



Consiglio Nazionale delle Ricerche

**ISTITUTO DI ELABORAZIONE
DELLA INFORMAZIONE**

PISA

**METODI ITERATIVI PER LA SOLUZIONE DI
PROBLEMI DI SCATTERING INVERSO**

EMANUELE SALERNO, ANNA TONAZZINI

Nota Interna B4 - 28
Giugno 1989

METODI ITERATIVI PER LA SOLUZIONE DI PROBLEMI DI SCATTERING INVERSO

EMANUELE SALERNO, ANNA TONAZZINI

IEI - CNR, Via S.Maria, 46

56126 Pisa

1. Introduzione

Saranno qui analizzati alcuni metodi, basati sulla minimizzazione iterativa di funzionali [4][5][7][10][11][19], applicati alla ricostruzione delle caratteristiche elettromagnetiche di un corpo dielettrico in base alla misura del campo elettromagnetico diffratto dal corpo stesso (problema dello scattering inverso [1] [6] [12][13] [18] [21] [22]). Il fenomeno dello scattering è stato inquadrato nello stesso ambito di validità considerato in [6] e [18], con le conseguenti equazioni fondamentali. In particolare, il campo elettromagnetico è considerato di tipo scalare, cioè non si tiene conto dei fenomeni di depolarizzazione che si verificano nell'iterazione tra un corpo e un'onda elettromagnetica. Il problema dello scattering inverso è di tipo non lineare, e l'applicazione di metodi di ottimizzazione per la sua soluzione nasce dall'esigenza di evitare le linearizzazioni che conducono a metodi di soluzione diretta (approssimazioni di Born e di Rytov [6] [20]), ma sono valide solo sotto ipotesi abbastanza restrittive sulle dimensioni e la natura materiale del corpo diffrangente.

I funzionali da minimizzare sono in ogni caso legati allo scarto quadratico tra il campo elettromagnetico misurato e quello calcolato in base alla stima corrente della soluzione. Per implementare algoritmi iterativi che convergano abbastanza rapidamente alla soluzione è necessario disporre dei valori delle "derivate" di tali funzionali; anche a questo scopo si potrebbe ricorrere a linearizzazioni del problema, ma l'algoritmo iterativo rischierebbe di diventare non convergente se il corpo fosse fortemente disadattato rispetto al mezzo in cui è immerso, altrimenti si ricadrebbe in un caso analogo a quello delle approssimazioni di Born o di Rytov [22]. In [12], [13] e, per condizioni particolari del campo elettromagnetico incidente, in [22], sono calcolate le espressioni esatte delle necessarie derivate (o *gradienti*), in base alle quali è possibile applicare metodi iterativi che assicurano la convergenza a punti stazionari dei funzionali obiettivo. In generale tali funzionali non sono di tipo convesso, quindi la soluzione non sarà unica, e ciò corrisponde a un risultato già giustificato teoricamente in [6], a meno che non si abbiano dati sufficienti a confinare la soluzione in un intorno della funzione desiderata.

L'ampiezza del suddetto intorno sarà in questo caso determinata dall'entità degli errori di misura, che costituiscono il *rumore* del sistema [22].

Nel prossimo paragrafo sarà sinteticamente presentato il modello matematico dello scattering utilizzato per tutti gli sviluppi della presente trattazione, mentre nel terzo paragrafo il problema dello scattering inverso sarà formulato come problema di ottimo (cioè di minimizzazione), mostrando il funzionale obiettivo utilizzato in [12] e [13] e, dopo le necessarie particolarizzazioni, in [22]. Il paragrafo 4 Mostra l'applicazione al problema della minimizzazione dei metodi dello steepest descent e del gradiente coniugato, con il calcolo esatto delle necessarie derivate. Questi metodi, se non garantiscono una elevata velocità di convergenza, hanno il grande vantaggio di essere molto stabili e di assicurare la convergenza indipendentemente dalla stima iniziale.

L'approccio qui presentato porta a una formulazione di applicabilità abbastanza generale che, dal punto di vista dei risultati teoricamente ottenibili, promette notevoli miglioramenti rispetto ai noti metodi diretti. Infatti, come si è già detto, si può evitare di ricorrere a linearizzazioni e, nello stesso tempo, si possono inglobare nel funzionale obiettivo e negli eventuali vincoli tutte le conoscenze a priori che si hanno sul modello e sulle statistiche degli errori di misura. Lo svantaggio fondamentale, di ordine pratico, di questi metodi, consiste nella enorme complessità computazionale. Infatti nel funzionale obiettivo sono sempre contenuti i dati sul campo diffratto calcolato in base alla stima corrente; come si vedrà al paragrafo 2 anche questo problema (dello *scattering diretto*), in generale, non è facilmente risolvibile in maniera esatta, se non impiegando elaborati algoritmi basati sulla discretizzazione della soluzione [9][17] oppure, a loro volta, di tipo iterativo [21][14]. Ogni iterata dei metodi qui descritti dunque presuppone ripetute soluzioni di problemi di scattering diretto e ciò complica tanto più le cose quanto più il corpo è esteso rispetto alla lunghezza d'onda della radiazione elettromagnetica incidente. In [13] sono presentati degli esempi di applicazione a problemi monodimensionali. Applicazioni pluridimensionali richiederanno senz'altro le enormi risorse di memoria e di velocità di calcolo di cui solo gli attuali supercalcolatori possono disporre.

2. Modello matematico dello scattering

Ogni onda (sia essa di tipo scalare, come le onde di pressione sonora, o vettoriale, come le onde elettromagnetiche) che si propaga in un mezzo in cui incontra degli ostacoli (cioè delle disomogeneità nelle caratteristiche propagative) subisce degli effetti che si indicano globalmente come fenomeni di *diffrazione*. In generale i fenomeni che hanno luogo possono essere studiati esattamente solo per mezzo delle equazioni delle onde, ovvero solo risolvendo le equazioni di Maxwell

per le date condizioni al contorno. Il problema generale della diffrazione (o *scattering*) assume un duplice aspetto a seconda delle incognite che si vogliono determinare. Si parlerà di scattering diretto quando, essendo nota l'onda incidente e le caratteristiche del mezzo in cui si propaga, si vuole determinare la perturbazione subita dall'onda stessa; si parlerà di scattering inverso quando è nota l'onda incidente e alcuni dati sulla perturbazione che subisce e si vogliono stimare le caratteristiche del mezzo di propagazione.

Si supponga di avere un corpo dielettrico, occupante un dominio compatto, con caratteristiche elettromagnetiche descritte dalla funzione reale *potenziale di scattering* [6] o *modello* [12] $v(\mathbf{r})$, identicamente nulla al di fuori di tale dominio (funzione a supporto compatto) e data da:

$$v(\mathbf{r}) = n^2(\mathbf{r}) - 1; \quad n^2 = \frac{\epsilon(\mathbf{r})}{\epsilon_0} \quad (2.1)$$

in cui ϵ è la permittività del corpo (funzione del punto) e ϵ_0 la permittività (costante) del mezzo in cui esso è immerso. Il corpo è illuminato da un campo elettromagnetico monocromatico polarizzato linearmente e originato da una sorgente puntiforme localizzata in \mathbf{r}_s .

A partire dalle equazioni vettoriali di Maxwell, se sono verificate certe condizioni, riassunte in [18] e corrispondenti essenzialmente all'isotropia e alla stazionarietà del mezzo di propagazione e all'assenza di carica libera, il sistema differenziale dalla cui soluzione si ricava il campo elettrico totale si presenta sotto forma di tre equazioni indipendenti, quindi, se il campo incidente è polarizzato linearmente, con una semplice rotazione degli assi coordinati ci si può ricondurre alla soluzione di un problema scalare. Come già detto nell'introduzione, si supporrà sempre che ciò possa essere fatto.

In ogni punto \mathbf{r} dello spazio il campo elettromagnetico totale sarà dato dalla somma tra il campo incidente noto e il *campo di scattering*, u^s , prodotto dall'interazione dell'onda incidente con il corpo diffrangente.

$$u(\mathbf{r}, \mathbf{r}_s; \nu) = u^i(\mathbf{r}, \mathbf{r}_s) + u^s(\mathbf{r}, \mathbf{r}_s; \nu) \quad (2.2),$$

in cui sono state esplicitate, oltre alla dipendenza dal punto dello spazio, anche quella dalla locazione \underline{r}_s della sorgente e quella dal modello v . Il campo u è legato al modello per mezzo della seguente equazione scalare di Helmholtz omogenea, in cui n ha il valore dato in (2.1) e $k^{(1)}$ è il numero d'onda della radiazione illuminante:

$$\nabla^2 u + n^2 k^2 u = 0 \quad (2.3 \text{ a});$$

inoltre il campo incidente u^i soddisfa alla seguente:

$$\nabla^2 u^i + k^2 u^i = 0 \quad (2.3 \text{ b}).$$

Sfruttando (2.2) e sottraendo membro a membro la (2.3 b) dalla (2.3 a) si ha:

$$\nabla^2 u^s + k^2 u^s = k^2 (1 - n^2) u = -k^2 v u \quad (2.4)$$

la cui soluzione, ricavata con il metodo della funzione di Green e calcolata in un punto \underline{r} , soddisfa anche all'equazione integrale:

$$u^s(\underline{r}, \underline{r}_s; v) = \frac{k^2}{4\pi} \int G(\underline{r}, \underline{r}') v(\underline{r}') u(\underline{r}', \underline{r}_s; v) d\underline{r}' \quad (2.5)$$

in cui G è la funzione di Green per il particolare problema trattato. Sostituendo la (2.5) nella (2.2) si ottiene:

$$u(\underline{r}, \underline{r}_s; v) = u^i(\underline{r}, \underline{r}_s) + \frac{k^2}{4\pi} \int G(\underline{r}, \underline{r}') v(\underline{r}') u(\underline{r}', \underline{r}_s; v) d\underline{r}' \quad (2.6).$$

Definiamo ora la funzione $\hat{u}(\underline{r}, \underline{r}; v)$ come la soluzione della seguente equazione:

$$\hat{u}(\underline{r}, \underline{r}; v) = G(\underline{r}, \underline{r}) + \frac{k^2}{4\pi} \int G(\underline{r}, \underline{r}') v(\underline{r}') \hat{u}(\underline{r}', \underline{r}; v) d\underline{r}' \quad (2.7).$$

Confrontando la (2.6) con la (2.7) si vede che sono due espressioni formalmente analoghe; ciò significa che la funzione \hat{u} può essere interpretata come un'onda che si origina nel punto \underline{r} (il campo incidente ha la forma della funzione di Green), viene diffratta dal corpo e si propaga verso il punto \underline{r} interno al corpo [13].

(1) Nel nostro caso (corpo dielettrico) si ha $k = 2\pi/\lambda$, in cui λ è la lunghezza d'onda della radiazione incidente nel mezzo di permittività uniforme ϵ_0 .

In [13] si dimostra che, dati due differenti modelli $v_1(\underline{r})$ e $v_2(\underline{r})$, si ha:

$$u^s(\underline{r}_r, \underline{r}_s; v_1) - u^s(\underline{r}_r, \underline{r}_s; v_2) = \int [v_1(\underline{r}') - v_2(\underline{r}')] \hat{u}(\underline{r}', \underline{r}_r; v_1) u(\underline{r}', \underline{r}_s; v_2) d\underline{r}' \quad (2.8);$$

sfruttando questa formula si può calcolare il gradiente dei dati u^s rispetto al modello v , che tornerà utile nel seguito, applicando la definizione di derivata di Gateaux [10]:

$$u^{s,v}(\underline{r}_r, \underline{r}_s; h) = \int h(\underline{r}') \hat{u}(\underline{r}', \underline{r}_r; v) u(\underline{r}', \underline{r}_s; v) d\underline{r}' \quad (2.9).$$

Detto $\underline{r}_0 = (\underline{r}_r, \underline{r}_s)$ il vettore formato dalle coordinate della sorgente e del punto di misura del campo, e posto

$$U(\underline{r}, \underline{r}_0; v) = \hat{u}(\underline{r}, \underline{r}_r; v) \cdot u(\underline{r}, \underline{r}_s; v) \quad (2.10)$$

la (2.9) può essere scritta nella seguente forma:

$$u^{s,v}(\underline{r}_0; h) = \int h(\underline{r}') U(\underline{r}', \underline{r}_0; v) d\underline{r}' \quad (2.11)$$

che ci mostra come la funzione U possa essere vista come gradiente di u^s nel punto v .

Supponiamo ora che il campo incidente, oltre a essere monocromatico, sia anche rappresentabile sotto forma di un'onda piana e uniforme, cioè, a meno della dipendenza temporale $e^{j\omega t}$,

$$u^i(\underline{r}) = e^{j\mathbf{k}^i \cdot \underline{r}} \quad (2.12);$$

\mathbf{k}^i è il vettore di propagazione incidente, che ha come modulo k e come versore quello della direzione di propagazione dell'onda incidente. supponiamo inoltre che il punto di misura \underline{r}_r si trovi in zona di campo lontano, cioè $|\underline{r}_r| \rightarrow \infty$. In questo caso, nello spazio tridimensionale, la (2.5) si trasforma nella seguente:

$$u^s(\underline{r}) \approx \frac{e^{jk|\underline{r}|}}{|\underline{r}|} g(\mathbf{k}^i, \mathbf{k}^s; v), \quad \mathbf{k}^s = \frac{k \underline{r}}{|\underline{r}|} \quad (2.13)$$

in cui l'ampiezza complessa di scattering, o matrice di scattering [6] $g(\mathbf{k}^i, \mathbf{k}^s; v)$ è data da:

$$g(\mathbf{k}^i, \mathbf{k}^s; v) = \frac{k^2}{4\pi} \int e^{-j\mathbf{k}^s \cdot \underline{r}} v(\underline{r}) u(\underline{r}) d\underline{r} \quad (2.14)$$

Il vettore \underline{k}^s ha modulo k e direzione coincidente con quella di scattering. In questo caso particolare le espressioni analoghe alle (2.6) e (2.7) sono le seguenti:

$$u(\underline{k}^i, v) = e^{j \underline{k}^i \cdot \underline{r}} + \frac{k^2}{4\pi} \frac{e^{jk|\underline{r}|}}{|\underline{r}|} \int e^{-j \underline{k}^s \cdot \underline{r}} v(\underline{r}) u(\underline{k}^i, v) d\underline{r} \quad (2.15)$$

$$u(-\underline{k}^s, v) = e^{-j \underline{k}^s \cdot \underline{r}} + \frac{k^2}{4\pi} \frac{e^{jk|\underline{r}|}}{|\underline{r}|} \int e^{-j \underline{k}^s \cdot \underline{r}} v(\underline{r}) u(-\underline{k}^s, v) d\underline{r} \quad (2.16).$$

In [22] si dimostra la validità di una espressione alle differenze finite, analoga alla (2.8), in base alla quale si ottiene il gradiente dei dati $g(\underline{k}^i, \underline{k}^s; v)$ rispetto al modello v . Le formule che ne derivano, analoghe alle (2.10) e (2.11), per il gradiente dei dati rispetto al modello e per la derivata di Gateaux di g in v , in direzione h , sono le seguenti:

$$G(\underline{k}^i, \underline{k}^s; v) = \frac{k^2}{4\pi} u(\underline{k}^i, v) u(-\underline{k}^s, v) \quad (2.17)$$

$$g'v(\underline{k}^i, \underline{k}^s; h) = \int G(\underline{k}^i, \underline{k}^s; v) h(\underline{r}) d\underline{r} \quad (2.18).$$

Il calcolo di u^s dalla conoscenza di v e di u^i costituisce la soluzione del problema dello scattering diretto. La soluzione della (2.5) o della (2.14) non è semplice, perché essa presuppone la conoscenza del campo totale u all'interno del corpo diffrangente, che invece in generale non è noto. Un mezzo per risolvere direttamente la (2.5) e la (2.14), nella restrittiva ipotesi in cui il corpo perturbi poco l'onda incidente, ossia risulti *debolmente diffrangente*, è quello di assumere il campo totale u all'interno del corpo uguale al campo incidente u^i (prima approssimazione di Born). Nella realtà questa ipotesi è difficilmente verificata e, nella maggior parte dei casi pratici, non risolve il problema elettromagnetico del calcolo del campo diffratto da un corpo.

Il problema inverso consiste nel ricavare la funzione v dato il campo incidente u^i e il campo diffratto u^s . Quando le misure sono fatte in zona lontana sono disponibili i valori della matrice di scattering per un insieme (discreto) di vettori di incidenza \underline{k}^i e vettori di scattering \underline{k}^s , per un singolo numero d'onda k o per un insieme di numeri d'onda (cioè per un insieme di frequenze); se invece non si è in zona lontana si hanno a disposizione, per ogni numero d'onda considerato, i valori del campo in un insieme discreto di punti dello spazio. Come si dimostra in [6], la soluzione del problema inverso non è, in generale, unica, sia che ci si ponga in zona lontana che in zona vicina, che si ammetta valida o meno

l'approssimazione di Born. Lo scattering inverso è quindi un problema *mal posto* nel senso di Hadamard [2]. Se si ammette valida l'approssimazione di Born la soluzione del problema può essere resa unica, per esempio, supponendo che il modello da ricostruire abbia trasformata di Fourier spaziale limitata in banda. Su questo principio si basano diversi metodi di inversione (detti appunto *di Born*) [18][20] le cui prestazioni teoriche sono state ampiamente saggiate, ma che non hanno trovato ancora applicazioni pratiche, sia per le difficoltà di realizzazione degli apparati di misura del campo elettromagnetico, sia perché le loro prestazioni si degradano rapidamente all'aumentare del disadattamento del corpo rispetto al mezzo circostante. Un'ulteriore problema è originato dal fatto che il campo diffratto non è noto esattamente, ma deriva da un insieme di misure e pertanto è affetto da un certo errore, qui chiamato *rumore del sistema*.

3. Problema inverso come problema di ottimo

Il fatto che il problema dello scattering inverso ammetta infinite soluzioni fa pensare a una perdita di informazione su v che ha luogo nel processo di scattering, descritto dalla (2.5). L'idea fondamentale per la regolarizzazione del problema, cioè per il recupero di una soluzione unica e stabile al rumore, è quella di reintrodurre l'informazione perduta sotto forma di conoscenze a priori sul modello v e sul rumore del sistema, la cui presenza non è neanche considerata negli schemi di inversione tipo Born. Un approccio seguendo il quale l'introduzione delle conoscenze a priori sul sistema risulta abbastanza immediata è quello probabilistico, in cui tali conoscenze sono espresse come densità di probabilità sulla soluzione e sul rumore. La soluzione del nostro problema potrà essere vista come quella funzione v la cui probabilità, condizionata alla disponibilità di un certo insieme di dati, è resa massima, cioè

$$v : p(v | d) \geq p(x | d) \quad (3.1)$$

per ogni altra stima x del modello. Nella (3.1) la funzione d rappresenta i dati del problema, cioè i valori del campo diffratto misurato in un certo insieme di punti dello spazio \mathbf{r} , che, se si indica con $V(\mathbf{r})$ il modello reale e con $\varepsilon(\mathbf{r})$ il rumore del sistema, sono dati da:

$$d(\mathbf{r}, \mathbf{I}_s, V) = u^s(\mathbf{r}, \mathbf{I}_s, V) + \varepsilon(\mathbf{r}) \quad (3.2)$$

in cui il campo diffratto è dato dall'espressione (2.5). La *Probabilità a posteriori* $p(v | d)$, per il noto teorema di Bayes, può essere espressa come:

$$p(v | d) = \text{costante} \cdot p(d | v) p(v) \quad (3.3)$$

dove $p(v)$ è la *probabilità a priori* della funzione v ed esprime le conoscenze sulla soluzione; $p(d|v)$ esprime, nota una trasformazione tipo (2.5), le conoscenze sul rumore del sistema. Supponiamo che questo sia additivo e gaussiano a media nulla e che la funzione v sia anch'essa distribuita gaussianamente attorno a una sua stima *a priori* $v'(\underline{r})$. Le densità di probabilità di ϵ e di v saranno rispettivamente:

$$p(\epsilon) = \text{costante} \cdot \exp\left\{-\frac{1}{2} \iint \epsilon(\underline{r}) W(\underline{r}, \underline{r}') \epsilon^*(\underline{r}') d\underline{r} d\underline{r}'\right\} \quad (3.4 a)$$

$$p(v) = \text{costante} \cdot \exp\left\{-\frac{1}{2} \iint [v(\underline{r}) - v'(\underline{r})] W(\underline{r}, \underline{r}') [v(\underline{r}') - v'(\underline{r}')] d\underline{r} d\underline{r}'\right\}$$

(3.4 b).

Posti:

$$e(\underline{r}, \underline{r}_s, v) = d(\underline{r}, \underline{r}_s, V) - u^s(\underline{r}, \underline{r}_s, v) \quad (3.5 a);$$

$$e'(\underline{r}, v) = v(\underline{r}) - v'(\underline{r}) \quad (3.5 b),$$

la densità di probabilità della variabile d , data una certa v , sarà:

$$p(d|v) = \text{costante} \cdot \exp\left\{-\frac{1}{2} \iint e(\underline{r}_o, v) W(\underline{r}_o, \underline{r}_o') e^*(\underline{r}_o', v) d\underline{r}_o d\underline{r}_o'\right\} \quad (3.6)$$

avendo indicato, come al paragrafo precedente, con \underline{r}_o l'insieme dei due vettori \underline{r} e \underline{r}_s ; per la (3.6), la (3.3) si trasforma nella seguente:

$$p(v|d) = \text{cost.} \cdot \exp\left\{-\frac{1}{2} \left[\iint e(\underline{r}_o, v) W(\underline{r}_o, \underline{r}_o') e^*(\underline{r}_o', v) d\underline{r}_o d\underline{r}_o' + \iint e'(\underline{r}, v) W'(\underline{r}, \underline{r}') e'(\underline{r}', v) d\underline{r} d\underline{r}' \right] \right\} \quad (3.7).$$

Massimizzare l'espressione (3.7) in v è equivalente a minimizzare

$$E(v) = \iint e(\underline{r}_o, v) W(\underline{r}_o, \underline{r}_o') e^*(\underline{r}_o', v) d\underline{r}_o d\underline{r}_o' + \iint e'(\underline{r}, v) W'(\underline{r}, \underline{r}') e'(\underline{r}', v) d\underline{r} d\underline{r}'$$

(3.8)

in cui le funzioni W e W' sono gli analoghi continui degli inversi delle matrici di correlazione per il corrispondente problema discreto, sia nei dati che nella soluzione. La funzione W esprime la correlazione tra i valori $\epsilon(\underline{r})$ e $\epsilon(\underline{r}')$. Se il

rumore del sistema è considerato stazionario e bianco la W assume una forma "diagonale":

$$W(\mathbf{I}_0, \mathbf{I}_0') = W(\mathbf{I}_0) \delta(\mathbf{I}_0 - \mathbf{I}_0') \quad (3.9);$$

in questo caso $W(\mathbf{I}_0)$ è l'inverso dell'autocovarianza del rumore nel punto \mathbf{I}_0 . W può essere interpretata anche come una funzione *peso*, che attribuisce maggiore o minore importanza al valore dello scarto in punti diversi dipendentemente dalla maggiore o minore affidabilità della misura. Analogamente la W' esprime la correlazione tra i valori di v in punti diversi e la sua non diagonalità starebbe a indicare che *a priori* si sa che il modello non presenta brusche variazioni con \mathbf{r} . Nel caso in cui si fa assumere anche alla W' forma diagonale:

$$W'(\mathbf{r}, \mathbf{r}') = W'(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \quad (3.10),$$

il suo valore nel punto \mathbf{r} può essere interpretato come il grado di "fiducia" che, in quel punto, si attribuisce alla stima a priori del modello. Se si assumono valide le (3.9) e (3.10), i valori delle funzioni W e W' possono essere scelti in maniera tale da regolare l'influenza relativa delle misure e della stima a priori sul valore del funzionale $E(v)$. Il valore di v che minimizza $E(v)$ rappresenterà quindi una soluzione di compromesso che sia sufficientemente consistente con i dati e, nello stesso tempo, non troppo lontana da una stima predeterminata [13]. In questo caso la (3.8) si trasforma come segue:

$$E(v) = \int W(\mathbf{I}) |e(\mathbf{I}_0)|^2 d\mathbf{I}_0 + \int W'(\mathbf{r}) [e'(\mathbf{r})]^2 d\mathbf{r} \quad (3.11).$$

Si vede quindi che, nell'ipotesi di rumore e soluzione gaussiani, e valendo le (3.9) e (3.10), il funzionale da minimizzare nell'approccio *massima probabilità a posteriori* (MAP) risulta formalmente identico a quello derivato dall'approccio deterministico *minimo errore quadratico medio*.

Il criterio individuato dalla (3.11) può essere applicato alla determinazione della soluzione del problema dello scattering inverso nella sua formulazione più generale; se poi si assumono valide la (2.12) e le (2.13), la (3.11) rimane formalmente identica, salvo il fatto che in questo caso sarà:

$$d(\mathbf{k}^i, \mathbf{k}^s; V) = g(\mathbf{k}^i, \mathbf{k}^s; V) + \varepsilon(\mathbf{k}^i, \mathbf{k}^s) \quad (3.12 a)$$

$$e(\mathbf{k}^i, \mathbf{k}^s; v) = d(\mathbf{k}^i, \mathbf{k}^s; V) - g(\mathbf{k}^i, \mathbf{k}^s; v) \quad (3.12 b)$$

e quindi la variabile \mathbf{L}_0 nel primo integrale in (3.11) si trasformerà in

$$\mathbf{L}_0 = (\mathbf{k}^i, \mathbf{k}^s) \quad (3.13)$$

mentre il secondo integrale rimarrà invariato.

Resta ora il problema di minimizzare la (3.11). Data la non linearità del problema la via seguita per la soluzione sarà quella iterativa a partire da una stima iniziale. A tale scopo possono essere usati metodi di tipo Newton, che hanno il vantaggio di convergere abbastanza rapidamente; tuttavia tali metodi sono molto sensibili alla stima iniziale, risultano spesso instabili in dipendenza dal particolare problema, e richiedono la conoscenza della derivata seconda del funzionale [13]. Seguendo [12] [13] [22], in questo lavoro verranno utilizzati i metodi detti di *descent*, in particolare quelli dello steepest descent e del gradiente coniugato. Tali algoritmi, pur essendo a convergenza più lenta rispetto a quelli di tipo Newton, sono più stabili, meno sensibili alla stima iniziale e soprattutto richiedono la conoscenza della sola derivata prima del funzionale. Tutti i metodi iterativi si basano sull'aggiornamento di una stima della soluzione, secondo lo schema generale;

$$\mathbf{v}_n = \mathbf{v}_{n-1} + a_n \mathbf{p}_n \quad (3.14)$$

in cui la funzione \mathbf{p}_n è la *direzione* lungo la quale ci si sposta partendo da \mathbf{v}_{n-1} e il numero a_n è l'ampiezza dello spostamento. Nei metodi di descent \mathbf{p}_n dipende dalle derivate del funzionale obiettivo. Nel prossimo paragrafo mostreremo le espressioni esatte dei gradienti e quindi le formule iterative per la minimizzazione nel nostro caso particolare.

4. Procedura iterativa di minimizzazione

Si è accennato, al paragrafo precedente, al fatto che nei metodi di descent la direzione di incremento ad ogni iterata è legata al gradiente del funzionale obiettivo. Ciò trova la sua giustificazione intuitiva nel fatto che, dovendo minimizzare il funzionale, occorre conoscere le direzioni lungo le quali questo decresce. Nel nostro caso calcoliamo la derivata di Gateaux del funzionale E in (3.11):

$$E'v(h) = 2\Re \int W(\mathbf{L}_0) e(\mathbf{L}_0; v) [u^s v(h; \mathbf{L}_0)]^* d\mathbf{r}_0 + 2 \int W'(\mathbf{r}) e'(\mathbf{r}; v) h(\mathbf{r}) d\mathbf{r} \quad (4.1).$$

La difficoltà sta a questo punto nel calcolo della derivata di Gateaux $u^s v$. In [13] tale difficoltà è superata facendo uso della (2.11), che trasforma la (4.1) nella seguente:

$$E'v(h) = 2 \int h(\mathbf{r}) g(\mathbf{r}; v) d\mathbf{r} \quad (4.2)$$

con

$$g(\mathbf{r}; v) = \mathfrak{R} \int W(\mathbf{r}_0) e(\mathbf{r}_0; v) U^*(\mathbf{r}, \mathbf{r}_0; v) d\mathbf{r}_0 + W'(\mathbf{r}) e'(\mathbf{r}; v) \quad (4.3).$$

Vista la struttura della (4.2) la funzione $g(\mathbf{r}, v)$ può essere considerata il gradiente di E .

Definiamo ora, in base alle (3.5) e alla (4.3),

$$e_n(\mathbf{r}_0) = e(\mathbf{r}_0; v_n) \quad (4.4 a)$$

$$e'_n(\mathbf{r}) = e'(\mathbf{r}; v_n) \quad (4.4 b)$$

$$g_n(\mathbf{r}) = g(\mathbf{r}; v_n) \quad (4.4 c)$$

essendo v_n la stima n -esima del modello v . Considerando adesso lo schema iterativo (3.14), che qui riportiamo per comodità

$$v_n(\mathbf{r}) = v_{n-1}(\mathbf{r}) + a_n p_n(\mathbf{r}) \quad (4.5),$$

il valore dello scalare a_n e della funzione $p_n(\mathbf{r})$ devono essere determinati in maniera tale che, per ogni n , risulti $E(v_n) < E(v_{n-1})$. Supponendo di aver determinato la funzione p_n , il valore di a_n sarà scelto in modo da minimizzare il funzionale E lungo la direzione individuata da p_n , cioè in modo da minimizzare la seguente funzione della variabile reale a :

$$f(a) = E(v_{n-1} + a p_n) \quad (4.6).$$

Questo risultato può essere sempre raggiunto per mezzo di una ricerca numerica, che tuttavia risulta computazionalmente dispendiosa, richiedendo ripetute valutazioni del funzionale E. Nel caso in cui il funzionale è approssimativamente quadratico, e ciò avviene sempre nelle vicinanze di un minimo [5] [11], a_n può essere valutata analiticamente con errore trascurabile. Nel nostro caso l'espressione appropriata, ricavata in [13] è la seguente:

$$a_n = \frac{P_n}{Q_n} \quad (4.7)$$

con

$$P_n = - \int P_n(\underline{x}) g_{n-1}(\underline{x}) d\underline{x} \quad (4.8 a)$$

$$Q_n = \iint P_n(\underline{x}) M_n(\underline{x}, \underline{y}) P_n(\underline{y}) d\underline{x} d\underline{y} \quad (4.8 b)$$

$$M_n(\underline{x}, \underline{y}) = \int W(\underline{r}_o) U(\underline{x}, \underline{r}_o; v_{n-1}) U^*(\underline{y}, \underline{r}_o; v_{n-1}) d\underline{r}_o + W'(\underline{x}) \delta(\underline{x} - \underline{y}) \quad (4.8 c);$$

in [13] si dimostra che Q_n è positivo. Con questa scelta di a_n risulta

$$E(v_n) = E(v_{n-1}) - \frac{P_n^2}{Q_n} \quad (4.9).$$

Il valore $E(v_n)$ sarà minore di $E(v_{n-1})$ per ogni scelta di p_n che non annulli P_n . La scelta di p_n sarà legata al particolare metodo di descent adottato.

Una possibile scelta, che massimizza il valore assoluto di P_n nella (4.8 a), è la seguente:

$$p_n(\underline{r}) = -g_{n-1}(\underline{r}) \quad (4.10)$$

e corrisponde al metodo della massima decrescenza (o steepest descent). Non si possono fare, in generale, considerazioni sulla velocità di convergenza di questo metodo. Nel caso in cui il funzionale da minimizzare è di tipo quadratico si può dimostrare [10] che la convergenza è lineare. Un tale tipo di convergenza sconsiglierebbe l'uso del metodo di steepest descent. Il fatto però che questo metodo risulti sempre stabile, e solo raramente [10] non convergente ad un minimo almeno locale, fa sì che esso costituisca sempre la base per algoritmi iterativi con ordine di convergenza superiore [19]. Uno di tali metodi è il metodo del gradiente coniugato, che corrisponde a una scelta di p_n tale da ottenere un

valore di P_n uguale al caso dello steepest descent, minimizzando contemporaneamente il valore di Q_n . In tal modo si vede, anche intuitivamente, che la convergenza del metodo risulta migliore, perché (vedere la (4.9)) lo scarto tra $E(v_n)$ e $E(v_{n-1})$ è maggiore che nel caso precedente. Si dimostra che in questo caso la scelta opportuna per p_n è la seguente:

$$P_1 = -g_0 \quad (4.11 \text{ a})$$

$$P_n = -g_{n-1} + b_n P_{n-1} \quad \text{per } n > 1 \quad (4.11 \text{ b})$$

con

$$b_n = P_n / P_{n-1} \quad (4.11 \text{ c}).$$

Le (4.11) rappresentano la versione continua dell'algoritmo del gradiente coniugato, nell'implementazione di Fletcher e Reeves [13]. Il metodo del gradiente coniugato presenta il vantaggio che la formula per la determinazione delle direzioni di ricerca, particolarmente semplice, lo rende poco più complicato del metodo dello steepest descent; inoltre, essendo le direzioni di ricerca basate sui gradienti, la convergenza verso la soluzione è abbastanza uniforme per tutte le iterazioni, a differenza di altri metodi di descent per i quali le direzioni di ricerca sono predeterminate.

5. Conclusioni

In questo lavoro è stata analizzata una soluzione del problema mal posto dello scattering inverso, basata sulla minimizzazione di un funzionale costo che incorpora conoscenze a priori sulla soluzione stessa e sulle statistiche del rumore di misura. In particolare tale funzionale costo rappresenta la somma tra l'errore quadratico medio tra i dati misurati e quelli ottenuti in base alla stima della soluzione, e l'errore quadratico medio tra la stima della soluzione e una stima a priori ragionevole. Il problema di ottimo che ne risulta, altamente non lineare, viene risolto con le due tecniche di descent iterativo dette dello steepest descent e del gradiente coniugato. Un tale approccio presenta tra l'altro, rispetto ai metodi diretti, il vantaggio di non dover ricorrere a linearizzazioni. Per contro la complessità numerica degli algoritmi risulta certamente molto aumentata.

Gli schemi iterativi sono derivati con riferimento alla forma dello specifico funzionale costo che deve essere minimizzato. Tali derivazioni risultano matematicamente rigorose nell'ipotesi che il funzionale costo e la soluzione siano definiti nello spazio L^2 delle funzioni a quadrato sommabile. E' tuttavia di interesse far notare qui che il problema dello scattering inverso si posiziona in

modo più naturale nello spazio $C[D]$ delle funzioni continue su un certo supporto compatto D . Occorrerebbe allora poter dimostrare la validità dei metodi di descent sopra citati anche nell'ipotesi più generale di funzionali definiti su spazi di Banach o, quanto meno, sul particolare spazio $C[D]$. Tale problema è attualmente allo studio.

BIBLIOGRAFIA

- [1] M.Bertero, C.DeMol, G.A.Viano: "The Stability of Inverse Problems", in *Inverse Scattering Problems in Optics*, Springer-Verlag, 1980.
- [2] M.Bertero: "Linear Inverse and Ill-Posed Problems", Istituto Nazionale di Fisica Nucleare, Genova, INFN/TC-88/2, 19 gennaio 1988.
- [3] R.F.Curtain, A.J.Pritchard: "Functional Analysis in Modern Applied Mathematics", Academic Press, 1977.
- [4] J.W.Daniel: "The Approximate Minimization of Functionals", Prentice Hall, Englewood Cliffs, NJ, 1971
- [5] V.F.Demyanov, A.M.Rubinov: "Approximate Methods in Optimization Problems", American Elsevier, New York, 1970.
- [6] A.J.Devaney: "Nonuniqueness in the Inverse Scattering Problem", *J.Math. Phys.*, Vol. 19, no. 7, 1978, pp. 1526-1531.
- [7] R.Fletcher, C.M.Reeves: "Function Minimization by Conjugate Gradients", *Computer J.*, Vol. 7, no. 2, 1964, pp. 149-154.
- [8] D.H.Griffel: "Applied Functional Analysis", Ellis Horwood Ltd., Chichester, UK, 1981.
- [9] R.F.Harrington: "Field Computation by Moment Methods", MacMillan, New York, 1968.
- [10] D.G.Luenberger: "Optimization by Vector Space Methods", Wiley, 1969.
- [11] D.G.Luenberger: "Linear and Nonlinear Programming", Addison-Wesley, 1984.
- [12] S.J.Norton, J.A.Simmons, A.H.Kahn, H.N.G.Wadley: "Research on Inverse Problems in Materials Science and Engineering", in *Signal Processing and Pattern Recognition in Nondestructive Evaluation of Materials*, NATO ASI Series, F: Computer and Systems Sciences, Vol. 44, Springer Verlag, 1988, C.H.Chen editor, pp. 1-21.
- [13] S.J.Norton: "Iterative Seismic Inversion", *Geophys. J. Roy. Astr. Soc.*, Vol. 94, no. 3, 1988, pp. 457-468.
- [14] T.J.Peters, J.L.Volakis: "Application of a Conjugate Gradient FFT Method to Scattering from Thin Planar Material Plates", *IEEE Trans.*, Vol. AP-36, no. 4, April 1988, pp. 518-526.
- [15] T.J.Peters, J.L.Volakis: "Computation of the Scattering by Planar and Non-Planar Plates Using a Conjugate Gradient FFT Method", Radiation Laboratory, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Harbor, MI, 389604-4-T, Sept., 1988.

- [16] G.Prodi, A.Ambrosetti: "Analisi non Lineare", Scuola Normale Superiore, Pisa, 1973.
- [17] J.H.Richmond: "Scattering by a Dielectric Cylinder of Arbitrary Cross-Section Shape", IEEE Trans., Vol. AP-13, no. 3, May 1965, pp. 334 - 341.
- [18] E.Salerno: "Caratterizzazione del Fenomeno dello Scattering con Particolare Riferimento al Problema Inverso ed Applicazioni all'Imaging Diffrattivo", IEI-CNR, Pisa, N.I.B4-04, Feb. 1988.
- [19] L.E.Scales: "Introduction to Non-Linear Optimization", MacMillan, 1985.
- [20] C.F.Schueler, H.Lee, G.Wade: "Fundamentals of Digital Ultrasonic Imaging", IEEE Trans., Vol. SU-31, no. 4, July 1984, pp. 195 - 217.
- [21] P.M.VanDenBerg: "Iterative Computational Techniques in Scattering Based Upon the Integrated Square Error Criterion", IEEE Trans., Vol. AP-32, no. 10, Oct. 1984, pp. 1063-1071.
- [22] V.H.Weston: "Nonlinear Approach to Inverse Scattering", J. Math. Phys., Vol. 20, 1979, pp. 53-59.