Bollettino
Unione Matematica Italiana

Sezione A – La Matematica nella Società e nella Cultura

Mario Bertero

Matematica e immagini: alcuni esempi di applicazioni

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **2-A**—*La Matematica nella Società e nella Cultura* (1999), n.1, p. 47–67.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_1999_8_2A_1_47_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 1999.

Bollettino U. M. I.
La matematica nella Società e nella Cultura
(8) 2-A (1999), pag. 47-67

Matematica e immagini: alcuni esempi di applicazioni.

MARIO BERTERO

1. – Introduzione.

La nostra vita quotidiana è sempre più invasa da immagini di ogni tipo, ottenute con le tecniche più svariate. Quelle di cui si parla in questa nota richiedono l'uso della Matematica per essere generate, pur non essendo immagini completamente create al calcolatore come, ad esempio, quelle della realtà virtuale o della progettazione assistita da calcolatore (CAD, cioè *computer assisted design*), che costituiscono l'oggetto della *computer graphycs*. Sono infatti immagini ottenute elaborando, mediante sofisticati metodi ed algoritmi matematici, dati che sono prodotti da altrettanto sofisticati strumenti fisici e che non sono direttamente interpretabili. Mi riferisco, ad esempio, alle immagini dell'interno del corpo umano ottenute mediante la tomografia a raggi X o la risonanza magnetica, oppure alle immagini della struttura interna della Terra ottenute a partire da dati sismici (la cosiddetta tomografia sismica) od ancora alle immagini di oggetti estremamente lontani (astronomia) o di oggetti estremamente piccoli (microscopia), ottenute elaborando dati di complessi strumenti ottici. Per essere più precisi, mi concentrerò su due tipi particolari di immagini, apparentemente molto diverse tra di loro: quelle biomediche, che vengono usate nella diagnostica medica per immagini, e quelle astronomiche. L'elemento unificante è che in entrambi i casi occorre risolvere problemi matematici appartenenti alla stessa classe, quella dei *problemi mal posti*.

L'individuazione di questa classe è dovuta al grande matematico francese Jacques Hadamard e risale all'inizio del secolo. Studiando problemi al contorno per equazioni alle derivate parziali, Hadamard osserva che per alcuni di essi sono garantite le condizioni di esistenza, unicità e dipendenza continua della soluzione dai dati, mentre per altri l'una o l'altra di tali condizioni viene a mancare. Non è qui il caso di precisare più esattamente sotto quali condizioni si verificano tali circostanze. Una lettura della lucida opera originale è caldamente consigliabile (Hadamard, 1923), anche perchè permette di meglio comprendere il pensiero di Hadamard. Guidato da una visione deterministica ed essenzialmente ottocentesca della realtà, egli ipotizza che i problemi matematici corrispondenti ad una descrizione dei fenomeni fisici debbano godere di tutti e tre i requisiti sopra elencati: tali problemi sono da lui chiamati *ben posti* e di conseguenza sono detti *mal posti* tutti gli altri.

Per un lungo periodo i problemi mal posti sono considerati una curiosità matematica e pertanto il loro studio è trascurato. Solo negli anni cinquanta, grazie a lavori ed osservazioni di Fritz John, Carlo Pucci e Andrei N. Tikhonov, si incomincia a riconoscere che tali problemi si presentano in modo del tutto naturale nell'analisi di dati forniti da strumentazione fisica e più precisamente in tutti quei casi in cui si vuole misurare in modo indiretto una grandezza fisica a cui non si può accedere direttamente. Un esempio tipico è la tomografia a raggi X, che sarà discussa tra breve.

Problemi del tipo sopra indicato, che attualmente sono chiamati *problemi inversi*, sono solitamente mal posti ed il motivo di ciò risiede nel fatto che la strumentazione usata non fornisce un'informazione completa sulla grandezza fisica che si vuol determinare. Nel caso in cui il processo di misurazione può essere descritto mediante un modello lineare, nella versione discreta del problema sia la grandezza fisica da determinare che quella misurata sono rappresentate da vettori e il vettore di questa può essere ottenuto dal vettore di quella applicandovi un'opportuna matrice che descrive il processo di misurazione. Pertanto la risoluzione del problema inverso viene ricondotta alla risoluzione di un sistema algebrico lineare, in genere di grandi dimensioni (casi con milioni di dati sono oramai frequenti nelle applicazioni). Supponiamo per semplicità che la matrice sia simmetrica e quindi diagonalizzabile. Se la strumentazione fisica descritta da essa non fornisce un'informazione completa sulla grandezza incognita, ciò significa che la matrice ha un qualche autovalore nullo, e quindi la soluzione non è unica, oppure autovalori molto piccoli, ed in tal caso la soluzione è numericamente instabile (un piccolo errore sui dati produce un grande errore sulla soluzione): dunque il problema è mal posto.

Metodi generali per trattare questo particolare tipo di problemi sono introdotti, all'inizio degli anni sessanta, dal matematico russo Andrei N. Tikhonov e sono da lui chiamati *metodi di regolarizzazione*. L'idea base è sostanzialmente quella di cercare una soluzione che non riproduca esattamente i dati del problema (questa è numericamente instabile) ma solo approssimativamente, entro gli errori di misura, in modo da soddisfare a vincoli aggiuntivi, dedotti da proprietà della soluzione note *a priori*. Tutto ciò può essere solitamente formulato in termini di opportuni problemi variazionali. Non entreremo nei dettagli di questi metodi di regolarizzazione. Il lettore interessato può ricorrere ad alcuni libri che sono stati pubblicati sull'argomento e che forniscono rielaborazioni ed aggiornamenti del libro originario di Tikhonov (Tikhonov e Arsenin, 1977). Una versione piuttosto elementare con applicazioni a problemi di immagini è presentata in (Bertero e Boccacci, 1998), mentre una versione molto più completa e di livello matematico più elevato è esposta in (Engl e altri, 1996).

2. – Immagini biomediche e trasformata di Radon.

Agli inizi degli anni settanta l'invenzione della tomografia a raggi X, dovuta a G. H. Hounsfield, un ingegnere della EMI, la casa discografica dei Beatles, determina una rivoluzione nell'ambito della radiologia, paragonabile a quella dovuta alla scoperta dei raggi X da parte di W. C. Roengten nel 1896. Il primo tomografo viene installato nel 1971 presso l'Atkinson Morley's Hospital di Wimbledon, Inghilterra. Nel 1979 G. H. Hounsfield riceve il premio Nobel per la Fisiologia e Medicina assieme ad A. Cormack che aveva condotto alcuni studi teorici e sperimentali sui principi base.

Le immagini tomografiche (che vengono solitamente chiamate immagini TAC, da *tomografia assiale computerizzata*) appaiono come radiografie di sezioni del corpo umano, caratterizzate tuttavia da

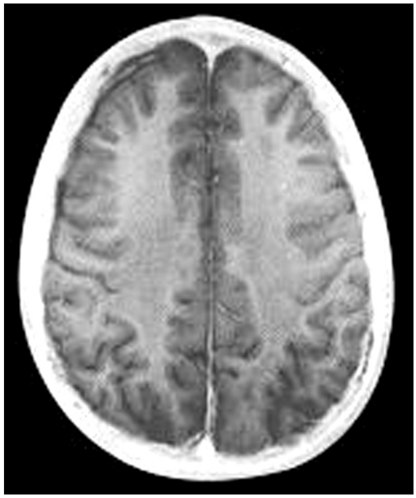


Figura 1. – Immagine tomografica di una sezione di cervello umano.

una risoluzione e da un contrasto nettamente superiori a quelli delle radiografie usuali, come si può dedurre dall' esempio di immagine cerebrale mostrato in Figura 1. Queste immagini sono ottenute elaborando dati relativi all'assorbimento subito dai raggi X durante l'attraversamento del corpo in esame. Mi propongo quindi di spiegare le idee e le tecniche che stanno alla base di tali metodiche.

Consideriamo un fascetto ben collimato di raggi X che si propaga lungo una retta che attraversa il corpo in esame. La sua intensità diminuisce lungo il cammino percorso secondo la legge seguente: se t indica una coordinata sulla retta, allora la derivata logaritmica dell'intensità $I(t)$ del fascio rispetto a t è eguale a $-\mu(t)$, essendo $\mu(t)$ il cosiddetto *coefficiente di attenuazione lineare* che è all'incirca proporzionale alla densità del corpo nel punto t. Pertanto se S ed R sono due punti della retta tali che il fascio si propaga da S a R e se si misura l'intensità sia in S che in R, il logaritmo del loro rapporto non è altro che l'integrale del coefficiente di attenuazione lineare lungo il segmento di retta da S ad R. Il problema è dunque quello di misurare tali integrali per un insieme opportuno di rette. Ciò può essere fatto mediante un procedimento di scansione che descriverò nel caso dei cosiddetti tomografi di

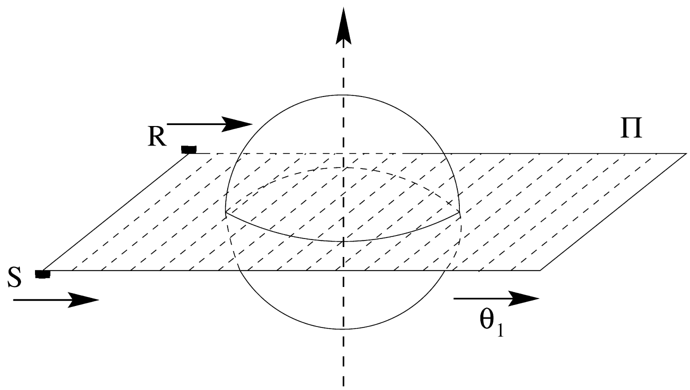


Figura 2. – Schematizzazione del processo di scansione nei tomografi di prima generazione. S indica la sorgente ed R il rivelatore. La descrizione del processo è data nel testo.

prima generazione, cioè quelli derivati dal primo tomografo di Hounsfield.

Facendo riferimento alla Figura 2, sia S la sorgente di un fascio di raggi X ben collimato e di intensità nota e sia R un rivelatore che misura l'intensità del fascio dopo che ha attraversato il corpo. Le posizioni di sorgente e rivelatore individuano una retta e, come si è precedentemente detto, i dati a disposizione permettono di ottenere l'integrale da S a R della densità del corpo. Il processo di scansione utilizzato è il seguente:

1) sorgente e rivelatore vengono mossi lungo due rette parallele e perpendicolari alla congiungente $S - R$; in tal modo viene individuato un piano Π che definisce la sezione del corpo di cui si produrrà l'immagine; per ogni posizione di sorgente e rivelatore si misura l'attenuazione e dunque si ottiene l'integrale lungo la retta corrispondente;

2) il sistema sorgente-rivelatore viene ruotato di un certo angolo (ad es. una qualche frazione di grado) attorno all'asse z perpendicolare al piano Π e viene ripetuto il procedimento precedente; si ruota di nuovo il sistema e così via fino a ricoprire tutto l'angolo di 180°;

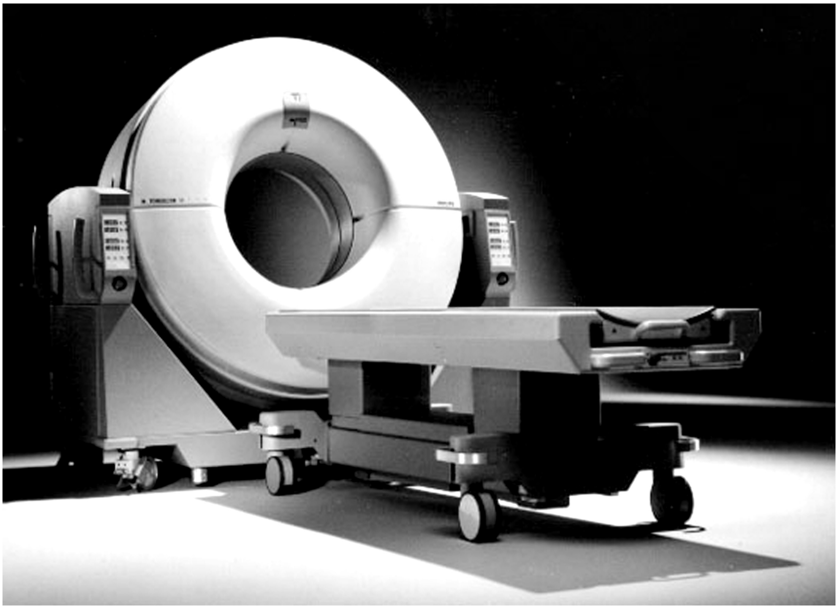


Figura 3. – Un esempio di tomografo a raggi X.

3) si sposta il sistema sorgente-rivelatore lungo l'asse z e si ripete tutto il procedimento precedente.

Nelle macchine attuali la scansione descritta nel punto 1) è eliminata utilizzando un sistema di più sorgenti e rivelatori che permettono di acquisire simultaneamente i dati relativi ad una direzione. In Figura 3 si vede un particolare tipo di macchina tomografica (sovente denominata *scanner* per evidenziare il fatto che il suo funzionamento è basato su un processo di scansione). La parte anulare della macchina contiene il sistema sorgenti-rivelatori nonchè tutti i meccanismi necessari per effettuare la scansione. Il paziente viene ovviamente introdotto all'interno dell'anello dove deve rimanere, il più possibile immobile, per tutto il tempo necessario ad effettuare l'acquisizione dei dati.

I passi 1) e 2) del procedimento di scansione permettono di ottenere i dati relativi al piano Π; il passo 3) permette di ottenere dati relativi ad altri piani: si hanno dunque dati relativi ad un volume sezionato mediante piani, da cui appunto il nome di *tomografia* per questo processo, cioè rappresentazione mediante sezioni. Nella fase di elaborazione dei dati i vari piani sono considerati separatamente e dunque ci limitiamo a trattare il caso di un singolo piano.

Indichiamo con $f(x)$ la densità del corpo nel piano Π, essendo x le coordinate di un punto del piano rispetto ad un opportuno sistema di riferimento (solitamente si sceglie come origine l'intersezione del piano con l'asse di rotazione della macchina). Indichiamo con θ il versore della retta lungo la quale si muovono sorgente e rivelatore e con s la distanza dall'origine, con segno, della retta, perpendicolare a θ e di versore $\theta^{\perp}$, lungo la quale si integra; indichiamo infine con t la coordinata sulla retta di integrazione. Tali coordinate sono precisate in Figura 4. Nel modello continuo del processo di scansione, i dati sono gli integrali lungo tutte le possibili rette, corrispondenti a

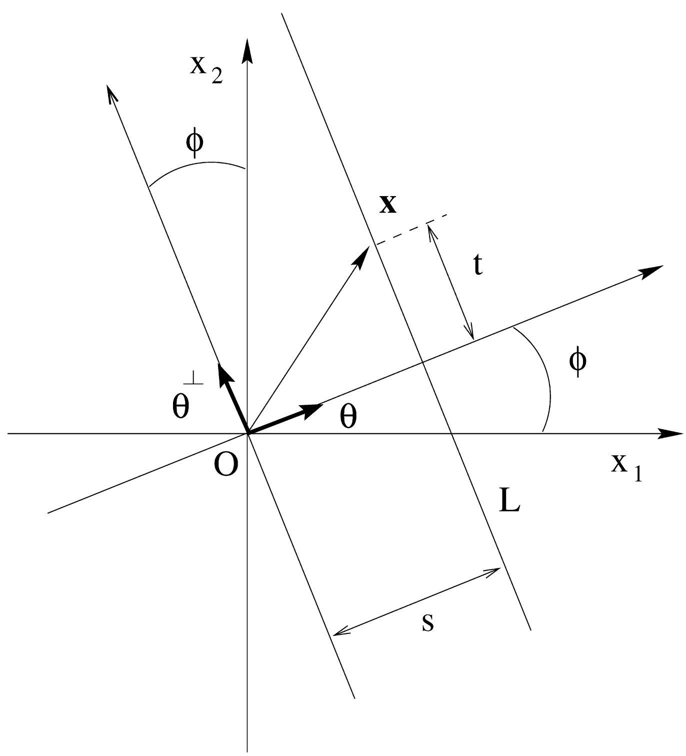


Figura 4. – Rappresentazione delle coordinate usate nella definizione della trasformata di Radon.

tutti i possibili valori di θ ed s. Essi sono qui indicati con $g(s, \theta)$ ed ammettono la seguente rappresentazione:

$$g(s, \theta) = \int_{-\infty}^{+\infty} f(s\theta + t\theta^{\perp})\, dt\,. \tag{1}$$

Il problema matematico che sta alla base della generazione di immagini tomografiche è dunque quello della determinazione di una funzione di due variabili, essendo assegnati i valori dei suoi integrali lungo tutte le possibili rette. Una formula risolutiva per tale problema fu ottenuta da Johann Radon nel 1917 (il suo risultato era più generale perchè si riferiva al caso di una funzione di n variabili, essendo dati i suoi integrali su tutti i piani di dimensione $n-1$) e pertanto, in suo onore, la funzione $g(s, \theta)$ è detta *trasformata di Radon* della funzione f. La funzione di s che si ottiene fissando θ nella (1) è detta *proiezione di f nella direzione θ*. È interessante osservare che, benchè il risultato di Radon non fosse noto all'inventore del tomografo che utilizzò un metodo algebrico per risolvere il problema, l'algoritmo che viene ora comunemente utilizzato è sostanzialmente un'implementazione intelligente della formula di Radon.

Quando si discute il problema dell'inversione della trasformata di Radon in due variabili, è sovente utile ricorrere al concetto di *sinogramma*, che è un'opportuna rappresentazione grafica della trasformata di Radon nel piano delle variabili s e ϕ (ϕ è l'angolo formato dal versore θ con l'asse x_1 nel piano Π - si veda ancora la Figura 4). Questa rappresentazione è ottenuta attribuendo al valore numerico della trasformata di Radon in un punto s, ϕ un livello di grigio in una scala che va dal nero (corrispondente allo zero) al bianco (corrispondente al valore massimo). In Figura 5 viene mostrato un esempio di sinogramma. La Figura 5 *a*) è una sezione di un fantoccio di cervello umano mentre la Figura 5 *b*) è il corrispondente sinogramma. Le ascisse corrispondono alla variabile s mentre le ordinate corrispondono alla variabile ϕ. Pertanto una riga del sinogramma è la rappresentazione di una proiezione. La figura fa comprendere l'origine del nome. Infatti il sinogramma di un oggetto puntiforme è un tratto di curva sinusoidale. Il sinogramma è, in un certo senso, l'im-

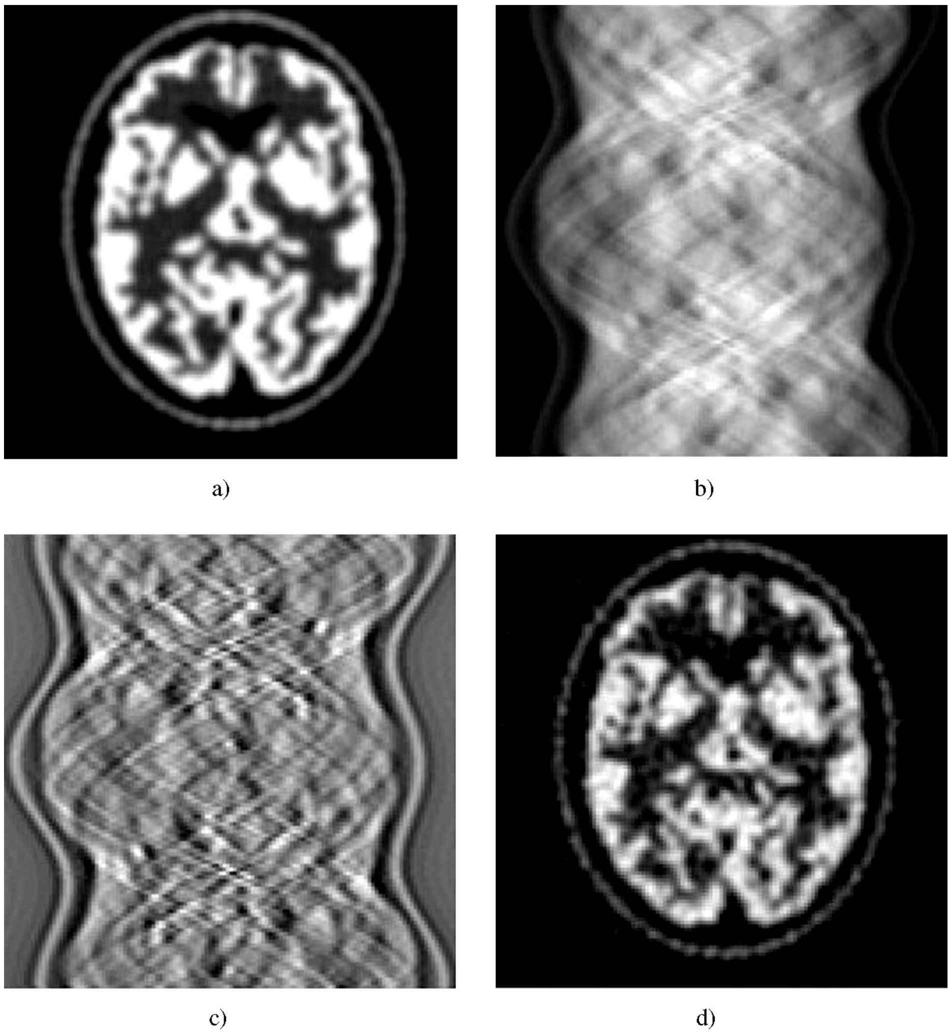


Figura 5. – *a*) Sezione di un fantoccio di cervello umano. *b*) Il suo corrispondente sinogramma, cioè la rappresentazione a livelli di grigio della sua trasformata di Radon. Ogni riga del sinogramma corrisponde ad una proiezione. *c*) Sinogramma filtrato, ottenuto da *b*) mediante l'applicazione del filtro a rampa. *d*) Ricostruzione ottenuta applicando l'operazione di retroproiezione al sinogramma filtrato.

magine grezza fornita da un tomografo. È chiaro che tale immagine non è direttamente interpretabile senza ulteriore elaborazione.

Il risultato fondamentale che porta all'inversione della trasformata di Radon può essere ottenuto in modo sostanzialmente elementare mediante il calcolo della trasformata di Fourier di ambo i membri dell'equazione (1), in quanto funzioni della variabile s per θ fissato. Esso può essere formulato dicendo che la trasformata di Fourier della proiezione di f nella direzione θ (primo membro della (1)) è eguale alla trasformata di Fourier della funzione f lungo la retta passante per l'origine ed avente direzione θ (secondo membro della (1)). Questo risultato, che va sotto il nome di *Fourier slice theorem*, riduce l'inversione della trasformata di Radon all'inversione della trasformata di Fourier. Tale inversione va effettuata in coordinate polari anzichè in coordinate cartesiane, poichè, calcolando la trasformata di Fourier di tutte le proiezioni, si ottiene la trasformata di Fourier di f lungo tutte le rette passanti per l'origine.

L'algoritmo di inversione che si ottiene a partire dalle suddette considerazioni viene solitamente indicato con il nome di *retroproiezione filtrata* (FBP, cioè *filtered backprojection*, in inglese) e consiste di operazioni piuttosto semplici: il primo passo è il calcolo delle trasformate di Fourier delle varie proiezioni, che sono poi moltiplicate per il cosiddetto filtro a rampa (che non è altro che lo Jacobiano della trasformazione da coordinate cartesiane a coordinate polari). Calcolando le trasformate di Fourier inverse si ottengono le proiezioni filtrate, la cui rappresentazione grafica è il sinogramma filtrato. Nell'esempio di Figura 5, il sinogramma filtrato è mostrato in Figura 5 *c*). Come si vede esso è molto più nitido del sinogramma originario. Infine l'ultimo passo è l'operazione di retroproiezione che consiste nell'attribuire ad un punto x, dove si vuol determinare la funzione f, la somma (l'integrale) di tutti i valori della trasformata di Radon corrispondenti a rette passanti per quel punto. Il risultato dell'applicazione della retroproiezione al sinogramma filtrato è mostrato in Figura 5 *d*).

Come detto in precedenza, il metodo qui descritto è equivalente alla formula di inversione di Radon anche se essa appare molto diversa poichè è scritta in termini di derivata dei dati, trasformata di Hilbert ecc. Tuttavia esso è molto più conveniente dal punto di vista del calcolo poichè tutte le operazioni che esso richiede sono estremamente veloci. In particolare per le operazioni di filtraggio si può ricorrere all'algoritmo veloce per il calcolo della trasformata di Fourier, la cosiddetta *Fast Fourier Transform* (FFT), anche se, sovente, vengono utilizzati circuiti dedicati per ridurre al massimo i tempi di elaborazione. In tal modo è possibile elaborare in pochi secondi milioni di dati, che sono appunto quelli richiesti se si vogliono ricostruire alcune decine di sezioni, utilizzando alcune centinaia di proiezioni per ogni sezione. Infatti da un'analisi della discretizzazione del problema basata sui teoremi di campionamento risulta che il numero ottimale di punti per proiezione è circa eguale al numero di proiezioni.

Un altro vantaggio del metodo FBP è che esso permette di introdurre facilmente metodi di regolarizzazione al fine di correggere effetti di instabilità numerica dovuti al filtro a rampa che amplifica il rumore (*noise*, in inglese) di alta frequenza. Occorre per altro dire che il carattere mal posto dell'inversione della trasformata di Radon è stato accuratamente studiato, giungendo alla conclusione che esso non impedisce di ottenere ottimi risultati nella maggior parte dei casi di interesse. Si veda ad esempio il libro di Natterer, che fornisce una presentazione completa dei problemi matematici legati alla tomografia.

Dalle precedenti considerazioni risulta che in un tomografo di prima generazione il tempo di calcolo è di molto inferiore al tempo richiesto per l'acquisizione dati, cioè per effettuare il processo di scansione precedentemente descritto. Poichè è opportuno ridurre il tempo durante il quale il paziente è sottoposto all'esame clinico, al fine di ridurre il suo disagio nonchè la dose di radiazione assorbita, e poichè l'incidenza del tempo di calcolo è irrilevante sul tempo complessivo, l'evoluzione dei tomografi è andata nella direzione di ridurre i tempi di acquisizione. Ciò è stato possibile utilizzando diversi sistemi di scansione. La maggior parte dei tomografi attuali sono del tipo *fan beam*, cioè con fascio a ventaglio, mentre si stanno sempre più diffondendo quelli del tipo *cone beam*, cioè con fascio a cono, che permettono di acquisire dati relativi ad un volume anzichè solamente ad un piano. Naturalmente ciò ha richiesto anche un'evoluzione degli algoritmi di calcolo, anzi talvolta gli algoritmi hanno preceduto le macchine su cui sono stati implementati. Bisogna anche dire che le macchine del tipo *cone beam* richiedono di uscire dall'ambito della trasformata di Radon poichè presentano problemi particolari legati all'inversione della *X-ray transform*, cioè della trasformata che associa ad una funzione di tre variabili i valori dei suoi integrali su rette che appartengono ad una varietà tridimensionale opportuna (ad esempio tutte le rette di coni con una certa apertura e vertice mobile su un cerchio o su una spirale, che è appunto la geometria dei tomografi tipo *cone beam*).

Occorre inoltre dire che il successo della tomografia a raggi X ha suggerito lo sviluppo di altre tecniche per uso clinico quali sono la *positron emission tomography*, cioè la tomografia a emissione di positroni o PET e la *single photon emission computed tomography*, cioè la tomografia ad emissione di fotone singolo o SPECT. Queste tecniche, che appartengono alla cosiddetta *tomografia ad emissione*, forniscono immagini che sono dette di tipo *funzionale* per il tipo di informazione che esse forniscono. Infatti, sia nel caso della PET che nel caso della SPECT, viene somministrato al paziente un radiofarmaco che contiene un tracciante radioattivo. Il radiofarmaco si distribuisce nei tessuti della parte di corpo da analizzare in base all'andamento del flusso ematico, dei processi metabolici ecc., e quindi la concentrazione del radiofarmaco nei vari tessuti dipende dalla funzionalità dei vari organi. Le immagini fornite dalle tecniche suddette sono una mappa di tale distribuzione di concentrazione. È quindi ovvio che, mentre si può fare una TAC ad una mummia, non si può fare una PET o una SPECT.

I traccianti utilizzati nella PET decadono emettendo un positrone che, a sua volta, si annichila con un elettrone emettendo due fotoni che si propagano in direzioni opposte; i traccianti usati nella SPECT decadono emettendo un fotone singolo. In entrambi i casi si tratta di selezionare le coppie di fotoni o i fotoni singoli che provengono da una ben determinata direzione. Ciò viene fatto mediante opportuni sistemi di collimazione e quindi i dati che si ottengono rappresentano gli integrali della concentrazione lungo ben determinate rette. In tal modo il problema viene sostanzialmente ricondotto a quello dell'inversione della trasformata di

Radon. Ne consegue che l'algoritmo solitamente implementato sulle macchine commerciali è la retroproiezione filtrata.

Va tuttavia osservato che il modello basato sulla trasformata di Radon è, in svariate circostanze, un modello troppo rozzo perchè non tiene conto di vari fenomeni fisici cui sottostanno i fotoni prima di giungere ai rivelatori nonchè del fatto che il sistema di collimazione non permette di ottenere gli integrali della concentrazione su rette, bensì su opportuni domini che dipendono appunto dalla geometria del sistema. Un'accurata modellizzazione del processo di misurazione fa uscire dall'ambito della trasformata di Radon ed occorre quindi ricorrere a metodi usati per la risoluzione di problemi mal posti con grandi masse di dati, solitamente metodi di tipo iterativo.

Prima di concludere questa parte dedicata alle immagini biomediche, occorre ricordare un'altra tecnica il cui uso è sempre più frequente in ambito clinico, cioè la *risonanza magnetica*. Essa si basa su principi fisici completamente diversi da quelli dei tomografi a raggi X. Mediante complessi sistemi di campi magnetici vengono individuati piani di cui si vuol formare l'immagine e vengono polarizzati i momenti magnetici dei nuclei di idrogeno appartenenti ad essi. Facendo rilassare il sistema in modo opportuno, si ottengono dati che rappresentano, con buona approssimazione, la trasformata di Fourier della densità di idrogeno nel piano considerato. Pertanto l'algoritmo di base per la formazione delle immagini è l'inversione della trasformata di Fourier. Ne consegue che anche in questo caso l'elaborazione dei dati è molto veloce. Va altresì osservato che la risonanza magnetica è molto più versatile della tomografia a raggi X poichè la giacitura dei vari piani di cui si forma l'immagine non è fissata come in quel caso. Si può inoltre intervenire su altri fattori legati ai tempi di rilassamento al fine di evidenziare particolari caratteristiche dei tessuti che si vogliono studiare.

3. – Immagini astronomiche e problemi di deconvoluzione.

Il primo utilizzo sistematico di metodi per la ricostruzione (o restauro) di immagini in Astronomia è stato effettuato nel caso delle immagini del telescopio spaziale Hubble. Mi propongo di spiegare il contesto in cui ciò avvenne.

Un telescopio consiste fondamentalmente di uno specchio di grandi dimensioni, il cosiddetto specchio primario, che focalizza la luce che riceve dal cielo in un piano situato al di sopra di esso. Mediante opportuni sistemi ottici, consistenti di uno specchio secondario ed altro, l'immagine viene poi formata in un piano che si trova al di sotto dello specchio primario (che prevede un foro centrale proprio per permettere una tale situazione), in modo da poter accedere facilmente a tale immagine con la strumentazione necessaria alla sua rilevazione.

Uno dei requisiti principali di un telescopio è il suo potere risolutivo, cioè la sua capacità di discriminare oggetti distanti corrispondenti ad angoli molto vicini. Come è noto dalle leggi dell'Ottica, la distanza angolare minima che può essere risolta mediante uno specchio di diametro D è data da $1.22\lambda/D$, essendo λ la lunghezza d'onda della luce osservata (lo stesso telescopio ha potere risolutivo diverso nel visibile, nell'infrarosso ecc.). Tale limite è dovuto ai fenomeni di diffrazione e, per tale motivo, viene anche detto il *limite di diffrazione* di un telescopio.

In pratica il potere risolutivo di un telescopio a terra è di molto inferiore al limite di diffrazione, principalmente a causa della turbolenza atmosferica che perturba i raggi provenienti dagli oggetti celesti, deviandoli in modo aleatorio. A causa di questi fenomeni, un telescopio di 5 metri, posto in un sito anche elevato, può avere un potere risolutivo equivalente a quello di un telescopio di mezzo metro di diametro!

Il problema di ridurre l'effetto della turbolenza atmosferica è dunque cruciale e per questo motivo i telescopi di grandi dimensioni vengono localizzati in siti il più possibile elevati. Ad esempio, i due più grandi telescopi fino ad oggi costruiti, quelli con specchio da 10 metri dell'Osservatorio W. M. Keck, sono situati a più di 4000 metri sul vulcano spento Manua Kea alle Hawaii. Il limite estremo è stato poi raggiunto con il telescopio spaziale Hubble, il cui schema è mostrato in Figura 6, che è stato posto al di fuori dell'atmosfera terrestre in modo da eliminare completamente gli effetti nocivi della tur-

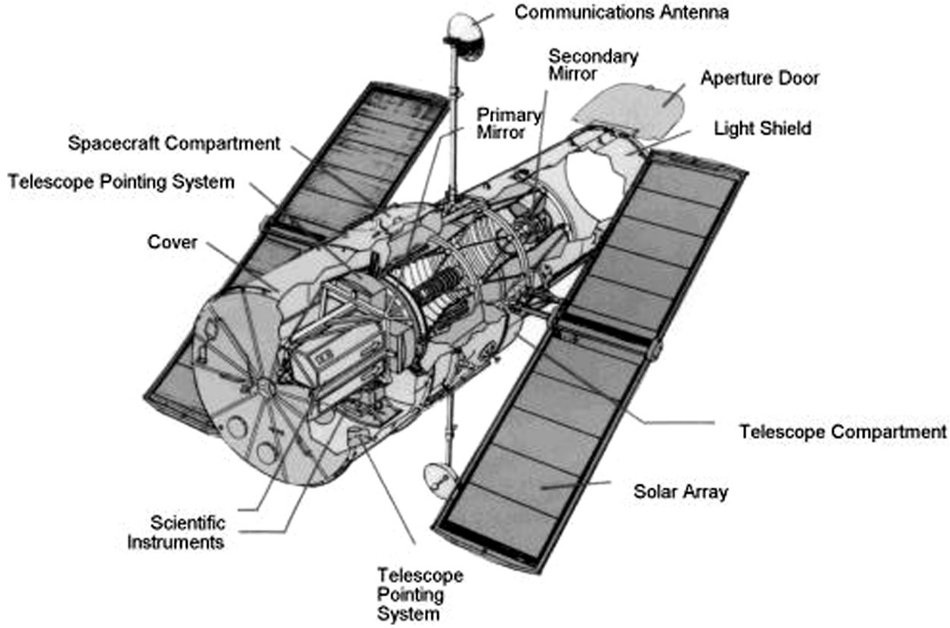


Figura 6. – Schema del telescopio spaziale Hubble.

bolenza. Proprio per questo motivo, pur avendo un primario con diametro di poco superiore ai due metri, lo Hubble ha fornito immagini di qualità e risoluzione mai prima raggiunti.

Tuttavia il telescopio spaziale, ai suoi inizi, fu al centro di una singolare vicenda. Subito dopo il lancio, avvenuto agli inizi del 1990, ci si accorse che, a causa di un errore di costruzione, non era possibile mettere a fuoco le immagini: il raggio di curvatura dello specchio primario non corrispondeva a quello previsto nel progetto. Era come se il telescopio fosse miope ed avesse bisogno di occhiali per poter vedere nitidamente. Gli occhiali furono costruiti (la cosiddetta ottica correttiva COSTAR) ed installati durante la missione dello Shuttle del Dicembre 1993. Tuttavia per quattro anni gli astronomi furono costretti a lavorare con le immagini fornite da un telescopio miope e quindi a tentare di correggerne la miopia mediante metodi numerici.

Il modello matematico che sta alla base del problema è piuttosto facile da formulare, almeno nella sua versione più semplificata. Se indichiamo con $f(x)$ il valore dell'immagine prodotta da uno Hubble ideale in un punto x del suo piano immagine e con $g(x)$ quello dell'immagine prodotta dallo Hubble miope nello stesso punto, allora la relazione tra f e g è data dal prodotto di convoluzione

$$g(x) = \int K(x - x')\, f(x')\, dx', \tag{2}$$

l'integrale essendo esteso a tutto il piano immagine. La funzione $K(x)$ è solitamente detta *point spread function* (PSF), in quanto rappresenta l'immagine di un punto luminoso posto al centro del piano immagine. In Figura 7 *a*) si mostra la PSF dello Hubble in una particolare configurazione mentre in Figura 7 *b*) si mostra un'immagine del pianeta Saturno corrispondente a tale PSF. Lo sfocamento dell'immagine è evidente. In pratica, l'immagine misurata è una versione contaminata di quella data dall'equazione (2), la contaminazione essendo causata da errori di misura, rumore ecc.

Il problema matematico è dunque quello di ottenere una valutazione di f a partire da una versione contaminata di g. Tale problema, che viene anche detto *problema di deconvoluzione*, è un tipico esempio di problema mal posto. Se si calcola la trasformata di Fourier di ambo i membri dell'equazione (2), si ottiene

$$\widehat{g}(\omega) = \widehat{K}(\omega)\, \widehat{f}(\omega). \tag{3}$$

Tuttavia non si può valutare la trasformata di Fourier di f, $\widehat{f}$, mediante il rapporto $\widehat{g}/\widehat{K}$. Se si fa questa operazione nel caso dell'esempio di Figura 7, se si divide cioè la trasformata di Fourier dell'immagine di Saturno per la trasformata di Fourier della PSF, e se si calcola la trasformata di Fourier inversa del risultato, si ottiene l'immagine di Figura 7 *c*), che è chiaramente peggiore dell'immagine di partenza. Ciò è dovuto al fatto che $\widehat{K}(\omega)$ tende a zero per grandi valori della frequenza ω, cosicchè la divisione per $\widehat{K}(\omega)$ amplifica in modo intollerabile il rumore di alta frequenza che corrompe l'immagine registrata.

Il comportamento di $\widehat{K}(\omega)$ ad alte frequenze è sostanzialmente una conseguenza del fatto che il telescopio trasmette male le alte

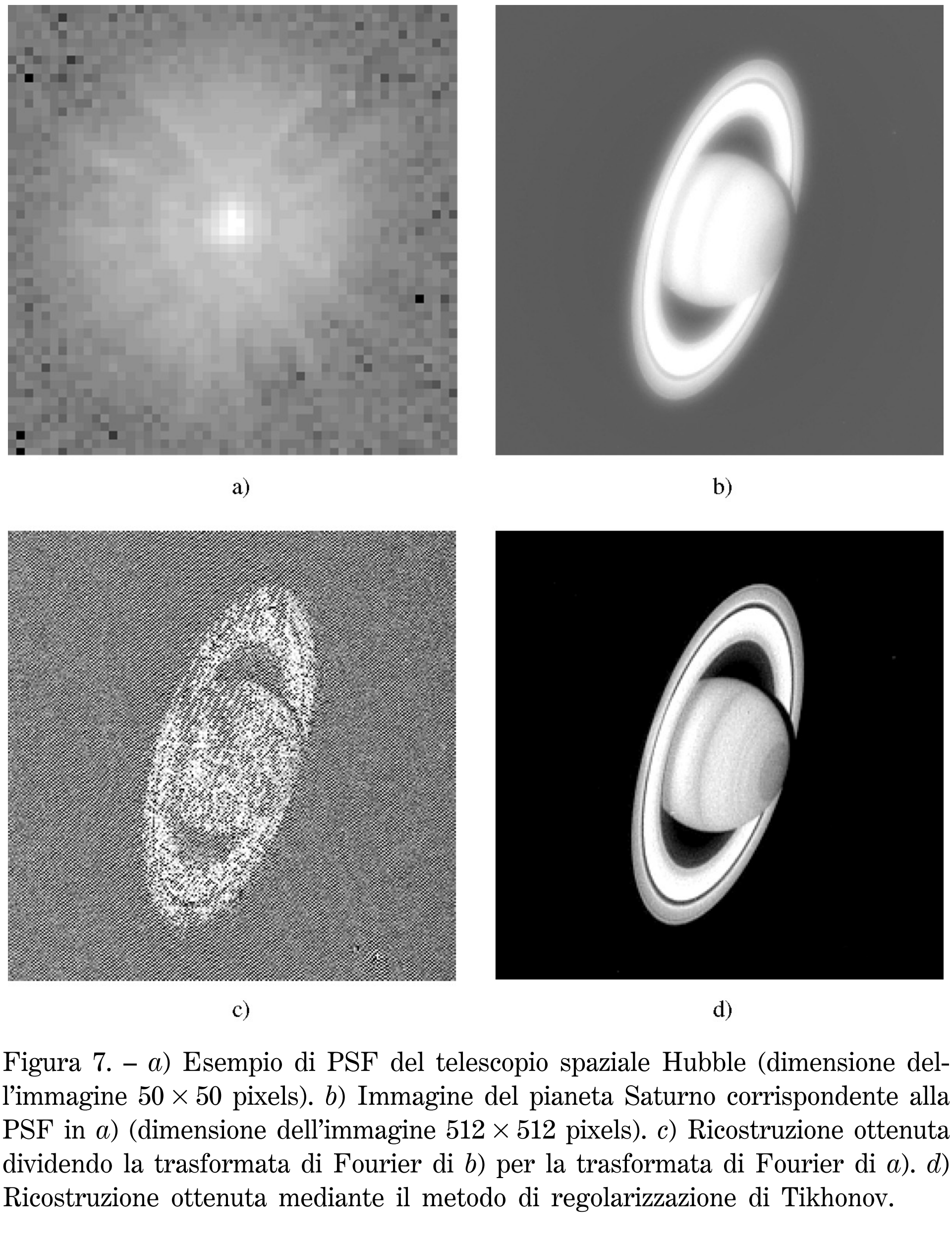


Figura 7. – *a*) Esempio di PSF del telescopio spaziale Hubble (dimensione dell'immagine 50×50 pixels). *b*) Immagine del pianeta Saturno corrispondente alla PSF in *a*) (dimensione dell'immagine 512×512 pixels). *c*) Ricostruzione ottenuta dividendo la trasformata di Fourier di *b*) per la trasformata di Fourier di *a*). *d*) Ricostruzione ottenuta mediante il metodo di regolarizzazione di Tikhonov.

frequenze dell'immagine e quindi non fornisce un'informazione completa su di esse. In una simile situazione occorre far ricorso ai metodi di regolarizzazione, cui si è accennato nell'Introduzione. In Figura 7 *d*) si mostra il risultato ottenuto mediante il cosiddetto metodo di Tikhonov. Questo esempio fa capire che è possibile ottenere risultati soddisfacenti con l'ausilio di tali metodi. Occorre tuttavia dire che le immagini fornite da Hubble dopo l'installazione dell'ottica correttiva sono decisamente superiori a quelle ottenute con metodi di ricostruzione a partire dalle vecchie immagini. Il motivo fondamentale è che le immagini dello Hubble miope contengono meno informazione di quelle dello Hubble corretto e tale perdita di informazione non può essere compensata mediante strumenti matematici.

È tuttavia interessante osservare che sono in fase di costruzione telescopi a terra che prevedono esplicitamente l'uso di tecniche di ricostruzione al fine di ottenere immagini di alta qualità e risoluzione. L'elemento tecnologico nuovo, che permette di andare al di là dei limiti imposti dalla turbolenza atmosferica, è la cosiddetta *ottica adattiva*. L'idea di base è la seguente: un sistema di sensori permette di misurare gli effetti della turbolenza sul fronte d'onda proveniente da una stella guida (naturale oppure artificiale, generata mediante una potente sorgente laser). Questa informazione viene utilizzata per modificare, in tempo reale, la superficie mobile di uno dei due specchi (in genere il secondario) in modo da correggere la perturbazione del fronte d'onda. Mediante questa tecnica è possibile compensare, a volte quasi totalmente con l'ausilio di opportuni metodi di deconvoluzione, gli effetti della turbolenza atmosferica così da ottenere risoluzioni vicine al limite di diffrazione precedentemente discusso.

Particolarmente interessante dal punto di vista dell'utilizzo di metodi di ricostruzione di immagini è il *Large Binocular Telescope* (LBT) che verrà installato a 3000 metri sul Mount Graham in Arizona. L' Italia contribuisce per il 25% al finanziamento di tale progetto e la comunità astronomica italiana è ampiamente coinvolta sia in fase di progettazione che in fase di futuro utilizzo del telescopio.

Il telescopio LBT sarà dotato di ottica adattiva. Tuttavia la sua caratteristica principale è che esso consisterà di due specchi di 8.4 metri di diametro con i centri posti alla distanza di 14.4 metri. Una rappresentazione pittorica di LBT viene mostrata in Figura 8. Quando le immagini dei due specchi vengono combinate interfero-

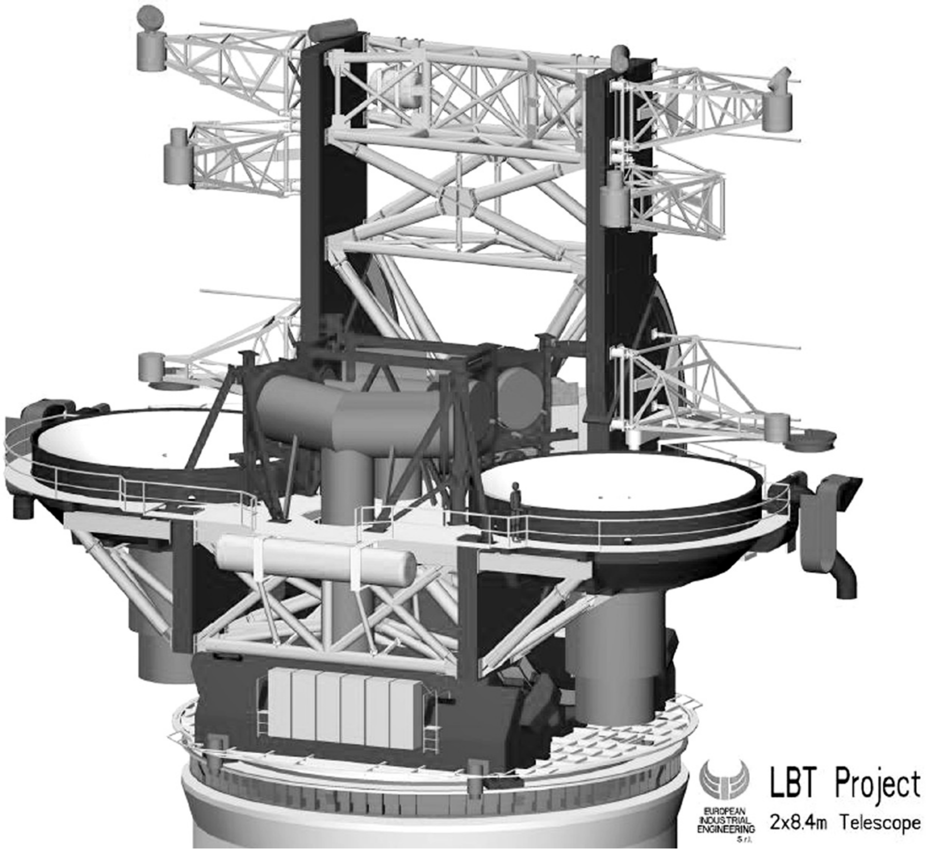
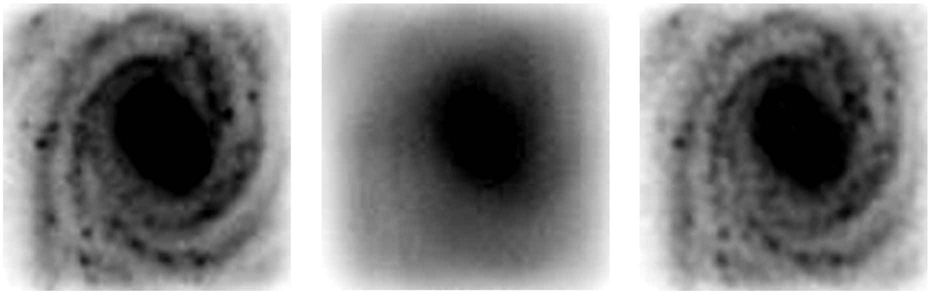


Figura 8. – Rappresentazione pittorica del telescopio LBT, in fase di costruzione sul Mount Graham in Arizona.

metricamente si ottiene un sistema la cui risoluzione nella direzione della congiungente i centri dei due specchi, la cosiddetta *baseline*, è equivalente a quella di uno specchio di 22.8 metri, mentre la risoluzione in direzione ortogonale rimane quella di uno specchio di 8.4 metri. Una singola immagine di LBT è dunque un'immagine che contiene un' informazione incompleta sull'oggetto celeste sotto osservazione e che non può essere facilmente interpretata. Al fine di ovviare a questo inconveniente è previsto di poter ruotare il telescopio in modo da poter riprendere più immagini dello stesso oggetto corrispondenti a diverse orientazioni della baseline. Si pone quindi



a) b) c)

Figura 9. – *a*) Immagine di galassia, vista da un ipotetico telescopio di 22.8 metri. *b*) Immagine della stessa galassia fornita da LBT nel caso in cui la baseline è situata orizzontalmente nel piano immagine. *c*) Ricostruzione della stessa galassia utilizzando sei immagini di LBT, corrispondenti a sei diverse orientazioni della baseline, ricoprenti l'angolo di 180°.

l'interessante problema di ricostruire l'oggetto celeste a partire da più immagini, ognuna delle quali contiene una parte di informazione su di esso. Tale problema presenta anche alcune interessanti analogie con il problema base della tomografia.

In Figura 9 si mostra il risultato di una simulazione numerica che fa capire la qualità delle immagini che è possibile ottenere con tale telescopio. In figura 9 a) si mostra l'immagine di una galassia quale sarebbe vista da un telescopio di 22.8 metri mentre in Figura 9 *b*) si mostra un'immagine quale sarebbe fornita da LBT ammettendo che la baseline sia disposta orizzontalmente. Finalmente in Figura 9 *c*) si mostra una ricostruzione ottenuta a partire da sei immagini corrispondenti a sei diverse direzioni equispaziate della baseline, tali da coprire l'angolo di 180°. Come si vede la ricostruzione è sostanzialmente equivalente all'immagine fornita da un telescopio da 22.8 metri. Il metodo di ricostruzione utilizzato è una generalizzazione del metodo di Tikhonov cui si era già fatto ricorso nel caso delle immagini dello Hubble.

Le due applicazioni discusse in questa nota, il tomografo a raggi X ed il telescopio LBT, sono due chiari esempi di complesse strumentazioni fisiche che richiedono l'uso di metodi di risoluzione di problemi inversi mal posti per poter fornire immagini significative.

Questi sono due tra i tanti esempi possibili. Si è dunque in presenza di un settore relativamente nuovo della matematica applicata che deve considerare tutti i vari aspetti dei problemi studiati: dall'accurata modellizzazione matematica della strumentazione fisica considerata, allo studio delle proprietà analitiche delle formulazioni continue dei relativi problemi matematici (che saranno un qualche esempio di problema inverso mal posto), alla discretizzazione dei metodi di risoluzione (tipicamente metodi di regolarizzazione) ed alla loro analisi numerica, fino alla scelta degli algoritmi e degli strumenti di calcolo più idonei ad affrontare il problema, tenendo conto delle masse di dati a disposizione e dei tempi entro i quali è necessario fornire una risposta. I due esempi considerati chiariscono quest'ultimo punto: mentre nel caso di un tomografo la risposta deve essere data entro pochi secondi, una tale esigenza non è particolarmente sentita nel caso del telescopio LBT.

Desidero ringraziare Patrizia Boccacci per alcuni utili suggerimenti circa la redazione di questa nota nonchè per l'assistenza nella preparazione delle figure.

BIBLIOGRAFIA

M. BERTERO - P. BOCCACCI, *Introduction to Inverse Problems in Imaging*, Bristol: Institute of Physics Publishing, 1998.

H. W. ENGL - M. HANKE - A. NEUBAUER, *Regularization of Inverse Problems*, Dordrecht: Kluwer.

J. HADAMARD, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, New Haven, CT: Yale University Press, 1923

F. NATTERER, *The Mathematics of Computerized Tomography*, New York: Wiley, 1986.

A. N. TIKHONOV - V. Y. ARSENIN, *Solutions of Ill-Posed Problems*, Washington: Winston/Wiley, 1977.