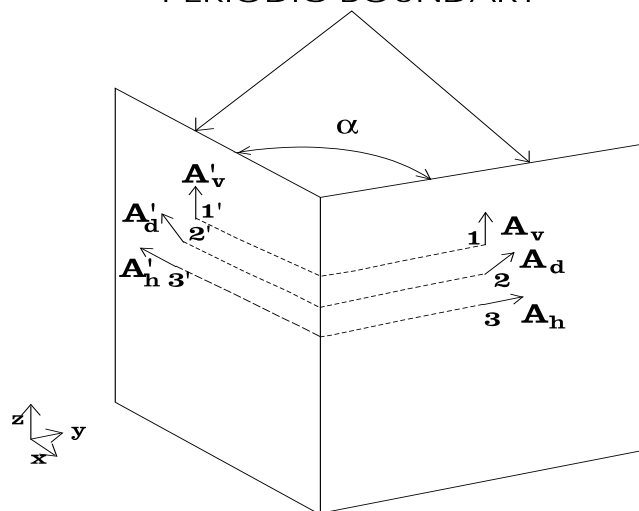


Fig.[14] Configurazione geometrica a simmetria rotazionale

Si supponga che il problema presenti simmetria rotazionale, o più in generale che la sua frontiera sia periodica di periodo α . In tal caso i potenziali vettori saranno ugualmente caratterizzati da questo tipo di simmetria. Si supponga di aver generato la mesh sfruttando la simmetria rotazionale, di modo che rotando uno dei piani di frontiera dell'angolo α , i suoi lati coincidano coi lati del piano rimanente (ad esempio, facendo riferimento alla fig.[14], si avrà che i lati 1,2,3 si sovrapporranno ai lati 1',2',3'). [48]

Considerate le coppie di lati corrispondenti che appartengono ai suddetti piani di simmetria, le condizioni al contorno vengono assegnate imponendo l'uguaglianza delle componenti tangente del potenziale vettore. In particolare, la soluzione deve soddisfare le seguenti condizioni:

-le *direzioni* dei gradi di libertà lungo lati simmetrici hanno simmetria rotazionale.

-i *moduli* dei gradi di libertà lungo lati simmetrici sono gli stessi:

$$|A_h|=|A'_h| \quad |A_v|=|A'_v| \quad |A_d|=|A'_d|$$

dove (A_h, A'_h) ; (A_v, A'_v) ; (A_d, A'_d) ; sono gli integrali di linea del potenziale vettore **A** lungo le coppie simmetriche di lati (1,1'); (2,2'); (3,3') riportate in fig.[14].

-i segni sono assegnati ad ogni coppia simmetrica in modo che, facendo sovrapporre attraverso rotazione i due piani di simmetria, i gradi di libertà associati al primo piano di frontiera puntino nelle stesse direzioni dei loro corrispondenti simmetrici appartenenti all'altra frontiera.

Un modo semplice per imporre operativamente le condizioni di simmetria rotazionale consiste nell'assegnare un unico numero d'ordine per ciascuna coppia di incognite simmetriche.

Imposizione delle condizioni al contorno attraverso il metodo delle Penalty Functions

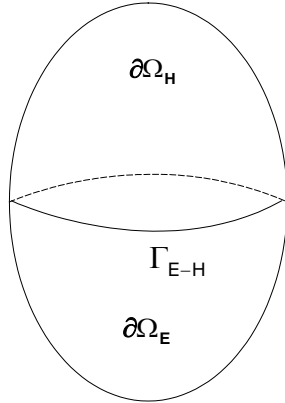


Fig.[16] Imposizione delle condizioni al contorno

Con il metodo delle Penalty functions il funzionale di errore si modifica nella forma seguente:

$$\Lambda_{Mod,k}(\mathbf{a}, \mathbf{w}) = \Lambda_k(\mathbf{a}, \mathbf{w}) + \frac{C_p}{2} \left(\sum_{r=1}^{N_A} (a_r - a_r^{(b)})^2 + \sum_{s=1}^{N_W} (w_s - w_s^{(b)})^2 \right)$$

dove per i lati appartenenti all'insieme $\partial\Omega_E$ e $\partial\Omega_H$ si ha semplicemente:

$$a_r^{(b)} = \int_{I_r} \mathbf{A}^{(b)}(\mathbf{x}, t_k + \Delta t_k) \cdot d\mathbf{x}; \quad w_r^{(b)} = \int_{I_r} \mathbf{W}^{(b)}(\mathbf{x}, t_k + \Delta t_k) \cdot d\mathbf{x};$$

cioè $a_r^{(b)}$ e $w_r^{(b)}$ sono rispettivamente gli integrali di linea dei potenziali vettori \mathbf{A} e \mathbf{W} lungo il lato r

In geometria tridimensionale per garantire l'unicità della soluzione è necessario imporre sulle superfici di frontiera uno solo dei campi -con riferimento alla fig.[16], quello elettrico sulla superficie $\partial\Omega_E$ oppure quello magnetico su $\partial\Omega_H$. Il vincolo imposto deve, naturalmente, essere soddisfatto da tutti i punti che appartengono alla superficie, compresi quelli che si trovano sulla curva di intersezione Γ_{E-H} tra le superfici $\partial\Omega_E$ e $\partial\Omega_H$. E' dunque chiaro che l'unicità della soluzione è garantita solo se si impone che tutti i lati appartenenti

all'intersezione di due diverse zone di frontiera rispettino i vincoli legati ad entrambe le superfici di frontiera.

Lo jacobiano e la matrice hessiana del funzionale di errore si modificano in maniera consistente con la nuova definizione del funzionale di errore.

Il maggiore vantaggio del metodo delle Penalty è rappresentato dalla facilità con la quale le condizioni al contorno vengono imposte. Operativamente bisogna, infatti, semplicemente aggiornare di volta in volta il termine noto, mentre i coefficienti della matrice non devono essere modificati.

Particolarmente critica nel metodo delle Penalty è la scelta del valore del coefficiente di penalità c_p . Questo deve essere abbastanza grande da dominare in maniera significativa su tutti i coefficienti dell'equazione che si intende vincolare. D'altra parte una scelta di c_p troppo elevata compromette il buon condizionamento della matrice e conduce a sistemi quasi-singolari: in questi casi il metodo fallisce oppure converge verso la soluzione solo dopo un elevatissimo numero di iterazioni di gradiente coniugato.

Imposizione delle condizioni al contorno attraverso il metodo di sostituzione

Nel metodo di sostituzione le condizioni di frontiera vengono imposte operando direttamente sulla matrice hessiana e sullo jacobiano del funzionale di errore. Ad esempio al k -mo passo di avanzamento nel tempo ed alla j -ma iterazione di Newton, il vincolo $a_r = a_r^{(b)}$ sul lato r viene imposto nella seguente maniera:

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \right)_r = a_r^{(j+1)} - a_r^{(b)} ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial a_r \partial a_r} = 1 \quad (6)$$

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \right)_s = \left(\frac{\partial \Lambda_k^{(j)}}{\partial \mathbf{a}} \right)_s + \frac{\partial^2 \Lambda_k^{(j)}}{\partial a_r \partial a_s} a_r^{(b)} ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial a_r \partial a_s} = 0 ; \quad s=1, \dots, Nedges ; s \neq r$$

$$\left(\frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \right)_s = \left(\frac{\partial \Lambda_k^{(j)}}{\partial \mathbf{w}} \right)_s + \frac{\partial^2 \Lambda_k^{(j)}}{\partial a_r \partial w_s} a_r^{(b)} ; \quad \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial a_r \partial w_s} = 0 \quad s=1, \dots, Nedges ;$$

A differenza del metodo di penalità, il buon condizionamento della matrice non risulta compromesso.

Condizioni al contorno di tipo miste:

In alcune classi di problemi, anzichè vincolare ad un particolare valore le componenti tangenti del campo elettrico oppure del campo magnetico, si impone sulla frontiera del dominio che questi campi, pur restando incogniti, siano legati l'un l'altro da una qualche relazione lineare. Un esempio è rappresentato dalle condizioni di assorbimento che vengono imposte in alcuni problemi elettromagnetici a carattere propagativo.

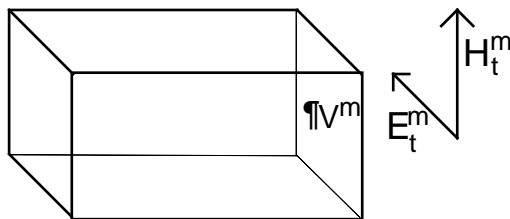


Fig [15] Imposizione delle condizioni di assorbimento

Con riferimento alla struttura di figura [15] la condizione da imporre sulla base sia la seguente $\forall \mathbf{x} \in \partial V^m$

$$a_{1m} \mathbf{E}_t^m + a_{2m} \mathbf{H}_t^m + a_{3m} = 0 \Rightarrow \mathbf{H}_t^m = -\frac{a_{3m}}{a_{2m}} - \frac{a_{1m}}{a_{2m}} \mathbf{E}_t^m; \quad (7)$$

dove i campi $(\mathbf{E}_t^m, \mathbf{H}_t^m)$ sono incognite del problema.

Le condizioni di tipo miste non possono essere imposte attraverso il metodo di sostituzione, pena la perdita di simmetria della matrice.

D'altra parte non è neanche conveniente adoperare le funzioni di penalità, in quanto la matrice risultante pur conservando la simmetria, sarebbe dominata da coefficienti fuori diagonale, con un sistema risultante molto mal

condizionato e scarse speranze di convergenza dello schema iterativo verso la soluzione.

La maniera più efficiente di affrontare il problema è quella di scegliere come variabili indipendenti tutte le incognite associate ai lati del dominio discretizzato, a^j e w^j , escluse quelle sul contorno ∂V^m , dove si assume come variabile solo una delle due incognite associate a ciascun lato. Ad esempio, scegliendo su ∂V^m come variabile indipendente la componente tangente del potenziale vettore \mathbf{A}_t , le seguenti espressioni di \mathbf{E}_t e \mathbf{H}_t verificano automaticamente la (7):

$$\mathbf{E}_t = -\frac{\partial \mathbf{A}_t}{\partial t} \quad (8a)$$

$$\mathbf{H}_t = -\frac{a_3}{a_2} + \frac{a_1}{a_2} \frac{\partial \mathbf{A}_t}{\partial t} \quad (8b)$$

il che si ottiene imponendo al contorno:

$$\mathbf{W}_t = \int_0^t \left(-\frac{a_3}{a_2} + \frac{a_1}{a_2} \frac{\partial \mathbf{A}_t}{\partial t} \right) dt \quad (9)$$

Di tali espressioni bisogna ovviamente tenere conto nel calcolo dello Jacobiano e dell'Hessiano, ma questo è solo un banale problema di programmazione.

Ad esempio se r è il lato su cui vanno imposte le condizioni di assorbimento, scegliendo come variabile indipendente a^r , la restrizione del funzionale di errore globale al k^{mo} time-step si trasforma nella maniera seguente:

$$\Lambda_k(\mathbf{a}, \mathbf{w}) = \Lambda_k(a^j, w^j, a^r, w^r) = \Lambda_k(a^j, w^j, a^r, w^r[a^r]) = \Lambda'_k(a^j, w^j, a^r)$$

$$\text{dove } w^r = \int_{l_r} \mathbf{W}_t \cdot d\mathbf{l} = \int_{l_r} \left[\int_0^t \left(-\frac{a_3}{a_2} + \frac{a_1}{a_2} \frac{\partial \mathbf{A}_t}{\partial t} \right) dt \right] \cdot d\mathbf{l} \quad (10)$$

Come appare chiaramente dalla (10) la relazione (7) viene soddisfatta solo in modo approssimato, nel senso che risulterà valida solo come integrale lungo gli spigoli delle facce di frontiera.

- Risoluzione numerica del sistema non lineare associato alla minimizzazione.

Il sistema algebrico finale può essere risolto per via numerica attraverso il metodo di Newton. Lo Jacobiano di $\Lambda_{Mod,k}$ viene approssimato in serie di Taylor:

$$\begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j+1)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j+1)}}{\partial \mathbf{w}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{w}} \\ \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{w}} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{a}^{(j+1)} \\ \Delta \mathbf{w}^{(j+1)} \end{bmatrix} = \mathbf{0}$$

dove $\Delta \mathbf{a}^{(j+1)} = \mathbf{a}^{(j+1)} - \mathbf{a}^{(j)}$, e $\mathbf{a}^{(j)}$ è il vettore, valutato alla j-ma iterazione, delle incognite associate al potenziale magnetico $\mathbf{A}(t_{k+1})$; simbologia analoga è utilizzata per la variabile \mathbf{w} .

Alla j-ma iterazione di Newton bisogna pertanto risolvere il seguente sistema lineare;

$$\begin{bmatrix} \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{w}} \\ \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{w}} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{a}^{(j+1)} \\ \Delta \mathbf{w}^{(j+1)} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \end{bmatrix} \quad (11)$$

Il sistema è sparso, simmetrico e definito positivo. I suoi coefficienti vengono immagazzinati attraverso il metodo di sparsità generale descritto nel capitolo precedente.

Il sistema viene prima preconditionato attraverso il metodo della fattorizzazione incompleta di Cholesky e, poi, risolto per via iterativa col metodo del gradiente coniugato. Per ridurre il rischio di patologie ed accelerare il processo di convergenza, si assume alla (j+1)^{ma} iterazione di Newton come soluzione il vettore:

$$\begin{bmatrix} \mathbf{a}^{(j+1)} \\ \mathbf{w}^{(j+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{a}^{(j)} \\ \mathbf{w}^{(j)} \end{bmatrix} + \beta \begin{bmatrix} \Delta \mathbf{a}^{(j+1)} \\ \Delta \mathbf{w}^{(j+1)} \end{bmatrix}$$

dove β è la soluzione del problema di minimizzazione monodimensionale del funzionale;

$$\min \Lambda_{\text{Mod},k}(\mathbf{a}^{(j)} + \beta \Delta \mathbf{a}^{(j+1)}, \mathbf{w}^{(j)} + \beta \Delta \mathbf{w}^{(j+1)}) ;$$

e viene ricavata attraverso il metodo delle secanti.

Attraverso una scelta opportuna dei coefficienti adimensionali α_H e α_E che compaiono nell'espressione del funzionale di errore -e precisamente, $\alpha_H = \alpha_E$ (si veda appendice B)-, i due campi incogniti $\mathbf{A}(t_{K+1})$ e $\mathbf{W}(t_{K+1})$ possono essere ottenuti come le soluzioni di due sistemi di equazione indipendenti. Vale la pena, tuttavia, di osservare che questo splitting non consente di separare i due problemi per \mathbf{A} e \mathbf{W} in tutto l'intervallo $(0, T)$ perchè nel nostro caso ogni potenziale a t_{K+1} dipende da entrambi i potenziali a t_K .

3.5) Rappresentazione grafica dei campi di errore.

Il metodo qui proposto provvede due soluzioni complementari che esattamente verificano le equazioni di Maxwell. Precisamente la soluzione in \mathbf{A} fornisce i campi \mathbf{E} e \mathbf{B} che verificano esattamente la legge di Faraday (2.2.1-7d), mentre la soluzione in termini di \mathbf{W} produce i campi \mathbf{D}_T e \mathbf{H} che soddisfano rigorosamente la legge di Ampere (2.2.1-7c). In questa maniera l'errore numerico risulta associato esclusivamente alle relazioni costitutive del mezzo (2.2.1-9a) e (2.2.1-9b).

Una visualizzazione della distribuzione della densità dell'errore locale può essere ottenuta, riportando la nuvola di punti

$$({}^0\mathbf{B}^0 = {}^0\mathbf{B}_0 + \nabla \times \mathbf{A}^0, {}^0\mathbf{H}^0 = {}^0\partial \mathbf{W} / \partial t^0)$$

in un piano B-H insieme con la curva rappresentante l'equazione costitutiva (2.2.1-9a).

Da un punto di vista pratico, il punto debole di questa rappresentazione è la mancanza dell'informazione correlata alla distribuzione spaziale della densità di errore; sebbene non esista una dimostrazione formale della diretta correlazione tra errore locale e raffinamento della mesh, l'esperienza pratica indica che una discretizzazione più raffinata localmente è generalmente utile per migliorare l'accuratezza locale. A questo scopo, due possibili rappresentazioni grafiche possono essere di supporto: un primo approccio può essere quello di disegnare nel dominio di interesse i vettori freccia che rappresentano i campi vettoriali $\mathbf{B}-\mathbf{B}(\mathbf{H})$ o alternativamente $\mathbf{D}_T-\mathbf{D}_T(\mathbf{E})$; la seconda possibilità è quella di disegnare la quantità scalare λ attraverso curve di livello in sezioni 2D oppure, in 3D, attraverso il modulo di vettori freccia che puntano lungo una direzione preassegnata.

E' opportuno rimarcare che gli errori vettoriali associati ai campi sono intimamente legati al funzionale di errore che si cerca di minimizzare per raggiungere la soluzione. Ad esempio nel caso lineare per il funzionale locale di errore si può adottare l'espressione (2.2.2-1) :

$$\lambda(\mathbf{x},t) = \alpha_H [\mathbf{B}-\mathbf{B}(\mathbf{H})]^2 + \alpha_E [\mathbf{D}_T-\mathbf{D}_T(\mathbf{E})]^2 \quad (1)$$

3.6) Implementazione numerica del metodo error-based per la rete RLC.

Posto :

$$A(t) = \frac{R}{(t-t_K) + RC}; \quad \forall t \in [t_K, t_{K+1}]; \quad \Delta t_m = t_{m+1} - t_m ;$$

$$Cost = q_0 - q_{C0} + \int_0^{t_K} \frac{V - \dot{\Phi}}{R} dt = q_0 - q_{C0} + \sum_{m=0}^{K-1} \left[V(t_m^+) - \dot{\Phi}(t_m^+) \right] \Delta t_m$$

Usando il time-stepping, e tenendo conto delle espressioni (2.2.4-6) e (2.2.4-7), il funzionale di errore (2.2.4-9) della rete RLC esaminata nel capitolo 2, può essere posto in forma esplicita come.

$$\Lambda_k(\Phi_{k+1}, q_{k+1}, \alpha) = \frac{1}{2} \sum_{j=1}^{N_G} w_j \left\{ \alpha_L \left[\frac{\Phi^2}{L} + L(\dot{q})^2 - 2\Phi \dot{q} \right] + \alpha_{RC} \left[\frac{(V - \dot{\Phi})^2}{A(t_j)} + A(t_j)(q - Cost)^2 - 2(V - \dot{\Phi})(q - Cost) \right] \right\}$$

Dove per l'integrale nell'intervallo di tempo si possono usare le formule di integrazione gaussiana oppure uno dei metodi monostep descritti nel paragrafo (3.1.2). La minimizzazione del funzionale di errore si ottiene uguagliando a zero la derivata di Λ_k rispetto alle incognite Φ_{k+1} e q_{k+1} :

$$\begin{bmatrix} \frac{\partial \Lambda_k}{\partial q_{k+1}} \\ \frac{\partial \Lambda_k}{\partial \Phi_{k+1}} \end{bmatrix} = \mathbf{0}$$

che conduce al sistema:

$$\begin{bmatrix} \sum_{j=1}^{N_G} w_G \left[\alpha_L \frac{L}{\Delta t_K} + \alpha_{RC} \left(\frac{A(t_j)(t_j - t_K)^2}{\Delta t_K} \right) \right] & \sum_{j=1}^{N_G} w_G \left[(\alpha_L - \alpha_{RC}) \left(\frac{t_j - t_K}{\Delta t_K} \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[(\alpha_L - \alpha_{RC}) \left(\frac{t_j - t_K}{\Delta t_K} \right) \right] & \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \frac{1}{A(t_j)\Delta t_K} + \alpha_L \left(\frac{(t_j - t_K)^2}{L\Delta t_K} \right) \right] \end{bmatrix} \begin{bmatrix} q_{k+1} \\ \Phi_{k+1} \end{bmatrix} =$$

$$\begin{aligned}
& \left[\begin{array}{cc} \sum_{j=1}^{N_G} w_G \left[\alpha_L \frac{L}{\Delta t_K} + \alpha_{RC} (t_j - t_K) A(t_j) \left(\frac{t_j - t_K}{\Delta t_K} - 1 \right) \right] & \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \frac{t_j - t_K}{\Delta t_K} + \alpha_L \left(1 - \frac{t_j - t_K}{\Delta t_K} \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left(\frac{t_j - t_K}{\Delta t_K} - 1 \right) - \alpha_L \frac{t_j - t_K}{\Delta t_K} \right] & \sum_{j=1}^{N_G} w_G \left[\frac{\alpha_{RC}}{A(t_j) \Delta t_K} + \alpha_L \frac{t_j - t_K}{L} \left(\frac{t_j - t_K}{L \Delta t_K} - 1 \right) \right] \end{array} \right] \begin{bmatrix} q_k \\ \Phi_k \end{bmatrix} \\
& + \left[\begin{array}{c} \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left((kA(t_j) + V(t_j))(t_j - t_K) \right) \right] \\ \sum_{j=1}^{N_G} w_G \left[\alpha_{RC} \left(k \frac{V(t_j)}{A(t_j)} \right) \right] \end{array} \right]
\end{aligned}$$

Si noti che la scelta $\alpha_L = \alpha_{RC}$ produce il disaccoppiamento delle equazioni. Ciò, ovviamente, non significa che i sistemi in q e Φ possono essere avanzati nel tempo indipendentemente. Difatti, in ogni caso, il termine noto di ciascuna equazione dipende dai valori assunti nel time-step precedente da entrambe le variabili. L'andamento temporale della corrente attraverso l'induttore e della tensione ai capi del condensatore può essere ricavato in termini di q e Φ come:

$$i_q(t_{k+1}) = \frac{q_{k+1} - q_k}{\Delta t_k}; \quad i_\Phi(t_{k+1}) = \frac{\Phi_{k+1}}{L}$$

$$v_q(t_{k+1}) = (q_{k+1} - Cost_k) \frac{R}{(RC + \Delta t_k)}; \quad v_\Phi(t_k) = V(t_k) - \frac{\Phi_{k+1} - \Phi_k}{\Delta t_k}$$

$$\text{dove: } Cost_{k+1} = Cost_k + \frac{[V(t_k) \Delta t_k - (\Phi_{k+1} - \Phi_k)]}{R}$$

mentre $V(t_k)$ è il valore all'istante t_k della tensione impressa dal generatore indipendente.

CAPITOLO IV

Casi test e risultati

Introduzione

Una delle caratteristiche più interessanti del metodo proposto è il trattamento unificato di problemi stazionari, quasi-stazionari e propagativi. La corrente di spostamento, infatti, è tenuta in conto in maniera automatica, senza alcuna complicazione addizionale, tanto nel caso in cui sia importante che in quello in cui sia numericamente trascurabile. Inoltre, i problemi stazionari possono essere trattati come la situazione di regime di un transitorio fittizio.

Si è già visto come i campi possano essere espressi attraverso due espressioni differenti ed indipendenti. Ad esempio, il campo magnetico di induzione \mathbf{B} si può scrivere rispettivamente in funzione di \mathbf{A} e \mathbf{W} come:

$$\mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A}$$
$$\mathbf{B} = \mathbf{f}_B(\partial \mathbf{W} / \partial t)$$

Inoltre, come visto, anche per grandezze integrali, quali ad esempio l'energia magnetica o la potenza ohmica, sono disponibili varie espressioni. Appare, tuttavia, opportuno ricordare che queste espressioni stabiliscono limiti superiori ed inferiori per alcune grandezze integrali, quali l'energia elettrostatica e magnetostatica (problemi stazionari), e non per altre come la potenza ohmica o l'energia elettrica e magnetica in problemi non-stazionari. Per questo motivo, il metodo proposto non può essere considerato, a rigore, una vera formulazione CVP (vedi par.1.3).

Il primo paragrafo sarà dedicato alla discussione dei risultati numerici ottenuti applicando la formulazione basata sulla minimizzazione dell'errore costitutivo alla particolare rete RLC introdotta nelle pagine precedenti. Sarà evidenziato, tra l'altro, come il raffinamento della discretizzazione temporale induca un miglioramento della convergenza nell'approssimazione della soluzione e nella minimizzazione dell'errore sul funzionale col raffinamento temporale di ordine quadratico.

Nei successivi paragrafi saranno esaminati tre casi test che si riferiscono, rispettivamente, a problemi di magnetostatica, correnti parassite e propagazione. Le soluzioni ottenute con l'approccio basato sulla minimizzazione dell'errore costitutivo saranno confrontate con quelle riportate in letteratura.

4.1) Circuiti a parametri concentrati: Rete RLC lineare.

Come si è detto nel paragrafo [2.4], l'approccio basato sulla minimizzazione dell'errore costitutivo può fallire se le scelte dei parametri operativi (pesi, ampiezza del time-step, schema di integrazione) non vengono effettuate in maniera adeguata. Diventa allora importante poter disporre di informazioni circa il comportamento qualitativo delle soluzioni al variare dei parametri in gioco. Queste possono essere ottenute attraverso un'analisi parametrica di un semplice modello zero-dimensionale del problema in esame.

A scopo di esempio, con riferimento al semplice problema di campo descritto nel par. [2.4], sono mostrati i risultati ottenuti per il corrispondente modello circuitale costituito da una rete RLC forzata attraverso un generatore indipendente di tensione. si è assunto che l'induttore e il condensatore partano da condizioni iniziali non nulle ($I_{l0}=4A, V_{c0}=-3V$), e che il generatore imponga una tensione costante pari a 1 V. I valori della resistenza, della capacità e dell'induttanza sono stati fissati rispettivamente pari a $1\Omega, 1e-4 F$ e $1H$. La durata dell'intervallo di studio è stata fissata uguale a $T=6s$, mentre per i pesi, in una prima serie di simulazioni, si è posto $\alpha_L = \alpha_{RC}=1$.

Gli errori puntuali nelle due formulazioni per la corrente nell'induttore, $\varepsilon_{L,q}$ e $\varepsilon_{L,\phi}$, e per la tensione ai capi del condensatore, $\varepsilon_{RC,q}$ e $\varepsilon_{RC,\phi}$, sono definiti rispettivamente come:

$$\varepsilon_{L,q}^{(p)}(t) = \frac{i_q(t) - i_{an.}(t)}{i_{an.}(t)}; \quad \varepsilon_{L,\phi}^{(p)}(t) = \frac{i_\phi(t) - i_{an.}(t)}{i_{an.}(t)}$$

$$\varepsilon_{RC,q}^{(p)}(t) = \frac{v_q(t) - v_{an.}(t)}{v_{an.}(t)}; \quad \varepsilon_{RC,\phi}^{(p)}(t) = \frac{v_\phi(t) - v_{an.}(t)}{v_{an.}(t)}$$

Mentre gli errori integrali sono stati posti uguali a:

$$E_{L,q}^{(g)} = \sqrt{\frac{\int_0^T (i_q(t) - i_{an.}(t))^2 dt}{\int_0^T (i_{an.}(t))^2 dt}}; \quad E_{L,\phi}^{(g)} = \sqrt{\frac{\int_0^T (i_\phi(t) - i_{an.}(t))^2 dt}{\int_0^T (i_{an.}(t))^2 dt}}$$

$$E_{RC,q}^{(g)} = \sqrt{\frac{\int_0^T (v_q(t) - v_{an.}(t))^2 dt}{\int_0^T (v_{an.}(t))^2 dt}}; \quad E_{RC,\phi}^{(g)} = \sqrt{\frac{\int_0^T (v_\phi(t) - v_{an.}(t))^2 dt}{\int_0^T (v_{an.}(t))^2 dt}}$$

Il periodo di studio $[0,T]$ viene partizionato in N intervalli. Ciascun integrale che compare nelle espressioni sopra riportate viene calcolato come la somma degli N integrali valutati, su ciascuno degli intervalli in cui il periodo di studio è stato diviso, attraverso formule numeriche di

integrazione del primo ordine (schema implicito, schema di Crank-Nicolson), oppure di ordine superiore (integrali Gaussiani).

Andamento dell'errore globale su I

Andamento dell'errore globale su V_c

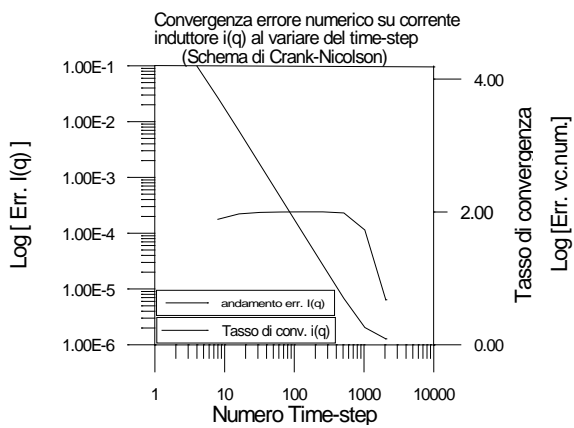


Fig.[1.A]

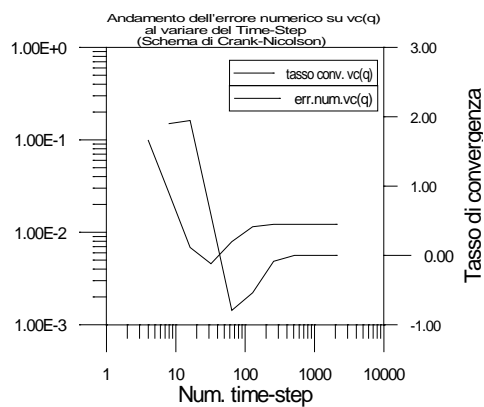


Fig.[2.A]

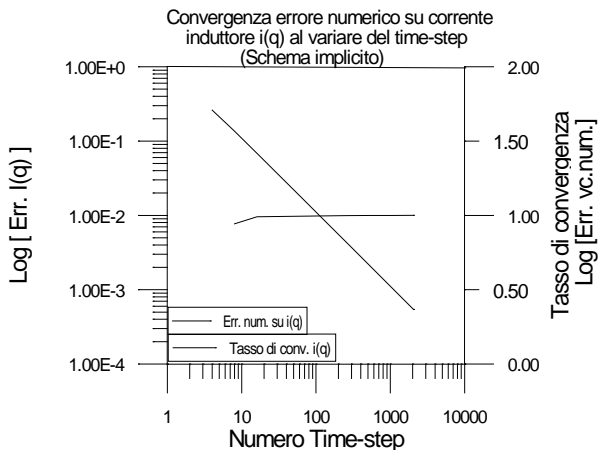


Fig.[1.B]

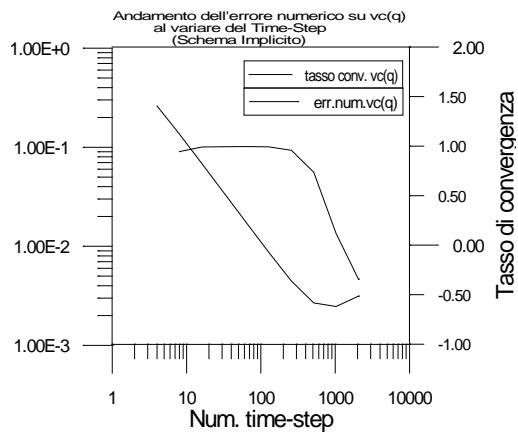


Fig.[2.B]

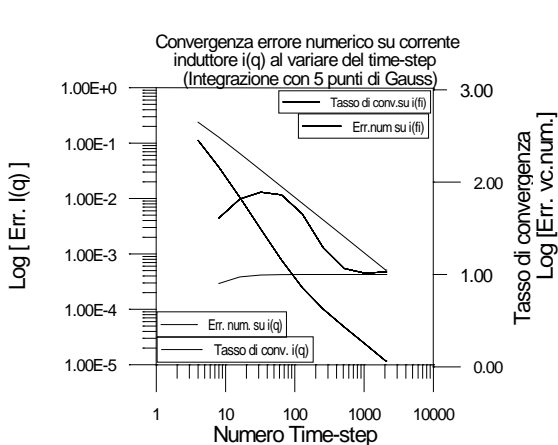


Fig.[1.C]

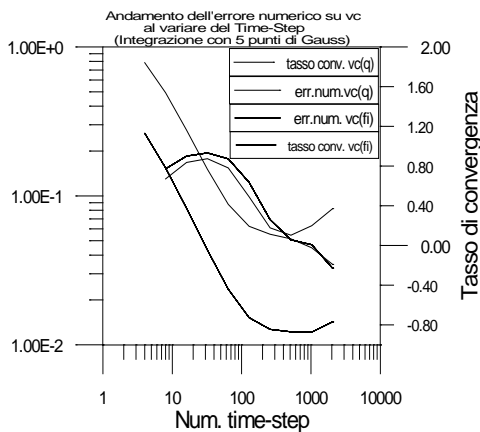


Fig.[2.C]

Andamento dell'errore puntuale su I
Time=3 s

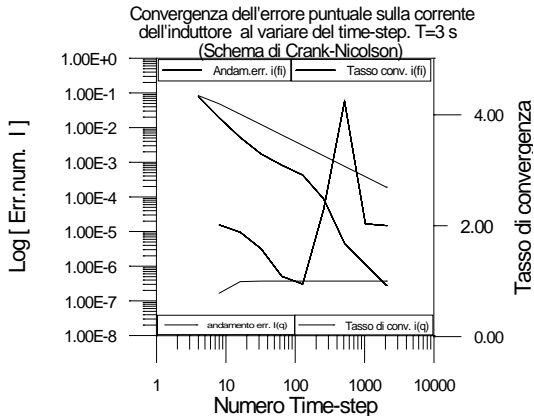


Fig.[3.A]

Andamento dell'errore puntuale su V_c
Time=3 s

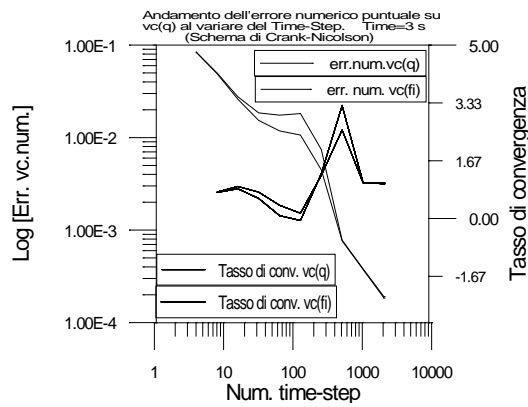


Fig.[4.A]

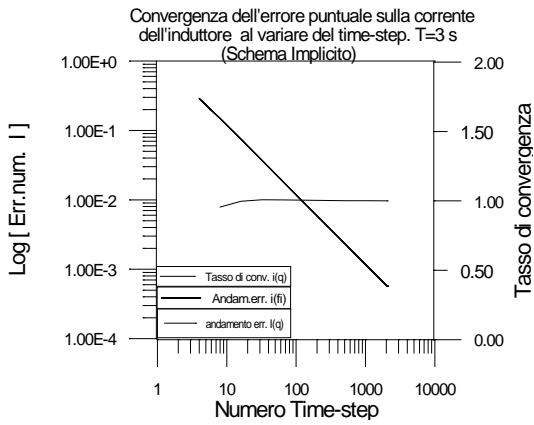


Fig.[3.B]

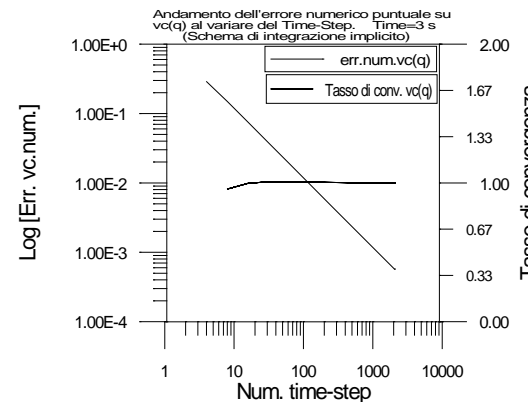


Fig.[4.B]

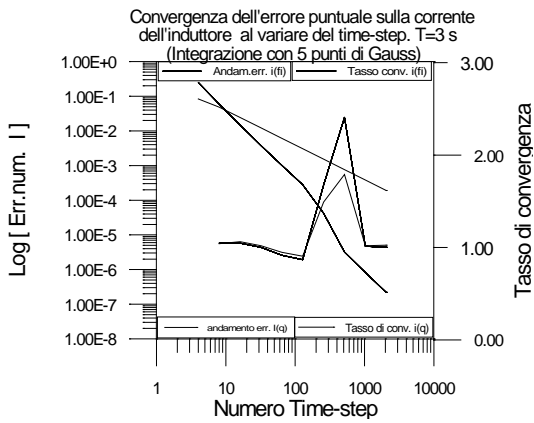


Fig.[3.C]

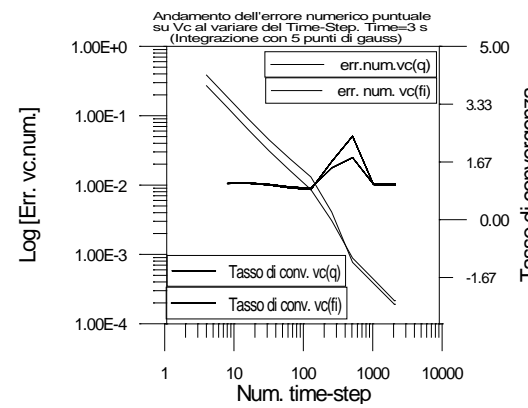


Fig.[4.C]

Andamento del funzionale di errore

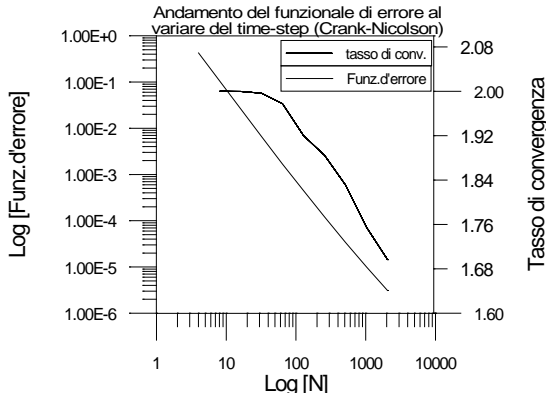


Fig.[5.A]

Evoluzione temporale delle incognite f_i e q e soluzioni numeriche

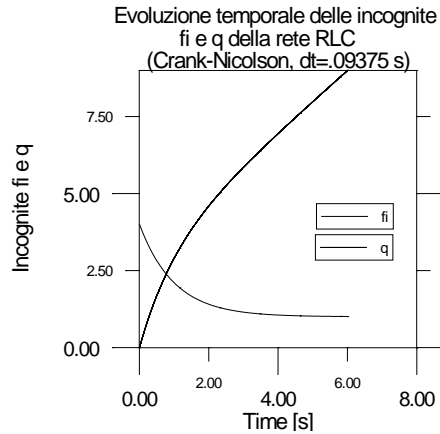


Fig.[6.A]

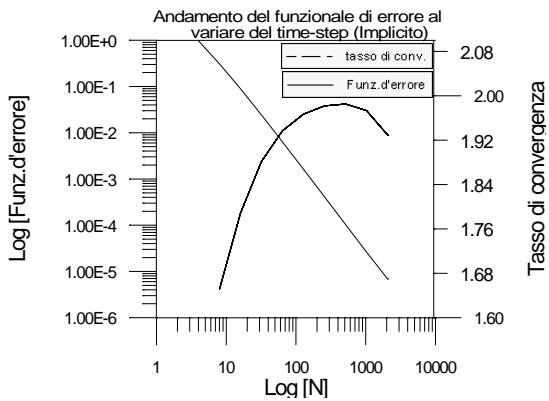


Fig.[5.B]

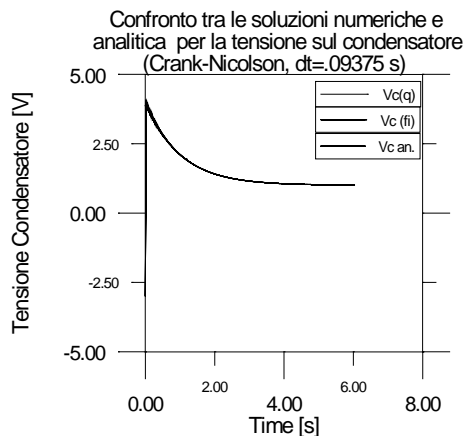


Fig.[6.B]

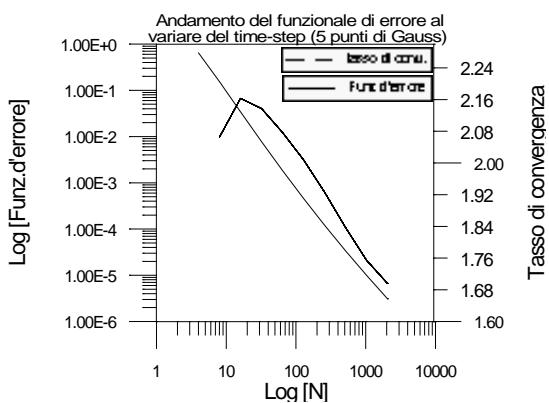


Fig.[5.C]

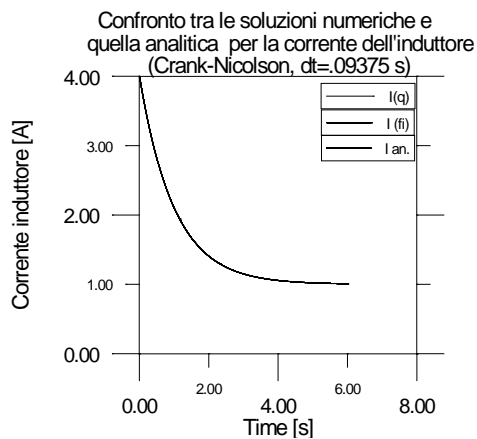


Fig.[6.C]

Comportamento soluzioni numeriche al variare della capacità C ($\alpha_{rc}=\alpha_1 dt=.047$)

Comportamento soluzioni numeriche variare della al variare del peso α_1 ($\alpha_{rc}=1, dt=.047$)

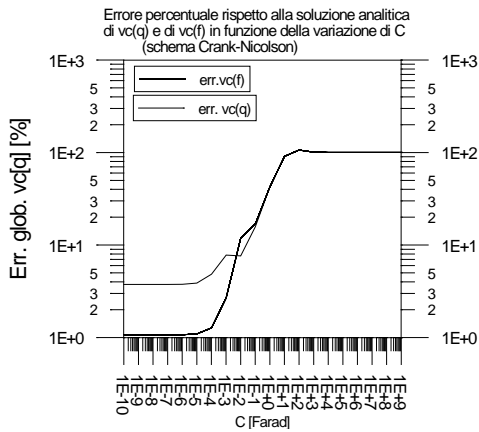


Fig.[7.A]

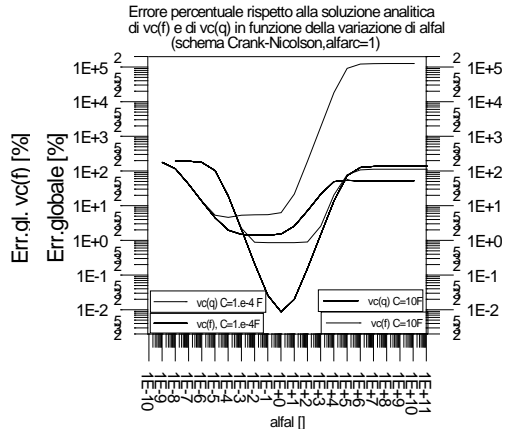


Fig.[8.A]

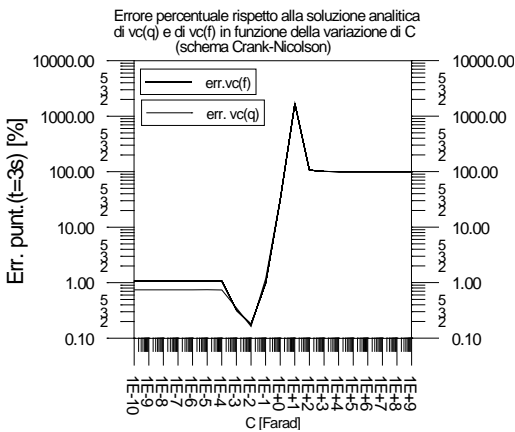


Fig.[7.B]

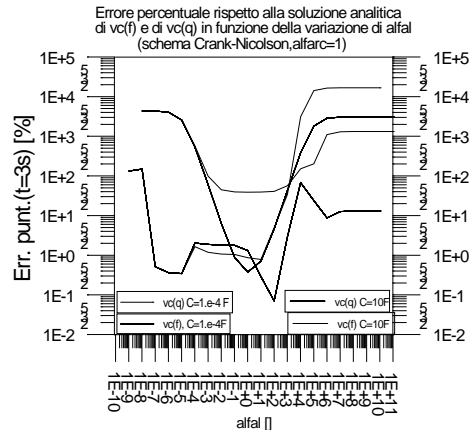


Fig.[8.B]

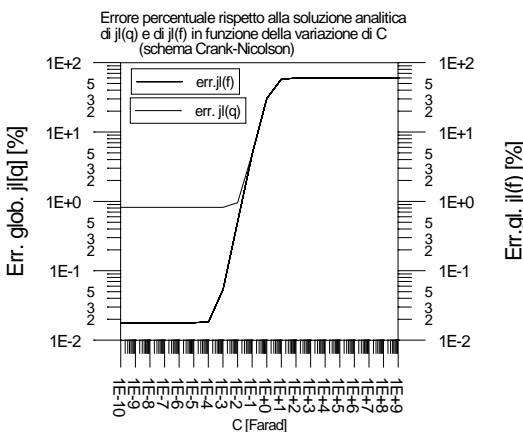


Fig.[7.C]

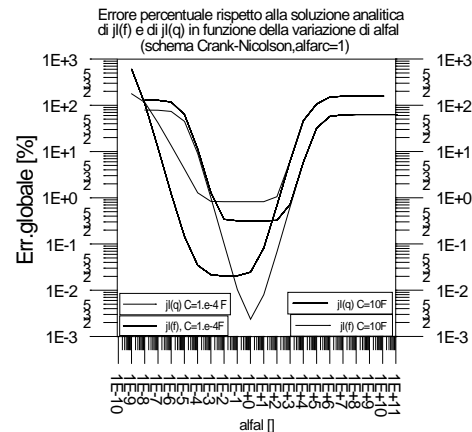


Fig.[8.C]

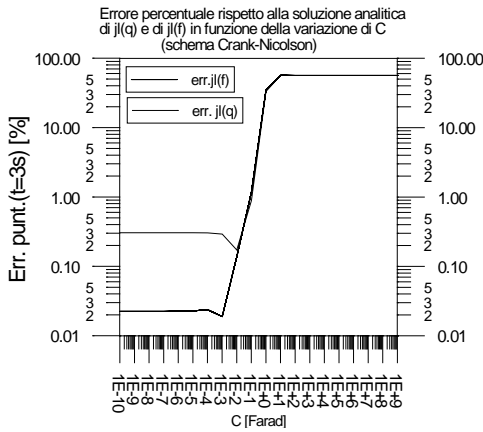


Fig.[7.D]

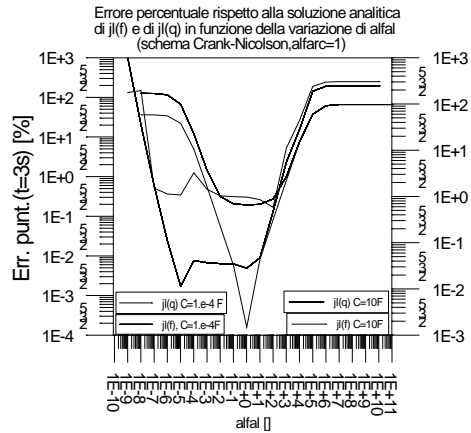


Fig.[8.D]

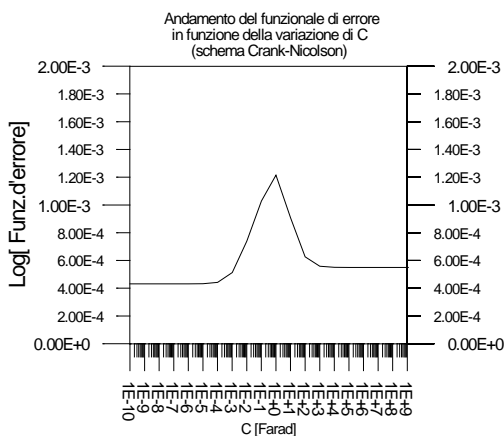


Fig.[7.E]

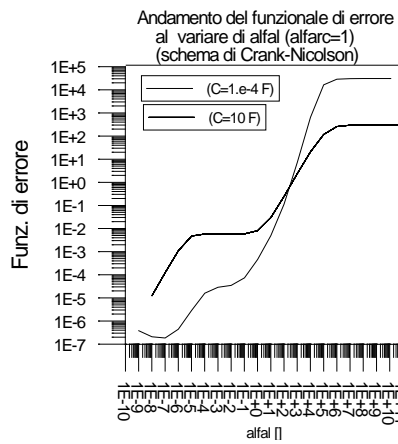


Fig.[8.E]

-Nelle figg.[1-2a,b,c] viene mostrato per le due formulazioni in q e φ quale sia, al variare del time-step, l'andamento degli errori relativi integrali commessi nella valutazione della corrente che attraversa l'induttore e della tensione ai capi del condensatore e i rispettivi tassi di convergenza. Sono stati usati differenti schemi di integrazione [a=Crank-Nicolson, b=implicito, c=Gaussiano (5 punti)].

Si noti che l'errore commesso nel valutare la tensione sul condensatore è molto più elevato rispetto a quello associato alla soluzione numerica per la corrente che attraversa l'induttore, ed inoltre non diminuisce uniformemente col diminuire del time-step. Questa circostanza indica che per il circuito analizzato la scelta dei pesi $\alpha_L = \alpha_{RC}$ penalizza la relazione costitutiva sul condensatore a vantaggio dell'altra. Per migliorare contemporaneamente la qualità di entrambe le soluzioni potrebbe essere opportuno scegliere i valori dei pesi in modo che $\alpha_{RC} > \alpha_L$.

Per gli schemi di integrazione mono-step in ogni figura appare una sola curva di errore essendo l'altra perfettamente coincidente: la cosa non deve stupire in quanto si è adoperato lo stesso schema di integrazione sia per il calcolo dei coefficienti della matrice da invertire sia per la valutazione degli errori integrali.

Si noti, infatti, che cercare i valori q_{k+1} e j_{k+1} che minimizzano il valore del funzionale di errore costitutivo all'interno dell'intervallo $[t_k, t_{k+1}]$ è equivalente a cercare i valori q_{k+1} e j_{k+1} che minimizzano la distanza integrale tra le coppie di soluzioni numeriche ottenute per $i_l(t)$ e $v_c(t)$ attraverso le incognite duali q e φ . In termini matematici il problema si esprime come:

$$\min_{q_{k+1}, \varphi_{k+1}} \left(\int_{t_k}^{t_{k+1}} \left[\alpha_l (i_l[q(t)] - i_l[\varphi(t)])^2 + \alpha_{RC} (v_c[q(t)] - v_c[\varphi(t)])^2 \right] dt \right)$$

Se si adotta un metodo mono-step il problema si riduce a cercare i valori q_{k+1} e j_{k+1} che minimizzano la distanza tra le soluzioni numeriche ottenute per i_l e v_c attraverso le incognite duali q e φ ad un particolare istante t^* $[t_k, t_{k+1}]$, dove t^* dipende dallo schema di integrazione mono-step utilizzato (ad es. $t^*=t_{k+1}$ se si adotta lo schema implicito, $t^*=(t_k+t_{k+1})/2$ se si usa Crank-Nicolson):

$$\min_{q_{k+1}, \varphi_{k+1}} \left(\left[\alpha_l (i_l[q(t^*)] - i_l[\varphi(t^*)])^2 + \alpha_{RC} (v_c[q(t^*)] - v_c[\varphi(t^*)])^2 \right] \right)$$

e questo equivale a richiedere:

$$\begin{cases} i_l[q(t^*)] = i_l[\varphi(t^*)] \\ v_c[q(t^*)] = v_c[\varphi(t^*)] \end{cases}$$

Le figg. [3-4a,b,c] mostrano le curve degli errori puntuali su $i_l(t)$ e $v_c(t)$ (calcolati all'istante $t=3s$) e i tassi di convergenza degli errori puntuali.

Le figg [5a,b,c] riportano l'andamento del funzionale di errore e il tasso di convergenza associato per ciascun schema di integrazione.

Dall'esame delle curve si evidenziano alcuni punti:

- Quando si utilizza lo schema di Crank-Nicolson, il tasso di convergenza dell'errore sulla corrente dell'induttore ha andamento quadratico per la soluzione $i(\varphi)$ e solo lineare per la soluzione $i(q)$.
- Quando si utilizza lo schema implicito, i tassi di convergenza degli errori numerici sulla corrente dell'induttore sono in ogni caso lineari.

La fig.[6.A] mostra l'andamento temporale delle incognite flusso e carica; mentre nelle figg.[6.B-C] vengono comparate soluzioni numeriche e analitiche per $i_l(t)$ e $v_c(t)$ (si noti come in entrambi i casi le curve numeriche risultano sovrapposte a quella analitica). Le soluzioni numeriche sono state ottenute attraverso lo schema di integrazione di Crank-Nicolson, e con un'ampiezza del time-step pari a $dt=.09375$ s.

Le figg.[7a-e] mostrano i risultati ottenuti bloccando l'ampiezza del time-step a $dt=.0078125$ s e variando la capacità. Si vede che il modello va in crisi non appena la costante di tempo più rapida del circuito diventa paragonabile con l'ampiezza del time-step: $dt/10 < \tau_1 \cong 1e-3$, cioè per $C > 1e-3F$. Si noti, inoltre, che, per $C > .25F$, i modi caratteristici del circuito inducono oscillazioni nelle soluzioni. In questo caso, perchè il metodo numerico risulti efficace è necessario che l'ampiezza del time-step soddisfi la relazione $dt/10 < 1/f$, dove f è la frequenza delle oscillazioni.

Mantenendo $\alpha_{RC}=1$ e assumendo $dt=.0078125s$, nelle figg. [8a-e] vien quindi mostrata l'influenza sulla soluzione della scelta del peso α_L per due valori della capacità: $C=1.e-4F$ e $C=10F$. Coerentemente con quanto era lecito aspettarsi, la qualità della soluzione ottenuta dalla computazione dipende pesantemente dalla scelta di questo parametro.

Le informazioni ricavate sul comportamento qualitativo della soluzione al variare dei parametri possono essere vantaggiosamente utilizzate nell'analisi 3D del problema elettromagnetico che il circuito esaminato modellava alquanto rozzamente.

4.2) Magnetostatica: Cubo attraversato da corrente uniforme

Come esempio di applicazione a problemi di magnetostatica, si è calcolato il campo magnetico bidimensionale prodotto da una corrente di densità uniforme parallela ad uno dei lati di un cubo di permeabilità uniforme. La lunghezza del lato sia $l=1$ m, mentre la permeabilità magnetica quella del vuoto; la densità di corrente sia fissata pari a $J=10^7$ A/m².

I risultati sono stati ottenuti usando due mesh: una più rozza costituita da due soli elementi per lato, e l'altra più raffinata con otto elementi per lato. Si sono sfruttate le simmetrie di riflessione e quindi si è discretizzato solo 1/8 del cubo. I risultati numerici in termini di energia magnetica sono riportati in tavola 1 insieme con quelli ottenuti in [49] dove si sono adoperate le formulazioni in **A** e in **T-Ω**.

Tavola 1
Calcolo del campo magnetostatico in un cubo
($l=1$ m, $J_x=J_y=0, J_z=10$ MA/m²)

<i>Mesh</i>		
elementi per lato	2	8
nodi	27	729
elementi	8	512
lati	54	1944
<i>Gradi di libertà</i>		
formulazione in A [49]	5	833
formulazione in T-Ω [49]	54	1944
formulazione in A-W	108	3888
<i>Occupazione di memoria</i>		
form. in A [49] (matrice bandata)	14	77914
form. in T-Ω [49](matrice bandata)	257	119718
form. in A-W (matrice sparsa)	510	27672
<i>energia magnetica (MJ)</i>		
Valore analitico	2.2082	2.2082
form.in A [49]	1.4726	2.1572
form.in T-Ω [49]	2.6180	2.2380
form.in A-W (int. volume $\mathbf{B}^2/2\mu$)	1.4726	2.1572
form.in A-W (int. volume $\mu\mathbf{H}^2/2$)	2.6180	2.2381
form.in A-W (int. volume $\mathbf{B}\cdot\mathbf{H}/2$)	1.4726	2.1572

4.3) Correnti parassite: il Felix-brick.

Come esempio di applicazione per il problema delle correnti parassite è stato scelto il cosiddetto Felix Brick, uno dei problemi usati come test nell'ambito dei TEAM Workshops, organizzati da un gruppo internazionale di ricercatori al fine di verificare la validità delle varie formulazioni numeriche e dei codici di calcolo associati [50].

Il dominio del problema è costituito da un conduttore massiccio di forma rettangolare con un foro, ugualmente rettangolare, che lo attraversa completamente. Un campo magnetico uniforme, che decade esponenzialmente col tempo, è applicato perpendicolarmente alle facce forate. Sono disponibili valori numerici ottenuti mediante altre procedure di calcolo, tra cui la potenza dissipata per effetto ohmico e il campo magnetico in varie posizioni.

Il materiale di cui è composto il conduttore è una lega di alluminio con valore di resistività $\rho=3.94 \mu\Omega\text{cm}$. Le dimensioni del mattone sono $0.1524\text{m}\times 0.1016\text{m}\times 0.0508\text{m}$. Il foro rettangolare che attraversa completamente il mattone dalla parte della facce più larghe ha dimensioni $0.0889\text{m}\times 0.0381\text{m}$. Il campo magnetico esterno ha la sola componente nella direzione z ed è uniforme nello spazio. Per $t>0$, decade in maniera esponenziale nel tempo con la seguente legge:

$$B_z = B_0 e^{-t/\tau} \quad (1)$$

dove $B_0=0.1 \text{ T}$ e $\tau=0.0119 \text{ s}$. Questa condizione può essere imposta attraverso la scelta di opportune condizioni al contorno.

Per la simulazione numerica è stato modellato il conduttore insieme con una zona di spazio vuoto ad esso circostante. Sfruttando le simmetrie del problema è stato possibile modellare solo 1/8 dell'intero dominio.

Sono state adoperate in una prima fase due discretizzazioni alquanto rozze: la prima costituita da 18 elementi esaedrali e 208 incognite, mentre la seconda formata da 48 elementi esaedrali e 470 incognite.

In fig.[8] per ciascuna discretizzazione sono rappresentate le distribuzioni delle seguenti grandezze vettoriali all'istante $t=10.5$ ms -momento in cui viene raggiunto il valore di picco della potenza dissipata per effetto Joule-:

- il campo magnetico \mathbf{B} calcolato attraverso il potenziale vettore magnetico \mathbf{A} (fig.[8a,1-2]);

- il campo magnetico $\mathbf{B(H)}$ calcolato attraverso il potenziale vettore elettrico \mathbf{W} (fig.[8b,1-2]);

- la differenza $\mathbf{B-B(H)}$ nella stessa scala usata per rappresentare i campi \mathbf{B} e $\mathbf{B(H)}$ (fig.[8c,1-2]);

- la differenza $\mathbf{B-B(H)}$ amplificata attraverso un fattore di scala pari a 25 (fig.[8d1,2]);

- e' stata, infine, riportata la distribuzione del funzionale di errore. Pur essendo quest'ultima una quantità scalare, la si è rappresentata attraverso un campo vettoriale con la sola componente nella direzione z diversa da zero(fig.[8e1,2]).

Dalle figure [8a-b] si rileva il comportamento qualitativo del campo magnetico, mentre le figure [8d-e] evidenziano che all'interno del conduttore l'errore costitutivo è massimo ed è pertanto necessario raffinare ulteriormente la discretizzazione.

Come discretizzazione finale è stata adoperata la mesh consigliata nella referenza[50] e mostrata con dovizia di particolari in figura [9]. Essa è composta da 1344 nodi, 1001 elementi esaedrali, 3656 lati e, pertanto, 7312 incognite per il metodo $\mathbf{A-W}$. La mesh, congruentemente con le indicazioni rilevate dalle simulazioni precedenti, è più fitta nella zona spaziale occupata dal conduttore e perde gradatamente in raffinatezza man mano che ce ne si allontana. I valori computati per la potenza ohmica e per il campo indotto nel centro del buco ($x=y=z=0$) sono riportati in tabella 2, e sono in buon accordo con i risultati ottenuti attraverso le altre formulazioni (vedi Tabella 3).

E' interessante riportare il valore incrementale di Λ in un time-step $\Delta t=0.5$ ms (1.66 mJs a $t=10.5$ ms, quando la potenza ohmica è massima). che è

sufficientemente piccolo quando comparato al valore della figura di riferimento (2) che è pari a 373.91 mJs:

$$\Lambda_{\text{ref}} = \alpha_H \Delta t \int_V \mathbf{B} \cdot \mathbf{H} / 2 \, dV + \alpha_E \Delta t \int_V (\mathbf{E} \cdot \mathbf{D} / 2 + \mathbf{E} \cdot \mathbf{J} \Delta t) \, dV \quad (2)$$

Tavola 2

Un problema di eddy currents: il Felix brick

($t=10.5 \text{ ms}, \Delta t=0.5 \text{ ms}$)

Potenza Ohmica

form.in A-W (int. volume \mathbf{J}^2/s)	106.98 W
form.in A-W (int. volume $s\mathbf{E}^2$)	106.03 W
form.in A-W (int. volume $\mathbf{E} \cdot \mathbf{J}$)	105.53 W
<i>Campo magnetico indotto in ($x=y=z=0$)</i>	
form.in A-W (\mathbf{B}_z)	0.04108 T
form.in A-W (n \mathbf{H}_z)	0.03907 T

Tavola 3

Risultati ottenuti con altre formulazioni ($t=10.5 \text{ ms}$) [50]

Codice	Form.	Time-step method	Variabili		Num. Elem.	Funzione di Forma	Numero Incognite	Bz indotto $z=(0,0,0)$	Potenza Ohmica
			Cond.	Vuoto					
Eddy3DT A-V-ψ	FEM	Crank-Nic	A-V	Ω	1001	8 nodi Esaedrale	1641	0.0388 T	110.2 W
Eddy3DT A*-ψ	FEM	Crank-Nic	A*	Ω	1001	8 nodi Esaedrale	1577	0.0391 T	110.0 W
Field/ T3DT	FEM	Implicito	T-Ω	Ω	6006	4 nodi Tetraedrale	1260	0.0413 T	87.1 W
Field/ A3DT	FEM	Implicito	A-V	A	6006	4 nodi Tetraedrale	3212	0.0365 T	127.9 W
CARIDDI	FEIM	Crank-Nic.	T		144	12 lati Esaedrale	216	0.0390 T	113.6 W
EDDY CUFF	FECM	Espansione automodi	T_n		260	3 nodi Triangolare	284	0.0386 T	113.7 W
MEGA	FEM	Crank-Nic	A-V	Ω	1404	8 nodi Esaedrale	2289	0.0392 T	111.0 W

Nelle figg.[10a-e] con riferimento alla sola zona spaziale occupata dal mattone e dal foro che l'attraversa, sono rappresentati in ordine:

- L'andamento del campo **B** calcolato attraverso il potenziale vettore magnetico **A** ed il potenziale vettore elettrico **W** (figg.[10a-b]).
- La differenza tra le due soluzioni ottenute per il campo **B** (fig.[10c]).

-La distribuzione della densità di corrente \mathbf{J} indotta nel conduttore massiccio (fig.[10d]).

-La mesh adoperata per la discretizzazione del solo conduttore massiccio (fig.[10e]).

La congruenza tra i risultati riportati in letteratura e quelli ottenuti attraverso la formulazione proposta mostra come questa non degeneri nelle situazioni in cui la corrente di spostamento risulta trascurabile rispetto a quella di conduzione.

4.4) Propagazione.

E' stata messa, infine, alla prova l'efficacia della formulazione proposta attraverso alcuni problemi di carattere propagativo. I mezzi attraverso i quali la propagazione dei campi elettromagnetici aveva luogo sono stati supposti lineari in una prima serie di esercizi. Si è quindi studiato un problema di propagazione di un'onda piana attraverso un mezzo dielettrico non lineare.

-Incidenza di un'onda piana su uno slab di materiale dielettrico lineare.

Nel primo esercizio, un'onda piana incide su una lastra di lunghezza indefinita nelle direzioni x e y e dello spessore di 20 metri. Si è assunto $\mu_r=1$, $\epsilon_r=10$ e $\sigma=0$. Sebbene il problema fosse di natura monodimensionale, è stato studiato come se fosse 3D usando una mesh consistente in una pila di elementi esaedrali lungo la direzione z (direzione di propagazione dell'onda). Sono state usate due discretizzazioni, ciascuna delle quali includeva 200 metri di spazio vuoto prima e 100 metri oltre la lastra. La mesh più rozza consiste di 128 elementi e 2056 incognite, mentre quella più raffinata di 320 elementi e 5128 incognite. Il passo spaziale vale per la mesh rozza $\Delta z=2.5m$ e per quella più raffinata $\Delta z=1m$. Per l'avanzamento della soluzione nel tempo si è adoperato il metodo implicito ed un'ampiezza del time step pari a $\Delta t=1e-9s$.

Si è supposto, in un primo momento, che sulla lastra incidesse un'onda piana il cui andamento temporale fosse rappresentato dalla semplice forma di un gradino. Si è assunto, inoltre, che il campo magnetico avesse ampiezza unitaria e fosse linearmente polarizzato lungo la direzione y . Questo caso,

nonostante la semplicità della sua soluzione analitica, è piuttosto critico dal punto di vista numerico a causa della brusca variazione del fronte d'onda. Le soluzioni ottenute attraverso formulazioni FEM di tipo tradizionale possono, infatti, mostrare accentuate oscillazioni che precedono e seguono il fronte d'onda.

Sono state assunte le seguenti condizioni al contorno: $\mathbf{E} \times \mathbf{n} = 0$, che implica $\mathbf{A} \times \mathbf{n} = 0$, sui due piani di simmetria a $x = \text{costante}$; $\mathbf{H} \times \mathbf{n} = 0$, che si traduce in $\mathbf{W} \times \mathbf{n} = 0$, sui piani di simmetria a $y = \text{costante}$; sulla frontiera $z = \text{cost}$ situata dietro la lastra si è imposto, attraverso l'incognita \mathbf{W} , H_x uguale a zero e H_y essere un gradino unitario; ed infine, sul piano di frontiera $z = \text{cost}$ collocato oltre la lastra, si è imposta la condizione $\mathbf{H} \times \mathbf{n} = 0$, ovvero $\mathbf{W} \times \mathbf{n} = 0$. In questo modo il problema dell'incidenza dell'onda piana sulla lastra dielettrica nello spazio libero è stato rappresentato correttamente, almeno fino al momento in cui uno dei fronti d'onda uscenti dalla lastra raggiunge le superfici di frontiera $z = \text{cost}$.

-La fig.[11] si riferisce alla discretizzazione più rozza. Viene riportato l'andamento spaziale del campo \mathbf{B} come ottenuto attraverso le due formulazioni all'istante $t = 9.2e-7$ s; sulla stessa figura viene altresì rappresentata la differenza tra le due soluzioni numeriche. La parte della mesh usata per discretizzare lo slab è sovrapposta al campo \mathbf{B} .

Le figure [12-13] riportano invece le due soluzioni ottenute per il campo magnetico \mathbf{B} attraverso la discretizzazione più raffinata negli istanti $6.8e-7$ s e $9.2e-7$ s. Oltre alla differenza dei campi magnetici, sulla stessa figura è stato riportato attraverso una rappresentazione ugualmente vettoriale l'andamento del funzionale di errore.

Sulle stesse figg.[12-13] sono stati confrontati attraverso i campi $\mathbf{E}(\mathbf{A})$ ed $\mathbf{H}(\mathbf{W})$ i risultati numerici ottenuti attraverso le due discretizzazioni con la soluzione analitica. Le figure a fondo pagina hanno come ordinata gli errori percentuali delle soluzioni numeriche rispetto alla soluzione vera.

Al fronte d'onda delle soluzioni numeriche è vietato variare in maniera ripida dalle intrinseche proprietà di continuità sulle componenti tangenti delle edge

functions, attraverso le quali sono rappresentati gli andamenti spaziali dei potenziali vettori **A** e **W** e, conseguentemente, i campi **E** e **H**. Tuttavia, come si può osservare, il fenomeno di dispersione del fronte d'onda coinvolge un numero molto limitato di elementi (all'incirca 3).

Si evidenzia come la soluzione ottenuta attraverso la formulazione proposta non contenga eccessive oscillazioni in prossimità del fronte d'onda, ed in questo è qualitativamente superiore rispetto alle soluzioni ottenibili con metodi FEM tradizionali. Si può infine notare che il raffinamento della mesh produce una diminuzione del funzionale di errore Λ , che vale $0.22E-13$ Js per la mesh rozza e $0.19E-14$ per quella raffinata nell'intervallo di tempo $[0, 0.68]$ μ s. Purtroppo, nel caso dell'analisi nel dominio del tempo di un fenomeno transitorio di propagazione, l'informazione sulla localizzazione dell'errore massimo è poco significativa, perchè essa coincide ovviamente col fronte d'onda che si sta muovendo. Nondimeno, l'informazione può ridiventare preziosa se si adottano meshes mobili oppure nel caso dell'analisi nel dominio complesso di fenomeni propagativi pervenuti nella condizione di regime.

Mantenendo inalterati i dati geometrici del problema e le caratteristiche dei materiali si è quindi provato ad introdurre nello slab un'onda piana descritta nel piano $z=0$ attraverso il campo H_y dalla seguente espressione analitica:

$$H_y(z=0,t)=(1-\cos(\omega t)) [1(t)-1(t-T/2)] \quad t \in [0, \Delta T], \quad T=2\pi/\omega=60 \text{ Ns}$$

dove ΔT è l'intervallo totale di studio.

In questo caso, si rilevano a priori, rispetto al caso esaminato precedentemente, essenzialmente due differenze :

I) Prima, imponendo sulla frontiera $z=0$ un forzamento a gradino, si pompava energia all'interno della struttura in tutto il periodo di studio; nel problema attuale ciò avviene solo all'interno della finestra temporale $[0, T/2]$

II) La forma d'onda a coseno rialzato, a differenza del gradino, non impone brusche discontinuità sulle componenti tangenti dei campi e per questo è lecito attendersi che le oscillazioni sulle soluzioni, già ridotte nel caso precedente, siano ora completamente assenti.

Il dominio del problema è stato discretizzato con due differenti mesh una di 320 elementi ed un'altra più raffinata che constava di 1280 elementi. Per l'integrazione temporale si sono adottati gli schemi implicito e Crank-Nicolson con differenti ampiezze del passo temporale.

Per ciascuno dei casi eseguiti, vengono riportati in tabella [4] i valori dell'energia magnetica (calcolata attraverso il potenziale vettore magnetico) in alcuni istanti temporali multipli del periodo ($T=60$ Ns) della funzione in ingresso.

Tempo (s)	Valore Analitico	$\Delta t_1=1.e-9$ s $\Delta l_1=1$ m (implicito) (I)	$\Delta t_1=1.e-9$ s $\Delta l_1=.25$ m (implicito) (II)	$\Delta t_1=.5e-9$ s $\Delta l_1=.25$ m (implicito) (III)	$\Delta t_1=1.e-10$ s $\Delta l_1=.25$ m (implicito) (IV)	$\Delta t_1=1.e-9$ s $\Delta l_1=1.$ m (Crank-Nic.) (V)
30e-9 (.5 T)	.848e-5 J	.845e-5 J	.804e-5 J	.824e-5 J	.842e-5 J	.847e-5 J
90e-9 (1.5 T)	1.696e-5 J	1.67e-5 J	1.42e-5 J	1.54e-5 J	1.66e-5 J	1.698e-5 J
210e-9 (3.5 T)	1.696e-5 J	1.60e-5 J	1.10e-5 J	1.31e-5 J	1.59e-5 J	1.691e-5 J
270e-9 (4.5 T)	1.696e-5 J	1.57e-5 J	1.01e-5 J	1.23e-5 J	1.57e-5 J	1.688e-5 J

Tabella[4] Andamento temporale dell'energia magnetica (Formulazione in A) per differenti ampiezze dei passi di integrazione spaziali e temporali

Dai dati si rileva un decadimento della quantità di energia associata all'onda elettromagnetica quando si adotta lo schema di integrazione implicito che si va attenuando a mano a mano che aumenta il rapporto tra i passi di integrazione spaziale e temporale.

I fenomeni di amplificazione e di dissipazione numerica dell'energia sono ben noti in letteratura[51] e dipendono tanto dalla scelta del metodo di integrazione quanto appunto dal rapporto tra le ampiezze dei passi di integrazione.

Tali fenomeni scompaiono quando si adotti lo schema Crank-Nicolson per il quale vale il cosiddetto criterio di Courant[52] che impone alla velocità di propagazione numerica di essere maggiore-uguale della velocità di propagazione fisica dei campi:

$$c \frac{\Delta t}{\Delta l} \leq \sqrt{\frac{\mu_r \epsilon_r}{n}}$$

dove c è la velocità della luce nel vuoto, e $n=3$ è la dimensione del problema. Ad esempio, nel caso (V) riportato in tabella, Δt vale $1.e-9s$ mentre Δl è pari ad 1m. In queste condizioni, nel vuoto, il primo membro della disuguaglianza vale circa .3 mentre il secondo membro è $\cong .6$ e dunque il criterio di Courant è rispettato.

La fig.[14] mostra l'andamento spaziale del campo H fotografato in diversi istanti. Si parte dal momento in cui l'onda entra nello slab dielettrico, e si arriva all'istante in cui il campo elettromagnetico giunto sulla seconda parete dello slab si divide in un'onda trasmessa ed in una riflessa. Si può notare l'assenza di oscillazioni negli elementi che istante per istante sono toccati dal fronte anteriore e posteriore dell' onda.

In tabella [5] vengono riportati i tempi di elaborazione richiesti dal programma per i casi (I-IV).

Caso	Numero Time-step	Numero incognite	CPU time (s)	Numero medio di iterazioni CG
(I)	600	5128	1443.6	3÷4
(II)	600	20488	5716.5	3÷4
(III)	1200	20488	11072.6	3÷4
(IV)	6000	20488	53832.4	3÷4

Tabella[5] Tempi di Cpu richiesti per i casi (I-IV) su una macchina IBM RISC 6000-modello 350.

-Incidenza di un'onda piana su uno slab di materiale dielettrico non lineare.

Come secondo esercizio si è studiato un problema di propagazione di un'onda piana all'interno di un mezzo dielettrico non lineare[53]. La soluzione ottenuta coll'approccio basato sull'errore costitutivo è stata messa a confronto con un'altra ottenuta con una formulazione monodimensionale di tipo tradizionale alla Galerkin attraverso il potenziale vettore elettrico. La precisione di entrambi gli approcci è stata valutata attraverso il confronto con la soluzione analitica disponibile se si assume una particolare non linearità per la

caratteristica del mezzo e una determinata forma del campo incidente. Il problema affrontato appare particolarmente significativo in quanto comprende la maggior parte dei fenomeni basilari di propagazione elettromagnetica all'interno di mezzi non lineari: biforcazione e caos, brusche discontinuità del fronte d'onda, etc.

Anche se il problema è a carattere mono-dimensionale, quando si è usato l'approccio basato sull'errore costitutivo, il caso è stato studiato come se fosse di tipo 3D, usando delle meshes che consistevano di una pila di elementi esaedrici lungo la direzione z (la direzione di propagazione), imponendo adeguate condizioni di simmetria ($\mathbf{E} \times \mathbf{n} = \mathbf{0}$ da cui $\mathbf{A} \times \mathbf{n} = \mathbf{0}$ sui due piani di simmetria $x = \text{costante}$; e $\mathbf{H} \times \mathbf{n} = \mathbf{0}$ da cui $\mathbf{W} \times \mathbf{n} = \mathbf{0}$ sui due piani di simmetria $y = \text{costante}$) ed usando lo schema di Crank-Nicolson per l'integrazione nel tempo. Nel metodo di Galerkin, invece, ci si è serviti per la soluzione del carattere mono-dimensionale del problema. In questo caso l'integrazione temporale del sistema di equazioni differenziali del secondo ordine associato al problema è stata effettuata attraverso il metodo del quarto ordine di Runge-Kutta-Nyström .

Nella formulazione basata sull'errore costitutivo sono state assunte inoltre le seguenti condizioni al contorno:

- nel piano di frontiera $z = \text{costante}$ nello spazio vuoto prima del dielettrico si è imposto il campo magnetico \mathbf{H} : $H_y = H_0(t)$ dove $H_0(t)$ è un'assegnata funzione del tempo, e H_x uguale a zero;

- nel piano $z = \text{costante}$ situato oltre lo slab si è imposto $\mathbf{H} \times \mathbf{n} = \mathbf{0}$ da cui $\mathbf{W} \times \mathbf{n} = \mathbf{0}$. In questo modo il problema dell'incidenza dell'onda piana sullo slab dielettrico è correttamente rappresentato fino a quando uno dei fronti d'onda uscenti dal dielettrico non raggiunga uno dei piani di frontiera .

Per la formulazione di Galerkin, invece è stato necessario modellare la sola regione interna allo slab dielettrico ed imporre sul contorno le condizioni di impedenza (dette anche di assorbimento) che legano tra loro i campi \mathbf{E} ed \mathbf{H} senza imporre vincoli sui valori.

Anche in questo caso, si sono studiati due distinti sotto-problemi.

In un primo momento si è assunto che la relazione costitutiva dello slab dielettrico fosse di tipo lineare con $\epsilon_r=5$. Si è posto che lo spessore dello slab fosse di 1m. La forma del campo elettrico incidente è:

$$E_x = \frac{(1 + \sqrt{\epsilon_r}) \sin(\pi t / T_0)}{2\epsilon_r \epsilon_0}$$

ed è applicata nell'intervallo temporale $(0, T_0=1\text{ns})$. La soluzione attraverso Galerkin è stata ottenuta usando 199 elementi lineari nello slab e un time-step $\Delta t=0.002\text{ns}$ per l'integrazione temporale effettuata usando, come detto, il metodo esplicito del quarto ordine di Runge-Kutta-Nyström.

La soluzione data dal metodo basato sull'errore è stata prodotta usando 199 elementi nello slab e 99 elementi nella regione di vuoto anteposta allo slab $(-1\text{m}<z<0)$. L'integrazione temporale è effettuata attraverso il metodo di Crank-Nicolson ed è stata assunta un'ampiezza del time-step pari a $\Delta t=0.005\text{ns}$. Le condizioni al contorno sono state imposte in termini del campo magnetico incidente, mentre un campo magnetico tangenziale uguale a zero è stato imposto sul piano di frontiera $z=1\text{m}$, corrispondente all'altra estremità dello slab. Il confronto tra i risultati numerici e la soluzione analitica è mostrato in fig.[15], dove si constata che per entrambi i metodi il fronte d'onda è disperso in un numero estremamente contenuto di elementi.

Si può altresì notare che la soluzione basata sulla formulazione agli edge elements in **A-W**, mostra un differente comportamento rispetto alla soluzione agli elementi finiti alla Galerkin. Non ci sono, infatti, oscillazioni prima che il fronte d'onda raggiunga il punto, mentre qualche leggera oscillazione compare dopo il passaggio dell'onda. Inoltre, il valore di picco è un po' più piccolo rispetto al valore analitico. Questo può essere imputato alle differenti tecniche di integrazione nel tempo.

Per il secondo sotto-caso con entrambe le formulazioni si è impiegata una discretizzazione del solo slab dielettrico, ancora supposto dello spessore di 1m attraverso 499 elementi.

L'espressione generale della famiglia di curve cui appartiene la caratteristica del dielettrico è:

$$E = \frac{(D + \eta D^3)}{\epsilon_0 \epsilon_{r0}} \quad (3)$$

La forma d'onda del campo elettrico incidente, racchiuso nella finestra temporale $[0, T_0=1 \text{ ns}]$ è:

$$E_i(-ct) = \frac{1}{2\epsilon_0 \epsilon_{r0}} \left[D_0 (1 + \eta D_0^2) + \frac{\sqrt{3\eta \epsilon_{r0}}}{2} \left\{ D_0 \sqrt{D_0^2 + \frac{1}{3\eta}} + \frac{1}{3\eta} \ln \left(\frac{D_0 + \sqrt{D_0^2 + \frac{1}{3\eta}}}{\sqrt{\frac{1}{3\eta}}} \right) \right\} \right]$$

dove: $\epsilon_{r0}=5$; $\eta=0.2$;

$$D(z=0^+, t) = D_0(t) = \sin\left(\frac{\pi t}{T_0}\right);$$

Per ottenere la soluzione con Galerkin si è adoperato un time step pari a $\Delta t=0.002\text{ns}$, mentre per la formulazione basata sull'errore costitutivo è stata scelta un'ampiezza del time-step pari a $\Delta t=0.5\text{ps}$. In quest'ultimo caso, come condizione al contorno sul piano $z=0$ è stata imposta la componente tangente del campo \mathbf{H} legata al campo elettrico incidente attraverso l'espressione:

$$H(0^+, t) = \frac{2E_i(t) - E(0^+, t)}{\zeta_0}$$

In queste condizioni la soluzione analitica può essere calcolata facilmente[53].

Questo esempio è particolarmente interessante in quanto la soluzione analitica non è a valore singolo nell'intero dominio di interesse. In altre parole,

è possibile avere una soluzione a valore singolo soltanto nel dominio del piano zt dove non si verificano intersezioni tra diverse curve caratteristiche appartenenti alla famiglia (3); nella parte restante del dominio una soluzione può essere trovata solo ammettendo la presenza di onde d'urto elettromagnetiche (*electromagnetic shock waves*)[54] Come si può vedere in fig.[16-17] il comportamento delle soluzioni numeriche riflette queste peculiarità della soluzione analitica. Infatti, in alcuni punti i risultati numerici sono in buon accordo con la soluzione analitica nell'intero dominio di interesse. In altri punti, la soluzione numerica presenta fortissime oscillazioni che appaiono dopo un certo istante che si può immaginare segni la nascita di un'onda d'urto elettromagnetica.

Nelle figure [18-19-20] sono mostrate le soluzioni numeriche ottenute attraverso il metodo basato sull'errore costitutivo per i campi e per i funzionali di errore in tre istanti appartenenti all'intervallo di studio ($t_1=1.451e-9s$, $t_2=2.376e-9s$, $t_3=3.401e-9s$). In tab.[6] vengono riportati per i medesimi istanti i valori numerici ottenuti attraverso le due formulazioni di alcune grandezze integrali di particolare interesse: le energie magnetica ed elettrica; i valori assoluti dei funzionali di errore elettrico, magnetico e globale nel time-step $[t_k, t_k+\Delta t_k]$ $k=(1,2,3)$ ed i valori degli stessi funzionali (che ricordiamo si esprimono in [Js]) normalizzati rispetto ai prodotti delle corrispondenti energie per il time -step sono invece riportati nelle tabelle [7-8]. Nonostante il fatto che il criterio di Courant sia stato rispettato (infatti $\Delta t/\Delta l=2.5e-10 < (\Delta t/\Delta l)_{crit}=1.92e-9$) salta agli occhi la dissipazione numerica di energia che nell'istante $t_2=2.376e-9s$, è pari al 2.5% e, nell'istante $t_3=3.401e-9s$, arriva addirittura al 12% del valore dell'energia complessiva immessa nel sistema. Si noti, inoltre la corrispondenza che esiste tra la dissipazione numerica ed i valori percentuali dei funzionali di errore. Si giustificano così anche numericamente le conclusioni cui si era giunti nel paragrafo [2.2.3] circa la non conservazione dell'energia per le soluzioni numeriche ottenute attraverso l'approccio presentato: quanto meno sono soddisfatte le relazioni costitutive (errore

costitutivo più alto) tanto più rilevanti possono essere i fenomeni di amplificazione o dissipazione numerica. Questa conclusione resta vera anche se i passi di integrazione spaziale e temporale sono stati scelti in maniera da soddisfare il criterio di Courant.

D'altra parte, se si riporta su un grafico, per ciascun time-step, il valore dell'errore costitutivo commesso all'interno del time-step stesso, si hanno chiare indicazioni circa gli intervalli temporali compresi nel periodo di studio nei quali risulta opportuno migliorare la qualità della soluzione attraverso la riduzione dell'ampiezza del passo di integrazione temporale.

Tabella[6]: Propagazione di un'onda piana in un dielettrico non-lineare. Valori dell'energia magnetica W_{Magn} ed elettrica W_{El} come ottenuti attraverso il potenziale vettore elettrico \mathbf{A} ed il potenziale vettore magnetico \mathbf{W} .

Istante [s]	En. Elettrica (A)	En. Elettrica (W)	En. Magnetica (A)	En. Magnetica (W)
1.e-9	.902E+5 [J]	.902E+5 [J]	.963E+5 [J]	.963E+5 [J]
1.451e-9	.902E+5 [J]	.902E+5 [J]	.963E+5 [J]	.963E+5 [J]
2.376e-9	.869E+5 [J]	.878E+5 [J]	.942E+5 [J]	.932E+5 [J]
3.401e-9	.795E+5 [J]	.801E+5 [J]	.854E+5 [J]	.845E+5 [J]

Tabella[7]: Propagazione di un'onda piana in un dielettrico non-lineare. Valore assoluto dei funzionali di errore elettrico (Λ_{El}), magnetico (Λ_{Magn}) e globale ($\Lambda_{Tot} = \alpha_E \Lambda_{El} + \alpha_H \Lambda_{Magn}$) nell'intervallo di tempo $[t_k, t_k + dt_k]$, ($dt_k = 5e-13s$), ($\alpha_E = \alpha_H = 1$)

Istante [s]	Funz. d'errore elettrico Λ_{El}	Funz. d'errore magnetico Λ_{Magn}	Funz. d'errore globale Λ_{Tot}
1.451e-9	.2135E-10 [Js]	.2135E-10 [Js]	.427E-10 [Js]
2.376e-9	.471E-9 [Js]	.462E-9 [Js]	.933E-9 [Js]
3.401e-9	.456E-9 [Js]	.453E-9 [Js]	.919E-9 [Js]

Tabella[8]: Propagazione di un'onda piana in un dielettrico non-lineare. Valori percentuali dei funzionali di errore rapportati ai valori di riferimento, costituiti dal prodotto delle rispettive energie per il passo temporale: funzionale di errore elettrico $\Lambda_{El}^{(r)} = \Lambda_{El} / (\Delta t W_{El})$, magnetico $\Lambda_{Magn}^{(r)} = \Lambda_{Magn} / (\Delta t W_{Magn})$ e globale $\Lambda_{Tot}^{(r)} = \Lambda_{Tot} / (\Delta t (W_{El} + W_{Magn}))$, nell'intervallo di tempo $[t_k, t_k + \Delta t]$; ($\Delta t = 5e-13s$), ($\alpha_E = \alpha_H = 1$)

Istante [s]	$\Lambda_{EI}^{(r)} = \Lambda_{EI} / W_{EI}$	$\Lambda_{Magn}^{(r)} = \Lambda_{Magn} / W_{Magn}$	$\Lambda_{Tot}^{(r)} = \Lambda_{Tot} / (W_{EI} + W_{Magn})$
1.451e-9	.474E-3 [%]	.442E-3 [%]	.457E-3 [%]
2.376e-9	.108E-1 [%]	.098E-1 [%]	.103E-1 [%]
3.401e-9	.115E-1 [%]	.106E-1 [%]	.110E-1 [%]

Verificata l'efficacia del metodo basato sull'errore costitutivo in problemi non lineari e non-dissipativi, uno dei prossimi obiettivi è quello di studiare il transitorio di un impulso elettromagnetico che propagantesi attraverso lastre di materiale ferro-magnetico. Se il campo incidente non ha ampiezza tale da condurle in saturazione, queste, a parità di spessore, risultano superiori nello schermaggio di interferenze elettromagnetiche alle basse frequenze rispetto alle classiche lastre in rame o in alluminio[55]. Questa ipotesi non viene soddisfatta nello studio di problemi che coinvolgono fenomeni transitori violenti, che pure, per ovvi scopi di sicurezza, risultano di particolare interesse. Le maggiori difficoltà nell'analisi di questa classe di problemi deriva dalla circostanza che il mezzo, seppure supposto non isteretico -ipotesi peraltro realistica nel range di frequenza in cui ci si trova-, è non lineare ed ha conducibilità non nulla.

In passato si è ricorso a formulazioni basate sulle differenze finite [56,57], con gli inevitabili limiti associati nella discretizzazione del dominio nei casi 3D con geometrie complesse, nella stima dell'errore e nella possibilità di un eventuale raffinamento locale che queste metodologie comportano. Alla luce dei risultati presentati, è lecito attendersi che il metodo dell'errore costitutivo risulti adeguato al trattamento del problema.

CAPITOLO V

Minimizzazione dell'errore costitutivo nella MHD e applicazioni ai modelli dinamici di plasma.

Introduzione

L'energia costituisce uno degli ingredienti essenziali del processo di civilizzazione. Mentre il fabbisogno energetico strettamente vitale di un adulto è di circa $1e7$ J/giorno, il consumo procapite di energia nei paesi tecnologicamente avanzati è pari a circa 100 volte tale quantità, il che in termini di potenza si traduce in un fabbisogno di circa $1e4$ W/persona. Partendo da queste cifre, si arriva a stimare la domanda mondiale di energia pari a qualcosa come $5e20$ J/anno. In tabella 1 sono riportati alcuni dati circa le quantità di energia ricavabili attraverso fonti convenzionali[58]. Dai dati riportati si arriva alla conclusione che le sorgenti energetiche mondiali attualmente sfruttate saranno completamente esaurite nell'arco di circa un secolo. E', dunque, essenziale per l'uomo individuare nelle prossime decine di anni nuove fonti -alternative e non- capaci di soddisfare ad una così vasta sete di energia. Una possibile alternativa è costituita dalla fusione nucleare, lo stesso meccanismo con il quale le stelle generano l'energia che viene irradiata nell'universo.

Fonte energetica	Energia mondiale (J)
Carbone	(2-3) $e22$
Petrolio	(3-7) $e21$
Gas	(2-6) $e21$
Uranio 235	$\approx 1e23$
Deuterio (Fusione)	$\approx 1e31$

Tavola [1] Stima delle risorse energetiche mondiali

5.1) Generalita' sulla fusione nucleare.

Il principio fisico fondamentale alla base del processo di fusione nucleare è che la combinazione di atomi leggeri (deuterio, tritio) in atomi più pesanti genera ingenti quantità di energia. L'ultima voce della tabella [1] si riferisce, ad esempio, al quantitativo di energia che sarebbe disponibile se si riuscisse a fondere in elio tutto il deuterio disponibile sulla terra. All'enorme quantità di energia messa a disposizione, la fusione nucleare unisce gli ulteriori vantaggi di avere uno scarso impatto ambientale e di non prestarsi facilmente ad applicazioni militari (ciò è vero in particolar modo per l'approccio basato sul confinamento magnetico).

D'altra parte la generazione di energia attraverso il meccanismo della fusione si pone ai limiti dell'attuale livello tecnologico. Difatti, per arrivare alla fusione nucleare auto-sostenuta (ovvero nella situazione in cui il processo generi più energia di quanta ne richieda per sostenersi), è necessario confinare una mistura di gas leggeri (ad esempio deuterio e tritio) completamente ionizzati (plasma), con una densità n delle particelle, ad una certa temperatura T e per un certo tempo τ_E , dove questi parametri devono soddisfare il cosiddetto criterio di Lawson:

$$n T \tau_E > 5 \times 10^{21} \text{ m}^{-3} \text{ s keV}$$

$$T > 10 \text{ keV}$$

I due principali metodi di confinamento sui quali si sta attualmente lavorando sono quello inerziale e quello magnetico.

Con l'approccio inerziale si intende confinare un plasma ad alta densità ($n \cong 1 \times 10^{30} \text{ m}^{-3}$) per un tempo estremamente breve ($\tau_E \cong 1 \times 10^{-10} \text{ s}$). Viceversa, attraverso l'approccio magnetico, si intende confinare per tempi lunghi ($\tau_E \cong 1 \text{ s}$) plasma a bassa densità ($n \cong 1 \times 10^{20} \text{ m}^{-3}$).

Nell'ambito dell'approccio magnetico, sono state concepite diverse possibili strutture di reattore. Tra queste, quella che ha fornito i migliori risultati sperimentali è la configurazione magnetica toroidale (Tokamak).

Dal punto di vista elettromagnetico, è possibile individuare nella struttura Tokamak tre diversi sistemi che interagiscono reciprocamente: *le bobine alimentate, le strutture conduttrici e il plasma.*

Le *bobine*, alcune delle quali superconduttrici, vengono differenziate in:

-bobine di riscaldamento ohmico (OH), che generano la corrente di plasma per effetto trasformatore (le bobine OH costituiscono il primario, mentre il plasma forma il circuito secondario);

-bobine di campo toroidale (TF) che creano il campo toroidale necessario per il confinamento e la stabilità Magneto-idro-dinamica (MHD);

-bobine di campo poloidale, che provvedono il campo poloidale per il controllo a ciclo aperto e chiuso della posizione e della forma del plasma.

Le principali *strutture conduttrici* che circondano il plasma sono:

-i componenti a diretto contatto col plasma (First Wall e moduli Blanket), che hanno anche il ruolo di scambiatori di calore;

-il Vacuum Vessel, la vera e propria camera di plasma, che ha tra l'altro un compito di protezione;

-il rame nel quale sono immerse le bobine superconduttrici.

Il comportamento elettromagnetico del *plasma* è sostanzialmente quello di un conduttore assi-simmetrico deformabile, la cui evoluzione è determinata dalla densità di forza elettromagnetica $\mathbf{J} \times \mathbf{B}$ e dai fenomeni di trasporto. Lo studio dei fenomeni elettromagnetici in una macchina tokamak è essenziale per la comprensione dei fenomeni fisici che si verificano ivi, e per il progetto dei reattori delle generazioni future.

Sono diversi i problemi elettromagnetici che si pongono nello studio delle macchine Tokamak.

Un primo problema è quello di progettare -soddisfacendo ad un insieme di vincoli geometrici, meccanici, termici, tecnologici- gli avvolgimenti capaci di mantenere il plasma in una desiderata configurazione di equilibrio (Flat Top). Tra le altre importanti fasi dell'analisi elettromagnetica compare il progetto

dello scenario e del sistema di controllo e l'analisi dei rapidi transitori elettromagnetici che si verificano per cause accidentali.

-Il progetto dello scenario consiste nella determinazione degli andamenti temporali delle correnti che devono circolare nelle bobine durante le fasi transitorie normali: la conduzione del plasma dallo stato iniziale a quello di Flat top (processo di Startup), quella di mantenimento del plasma nello stato di equilibrio (progetto del sistema di controllo a ciclo aperto), ed infine il passaggio dalla situazione di Flat top a quella di disinnescamento del processo (processo di Shut-down).

-diverse possibili condizioni accidentali -plasma disruptions, guasto di una bobina superconduttrice- possono verificarsi nel corso di una scarica. In tali casi vengono indotte nelle strutture conduttrici e nelle bobine correnti e tensioni che possono dare origine a danni negli isolamenti, archi elettrici, riscaldamento e (la cosa peggiore) forze di notevole intensità capaci di danneggiare, irreparabilmente, l'intera struttura del reattore. Un loro studio è essenziale per lo sviluppo di un sistema di controllo a ciclo chiuso che possa, se non eliminare, per lo meno ridurre gli effetti di tali incidenti.

Questi fenomeni elettromagnetici transitori sono così complessi sia dal punto di vista geometrico sia da quello fisico, che finora è stato praticamente impossibile sviluppare una singola metodologia che risulti soddisfacente per ciascun particolare aspetto. Una piena descrizione del problema dovrebbe, infatti, portare in conto l'evoluzione simultanea di tre fenomeni differenti ma interagenti nei tre diversi sotto-sistemi[59]:

- bobine conduttrici alimentate da amplificatori a ciclo aperto e/o chiuso.
- correnti parassite indotte nelle strutture conduttrici.
- il comportamento elettromagnetico del plasma.

Pertanto, sono stati proposti modelli semplificati per lo studio dei suddetti fenomeni transitori che pongono di volta in volta l'attenzione su uno solo dei tre fenomeni indicati. Ad esempio, un codice circuitale può provvedere uno strumento approssimato, ma agile per il progetto del sistema di controllo; in

questo caso il plasma e le strutture conduttrici sono schematizzate attraverso un insieme di parametri concentrati, che possono essere determinati attraverso metodologie centrate sugli altri due sistemi.

La simulazione dell'esperimento nei suoi aspetti globali può essere effettuata rivolgendo l'attenzione al sottosistema plasma e portando in conto in modo semplificato i circuiti e le strutture passive.

Se, infine, l'obiettivo da raggiungere è la determinazione dei carichi che agiscono sulle strutture conduttrici passive, queste vengono modellate in modo accurato, mentre il plasma viene rappresentato attraverso un filamento nel quale circola una corrente la cui evoluzione è assegnata.

Il maggior svantaggio di questo tipo di approccio è che comporta l'obbligo di doversi accontentare di una consistenza della soluzione solo approssimata.

La formulazione basata sull'errore può costituire una valida alternativa per superare queste pesanti limitazioni. Rinviando al futuro l'approfondimento di questi aspetti, in quest'ultimo capitolo ci si limiterà, attraverso un esempio basato su un'esemplificazione del modello MHD, a mostrare come con la formulazione proposta possano essere trattati i problemi accoppiati.

5.2) Il modello MHD

-Modello matematico del plasma: le equazioni MHD

Il plasma costituisce l'elemento fondamentale di un reattore per la fusione nucleare. Il suo comportamento macroscopico può essere descritto attraverso il modello MHD[60]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (1a)$$

$$\rho \frac{d\mathbf{v}}{dt} + \mathbf{J} \times \mathbf{B} = \nabla p \quad (1b)$$

$$p = p(\rho) \quad (1c)$$

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = -\nabla \phi \quad (1d)$$

che va accoppiato alle equazioni quasi-stazionarie magnetiche di Maxwell:

$$\nabla \cdot \mathbf{B} = 0 \quad (2a)$$

$$\nabla \times \mathbf{H} = \mathbf{J} \quad (2b)$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t \quad (2c)$$

e alle relazioni costitutive:

$$\mathbf{B} = \mu \mathbf{H} \quad (3a)$$

$$\mathbf{J} = \sigma (\mathbf{E} + \mathbf{E}_{\text{ext}}) \quad (3b)$$

con opportune condizioni iniziali e al contorno.

In queste equazioni:

\mathbf{E}_{ext} è il campo elettrico applicato dall'esterno.

p è la pressione del plasma

ρ è la densità di massa

\mathbf{v} è la velocità media del fluido

$\sigma = \eta^{-1}$ è la conducibilità elettrica

La relazione (3b) è stata scritta per i conduttori fuori dalla zona di plasma.

Nel modello MHD ideale il plasma viene assunto essere un conduttore perfetto e la (1d) si semplifica nella:

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{0} \quad (1e)$$

A dispetto della sua elegante semplicità, il modello MHD ideale è ancora troppo difficoltoso per essere risolto nella maggior parte delle geometrie di interesse e richiede, pertanto, l'assunzione di ipotesi maggiormente semplificative.

-Il modello del plasma all'equilibrio

Come si è detto, una prima classe di problemi di interesse fusionistico consiste nel determinare le configurazioni magnetiche capaci di mantenere il

plasma in condizioni di equilibrio stabile. In queste condizioni, il modello sopra scritto diventa indipendente dal tempo e il termine di velocità \mathbf{v} scompare:

$$\mathbf{J} \times \mathbf{B} = \nabla p \quad (4a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (4b)$$

$$\nabla \times \mathbf{H} = \mathbf{J} \quad (4c)$$

$$\mathbf{B} = \mu \mathbf{H} \quad (4d)$$

Sotto l'ipotesi di assisimmetria (peraltro ben verificata nei problemi di interesse) le equazioni (4b,c,d) assumono il seguente aspetto:

$$\mathbf{B}(R, z) = \frac{\nabla \psi}{R} \times \hat{e}_\varphi + \frac{f}{R} \hat{e}_\varphi \quad (5a)$$

$$\mathbf{J}(R, z) = \frac{\nabla \left(\frac{f}{\mu} \right)}{R} \times \hat{e}_\varphi + L\psi \hat{e}_\varphi \quad (5b)$$

dove:

\hat{e}_φ è il versore nella direzione toroidale;

$\psi(R, z) = \int_0^R R B_z(R, z) dR$ è il flusso poloidale per radiante.

$\frac{f(R, z)}{\mu} = \int_0^R R J_z(R, z) dR$ è il flusso per radiante della densità di corrente poloidale e

$$L\psi = -\frac{\partial}{\partial R} \left(\frac{1}{\mu} \frac{\partial \psi}{R \partial R} \right) - \frac{\partial}{\partial z} \left(\frac{1}{\mu} \frac{\partial \psi}{R \partial z} \right) \quad (5c)$$

Se si sostituiscono le relazioni (5a-b) nella (4a), tenendo conto che nel plasma si può assumere $\mu = \mu_0$ si ottiene:

$$\nabla p = \frac{L\psi}{R} \nabla \psi - \frac{f \nabla f}{\mu_0 R^2} \quad (5d)$$

Inoltre sempre dalla (4a) si deduce:

$$\nabla p \cdot \mathbf{J} = 0 \qquad \nabla p \cdot \mathbf{B} = 0 \qquad (5e)$$

E, pertanto, le linee del campo densità di corrente e le linee del campo di induzione magnetica giacciono sulle superfici descritte da $p=\text{costante}$, che sono pertanto contemporaneamente anche superfici magnetiche e di corrente. Inoltre, per l'assunta simmetria assiale i campi scalari ψ e f sono costanti su una superficie magnetica. Pertanto, se ci si limita a considerare topologie di plasma a singolo asse magnetico, e si sceglie come parametro descrittivo della famiglia di superfici magnetiche la funzione ψ si ottiene:

$$\nabla p(\psi) = \frac{dp}{d\psi} \nabla \psi ; \qquad \nabla f(\psi) = \frac{df}{d\psi} \nabla \psi; \qquad (5f)$$

Ed inserendo le espressioni (5f) nell'equazione di equilibrio (5d) si ottiene l'equazione di Grad-Shafranov:

$$L_\psi = \mathbf{J}_\varphi = R \frac{dp}{d\psi} + \frac{f}{\mu_0} \frac{df}{R d\psi} \qquad (6)$$

dove $p(\psi)$ e $f(\psi)$ sono due funzioni che possono essere prescritte sulla base di osservazioni sperimentali e dipendono dalla storia pregressa della scarica.

Si può dimostrare (*teorema del viriale*) che il plasma non può essere magneticamente confinato senza il contributo di sorgenti esterne. Dal punto di vista del progetto elettromagnetico dell'equilibrio del plasma, possono pertanto essere identificati due distinti problemi:

1) *Assegnata una desiderata configurazione di plasma (cioè, la forma della sezione trasversale S_p , e il profilo della densità di corrente $J_\varphi(R, \psi)$ del plasma), determinare il sistema di correnti che deve attraversare le bobine PF per mantenere l'equilibrio.* (Problema inverso o del Fixed Boundary).

II) *Assegnato un sistema di correnti che passa attraverso le bobine PF, il flusso massimo sulla frontiera e la corrente totale di plasma, determinare la forma e la posizione assunta dalla frontiera di plasma. (Problema diretto o del Free Boundary).*

-Il modello dinamico del plasma

Se si trascurano i rapidi effetti inerziali (cioè se ci si pone su scale temporali lunghe rispetto al tempo caratteristico di assestamento dei fenomeni MHD $\tau \approx 1 \mu\text{s}$), si può approssimare il processo evolutivo del plasma come una successione di equilibri quasi-statici ancora descritti dall'equazione di Grad-Shafranov (6). In questo caso, naturalmente, il profilo della densità di corrente \mathbf{J} , correlata ai profili di $p(\psi, t)$ ed $f(\psi, t)$, varierà in funzione del tempo e, per chiudere il sistema, occorrerà aggiungere altre equazioni (*equazioni del trasporto*[61])

Diversi modelli alternativi possono essere proposti. Una prima scelta è quella di prescrivere l'evoluzione del profilo di corrente. Un'altra possibilità, più consistente dal punto di vista MHD, è quella di ricavare la $f(\psi, t)$ sulla base della conoscenza del profilo della resistività del plasma nella direzione parallela alle linee di flusso, η_{\parallel} . In questo approccio il profilo della pressione $p(\psi, t)$ è prescritto, e il modello risultante è disaccoppiato dalle equazioni di trasporto delle particelle e dell'energia[62].

Nel modello MHD ideale il plasma è assunto essere un conduttore perfetto e pertanto, nel corso dell'evoluzione, i flussi magnetici toroidali e poloidali saranno conservati (modello a flusso adiabatico). Da questa condizione è possibile determinare $f(\psi, t)$. Infatti nel limite di resistività zero, può essere mostrato che la quantità $\frac{d\Phi}{d\psi} = -2\pi q(\psi)$ è costante nel tempo, dove si è definito con $\Phi(\psi)$ il flusso del campo magnetico toroidale attraverso la superficie $\psi = \text{costante}$. Inoltre $f(\psi, t)$ è legato a $q(\psi)$ dalla seguente relazione:

$$f(\psi) = -\frac{2\pi q(\psi)}{\oint \frac{dl_p}{R|\nabla\psi|}} \quad (7)$$

in cui la circuitazione è eseguita lungo il contorno della sezione poloidale della superficie $\psi=\text{costante}$.

Sostituendo l'espressione (7) nella (6) si ottiene un'equazione integro-differenziale che descrive l'evoluzione di un plasma a flusso adiabatico[63].

Le correnti nei conduttori assi-simmetrici passivi può essere determinato attraverso la legge di Faraday:

$$\mathbf{J}_\phi = -\frac{\sigma}{R} \frac{\partial \psi}{\partial t} \quad (8)$$

Le correnti nelle bobine PF possono essere portate in conto attraverso il metodo delle correnti di maglia. Sia $\mathbf{J}_k(R,\varphi,z)$ la densità di corrente associata alla k-ma corrente di maglia, dove la dipendenza da φ sottolinea la possibilità di avere connessioni elettriche non-assisimmetriche. La densità toroidale di corrente nelle regioni delle bobine PF è allora data da:

$$\mathbf{J}_\phi = \sum_{k=1}^N I_k(t) \mathbf{q}_k(R,z) \quad (9)$$

dove N è il numero di maglie indipendenti e \mathbf{q}_k è pari al valor medio di $\hat{\mathbf{e}}_\phi \cdot \mathbf{J}_k$ per $0 \leq \varphi \leq 2\pi$.

Per determinare il comportamento delle nuove incognite $I_k(t)$ sarà necessario introdurre altre equazioni che correlino le $I_k(t)$ all'equazione del flusso poloidale nella regione delle bobine (accoppiamento campi-circuiti).

Queste equazioni [64] possono essere dedotte attraverso la legge di Ohm:

$$\mathbf{J} / \sigma = \mathbf{E} + \mathbf{E}_{\text{ext}} \quad (3b)$$

Difatti, moltiplicando entrambi i membri della (3b) per \mathbf{J}_i e integrando su tutta la regione occupata dalle bobine poloidali, Ω_{PF} , risulta:

$$-\int_{\Omega_{PF}} \mathbf{J}_i \cdot \mathbf{E} dV + \sum R_{ik} I_k(t) = V_i \quad (10)$$

dove:

$V_i = \int_{\Omega_{PF}} \mathbf{J}_i \cdot \mathbf{E}_{Ext} dV$ è la forza elettromotrice impressa all'i-ma maglia;

$R_{ik} = \int_{\Omega_{PF}} \frac{\mathbf{J}_i \cdot \mathbf{J}_k}{\sigma} dV$ è la resistenza mutua tra le maglie i e k.

Il termine della (10) che tiene conto del contributo induttivo può essere espresso come:

$$-\int_{\Omega_{PF}} \mathbf{J}_i \cdot \mathbf{E} dV = 2\pi \int_{\Sigma_{PF}} q_i \frac{\partial \psi}{\partial t} dS + \sum M_{ik}^{Ext} \frac{dl_k}{dt} \quad (11)$$

dove il primo addendo al secondo membro si riferisce alla sola parte assisimmetrica degli avvolgimenti, mentre il secondo addendo porta in conto le connessioni esterne non-assisimmetriche (infatti M_{ik}^{Ext} rappresenta il contributo apportato alla mutua induttanza tra le maglie i e k dai rami non assisimmetrici).

La presenza di strutture passive non-assisimmetriche può essere portata in conto attraverso una conveniente schematizzazione circuitale simile a quella associata alle bobine PF. I parametri concentrati equivalenti possono essere dedotti attraverso un codice 3D per il calcolo delle correnti parassite e successive tecniche di riduzione del modello[65].

Si perviene, in questo modo, ad una modellizzazione del problema dell'evoluzione temporale della distribuzione della corrente di plasma attraverso una successione di equilibri quasi-statici in geometria 2D assi-simmetrica. L'insieme di equazioni integro-differenziali non-lineari associate al modello è stato implementato in un programma di calcolo (denominato PROTEUS [61]) basato sui metodi FEM, che usa una procedura alla Galerkin per ottenere un insieme di equazioni non-lineari differenziali ordinarie che avanzano nel tempo attraverso una tecnica di time-stepping.

La validità del modello è stata dimostrata attraverso il confronto dei risultati numerici con misure sperimentali rilevate nel corso di alcune scariche tipiche della macchina sperimentale ASDEX-U[65-68].

5.3) Il problema MHD in assenza di forze di pressione: un approccio basato sulla minimizzazione dell'errore costitutivo.

Le equazioni costitutive del plasma sono:

$$\mathbf{J} \times \mathbf{B} = \nabla p \quad (1)$$

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \eta \mathbf{J} \quad (2)$$

dove:

$p=p(\psi)$ è la pressione del plasma sulla superficie di flusso ψ

\mathbf{v} è la velocità

$\eta=\sigma^{-1}$ è la matrice di resistività del plasma

Moltiplicando scalarmente l'equazione (2) per \mathbf{B} si può eliminare dall'equazione il vettore velocità \mathbf{v} :

$$(\mathbf{E} - \eta \mathbf{J}) \cdot \mathbf{B} = 0 \rightarrow (\nabla_{\parallel} \mathbf{E} - \mathbf{J}) \cdot \mathbf{B} = 0 \rightarrow (\nabla_{\parallel} \mathbf{E} - \mathbf{J}) \cdot \mathbf{B} = 0 \quad (3)$$

dove σ_{\parallel} è il valore della conducibilità nella direzione parallela alle superfici di flusso

Perchè il modello del problema evolutivo sia MHD consistente, bisogna ricavare l'evoluzione del profilo della pressione $p(\psi)$ attraverso le equazioni del trasporto[61]. Si perviene in questa maniera ad un problema accoppiato che richiede l'introduzione di nuove variabili associate a grandezze di natura non elettromagnetica.

Rinviando al prossimo futuro lo sviluppo, attraverso la formulazione basata sull'errore costitutivo, del modello per il problema accoppiato, ci si limiterà qui ad osservare che, se si suppone di trascurare le forze di pressione, il problema ridiventa puramente elettromagnetico. Difatti la (1) assume la forma:

$$\mathbf{J} \times \mathbf{B} = \mathbf{0} \quad (4)$$

ed impone che all'interno del plasma la densità di corrente \mathbf{J} risulti parallela a \mathbf{B} . Questa semplificazione risulta accettabile qualora si intenda modellare il comportamento di plasmi a bassa temperatura (*plasmi freddi*), ad esempio nella fase di startup nei Tokamak (tipicamente a basso β)

Alle equazioni (3-4) vanno aggiunte le equazioni di Maxwell

$$\nabla \cdot \mathbf{B} = 0 \quad (5a)$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \partial \mathbf{D} / \partial t \quad (5b)$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t \quad (5c)$$

e le relazioni costitutive:

$$\mathbf{B} = \mu \mathbf{H} \quad (6a)$$

$$\mathbf{D} = \epsilon \mathbf{E} \quad (6b)$$

Le equazioni di Maxwell possono essere soddisfatte esattamente introducendo i potenziali vettori magnetico \mathbf{A} ed elettrico \mathbf{W} , definiti come:

$$\mathbf{A} = - \int_0^t \mathbf{E} dt \quad (7a)$$

$$\mathbf{W} = \int_0^t \mathbf{H} dt \quad (7b)$$

il che comporta:

$$\mathbf{E} = -\partial \mathbf{A} / \partial t \quad (8a)$$

$$\mathbf{H} = \partial \mathbf{W} / \partial t \quad (8b)$$

Si può inoltre porre:

$$\mathbf{D}_T(\mathbf{x}, t) = \mathbf{D}_{T_0}(\mathbf{x}) + \int_0^t (\mathbf{J}(\mathbf{x}, t) + \partial \mathbf{D}(\mathbf{x}, t) / \partial t) dt \quad (9)$$

$$\mathbf{B} = \mathbf{B}_0 + \nabla \times \mathbf{A} \quad (10a)$$

$$\mathbf{D}_T = \mathbf{D}_{T_0} + \nabla \times \mathbf{W} \quad (10b)$$

Se si suppone che la corrente di spostamento sia trascurabile rispetto alla corrente di conduzione (limite quasi stazionario magnetico), dalle relazioni (9) e (10) si deduce:

$$\mathbf{J}(\mathbf{x}, t) = \dot{\mathbf{D}}_T(\mathbf{x}, t) = \nabla \times \dot{\mathbf{W}}(\mathbf{x}, t) \quad (11)$$

In queste condizioni la relazione (6b) risulta disaccoppiata dal resto del problema.

La soluzione nella zona del plasma viene quindi cercata attraverso la minimizzazione di un funzionale associato all'errore sulle relazioni costitutive (3,4,6a). Il funzionale deve essere definito positivo ed annullarsi solo nel caso in cui le relazioni costitutive siano esattamente verificate. Tra le infinite definizioni possibili, una scelta ammissibile per il funzionale di errore è, ad esempio:

$$\lambda_p(R, z) = \left\{ I_{C1} \|\mathbf{J} \times \mathbf{B}\| + I_{C2} |(\nabla \times \mathbf{E} - \mathbf{J}) \cdot \mathbf{B}| \right\} + \frac{(\mathbf{B} - \mu \mathbf{H})^2}{2\mu} \quad (12)$$

dove I_{C1} e I_{C2} sono i pesi sulle relazioni costitutive (3) e (4) che devono essere scelti in modo opportuno per bilanciare l'errore tra i vari termini della (12). Si noti che in questo caso i pesi sono dei parametri dimensionali ed hanno la dimensione fisica di una lunghezza. Sostituendo ai campi le espressioni dei potenziali vettori la (12) diventa:

$$\lambda_p(R, z) = \left\{ I_{C1} \|(\nabla \times \dot{\mathbf{W}}) \times (\nabla \times \mathbf{A})\| + I_{C2} |(\nabla \times \dot{\mathbf{A}} - \nabla \times \dot{\mathbf{W}}) \cdot (\nabla \times \mathbf{A})| \right\} + \frac{(\nabla \times \mathbf{A} - \mu \dot{\mathbf{W}})^2}{2\mu} \quad (13)$$

L'espressione (13) può essere adoperata in tutta la regione accessibile al plasma. Il dominio occupato dalle strutture passive -eventualmente anche realizzati con materiali non lineari, come ad esempio le scarpe polari ferromagnetiche adoperate per il trasformatore del Tokamak JET[69]- e dai conduttori può essere invece trattato attraverso il funzionale di errore (2.2.2-2).

L'adozione degli edge-elements garantisce la continuità delle componenti tangenti del campo elettrico e magnetico sulle superfici di interfaccia tra il plasma e le strutture passive. Con la presente formulazione, l'accoppiamento plasma-strutture passive viene, pertanto, realizzato in maniera assolutamente automatica.

Questa considerazione ci porta a ritenere che il modello presentato risulta particolarmente adeguato alla comprensione del fenomeno delle *Halo currents*, che comporta un contatto elettrico fra plasma e prima parete.

Allorchè il plasma, a causa di instabilità di tipo verticale, viene a contatto con la struttura metallica della prima parete si ipotizza che si stabiliscano delle correnti che interessano entrambi i sistemi. A causa del contatto con la parete, tenuto conto del rapporto delle masse fra plasma e strutture esterne, il plasma si raffredda, la temperatura scende quasi istantaneamente e le forze di pressione diventano trascurabili quantomeno nella regione di plasma interessata dalle linee di campo che interessano la parete (halo). Ovviamente, data la rapidità del fenomeno (sulla scala delle centinaia di μs per il tokamak DIII-D) e date le costanti di tempo del meccanismo di diffusione termica in direzione ortogonale alle linee di flusso, può tuttavia rimanere una regione interna del plasma (core) ad alta temperatura.

CONCLUSIONI

L'approccio basato sulla minimizzazione dell'errore costitutivo è stato esteso allo studio di problemi tridimensionali generali di natura elettromagnetica. Attraverso un problema di propagazione si è dimostrato che l'approccio è in grado di produrre risultati soddisfacenti anche in presenza di mezzi non lineari.

I principali vantaggi di questa estensione sono i seguenti:

a) la soluzione è ottenuta attraverso la minimizzazione di un funzionale di errore globale e, da questo, la matrice delle equazioni di sistema è simmetrica e definita positiva;

b) è facilmente disponibile una stima dell'errore numerico locale e globale anche senza conoscere la soluzione esatta;

c) è possibile visualizzare graficamente le zone del dominio ove maggiormente si concentra l'errore sulla soluzione;

d) le informazioni qualitative ottenute attraverso la visualizzazione della mappa d'errore possono essere utilizzate per estendere nella maniera più opportuna lo spazio funzionale all'interno del quale ricercare la soluzione;

e) la formulazione consente un trattamento unificato di problemi stazionari, quasi-stazionari e propagativi;

f) E' stata infine mostrata la possibilità di estendere l'approccio basato sull'errore costitutivo a problemi di natura non esclusivamente elettromagnetica (problemi accoppiati). In particolare, per il modello MHD, sotto l'ipotesi di poter trascurare le forze di pressione, è stata derivata un'espressione ammissibile per il funzionale di errore. Si è evidenziato come l'adozione degli edge elements consenta di realizzare l'accoppiamento in modo completamente automatico.

D'altra parte, le incognite sono in numero doppio rispetto alle classiche formulazioni in termini di **B** o di **E**. Tuttavia, attraverso una scelta conveniente dei pesi, è possibile ottenere lo splitting del modello discretizzato; in tal modo ci si riconduce a dover risolvere, ad ogni passo nel tempo, due sistemi indipendenti di equazioni rispettivamente nelle incognite **A** e **W**.

Appendice A Edge functions

TETRAEDRO

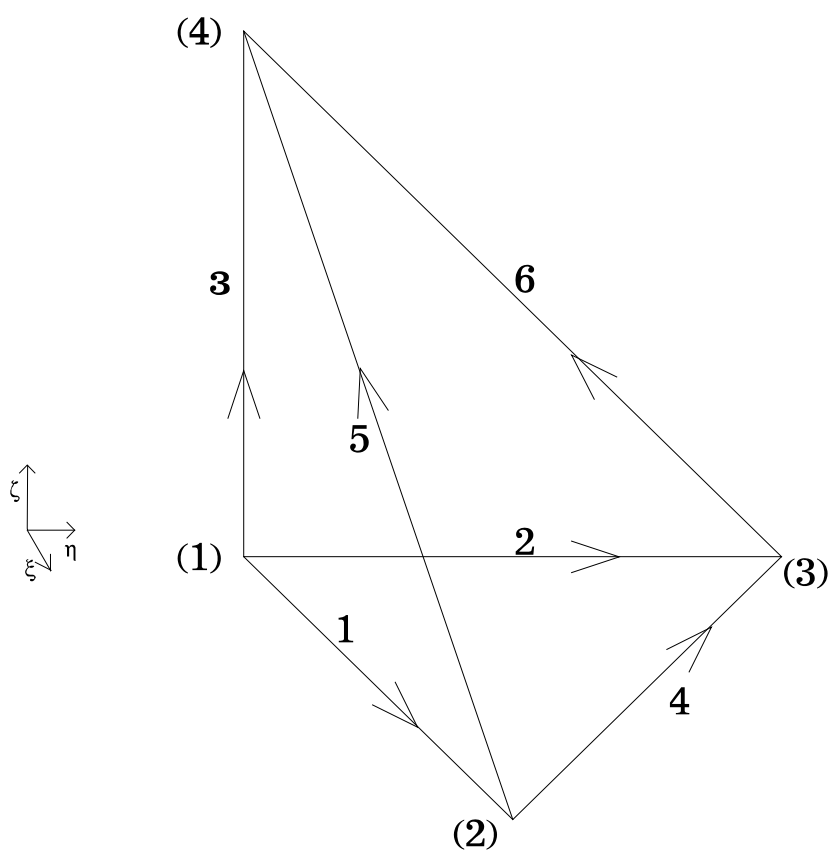


Fig.[1] Numerazione ed orientamento dei lati del tetraedro master per la valutazione delle edge functions.

Le funzioni di edge e i rispettivi rotori si ricavano dalle funzioni nodali come:

$$\mathbf{N}_{IJ} = \psi_I \nabla \psi_J - \psi_J \nabla \psi_I; \quad (6a)$$

$$\nabla \times \mathbf{N}_{IJ} = 2 \nabla \psi_I \times \nabla \psi_J \quad (6b)$$

Con riferimento alla fig.[1] vengono tabellate di seguito le espressioni delle funzioni nodali, delle funzioni di lato e dei rispettivi rotori[44]:

Nodo	Funzione locale di nodo $\psi^{(l)}$
(1)	ξ
(2)	η
(3)	ζ
(4)	$1-\xi-\eta-\zeta$

Lato	Funzione locale di lato: $\mathbf{N}^{(l)}$		
	$N_\xi^{(l)}(\xi)$	$N_\eta^{(l)}(\xi)$	$N_\zeta^{(l)}(\xi)$
1	$1-\eta-\zeta$	ξ	ξ
2	η	$1-\xi-\zeta$	η
3	ζ	ζ	$1-\xi-\eta$
4	$-\eta$	ξ	0
5	$-\zeta$	0	ξ
6	0	$-\zeta$	η

Lato	Funzione globale di lato: $\mathbf{N}^{(e)}$		
	$N_x^{(e)}(\xi)$	$N_y^{(e)}(\xi)$	$N_z^{(e)}(\xi)$
1	$c_{23} \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1} + \xi \mathbf{J}_{31}^{-1}$	$c_{23} \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1} + \xi \mathbf{J}_{32}^{-1}$	$c_{23} \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1} + \xi \mathbf{J}_{33}^{-1}$
2	$\eta \mathbf{J}_{11}^{-1} + c_{13} \mathbf{J}_{21}^{-1} + \eta \mathbf{J}_{31}^{-1}$	$\eta \mathbf{J}_{12}^{-1} + c_{13} \mathbf{J}_{22}^{-1} + \eta \mathbf{J}_{32}^{-1}$	$\eta \mathbf{J}_{13}^{-1} + c_{13} \mathbf{J}_{23}^{-1} + \eta \mathbf{J}_{33}^{-1}$
3	$\zeta \mathbf{J}_{11}^{-1} + \zeta \mathbf{J}_{21}^{-1} + c_{12} \mathbf{J}_{31}^{-1}$	$\zeta \mathbf{J}_{12}^{-1} + \zeta \mathbf{J}_{22}^{-1} + c_{12} \mathbf{J}_{32}^{-1}$	$\zeta \mathbf{J}_{13}^{-1} + \zeta \mathbf{J}_{23}^{-1} + c_{12} \mathbf{J}_{33}^{-1}$
4	$-\eta \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1}$	$-\eta \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1}$	$-\eta \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1}$
5	$-\zeta \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{31}^{-1}$	$-\zeta \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{32}^{-1}$	$-\zeta \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{33}^{-1}$
6	$-\zeta \mathbf{J}_{12}^{-1} + \eta \mathbf{J}_{31}^{-1}$	$-\zeta \mathbf{J}_{22}^{-1} + \eta \mathbf{J}_{32}^{-1}$	$-\zeta \mathbf{J}_{23}^{-1} + \eta \mathbf{J}_{33}^{-1}$

Avendo posto: $c_{12} = 1-\xi-\eta$; $c_{13} = 1-\xi-\zeta$; $c_{23} = 1-\eta-\zeta$;

Lato	Rotore della funzione locale di lato: $\nabla_\xi \times \mathbf{N}_i^{(l)}$		
	$[\nabla_\xi \times \mathbf{N}_i^{(l)}(\xi)]_\xi$	$[\nabla_\xi \times \mathbf{N}_i^{(l)}(\xi)]_\eta$	$[\nabla_\xi \times \mathbf{N}_i^{(l)}(\xi)]_\zeta$
1	0	-2	2
2	2	0	-2
3	-2	2	0
4	0	0	2
5	0	-2	0
6	2	0	0

Lato	Rotore della funzione globale di lato: $\nabla_{\mathbf{x}} \times \mathbf{N}_i^{(\mathbf{e})}$		
	$[\nabla_{\mathbf{x}} \times \mathbf{N}_i^{(\mathbf{e})}(\xi)]_x$	$[\nabla_{\mathbf{x}} \times \mathbf{N}_i^{(\mathbf{e})}(\xi)]_y$	$[\nabla_{\mathbf{x}} \times \mathbf{N}_i^{(\mathbf{e})}(\xi)]_z$
1	$2 (J_{13} - J_{12}) / \text{Det}(\mathbf{J})$	$2 (J_{23} - J_{22}) / \text{Det}(\mathbf{J})$	$2 (J_{33} - J_{32}) / \text{Det}(\mathbf{J})$
2	$2 (J_{11} - J_{13}) / \text{Det}(\mathbf{J})$	$2 (J_{21} - J_{23}) / \text{Det}(\mathbf{J})$	$2 (J_{31} - J_{33}) / \text{Det}(\mathbf{J})$
3	$2 (J_{12} - J_{11}) / \text{Det}(\mathbf{J})$	$2 (J_{22} - J_{21}) / \text{Det}(\mathbf{J})$	$2 (J_{32} - J_{31}) / \text{Det}(\mathbf{J})$
4	$2 J_{13} / \text{Det}(\mathbf{J})$	$2 J_{23} / \text{Det}(\mathbf{J})$	$2 J_{33} / \text{Det}(\mathbf{J})$
5	$-2 J_{12} / \text{Det}(\mathbf{J})$	$-2 J_{22} / \text{Det}(\mathbf{J})$	$-2 J_{32} / \text{Det}(\mathbf{J})$
6	$2 J_{11} / \text{Det}(\mathbf{J})$	$2 J_{21} / \text{Det}(\mathbf{J})$	$2 J_{31} / \text{Det}(\mathbf{J})$

PRISMA

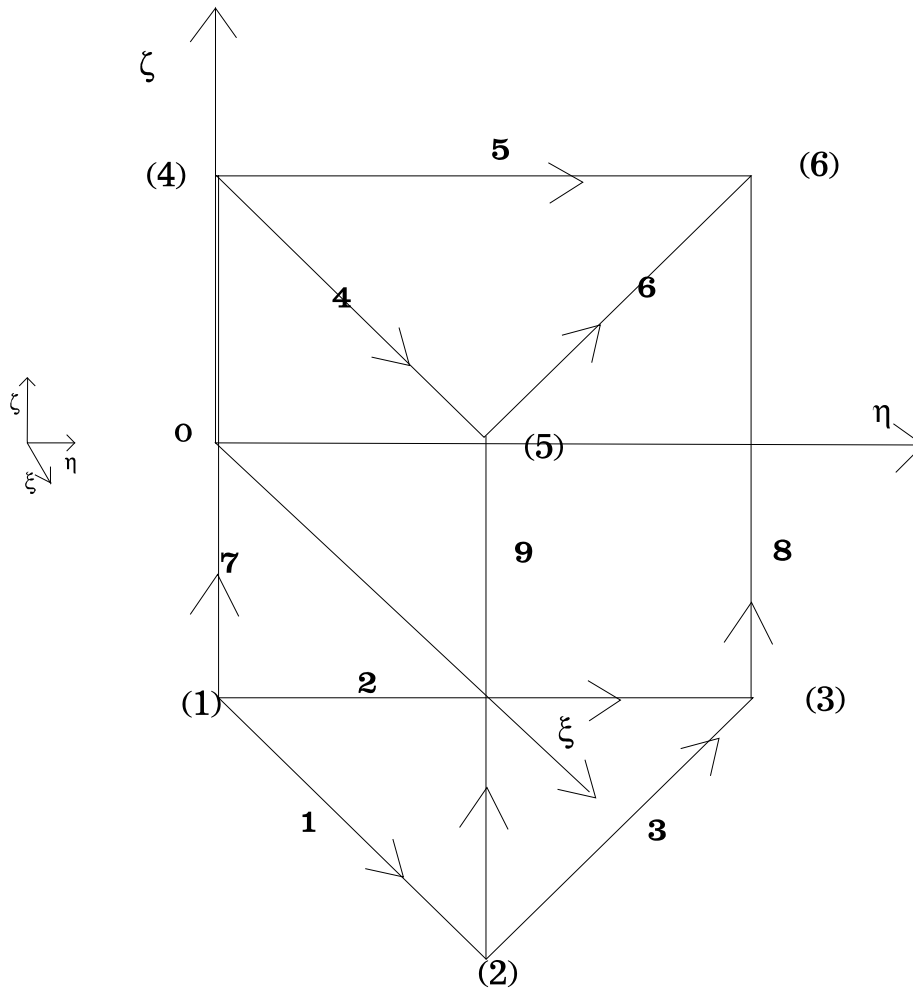


Fig.[2] Numerazione ed orientamento dei lati del pentaedro master per la valutazione delle edge functions.

Con riferimento alla fig.[2] vengono tabellate di seguito le espressioni delle funzioni nodali, delle funzioni di lato e dei rispettivi rotori calcolate attraverso il procedimento indicato[44]:

Nodo	Funzione locale di nodo $\psi^{(l)}$
(1)	$(1-\xi-\eta) (1-\zeta) / 2$
(2)	$\xi (1-\zeta) / 2$
(3)	$\eta (1-\zeta) / 2$
(4)	$(1-\xi-\eta) (1+\zeta) / 2$
(5)	$\xi (1+\zeta) / 2$
(6)	$\eta (1+\zeta) / 2$

Lato	Funzione locale di lato: $\mathbf{N}^{(l)}$		
	$N_\xi^{(l)}(\xi)$	$N_\eta^{(l)}(\xi)$	$N_\zeta^{(l)}(\xi)$
1	$(1-\eta)(1-\zeta)/2$	$\xi(1-\zeta)/2$	0
2	$\eta(1-\zeta)/2$	$(1-\xi)(1-\zeta)/2$	0
3	$-\eta(1-\zeta)/2$	$\xi(1-\zeta)/2$	0
4	$(1-\eta)(1+\zeta)/2$	$\xi(1+\zeta)/2$	0
5	$\eta(1+\zeta)/2$	$(1-\xi)(1+\zeta)/2$	0
6	$-\eta(1+\zeta)/2$	$\xi(1+\zeta)/2$	0
7	0	0	$(1-\xi-\eta)/2$
8	0	0	$\xi/2$
9	0	0	$\eta/2$

Lato	Funzione globale di lato: $\mathbf{N}^{(e)}$		
	$N_x^{(e)}(\xi)$	$N_y^{(e)}(\xi)$	$N_z^{(e)}(\xi)$
1	$(c_{2-} \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1}) c_{3-} / 2$	$(c_{2-} \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1}) c_{3-} / 2$	$(c_{2-} \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1}) c_{3-} / 2$
2	$(\eta \mathbf{J}_{11}^{-1} + c_{1-} \mathbf{J}_{21}^{-1}) c_{3-} / 2$	$(\eta \mathbf{J}_{12}^{-1} + c_{1-} \mathbf{J}_{22}^{-1}) c_{3-} / 2$	$(\eta \mathbf{J}_{13}^{-1} + c_{1-} \mathbf{J}_{23}^{-1}) c_{3-} / 2$
3	$(-\eta \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1}) c_{3-} / 2$	$(-\eta \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1}) c_{3-} / 2$	$(-\eta \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1}) c_{3-} / 2$
4	$(c_{2-} \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1}) c_{3+} / 2$	$(c_{2-} \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1}) c_{3+} / 2$	$(c_{2-} \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1}) c_{3+} / 2$
5	$(\eta \mathbf{J}_{11}^{-1} + c_{1-} \mathbf{J}_{21}^{-1}) c_{3+} / 2$	$(\eta \mathbf{J}_{12}^{-1} + c_{1-} \mathbf{J}_{22}^{-1}) c_{3+} / 2$	$(\eta \mathbf{J}_{13}^{-1} + c_{1-} \mathbf{J}_{23}^{-1}) c_{3+} / 2$
6	$(-\eta \mathbf{J}_{11}^{-1} + \xi \mathbf{J}_{21}^{-1}) c_{3+} / 2$	$(-\eta \mathbf{J}_{12}^{-1} + \xi \mathbf{J}_{22}^{-1}) c_{3+} / 2$	$(-\eta \mathbf{J}_{13}^{-1} + \xi \mathbf{J}_{23}^{-1}) c_{3+} / 2$
7	$c_{12} \mathbf{J}_{31}^{-1} / 2$	$c_{12} \mathbf{J}_{32}^{-1} / 2$	$c_{12} \mathbf{J}_{33}^{-1} / 2$
8	$\xi \mathbf{J}_{31}^{-1} / 2$	$\xi \mathbf{J}_{32}^{-1} / 2$	$\xi \mathbf{J}_{33}^{-1} / 2$
9	$\eta \mathbf{J}_{31}^{-1} / 2$	$\eta \mathbf{J}_{32}^{-1} / 2$	$\eta \mathbf{J}_{33}^{-1} / 2$

Avendo posto: $c_{12} = 1-\xi-\eta$; $c_{1-} = 1-\xi$; $c_{1+} = 1+\xi$;
 $c_{2-} = 1-\eta$; $c_{2+} = 1+\eta$; $c_{3-} = 1-\zeta$; $c_{3+} = 1+\zeta$;

Lato	Rotore della funzione locale di lato: $\nabla_{\xi} \times \mathbf{N}_i^{(0)}$		
	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\xi}$	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\eta}$	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\zeta}$
1	$\xi/2$	$-(1-\eta)/2$	$(1-\zeta)$
2	$(1-\xi)/2$	$-\eta/2$	$-(1-\zeta)$
3	$\xi/2$	$\eta/2$	$(1-\zeta)$
4	$-\xi/2$	$(1-\eta)/2$	$(1+\zeta)$
5	$-(1-\xi)/2$	$\eta/2$	$-(1+\zeta)$
6	$-\xi/2$	$-\eta/2$	$(1+\zeta)$
7	$-1/2$	$1/2$	0
8	0	$-1/2$	0
9	$1/2$	0	0

Esaedro

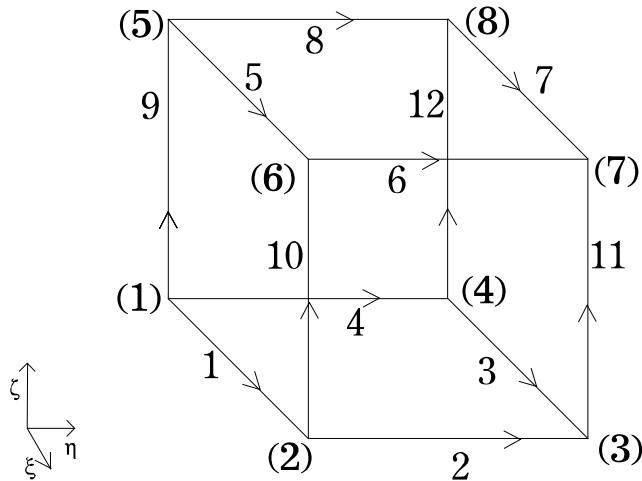


Fig.[3] Numerazione ed orientamento dei lati dell'esaedro master del primo ordine per la valutazione delle edge functions.

Le funzioni di edge e i rispettivi rotori si ricavano dalle funzioni nodali col procedimento precedentemente indicato. Con riferimento alla fig.[3] vengono tabellate di seguito le espressioni delle funzioni nodali, delle funzioni di lato e dei rispettivi rotori[44]:

Nodo	Funzione locale di nodo $\psi^{(0)}$		
(1)	$(1-\xi)(1-\eta)(1-\zeta) / 8$		
(2)	$(1+\xi)(1-\eta)(1-\zeta) / 8$		
(3)	$(1+\xi)(1+\eta)(1-\zeta) / 8$		
(4)	$(1-\xi)(1+\eta)(1-\zeta) / 8$		
(5)	$(1-\xi)(1-\eta)(1+\zeta) / 8$		
(6)	$(1+\xi)(1-\eta)(1+\zeta) / 8$		
(7)	$(1+\xi)(1+\eta)(1+\zeta) / 8$		
(8)	$(1-\xi)(1+\eta)(1+\zeta) / 8$		
Lato	Funzione locale di lato: $\mathbf{N}^{(0)}$		
	$N_{\xi}^{(0)}(\xi)$	$N_{\eta}^{(0)}(\eta)$	$N_{\zeta}^{(0)}(\zeta)$
1	$(1-\eta)(1-\zeta)/8$	0	0
2	0	$(1+\xi)(1-\zeta)/8$	0

3	$(1+\eta)(1-\zeta)/8$	0	0
4	0	$(1-\xi)(1-\zeta)/8$	0
5	$(1-\eta)(1+\zeta)/8$	0	0
6	0	$(1+\xi)(1+\zeta)/8$	0
7	$(1+\eta)(1+\zeta)/8$	0	0
8	0	$(1-\xi)(1+\zeta)/8$	0
9	0	0	$(1-\eta)(1-\xi)/8$
10	0	0	$(1-\eta)(1+\xi)/8$
11	0	0	$(1+\eta)(1+\xi)/8$
12	0	0	$(1+\eta)(1-\xi)/8$

Lato Funzione globale di lato: $\mathbf{N}^{(e)}$

	$N_x^{(e)}(\xi)$	$N_y^{(e)}(\xi)$	$N_z^{(e)}(\xi)$
1	$c_{2-} c_{3-} \mathbf{J}_{11}^{-1}/2$	$c_{2-} c_{3-} \mathbf{J}_{12}^{-1}/2$	$c_{2-} c_{3-} \mathbf{J}_{13}^{-1}/2$
2	$c_{1+} c_{3-} \mathbf{J}_{21}^{-1}/2$	$c_{1+} c_{3-} \mathbf{J}_{22}^{-1}/2$	$c_{1+} c_{3-} \mathbf{J}_{23}^{-1}/2$
3	$c_{2-} c_{3+} \mathbf{J}_{11}^{-1}/2$	$c_{2-} c_{3+} \mathbf{J}_{12}^{-1}/2$	$c_{2-} c_{3+} \mathbf{J}_{13}^{-1}/2$
4	$c_{1-} c_{3-} \mathbf{J}_{21}^{-1}/2$	$c_{1-} c_{3-} \mathbf{J}_{22}^{-1}/2$	$c_{1-} c_{3-} \mathbf{J}_{23}^{-1}/2$
5	$c_{2+} c_{3-} \mathbf{J}_{11}^{-1}/2$	$c_{2+} c_{3-} \mathbf{J}_{12}^{-1}/2$	$c_{2+} c_{3-} \mathbf{J}_{13}^{-1}/2$
6	$c_{1+} c_{3+} \mathbf{J}_{21}^{-1}/2$	$c_{1+} c_{3+} \mathbf{J}_{22}^{-1}/2$	$c_{1+} c_{3+} \mathbf{J}_{23}^{-1}/2$
7	$c_{2+} c_{3+} \mathbf{J}_{11}^{-1}/2$	$c_{2+} c_{3+} \mathbf{J}_{12}^{-1}/2$	$c_{2+} c_{3+} \mathbf{J}_{13}^{-1}/2$
8	$c_{1-} c_{3+} \mathbf{J}_{21}^{-1}/2$	$c_{1-} c_{3+} \mathbf{J}_{22}^{-1}/2$	$c_{1-} c_{3+} \mathbf{J}_{23}^{-1}/2$
9	$c_{1-} c_{2-} \mathbf{J}_{31}^{-1}/2$	$c_{1-} c_{2-} \mathbf{J}_{32}^{-1}/2$	$c_{1-} c_{2-} \mathbf{J}_{33}^{-1}/2$
10	$c_{1+} c_{2-} \mathbf{J}_{31}^{-1}/2$	$c_{1+} c_{2-} \mathbf{J}_{32}^{-1}/2$	$c_{1+} c_{2-} \mathbf{J}_{33}^{-1}/2$
11	$c_{1+} c_{2+} \mathbf{J}_{31}^{-1}/2$	$c_{1+} c_{2+} \mathbf{J}_{32}^{-1}/2$	$c_{1+} c_{2+} \mathbf{J}_{33}^{-1}/2$
12	$c_{1-} c_{2+} \mathbf{J}_{31}^{-1}/2$	$c_{1-} c_{2+} \mathbf{J}_{32}^{-1}/2$	$c_{1-} c_{2+} \mathbf{J}_{33}^{-1}/2$

Avendo posto: $c_{1-} = (1-\xi)/2$; $c_{1+} = (1+\xi)/2$;

$$c_{2-} = (1-\eta)/2; \quad c_{2+} = (1+\eta)/2; \quad c_{3-} = (1-\zeta)/2; \quad c_{3+} = (1+\zeta)/2;$$

Lato Rotore della funzione globale di lato: $\nabla_x \times \mathbf{N}_i^{(e)}$

	$[\nabla_x \times \mathbf{N}_i^{(e)}(\xi)]_x$	$[\nabla_x \times \mathbf{N}_i^{(e)}(\xi)]_y$	$[\nabla_x \times \mathbf{N}_i^{(e)}(\xi)]_z$
1	$-c_{2-} \mathbf{J}_{12} + c_{3-} \mathbf{J}_{13}$	$-c_{2-} \mathbf{J}_{22} + c_{3-} \mathbf{J}_{23}$	$-c_{2-} \mathbf{J}_{32} + c_{3-} \mathbf{J}_{33}$
2	$c_{3-} \mathbf{J}_{13} + c_{1+} \mathbf{J}_{11}$	$c_{3-} \mathbf{J}_{23} + c_{1+} \mathbf{J}_{21}$	$c_{3-} \mathbf{J}_{33} + c_{1+} \mathbf{J}_{31}$

3	$-C_{2+} J_{12} - C_{3-} J_{13}$	$-C_{2+} J_{22} - C_{3-} J_{23}$	$-C_{2+} J_{32} - C_{3-} J_{33}$
4	$-C_{3-} J_{13} + C_{1-} J_{11}$	$-C_{3-} J_{23} + C_{1-} J_{21}$	$-C_{3-} J_{33} + C_{1-} J_{31}$
5	$C_{2-} J_{12} + C_{3+} J_{13}$	$C_{2-} J_{22} + C_{3+} J_{23}$	$C_{2-} J_{32} + C_{3+} J_{33}$
6	$C_{3+} J_{13} - C_{1+} J_{11}$	$C_{3+} J_{23} - C_{1+} J_{21}$	$C_{3+} J_{33} - C_{1+} J_{31}$
7	$C_{2+} J_{12} - C_{3+} J_{13}$	$C_{2+} J_{22} - C_{3+} J_{23}$	$C_{2+} J_{32} - C_{3+} J_{33}$
8	$-C_{3+} J_{13} - C_{1-} J_{11}$	$-C_{3+} J_{23} - C_{1-} J_{21}$	$-C_{3+} J_{33} - C_{1-} J_{31}$
9	$-C_{1-} J_{11} + C_{2-} J_{12}$	$-C_{1-} J_{21} + C_{2-} J_{22}$	$-C_{1-} J_{31} + C_{2-} J_{32}$
10	$-C_{1+} J_{11} - C_{2-} J_{12}$	$-C_{1+} J_{21} - C_{2-} J_{22}$	$-C_{1+} J_{31} - C_{2-} J_{32}$
11	$C_{1+} J_{11} - C_{2+} J_{12}$	$C_{1+} J_{21} - C_{2+} J_{22}$	$C_{1+} J_{31} - C_{2+} J_{32}$
12	$C_{1-} J_{11} + C_{2+} J_{12}$	$C_{1-} J_{21} + C_{2+} J_{22}$	$C_{1-} J_{31} + C_{2+} J_{32}$

Avendo posto: $c_{1-} = (1-\xi) / (8^* \text{Def}[\mathbf{J}]);$ $c_{1+} = (1+\xi) / (8^* \text{Def}[\mathbf{J}]);$
 $c_{2-} = (1-\eta) / (8^* \text{Def}[\mathbf{J}]);$ $c_{2+} = (1+\eta) / (8^* \text{Def}[\mathbf{J}]);$
 $c_{3-} = (1-\zeta) / (8^* \text{Def}[\mathbf{J}]);$ $c_{3+} = (1+\zeta) / (8^* \text{Def}[\mathbf{J}]);$

Esaedro quadratico

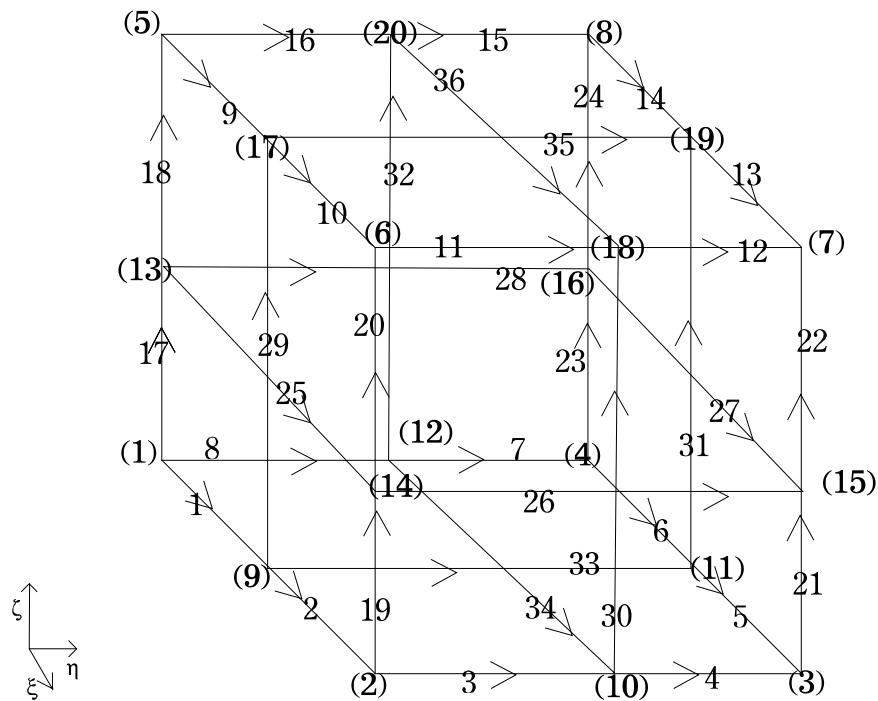


Fig.[4] Numerazione ed orientamento dei lati dell'esaedro master del secondo ordine per la valutazione delle edge functions.

Le funzioni di edge e i rispettivi rotori si ricavano dalle funzioni nodali col procedimento precedentemente indicato. Con riferimento alla fig.[4] vengono tabellate di seguito le espressioni delle funzioni nodali, delle funzioni di lato e dei rispettivi rotori [45]:

Nodo	Funzione locale di nodo $\psi^{(l)}$
(1)	$\xi\eta\zeta(1-\xi)(1-\eta)(1-\zeta) / 8$
(2)	$\xi\eta\zeta(1+\xi)(1-\eta)(1-\zeta) / 8$
(3)	$\xi\eta\zeta(1+\xi)(1+\eta)(1-\zeta) / 8$
(4)	$\xi\eta\zeta(1-\xi)(1+\eta)(1-\zeta) / 8$
(5)	$\xi\eta\zeta(1-\xi)(1-\eta)(1+\zeta) / 8$
(6)	$\xi\eta\zeta(1+\xi)(1-\eta)(1+\zeta) / 8$
(7)	$\xi\eta\zeta(1+\xi)(1+\eta)(1+\zeta) / 8$
(8)	$\xi\eta\zeta(1-\xi)(1+\eta)(1+\zeta) / 8$
(9)	$\eta\zeta(1-\xi^2)(1-\eta)(1-\zeta) / 16$
(10)	$\xi\zeta(1+\xi)(1-\eta^2)(1-\zeta) / 16$
(11)	$\eta\zeta(1-\xi^2)(1+\eta)(1-\zeta) / 16$
(12)	$\xi\zeta(1-\xi)(1-\eta^2)(1-\zeta) / 16$
(13)	$\xi\eta(1-\xi)(1-\eta)(1-\zeta^2) / 16$
(14)	$\xi\eta(1+\xi)(1-\eta)(1-\zeta^2) / 16$
(15)	$\xi\eta(1+\xi)(1+\eta)(1-\zeta^2) / 16$
(16)	$\xi\eta(1-\xi)(1+\eta)(1-\zeta^2) / 16$
(17)	$\eta\zeta(1-\xi^2)(1-\eta)(1+\zeta) / 16$
(18)	$\xi\zeta(1+\xi)(1-\eta^2)(1+\zeta) / 16$
(19)	$\eta\zeta(1-\xi^2)(1+\eta)(1+\zeta) / 16$
(20)	$\xi\zeta(1-\xi)(1-\eta^2)(1+\zeta) / 16$

Lato Funzione globale di lato: $\mathbf{N}^{(e)}$

	$N_x^{(\theta)}(\xi)$	$N_v^{(\theta)}(\xi)$	$N_z^{(\theta)}(\xi)$
1	$(-\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{11}^{-1}/2$	$(-\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{12}^{-1}/2$	$(-\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{13}^{-1}/2$
2	$(\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{11}^{-1}/2$	$(\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{12}^{-1}/2$	$(\xi/2-\eta-\zeta-1)c_{2-}c_{3-}J_{13}^{-1}/2$
3	$(\xi-\eta/2-\zeta-1)c_{1+}c_{3-}J_{21}^{-1}/2$	$(\xi-\eta/2-\zeta-1)c_{1+}c_{3-}J_{22}^{-1}/2$	$(\xi-\eta/2-\zeta-1)c_{1+}c_{3-}J_{23}^{-1}/2$
4	$(\xi+\eta/2-\zeta-1)c_{1+}c_{3-}J_{21}^{-1}/2$	$(\xi+\eta/2-\zeta-1)c_{1+}c_{3-}J_{22}^{-1}/2$	$(\xi+\eta/2-\zeta-1)c_{1+}c_{3-}J_{23}^{-1}/2$
5	$(\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{11}^{-1}/2$	$(\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{12}^{-1}/2$	$(\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{13}^{-1}/2$
6	$(-\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{11}^{-1}/2$	$(-\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{12}^{-1}/2$	$(-\xi/2+\eta-\zeta-1)c_{2+}c_{3-}J_{13}^{-1}/2$
7	$(-\xi+\eta/2-\zeta-1)c_{1-}c_{3-}J_{21}^{-1}/2$	$(-\xi+\eta/2-\zeta-1)c_{1-}c_{3-}J_{22}^{-1}/2$	$(-\xi+\eta/2-\zeta-1)c_{1-}c_{3-}J_{23}^{-1}/2$
8	$(-\xi-\eta/2-\zeta-1)c_{1-}c_{3-}J_{21}^{-1}/2$	$(-\xi-\eta/2-\zeta-1)c_{1-}c_{3-}J_{22}^{-1}/2$	$(-\xi-\eta/2-\zeta-1)c_{1-}c_{3-}J_{23}^{-1}/2$
9	$(-\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{11}^{-1}/2$	$(-\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{12}^{-1}/2$	$(-\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{13}^{-1}/2$
10	$(\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{11}^{-1}/2$	$(\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{12}^{-1}/2$	$(\xi/2-\eta+\zeta-1)c_{2-}c_{3+}J_{13}^{-1}/2$
11	$(\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{21}^{-1}/2$	$(\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{22}^{-1}/2$	$(\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{23}^{-1}/2$
12	$(\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{21}^{-1}/2$	$(\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{22}^{-1}/2$	$(\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{23}^{-1}/2$
13	$(\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{11}^{-1}/2$	$(\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{12}^{-1}/2$	$(\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{13}^{-1}/2$
14	$(-\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{11}^{-1}/2$	$(-\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{12}^{-1}/2$	$(-\xi/2+\eta+\zeta-1)c_{2+}c_{3+}J_{13}^{-1}/2$
15	$(-\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{21}^{-1}/2$	$(-\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{22}^{-1}/2$	$(-\xi+\eta/2+\zeta-1)c_{1-}c_{3+}J_{23}^{-1}/2$
16	$(-\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{21}^{-1}/2$	$(-\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{22}^{-1}/2$	$(-\xi-\eta/2+\zeta-1)c_{1-}c_{3+}J_{23}^{-1}/2$
17	$(-\xi-\eta-\zeta/2-1)c_{1-}c_{2-}J_{31}^{-1}/2$	$(-\xi-\eta-\zeta/2-1)c_{1-}c_{2-}J_{32}^{-1}/2$	$(-\xi-\eta-\zeta/2-1)c_{1-}c_{2-}J_{33}^{-1}/2$
18	$(-\xi-\eta+\zeta/2-1)c_{1-}c_{2-}J_{31}^{-1}/2$	$(-\xi-\eta+\zeta/2-1)c_{1-}c_{2-}J_{32}^{-1}/2$	$(-\xi-\eta+\zeta/2-1)c_{1-}c_{2-}J_{33}^{-1}/2$
19	$(\xi-\eta-\zeta/2-1)c_{1+}c_{2-}J_{31}^{-1}/2$	$(\xi-\eta-\zeta/2-1)c_{1+}c_{2-}J_{32}^{-1}/2$	$(\xi-\eta-\zeta/2-1)c_{1+}c_{2-}J_{33}^{-1}/2$
20	$(\xi-\eta+\zeta/2-1)c_{1+}c_{2-}J_{31}^{-1}/2$	$(\xi-\eta+\zeta/2-1)c_{1+}c_{2-}J_{32}^{-1}/2$	$(\xi-\eta+\zeta/2-1)c_{1+}c_{2-}J_{33}^{-1}/2$
21	$(\xi+\eta-\zeta/2-1)c_{1+}c_{2+}J_{31}^{-1}/2$	$(\xi+\eta-\zeta/2-1)c_{1+}c_{2+}J_{32}^{-1}/2$	$(\xi+\eta-\zeta/2-1)c_{1+}c_{2+}J_{33}^{-1}/2$
22	$(\xi+\eta+\zeta/2-1)c_{1+}c_{2+}J_{31}^{-1}/2$	$(\xi+\eta+\zeta/2-1)c_{1+}c_{2+}J_{32}^{-1}/2$	$(\xi+\eta+\zeta/2-1)c_{1+}c_{2+}J_{33}^{-1}/2$
23	$(-\xi+\eta-\zeta/2-1)c_{1-}c_{2+}J_{31}^{-1}/2$	$(-\xi+\eta-\zeta/2-1)c_{1-}c_{2+}J_{32}^{-1}/2$	$(-\xi+\eta-\zeta/2-1)c_{1-}c_{2+}J_{33}^{-1}/2$
24	$(-\xi+\eta+\zeta/2-1)c_{1-}c_{2+}J_{31}^{-1}/2$	$(-\xi+\eta+\zeta/2-1)c_{1-}c_{2+}J_{32}^{-1}/2$	$(-\xi+\eta+\zeta/2-1)c_{1-}c_{2+}J_{33}^{-1}/2$
25	$2c_{2-}c_{3-}c_{3+}J_{11}^{-1}$	$2c_{2-}c_{3-}c_{3+}J_{12}^{-1}$	$2c_{2-}c_{3-}c_{3+}J_{13}^{-1}$
26	$2c_{1+}c_{3-}c_{3+}J_{21}^{-1}$	$2c_{1+}c_{3-}c_{3+}J_{22}^{-1}$	$2c_{1+}c_{3-}c_{3+}J_{23}^{-1}$
27	$2c_{2+}c_{3-}c_{3+}J_{11}^{-1}$	$2c_{2+}c_{3-}c_{3+}J_{12}^{-1}$	$2c_{2+}c_{3-}c_{3+}J_{13}^{-1}$
28	$2c_{1-}c_{3-}c_{3+}J_{21}^{-1}$	$2c_{1-}c_{3-}c_{3+}J_{22}^{-1}$	$2c_{1-}c_{3-}c_{3+}J_{23}^{-1}$
29	$2c_{1-}c_{1+}c_{2-}J_{31}^{-1}$	$2c_{1-}c_{1+}c_{2-}J_{32}^{-1}$	$2c_{1-}c_{1+}c_{2-}J_{33}^{-1}$
30	$2c_{1+}c_{2-}c_{2+}J_{31}^{-1}$	$2c_{1+}c_{2-}c_{2+}J_{32}^{-1}$	$2c_{1+}c_{2-}c_{2+}J_{33}^{-1}$
31	$2c_{1-}c_{1+}c_{2+}J_{31}^{-1}$	$2c_{1-}c_{1+}c_{2+}J_{32}^{-1}$	$2c_{1-}c_{1+}c_{2+}J_{33}^{-1}$
32	$2c_{1-}c_{2-}c_{2+}J_{31}^{-1}$	$2c_{1-}c_{2-}c_{2+}J_{32}^{-1}$	$2c_{1-}c_{2-}c_{2+}J_{33}^{-1}$
33	$2c_{1+}c_{1-}c_{3-}J_{21}^{-1}$	$2c_{1+}c_{1-}c_{3-}J_{22}^{-1}$	$2c_{1+}c_{1-}c_{3-}J_{23}^{-1}$
34	$2c_{2-}c_{2+}c_{3-}J_{11}^{-1}$	$2c_{2-}c_{2+}c_{3-}J_{12}^{-1}$	$2c_{2-}c_{2+}c_{3-}J_{13}^{-1}$
35	$2c_{1+}c_{1-}c_{3+}J_{21}^{-1}$	$2c_{1+}c_{1-}c_{3+}J_{22}^{-1}$	$2c_{1+}c_{1-}c_{3+}J_{23}^{-1}$
36	$2c_{2-}c_{2+}c_{3+}J_{11}^{-1}$	$2c_{2-}c_{2+}c_{3+}J_{12}^{-1}$	$2c_{2-}c_{2+}c_{3+}J_{13}^{-1}$

Avendo posto: $c_{1-} = (1-\xi)/2$; $c_{1+} = (1+\xi)/2$;
 $c_{2-} = (1-\eta)/2$; $c_{2+} = (1+\eta)/2$;

$$c_{3-} = (1-\zeta)/2; \quad c_{3+} = (1+\zeta)/2;$$

Lato	Rotore della funzione locale di lato: $\nabla_{\xi} \times \mathbf{N}_i^{(0)}$		
	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\xi}$	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\eta}$	$[\nabla_{\xi} \times \mathbf{N}_i^{(0)}(\xi)]_{\zeta}$
1	0	$-c_{2-} (-\xi/2 - \eta - 2\zeta)$	$c_{3-} (-\xi/2 - 2\eta - \zeta)$
2	0	$-c_{2-} (\xi/2 - \eta - 2\zeta)$	$c_{3-} (\xi/2 - 2\eta - \zeta)$
3	$c_{1+} (\xi - \eta/2 - 2\zeta)$	0	$c_{3-} (2\xi - \eta/2 - \zeta)$
4	$c_{1+} (\xi + \eta/2 - 2\zeta)$	0	$c_{3-} (2\xi + \eta/2 - \zeta)$
5	0	$-c_{2+} (\xi/2 + \eta - 2\zeta)$	$-c_{3-} (\xi/2 + 2\eta - \zeta)$
6	0	$-c_{2+} (-\xi/2 + \eta - 2\zeta)$	$-c_{3-} (-\xi/2 + 2\eta - \zeta)$
7	$c_{1-} (-\xi + \eta/2 - 2\zeta)$	0	$-c_{3-} (-2\xi + \eta/2 - \zeta)$
8	$c_{1-} (-\xi - \eta/2 - 2\zeta)$	0	$-c_{3-} (-2\xi - \eta/2 - \zeta)$
9	0	$c_{2-} (-\xi/2 - \eta + 2\zeta)$	$c_{3+} (-\xi/2 - 2\eta + \zeta)$
10	0	$c_{2-} (\xi/2 - \eta + 2\zeta)$	$c_{3+} (\xi/2 - 2\eta + \zeta)$
11	$-c_{1+} (\xi - \eta/2 + 2\zeta)$	0	$c_{3+} (2\xi - \eta/2 + \zeta)$
12	$-c_{1+} (\xi + \eta/2 + 2\zeta)$	0	$c_{3+} (2\xi + \eta/2 + \zeta)$
13	0	$c_{2+} (\xi/2 + \eta + 2\zeta)$	$-c_{3+} (\xi/2 + 2\eta + \zeta)$
14	0	$c_{2+} (-\xi/2 + \eta + 2\zeta)$	$-c_{3+} (-\xi/2 + 2\eta + \zeta)$
15	$-c_{1-} (-\xi + \eta/2 + 2\zeta)$	0	$-c_{3+} (-2\xi + \eta/2 + \zeta)$
16	$-c_{1-} (-\xi - \eta/2 + 2\zeta)$	0	$-c_{3+} (-2\xi - \eta/2 + \zeta)$
17	$-c_{1-} (-\xi - 2\eta - \zeta/2)$	$c_{2-} (-2\xi - \eta - \zeta/2)$	0
18	$-c_{1-} (-\xi - 2\eta + \zeta/2)$	$c_{2-} (-2\xi - \eta + \zeta/2)$	0
19	$-c_{1+} (\xi - 2\eta - \zeta/2)$	$-c_{2-} (2\xi - \eta - \zeta/2)$	0
20	$-c_{1+} (\xi - 2\eta + \zeta/2)$	$-c_{2-} (2\xi - \eta + \zeta/2)$	0
21	$c_{1+} (\xi + 2\eta - \zeta/2)$	$-c_{2+} (2\xi + \eta - \zeta/2)$	0
22	$c_{1+} (\xi + 2\eta + \zeta/2)$	$-c_{2+} (2\xi + \eta + \zeta/2)$	0
23	$c_{1-} (-\xi + 2\eta - \zeta/2)$	$c_{2+} (-2\xi + \eta - \zeta/2)$	0
24	$c_{1-} (-\xi + 2\eta + \zeta/2)$	$c_{2+} (-2\xi + \eta + \zeta/2)$	0
25	0	$-4\zeta c_{2-}$	c_{33}
26	$4\zeta c_{1+}$	0	c_{33}
27	0	$-4\zeta c_{2+}$	$-c_{33}$
28	$4\zeta c_{1-}$	0	$-c_{33}$
29	$-c_{11}$	$-4\xi c_{2-}$	0
30	$-4\eta c_{1+}$	$-c_{22}$	0
31	c_{11}	$-4\xi c_{2+}$	0
32	$-4\eta c_{1-}$	c_{22}	0
33	c_{11}	0	$-4\xi c_{3-}$
34	0	c_{22}	$4\eta c_{3-}$
35	$-c_{11}$	0	$-4\xi c_{3+}$
36	0	$-c_{22}$	$4\eta c_{3+}$

Avendo posto:

$$\begin{aligned}
 c_{1-} &= (1-\xi)/8; & c_{1+} &= (1+\xi)/8; & c_{11} &= (1-\xi^2)/4 \\
 c_{2-} &= (1-\eta)/8; & c_{2+} &= (1+\eta)/8; & c_{22} &= (1-\eta^2)/4 \\
 c_{3-} &= (1-\zeta)/8; & c_{3+} &= (1+\zeta)/8; & c_{33} &= (1-\zeta^2)/4
 \end{aligned}$$

Appendice B

Scelta dei pesi sulle relazioni costitutive per il disaccoppiamento delle variabili

Nel paragrafo (3.3) si è visto che si poteva ottenere il valore delle incognite discrete \mathbf{a} e \mathbf{w} associate ai potenziali vettori $\mathbf{A}(t_{k+1})$ e $\mathbf{W}(t_{k+1})$ al k-mo time-step risolvendo il sistema algebrico non lineare (3.3-11) attraverso il metodo di Newton.

Alla j^{ma} iterazione il metodo di Newton si riduce a risolvere il seguente sistema lineare:

$$\begin{bmatrix} \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a} \partial \mathbf{w}} \\ \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{a}} & \frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w} \partial \mathbf{w}} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{a}^{(j+1)} \\ \Delta \mathbf{w}^{(j+1)} \end{bmatrix} = - \begin{bmatrix} \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}} \\ \frac{\partial \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{w}} \end{bmatrix} \quad (1)$$

dove al k^{mo} time-step la matrice del sistema si ottiene modificando attraverso le relazioni (3.3-6) -che tengono conto delle condizioni al contorno- l'Hessiano del funzionale Λ_k (cioè della restrizione all'intervallo $[t_k, t_k + \Delta t_k]$ del funzionale globale di errore Λ).

Si riportano di seguito le espressioni dei blocchi dell'Hessiano fuori dalla diagonale principale che determinano l'accoppiamento tra le incognite associate ai due potenziali:

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \partial \mathbf{w}_s} = \int_V d\mathbf{x} \int_k^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k^2} \{ \alpha_E [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_H [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \} dt \quad (2a)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \partial \mathbf{a}_s} = - \int_V d\mathbf{x} \int_k^{t_{k+1}} \frac{(t-t_k)}{\Delta t_k^2} \{ \alpha_H [(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s)] - \alpha_E [(\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r)] \} dt \quad (2b)$$

dove l'integrale spaziale è esteso all'intero dominio di definizione del problema V .

Se i pesi sulle relazioni costitutive vengono scelti imponendo $\alpha_E = \alpha_H$, ricordando l'identità vettoriale:

$$(\mathbf{N}_r \cdot \nabla \times \mathbf{N}_s) - (\mathbf{N}_s \cdot \nabla \times \mathbf{N}_r) = \nabla \cdot (\mathbf{N}_s \times \mathbf{N}_r)$$

e che le funzioni di forma \mathbf{N}_r e \mathbf{N}_s non dipendono dal tempo, le espressioni (2a-b) diventano:

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \partial \mathbf{w}_s} = \alpha_E \int_{t_k}^{t_{k+1}} \frac{(t - t_k)}{\Delta t_k^2} dt \int_V [\nabla \cdot (\mathbf{N}_s \times \mathbf{N}_r)] d\mathbf{x} \quad (3a)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \partial \mathbf{a}_s} = \alpha_E \int_{t_k}^{t_{k+1}} \frac{(t - t_k)}{\Delta t_k^2} dt \int_V [\nabla \cdot (\mathbf{N}_r \times \mathbf{N}_s)] d\mathbf{x} \quad (3b)$$

Denotata con ∂V la frontiera di V , eseguendo l'integrale temporale ed applicando il teorema della divergenza alle espressioni (3a-b) si ottiene:

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{a}_r \mathbf{w}_s} = \frac{\alpha_E}{2} \int_{\partial V} (\mathbf{N}_s \times \mathbf{N}_r) \cdot \hat{\mathbf{n}} d\Sigma \quad (4a)$$

$$\frac{\partial^2 \Lambda_k}{\partial \mathbf{w}_r \mathbf{a}_s} = \frac{\alpha_E}{2} \int_{\partial V} (\mathbf{N}_r \times \mathbf{N}_s) \cdot \hat{\mathbf{n}} d\Sigma \quad (4b)$$

Facendo riferimento all'espressione (4a), per come sono definite le funzioni edge, l'integrale di superficie al secondo membro è nullo a meno che i lati r ed s non si trovino entrambi sulla frontiera ∂V ed inoltre appartengano alla stessa faccia di un elemento della mesh. In quest'ultimo caso, però, su quella faccia devono essere state assegnate le condizioni al contorno sulle componenti tangenti di \mathbf{A} o di \mathbf{W} e questo comporta che generalmente una delle due, \mathbf{a}_r oppure \mathbf{w}_s , non è un'incognita. Si supponga ad esempio che sia assegnata \mathbf{a}_r . L'imposizione delle condizioni al contorno attraverso il metodo di sostituzione, (equazioni (3.3-6)), comporta l'annullamento di tutti i termini misti dell'Hessiano che si trovano sulla riga r :

$$\frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{w}_m} = 0; \quad m=1, \dots, Nedges;$$

Ed in particolare di $\frac{\partial^2 \Lambda_{Mod,k}^{(j)}}{\partial \mathbf{a}_r \partial \mathbf{w}_s}$. Si è così dimostrato che la scelta dei pesi $\alpha_E = \alpha_H$,

congiunta all'adozione degli Edge-Elements per la discretizzazione spaziale e del metodo di sostituzione per l'imposizione delle condizioni al contorno, conduce al disaccoppiamento delle incognite \mathbf{a} e \mathbf{w} .

Appendice C
Imposizione delle condizioni di assorbimento in problemi monodimensionali

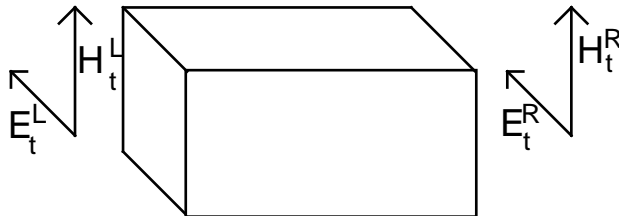


Fig [15] Imposizione delle condizioni di assorbimento

Ci si riferisca alla struttura mostrata in fig.[15]. Si assuma che i campi si propaghino lungo la direzione z e che il loro valore dipenda solo dalla coordinata z . Poichè il problema ha carattere monodimensionale, sulle superfici laterali del dominio (cioè sui quattro piani di frontiera $x=cost$ e $y=cost$) si possono imporre le condizioni di simmetria. Sulle basi devono essere imposte le condizioni di assorbimento, che hanno la seguente forma generale:

$$\begin{cases} a_{1L} \mathbf{E}_t^L + a_{2L} \mathbf{H}_t^L + a_{3L} = 0 \\ a_{1R} \mathbf{E}_t^R + a_{2R} \mathbf{H}_t^R + a_{3R} = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{H}_t^L = -\frac{a_{3L}}{a_{2L}} - \frac{a_{1L}}{a_{2L}} \mathbf{E}_t^L; & (a) \\ \mathbf{H}_t^R = -\frac{a_{3R}}{a_{2R}} - \frac{a_{1R}}{a_{2R}} \mathbf{E}_t^R; & (b) \end{cases} \quad (1)$$

dove i campi $(\mathbf{E}_t^L, \mathbf{H}_t^L), (\mathbf{E}_t^R, \mathbf{H}_t^R)$ sono incognite del problema.

Esiste, tuttavia, la possibilità di imporre condizioni miste con uno schema numerico efficace che conduce alla soluzione del problema attraverso quella di tre problemi simili a quello di interesse, ma con le condizioni al contorno miste sostituite da quelle di tipo standard elencate di seguito,

- | | |
|---|---|
| (I): $\mathbf{E}_{t,1}^{*L}=1;$ | $\mathbf{E}_{t,1}^{*R}=0$ |
| (II) $\mathbf{E}_{t,2}^{*L}=0;$ su ∂V^L | $\mathbf{E}_{t,2}^{*R}=1$ su ∂V^R |
| (III) $\mathbf{E}_{t,3}^{*L}=0;$ | $\mathbf{E}_{t,3}^{*R}=0$ |

Risolvendo i problemi (I), (II), (III) si ottengono sulle frontiere L e R rispettivamente i seguenti valori del campo magnetico \mathbf{H} :

$$\begin{array}{ll}
 \text{(I):} & \mathbf{H}_{t,1}^{*L}, & & \mathbf{H}_{t,1}^{*R} \\
 \text{(II)} & \mathbf{H}_{t,2}^{*L}, & \text{su } \partial V^L & \mathbf{H}_{t,2}^{*R} & \text{su } \partial V^R \\
 \text{(III)} & \mathbf{H}_{t,3}^{*L}, & & \mathbf{H}_{t,3}^{*R}
 \end{array}$$

I campi \mathbf{E} ed \mathbf{H} soluzioni sulle frontiere L ed R del problema originale si possono scrivere come combinazione lineare dei campi \mathbf{E} ed \mathbf{H} soluzioni dei problemi (i),(ii) e (iii) :

$$\mathbf{E}_t^L = c_1 \mathbf{E}_{t,1}^{*L} + c_2 \mathbf{E}_{t,2}^{*L} + \mathbf{E}_{t,3}^{*L} = c_1 \quad \text{su } \partial V^L \quad \mathbf{(2.a)}$$

$$\mathbf{E}_t^R = c_1 \mathbf{E}_{t,1}^{*R} + c_2 \mathbf{E}_{t,2}^{*R} + \mathbf{E}_{t,3}^{*R} = c_2 \quad \text{su } \partial V^R \quad \mathbf{(2.b)}$$

$$\mathbf{H}_t^L = c_1 \mathbf{H}_{t,1}^{*L} + c_2 \mathbf{H}_{t,2}^{*L} + \mathbf{H}_{t,3}^{*L} \quad \text{su } \partial V^L \quad \mathbf{(3.a)}$$

$$\mathbf{H}_t^R = c_1 \mathbf{H}_{t,1}^{*R} + c_2 \mathbf{H}_{t,2}^{*R} + \mathbf{H}_{t,3}^{*R} \quad \text{su } \partial V^R \quad \mathbf{(3.b)}$$

Sostituendo la (1.a) nella (2.b) e la (1.b) nella (3.b) si ottiene:

$$\left\{ \begin{array}{l}
 \mathbf{E}_t^L = -\frac{a_{3L}}{a_{1L}} - c_1 \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,1}^{*L} - c_2 \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,2}^{*L} - \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,3}^{*L}, \quad \mathbf{(4.a)} \\
 \mathbf{E}_t^R = -\frac{a_{3R}}{a_{2R}} - c_1 \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,1}^{*R} - c_2 \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,2}^{*R} - \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,3}^{*R}, \quad \mathbf{(4.b)}
 \end{array} \right.$$

Sostituendo la (4.a) nella (2.a) e la (4.b) nella (3.a) si ottiene il seguente sistema nelle incognite c_1 e c_2 :

$$\begin{bmatrix}
 1 + \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,1}^{*L} & \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,2}^{*L} \\
 \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,1}^{*R} & 1 + \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,2}^{*R}
 \end{bmatrix}
 \begin{bmatrix}
 c_1 \\
 c_2
 \end{bmatrix}
 =
 \begin{bmatrix}
 -\frac{a_{3L}}{a_{1L}} - \frac{a_{2L}}{a_{1L}} \mathbf{H}_{t,3}^{*L} \\
 -\frac{a_{3R}}{a_{2R}} - \frac{a_{2R}}{a_{1R}} \mathbf{H}_{t,3}^{*R}
 \end{bmatrix}$$

Calcolati così i coefficienti c_1 e c_2 è possibile attraverso le (2.a) e (2.b) calcolare i campi \mathbf{E}_t^L ed \mathbf{E}_t^R da imporre sulle frontiere ∂V^L e ∂V^R . La soluzione del problema originale si può, infine, trovare risolvendo un quarto problema di tipo standard in cui sulla frontiera, al posto delle condizioni di assorbimento, si impongono i valori dei campi \mathbf{E}_t^L ed \mathbf{E}_t^R .

Questo metodo sfrutta il principio di sovrapposizione degli effetti, ma può essere adoperato anche per problemi in cui siano presenti relazioni costitutive non lineari, a patto che queste siano linearizzate all'interno del time-step.

4.2) Magnetostatica. Caso test : Cubo percorso da corrente . Confronto con risultati precedenti.

Il primo esempio che viene riportato si riferisce al calcolo del campo magnetico bidimensionale prodotto da una corrente di densità uniforme parallela ad uno dei lati di un cubo di permeabilità uniforme. La lunghezza del lato sia $l=1$ m, mentre la permeabilità magnetica quella del vuoto; la densità di corrente sia fissata pari a $J=10^7$ A/m².

I risultati sono stati ottenuti usando due mesh: una più rozza costituita da due soli elementi per lato, e l'altra più raffinata con otto elementi per lato. I risultati numerici in termini di energia magnetica sono riportati in tavola 1 insieme con quelli ottenuti in [] dove si sono adoperate le formulazioni in **A** e in **T-Ω**.

Tavola 1

Calcolo del campo magnetostatico in un cubo
($l=1\text{m}$, $J_x=J_y=0$, $J_z=10\text{MA/m}^2$)

<i>Mesh</i>		
elementi per lato	2	8
nodi	27	729
elementi	8	512
lati	54	1944
<i>Gradi di libertà</i>		
formulazione in A []	5	833
formulazione in T-Ω []	54	1944
formulazione in A-W	108	3888
<i>Occupazione di memoria</i>		
form. in A [] (matrice bandata)	14	77914

form. in $\mathbf{T}-\Omega$ (matrice bandata)	257	119718
form. in $\mathbf{A}-\mathbf{W}$ (matrice sparsa)	510	27672
<i>energia magnetica (MJ)</i>		
Valore analitico	2.2082	2.2082
form. in \mathbf{A}	1.4726	2.1572
form. in $\mathbf{T}-\Omega$	2.6180	2.2380
form. in $\mathbf{A}-\mathbf{W}$ (int. volume $\mathbf{B}^2/2\mu$)	1.4726	2.1572
form. in $\mathbf{A}-\mathbf{W}$ (int. volume $\mu\mathbf{H}^2/2$)	2.6180	2.2381
form. in $\mathbf{A}-\mathbf{W}$ (int. volume $\mathbf{B}\cdot\mathbf{H}/2$)	1.4726	2.1572

[R.Albanese,G.Rubinacci,Magnetostatic field computations in terms of two-component vector potentials, Int. j. num.meth. eng.,29,515-532,1990

4.3) Eddy currents. Caso test: Felix-brick. Confronto con risultati precedenti.

Come esempio di applicazione per il problema delle eddy currents è stato scelto il cosiddetto Felix Brick, uno dei problemi usati come test da una associazione internazionale di ricercatori, che va sotto il nome di TEAM Workshop, per misurare le prestazioni di differenti codici numerici.

Il dominio del problema è costituito da un mattone di alluminio di forma rettangolare con un foro ugualmente rettangolare che lo attraversa completamente. Un campo magnetico uniforme, che decade esponenzialmente col tempo, è applicato perpendicolarmente alle facce forate. Il confronto tra i vari codici viene effettuato comparando la perdita totale di potenza per effetto ohmico, e il campo magnetico in varie posizioni.

Il materiale di cui è composto il mattone è una lega di alluminio con valore di resistività $\rho=3.94 \mu\Omega\text{cm}$. Le dimensioni del mattone sono $0.1524\text{m}\times 0.1016\text{m}\times 0.0508\text{m}$. Il foro rettangolare che attraversa completamente il mattone dalla parte della facce più larghe ha dimensioni $0.0889\text{m}\times 0.0381\text{m}$. Il campo magnetico esterno ha la sola componente nella direzione z ed è uniforme nello

spazio. Per $t > 0$, decade in maniera esponenziale nel tempo con la seguente legge:

$$B_z = B_0 e^{-t/\tau} \quad (1)$$

dove $B_0 = 0.1$ T e $\tau = 0.0119$ s. Questa condizione può essere imposta attraverso la scelta di opportune condizioni al contorno.

Grazie alle condizioni di simmetria, è necessario modellare solo un ottavo del mattone.

Per i calcoli effettuati con la formulazione proposta è stata adoperata la mesh consigliata nella referenza e mostrata nelle figure 1 e 2. Essa è composta da 1344 nodi, 1001 elementi esaedrali, 3656 lati e, pertanto, 7312 incognite per il metodo **A-W**. I valori computati per la potenza ohmica e per il campo indotto nel centro del buco ($x=y=z=0$) sono riportati in tabella 2, e sono in buon accordo con i risultati ottenuti attraverso le altre formulazioni (vedi Tabella 3).

È interessante riportare il valore incrementale di Λ in un time-step $\Delta t = 0.5$ ms (1.66 mJs a $t = 10.5$ ms, quando la potenza ohmica è massima), che è sufficientemente piccolo quando comparato al valore della figura di riferimento (2) che è pari a 373.91 mJs:

$$\Lambda_{\text{ref}} = \alpha_H \Delta t \int_V \mathbf{B} \cdot \mathbf{H} / 2 \, dV + \alpha_E \Delta t \int_V (\mathbf{E} \cdot \mathbf{D} / 2 + \mathbf{E} \cdot \mathbf{J} \Delta t) \, dV \quad (2)$$

Tavola 2

Un problema di eddy currents: il Felix brick
($t = 10.5$ ms, $\Delta t = 0.5$ ms)

Potenza Ohmica

form.in A-W (int. volume \mathbf{J}^2/σ)	106.98 W
form.in A-W (int. volume $\sigma \mathbf{E}^2$)	106.03 W
form.in A-W (int. volume $\mathbf{E} \cdot \mathbf{J}$)	105.53 W

Campo magnetico indotto in ($x=y=z=0$)

form.in $\mathbf{A-W}$ (\mathbf{B}_z)	0.04108 T
form.in $\mathbf{A-W}$ ($\mu\mathbf{H}_z$)	0.03907 T

[] **A.Kameari**. *Results for benchmark calculations of problem 4 (the Felix brick)*. COMPEL, vol.7, N.1 & 2, 65-80 (1988)

4.4) Propagazione. Caso test : Incidenza di un'onda su uno slab di materiale dielettrico lineare. Confronto con la soluzione analitica. Incidenza di un'onda su uno slab di materiale dalle caratteristiche non lineari.

Come ultimo esempio si è scelto un semplice problema di propagazione: l'incidenza di un'onda piana su una lastra di lunghezza indefinita nelle direzioni x e y e dello spessore di 20 metri. Si è assunto $\mu_r=1$, $\varepsilon_r=10$ e $\sigma=0$). Sebbene il problema sia di natura monodimensionale, è stato studiato come se fosse 3D usando una mesh consistente in una pila di elementi esaedrali lungo la direzione z (direzione di propagazione dell'onda). Sono state usate due discretizzazioni, ciascuna delle quali includeva 200 metri di spazio vuoto prima e 100 metri oltre la lastra. La mesh più rozza consiste di 128 elementi e 2056 incognite, mentre quella più raffinata di 320 elementi e 5128 incognite.

Si è supposto che sulla lastra incidesse un'onda piana il cui andamento temporale fosse rappresentato dalla semplice forma di un gradino. Si è assunto inoltre che il campo magnetico avesse ampiezza unitaria e fosse linearmente polarizzato lungo la direzione y . Questo caso, nonostante la semplicità della sua soluzione analitica è piuttosto critico dal punto di vista numerico a causa della brusca variazione del fronte d'onda. Le soluzioni ottenute attraverso formulazioni FEM di tipo tradizionale possono, infatti, mostrare accentuate oscillazioni che precedono e seguono il fronte d'onda.

Le seguenti condizioni al contorno sono state assunte: $\mathbf{E} \times \mathbf{n} = 0$, che implica $\mathbf{A} \times \mathbf{n} = 0$, sui due piani di simmetria a $x = \text{costante}$; $\mathbf{H} \times \mathbf{n} = 0$, che si traduce in $\mathbf{W} \times \mathbf{n} = 0$, sui piani di simmetria a $y = \text{costante}$; sulla frontiera $z = \text{cost}$ situata dietro la

lastra si è imposto, attraverso l'incognita \mathbf{W} , H_x uguale a zero e H_y essere un gradino unitario; ed infine, sul piano di frontiera $z=\text{cost}$ collocato oltre la lastra si è imposta la condizione $\mathbf{H} \times \mathbf{n} = 0$, ovvero $\mathbf{W} \times \mathbf{n} = 0$. In questo modo il problema dell'incidenza dell'onda piana sulla lastra dielettrica nello spazio libero è stato rappresentato correttamente, almeno fino al momento in cui uno dei fronti d'onda uscenti dalla lastra raggiunge le superfici di frontiera $z=\text{cost}$.

In figura, è mostrato il confronto tra i risultati numerici e la soluzione analitica. Al fronte d'onda delle soluzioni numeriche è vietato variare in maniera ripida dalle intrinseche proprietà di continuità sulle componenti tangenti delle edge functions attraverso le quali sono rappresentati gli andamenti spaziali dei potenziali vettori \mathbf{A} e \mathbf{W} , e conseguentemente i campi \mathbf{E} e \mathbf{H} . Tuttavia, come si può osservare il fenomeno di dispersione del fronte d'onda coinvolge un numero molto limitato di elementi (all'incirca 3).

Si sottolinea che la formulazione basata sull'uso congiunto dei potenziali \mathbf{A} e \mathbf{W} e degli edge elements non comporta oscillazioni nei campi in prossimità del fronte d'onda, ed in questo si dimostra superiore rispetto alle tradizionali soluzioni agli elementi finiti. Si può infine notare che il raffinamento della mesh produce una diminuzione del funzionale di errore Λ , che vale $0.22\text{E-}13$ Js per la mesh rozza e $0.19\text{E-}14$ nell'intervallo di tempo $(0,0.68 \mu\text{s})$. Sfortunatamente, nel caso dell'analisi nel dominio del tempo di un fenomeno transitorio di propagazione, l'informazione sulla localizzazione dell'errore massimo è poco significativa, perchè essa coincide ovviamente col fronte d'onda che si sta muovendo. Nondimeno, l'informazione può ridiventare preziosa se si adottano mesh mobili oppure nel caso dell'analisi nel dominio complesso di fenomeni propagativi pervenuti nella condizione di regime.

CAPITOLO V

Minimizzazione dell'errore costitutivo nella MHD e applicazioni ai modelli dinamici di plasma.

5.1) Generalita' sulla fusione nucleare. Plasma. Equazioni MHD. Problema aperti : equilibrio e evoluzione del confinamento del plasma in strutture tridimensionali quali i Tokamak . Halo currents, diffusione resistiva.

(Ref. Tesi Phd . Albanese,...)

5.2) Review delle metodologie e dei codici attualmente adoperati. Il codice Proteus. Limite della quasi assi-simmetria .

(Ref. Albanese-Rubinacci...)

5.3) Modello matematico per le equazioni MHD col metodo della minimizzazione dell'errore costitutivo.

CONCLUSIONI

L'approccio basato sulla minimizzazione dell'errore costitutivo è stato esteso allo studio di problemi tridimensionali generali di natura elettromagnetica. I principali vantaggi di questa estensione sono i seguenti:

a) la soluzione è ottenuta attraverso la minimizzazione di un funzionale di errore globale e, da questo, la matrice delle equazioni di sistema è simmetrica e definita positiva;

b) è facilmente disponibile una stima dell'errore numerico che si sta commettendo.

c) consente un trattamento unificato di problemi stazionari, quasi-stazionari e propagativi.

D'altra parte, le incognite sono in numero doppio rispetto alle classiche formulazioni in termini di **B** o di **E**. Tuttavia, attraverso una scelta conveniente dei pesi, è possibile ottenere lo splitting del modello discretizzato; in tal modo, ci si riconduce a dover risolvere, ad ogni passo nel tempo, due sistemi indipendenti di equazioni rispettivamente nelle incognite **A** e **W**.

Bibliografia

- [1] **Jackson**, *Elettrodinamica classica*, Zanichelli
- [2] *Mathematical Handbook for Scientists and Engineers*, Korn & Korn, p.10-1-2
- [3] **K.R.Richter, W.M.Rucker**, *Numerical foundations, integral methods and applications*, in "Industrial Application of Electromagnetic Computer Codes", (Crutzen, Molinari, Rubinacci), p.29-50, Kluwer Academic Publishers, **1990**.
- [4] **R.F.Harrington**. *Field computation by moment methods*. The Macmillan Company, New York, (**1968**)
- [5] **C.A.Brebbia, S.Walker**, *Boundary element techniques in Engineering*, Newnes-Butterworths, London (**1980**)
- [6] **W.H.Mc Donald, A.Wexler**, IEEE Trans.Microw.Th.Tech, MTT-20(**1972**) **841**
- [7] **R.Courant**, *Variational methods for the solutions of Problems of Equilibrium and Vibrations*. Bull Am.Math.Soc., 49:1-23, (**1943**)
- [8] **R.W.Clough**, *The Finite Element Method in Plane Stress Analysis*, Proc.2d ASCE Conf. Electronic Computation, Pittsburgh, PA., (**Sept.1960**)
- [9] **J.Simkin, C.W.Trowbridge**, *On the use of the total scalar potential in the numerical solution of field problems in electromagnets*, IJNME, vol.14, p.423, (**1979**)
- [10] **T.Morisue**, *Magnetic Vector Potential and Electric Scalar Potential in Three-Dimensional Eddy Current problem*, IEEE Trans.on Mag., Vol.18, p.531, (**1982**)
- [11] **C.F.Bryant, C.R.I.Emson, C.W.Trowbridge**, *A comparaision of Lorentz Gauge Formulations in Eddy Current Computations*, IEEE Trans.on Mag., Vol.26(2) (**1990**)
- [12] **D.Rodger, J.F.Eastham**, *A Formulation for Low Frequency Eddy Current Solutions*, IEEE Trans.on Mag., Vol.18, p.2443, (**1983**)
- [13] **A.Bossavit, J.C.Verite**, *A fixed FEM-BIEM method to solve 3D eddy current problems*, IEEE Trans.On Magn., Vol.18, p.431, (**1982**)
- [14] **R.Albanese, R.Martone, G.Miano, G.Rubinacci**, *A T Formulation for 3D FEM Eddy Current Computations*, IEEE Trans.on Mag., Vol.21, p.2299, (**1985**)
- [15] **R.Albanese, G.Rubinacci**, *Error estimation in a transient non linear 3D Eddy Current Problem (Team problem 10)*, TEAM Workshop, Miami, **1993**
- [16] **O.Biro, K.Preis**, *FE analysis of 3D eddy currents*, IEEE Trans.on Magn., Vol.MAG-26, p.418, (**1990**)
- [17] **C.F.Bryant, C.R.I.Emson, C.W.Trowbridge**, *A general purpose 3D formulation for eddy currents using the Lorentz gauge*, IEEE Trans.on Magn., Vol. MAG-26, p.2373, (**1990**)
- [18] **C.F.Bryant, B.Dillon, C.R.I.Emson, J.Simkin, C.W.Trowbridge**, *Solving high frequency problems using the magnetic vector potential with Lorentz gauge*, IEEE Trans.on Magn., Vol. MAG-28, p.1182, (**1992**)

- [19]**F.A.Fernandez,Y.Lu**, *A variational FE formulation for dielectric waveguides in terms of transverse magnetic fields*, IEEE Trans.on Magn., vol.27,pp.3864-3867, (1991)
- [20]**B.Davies,A.Fernandez,G.Y.Philippou**,*FE analysis of all modes in cavities with circular symmetry*,IEEE Trans.on Microwave Theory and Techn., Vol.30 pp.1975-1979,(1982)
- [21]**J.Webb**,*The FEM for finding modes of dielectric loaded cavities*,IEEE Trans. on Microwave Theory and Tech.,Vol.68-A,N.6,pp.533-540,(1985)
- [22]**I.Bardi,O.Biro,K.Preis,G.Vrisk,K.R.Richter**, *Nodal and edge element analysis of inhomogeneously loaded 3D cavities*, IEEE Trans.on Magn., Vol.28, pp.1142-1145, (1992)
- [23]**J.Webb**,*Efficient generation of divergence-free fields for the FE analysis of 3D cavity resonances*, IEEE Trans.on Magn.,Vol.24,pp.162-165 (1988)
- [24]**A.Bossavit**,*Solving Maxwell equations in a closed cavity, and the question of spurious modes*,IEEE Trans.On Magn.,Vol.26,pp.702-5,(1990)
- [25]**R.Albanese,G.Rubinacci**,*Two-component vector potential formulations for solving Maxwell Equations in closed cavities*.Rev.Roum.Sci.Techn.-Electrotechn.et Energ.,38,1, p.45-52, Bucarest (1993)
- [26]**P.Hammond**, *Energy Methods in Electromagnetism*, Clarendon Press, Oxford,(1981).
- [27]**J.Penman,J.R.Fraser**, '*Dual and complementary methods in electromagnetism*', IEEE Trans., MAG-19, 2311-2316,(1983).
- [28]**J.Rikabi,C.F.Bryant,E.M.Freeman**,'*An error-based approach to complementary formulations of static field solutions*', Int.j.numer.methods eng., 26, (1988)
- [29]**J.Rikabi,C.Bryant,E.Freeman**,*Error-based derivation of complementary formulations for the eddy-current problem*, IEE Proc. Vol.135, Pt.A, N.4, (1988) .
- [30]**R.Albanese,G.Rubinacci**,*Analysis of Three-dimensional Electromagnetic Fields Using Edge Elements*, Journal of Comp.Phys. 108,236-245 (1993)
- [31]**L.Pichon,A.Bossavit**, *A New Variational Formulation Free of Spurious Modes, for the Problem of Loaded Cavities*, IEEE Trans.on Magn.Vol.29, N.2, (marzo 1993)
- [32]**C.W.Trowbridge**, "*Overview of the 'State of the Art' in Electromagnetic analysis and design*", in "*Industrial Application of Electromagnetic Computer Codes*", (Crutzen,Molinari,Rubinacci), p.4, Kluwer Academic Publishers,1990.
- [33] Computer Physics Communications, Vol.68 (Special Issue),1991
- [34]**C.A.Desoer-E.S.Kuh**,*Fondamenti di teoria dei circuiti*,Franco Angeli,(1989)
- [35]**D.Burnett**, *Finite Element Analysis*, Addison Wesley 1987
- [36]**G.F.Carey-J.T.Oden**, *Finite elements vol.III*, Prentice-Hall 1984
- [37]**V.Comincioli**, *Analisi Numerica*, Mc Graw-Hill,1990
- [38]**V.Comincioli**,*Fortran 77.Introduzione e applicazioni numeriche*,Mc Graw-Hill,1990
- [39]**W.H.Press et al.**, *Numerical Recipes*, Cambridge University Press,1986

- [40]**D.M.Young**, *Iterative Solution of Large Linear Systems*, Academic Press, New York, **1971**
- [41]**M.R.Hestenes,E.L.Stieffel**,*Methods of Conjugate Gradients for solving Linear Systems*, NBS J.Res., 49, 409-436,(**1952**)
- [42]**A.Bossavit**,*Edge elements for scattering problems*,E.D.F.Bulletin de la direction des etudes et recherches serie C Mathematiques, Informatique n.2 pp.17-33 (**1989**)
- [43]**J.S.van Welij**,*Calculation of eddy currents in terms of H on hexaedra*, IEEE Trans.on magn. vol.21.n.6,(**novembre 1985**)
- [44]**P.Dular et al.**,*Calculation of 3D eddy currents using a coupling between edge finite elements, nodal finite elements and boundary elements.* (1993)
- [45]**A.Kameari**,*Calculation of transient 3D eddy currents using edge elements*, IEEE Trans.on magn. vol.26.n.2,(**marzo 1990**)
- [46]**T.Nakata, N.Takahashi,K.Fujiwara,Y.Shiraki**, *Comparaison of different finite elements for 3-D eddy current analysis*,IEEE Trans.on Magn.,Vol.26,n.2,(**1990**)
- [47]**R.Albanese,G.Rubinacci**,*Integral Formulation for 3D eddy-current computation using edge elements*, IEE Proc., Vol.135,Pt.A,n.7,(**1988**)
- [48]**P.Olszewski**,*Expansion of periodic boundary condition for 3-D FEM analysis using edge elements*,IEEE Trans.on Magn.,Vol.28,n.2,(**1992**)
- [49]**R.Albanese,G.Rubinacci**,*Magnetostatic field computations in terms of two-component vector potentials*, Int. j. num.meth. eng.,29,515-532,**1990**
- [50] **A.Kameari**.*Results for benchmark calculations of problem 4 (the Felix brick)*. COMPEL, vol.7, N.1 & 2, 65-80 (**1988**)
- [51]**I.S.Kim,W.J.R.Hoefer**, *The numerical energy conservation of the TD-FD method*, IEEE Trans.on Magn.,Vol.27,n.5,pp.4056-4059,(**1991**)
- [52]**R.Courant,K.O.Friedrichs,H.Lewy**, *On the partial difference Equations of Mathematical Physics*,(Traduzione inglese dall'originale apparso in Math.Ann.100 (**1928**), IBM Jour.Res.Devel.,n.11,pp.215-234 (**1967**)
- [53]**R.Albanese,R.Fresa,G.Miano,G.Rubinacci,L.Verolino**,*F.E.solution of nonlinear Maxwell Equations in the Time Domain*, 2° Workshop on Finite Element Method in Electromagnetic Wave Problems,(**1994**)
- [54]**I.G.Katayev**,*Electromagnetic Shock Waves*,London Ilife Books LTD (**1966**)
- [55]**C.W.Harrison,Jr.**,*Transient electromagnetic field propagation through infinite sheets,into spherical shells and into hollow cylinders*,IEEE Trans. Antennas and Propagation, vol. AP-12,pp.319-334 (**1964**)
- [56]**D.E.Merewether**,*Electromagnetic Pulse Transmission Through a Thin Sheet of Saturable Ferromagnetic Material of Infinite Surface Area*, IEEE Trans.On Elec.Comp., Vol.EMO-11,n.4, (**1969**)

- [57] **R.Luebbers, K.Kumagai, S.Adachi, T.Uno**, *FDTD Calculation of Transient Pulse Propagation through a Nonlinear Magnetic Sheet*, IEEE Trans.on Elec.Comp., Vol.35,n.1,1993
- [58] *Resources and Man. A study and recommendations by a Committee of the U.S. National Academy of Sciences*, Freeman, San Francisco (1969)
- [59] **R.Albanese, G.Rubinacci**, *Computational methods for the electromagnetic analysis in fusion devices* in Industrial Application of Electromagnetic Computer Codes, Kluwer Academic Publishers (1990)
- [60] **J.P.Freidberg**, *Ideal Magnetohydrodynamics*, Plenum Press, New York (1987)
- [61] **R.Albanese**, *Evoluzione temporale di un plasma confinato magneticamente*, PhD Thesis, (Dip.di Ing.Elettrica Un.di Napoli)(1986)
- [62] **J.Blum, J.Le Foll**, *Plasma Equilibrium Evolution at the Resistive Diffusion Time Scale*, Computer Physics Reports 1, pp.465-494, (1984)
- [63] **H.Grad, P.N.Hu, D.C.Stevens**, *Adiabatic evolution of plasma equilibrium*, Proc.US Nat.Acad.Science, 72, pp. 3789-3793, (1975)
- [64] **R.Albanese**, *Analysis of the Plasma Equilibrium Evolution in the Presence of Circuits and Massive Conducting Structures*, Proc. of XV SOFT Symp.on Fusion Tech., pp.281-286, Elsevier Science Publishers B.V., (1989)
- [65] **E.Coccorese, S.Coco, R.Martone**, *A model reduction technique for high order eddy current problems*, IEEE Trans. MAG-26, pp.690-693, (1990)
- [66] **R.Albanese, L.Bottura, S.Chiocchio, E.Coccorese, R.Fresa, R.Martone, A.Portone**, *Validation of Numerical Codes for the Analysis of Plasma Discharges and the Design of Fusion Reactors*, Fusion Eng.and Des.(1993)
- [67] **L.Bottura, R.Albanese, L.Boccaccini, R.Fresa et al.** *Eddy current codes validation against electromagnetic transients*, Proc. of 17th SOFT Symp.on Fusion Tech., pp.699-703, H.Knoepfel-North-Holland, (1993)
- [68] **A.Portone, R.Albanese, E.Coccorese, R.Fresa et al.** *Sensitivity of a Tokamak plasma simulation to some model uncertainties*, Proc. of 17th SOFT Symp.on Fusion Tech., pp.732-736, H.Knoepfel-North-Holland, (1993)
- [69] **S.Bobbio, R.Martone, G.Miano**, *Analysis of Vertical Instabilities occurring in JET*, JET REPORT, JET- IR(88)01, Abington, (1988)