
Inverse problems in biomedical imaging: modeling and methods of solution

M. Bertero¹ and M. Piana²

¹ Dipartimento di Informatica e Scienze dell'Informazione, Università di Genova,
Via Dodecaneso 35, 16146 Genova

`bertereo@disi.unige.it`

² Dipartimento di Informatica, Università di Verona, Ca' Vignal 2, Strada le
Grazie 15, 37134 Verona

`piana@sci.univr.it`

Imaging techniques are a powerful tool for the analysis of human organs and biological systems and they range from different kinds of tomography to different kinds of microscopy. Their common feature is that they require mathematical modeling of the acquisition process and numerical methods for the solution of the equations relating the data to the unknown object. These problems are usually named *inverse problems* and their main feature is that they are ill-posed in the sense of Hadamard, so that their solutions require special care. In this chapter we sketch the main issues which must be considered when treating inverse problems of interest in biomedical imaging.

1 Introduction

The invention of computed tomography (CT) by G. H. Hounsfield at the beginning of the seventies was a breakthrough in medical imaging comparable to the discovery of X-rays by W. C. Roentgen in 1895. Even if CT and radiography derive from the same physical phenomenon, the conception of CT was based on ideas which opened new and wide perspectives. Indeed, CT requires mathematical modeling of X-ray absorption in order to provide equations which relate the observed data to the unknown physical parameters, and methods for the solution of these equations. In such a way it is possible to exploit the tremendous amount of information contained in radiographic data: a 3D image of the human body can be obtained, discerning variations in soft tissue such as the liver and pancreas, and measuring in a quantitative way the density variations of the different tissues. An accuracy of few percent can be obtained with a resolution of the order of 1 mm.

The new ideas introduced in CT were soon transferred to other methods for imaging human tissues. The first was magnetic resonance (MR), which is

based on the detection of the signals emitted by the magnetic moments of hydrogen nuclei when polarized by means of suitable static magnetic fields and excited by radiofrequency signals under resonance conditions. Moreover, earlier scintigraphic methods evolved into the functional imaging techniques known as positron emission tomography (PET) and single photon emission computed tomography (SPECT). In these cases a radio-pharmaceutical is administered to the patient and its distribution, due to metabolic processes, is investigated by detecting the γ -rays emitted by the radionuclides. As we briefly discuss at the end of this chapter, the development of other techniques, based, e. g., on microwaves and on infrared radiation, is in progress.

In general, the new techniques of medical imaging are based on the interrogation of the human body by means of radiation transmitted, reflected or emitted by the body: the effect of the body on the radiations is observed, a mathematical model for the body-radiation interaction is developed and the equations provided by this model are solved in post-processing of the observed data. The same approach applies to cell imaging by means of fluorescence or electron microscopy.

We emphasize a specific requirement of medical imaging, namely, the need for a solution in almost real time. In general a refined model of body-radiation interaction leads to complex non-linear equations, whose solution may require hours of computation time on a powerful computer. Hence the need to develop sufficiently accurate linear models, whenever this is possible, or also to design the observation process in such a way that a linear approximation is feasible. For this reason linearity is the first issue we discuss in this chapter (Sect. 2).

A second specific feature of biomedical imaging is that the problems to be solved are ill-posed in the sense of Hadamard. As we discuss in Sect. 3, being ill-posed implies that it is meaningless to look for exact solutions and that, nevertheless, the set of approximate solutions is too broad to be significant. In other words, although the data at our disposal can contain a tremendous amount of information, the fact that the problem is ill-posed, combined with the presence of noise, implies that the extraction of this information is not trivial.

A very important consequence of being ill-posed is that mathematical modeling of the medical imaging process cannot uniquely consist in establishing the equations relating the data to the solution; it must also include a model of the noise perturbing the data and, as far as possible, a model of known properties of the solution. Indeed the modeling of the noise is needed in order to clarify in what sense one is looking for approximate solutions; on the other hand the modeling of the solution properties must be used for extracting meaningful solutions from the broad set of approximate ones. Therefore noise and “a priori” information on the solution are two other important issues to be considered in biomedical imaging. These are discussed in Sect. 4 and Sect. 5 respectively. In Sect. 6 we outline the main computational issues concerned with the solution procedure and the solution methods which are most fre-

quently used in practice and lastly in Sect. 7 we provide a brief description of some of the current medical imaging techniques in progress.

Before concluding this introduction we briefly describe two important examples which can be used as reference cases for the general treatments described in subsequent sections: the first is X-ray tomography and the second fluorescence microscopy.

1.1 X-ray tomography

In the case of X-ray tomography we adopt a tutorial approach which does not correspond exactly to the data acquisition geometry in CT scanners. Therefore we assume that we have a source S emitting a well collimated X-ray beam; the beam crosses the body to be imaged and, at exit, its intensity is measured by a detector D . The attenuation of the X-rays across the body is described by the following simple model: let $f(x)$ be the attenuation coefficient at point x (roughly proportional to the density of the tissue at x); then, if u is a coordinate along the straight line L joining S and D (see Figure 1), the intensity loss at x is given by:

$$\frac{dI}{du}(x) = -f(x)I(x) \quad , \quad (1)$$

where I is the intensity measured by D . It follows that, if I_0 is the intensity emitted by S , then

$$I = I_0 \exp \left\{ - \int_L f(x) du \right\} \quad , \quad (2)$$

so that the logarithm of the ratio between the intensities of the emitted and detected radiation is just the line integral of the attenuation coefficient. By moving the $S - D$ system along two parallel lines, the plane to be imaged is defined and, by measuring the intensity for all the positions, one gets what is called a projection of the unknown function $f(x)$. More precisely, if θ is the unit vector in the direction of the movement of the $S - D$ system (linear scanning), s the distance (with sign) of L from the origin of the coordinate system (see Fig. 1), and θ' the unit vector in the orthogonal direction, then the *projection of f in the direction θ* is given by

$$(P_\theta f)(s) = \int f(s\theta + u\theta') du \quad . \quad (3)$$

By rotating the $S - D$ system and repeating the linear scanning for all possible angles (angular scanning) one obtains all possible projections and the result is the (two-dimensional) *Radon transform* of the function f : $(Rf)(s, \theta) = (P_\theta f)(s)$. These are just the data of X-ray CT, obtained by combining the linear and angular scanning as described above. Then, in order to get the function f , one has to solve the linear equation

$$g(s, \theta) = (Rf)(s, \theta) \quad , \quad (4)$$

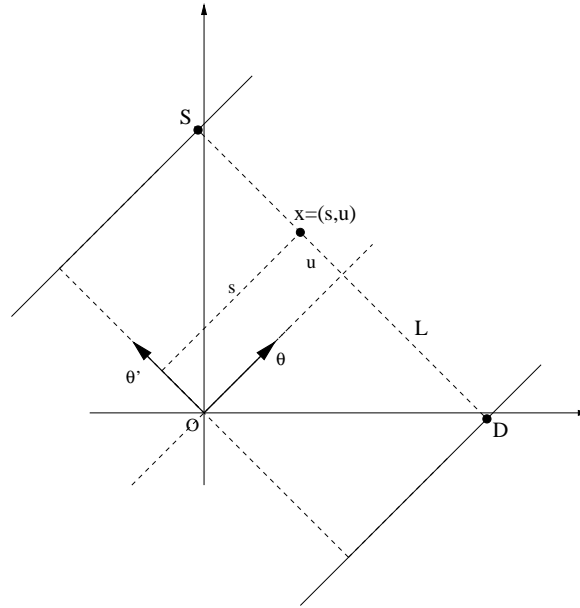


Fig. 1. Geometry of data acquisition in X-ray CT. The source S and the detector D move along two parallel straight lines with direction θ . The line L , joining S and D , is the integration line, with direction θ' , orthogonal to θ . A point x of L has coordinates $\{s, u\}$ with respect to the system formed by θ, θ'

where $g(s, \theta)$ denotes the measured data. This problem was solved by Radon in 1917 [59] and its inversion formula in the 2D case can be written as follows [57]:

$$f(x) = \frac{1}{4\pi^2} \int_{S^1} P \int_{R^1} \frac{1}{x \cdot \theta - s} \frac{\partial g}{\partial s}(s, \theta) ds d\theta \quad , \quad (5)$$

where P denotes the principal value. This formula clearly shows that the inversion of the Radon transform is an ill-posed problem since it requires the computation of the derivative of (noisy) data. Moreover the *filtered backprojection algorithm*, first introduced by Bracewell and Riddle [9] in radio astronomy and now widely used in medical imaging, is just a clever implementation of this formula.

The 3D imaging is obtained by repeating the previous procedure for different planes, namely, by scanning in the z -direction also, orthogonal to the imaging plane. Therefore the data of the problem depend on the variables $\{s, \theta, z\}$, which essentially characterize the position of the $S-D$ system. These data can be called the *image* of f , as provided by the CT scanner. For a given z the representation of g in the plane $\{s, \theta\}$ is the so-called *sinogram*. We give an example in Figure 2. It is obvious that the interpretation of these data without the help of a reconstruction algorithm is impossible. As text books in tomography we mention the books of Kac and Slaney [42] and Natterer [56].

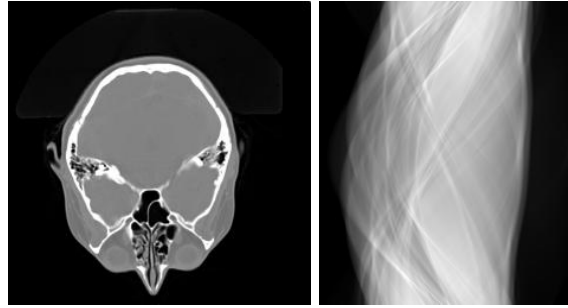


Fig. 2. Left-hand panel: tomographic reconstruction of a section of a human head. Right-hand panel: the corresponding sinogram.

1.2 Fluorescence microscopy

Fluorescence microscopy is a technique which is used for the investigation of living cells. The cell is treated with a fluorescent marker and its 3D image is formed by means of a technique known as *optical sectioning* [70], which can be applied to different kinds of optical microscopes (wide-field, confocal, multiphoton, etc. [28]). In all cases a 3D image is formed by acquiring a set of 2D images corresponding to different planes of focus.

Thanks to geometric optics there is a one-to-one correspondence between the points of the image plane where the detector is located and the points of the focal plane where the section of the object to be imaged is located. From this correspondence we can identify a point of the image plane with the corresponding point of the focal plane, hence with a point $x \in \mathbb{R}^3$ of the volume of the sample. In such a way the 3D image can also be considered as a function of x . However the image of one point receives contributions not only from the corresponding point in the focal plane but also from neighboring points both in the plane which is in focus and in the other planes. The result is an integral relationship between the image and the object.

By assuming perfect spatial incoherence of the radiation emitted by the sample it turns out that the relationship between the intensities of the detected and emitted radiation is linear. Moreover, by neglecting the spherical aberration of lenses, this relationship is also translation invariant. In such a case the imaging system is called *isoplanatic* or *space invariant*. As a result, the image $g(x)$ is given by the convolution product of the object $f(x)$ (proportional to the density of fluorescent atoms at x) with a function $h(x)$, which is the image of a point-source and is called the *point spread function* (PSF) of the imaging system:

$$g(x) = \int_{\mathbb{R}^3} h(x - x')f(x')dx' \quad . \quad (6)$$

The effect of the PSF is usually denoted as *blurring*; moreover the data are obviously corrupted by noise. It may be interesting to remark that (3) has a

similar structure but with a PSF which is a distribution and is not translation invariant.

In conclusion, the problem of reconstructing the object is equivalent to that of solving the convolution equation of (6). Such a problem is called *deconvolution* or *deblurring* and is another classical example of an ill-posed problem. An introduction to deconvolution methods is given in [4].

2 Linearity

As outlined in the introduction, a basic requirement in medical imaging is the availability of a linear relationship between the properties of the tissues to be imaged and the data provided by the medical equipment. All the algorithms implemented in commercial machines are based on such an assumption, because, in spite of the increasing power of computers, only in the case of linear problems it is possible to get a solution in almost real time for large-scale problems such as those arising in medical imaging. Of course the situation can change in the future and, for this reason, the investigation of non-linear models, which provide a more accurate description of the radiation-body interaction, is an exciting research topic.

In the mathematics of image processing the object is that property which is distributed in a volume and which has to be estimated by means of indirect measurements; it belongs to a well-characterized functional space (typically, for sake of simplicity, a Hilbert space) and, in what follows, it is denoted by f , a function of the variable $x \in \mathbb{R}^3$. On the other hand, the image, denoted from now on by g , is that measurement (again, in a suitable Hilbert space) which is provided by the imaging device and is regarded as best representing the object given the specific hardware. The image g is a function of the parameters, overall denoted by ξ , which characterize the acquisition process. The two examples outlined in the introduction show that these parameters can have quite different physical meanings: in the case of X-ray CT they characterize the position of the detector and the direction of the incoming radiation, while in the case of microscopy they can be identified with the coordinates of a point in the object volume. In the first case the 3D image is a set of sinograms, one for each section of the volume; in the second case, it is a “blurred” version of the object.

The relationship between object and image can be obtained by a mathematical model of the physical phenomena which provide the basis of the acquisition process. The most general representation of such a process is given by the non-linear integral equation

$$g(\xi) = \int h(\xi, x, f(x)) dx \quad , \quad (7)$$

where h is a continuous function of all its variables. This equation provides the solution of the so-called *direct* problem, namely, the problem which must

be solved for computing the data g related to a given object f . The continuous dependence on f is just due to the fact that this problem is well-posed, namely, the solution exists, is unique and depends continuously on the data (in this case, the object f). The inverse problem is obtained by exchanging the roles of the data and the solution: in such a case one must find the object f for a given image g .

The solution of (7) is very difficult from both the theoretical and practical points of view. No general theory exists for such a non-linear integral equation: each problem requires specific analysis. Moreover the problem may be ill-posed and no general regularization theory exists for wide classes of non-linear problem. A few general results applying to non-linear problems are described in the book of Engl et al. [29]. Most of these results are inspired by classical Newton-like optimization schemes, according to which stable approximate solutions of (7) can be obtained by stopping an iterative procedure appropriately initialized. The idea is to assume that a non-linear differentiable map \mathcal{F} can be defined between two functional spaces, so that

$$g = \mathcal{F}(f) \quad , \quad (8)$$

where \mathcal{F} is here the integral operator on the right-hand side of (7). With \mathcal{F}' denoting the Frechet derivative of \mathcal{F} , (8) can be replaced by the linearized equation

$$\mathcal{F}(f) + \mathcal{F}'(f)q = g \quad , \quad (9)$$

which has to be solved for the perturbation q to obtain the upgrade $f + q$. The method consists in solving iteratively the equations:

$$\mathcal{F}'(f^{(k)})q^{(k+1)} = g - \mathcal{F}(f^{(k)}) \quad , \quad f^{(k+1)} = f^{(k)} + q^{(k+1)}, \quad k = 0, 1, 2, \dots \quad (10)$$

starting from an initial guess $f^{(0)}$ for the function f . This approach has two fundamental drawbacks. First of all, it requires, at each step of the iteration procedure, the solution of the direct problem in order to compute the map \mathcal{F} (this point is clarified by the example of the inverse scattering problem discussed below) as well as its Frechet derivative; therefore it is extremely demanding from the numerical point of view and an accurate approximation of the solution can be achieved only after such a long computational time that no realistic technological application is possible at the moment. Second, the sequence of iterates defined in (10) may converge only if a sufficiently accurate initialization of f is to hand. This can be a serious bias in applications to medical imaging.

The situation is quite different if the non-linear equation (7) can be approximated by a linear one:

$$g(\xi) = \int h(\xi, x)f(x)dx \quad , \quad (11)$$

where h is the impulse response of the imaging instrument, which can be a distribution, as in tomography, or a function, as in microscopy. Indeed, in

the case of linear inverse problems one has at one's disposal very powerful theoretical tools.

However, the key question is to understand what kind of approximation can lead to (11). Indeed, two possibilities may occur. In one case, (11) is a brand new model where linearity is obtained by a precise technological realization or is the consequence of physical approximations. For instance, in magnetic resonance (MR), data acquisition is designed in such a way that the data are just the values of the Fourier transform of the object to be imaged. On the other hand, in fluorescence microscopy, linearity is obtained by neglecting a physical phenomenon. Indeed, it is well-known that an optical system, namely, a system of lenses, is linear in the amplitude of the wave field in the case of perfectly coherent radiation and in intensity in the case of perfectly incoherent radiation (see, e. g., Goodman [32]). In the case of fluorescence, the emitted radiation is partially coherent but the degree of coherence is not very high so that the approximation of perfectly incoherent radiation is assumed to be satisfactory and one gets a linear integral equation relating the intensities of the emitted and detected radiation. Even the model used in X-ray CT corresponds to a rather simplified (but sufficiently accurate) description of the absorption of photons due to the interaction with the material and moreover, in this case, linearity is obtained by considering the logarithm of the data and not the data itself.

A completely different situation occurs when (11) is obtained by means of a sort of perturbation theory applied to (7), i.e., through a linearization of Eq. (7) around zero or an approximate object $f^{(0)}$. This approach can be illustrated by the following inverse scattering problem which represents a reasonable model for microwave tomography. For simplicity we consider a two-dimensional approximation and we assume that the body is illuminated by means of a set of plane waves with fixed wave number k , coming from different directions given by unit vectors θ . The body is characterized by a refractive index $n(x) \neq 1$ and is immersed in a medium with $n(x) = 1$; therefore it is described by the function

$$f(x) = n^2(x) - 1 \quad , \quad (12)$$

whose support Ω is the domain of the body. The direct problem consists in determining the wave function $u(x) \in C^2(\mathbb{R}^2 \setminus \partial D) \cap C^1(\mathbb{R}^2)$ which solves the problem:

$$\Delta_2 u(x) + k^2 n^2(x) u(x) = 0 \quad , \quad (13)$$

$$u(x) = e^{ik\theta \cdot x} + u^s(x) \quad , \quad (14)$$

$$\lim_{r \rightarrow \infty} \sqrt{r} \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0 \quad . \quad (15)$$

In (14) the first term is the incident plane wave with direction θ and the second term is the scattered wave. Potential theory allows us to prove that

the elliptic differential problem (12)-(15) is equivalent to the integral equation [21]:

$$u(x) = e^{ik\theta \cdot x} - k^2 \int_{\mathbb{R}^2} \Phi(x, y) f(y) u(y) dy \quad , \quad x \in \mathbb{R}^2 \quad , \quad (16)$$

known as the Lippmann-Schwinger equation, in which $\Phi(x, y)$ is the fundamental solution of the Helmholtz equation in \mathbb{R}^2 given by

$$\Phi(x, y) = \frac{i}{4} H_0^{(1)}(k|x - y|) \quad , \quad (17)$$

where $H_0^{(1)}$ is the zero-order Hankel function of the first kind. From this integral formulation, together with considerations based on the unique continuation principle, one can prove that the direct problem is well-posed. Furthermore, the Sommerfeld radiation condition (15) implies that the scattered field, propagating in the direction θ' , can be asymptotically factorized in the form

$$u^s(r, \theta', \theta) = \frac{e^{ikr}}{\sqrt{r}} u_\infty(\theta', \theta) + O(r^{-3/2}) \quad , \quad (18)$$

where $u_\infty(\theta', \theta)$ is the far-field pattern of the scattered field u^s . By using Green's theorems one can easily show that

$$u_\infty(\theta', \theta) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\mathbb{R}^2} e^{-ik\theta' \cdot y} f(y) u(y) dy \quad , \quad (19)$$

where the dependence on θ on the right-hand side comes from the fact that u is given by (14). Equation (19) is an actual realization of (7) and allows us to define the non-linear inverse problem we are interested in. In such a problem the data function g is represented by the measured values of the far-field pattern u_∞ (with ξ represented by the unit vectors θ, θ'), the unknown object is the contrast function (12) and the non-linearity is due to the fact that $u(y)$ depends on $f(y)$ itself. In other words the mapping \mathcal{F} of (8) is defined by the right-hand-side of (19) and therefore its computation requires the solution of the direct problem (13)-(15).

However, under the low-frequency, weak-scattering assumption,

$$kMa < \pi \quad , \quad (20)$$

where $f(x) \neq 0$ for $|x| \geq a$ and $M = \sup_{|x| \leq a} |f(x)|$, a linearization of (19) can be obtained by applying the successive approximations method to the Lippmann-Schwinger equation. If we take the first-order approximation, i. e. we neglect the scattered wave with respect to the incident plane wave, we obtain the so-called Born approximation and the result is the diffraction tomography equation [27]

$$u_\infty(\theta', \theta) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\mathbb{R}^2} e^{-ik(\theta' - \theta) \cdot y} f(y) dy \quad . \quad (21)$$

In other words, in the Born approximation regime, the non-linear image reconstruction problem becomes a linear Fourier transform inversion problem with limited data, where the low-frequency cut-off which determines the data limitation is given by (20).

Linearizations like the one induced by Born approximation are unreliable in those cases, as for microwaves, in which the wavelength $\lambda = 2\pi/k$ of the field propagating in the biological tissue is of the same order of magnitude as the tissue dimensions and therefore diffraction effects are predominant. However, even in this case a linear approach is possible if the aim is the reconstruction of the support of the object f and not the object itself. This is the basic idea underlying visualization techniques based on the so-called *linear sampling method* [20, 23].

The starting point of this approach is to consider a one-parameter family of linear integral equations of the first kind which, in a sense, provide exactly the same estimate of the support of the object as the complete solution of the non-linear inverse scattering problem (12)-(15). In order to clarify this statement, we introduce the parameter x , which is just a point in a region of \mathbb{R}^2 (or \mathbb{R}^3 , if the problem is 3D) containing the support Ω of f and, for each x , we write the far-field equation [20, 18, 19]

$$\int_{S^1} u_\infty(\theta', \theta) g_x(\theta) d\theta = \Phi_{\infty, x}(\theta') \quad , \quad (22)$$

where:

- $u_\infty(\theta', \theta)$ is the far-field pattern of (19); in practical applications of the method it is approximated by means of its measured values;
- $\Phi_{\infty, x}(\theta')$ is the far-field pattern of the fundamental solution of the Helmholtz equation, which, in the 2D case discussed above, is given by the far-field pattern of (17), i. e.

$$\Phi_{\infty, x}(\theta') = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} e^{-ik\theta' \cdot x} \quad . \quad (23)$$

A solution $g_x(\theta)$ of (22) does not exist for any scattering data and, when it does exist, has no physical meaning. However it can be proved [12] that an approximate solution exists which has the property that it becomes unbounded for x on and in the exterior of the boundary of the scatterer, thus acting as an indicator for the boundary itself. Such behavior naturally inspires a visualization algorithm [23] where, for each point of the grid containing the inhomogeneity, an approximate stable solution of the far-field equation is constructed and its norm is plotted: the contour of the scatterer is given by all points x where this norm is greater than a given threshold.

This approach has two main advantages: first, the implementation is computationally simple and a notable computational speed is achieved (2D objects can be visualized in a few minutes, which become a couple of hours in the case

of very complicated 3D objects); second, very little a priori information on the inhomogeneity is necessary for the method to work (no knowledge of the number of scatterers, of their physical nature and of possible boundary conditions is required). A vast literature on the linear sampling method is at our disposal. Besides the more theoretical papers already cited, applications to biomedical imaging problems involve microwave tomography [51], impedance tomography [10] and the detection of leukemia in human bone marrow using microwaves [22]. Furthermore, a variety of similar inversion schemes has been formulated, involving the factorization method of Kirsch [44] and the indicator sampling method of You et al. [71]. Of course, such approaches also have significant drawbacks. As already mentioned, the linear sampling method is not a reconstruction method, in the sense that it simply allows a visualization of the object, without providing information on the point values of the refractive index. Moreover the spatial resolution achieved is not yet satisfactory, particularly if medical imaging applications are considered. However these techniques allow us to visualize, although coarsely, even very complex objects in a reasonable time, and therefore can be very helpful to provide initializations for Newton-like schemes or information on the support of the inhomogeneity in the case of the application of constrained iterative algorithms.

3 Ill-posed problems and uncertainty of solution

The availability of a reliable linearized mathematical model for image formation is not sufficient for a straightforward solution of the reconstruction problem. A crucial difficulty is due to the fact that most image reconstruction problems arising in biomedical imaging are ill-posed. The concept of being ill-posed was introduced by Hadamard as a mathematical anomaly in the solution of particular boundary value problems for partial differential equations. The discussion of the Cauchy problem for the Laplace equation is a classical example [33]. However, an exhaustive definition of ill-posed is given by a negative characterization: ill-posed problems do not satisfy at least one of the three conditions required for being well-posed, i.e., existence, uniqueness and continuous dependence of the solution on the data. As pointed out in [24], the lack of the third requirement has particularly important consequences for the solution of ill-posed problems modeling physical situations: indeed, in the case of practical applications, the presence of measurement noise in the data may imply the presence of strong numerical instabilities in the solution when obtained by means of a straightforward approach.

Inverse problems are very often ill-posed in the sense of Hadamard. In fact, in most cases of interest for applications, the linear operators modeling the problems are compact and being ill-posed is a consequence of this functional property of the model. This is the case, for example, of the two image reconstruction problems described in the introduction. Indeed, if the object has bounded support, for suitable choices of the source and data functional

spaces, both the Radon transform describing X-ray tomography and the convolution operator at the basis of fluorescence microscopy become compact linear operators.

In order to formally discuss the issue of being ill-posed and its relationship with compactness we consider linear inverse problems characterized by the following general structure [3]. The first step is to define the corresponding direct problem, whose solution allows us to define a linear operator A from the (Hilbert) space X containing all functions characterizing the unknown objects to the (Hilbert) space Y containing all functions describing the corresponding measurable images. Therefore the measured image $g \in Y$ is related to the true physical object $f^\circ \in X$ by

$$g = Af^\circ + h \quad , \quad (24)$$

where Af° is the (exact, or computable) image of the object f° and h is a term containing all possible experimental errors. When the signal-to-noise ratio associated with g is sufficiently high, the norm of Af° is larger than the norm of h , the noise term in (24) can be neglected and the linear inverse problem that we are interested in reduces to that of finding $f \in X$ such that

$$g = Af \quad . \quad (25)$$

Equations (24) and (25) display a clear pathological feature of ill-posed inverse problems. In fact, if we represent the problem by (24), we have only one equation for two unknowns: the true object f° and the function h . On the other hand, if we use (25), since the correct representation of g is given by (24), then a solution may not exist for all noise realizations.

According to a more formal approach, being ill-posed can be viewed as a property of the triple $\{A, X, Y\}$. Indeed, the data space Y must be broad enough to contain both the exact image Af° and the noisy image g , and therefore the range of A , $\mathcal{R}(A)$, is strictly contained in Y (typically the functions in $\mathcal{R}(A)$ are much smoother than the functions describing noisy images). Furthermore, when the kernel of A is not empty, the solution of the problem is not unique; finally, if A^{-1} , when it can be defined, is an unbounded mapping from Y to X , then the dependence of the solution on the data is not continuous. The only way out from this puzzling situation is, first of all, to give up looking for an exact solution of (25) and, for example, to consider the least-squares problem of determining all functions $f \in X$ such that

$$\|Af - g\|_Y = \min \quad . \quad (26)$$

It is easy to prove that the set of least-squares solutions coincides with the set of solutions of the Lagrange-Euler equation

$$A^*Af = Ag \quad , \quad (27)$$

where A^* is the adjoint operator of A . If P is the linear projection operator onto the closure $\overline{\mathcal{R}(A)}$ of the range of A , it can be shown that the Euler equation is in its turn equivalent to

$$Af = Pg \quad . \quad (28)$$

It follows that there exists a class of linear operators for which the set of least-squares solutions is not empty, namely, the operators whose range is closed. Furthermore, the set of least-squares solutions is a closed and convex subset of the Hilbert space X and therefore there exists only one least-squares solution of minimal norm, which is called the generalized solution and denoted by f^\dagger . It is natural to introduce the generalized inverse operator $A^\dagger : Y \rightarrow X$ mapping g to f^\dagger and, since A^\dagger is continuous if and only if $\mathcal{R}(A)$ is closed, then, in this case, the problem of determining the generalized solution is well-posed (these results apply to the case of discretized problems, as discussed later). However, as already stated, most inverse problems of interest in biomedical imaging are modeled by compact operators and easy considerations essentially based on the open mapping theorem show that, if A is compact, then $\mathcal{R}(A)$ is not closed (unless it is finite-dimensional). In other words, the search for a generalized solution of problem (25) for compact operators is still an ill-posed problem.

According to a different approach, the best one can do is to look for an f reproducing the given g within a tolerable uncertainty [6]. Since the image (24) can contain a term in $\mathcal{R}(A)^\perp = \ker(A^*)$ due to the noise, then the idea is to look for functions in X such that

$$\|Af - Pg\|_Y \leq \epsilon \quad , \quad (29)$$

where ϵ measures the magnitude of the noise. In order to show that, in the case of compact A , even this attempt is unsuccessful, we introduce the singular system of A [43], defined as the set of triples $\{\sigma_k; u_k, v_k\}_{k=1}^\infty$ solving the shifted eigenvalue problem

$$Au_k = \sigma_k v_k \quad , \quad A^* v_k = \sigma_k u_k \quad . \quad (30)$$

Before going on, we recall that a complete characterization of the singular system of the Radon transform in any dimension is known [26, 49] (also see [56, 4]). These results show that the singular values tend to zero very slowly, so that thousands of singular values are necessary for an accurate reconstruction. Moreover, the singular functions, which are related to orthogonal polynomials, become highly oscillating when associated to small singular values. This is a rather general property of the singular functions of operators involved in biomedical imaging.

Now, since the set of the singular functions $\{u_k\}_{k=1}^\infty \subset X$ is an orthonormal basis in $\ker(A)^\perp$ while the set of the singular functions $\{v_k\}_{k=1}^\infty \subset Y$ is an orthonormal basis in $\overline{\mathcal{R}(A)} = \ker(A^*)^\perp$, elementary computations based on (30) lead to the following expression of (29):

$$\sum_{k=1}^{\infty} \frac{\sigma_k^2}{\epsilon^2} \left| (f, u_k)_X - \frac{(g, v_k)_Y}{\sigma_k} \right|^2 \leq 1 \quad . \quad (31)$$

This equation defines the set of interior points of a sort of 'ellipsoid' in the infinite-dimensional solution space X . As the singular values of a compact operator are real positive numbers accumulating to zero when $k \rightarrow \infty$, then such an ellipsoid is unbounded, and, together with the true object, it also contains completely unreliable approximate solutions which, nevertheless, can reproduce the data within the prescribed accuracy. We observe in Section 5 that the main idea, common to most available methods for dealing with ill-posed problems, is just to restrict the class of admissible solutions by selecting a subset of this ellipsoid by exploiting a priori information available about the solution.

We conclude this section devoted to ill-posed problems, by discussing the case of finite-dimensional linear problems obtained from a sort of discretization of the original ill-posed linear inverse problem formulated in the infinite-dimensional Hilbert space setting.

Since, in general, we are interested in 3D images, we assume that the volume of the body is partitioned into N voxels, characterized by an index n , and we denote by f_n the average value of the quantity of interest $f(x)$ in the voxel n . Moreover we assume that the radiation transmitted, reflected or emitted by the body is measured by means of M detectors, characterized by an index m ; we denote by g_m the output of the detector m . We denote by \mathbf{f} and \mathbf{g} respectively, the vectors of the unknown parameters (the object) and of the outputs (the image). If a linear model has been developed for the imaging process, then the discretization of this model leads to a matrix A , $M \times N$, relating the unknown object \mathbf{f} to the image \mathbf{g} . In practical applications, typically we have $M \geq N$, so that the problem can be overdetermined. Then, in the absence of experimental errors, the output of the detector m should be given by

$$(A\mathbf{f})_m = \sum_{n=1}^N A_{m,n} f_n . \quad (32)$$

We can now reformulate a discrete least-squares problem, as in (26), with the norm of Y replaced, for instance, by the usual Euclidean norm of an M -dimensional vector space. The problem of determining the generalized solution is then well-posed in the sense of Hadamard. However, being well-posed is only a necessary condition for numerical stability. Indeed, if we introduce the singular system of the matrix A , the generalized solution is given by

$$f^\dagger = \sum_{k=1}^p \frac{(\mathbf{g}, \mathbf{v}_k)_2}{\sigma_k} \mathbf{u}_k , \quad (33)$$

where p is the rank of the matrix A and the scalar product is the Euclidean one. The numerical stability of this solution is controlled by the condition number $\alpha = \sigma_1/\sigma_p$ and, since the problem derives from the discretization of an ill-posed problem where the singular values accumulate to zero, it is quite natural to expect that this condition number is quite large. Moreover, it is also

clear that the value of the condition number increases for increasing accuracy of the discretization. Lastly, since the instability is due to the propagation of the noise corrupting the components associated with small singular values and since these components are associated with singular functions which are, in general, highly oscillating, as remarked above, it should be clear that the generalized solution is characterized by wild oscillations.

If we now consider that, analogous to (31) in the case of infinite-dimensional spaces, in a finite-dimensional framework the set of approximate solution vectors which reproduce the data vector within an uncertainty ϵ is represented by the ellipsoid

$$\sum_{k=1}^p \frac{\sigma_k^2}{\epsilon^2} |(f, u_k)_2 - \frac{(g, v_k)_2}{\sigma_k}|^2 \leq 1 \quad , \quad (34)$$

we find that the center of this ellipsoid is the generalized solution and its half-axes have lengths ϵ/σ_k . Since for a refined discretization the singular values of the matrix become closer and closer to those of the corresponding (compact) operator, it follows that, from some principal directions on, the ellipsoid also contains, together with the true and generalized solutions, approximate solutions characterized by huge norms and therefore completely unreliable. Lastly, we note that the ratio between the lengths of the longest and shortest half-axes is just the condition number.

4 Noise modeling

In the previous section we emphasized the idea that, in the case of an ill-posed problem or of a discrete ill-conditioned problem, one must look for approximate solutions, namely, for objects which do not reproduce the detected image exactly. Indeed, an exact reproduction of the image should be a reproduction not only of the signal contained in the image but also of the noise affecting the signal. In other words one must look for objects approximating the image “within the noise” and it is obvious that this expression becomes more significant if the structure of the noise is known. Therefore in this section we concentrate on the noise models which are most frequently used in applications, we provide a definition of the set of approximate solutions and we ignore, for a moment, the difficulty due to the fact that this set is too broad. Moreover, in this and subsequent sections we continue to consider discrete versions of the imaging problems.

We already stated that the outputs of the detectors are affected by perturbations which are usually denoted as *noise*; randomness is their main feature. As a consequence, the output of a detector must be viewed as the realization of a random variable and, if a measurement is repeated several times, the results will always be different. As a consequence of the ill-conditioning of the imaging matrix A , the solutions of the linear equation

$$A\mathbf{f} = \mathbf{g} \quad , \quad (35)$$

corresponding to different realizations \mathbf{g} of the image are completely different and this is another way of stating that the set of approximate solutions is too broad; indeed the solution associated to one realization is an approximate solution for another realization of the image of the same object.

These remarks indicate that a statistical approach is a quite natural setting for discussing these questions and, in the following, we provide an attempt at quantifying the concept of approximate solution in a general way, for any given noise; it is based on the so-called *likelihood function*, which is related to the randomness of the images. As we show, the least-squares approach, already discussed in the previous section, is obtained as a particular case.

We assume that \mathbf{g} is the realization of a vector-valued random variable (RV) \mathbf{G} and that we know the probability distribution of \mathbf{G} for a given object \mathbf{f} ; for simplicity, we assume that it can be given in terms of a probability density, which is denoted by $P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})$. If a particular realization \mathbf{g} of \mathbf{G} is given and if we insert this value in $P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})$, we obtain a function of \mathbf{f} which is called the *likelihood*, or the *likelihood function*, and is denoted as

$$L_{\mathbf{g}}(\mathbf{f}) = P_{\mathbf{G}}(\mathbf{g}|\mathbf{f}) \quad . \quad (36)$$

A careful discussion of the statistical meaning of this function is beyond the scope of this work; we only observe that, if we consider two objects, \mathbf{f}_1 and \mathbf{f}_2 , and if $L_{\mathbf{g}}(\mathbf{f}_1) > L_{\mathbf{g}}(\mathbf{f}_2)$, then \mathbf{f}_1 is “more likely” than \mathbf{f}_2 to be the object which has generated the image \mathbf{g} . Since, in general, the RVs G_m are independent, so that $P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})$ is the product of a large number of density functions, it is convenient to introduce the logarithm of the likelihood function

$$J_{\mathbf{g}}(\mathbf{f}) = -\ln P_{\mathbf{G}}(\mathbf{g}|\mathbf{f}) \quad . \quad (37)$$

Then we can define a set of approximate solutions as the set of objects with a likelihood greater than a given value (obviously smaller than the maximum value of the likelihood) or, equivalently, as the set of objects defined by the condition

$$S_{\epsilon, \mathbf{g}} = \{\mathbf{f} | J_{\mathbf{g}}(\mathbf{f}) \leq \epsilon\} \quad . \quad (38)$$

In the two examples we discuss in detail, this second definition is preferable, when combined with a suitable rescaling of the function $J_{\mathbf{g}}(\mathbf{f})$, since, in these cases, this function can also be interpreted as the discrepancy between the computed image $A\mathbf{f}$, associated with the object \mathbf{f} and the detected image \mathbf{g} .

Before discussing the two particular models, we point out that in both cases the basic assumption is that the expected value of the RV \mathbf{G} is given by the ideal (computed) image

$$E\{\mathbf{G}\} = A\mathbf{f} \quad . \quad (39)$$

The two models correspond, respectively, to so-called *additive Gaussian noise* and *Poisson noise*.

4.1 Additive Gaussian noise

In this model, the RV \mathbf{G} is given by

$$\mathbf{G} = \mathbf{A}\mathbf{f} + \mathbf{W} \quad , \quad (40)$$

where \mathbf{W} is a Gaussian vector-valued RV. The noise is called additive just because it is a random process which is added to the deterministic signal coming from the object. If all the RVs W_m have zero expected value and if C is their covariance matrix, then the joint probability density of these RVs is given by

$$P_{\mathbf{W}}(\mathbf{w}) = [(2\pi)^M |C|]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (C^{-1} \mathbf{w}, \mathbf{w})_2 \right\} \quad , \quad (41)$$

where $|C|$ is the determinant of the covariance matrix. If $C = \sigma^2 I$, with I the identity matrix, then we have so-called *white noise*. From (40) and (41) we obtain

$$P_{\mathbf{G}}(\mathbf{g}|\mathbf{f}) = [(2\pi)^M |C|]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (C^{-1}(\mathbf{g} - \mathbf{A}\mathbf{f}), \mathbf{g} - \mathbf{A}\mathbf{f})_2 \right\} \quad , \quad (42)$$

and therefore the functional (37), after multiplication by a factor of 2 and addition of a suitable constant, becomes

$$J_{\mathbf{g}}(\mathbf{f}) = (C^{-1}(\mathbf{g} - \mathbf{A}\mathbf{f}), \mathbf{g} - \mathbf{A}\mathbf{f})_2 \quad . \quad (43)$$

In the particular case of white noise, we get

$$J_{\mathbf{g}}(\mathbf{f}) = \|\mathbf{A}\mathbf{f} - \mathbf{g}\|_2^2 \quad , \quad (44)$$

and this is just the discrete version of the least-squares approach discussed in the previous section in a continuous setting. On the other hand the functional (43) is that used in the so-called *weighted* least-squares approach. In other words, from a statistical point of view, all the different forms of least-squares approach derive from a specific assumption on the noise perturbing the data. We remark that these approaches are also the starting points of the regularization theory of ill-posed problems [29].

4.2 Poisson noise

The second model we consider applies to the case of so-called *photon noise*, namely, the noise due to fluctuations in the emission and counting of the photons involved in the imaging process. This noise is relevant both for transmission and emission CT as well as for fluorescence microscopy. The treatment of emission CT and fluorescence microscopy is very similar and, for simplicity, we discuss a model which applies to both cases. The model for transmission CT is discussed, e. g., by Lange and Carson [47].

The basic assumption is that each voxel n is a source of photons. This is a statistical process and we denote by F_n the RV describing the statistical distribution of the number of photons emitted at voxel n and collected by the detectors of the CT scanner or of the microscope during a given acquisition time T . Then the *first basic assumption* is the following:

- F_n is a Poisson RV, with expected value f_n , i. e., the probability of the emission of k photons at voxel n is given by

$$P_{F_n}(k) = \frac{e^{-f_n} f_n^k}{k!} \quad , \quad k = 0, 1, 2, \dots; \quad (45)$$

- the RVs F_n and $F_{n'}$, corresponding to different voxels, are statistically independent.

Next, we denote by $A_{m,n}$ the probability that a photon emitted at voxel n is collected by the detector m . This probability is a crucial quantity in the modeling of the imaging process. Its computation must take into account both the geometry of the acquisition system (for instance, the geometry of the collimating devices) and the physical processes perturbing the photon before arriving at the detector m . In the case of emission tomography, for instance, one must take into account the scattering of the photons by the constituents of the tissues (generating the effects known as attenuation and scatter in PET and SPECT imaging), while, in the case of microscopy, one must consider the diffraction effects. In the latter case the matrix $A_{m,n}$ is given essentially by the PSF of the optical system.

Let $F_{m,n}$ be the RV corresponding to the number of photons emitted at voxel n and collected by the detector m . Then, the *second basic assumption* is the following:

- $F_{m,n}$ is a Poisson RV with expected value given by $A_{m,n} f_n$;
- for any fixed n and $m \neq m'$, the RVs $F_{m,n}$ and $F_{m',n}$ are statistically independent.

If we now denote by G_m the RV corresponding to the number of photons collected by the detector m , and if we assume efficiency 1 (i. e., all the photons arriving at the detector are detected), then it is obvious that G_m is given by

$$G_m = \sum_{n=1^N} F_{m,n} \quad . \quad (46)$$

Thanks to the previous assumptions this RV is also a Poisson process with an expected value given by

$$E\{G_m\} = \sum_{n=1}^N A_{m,n} f_n = (\mathbf{A}\mathbf{f})_m \quad . \quad (47)$$

Moreover, the RVs associated to different detectors are statistically independent. It follows that the probability distribution of the vector-valued RV \mathbf{G} is given by

$$P_{\mathbf{G}}(\mathbf{g}|\mathbf{f}) = \prod_{m=1}^M e^{-(A\mathbf{f})_m} \frac{(A\mathbf{f})_m^{g_m}}{g_m!} , \quad (48)$$

where we denote by \mathbf{g} the set of whole numbers corresponding to the outputs of the detectors.

In the framework of the likelihood approach outlined above, it is easy to see that the functional (37) associated with (48) is equivalent to

$$J_{\mathbf{g}}(\mathbf{f}) = \sum_{m=1}^M \left\{ g_m \ln \frac{g_m}{(A\mathbf{f})_m} + (A\mathbf{f})_m - g_m \right\} , \quad (49)$$

since their difference does not depend on \mathbf{f} . This is the Csiszàr I-divergence which has the properties of a discrepancy functional [25]. It is a convex and non-negative functional. Its level sets can be used for defining the sets of approximate solutions. Their investigation is not easy. However experimental results obtained on the minimum points of this functional indicate that these sets of approximate solutions are also presumably very broad.

We conclude this section with a generalization and refinement of the previous model. In the case of fluorescence microscopy where photons are detected by means of a charge-coupled-device (CCD), in addition to photon noise, described above, one should also take into account so-called *read-out noise* (RON) [64]. This is a white additive Gaussian noise and is statistically independent of the photon noise, so that we have a combination of the two types of noise described above. Since each G_m is the sum of two independent RVs, one with a Poisson distribution and the other with a Gaussian one, it follows that the probability density of the detected signals is given by

$$P_{\mathbf{G}}(\mathbf{g}|\mathbf{f}) = \prod_{m=1}^M \sum_{k=0}^{+\infty} e^{-(A\mathbf{f})_m} \frac{(A\mathbf{f})_m^k}{k!} P_{\text{RON}}(g_m - k) , \quad (50)$$

with

$$P_{\text{RON}}(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(u-r)^2}{2\sigma^2}} \quad (51)$$

if the read-out-noise has expected value r and variance σ^2 .

We remark that the functional (37) derived from this probability density is bounded from below but is not convex. The utility of such an approach has still to be demonstrated; however, its investigation is an interesting mathematical problem.

5 The use of prior information

In the framework of the likelihood approach outlined in the previous section, it is quite natural to consider the object which most likely reproduces the detected image \mathbf{g} as a possible approximate solution of the inverse problem.

This is the *maximum likelihood* (ML) *estimate* and, according to (37), it can be defined by

$$\mathbf{f}_{\text{ML}} = \arg \min_{\mathbf{f}} J_{\mathbf{g}}(\mathbf{f}) \quad . \quad (52)$$

In the case of additive white noise, we again obtain the least-squares problem discussed in Sect. 3. As we know, the solution of this problem (which, in general, is unique since the problem is overdetermined) is ill-conditioned and therefore is affected by strong noise propagation from the data to the solution. The situation is not so clear in the case of photon noise. However there are strong experimental indications that the minimization of the Csiszàr I-divergence (49) also does not provide sound solutions. Indeed the minimum (or the minima; uniqueness is not proved) lies on the boundary of the closed cone of the non-negative vectors and, as a result, minima must have several zero values. This effect appears in the use of iterative methods converging to these minima and is known as *checkerboarding effect*. From these remarks one can draw the conclusion that, if one defines a set of approximate solutions as (38), then this set is too broad. Indeed, it contains both the minima of the Csiszàr I-divergence and the correct solution (if ϵ is correctly chosen) and therefore it contains very different objects.

A very general idea underlying all approaches to the definition of meaningful approximate solutions of inverse problems consists in introducing criteria for extracting these solutions from the broad set of all approximate solutions by means of additional information on the solution itself. This additional information, which is sometimes called *a priori information*, derives from knowledge of expected properties of the solution. For instance, in almost all the problems considered in this chapter, the solutions must be non-negative; we also know that they cannot be too large and so on. This information can be expressed in the form of constraints on the solution of the minimization problems outlined above. For example, constraints on a norm of derivatives of the solution lead, through the method of Lagrange multipliers, to the minimization of functionals which are the sum of the discrepancy and of a regularization functional derived from the smoothness conditions. This is just the basis of Tikhonov regularization theory. In general terms, the problem becomes the minimization of a functional with structure

$$\Phi_{\mathbf{g},\mu}(\mathbf{f}) = J_{\mathbf{g}}(\mathbf{f}) + \mu\Omega(\mathbf{f}) \quad , \quad (53)$$

where $\mu > 0$ is the so-called *regularization parameter* controlling the trade-off between data fitting (the first term) and smoothness of solution (second term).

However we prefer to provide here a probabilistic justification of this approach, based on Bayes formula, which is probably more general than the regularization approach even if its formulation is, in general, restricted to the discrete case in order to avoid excessive mathematical technicalities.

The basic point in this approach is that the object \mathbf{f} is also considered as a realization of a vector-valued RF \mathbf{F} ; moreover, the probability density

$P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})$, introduced in the previous section, is viewed as the *conditional probability of \mathbf{g} , given \mathbf{f}* . Therefore, if the marginal probability density of \mathbf{F} , $P_{\mathbf{F}}(\mathbf{f})$, is also given, then the conditional probability density of \mathbf{F} for a given \mathbf{g} can be obtained by means of *Bayes formula*

$$P_{\mathbf{F}}(\mathbf{f}|\mathbf{g}) = \frac{P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})P_{\mathbf{F}}(\mathbf{f})}{P_{\mathbf{G}}(\mathbf{g})} , \quad (54)$$

where $P_{\mathbf{G}}(\mathbf{g})$ is the marginal probability density of \mathbf{G} , which can be obtained from the joint probability density $P_{\mathbf{FG}}(\mathbf{g}, \mathbf{f}) = P_{\mathbf{G}}(\mathbf{g}|\mathbf{f})P_{\mathbf{F}}(\mathbf{f})$. Equation (54) is the basis of the so-called *Bayesian approach* to inverse problems; the marginal density $P_{\mathbf{F}}(\mathbf{f})$ is usually called the *prior*, while the conditional probability $P_{\mathbf{F}}(\mathbf{f}|\mathbf{g})$ is also called the *a posteriori* conditional probability of the object for a given image. This function provides a complete solution of the inverse problem in the sense of the Bayesian approach. Indeed, from (54) one can compute, in principle, everything about the unknown object corresponding to the detected image: expected value, maximum probability value, probability of subsets of objects, etc.

The difficulty in this approach is that the marginal probability distribution of \mathbf{F} , the prior, is not known, even if in some specific medical applications (e. g., the image of a human organ), one could use data bases of previously obtained images for estimating the prior. This example suggests that the prior is just what is needed for expressing our *a priori information* about the object. In other words we must use our knowledge, or ignorance, about the object for selecting this marginal distribution which restates, in a probabilistic setting, the need, mentioned above, of criteria to be used for selecting meaningful objects from the broad set of all objects compatible with the given image.

However, in any application of inverse problems and, in particular, in medical imaging, it is necessary to show at least one reconstructed object and this can be provided by the *maximum a posteriori* (MAP) estimate, which is an object maximizing the a posteriori conditional probability

$$\mathbf{f}_{MAP} = \arg \max_{\mathbf{f}} P_{\mathbf{F}}(\mathbf{f}|\mathbf{g}) . \quad (55)$$

It is possible to replace this problem with a minimization problem by taking the logarithm of the a posteriori density and changing its sign, so that, on recalling the definition (37), we obtain

$$\mathbf{f}_{MAP} = \arg \max_{\mathbf{f}} \{J_{\mathbf{g}}(\mathbf{f}) - \log P_{\mathbf{F}}(\mathbf{f})\} , \quad (56)$$

where we have neglected the contribution of $P_{\mathbf{G}}(\mathbf{g})$ since it is independent of \mathbf{f} . Therefore the term $-\log P_{\mathbf{F}}(\mathbf{f})$ plays the role of a regularization functional.

The most frequently used priors are of Gibbs type

$$P_{\mathbf{F}}(\mathbf{f}) = C \exp\{-\mu\Omega(\mathbf{f})\} , \quad (57)$$

where $\Omega(\mathbf{f})$ is, in general, a convex and non-negative functional expressing prior information about the object and μ is a positive parameter (which is just

the regularization parameter in regularization theory). In such a case we find that the functional to be minimized for determining the MAP estimate is just that given in (53). In particular we obtain precisely the standard functional of Tikhonov regularization theory if we assume that the image is perturbed by additive white noise, so that the discrepancy functional is given by (44), and also that the prior of the object corresponds to a white Gaussian process with zero expected value and variance $1/2\mu$, i.e., $\Omega(\mathbf{f}) = \|\mathbf{f}\|_2^2$. As a result the functional (53) becomes

$$\Phi_{\mathbf{g},\mu}(\mathbf{f}) = \|\mathbf{A}\mathbf{f} - \mathbf{g}\|_2^2 + \mu\|\mathbf{f}\|_2^2. \quad (58)$$

As it is well known, for each value of μ this functional has a unique minimum which is just the classical Tikhonov regularized solution

$$\mathbf{f}_\mu = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T \mathbf{g}, \quad (59)$$

where \mathbf{I} denotes the identity matrix. This regularized solution, however, is not frequently used in medical imaging, firstly because it is not suitable for the solution of large-scale problems, as we discuss in the next section and, secondly because, in the case of tomography, it is affected by aliasing effects, as discussed, e.g., in [14].

6 Computational issues and reconstruction methods

If a linear model is available, then one has at one's disposal the very powerful theoretical tools outlined in the previous Sections for the analysis and the solution of the problem. However its practical solution can require a considerable computational burden both in the 2D and in the 3D case. If the problem is treated as a sequence of 2D problems as in standard X-ray CT and also, in general, in emission tomography, then, for each section, the number of unknowns is 256×256 or 512×512 . If the problem is genuinely 3D, then the number of unknowns becomes $256 \times 256 \times 64$ or $512 \times 512 \times 64$ and therefore it is of the order of millions. In such a situation it is clear that the discretization of (11) leads to a matrix which, in general, cannot be stored, even if it is sparse. Therefore further approximations are, in general, introduced in the model in order to make the problem tractable from the numerical point of view.

For instance, in the case of SPECT imaging, one neglects the so-called "collimation blur". The discrimination of the photons coming from a given direction is obtained by means of a hole in a slab; the detector counts all the photons crossing the hole and therefore integrates over the acceptance cone of this hole. It follows that the problem is moderately 3D. However, if one assumes that it is reasonable to approximate the cone with a circular cylinder, one has a situation close to that of X-ray CT (integration over straight lines or, more precisely, over tubes). In such a case the problem can be approximated

by a sequence of 2D problems (as in CT) and one can use the fast algorithm of filtered back-projection. A more refined model, also leading to a sequence of 2D problems, is the so-called 2D+1 model developed in [8]. However, in such a case, the matrix does not come from the discretization of the Radon transform and, since it is too large, it must be computed whenever it is required. It is also worth mentioning the case of magnetic resonance (MR), where the acquisition process is designed in such a way that the reconstruction process can be reduced to Fourier transform inversion and therefore is extremely fast (and well-posed as well). For an early but excellent tutorial on MR we suggest the paper of Hinshaw and Lent [35], while, for a good description of the physics of radioisotope imaging such as PET and SPECT, see [69].

Another example is provided by fluorescence microscopy. Given that, in such a case, the PSF in (11) describes the spatial dependence of the point process, this dependence is, in general, space variant, i.e., it is not the same for all locations of the point source in the object space. Such an effect is a consequence, for instance, of the spherical aberration of the lenses of the optical system. However, the manufacturer attempts to correct this effect as far as possible. As a consequence, it is reasonable to assume that the PSF is space-invariant so that (11) can be replaced by (6). Moreover, practitioners approximate the convolution operator by means of a 3D circulant matrix, so that the Fast Fourier Transform (FFT) can be used for the computation of the matrix. It is also obvious that the storage of the matrix can be reduced to the storage of the PSF and therefore is just that of one image. If the condition of space-invariance is not satisfied, then one in general assumes that it is satisfied in subdomains of the image volume so that in each of them the reconstruction techniques developed for the space-invariant case can be used.

In general it is assumed that the problem can be solved in almost real time if one of the following conditions is satisfied:

- the 3D problem can be approximated by a sequence of 2D problems, each implying a Radon transform inversion;
- the matrix A is sparse and is not stored since it can be computed by means of a look-up table of given values or by means of simple rules;
- the matrix A is not sparse but is given by a space-invariant PSF which can be stored.

In the first case the standard algorithm is *filtered back-projection* (FBP) while, in the other cases, the most frequently used approaches are based on iterative algorithms with regularization properties. For the convenience of the reader we briefly describe FBP, also because it provides terminology which is frequently used in medical imaging.

As discussed in the introduction the 2D Radon transform is given by $(Rf)(s, \theta) = (P_\theta f)(s)$, with $(P_\theta f)(s)$ defined in (3). The inversion formula (5) can be decomposed in two steps: the first is the computation of the Hilbert transform of the derivative, with respect to s , of the Radon transform of f ; the second consists in applying to the result the *back-projection operator* defined

by

$$(R^\#g)(x) = \int_{S^1} g(x \cdot \theta, \theta) d\theta . \quad (60)$$

This operator is, in a sense, the dual of the Radon operator R because, while R corresponds to integrating over the points of a line, $R^\#$ corresponds to integrate over the lines through a point. It is also the (formal) adjoint of R and, for this reason, in medical imaging it is usual to denote the imaging matrix A , introduced in the previous sections, as the *projection matrix* and the matrix A^T as the *backprojection matrix*.

Now, the FBP algorithm consists of the following two steps, corresponding to the two steps indicated above.

- Step 1 (filtering) . Compute the 1D Fourier transform of each projection $g_\theta = (Pf)_\theta(s)$, namely, $\hat{g}_\theta(\omega)$, multiply the result by the ramp filter $|\omega|$ and take the inverse Fourier transform; the result is the filtered projection in the direction θ :

$$G_\theta(s) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\omega| \hat{g}_\theta(\omega) e^{is\omega} ds . \quad (61)$$

This step is just the computation, except for a constant, of the Hilbert transform of the derivative of g_θ . The set of filtered projections provides the filtered sinogram, which is the representation of the 2D function $G(s, \theta) = G_\theta(s)$.

- Step 2 (back-projection) . This step is just the application of the back-projection operator to the filtered projections; it provides the function f :

$$f(x) = \frac{1}{4\pi} (R^\#G)(x) . \quad (62)$$

It is easy to verify that this is a different way of writing (5).

In Figure 3 we give a pictorial representation of the FBP algorithm, also showing that, if we apply the back-projection operator directly to the projections (without the ramp filter), then we get a blurred version of the object. In this figure the sharpening of the sinogram provided by the ramp filter is also evident.

It is obvious that the FBP algorithm, here described in a continuous setting, allows for fast implementations. Indeed the computation of the filtered projections can be performed by means of the FFT algorithm and therefore its computational cost is of the order of $M \log_2 M$, if M is the number of data. More expensive is the computation of the back-projection even though fast algorithms have also been designed in this case (see, e.g., [57]). We also note that the ill-posed nature of the problem manifests itself in the multiplication of $\hat{g}_\theta(\omega)$ by the ramp filter. Indeed this filter amplifies the high-frequency noise, i.e., the noise affecting the values of $\hat{g}_\theta(\omega)$ for large values of ω . This effect is corrected, in practice, by attenuating the ramp filter at the higher

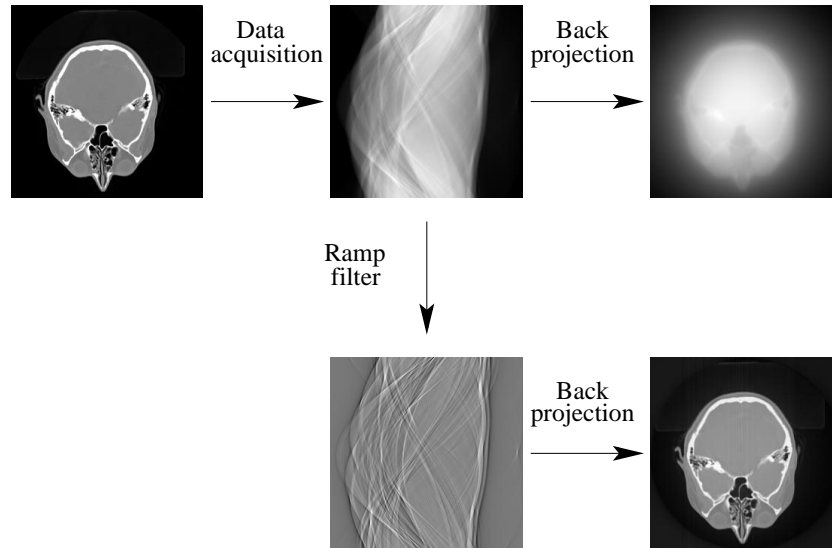


Fig. 3. Pictorial representation of the FBP algorithm.

frequency. An example is provided by the *Shepp-Logan filter* [42] and such a procedure is basically a regularizing one.

The efficiency of the FBP algorithm makes clear why it is the favorite in commercial machines: very often the acquisition processes are designed in such a way that the data approximately provides line integrals of the unknown object. If this approximation is not satisfactory or if one intends to improve the results provided by FBP, then one has to deal with a large-scale projection matrix A and one must solve one of the large-scale minimization problems discussed in the previous sections. It is obvious that direct methods such as those provided by the Tikhonov regularized solution (59) are not feasible in practice. Therefore it is quite natural to look for iterative methods such that at each iteration the main computational burden is a matrix-vector multiplication.

It is also important to observe that the minimization of regularized functionals, such as those introduced in the previous Section, faces the problem of the choice of regularization parameter. Several criteria have been investigated in the case of Tikhonov regularization method (see, Engl et al. [29], Bertero and Boccacci [4]) but not in the other cases. In general one should do experiment with the method on sets of simulated images for establishing the best possible values of these parameters. Such an approach is very costly from the computational point of view.

A practical solution to these problems is provided by iterative methods for the minimization of a discrepancy functional, such as that defined in (44) or (49), with a property which is called *semiconvergence*: if we define a restora-

tion error as a distance between the result of iteration k and the true object, then this restoration error first decreases, goes through a minimum and then increases up to very large values (remember that the minima of the discrepancy functionals are strongly affected by noise propagation). It turns out that the solution provided by the iteration corresponding to the minimum of the restoration error is, in general, a reliable solution of the reconstruction problem. In such a case, we do not have the problem of choice of regularization parameter, but rather a problem of optimal stopping. Again criteria can be obtained by experimenting with the algorithm.

In the case of the least-squares functional (44) the prototype of these methods is a gradient method known, in the inverse problem literature, as the *Landweber method*:

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \tau A^T (\mathbf{g} - A\mathbf{f}^{(k)}) \quad , \quad (63)$$

where τ is a relaxation parameter (the step in the direction of steepest descent) satisfying the condition

$$0 < \tau < \frac{2}{\sigma_1^2} \quad , \quad (64)$$

where σ_1 is the maximum singular value of the matrix A . It is easy to prove (see, e.g., [4]) that this iterative method, initialized with the zero object, converges to the minimum norm least-squares solution and that the effect of the iteration k is essentially a filtering of this solution, keeping the components corresponding to the largest singular values and dropping the others: as the iteration goes on, the components corresponding to smaller singular values also appear, hence introducing the instability due to noise amplification. By such an analysis it is possible to prove the semiconvergent behavior of the restoration error.

The main drawback of this method is that it is too slow: the minimum of the reconstruction error can be reached only after hundreds of iterations. A much more convenient method is provided by the *conjugate gradient* (CG). It is proved that this method has also the semiconvergence property (see, e.g., Engl et al. [29]), but it is much faster than the Landweber method, especially if preconditioning techniques are used. For instance, it was shown, in an application to SPECT imaging [8], that CG, equipped with a very simple preconditioner, provides reliable solutions after 10 to 15 iterations.

However, in the case of emission tomography or fluorescence microscopy, the least-squares discrepancy (44) is not the appropriate one and one must use that derived from the appropriate noise model, namely, (49). An iterative method for the minimization of this discrepancy in the case of emission tomography was proposed by Shepp and Vardi [63]. It is called *expectation maximization* (EM) since it is a particular case of a general method, with the same name, for the solution of maximum likelihood problems. It must be remarked that the same iterative algorithm was previously proposed by Richardson [60] and Lucy [50] for deconvolution problems and, for this reason,

the algorithm is known as the *Richardson-Lucy method* (RLM) in astronomy and microscopy. The algorithm is as follows

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} A^T \frac{\mathbf{g}}{A\mathbf{f}^{(k)}} , \quad (65)$$

where product and quotient of vectors are to be interpreted as component-wise (Hadamard product and quotient). We note an interesting property of this algorithm: since the elements of the matrix A and the components of the image \mathbf{g} are non-negative, if the initial guess is also non-negative, then all the iterates are automatically non-negative. It must also be pointed out that, in general, a positive initial guess is chosen (as a rule of thumb, a constant vector), because, as a consequence of the multiplicative structure of the algorithm, if a component of the initial guess is zero, then the same component of all the iterates is zero.

It was proved by Shepp and Vardi (a more complete proof is given by Lange and Carson [47]) that, for any positive initial guess, the iterates converge to a maximum of the likelihood function, hence, to a minimum of the Csiszàr I-divergence. However, after a number of iterations, the iterates show the checkerboard effect mentioned earlier, indicating that the minima of the Csiszàr I-divergence are not reliable solutions. The utility of the algorithm is due to the fact that it has the semiconvergence property (see [4] for a discussion), an experimental result derived from numerical practice. Some theoretical insight is derived from an analysis of the filtering effect of the iterations [58]. As a consequence reliable solutions can be obtained by suitable stopping of the iteration.

EM, as Landweber, is very slow and, in general, requires a large number of iterations. In the case of emission tomography, an accelerated version known as *ordered subset - expectation maximization* (OS-EM) was proposed by Hudson and Larkin [36]. This approach improves considerably the efficiency of EM; it was implemented for the reconstruction of SPECT data based on the 2D+1 model [8] and it provides reconstructions in almost real time (a few minutes) so that it is currently used by medical doctors of the Universities of Genoa and Florence for the reconstruction and analysis of SPECT data.

In Figure 4 we give an example of the improvement which can be obtained by means of this method with respect to FBP, by comparing the results obtained with the two methods in the reconstruction of a 3D brain image. The comparison is performed by showing the results for a transaxial section (upper panels), a coronal section (middle panels) and a sagittal section (lower panels). The left-hand panels correspond to FBP. Since the FBP reconstruction is simply a set of 2D reconstructions, a smoothing in the axial direction is needed and this is obtained by means of a suitable Gaussian filter. In the right-hand panels we show the result of a 3D reconstruction, based on the 2D+1 model for the collimator blur, obtained by grouping the projections into 12 subsets and using 9 OS-EM iterations. The improvement is evident not only in the coronal and sagittal sections, as is obvious since we use a 3D

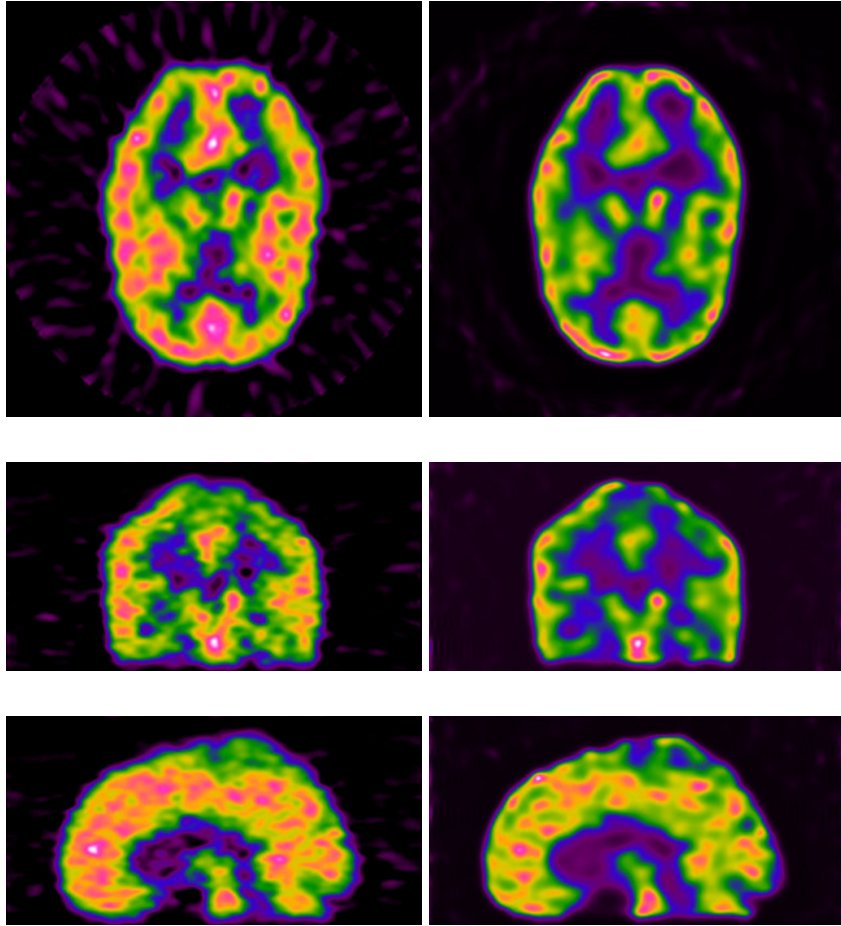


Fig. 4. Left-hand panels: reconstructions provided by the FBP algorithm, using a Gaussian filter in the axial direction. Right-hand panels: reconstructions obtained by means of the 2D+1 model for the collimator blur and the OS-EM algorithm for data inversion. Upper panels: transaxial sections. Middle panels: coronal sections. Lower panel: sagittal sections.

method, but also in the transaxial section. This improvement is mainly due to the improved model of the imaging matrix.

7 Perspectives

The different kinds of tomography and microscopy, as well as MR, represent well-established imaging techniques, characterized by a high degree of tech-

nological development and validated through a large variety of biomedical applications. Nevertheless, recent decades have seen an impressive growth of the investigation of new modalities able to provide information on both the structural characteristics and the functional status of the tissues under examination. The theoretical modeling of these procedures often utilizes sophisticated mathematical tools, thus providing actual motivations to investigations in important areas of pure mathematics. On the other hand, the necessity of reliable reconstruction methods on a real-time scale, has inspired the formulation of numerical algorithms of a more general flavor and has significantly contributed to the development of a new applied science, characterized by the integration of concepts and techniques originally formulated in different areas of mathematics, physics, biomedicine and computer science. In the following we provide outlines of some of these new imaging techniques, indicating both their roles as diagnostic tools and the main theoretical connections with different aspects of applied mathematics.

7.1 Electrical Impedance Tomography

The inverse problem at the basis of *electrical impedance tomography* (EIT) is that of imaging the electrical properties in the interior of a body given measurements of electric currents and voltages at the boundary. Medical problems for which knowledge of internal electrical properties are helpful are related to: lung medicine, for the detection of pulmonary emboli or blood clots in the lung; non-invasive monitoring of heart functions; non-invasive detection of breast cancer. Within the highly idealized framework of the continuum model [16], the electric potential u in the domain Ω satisfies the boundary value problem:

$$\nabla \cdot \gamma(x, \omega) \nabla u = 0 \quad , \quad x \in \Omega \quad , \quad (66)$$

$$\gamma(x) \frac{\partial u(x)}{\partial \nu} = j(x) \quad , \quad x \in \partial \Omega \quad , \quad (67)$$

where

$$\gamma(x, \omega) = \sigma(x, \omega) + i\omega\epsilon(x, \omega) \quad (68)$$

is the admittivity (proportional to the inverse of the impedance of the body), σ the electric conductivity, i the imaginary unit, ω the angular frequency of the applied current, ϵ the electric permittivity, j the surface current density and ν the inward-pointing normal unit vector. This Neumann problem is well-posed under very general conditions on γ , provided that the charge conservation is assured and a choice for $\int_{\partial \Omega} u$ is fixed. The corresponding inverse problem is much more complicated and can be related to that of determining γ when the Dirichlet-to-Neumann map relating the voltage and the current on the surface is known. Many important mathematical issues arise from this problem, essentially concerned with both the injectivity of the forward map and the stability of the inverse operator [46, 67, 55]. In [13] an imaging algorithm is provided for the linearized problem, which was numerically tested

in [37] in the case of an elementary geometry. However, in [17] and [65] it is pointed out that the continuum model is very poor for real experiments, where the current is known only in wires attached to discrete electrodes and significant electrochemical effects at the contact between the electrodes and the body must be taken into account. On the basis of a more realistic and complete model, an adaptive current tomography system has been designed and realized by the EIT group at the Rensselaer Polytechnic Institute [16] and reconstruction algorithms have been implemented and tested in the case of both simulated and real experiments.

7.2 Optical Tomography

Optical tomography (OT) [2] is a functional medical imaging modality whose aim is to reconstruct images of optical coefficients, such as tissue absorption, from boundary measurements of light transmission by using light in the infrared band between 700 and 1000 nm. The medical motivations of this technique [15] are essentially related to brain function monitoring, as in the case of the early detection of after-birth oxygen deficiency or localization of cortical activation areas during stimulated tasks, although structural reconstructions can be obtained in the case of the detection of brain and breast cancer. According to a very schematic description, the mathematical model for OT is based on the integro-differential equation for photon transport theory [39], which, in the frequency domain, reduces to the elliptic equation [48]:

$$\nabla \cdot (A \nabla \hat{\Phi}) - (B + iC) \hat{\Phi} = \hat{q} \quad . \quad (69)$$

Here $\hat{\Phi}$ and \hat{q} are, respectively, the Fourier transform of the isotropic photon density Φ and of a function q of space and time representing the source distribution; A and B are coefficients related to the tissue absorption and scattering properties and C is the reciprocal of the speed of light. The inverse problem of OT is that of determining estimates of A and B , given C and complete data at the boundaries for two different frequencies. A proof of uniqueness of the solution can be found in [38] while a rather complete review of the most frequently used reconstruction methods is in [2]. A prototype scanner for OT has been constructed at the Biomedical Optics Research Laboratory, University College, London (<http://www.medphys.ucl.ac.uk/research/borg/index.htm>). The primary clinical aim of this device is to detect oxygenation failures in newborn infant brains.

7.3 Microwave Tomography

Microwave tomography (MT) is a structural imaging technique whose aim is to reconstruct the electrical parameters of the tissues from measurements of the scattered electromagnetic field in the gigahertz frequency range [53]. This low-resolution, high-contrast methodology is particularly helpful both when a

low invasivity is required, as in mammography [30], and when the cancerous tissues provoke significant variations of the refractive index, as in the diagnosis of leukemia in human bone marrow [1]. From a mathematical point of view, in a two-dimensional setting MT can be described by the scattering problem introduced in Sect. 2 (Eq. (12)-(15)) and, in three dimensions, by an inverse problems for Maxwell equations [11]. In MT no linearization can be performed, since the microwave wavelength in biological tissues is of some centimeters and therefore genuine resonance conditions hold. This implies that most image reconstruction methods in this framework must address a notable computational expense which is often reduced by adopting oversimplified approximations in the model. Typical approaches are based on generalization of methods originally formulated for diffraction tomography [68], on Newton-Kantorowich algorithms [40] and on gradient methods [45]. Important experimental results have been obtained at the Biophysical Laboratory, Carolinas Medical Center, Charlotte, NC (USA) for various biomedical applications [62]. At the Department of Biocybernetics, Niigata University, Niigata, an MT prototype has been realized, with the aim of obtaining hardware-type linearization of the image restoration problem by means of a mixing/filtering procedure [52]. In this application, named *chirp-pulse microwave computerized tomography* (CP-MCT), the input is a chirp signal characterized by increasing frequency, typically from 1 GHz to 2 GHz, the geometry adopted is that of X-ray tomography in its parallel or fan beam setup and the basic reconstruction algorithm is filtered back-projection. A refined mathematical model for image formation in CP-MCT was proposed in [7] and validated in [54] while a theoretical computation of the PSF of the device by means of scattering theory methods was performed in [5]. Applications of this technique to the reconstruction of real simple objects and to simulated mammographical experiments are currently in progress.

7.4 Magnetoencephalography

Magnetoencephalography (MEG) [34] is a brain-imaging technique measuring the weak magnetic fields related to neural currents, with the aim of investigating cerebral functional behavior, particularly in the study of auditory and visual cortex, for both healthy subjects and for patients affected by epilepsy or developmental dyslexia. When compared to other functional modalities such as PET, SPECT or fMRI, MEG is characterized by very low invasivity and by a high temporal resolution, which allows us to infer information about the temporal hierarchy of physiological responses. Furthermore current MEG scanners provide a spatial resolution of some millimeters, particularly for sources in the brain cortex, and reliable coregistration procedures with MR structural data allows us to reproduce even finer anatomical details. The MEG forward problem [61] is that of computing the magnetic field $\mathbf{B}(\mathbf{r})$ of a given continuous current distribution $\mathbf{J}(\mathbf{r})$ when $\mathbf{B}(\mathbf{r})$ and $\mathbf{J}(\mathbf{r})$ are related by the Biot-Savart law

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\Omega} \mathbf{J}(\mathbf{r}') \times \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} d\mathbf{r}' \quad , \quad (70)$$

and Ω is a bounded domain corresponding to the human brain. The MEG inverse problem is ill-posed in the sense of Hadamard and non-linear. Being ill-posed is a consequence of the fact that the integral operator relating the current density and magnetic field has a non-empty kernel [31] while non-linearity arises when one models the neuronal source as a current dipole and wants to determine both its position and amplitude. Standard approaches to the solution of this problem need to deal with time series characterized by a sufficiently high signal-to-noise ratio, and this result is typically accomplished by averaging several magnetic responses to repeated realizations of the same external stimulus. However, the main drawback of this procedure is that the averaging process strongly dissipates the temporal information content of the MEG signal. Furthermore, averaging such signals would be reliable if the stimulus responses were the same for each repetition, but this is not necessarily true, since, in order to avoid habituation effects, typical stimulus realizations involve small modifications of the stimulus characteristics at each repetition. More effective approaches utilizing raw non-averaged time series are based on a probabilistic Bayesian approach [66], whereby the source localization at each sampled time is obtained by building up a density probability function conditioned on the experimental measurements, and the time evolution of the process is realized by exploiting the form of the density function constructed at the previous time step.

Acknowledgement. We thank our colleagues Patrizia Boccacci and Piero Calvini for providing some of the images displayed in this chapter and for discussing related topics.

References

- [1] Albanese, R. A., Medina, R. L., Penn, J. W.: Mathematics, medicine and microwaves. *Inverse Problems* **10**, 995–1007 (1994)
- [2] Arridge, S. R.: Optical tomography in medical imaging. *Inverse Problems* **15**, R41-R93 (1999)
- [3] Bertero, M.: Linear inverse and ill-posed problems. In: Hawkes, P. W. (ed) *Advances in electronics and electron physics*. Vol. 75, New York: Academic 1989, pp. 1-120
- [4] Bertero, M., Boccacci, P.: *Introduction to inverse problems in imaging*. Institute of Physics: Bristol 1998
- [5] Bertero, M., Conte, F., Miyakawa, M., Piana, M.: Computation of the response function in chirp-pulse microwave computerized tomography. *Inverse Problems* **17**, 485-500 (2001)
- [6] Bertero, M., De Mol, C., Viano, G.: The stability of inverse problems. In: Baltes, H. P. (ed) *Inverse scattering Problems in Optics*. Berlin: Springer 1980, pp. 161-204

- [7] Bertero, M., Miyakawa, M., Boccacci, P., Conte, F., Orikasa, K., Furutani, M.: Image restoration in chirp-pulse microwave CT (CP-MCT). *IEEE Trans. Biomed. Eng.* **47**, 690–609 (2000)
- [8] Boccacci, P., Bonetto, P., Calvini, P., Formiconi, A. R.: A simple model for the efficient correction of collimator blur in 3D SPECT imaging. *Inverse Problems* **15**, 907-930 (1999)
- [9] Bracewell, R. N., Riddle, A. C.: Inversion of fan-beam scans in radio astronomy. *Astrophys. J.* **150**, 427-434 (1967)
- [10] Brühl, M., Hanke, M.: Numerical implementation of two noniterative methods for locating inclusions by impedance tomography. *Inverse Problems* **16**, 1029-1042 (2000)
- [11] Bulyshev, A. E., Souvorov, A. E., Semenov, S. Y., Svenson, R. H., Nazarov, A. G., Sizov, Y. E., Tatsis, G. P.: Three-dimensional microwave tomography. Theory and computer experiments in scalar approximation. *Inverse Problems* **16**, 863-875 (2000)
- [12] Cakoni, F., Colton, D.: On the mathematical basis of the linear sampling method. *Georgian Math. J.* **10**, 411-425 (2003)
- [13] Calderon, A. P.: On an inverse boundary value problem, Seminar on Numerical Analysis and Its Applications to Continuum Physics. Rio de Janeiro: Soc. Brasileira de Matematica 1980, pp. 65-73
- [14] Caponnetto, A., Bertero, M.: Tomography with a finite set of projections: singular value decomposition and resolution. *Inverse Problems* **13**, 1191-1205 (1997)
- [15] Chance, B., Delpy, D. T., Cooper, C. E., Reynolds, E. O. R. (eds.): Near-infrared spectroscopy and imaging of living systems. *Philos. Trans. Roy. Soc. London ser. B* **352**, 643-761 (1997)
- [16] Cheney, M., Isaacson, D., Newell, J. C.: Electrical Impedance Tomography. *SIAM Rev.* **41**, 85-101 (1999)
- [17] Cheng, K. S., Isaacson, D., Newell, J. C., Gisser, D. G.: Electrode models for electric current computed tomography. *IEEE Trans. Biomed. Engr.* **36**, 918-924 (1989)
- [18] Colton, D., Coyle, J., Monk, P.: Recent developments in inverse acoustic scattering theory. *SIAM Rev.* **42**, 369-414 (2000)
- [19] Colton, D., Haddar, H., Piana, M.: The linear sampling method in electromagnetic scattering theory. *Inverse Problems* **19** S105–S138 (2003)
- [20] Colton, D., Kirsch, A.: A simple method for solving inverse scattering problems in the resonance region. *Inverse Problems* **12**, 383-393 (1996)
- [21] Colton, D., Kress, R.: *Inverse acoustic and electromagnetic scattering theory*. 2nd ed. Berlin: Springer 1998
- [22] Colton, D., Monk, P.: A linear sampling method for the detection of leukemia by using microwaves. *SIAM J. Appl. Math.* **58**, 926-941 (1998)
- [23] Colton, D., Piana, M., Potthast, R.: A simple method using Morozov's discrepancy principle for solving inverse scattering problems. *Inverse Problems* **13**, 1477-1493 (1997)
- [24] Courant, R., Hilbert, D.: *Methods of mathematical physics. II. Partial differential equations*. New York: Interscience 1962
- [25] Csiszár, I.: Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Ann. Statist.* **19**, 2032-2066 (1991)

- [26] Davison, M. E.: A singular value decomposition for the Radon transform in n -dimensional euclidean space. *Numer. Funct. Anal. Optim.* **3**, 321-340 (1981)
- [27] Devaney, A. J.: Diffraction tomography. In: W. M. Boerner et al. (Eds.): *Inverse methods in electromagnetic imaging, Part 2*. Boston: Reidel 1985, pp. 1107-1136
- [28] Diaspro, A. (ed): *Confocal and two-photon microscopy: foundations, applications, and advances*. New York: Wiley 2001
- [29] Engl, H. W., Hanke, M., Neubauer, A.: *Regularization of Inverse Problems*. Dordrecht: Kluwer 1996
- [30] Fear, E. C., Li, X., Hagness, S. C., Stuchly, M. A.: Confocal microwave imaging for breast cancer detection: localization of tumors in three dimensions. *IEEE Trans. Biomed. Eng.* **49** 812-823 (2002)
- [31] Fokas, A. S., Kurylev, Y., Marinakis, V.: The unique determination of neuronal currents in the brain via magnetoencephalography, *Inverse Problems* **20**, 1067-1082 (2004)
- [32] Goodman, J. W.: *Introduction to Fourier optics*. San Francisco: McGraw-Hill 1968
- [33] Hadamard, J.: *Lectures on Cauchy's problem in partial differential equations*. New Haven: Yale University Press 1923
- [34] Hamalainen, M., Hari, R., Ilmoniemi, R., Knuutila, J., Lounasmaa, O. V.: Magnetoencephalography - theory, instrumentation and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* **65**, 413-497 (1993)
- [35] Hinshaw, W. S., Lent, A. H.: An introduction to NMR imaging: from the Bloch equation to the imaging equation. *Proc. IEEE* **71**, 338-350 (1983)
- [36] Hudson, H. M., Larkin, R. S.: Accelerated image reconstruction using ordered subsets of projection data, *IEEE Trans. Med. Imag.* **13**, 601-609
- [37] Isaacson, D., Isaacson, E.: Comments on Calderon's paper: 'On an inverse boundary value problem'. *Math. Comp.* **52**, 553-559 (1989)
- [38] Isakov, V.: *Inverse problems for partial differential equations*. New York: Springer 1998
- [39] Ishimaru, A.: *Wave propagation and scattering in random media*. New York: Academic Press 1978
- [40] Joachimowicz, N., Pichot, C., Hugonin, J. P.: Inverse scattering: an iterative numerical method for electromagnetic imaging. *IEEE Trans. Antennas Propagat.* **39**, 1742-1753 (1991)
- [41] Kaipio, J. P., Somersalo, E.: *Statistical and computational inverse problems*. Berlin: Springer 2004
- [42] Kak, A. C., Slaney, M.: *Principles of computerized tomographic imaging*. Philadelphia: SIAM 2001
- [43] Kato, T.: *Perturbation theory for linear operators*. Berlin: Springer 1980
- [44] Kirsch, A.: Characterization of the shape of a scattering obstacle using the spectral data of the farfield operator. *Inverse Problems* **14**, 1489-1512 (1998)
- [45] Kleinman, R. E., van den Berg, P. M.: A modified gradient method for two-dimensional problems in tomography. *J. Comput. Appl. Math.* **42**, 17-35 (1992)
- [46] Kohn, R. V., Vogelius, M.: Determining conductivity by boundary measurements. *Comm. Pure Appl. Math.* **37**, 289-298 (1984)

- [47] Lange, K. L., Carson, R.: EM reconstruction algorithms for emission and transmission tomography. *J. Comput. Assist. Tomogr.* **8**, 306-316 (1984)
- [48] Lionheart, W. R. B, Arridge, S. R., Schweiger, M., Vauhkonen, M., Kaipio, J. P.: Electrical impedance and diffuse optical tomography reconstruction software. In: *Proceedings of the 1st World Congress on Industrial Process Tomography*. Wilmslow, UK: Process Tomography 1999, pp. 474-479
- [49] Louis, A. K.: Laguerre and computerized tomography: consistency conditions and stability of the Radon transform. In: Brezinsky C. et al. (Eds): *Polynomes orthogonaux et applications*. (Lecture Notes in Math. **1171**) Berlin: Springer 1985, pp. 524-531
- [50] Lucy, L.: An iterative technique for the rectification of observed distributions. *Astron. J.* **79**, 745-754 (1974)
- [51] Mazzone, F., Coyle, J., Massone, A. M., Piana, M.: FIST: A fast visualizer for fixed-frequency acoustic and electromagnetic inverse scattering problems. *Simulation Modeling Practice Theory* **14**, 177-187 (2006)
- [52] Miyakawa, M.: Tomographic measurements of temperature change in the phantoms of the human body by chirp radar-type microwave computed tomography. *Med. Biol. Eng. Comput.* **31**, S31-S36 (1993)
- [53] Miyakawa, M., Bolomey, J. C.: *Non-invasive thermometry of human body*. Boca Raton: CRC Press 1996
- [54] Miyakawa, M., Orikasa, K., Bertero, M., Boccacci, P., Conte, F., Piana, M.: Experimental validation of a linear model for data reduction in chirp-pulse microwave CT. *IEEE Trans. Med. Imag.* **21**, 385-395 (2002)
- [55] Nachman, A. I.: Global uniqueness for a two-dimensional inverse boundary value problem. *Ann. of Math.* **143**, 71-96 (1996)
- [56] Natterer, F.: *The mathematics of computerized tomography*. Philadelphia: SIAM 2001
- [57] Natterer, F., Wübbeling, F.: *Mathematical methods in image reconstruction*. Philadelphia: SIAM 2001
- [58] Prasad, S.: Statistical-information-based performance criteria for Richardson-Lucy image deblurring. *J. Opt. Soc. Amer. A* **19**, 1286-1296 (2002)
- [59] Radon, J.: Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Ber. Verh. Königl. Sächs. Ges. Wiss., Leipzig, Math.-Phys. Kl.*, **69**, 262-277 (1917)
- [60] Richardson, W. H.: Bayesian-based iterative method of image restoration. *J. Opt. Soc. Amer.* **62**, 55-59 (1972)
- [61] Sarvas, J.: Basic mathematical and electromagnetic concepts of the bio-magnetic inverse problem. *Phys. Med. Biol.* **32**, 11-22 (1987)
- [62] Semenov, S. Y., Svenson, R. H., Boulyshev, A. E., Souvorov, A. E., Borisov, V. Y., Sizov, Y., Starostin, A. N., Dezern, K. R., Tatsis, G. P., Baranov, V. Y.: Microwave tomography: two-dimensional system for biological imaging. *IEEE Trans. Biomed. Eng.* **43**, 869-877 (1996)
- [63] Shepp, L. A., Vardi, Y.: Maximum likelihood reconstruction for emission tomography, *IEEE Trans. Med. Imag.* **1**, 113-122 (1982)
- [64] Snyder, D. L., Hammoud, A. M., White, R. L.: Image recovery from data acquired with a charge-coupled-device camera. *J. Opt. Soc. Amer. A* **10**, 1014-1023 (1993)

- [65] Somersalo, E., Cheney, M., Isaacson, D.: Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.* **52**, 1023-1040 (1992)
- [66] Somersalo, E., Voutilainen, A., Kaipio, J. P. : Non-stationary magnetoencephalography by Bayesian filtering of dipole models. *Inverse Problems* **19**, 1047-1063 (2003)
- [67] Sylvester, J., Uhlmann, G.: A uniqueness theorem for an inverse boundary value problem in electrical prospection. *Comm. Pure Appl. Math.* **39**, 91-112 (1986)
- [68] Tabbara, W., Duchêne, B., Pichot, C., Lesselier, L., Chommeloux, L., Joachimowicz, N.: Diffraction tomography: Contribution to the analysis of some applications in microwaves and ultrasonics. *Inverse Problems* **4**, 305-331 (1988)
- [69] Webb, S.: *The physics of edical imaging*. Bristol: Institute of Physics 1988
- [70] Weinstein, M., Castleman, K. R.: Reconstructing 3-D specimens from 2-D section images. In: Herron R. E. (ed) *Quantitative imagery in the biomedical sciences I*. (Proceedings SPIE, **26**), Redondo Beach, CA: Soc. Photo-Optical Instrum. Engineers 1971, pp. 131-138
- [71] You, Y. R., Miao, G. P., Liu, Y. Z.: A fast method for acoustic imaging of multiple three-dimensional objects. *J. Acoust. Soc. Amer.*, **101**, 31-37 (2000)