CONTRIBUTIONS TO IMAGING

A thesis

presented for the degree of

Doctor of Philosophy in Electrical Engineering

in the

University of Canterbury,

Christchurch, New Zealand

by

G.C. McKinnon   B.E. (Hons)

1980

ABSTRACT

Four topics are considered - each associated with a different aspect of imaging.

Using X-ray diffraction it is possible to say much about the structure of molecules. Two models for the D.N.A. molecule are analysed with respect to the diffraction data. The models are Watson and Crick's "Double Helix" and Rodley's, recently developed, "Side-by-Side". It is demonstrated that the side-by-side is a viable alternative model for D.N.A. However the low quality of the data precludes a definitive decision as to the actual structure of the molecule.

Conventional X-ray computed tomography body scanners, while producing impressive results when imaging stationary objects, can not image rapidly moving organs such as the beating heart. As the heart motion is periodic, it has been suggested that stroboscopic techniques be employed. However the resulting imaging quality is poor when standard image reconstruction methods are used. By taking account of the fact that the region surrounding the heart is stationary, though, a significant improvement in image quality can be obtained. A simple procedure for achieving this improvement is presented here.

Ultrasonic transmission tomography is more complicated than the X-ray case because ultrasonic rays, unlike X-rays, are diffracted as they pass through a body. Therefore they are generally curved. It is shown how ray curvature makes it impossible to image certain types of objects exactly. Nevertheless it seems that useful results can be obtained by treating the rays as being straight, and using X-ray computed tomography image reconstruction algorithms.

Imaging using electric currents is examined. The types of independent measurements that can be made are discussed. The imaging problem is far from trivial and, in the general case, largely unsolved. Here a method for uniquely imaging circularly symmetric conductivity distributions is outlined.

ACKNOWLEDGEMENTS

PREFACE

In certain imaging fields it is either impractical, undesirable, or impossible to directly visualize or measure the internal structure of an object. In these situations one must resort to indirect imaging techniques. Usually the object is irradiated and the manner in which the radiation is scattered or absorbed is used to infer the structure. This is the essence of non-invasive structure determination procedures. They are used in such areas as medical imaging, non-destructive testing, X-ray crystallography, and geophysical investigations.

This thesis, consisting of six chapters, is concerned with non-invasive methods for imaging the volume structure of an object. The first chapter reviews the problem of imaging the interior of an object using wave like radiation. In this chapter various theoretical solutions to the imaging problem are discussed. Most of the necessary theoretical preliminaries to the imaging procedures presented in later chapters are to be found here.

Chapter 2 delves into X-ray crystallography. This is the field concerned with the imaging of molecules. By analysing the manner in which X-rays are diffracted by the molecules it is possible to infer much about their structure. Unfortunately X-ray crystallographic imaging is subject to ambiguity. It is only by imposing stereochemical constraints that a correct "image" can be found. (Stereochemical constraints relate to the physical structure of the molecule such as types of atoms present, the allowable bond lengths, and bond angles etc.) It is impractical to systematically check all the "images" consistent with the data, and then eliminate those which are stereo-chemically unsound. The general method of "imaging" large molecules is to build a model which is consistent stereochemically, and then see if the diffraction pattern calculated from the model agrees with the actual diffraction pattern. In Chapter 2, two models of an important biological molecule, D.N.A., are compared with the diffraction data. These models are the, now familiar, "Double Helix" model proposed by Drs Watson and Crick in the mid 1950's; and the alternative "Side-by-side" model recently proposed by Dr Rodley of the Chemistry Department of this University. The side-by-side model has been subject to a certain amount

of controversy. However the analysis of D.N.A. diffraction data shows
that (within the limitations on the data) the side-by-side model is a
viable alternative. In one particular test it appears to be in better
agreement with the data than the double helix model is.

The third chapter is concerned with one of the more impressive
medical imaging devices, the X-ray computed tomography body scanner.
This machine takes projections (X-ray views) of the patient from many
different directions. Then, using a type of "triangulation" procedure,
the projections are processed by computer to produce an image showing
the three dimensional structure of the patient's internal organs.
Despite the acclaim of X-ray computed tomography there are still some
limitations. One of these is that it takes a reasonable amount of time
to measure all the different projections. A consequence of this is that
images of moving organs appear blurred. In Chapter 3 the problem of
imaging the heart using X-ray computed tomography is examined. The
heart moves in a periodic fashion, therefore it is possible to obtain
"stationary" images using a stroboscopic procedure. Two schemes are
discussed. The first consists of some design considerations for a new
type of scanner specifically for imaging the heart. With this device
it should be possible to obtain images of the heart comparable in
quality to those produced by existing scanners when imaging stationary
objects. The second scheme is a procedure which could be implemented
on existing scanners. The main problem, with stroboscopic imaging of
the heart, is that the heart motion can only be treated as being
periodic for a short time. This means that (with existing scanners)
not very many projections, corresponding to a particular phase of the
heart cycle, can be measured. Therefore, when using standard
reconstruction techniques, the resulting image quality is poor.
However, in the case of heart imaging, it is possible to improve image
quality. The second scheme is a method for achieving this improvement.
It involves taking account of the fact that the material surrounding the
heart is stationary, and so can be estimated using a large number of
projections. It is shown how this estimate can be used to pre-process
the original projections in such a way that errors over the heart region
of the image are significantly reduced.

Chapter 4 discusses an aspect of ultrasonic transmission
tomography. Conventional ultrasonic imaging is based on the "radar
like" pulse-echo technique. However, with the success of X-ray computed

tomography, and the simplicity of the image reconstruction algorithms, increasing interest is being shown in ultrasonic transmission imaging. Unfortunately, even though the procedure is very similar, it is not enjoying the same success as its X-ray counterpart. One of the problems is that ultrasonic energy transmission can not be analysed as simply as X-ray transmission. Both ultrasonic rays and X-rays can be treated in terms of thin rays. X-rays travel in straight lines. However, ultrasonic rays are refracted as they pass through an object, and so are, in general, curved. The simplest method of reconstructing an "image" from ultrasonic transmission measurements is to assume that the rays are straight and then use the conventional X-ray computed tomography algorithms. While studying the effects this approximation has on the resulting image, and also pondering over modifications to standard computed tomography algorithms which might take account of ray curvature, a particular limitation of ultrasonic transmission tomography became apparent. This is the main result presented in Chapter 4. Although this limitation precludes the reconstruction of an exact image (even if ray curvature is allowed for!), it appears that "usefully representative" images can often be obtained using algorithms that treat the rays as being straight.

Chapter 5 is concerned with electrical impedance imaging. The idea is to determine the internal conductivity distribution of an object from measurements of the impedance between pairs of points on the surface of the object. Imaging using static electric fields is more complicated than those cases in which the energy can be treated as travelling along rays. Rays act almost independently of each other. Therefore the measured quantity along a ray path reflects only the characteristics of the medium in the immediate neighbourhood of the ray. In the case where electric current is used for imaging purposes, the impedance between any two points depends on the conductivity at every point within the body. Present methods of imaging using electric currents are largely empirical. In Chapter 5 the types of independent measurements which can be made are discussed. The problem of imaging general conductivity distributions is outlined; and a method for imaging circularly symmetric objects is presented.

The final chapter, Chapter 6, contains conclusions and suggestions for further research.

The following papers (including some only recently submitted for publication) have been prepared during the course of this work:

R.H.T. Bates, G.C. McKinnon, and R.P. Millane, "A new look at B-DNA diffraction data", Research Report, Electrical Engineering Department, University of Canterbury, N.Z., Nov. 1978.

R.H.T. Bates, and G.C. McKinnon, "Towards improving images in ultrasonic transmission tomography", Australasian Physical Sciences in Medicine, Vol. 2-3, March, April 1979, pp. 134-140.

R.H.T. Bates, and G.C. McKinnon, "Possible applications of inverse scattering concepts to ultrasonic imaging", In: Research Techniques in Non-destructive Testing Vol. IV, R.S. Sharpe, editor, Academic Press, London, 1980, pp. 71-85.

R.H.T. Bates, G.C. McKinnon, R.P. Millane, and G.A. Rodley, "Revised interpretations of the available X-ray data for B-DNA", Pramāna, Vol. 14, No. 4, April 1980, pp. 233-252.

G.C. McKinnon, and R.H.T. Bates, "A limitation on ultrasonic transmission tomography", Ultrasonic Imaging, Vol. 2, 1980, pp. 48-54.

R.H.T. Bates, G.C. McKinnon, and A.D. Seagar, "A limitation on systems for imaging electrical conductivity distributions, Trans. IEEE, Vol. BME-27, No. 7, July 1980, pp. 418-420.

G.C. McKinnon and R.H.T. Bates, "Towards imaging the beating heart usefully with a conventional C.T. scanner", Submitted to Trans. IEEE on Biomedical Engineering - Special issue on computerized medical imaging, August 1980.

R.H.T. Bates, G.A. Rodley, and G.C. McKinnon, "DNA X-ray data", Nature, Vol. 287, 2nd October 1980, p. 384.

G.J. Berzins, G.C. McKinnon, R.H.T. Bates, and A.H. Lumpkin, "Application of computed tomography techniques to pinhole images of test fuel rods", Submitted to Nuclear Science and Engineering, October 1980.

TABLE OF CONTENTS

1   IMAGE FORMATION

1.1   INTRODUCTION

The simplest and, at the present time, most useful methods for imaging the internal structure of an object are those based on echo location procedures. A pulse (or, alternatively, a noise-like signal or a linear-frequency-modulation transmission) is emitted in a narrow beam, and echoes scattered back from the object are received. By measuring the time between emitted and received pulses and the strength of the received pulse, and assuming a value for the pulse propagation speed and also that the beam is in a straight line through the object, the position and "size" of the scatterer can be estimated. Impressive and useful results are obtained using these simple ideas. However, beam (or ray) curvature is appreciable in many of the applications for which there is an effective continuum of reflecting bodies. Further to this,the propagation speed varies in an, initially, unknown manner. Consequently it is worth examining the feasibility of instituting more rigorous approaches to the imaging problem.

In this chapter methods for imaging the internal structure of objects are reviewed. These "methods" are computational procedures developed from the basic mathematical physics describing the interaction between various kinds of wave motion and different types of object. The concepts lie within the realm of inverse scattering theory. Unfortunately the more rigorous inverse scattering solutions to non-trivial imaging problems are far from completely understood, and much reliance is placed in practice on approximate solutions. However,with the increasing availability of quite powerful computers it is becoming feasible to develop more of these theoretical notions into useful imaging techniques.

As an introduction to inverse scattering concepts a generalised imaging problem is discussed. Consider a space T consisting of two regions T_+ and T_-, where T_+ surrounds T_- (see Fig. 1.1). T_- contains an object χ .   Suppose that in the absence of χ a field ψ_o (either electromagnetic or acoustic wave) is present throughout all of T . ψ_o is under control of the investigator and represents the radiation

$$\mathcal{T}_+$$

$$\mathcal{T}_-$$

$$\chi$$

$$\mathcal{T} = \mathcal{T}_+ \cup \mathcal{T}_-$$
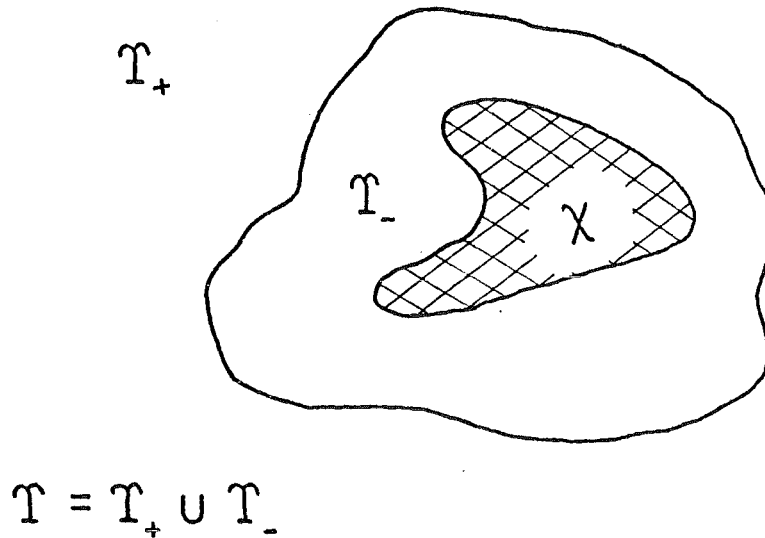
Fig. 1.1   The imaging space. The object χ lies with T_- .

T_+, surrounding T_-, is the "free space" region.

used to probe χ .   Introducing χ into T_- causes the field in T to change
(provided that χ interacts with ψ_o - if it does not then χ can not be
imaged). The resulting field ψ is then said to be the sum of the
original (incident) field ψ_o and the perturbing (scattered) field ψ_s,

$$\psi = \psi_o + \psi_s \quad . \tag{1.1}$$

This equation defines the scattered field ψ_s, which contains information
about χ .

In the analysis of physical systems there are basically two types
of problems. Firstly, direct problems where one tries to predict the
effect of an interaction from knowledge of the causes; and secondly,
indirect or inverse problems where the causes are deduced from the
observed effects. The direct problem, in our example, corresponds to
determining the scattered field ψ_s using knowledge of the object χ, and
the original undisturbed field ψ_o . This relationship can be expressed
in operator notation as

$$\psi_s = \Lambda \{ \chi, \psi_o \} \quad . \tag{1.2}$$

The operator Λ denotes a set of mathematical operations on χ and ψ_o. Λ is a mathematical descriptor of the "physical laws" relevant to the particular interaction. The converse, or inverse, problem is to determine the characteristics of the object χ from knowledge of ψ_o and ψ_s. In operator notation this is

$$\chi = \Lambda^{-1}\{ \psi_s, \psi_o \} \quad . \tag{1.3}$$

The inverse problem is synonymous with the imaging problem. Imaging is the process of determining the structure of an object from a remote position. In our imaging example ψ_s can only be measured within T_+ (i.e. remote from χ). To obtain an exact solution for χ, both ψ_s and χ have to be inferred within T_- . In many situations the inverse operator Λ^{-1} is either not known exactly or is impractical to use, and (1.3) has to be replaced by

$$\chi \simeq \tilde{\Lambda}^{-1}\{ \psi_s, \psi_o \} \quad , \tag{1.4}$$

where $\tilde{\Lambda}^{-1}$ is only approximately equivalent to Λ^{-1}. Also relevant to the imaging problem is the question of uniqueness - i.e. is the knowledge of ψ_s and ψ_o sufficient to determine χ ?

In the first part of this chapter the direct problem is discussed. Most imaging techniques use wavelike radiation to probe the object. In §§1.2 and 1.3 the scattering of acoustic and electromagnetic radiation is examined. The general methods for analysing the propagation of both electromagnetic and acoustic energy are based on the wave equation. These two sections review the approximations and assumptions inherent in this equation. Sections 1.4 to 1.7 discuss various solutions to the inverse scattering problem for the wave equation. The emphasis is on the underlying approximations, and hence applicability, of each method. In §1.8 the procedure for reconstructing an image from its projections is outlined. This imaging procedure is the simplest method of "solving" the inverse problem in those cases where acoustic and electromagnetic energy can be treated in terms of rays.

1.2 ACOUSTIC SCATTERING

The term "acoustic waves" is used here to include any sort of pressure or elastic wave. The propagation of acoustic waves is complicated by the fact that the waves are composed of two parts which propagate at different speeds.

The wave equation describing the propagation of acoustic energy is derived from the dynamics of an elastic medium (Morse and Feshback, §2.2, 1953). Assume that the displacement of the medium due to the wave motion is small and that there is no translation or rotation of the medium as a whole. Also assume that the material is perfectly elastic - can remain in equilibrium under a shearing stress, and isotropic - the elastic properties are independent of direction. Let $\underline{s}$ represent the displacement of an element of the medium. The inertial reaction of this element, of density ρ, to an accelerating force is $\rho \partial^2 \underline{s}/\partial t^2$, where t represents time. To write this force in terms of $\underline{s}$, and the elastic properties of the material, it is convenient to use tensor notation. Elastic media support both a twisting and a compression stress which can be described by a tensor quantity $\underline{\underline{T}}$, where

$$\underline{\underline{T}} = \lambda |\underline{\underline{G}}| \underline{\underline{I}} + 2\mu\underline{\underline{G}} \quad . \qquad (1.5)$$

This equation relates the stress to the strain tensor $\underline{\underline{G}}$. λ and μ represent the elastic properties of the material; $|\underline{\underline{G}}|$ is the spur of the tensor (sum of the diagonal elements); and $\underline{\underline{I}}$ is the unit tensor. The "strain" describes the distortion of the material, and "stress" is the system of forces required to sustain that distortion. The first term on the R.H.S. (right-hand side) relates to the difficulty in trying to pack more atoms into less space, and the second term is a measure of the bonding forces holding the atoms together. The quantity $(\lambda + 2\mu/3)$ is known as the compression modulus of the medium and relates the change in volume to an isotropic pressure. μ is the shear modulus. The strain tensor is related to the displacement by

$$\underline{\underline{G}} = (\nabla\underline{s} + \underline{s}\nabla)/2 \qquad (1.6)$$

$\nabla\underline{s}$ is the tensor quantity denoting the vector grad of $\underline{s}$; $\underline{s}\nabla$ is the transpose of $\nabla\underline{s}$). There is nothing very startling about (1.6), it is just a concise way of expressing linear and rotational strain.

The force over a surface element is given by $\underset{=}{T}\cdot\hat{n}$, where $\hat{n}$ is the unit vector normal to the surface. Therefore, using the divergence theorem, the net force on the volume element is $\nabla\cdot\underset{=}{T}$ ($\nabla\cdot$ is the vector divergence operator - operates on a tensor producing a vector). These results are combined to obtain the equation of motion for the medium acting under the influence of its own restoring forces. This is

$$\rho\partial^2\underline{s}/\partial t^2 = \nabla\cdot[\ \lambda\underset{=}{I}\nabla\cdot\underline{s} + \mu\nabla\underline{s} + \mu\underline{s}\nabla\ ]\ \ ,\tag{1.7}$$

where $\nabla\cdot$ is the ordinary divergence operator. As can be appreciated, obtaining the solution of this equation is not trivial especially when λ and μ vary significantly with position. However, when the variation of λ and μ is small compared to that of $\underline{s}$, (1.7) can be simplified. Neglecting terms containing $\nabla\lambda$ and $\nabla\mu$ (where ∇ is the gradient operator) results in

$$\rho\partial^2\underline{s}/\partial t^2 = (\lambda+\mu)\ \nabla(\ \nabla\cdot\underline{s}\ ) + \mu\nabla\cdot(\ \nabla\underline{s}\ )\ \ .\tag{1.8}$$

Using the identity

$$\nabla\cdot(\ \nabla\underline{s}\ ) \equiv \nabla(\ \nabla\cdot\underline{s}\ ) - \nabla\times\nabla\times\underline{s}\tag{1.9}$$

($\nabla\times$ is the curl operator), (1.8) can be written as

$$\rho\partial^2\underline{s}/\partial t^2 = (\lambda+2\mu)\ \nabla(\ \nabla\cdot\underline{s}\ ) - \mu\nabla\times\nabla\times\underline{s}\ \ .\tag{1.10}$$

Now any vector field can be represented as the sum of the gradient of a scalar and the curl of a zero divergence vector (Morse and Feshback, §1.5, 1953). Therefore write

$$\underline{s} = \nabla\psi + \nabla\times\underline{A}\ \ ,\tag{1.11}$$

where

$$\nabla\cdot A = 0\ \ .\tag{1.12}$$

Substituting this into (1.10) and separating terms in ψ and $\underline{A}$ results in two equations. Firstly

$$\nabla^2\psi - \rho/(\lambda+2\mu)\ \partial^2\psi/\partial t^2 = 0\ \ ,\tag{1.13}$$

where ∇^2 is the del operator. This is the scalar wave equation and the wave velocity is $(\lambda+2\mu)/\rho$. Solutions of (1.10), corresponding to the gradient of a scalar quantity, represent motion containing no

twisting of the medium, only stretching or compression (the curl of the gradient is always zero). Waves of this sort are called longitudinal waves or compression waves (P waves in seismology). The second solution is

$$\nabla \times \nabla \times \underline{A} + \rho/\mu \ \partial^2 \underline{A}/\partial t^2 = 0 \quad . \tag{1.14}$$

Using (1.12), and the vector identity

$$\nabla \times \nabla \times \underline{A} = \nabla(\ \nabla \cdot \underline{A} \ ) - \mathbf{\nabla}^2 \underline{A} \quad , \tag{1.15}$$

(1.14) becomes

$$\mathbf{\nabla}^2 \underline{A} - \rho/\mu \ \partial^2 \underline{A}/\partial t^2 = 0 \quad . \tag{1.16}$$

which again is the wave equation. ($\mathbf{\nabla}^2$ is the vector del operator.) This wave propagates at a speed given by μ/ρ . The vector components of $\underline{A}$, in (1.16), are uncoupled. However, because of (1.12), there are only two independent solutions. Solutions of (1.10), corresponding to the curl of a divergenceless vector, represent motion where there is no expansion or contraction of the medium. Therefore the strain is a type of shear. These waves are known as transverse or shear waves (S waves in seismology). By defining $\underline{s}$ as in (1.11) and (1.12), the solution has been separated into two components which travel at different speeds. It should be remembered that to achieve this, terms in $\nabla\lambda$ and $\nabla\mu$ have been neglected. Accordingly wherever the elastic parameters undergo abrupt changes, (1.10) is no longer a reasonable approximation to (1.7). Under these conditions there is coupling between the two solutions and the constitution of the wave motion usually changes. In some seismological work both P and S waves are detected and used to obtain information about the earth's structure (Bullen, 1963). But in most applications employing ultrasonic waves, such as medical imaging, it is assumed that only the faster compression waves propagate. In fluids the shear modulus μ is negligible. Therefore only compression waves are to be found.

1.3   ELECTROMAGNETIC SCATTERING

The wave equation describing the propagation of electromagnetic energy is derived from Maxwell's equations (Jones, chapter 1, 1964).

Maxwell's equations describe the interaction between electric field intensity $\underline{E}$, magnetic field intensity $\underline{H}$, electric flux density $\underline{D}$, and magnetic flux density $\underline{B}$. These relationships are;

$$\nabla \times \underline{E} + \partial \underline{B}/\partial t = 0 \quad , \tag{1.17}$$

$$\nabla \times \underline{H} - \partial \underline{D}/\partial t = \underline{J} \quad , \tag{1.18}$$

$$\nabla \cdot \underline{D} = \rho \quad , \tag{1.19}$$

and $\qquad \nabla \cdot \underline{B} = 0 \quad . \tag{1.20}$

$\underline{J}$ is the electric current density, and ρ is the charge density (not to be confused with the ρ in §1.2). The magnetic flux density $\underline{B}$ and electric flux density $\underline{D}$ depend on magnetic field intensity $\underline{H}$ and electric field intensity $\underline{E}$ respectively. In some materials such as ferromagnetics, ferrites, and ferroelectrics, the relationships between these quantities are non-linear; and also depend on the "history" of the material (Wert and Thomson, 1970; also Jones, chapter 1, 1953). In most materials, however, the relationships are linear and can be described by the following two equations;

$$\underline{D} = \varepsilon\, \underline{E} \quad , \tag{1.21}$$

and $\qquad \underline{B} = \mu\, \underline{H} \quad , \tag{1.22}$

where ε and μ are scalar or tensor quantities depending upon whether the material is istropic or not. Denote the values of ε and μ in free space by ε_o and μ_o respectively. The ratio $\varepsilon/\varepsilon_o$ is known as the dielectric constant of the material; and μ/μ_o as the relative permeability. For most materials the relative permeability is very close to unity. The dielectric constant, on the other hand, varies over a wide range - gases ~1 ; liquids 2-81 ; and solids 2-10 (Jones, 1964). [Note that for some ceramics, such as titania and titanate bodies, the dielectric constant may be in the thousands - see Von Hippel, 1954.]

In conductors, $\underline{J}$ is related to $\underline{E}$ by

$$\underline{J} = \sigma\, \underline{E} \quad . \tag{1.23}$$

This is Ohm's law (σ is known as the conductivity of the material).

Assuming the time dependence of $\underline{E}$, $\underline{H}$, $\underline{D}$, $\underline{B}$ and $\underline{J}$ to be of the form $\exp(j\omega t)$, and also that μ is constant ($\mu=\mu_o$), (1.17)-(1.23) can be manipulated to obtain (see for example Jones, 1964)

$$\nabla\times\nabla\times\underline{E} - \omega^2\mu_o(\varepsilon-j\sigma/\omega)\underline{E} = 0 \quad . \tag{1.24}$$

(ω is the frequency of the time variation, and j denotes the unit imaginary number $\sqrt{-1}$.) To simplify the notation, redefine ε as the complex quantity

$$\varepsilon = \varepsilon-j\sigma/\omega \quad , \tag{1.25}$$

where the imaginary part corresponds to a lossy dielectric. Suppose that there are no free charges (i.e. $\rho=0$), and that ε is a scalar (i.e. the medium is isotropic), then (1.19) and (1.21) result in

$$\varepsilon\nabla\cdot\underline{E} + \underline{E}\cdot\nabla\varepsilon = 0 \quad . \tag{1.26}$$

Using this, and the vector identity

$$\nabla\times\nabla\times\underline{E} \equiv \nabla(\nabla\cdot\underline{E}) - \mathbf{\nabla}^2\underline{E} \quad , \tag{1.27}$$

(1.24) becomes

$$\nabla^2\underline{E} + \omega^2\mu_o\varepsilon\underline{E} = -\nabla(\underline{E}\cdot\nabla\ln\varepsilon) \quad . \tag{1.28}$$

This is a vector equation for $\underline{E}$ (at a given frequency) in terms of the permeability and dielectric properties of a medium. The term on the R.H.S. of (1.28) causes difficulties in the analysis as it couples the different vector components to each other. In other words (1.28) is actually three simultaneous second order differential equations. However when certain assumptions can be made about ε the analysis simplifies. Firstly define

$$k = \omega\sqrt{\mu_o\varepsilon_o} \quad , \tag{1.29}$$

where k is known as the wave number. Also define the refractive index ν as

$$\nu = \sqrt{\varepsilon/\varepsilon_o} \quad . \tag{1.30}$$

Now, in the special case where ε is varying in one dimension, it is possible to "choose" $\underline{E}$ such that $\underline{E}\cdot\nabla\ln\varepsilon$ is zero. However, in the

general case where ε is varying in an arbitrary manner, (1.28) can only be simplified when $\nabla \ln \varepsilon$ is small. Suppose, then, that the spatial change in $\ln \varepsilon$ per unit wavelength is small, i.e.

$$\left| \nabla \ln \varepsilon \right| \ll \nu k \quad . \tag{1.31}$$

Then

$$\nabla \ (\ \underline{E} \cdot \nabla \ln \varepsilon \ ) \ \simeq \ k \nu \ \left| \nabla \ln \varepsilon \right| \ \underline{E} \tag{1.32}$$

$$\ll \ k^2 \nu^2 \underline{E} \quad . \tag{1.33}$$

Under these conditions the term on the R.H.S. of (1.19) can be neglected, resulting in

$$\mathbf{\nabla}^2 \underline{E} \ + \ k^2 \nu^2 \underline{E} \ = \ 0 \quad . \tag{1.34}$$

The vector components are now uncoupled and the analysis of electromagnetic scattering reduces to solving three scalar wave equations. Solutions to (1.34) describe how the electric field propagates. The other vector quantities can be obtained from $\underline{E}$ via Maxwell's equations.

1.4 BATES SOLUTION TO THE INVERSE SCATTERING PROBLEM

It has been shown in §§1.2 and 1.3 that, under certain conditions, electromagnetic and acoustic scattering can be analysed using the scalar wave equation. The equation is valid, in both cases, if the variation in the refractive index (includes the elastic parameters) is small compared with the variation of the field or wave function. Notice that the wave equation can be written in two different forms. Equation (1.34), written in the "frequency domain", is of the form

$$\nabla^2 \psi \ + \ k^2 \nu^2 \ \psi \ = \ 0 \quad ; \tag{1.35}$$

and (1.13) and (1.16), written in the "time domain" representation, are of the form

$$\nabla^2 \psi \ - \ (\nu^2/c^2) \ \partial^2 \psi / \partial t^2 \ = \ 0 \quad . \tag{1.36}$$

c is the free space propagation velocity (i.e. the speed at which a pulse would travel when the refractive index ν equals unity). The two equations are equivalent - one is the temporal Fourier transform (time

and frequency being the transform variables) of the other.

The inverse scattering problem as it relates to the wave equation is now discussed. Again consider the space T, divided into two regions T_+ and T_- (see Fig. 1.2).
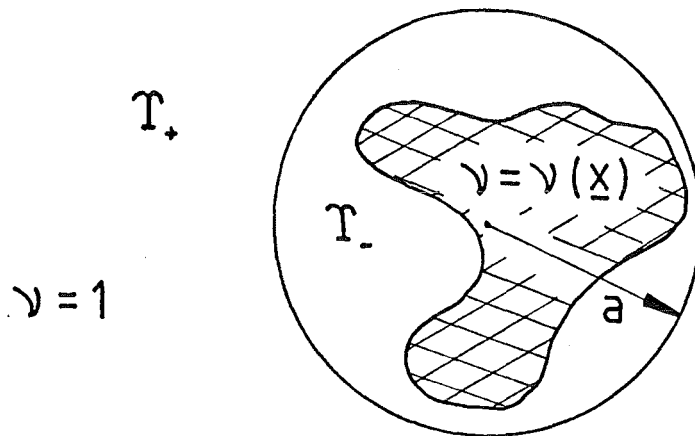
Fig. 1.2   The imaging space showing the variable refractive index confined to T_-. This time T_- is a circle, of radius a, just large enough to contain $\nu(\underline{x})$.

Within T_+ the refractive index is constant (equals unity). However within T_- it is a function of position $\nu(\underline{x})$. The imaging (or inverse scattering problem) is to determine $\nu(\underline{x})$ from measurements of Ψ, or ψ, made within T_+.

The first solution to be outlined in this chapter is that presented by Bates (1975). Conceptually the method is fairly straight-forward. However, it is a sort of "brute force" approach to the problem and the resulting solution is difficult to compute. Nevertheless it is the only exact solution to the inverse scattering problem associated with the wave equation, in cases where $\nu(\underline{x})$ varies in more than one dimension. It is presented here to illustrate the difficulties in trying to obtain exact solutions. Basically all the unknowns ($\nu(\underline{x})$ and

Ψ with T_-) are expressed in terms of a series of orthogonal functions. The orthogonal series for Ψ within T_- is chosen so that it is consistent with the field measured in T_+. Initially the coefficients for the orthogonal functions are unknown. However by substituting the series for both wavefunction and refractive index into the wave equation, conditions are found that the coefficients must satisfy. Unfortunately these conditions are a set of quadratic equations.

Bates' method is described here in terms of a two-dimensional space, although it is easily extended to three dimensions. Take a point in T_- as origin for the polar coordinates (r,θ). Redefine T_- to be a circle of radius a centered on the origin (see Fig. 1.2). The wavefunction $\Psi(r,\theta,k)$ satisfies the equation

$$\nabla^2 \Psi + k^2 \nu^2 \Psi = 0 \quad , \tag{1.37}$$

where

$$\nu = \begin{cases} \nu(r,\theta) & ; \ r < a \\ 1 & ; \ r \geq a \end{cases} \tag{1.38}$$

Ψ is measured for $r \geq a$. Write Ψ as an angular Fourier series

$$\Psi = \sum_{m=-\infty}^{\infty} \Psi_m(r,k) \exp(jm\theta) \quad . \tag{1.39}$$

The $\Psi_m(r,k)$ are known for $r \geq a$. The wave impedance vector $\underline{Z}$ at $r = a$ can be obtained from the measurements. The mth component is defined by

$$Z_m(k) = \frac{\Psi_m(a,k)}{\Psi'_m(a,k)} \quad , \tag{1.40}$$

where the prime denotes the derivative with respect to r. Bates poses the inverse scattering problem as follows; determine ν for all $r < a$, given $\underline{Z}$. Firstly a suitable set of orthogonal functions has to be chosen for the series representation of the Ψ_m over the range $0 \leq r < a$. The series has to be such that $Z_m(k)$, for each Ψ_m, is continuous at $r = a$. Define Ψ_m as

$$\Psi_m = \sum_{\ell=1}^{\infty} b_{\ell,m}(k) \, \phi_{\ell,m}(r,k) \quad ; \ r < a \quad . \tag{1.41}$$

Provided that the field at $r = 0$ is not zero or infinite, the $\phi_{\ell,m}$ must mimic the behaviour of the mth order cylindrical Bessel function $J_m(\cdot)$. On account of this and the impedance constraint (1.40), the $\phi_{\ell,m}$ are defined by

$$\phi_{\ell,m} = J_m(h_{\ell,m}(k) \, r/a) \quad , \tag{1.42}$$

where the $h_{\ell,m}(k)$ are defined by

$$J_m(h_{\ell,m}) = (h_{\ell,m}/a) \, Z_m \, J'_m(h_{\ell,m}) \quad . \tag{1.43}$$

The representation for the refractive index is obtained in a similar manner. ν equals unity at $r = a$. Therefore a suitable expansion is

$$\nu^2 = 1 + \sum_{n=-\infty}^{\infty} \nu_n(r) \, \exp(jn\theta) \quad ; \ r < a \quad , \tag{1.44}$$

where

$$\nu_n = \sum_{p=1}^{\infty} A_{n,p} \, J_n(g_{n,p} \, r/a) \quad , \tag{1.45}$$

and the $g_{n,p}$ are defined by

$$J_n(g_{n,p}) = 0 \quad . \tag{1.46}$$

The series for each Ψ_m and ν_n are orthogonal over $0 \leq r \leq a$. They are known as Dini series and Fourier-Bessel series respectively (Watson, 1966). Both series representations are (i) consistent with the known boundary conditions at $r = a$, and (ii) are sufficient to describe any physically realisable function within $0 \leq r \leq a$. Using (1.39) and (1.41)-(1.43), and the form of Bessel's differential equation, $\nabla^2 \Psi$ can be written as

$$\nabla^2 \Psi = - \sum_{\ell=1}^{\infty} \sum_{m=-\infty}^{\infty} (h_{\ell,m}/a)^2 \, b_{\ell,m} \, \phi_{\ell,m} \, \exp(jm\theta) \quad . \tag{1.47}$$

Substituting this expression and those for Ψ and ν^2 into the wave equation (1.37), and then using the orthogonality of the $\exp(jm\theta)$ and $\phi_{\ell,m}$, results in a system of simultaneous equations. The unknowns are the $A_{n,p}$ and the $b_{\ell,m}$ - but most of the terms are quadratic. However the $b_{\ell,m}$ and also the $h_{\ell,m}$ are functions of k, therefore taking measurements at different frequencies results in additional independent

sets of equations to aid in determining the $A_{n,p}$ (which are independent of frequency). Once the $A_{n,p}$ are determined an "image" of ν has been obtained. The quadratic equations are of the following form

$$\sum_{\ell=1}^{\infty} \sum_{m=-\infty}^{\infty} \left\{ [k^2 - (h_{\ell',m'}/a)^2] N_{\ell',m'} \delta_{\ell,\ell'} \delta_{m,m'} \right.$$

$$\left. + k^2 \sum_{p=1}^{\infty} A_{m'-m,p} \chi_{\ell,\ell',m,m',m-m,p} \right\} b_{\ell,m} = 0 \quad . \qquad (1.48)$$

Where

$$\chi_{\ell,\ell',m,m',n,p}(k) = \int_0^a J_n(g_{n,p}r/a) \phi_{\ell,m}(r,k) \phi_{\ell',m'}(r,k) \, r \, dr \quad , (1.49)$$

and

$$N_{\ell,m}(k) = \int_0^a [\phi_{\ell,m}(r,k)]^2 \, r \, dr \quad . \qquad (1.50)$$

ℓ' is any non-negative integer, m' is any integer and $\delta_{m,n}$ is the Kronecker delta. Equation (1.48) can be considered as a set of homogeneous linear algebraic equations for the $b_{\ell,m}$. It possesses a non-trivial solution only if the determinant of the coefficients equals zero. (The term inside the outer brackets of equation (1.48) represents an element in the coefficient matrix; ℓ',m' and ℓ,m denote row and column indices respectively.) The unknowns in the determinant are products of the $A_{n,p}$. Each determinant can be treated as a single linear algebraic equation, the unknowns being these products of the $A_{n,p}$. By taking measurements at different frequencies, the terms in the coefficient matrix change but the unknown products of the $A_{n,p}$ do not. Thus a set of simultaneous linear algebraic equations can be constructed for the products of the $A_{n,p}$. Once a value for each product is known the individual $A_{n,p}$ can be unravelled. Unfortunately this procedure is not as attractive as it initially appears. The summations in equation (1.48) have infinite limits. However if ν is reasonably well behaved only a finite number, M say, of the $A_{n,p}$ are needed to reconstruct ν to a given accuracy. If, now, the coefficient matrix is truncated to M×M, then the number of unknowns within the determinant is M!. Obviously all the unknowns are not independent, as there are only M different $A_{n,p}$, but there does not appear to be any

simple way of taking this into account. This reasoning shows that a direct approach would be quite laborious. It seems that iterative techniques which solve for the individual $A_{n,p}$ are necessary. The practical computational aspects of this procedure have not yet been investigated.

1.5 THE GELFAND-LEVITAN METHOD

The "Gelfand-Levitan method" is a procedure for solving the Sturm-Liouville (S-L) equation. This equation arises mainly in analysing non-relativistic quantum mechanical scattering (Newton, 1966), but can also be applied to the analysis of electromagnetic scattering by the ionosphere (Jordan and Ahn, 1979). The one-dimensional S-L equation can be written either, in terms of the frequency k, as

$$\partial^2 \Psi / \partial x^2 + k^2 \Psi - q \Psi = 0 \quad ; \qquad (1.51)$$

or, in the time domain, as

$$\partial^2 \psi / \partial x^2 + (1/c^2) \; \partial^2 \psi / \partial t^2 - q \psi = 0 \quad . \qquad (1.52)$$

The two equations are equivalent - one being the temporal Fourier transform of the other. c is a constant corresponding to the propagation speed of the wavefunction; and q is a function of x, known as the "scattering potential".

Compare (1.51) and (1.52) with the wave equations (1.35) and (1.36). The difference is that, in the S-L equation, the scattering potential is not associated with the second time derivative of ψ; whereas in the wave equation the scattering function (in this case ν) is associated with the second time derivative of ψ. From a physical point of view this means that a disturbance described by the S-L equation propagates at a constant speed c, whereas those of the wave equation travel at a speed depending on the refractive index ν. This feature results in the inverse scattering problem associated with the S-L equation being easier to solve than that of the wave equation. It can be seen to be easier by considering a short pulse being used to probe the scattering medium (either q or ν). The pulse is sent into the medium and the reflections are measured. If the propagation of the pulse can be described by the S-L equation then the pulse speed

is constant, and so the position of the pulse is known for all time.
However if the pulse propagation is described in terms of the wave
equation then the position is unknown, as it depends upon the unknown
function ν .   Therefore, in the case of the wave equation, the position
of the point within the medium causing the reflections is unknown, thus
making the reflections more difficult to interpret.   Equations (1.51)
and (1.52) have been written in terms of only one spatial coordinate.
It is possible to convert the one dimensional wave equation into this
form using an appropriate transformation (cf. Bates and Wall, 1976).
It should also be noted that the Gelfand-Levitan method is exact,
therefore this is a procedure for solving the inverse scattering problem
for the one-dimensional wave equation exactly.

The method for converting the wave equation to the S-L equation
is as follows.   Write the wave equation as

$$\partial^2\Psi/\partial x^2 + k^2\nu^2\Psi = 0 \quad , \tag{1.53}$$

where

$$\nu = \begin{cases} 1 & ; \ x \leq 0 \\ \nu(x) & ; \ x > 0 \end{cases} \tag{1.54}$$

Now define a new coordinate as

$$u = \begin{cases} x & ; \ x \leq 0 \\ \displaystyle\int_0^x \nu \, dx & ; \ x > 0 \end{cases} \tag{1.55}$$

Therefore

$$du = \nu dx \quad . \tag{1.56}$$

Also define a new wavefunction by

$$\Phi = \nu^{\frac{1}{2}} \Psi \quad . \tag{1.57}$$

Equations (1.55) and (1.56) insure that the propagation speed is
constant with respect to the coordinate u .   Equation (1.57) can be
understood in terms of the "density" of the wave.   Suppose that the
wavefunction describes a pulse; as the wave slows down (ν increases)

the width of the pulse decreases and so the energy density increases. Substituting (1.55), (1.56) and (1.57) into (1.53) results in

$$\partial^2 \Phi/\partial u^2 + k^2\Phi - q\Phi = 0 \quad , \tag{1.58}$$

where

$$q = \begin{cases} 0 & ; \; u \leq 0 \\ (1/4v^2) \; [2v \; \partial^2 v/\partial u^2 - (\partial v/\partial u)^2] & ; \; u > 0 \quad . \end{cases} \tag{1.59}$$

Once $q(u)$ is found it is possible to determine $v(x)$. For the remainder of this section the imaging problem will be described with respect to the one dimensional S-L equation.

Consider a one dimensional scattering problem. The field incident from the left impinges on a scattering body, which is confined to the half-space $x > 0$. The field satisfies the S-L equation, where

$$q = \begin{cases} 0 & ; \; x \leq 0 \\ q(x) & ; \; x > 0 \quad . \end{cases} \tag{1.60}$$

The incident field is a delta function impulse $\delta(x-ct)$. When this impulse penetrates $q(x)$ a reflected transient $R(x+ct)$ is produced which travels to the left (refer to Fig. 1.3). The field in the "free space" region $x \leq 0$ can be written in the form

$$\psi_f = \delta(x-ct) + R(x+ct) \quad ; \tag{1.61}$$

and the corresponding frequency dependent expression as

$$\Psi_f = \exp(jkx) + r(k) \exp(-jkx) \quad , \tag{1.62}$$

where $r(k)$ is the reflection coefficient. The inverse problem is to determine $q(x)$ from $R(x+ct)$ (or $r(k)$).

$$\psi_f = \delta(x - ct)$$
$$+ R(x + ct)$$
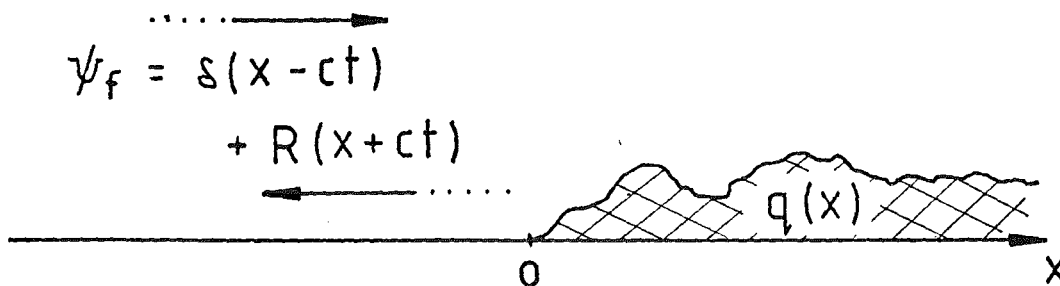
Fig. 1.3   An impulse, incident from the left,
            strikes q, causing a reflected wave.

The time dependent field in the inhomogeneous region $x > 0$ satisfies
the differential equation

$$\partial^2\psi/\partial x^2 - (1/c^2)\ \partial^2\psi/\partial t^2 - q\psi = 0 \quad . \qquad (1.63)$$

The Gelfand-Levitan "trick" is to write this field in terms of a
transformation of the free space field (1.61). This transformation is

$$\psi(x,t) = \psi_f(x,t) + \int_{-x}^{x} K(x,\xi)\ \psi_f(\xi,t)\ d\xi \quad . \qquad (1.64)$$

It can be shown (Newton, 1966) that the function $K(x,\xi)$ satisfies
equation (1.63) with the boundary conditions

$$K(x,-x) = 0 \quad , \qquad (1.65)$$

and

$$2dK(x,x)/dx = q(x) \quad . \qquad (1.66)$$

To further constrain $K(x,\xi)$, one has to resort to causality arguments.
For a pulse incident from the left, the field ψ must remain zero until
$x-ct > 0$. That is,

$$\psi(x,t) = 0 \quad ; \; ct < x \quad .\tag{1.67}$$

This requirement, together with equations (1.61) and (1.64), results
in the integral equation

$$R(x+ct) + K(x,ct) + \int_{-ct}^{x} K(x,\xi) R(\xi+ct) \; d\xi = 0 \quad .\tag{1.68}$$

In some ways this is similar to the "null field" (Bates and Wall, 1977)
or "extended boundary conditions" (Waterman, 1965) procedures adopted
in analysing the electromagnetic or acoustic scattering from perfectly
reflecting bodies. In these situations the field is zero within the
scattering body. However by treating the incident field and the field
scattered from the surface as though they separately exist both inside
and outside the body, and then requiring that the sum of the two fields
is zero within the body, the actual scattered field (outside the body)
can be determined. In the Gelfand-Levitan method $R(x+ct)$ is assumed
to exist everywhere, provided only that $x+ct > 0$ (for causality
requirements in the free space region), and $K(x,\xi)$ is chosen so that
the actual field is zero until $ct > x$. Once the integral equation (1.68)
is solved for $K(x,\xi)$, (1.66) gives the solution to the imaging problem.
The integral equation can be solved exactly when $r(k)$ is a rational
function of k (Jordan and Ahn, 1979), but there has been very little
reported on using the Gelfand-Levitan procedure to analyse real
scattering data. Equation (1.68) does, however, appear to be amenable to
numerical solution. $K(x,x)$ depends only upon $R(\mu)$ for values of μ
defined by $0 < \mu < 2x$; and, for a given value of x , $K(x,x)$ is
independent of previously calculated values - so errors do not
accumulate as x is increased.

The Gelfand-Levitan method for three-dimensional scattering is
largely restricted to fairly esoteric mathematical discussions (Newton,
1974). Even if it can be developed to handle the three dimensional S-L
equation it is going to be difficult to adapt it to electromagnetic and
acoustic imaging problems. The major hurdle is in generalizing equation
(1.55) to more than one dimension. This equation defines a metric over

which disturbances propagate with a constant speed. Another way of looking at this is that the "ray paths" in the metric for the S-L equation are straight lines. In the wave equation the "ray paths" are curved geodesics. [Geodesics define lines between pairs of points such that the integral of a function (in this case refractive index) over that line is minimised with respect to any other line between the same pair of points.] There is no way, in general, of transforming a space with curved geodesics into a space with straight line geodesics. In one dimension the problem is trivial as all geodesics are straight lines.

1.6 THE BORN AND RYTOV APPROXIMATIONS

The Born and Rytov approaches to the inverse scattering problem are both based upon approximations to the wave equation. The approximations inherent in each method are slightly different but basically they both require that the scattered field be much smaller than the incident field. The Born approximation is the basis for X-ray crystallographic diffraction pattern analysis (Cowley, 1975).

Once again consider the space T, comprised of T_+ and T_-, where T_- contains the scatterer (see Fig. 1.2). The wave equation (1.35) can be rewritten as

$$\nabla^2 \Psi + k^2 \Psi = - k^2 (\nu^2 - 1) \Psi \quad , \tag{1.69}$$

where ν is defined by

$$\nu = \begin{cases} 1 & ; \quad \underline{x} \in T_+ \\ \\ \nu(\underline{x}) & ; \quad \underline{x} \in T_- \end{cases} \tag{1.70}$$

Now a general solution of the differential equation

$$\nabla^2 \Psi + k^2 \Psi = - P \quad , \tag{1.71}$$

is (see for example Morse and Feshback, chapter 11, 1953)

$$\Psi(\underline{x}, k) = \Psi_0(\underline{x}, k) + \frac{1}{4\pi} \iiint P(\underline{x}') \frac{\exp(jk|\underline{x}-\underline{x}'|)}{|\underline{x}-\underline{x}'|} d\underline{x}' \quad , \tag{1.72}$$

where Ψ_o is a solution of the equation

$$\nabla^2 \Psi_o + k^2 \Psi_o = 0 \quad . \tag{1.73}$$

In our case Ψ_o represents the field being used to probe the medium. The total wavefunction is the sum of the incident field Ψ_o and the scattered field Ψ_s. That is

$$\Psi = \Psi_o + \Psi_s \quad . \tag{1.74}$$

Using (1.69), and (1.72)-(1.74), the scattered field can be written as

$$\Psi_s(\underline{x},k) = \frac{k^2}{4\pi} \iiint_{T_-} (\nu^2(\underline{x}')-1) \ \Psi(\underline{x}',k) \ \frac{\exp(jk|\underline{x}-\underline{x}'|)}{|\underline{x}-\underline{x}'|} \ d\underline{x}' \quad . \tag{1.75}$$

The difficulty in solving this equation for ν is that Ψ is also unknown within T_-. (Ψ_s is only measured in T_+.) However if $\Psi_s \ll \Psi_o$, then within the region T_-, Ψ can be approximated by Ψ_o. This is known as the Born (and also the Rayleigh-Gans) approximation (see for example Jones, 1964). Equation (1.75) then becomes

$$\Psi_s(\underline{x},k) = \frac{k^2}{4\pi} \iiint_{T} (\nu^2(\underline{x}')-1) \ \Psi_o(\underline{x}',k) \ \frac{\exp(jk|\underline{x}-\underline{x}'|)}{|\underline{x}-\underline{x}'|} \ d\underline{x}' \quad . \tag{1.76}$$

which can be solved (or inverted) to obtain $\nu(\underline{x})$.

One particularly simple method of inverting (1.76) is as follows. Suppose that the incident field is a plane wave. This can be written as

$$\psi_o(\underline{x},k) = \exp(j\underline{k}\cdot\underline{x}) \quad . \tag{1.77}$$

The vector wave number $\underline{k}$ denotes the direction as well as the frequency of the plane wave; $k=|\underline{k}|$. For simplicity define

$$F(\underline{x}) = \nu^2(\underline{x}) - 1 \quad . \tag{1.78}$$

If the scattered field is measured at a large distance from T_- (that is,

$$|\underline{x}| \gg |\underline{x}'| \quad), \tag{1.79}$$

then

$$|\underline{x}-\underline{x}'| \simeq |\underline{x}| - (\underline{x}\cdot\underline{x}')/|\underline{x}| \quad . \tag{1.80}$$

Using (1.77)-(1.80), (1.76) can be rewritten as

$$\psi_s(\underline{x},k) \simeq \frac{k^2}{4\pi} \frac{\exp(jk|\underline{x}|)}{|\underline{x}|} \iiint_{T} F(\underline{x}') \exp\left(j[\underline{k} - \frac{k\,\underline{x}}{|\underline{x}|}] \cdot \underline{x}'\right) d\underline{x}' \qquad . \quad (1.81)$$

Provided that ψ_s is measured at a constant radius from the origin (i.e. $|\underline{x}|$ = constant), and the vector $\underline{u}$ is defined by (see Fig. 1.4)

$$2\pi\underline{u} = \underline{k} - (k\underline{x})/|\underline{x}| \qquad , \qquad (1.82)$$

equation (1.81) becomes

$$H(\underline{u}) = \iiint_{T} F(\underline{x}') \exp(j2\pi\underline{u}\cdot\underline{x}') d\underline{x}' \qquad . \qquad (1.83)$$
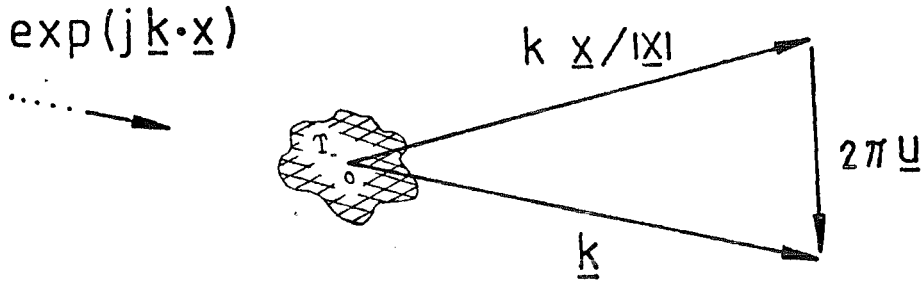

Fig. 1.4 $\underline{u}$ is the vector difference between the
direction of the incident and scattered
wave.

$H(\underline{u})$ incorporates the measured field Ψ_s and the constant term in front of the integral. Equation (1.83) is a Fourier transform; thus by taking the inverse Fourier transform of the scattered field, $F(\underline{x})$, and hence $\nu(\underline{x})$, can be obtained (see Wolf, 1969).

In order for Born's approximation

$$\Psi_s \ll \Psi_o \qquad (1.84)$$

to hold, changes in amplitude and phase of the incident wave, due to the variable refractive index, must be small. Since changes grow with the distance traversed, the integrated value of the deviation of the refractive index from unity must be small (Wade, 1979). This depends on the size of the domain T_- as well as the refractive index.

Rytov's method is very similar to Born's solution to the inverse scattering problem. However Rytov's approximation is claimed to be the more accurate of the two, as it is only the integrated deviation of the refractive index from unity per wavelength which has to be small (i.e. the approximation is independent of the size of the scattering domain) (Wade, 1979).

Write the wavefunction as

$$\Psi = \exp(\gamma) \quad , \qquad (1.85)$$

where

$$\gamma = \gamma_o + \gamma_s \quad . \qquad (1.86)$$

γ_o is defined by

$$\Psi_o = \exp(\gamma_o) \quad , \qquad (1.87)$$

where Ψ_o satisfies the free space wave equation (1.73). Therefore

$$\nabla\gamma_o \cdot \nabla\gamma_o + \nabla^2\gamma_o + k^2 = 0 \quad . \qquad (1.88)$$

Substituting (1.85) and (1.86) into the wave equation (1.69), and using (1.88), results in

$$\nabla^2\gamma_s + 2\nabla\gamma_o \cdot \nabla\gamma_s + \nabla\gamma_s \cdot \nabla\gamma_s = -k^2(\nu^2-1) \quad . \qquad (1.89)$$

Now define a new function as

$$\theta = \gamma_s \Psi_o \quad . \tag{1.90}$$

It is easily shown that this function, using (1.88) and (1.89), satisfies

$$\nabla^2 \theta + k^2 \theta = - [k^2 (\nu^2(x)-1) + \nabla\gamma_s \cdot \nabla\gamma_s] \Psi_o \quad . \tag{1.91}$$

The Rytov approximation involves neglecting the term $\nabla\gamma_s \cdot \nabla\gamma_s$, which is equivalent to requiring that

$$\nabla\gamma_s \ll \nabla\gamma_o \tag{1.92}$$

(in (1.89)). The function θ can therefore be written as

$$\theta(\underline{x},k) = \frac{k^2}{4\pi} \iiint_{T_-} F(\underline{x}') \ \Psi_o(\underline{x}',k) \ \frac{\exp(jk|\underline{x}-\underline{x}'|)}{|\underline{x}-\underline{x}'|} \ dx \quad . \tag{1.93}$$

This equation can be solved in the same manner as (1.76).

The difference between the two methods is the way in which the scattered field is defined. In the Born approach Ψ_s is defined by $\Psi = \Psi_o + \Psi_s$ whereas Rytov's scattered field θ is defined by

$$\Psi = \Psi_o \exp(\theta/\Psi_o) \quad . \tag{1.94}$$

In the limit as θ/Ψ_o becomes small, $\theta \to \Psi_s$, and the two approximate procedures for solving the inverse scattering problem become identical.

A method has been developed by Bates *et al*. (1976) which takes partial account of the $\nabla\gamma_s \cdot \nabla\gamma_s$ term neglected in Rytov's approach. It is assumed that the wave motion can be approximated by tubes of rays. The total wavefunction is then given by

$$\Psi \simeq \nu^{-\frac{1}{2}} \exp(jk \int_{\ell_0}^{\ell_1} \nu \, d\ell) \quad , \tag{1.95}$$

where the integral is evaluated along the ray path between points ℓ_0 and ℓ_1. This implies that

$$\gamma_o + \gamma_s = jk \int_{\ell_0}^{\ell_1} \nu \, d\ell - \tfrac{1}{2} \ln \nu \quad . \tag{1.96}$$

Now because

$$\nabla\gamma_0 = jk\hat{\underline{\tau}} \qquad (1.97)$$

(where $\hat{\underline{\tau}}$ is the unit vector pointing between ℓ_0 and ℓ_1),

$$\nabla\gamma_s = jk[\hat{\underline{s}}\nu - \hat{\underline{\tau}}] - \tfrac{1}{2}\ln \qquad (1.98)$$

(where $\hat{\underline{s}}$ is the unit vector pointing along the ray). It is also
assumed that the tangent to the curved ray is almost parallel to
the straight line, which is equivalent to

$$\hat{\underline{s}}\cdot\hat{\underline{\tau}} \simeq 1 \quad . \qquad (1.99)$$

Substituting (1.98) and (1.99) into (1.91) (rather than neglecting
the $\nabla\gamma_s\cdot\nabla\gamma_s$ term) means that the $F(\underline{x})$ in (1.93) has to be redefined
as

$$\tilde{F}(\underline{x},k) = (\nu(\underline{x}) - 1)\left[2 - (j/k)\,\frac{\partial\ln\nu}{\partial\tau}\right] + (1/4k^2)\,\nabla\ln\nu\cdot\nabla\ln\nu \quad . \qquad (1.100)$$

$\tilde{F}$ is a function of frequency as well as position. However the
different terms in $\tilde{F}$ can be separated out, as they have different
k dependence, enabling one to obtain the refractive index function ν .
This latter method has the disadvantage in that to solve for ν ,
scattering measurements must be made over several different
frequencies. Numerical results have been obtained (Dunlop *et al.*,
1976) which illustrate the increased accuracy of this method over
that of the conventional Rytov method, and also the Born procedure.

As mentioned at the start of this section, the Born
approximation is used very successfully in analysing X-ray diffraction
from crystals (Sherwood, 1976). The method is applicable in
crystallography because very little X-ray energy is scattered out
of the incident beam. Thus $\Psi_s << \Psi_o$.

Recently, interest has been shown as to the possibility of using
these techniques to interpret ultrasonic scattering in an imaging sense
(Kaveh *et al.*, 1979). It appears, though (from computer simulations

- Kaveh et al., 1979), that the Born and Rytov approximations
only give good results when:

(i)   the variation in refractive index is small (i.e. about 5%),
      and

(ii) the size of the scatterer is also small (less than about
      3 or 4 wavelengths).

However very few useful practical results have been reported.
This is unfortunate because the image reconstruction procedure,
via the Fourier inversion of (1.83), is so simple. Even if it is
difficult to image general refractive index distributions,
"representative images" could probably be obtained - especially when
$\nu(\underline{x})$ is composed of point like scatterers.

1.7   RAYS

The most widely used approaches to analysing wavelike scattering
are based on the assumption that the wave motion propagates as rays.
These rays are defined by geodesics in the "refractive index space".
That is, the ray path between two points $\underline{x}_0$ and $\underline{x}_1$ is such that the
value of the integral

$$s = \int_{\underline{x}_0}^{\underline{x}_1} \nu \, d\ell \qquad (1.101)$$

is minimised ($d\ell$ is along the ray path). s is known as the "optical
path length" between the points $\underline{x}_0$ and $\underline{x}_1$.

It is rather difficult to derive (1.101) from the frequency
domain wave equation without some sleight of hand. The standard
methods are of the following form. Assume that the wavefunction
can be represented as an asymptotic series (Born and Wolf, chapter 3,
1970)

$$\Psi(\underline{x},k) = \exp\bigl(jks(\underline{x})\bigr) \sum_{n=0}^{N} A_n(\underline{x}) \, k^{-n} \qquad . \qquad (1.102)$$

The quantity s($\underline{x}$) is called the phase of the wavefunction. Thus the wave appears to be travelling in the ∇s direction. This expression for Ψ is substituted into the wave equation (1.35). Terms associated with a particular power of k can be treated independently, producing a series of equations that s and the A_n must satisfy. The first of these equations (in the highest power of k) is

$$\nabla s \cdot \nabla s = \nu^2 \quad , \tag{1.103}$$

which is known as the eikonal equation. A solution to the eikonal equation is the integral equation (1.101). Thus the wave appears to propagate along ray paths defined by (1.101). The eikonal equation incorporates both Snell's law and Fermat's Principle.

The above argument has been presented using the frequency domain form of the wave equation. More solid reasoning behind using the concept of rays in refraction problems can be obtained by considering the behaviour of pulse type solutions to the time domain form of the wave equation (Friedlander, 1953).

The Geometry of Space-Time

Firstly consider a three dimensional hypersurface (Q), in a four dimensional space-time, defined by

$$F(\underline{x},t) = 0 \quad . \tag{1.104}$$

For such a space-time the components (n_1, n_2, n_3, n_0) of the unit normal to Q are defined by

$$n_1/(\partial F/\partial x_1) = n_2/(\partial F/\partial x_2) = n_3/(\partial F/\partial x_3) = n_0/(\partial F/\partial t) \quad , \tag{1.105}$$

and

$$n_\alpha n_\alpha + n_0^2 = 1 \quad . \tag{1.106}$$

[Note: $A_\alpha B_\alpha$ represents $\sum\limits_{m=1}^{3} A_m B_m$.]

Q represents a surface γ_t moving in space. A curve $\{\underline{x}(\beta),t(\beta)\}$ on Q therefore represents a point $\underline{x}(t)$ in space which lies at time t on γ_t . Denote the velocity of this point by V, then

$$V^2 = \dot{x}_\alpha \dot{x}_\alpha \quad ; \quad (\dot{x}_\alpha = \partial x_\alpha / \partial t) \quad . \tag{1.107}$$

Since $F(\underline{x}(t),t)=0$, using (1.105) and (1.106), it can be shown that

$$n_\alpha \dot{x}_\alpha + n_0 = 0 \quad . \tag{1.108}$$

This implies that

$$V^2 \geq n_0{}^2 / n_\alpha n_\alpha \quad ; \tag{1.109}$$

as (1.108) represents the dot product between the velocity vector and the vector normal to γ_t in space. ($V^2, n_\alpha n_\alpha$, and n_0^2 are the squares of the magnitudes of the velocity, normal, and resultant vectors respectively.) The equality, in (1.109), applies when V becomes the velocity υ of a point on γ_t moving in a direction normal to γ_t . In other words, when γ_t is propagating normal to itself,

$$\upsilon^2 = n_0{}^2 / n_\alpha n_\alpha \quad . \tag{1.110}$$

A surface Q for which $\upsilon = c/\nu$, that is

$$n_\alpha n_\alpha = n_0{}^2 \ (\nu^2/c^2) \quad , \tag{1.111}$$

is shown later in this section to be a "characteristic" of the wave equation (1.36). A characteristic represents a surface in space which propagates normal to itself - each point on the surface not necessarily propagating at the same speed. Such a surface is a wave front in the sense of geometric optics.

The Eikonal Equation and Optical Path Length

If (1.104) is replaced by

$$t = \tau(x) \tag{1.112}$$

(i.e. $F(\underline{x},t) = t - \tau(\underline{x})$), then, using (1.105),

$$\partial \tau / \partial x_\alpha = -n_\alpha / n_0 \quad ; \quad \alpha = 1,2,3 \quad . \tag{1.113}$$

Equations (1.111) and (1.113) can be combined to obtain

$$|\nabla \tau|^2 = \nu^2/c^2 \quad . \tag{1.114}$$

This is the eikonal equation. Geometrical optics rays are curves in space orthogonal to the wave front, and are hence in the direction of $\nabla \tau$.

The length of the optical path between two points $\underline{x}_0$ and $\underline{x}$ is the geodesic distance

$$s(\underline{x}_0, \underline{x}) = \int_{\underline{x}_0}^{\underline{x}} (\nu/c) \; |d\underline{x}| \qquad (1.115)$$

$$= s(\underline{x}, \underline{x}_0) \quad . \qquad (1.116)$$

The integral is taken along a geodesic. If $\underline{x} = \underline{x}(\sigma)$ is the equation of this geodesic in terms of a parameter , and $\underline{x}(0) = \underline{x}_0$, then

$$\nabla s = (\nu/c) \; \underline{x}'(\sigma)/|\underline{x}'(\sigma)| \quad . \qquad (1.117)$$

Thus the optical path length is a solution to the eikonal equation, and therefore the surface defined by $t = s$ also satisfies (1.111).

It now remains to be shown that "characteristics" of the wave equation (1.36) are described by (1.115).

Dependence and Influence Domains

A disturbance represented by a solution of the wave equation has the property that if it is confined at one instant, say $t = 0$, to a domain D_0, then the domain D_t affected by it at a subsequent time t is deducible from D_0 by geometrical optics. This seems intuitively obvious, but deriving it from the wave equation is another matter. It can, however, be shown to be the case in the following manner.

Multiplying the wave equation (1.36) by $2\partial\psi/\partial t$ results in

$$0 = 2 \; \partial\psi/\partial t \; \{(\nu^2/c^2) \; \partial^2\psi/\partial t^2 - \nabla^2\psi\} \qquad (1.118)$$

$$= \partial\{(\nu^2/c^2)(\partial\psi/\partial t)^2\}/\partial t - 2\nabla\cdot\{(\partial\psi/\partial t)\nabla\psi\} + 2\nabla\psi\cdot\nabla(\partial\psi/\partial t) \qquad (1.119)$$

$$= \partial\{(\nu^2/c^2)(\partial\psi/\partial t)^2 + |\nabla\psi|^2\}/\partial t - 2\nabla\cdot\{(\partial\psi/\partial t)\nabla\psi\} \quad . \qquad (1.120)$$

Now apply the divergence theorem. The integral of (1.120) over a domain in space time is equivalent to the integral

$$0 = \int_Q \{n_0[(v^2/c^2)(\partial\psi/\partial t)^2 + (\partial\psi/\partial x_\alpha)(\partial\psi/\partial x_\alpha)] - 2(\partial\psi/\partial t)n_\alpha(\partial\psi/\partial x_\alpha)\} \, dq$$

$$(1.121)$$

over the hypersurface bounding the domain. (n_1,n_2,n_3,n_0) are the direction cosines of the unit normal to Q. The integrand of (1.121) can also be written as

$$\{\sum_{\alpha=1}^{3}[(\partial\psi/\partial x_\alpha) - (n_\alpha/n_0)(\partial\psi/\partial t)]^2 + [(v^2/c^2) - n_\alpha n_\alpha/n_0^2](\partial\psi/\partial t)^2\}n_0 \quad .$$

$$(1.122)$$

Let $(\underline{x}_0,t_0)$ be a point of space-time with $t_0 > 0$. Then the characteristic conoid, defined by

$$C : \quad t = t_0 - s(\underline{x}_0,\underline{x}) \quad , \qquad (1.123)$$

the three dimensional hyperplane $t = 0$, and another hyperplane $t = t'$ (where $0 \le t' \le t_0$) bound a domain to which (1.121) can be applied (see Fig. 1.5). On $t=0$; $n_0 = -1$ and $n_1 = n_2 = n_3 = 0$. On $t = t'$; $n_0 = 1$ and $n_1 = n_2 = n_3 = 0$. And on C; $n_\alpha n_\alpha = n_0^2(v^2/c^2)$ as shown by (1.111)-(1.117). Using the integrand (1.122), it is seen that the contribution to the integral (1.121) from C is non-negative. Therefore it follows (since the second term in the integral disappears on $t = 0$, & $t = t'$) that

$$\int_{s \le t_0 - t' \atop t = t'} [(v^2/c^2)(\partial\psi/\partial t)^2 + (\partial\psi/\partial x_\alpha)(\partial\psi/\partial x_\alpha)] \, dx$$

$$\le \int_{s \le t_0 \atop t = 0} [(v^2/c^2)(\partial\psi/\partial t)^2 + (\partial\psi/\partial x_\alpha)(\partial\psi/\partial x_\alpha)] \, dx \quad . \qquad (1.124)$$

Thus it is seen that $\psi(\underline{x},t')$ within the region $s \le t_0 - t'$ depends on the value of $\psi(\underline{x},0)$ within the region $s \le t_0$, as both integrands are non-negative. Call the interior of C the dependence domain of $(\underline{x}_0,t_0)$. Denote the dependence domain by $\theta(\underline{x}_0,t_0)$. The points in θ satisfy the inequality $t \le t_0 - s(\underline{x}_0,\underline{x})$. Therefore $\psi(\underline{x}_0,t_0)$ is determined by all points within $\theta(\underline{x}_0,t_0)$. These points constitute the geodesic sphere

$$s(\underline{x}_0,\underline{x}) \le t_0 - t \quad . \qquad (1.125)$$

Imagine that ψ is modified by a localised impulse applied at $\underline{x}_1,t_1$. It is seen that $\psi(\underline{x},t)$ will be affected only if

$$(\underline{x}_1,t_1) \in \theta(\underline{x},t) \quad . \qquad (1.126)$$
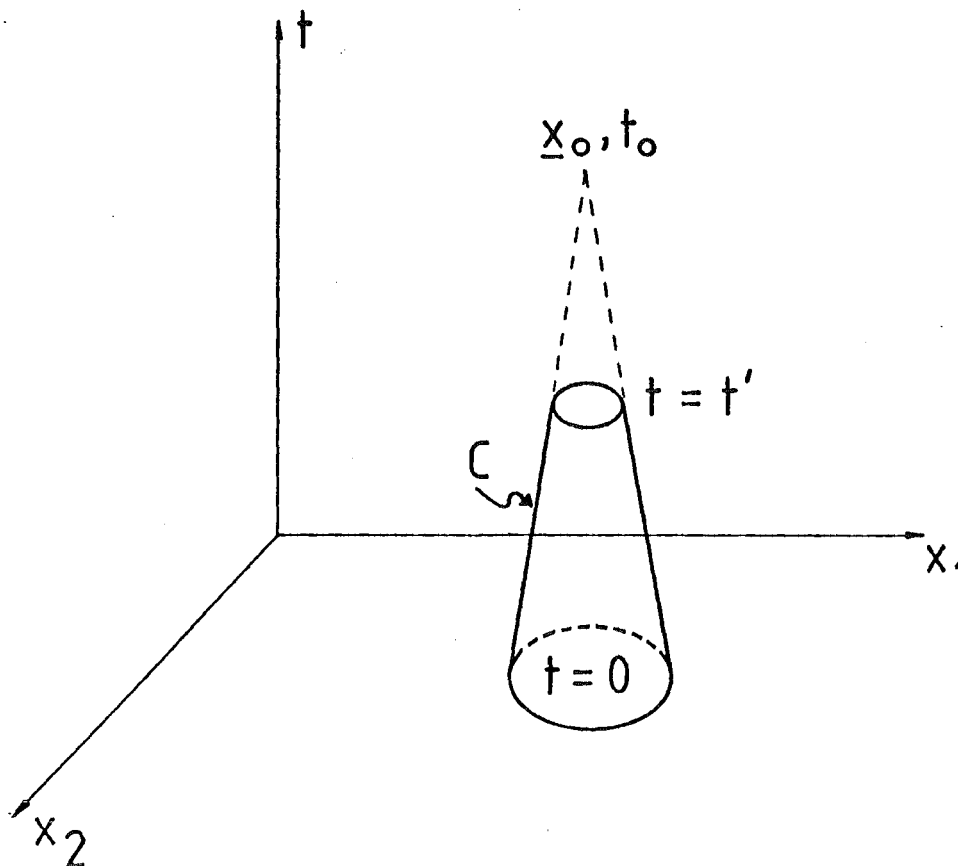
Fig. 1.5   A representation of the domain, in the four dimensional
            hyperspace, over which equation (1.121) is integrated.

The set of all points which satisfy this condition will be called the
influence domain $\theta'(\underline{x}_1, t_1)$ of $(\underline{x}_1, t_1)$, and is defined by $t \geq t_1 + s(\underline{x}_1, t)$.
θ' is the interior of the characteristic conoid with vertex $(\underline{x}_1, t_1)$
situated in $t \geq t_1$.   The frontier of $\theta'(\underline{x}_1, t_1)$ represents a wave front
diverging from $\underline{x}_1$, starting at time t_1 .   Thus the hypersurface defined
by $t = t_1 + S(\underline{x}_1, t)$ is a characteristic of the wave equation.

The above shows that the propagation of pulse fronts can be
described in terms of geodesic ray paths. When pulses are used to probe
refracting objects, the time taken for a pulse to travel between two
points is directly proportional to the optical path length. By measuring

the travel time, a direct measurement of the integral of the refractive index along a ray path is obtained. The problem now is how best to use these "integrals of the refractive index" to image the refractive index distribution.

1.8 IMAGE RECONSTRUCTION FROM PROJECTIONS

By considering the wavefunction to travel as rays the inverse scattering problem becomes very similar to another inverse problem which has a simple solution. In mathematical literature it is found under "Radon Transforms" (Gelfand et al., chapter 1, 1966), and in more applied works it is known as "Image Reconstruction from Projections Theory" (Bates and Peters, 1971). This theory is concerned with the reconstruction of a function $\lambda(\underline{x})$ from knowledge of the straight line integrals over λ, taken at all different orientations and positions. An integral of $\lambda(\underline{x})$ over a straight line is known as a projection. In two dimensions the projection $f(\rho\hat{\underline{\xi}})$ is defined by (see Fig. 1.6)

$$f(\rho\hat{\underline{\xi}}) = \int_{\hat{\underline{\xi}}\cdot\underline{x}=\rho} \lambda(\underline{x}) \, d\ell \quad . \tag{1.127}$$

The unit vector $\hat{\underline{\xi}}$ is perpendicular to the line of integration; ρ is the perpendicular distance between this line and the origin; and $d\ell$ is the differential element along this line (notice that $\hat{\underline{\xi}}\cdot\underline{x}=\rho$ defines the line of integration). Making use of delta function notation (1.127) can be rewritten as

$$f(\rho\hat{\underline{\xi}}) = \iint \lambda(\underline{x}) \, \delta(\rho - \hat{\underline{\xi}}\cdot\underline{x}) \, d\underline{x} \quad , \tag{1.128}$$

where the integration is taken over all $\underline{x}$. Now the one dimensional Fourier transform of $f(\rho\hat{\underline{\xi}})$ with respect to ρ is, using (1.128),

$$F(\alpha\hat{\underline{\xi}}) = \iiint \lambda(\underline{x}) \, \delta(\rho - \hat{\underline{\xi}}\cdot\underline{x}) \, \exp(j2\pi\alpha\rho) \, d\underline{x} \, \partial\rho \tag{1.129}$$

$$= \iint \lambda(\underline{x}) \, \exp(j2\pi\alpha\hat{\underline{\xi}}\cdot\underline{x}) \, d\underline{x} \quad . \tag{1.130}$$

Writing $\alpha\hat{\underline{\xi}} = \underline{\omega}$, it is apparent that $F(\underline{\omega})$ is the two-dimensional Fourier transform of $\lambda(\underline{x})$. Thus the object $\lambda(\underline{x})$ can be estimated from the
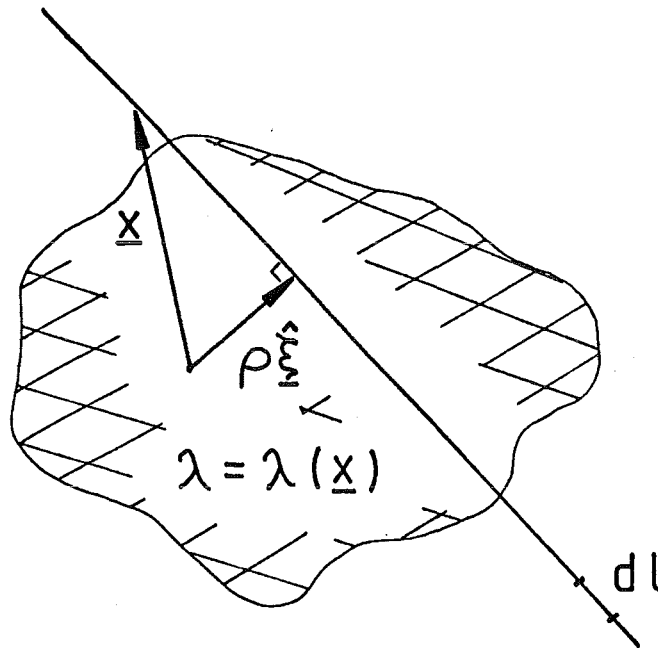
Fig. 1.6   The projection $f(\rho\hat{\xi})$ is the integral of λ

along the line perpendicular to $\rho\hat{\xi}$.

projections by firstly taking the one dimensional Fourier transform to obtain $F(\underline{\omega})$, and then taking the two dimensional inverse Fourier transform of $F(\underline{\omega})$. [In practice, λ is generally obtained using the modified back projection algorithm (see §3.2). The two methods are of course equivalent, but the modified back projection algorithm is numerically more efficient.]

This theory is used very successfully in computerised tomography X-ray body scanners. However there is a difficulty in extending the results to electromagnetic and acoustic transmission imaging. With X-rays the radiation can usefully be treated as though it travelled along straight line rays. By measuring the attenuation of the rays passing through the object being imaged, the integral of the absorbtivity along a straight line between the X-ray source and the detector is obtained. This,then,corresponds to a projection as per (1.127). From these projections an image of the X-ray absorbtivity can be produced.

When acoustic and electromagnetic propagation is treated as rays it has to be remembered that the rays are in general curved. This means that the resulting "projections" are not equivalent to straight line projections. However if ray curvature is not too great, reasonable images can be obtained by assuming the rays to be straight and using the reconstruction from straight line projections procedure. Further discussion on the effect of ray curvature, as regards the acoustic imaging problem, is given in chapter 4.

It is seen that, with the exception of computed tomography, there is still a long way to go before techniques such as those discussed in this chapter can be implemented into practical imaging schemes. As mentioned at the beginning of this chapter, the majority of acoustic imaging devices are based on the almost intuitive pulse echo location principles (see for example Wells, 1977). As soon as one enquires how to improve on these simple ideas, the whole business becomes much more complicated - both conceptually and procedurally.

2.   THE STRUCTURE OF THE D.N.A. MOLECULE

2.1   INTRODUCTION

Molecular structures can be imaged using X-rays. As the wavelength of X-rays is comparable to the bond lengths within molecules, atomic resolution is possible. However, unlike light rays, at the present time, no means is known by which X-rays may be focused. Hence direct visualization of an image, as would be formed in an "X-ray microscope", is beyond present technology. For this reason one must resort to analysing the X-ray scattering (or diffraction) in order to "image" the molecule. To determine the molecular structure immediately from the diffraction it is necessary to know the phase as well as the amplitude of the scattered X-rays. However the technology does not exist at present for measuring the phase of X-rays. There are two main reasons why this is so. Firstly, because of the very short wavelengths of X-rays, tolerances in the measuring equipment would have to be in the order of 10^{-11} m . And secondly, in presently available X-ray sources, the emission of an X-ray is a random event. Since, with these sources, there is no well-defined phase relationship between incident waves, the measuring of useful phase relationships between diffracted waves is prevented.

In the absence of *a priori* information about the structure, the lack of phase information causes ambiguity. That is, several structures may be compatible with the observed X-ray diffraction. Nevertheless, if constraints - such as the atoms are positive real, and approximations - such as treating the atoms as "point like" scatterers, are made, it is quite often possible to determine molecular structure in the absence of phase information (see for example Sherwood, 1976). However, these procedures tend to be unstable, especially in the presence of errors in the data. Therefore, they are usually only suitable for determining simple structures.

The general method of imaging large molecules is to propose a model for the structure. From the model the theoretical X-ray diffraction is calculated and compared with the actual diffraction pattern. If the agreement is high then the model is at least consistent with the diffraction data. To have any chance at all of obtaining the

"unique model" much use must be made of any available *a priori* knowledge about the molecule. Thus the model must contain the known chemical constituents in the correct proportions; all the bond lengths and angles between atoms must be within allowable tolerances; it must explain the observed diffraction data and other experimentally determined characteristics; and, in the case of "biological molecules", the structure must be consistent with the molecule's bio-chemical function. In this chapter, two different models for the D.N.A. molecule are compared with three sets of X-ray diffraction measurements. Even though there are many factors which go into determining the actual structure of a molecule, in general, the X-ray diffraction data is the main clue. However the available D.N.A. diffraction data is so poor that one can not really do much more than comment on the gross features of the molecule.

D.N.A. (deoxyribonucleic acid) is the biological molecule responsible for preserving and propagating the genetic code. In the mid 1950's Watson and Crick (1953) proposed their, now familiar, double helix (D.H.) model for D.N.A. The double helix model is a twisted ladder like arrangement (see Fig. 2.1).
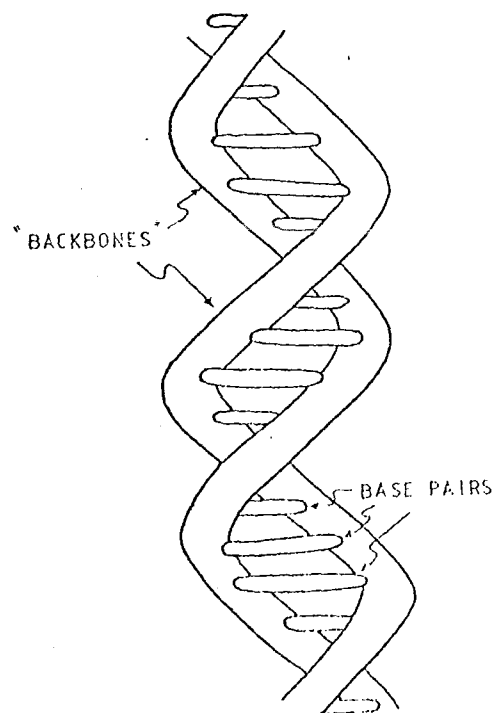

Fig. 2.1   A representation of the Double Helix Model for D.N.A.

The "rungs" consist of base pairs. These are strung together by two strands, consisting of alternating sugar and phosphate "groups", running the length of the molecule. There are four different types of bases denoted by A, T, G and C (adenine, thymine, guanine and cytosine). The structure of the bases is such that hydrogen bonds can link A with T and G with C to form the base pairs A-T and G-C. However A or T will not pair with G or C. The order in which the base pairs are arranged along the molecule determines the genetic code. The structures of both base pairs are very similar, and as far as X-ray diffraction is concerned they can be treated as being identical (see §2.5). The two strands, or backbones, twist around each other to form a right-handed double helix.

D.N.A. is found within the chromosomes of cells. When a cell divides, the D.N.A. also divides (along the longitudinal axis) - one half going to each of the two daughter cells. The D.H. model has one major shortcoming with respect to this splitting process. To illustrate take two strands of wire twisted around each other - the double helix. Now as the cells divide, the strands have to separate. Try to pull the strands apart. Unless the wires are short, a tangle will have formed. To successfully separate the wires the twisted part has to be supported and untwisted, while tension is maintained on the separated section of each strand. D.N.A. is a very long molecule - *E. coli* (a bacteria) D.N.A. is about 1 mm. Watson and Crick's model proposes one twist per 34 Å. Cells divide in approximately 30 minutes (these figures are from Górski, 1975). Therefore if the D.N.A. molecule was to unwind in order to separate, it would have to do so at 150 revs/sec. Obviously the analogy with the twisted wire is rather crude. However, if D.N.A. actually does have a D.H. conformation, the splitting mechanism must be reasonably sophisticated; and all sorts of enzymes have been postulated to this end (see Bates *et al.*, 1977, for a review of the separation problem and suggested mechanisms).

In the mid 1970's Gordon Rodley, of the Chemistry Department at this University, originated the "side-by-side" (S.B.S.) model for D.N.A. (Rodley *et al.*, 1976). This model contains the same base pairing scheme, and hence the same basic chemical structure, as Watson and Crick's. The difference is that the two backbones are not intertwined. They both twist first one way and then the other - hence they remain side-by-side (see Fig. 2.2).
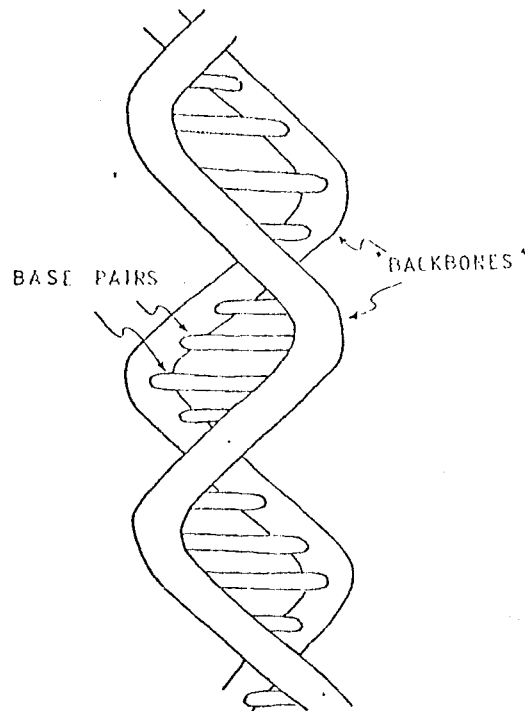
Fig. 2.2   A representation of the Side-By-Side model for D.N.A.

The S.B.S. model can be considered as a series of alternating left and right handed double helical sections. The net twist per section is less than 180°. Rodley's model divides by "unzipping" rather than untwisting. The results reported in this chapter arose out of the need to check the compatibility of the S.B.S. model with the X-ray diffraction data.

In §2.2 necessary theoretical preliminaries are collected; and the theoretical relationship between the molecular structure and the observed diffraction pattern is discussed. The diffraction pattern is influenced by both the degree of crystallisation and the relative humidity (or water content) of the D.N.A. specimens. The X-ray data used in this study (Feughelman et al., 1955; Bram, 1977; Zimmerman and Pheiffer, 1979) were obtained from fibrous B-DNA. "Fibrous" indicates that the molecules within the specimen are orientated with their long axes all in the same direction, but that the molecules are randomly rotated with respect to each other. The "B" relates to the water content. The specimen is prepared in vitro, and is comprised of many segments of D.N.A. which have been separated from other cellular matter. The specimen is dried, and as it dries each molecule is compressed more and more by its neighbours. These forces influence the arrangement adopted by the molecules without changing their gross conformation (see

Bloomfield *et al.*, 1974).

One of the problems, in dealing with structures containing a large number of atoms, is in developing computationally efficient methods of comparing the data with the model. Axial Patterson analysis, §2.3, is a particularly simple method of achieving this. The axial Patterson function reflects only the gross structure of the model and so is useful in comparing the general features (such as an S.B.S. or D.H. conformation) with the data. It is interesting that this simple test favours the S.B.S. model. Of course, the significance of this depends upon how reliable are the data. In §2.4 the full theoretical diffraction patterns, as calculated from the two models, are compared with the diffraction data. With respect to the diffraction patterns, both models appear satisfactory. There is, however, a large variability within the data. The possibility of the base pairing sequence influencing the diffraction pattern is examined, §2.5, but the effect is shown to be negligible. Despite the fact that the different sets of diffraction data differ greatly, the axial Pattersons calculated from each are remarkably similar. The significance of this is discussed in §2.6. Finally, in §2.7, the results presented in this chapter are assessed.

The initial work to determine the compatibility of an S.B.S. model with the D.N.A. diffraction data was carried out using idealised models. These models were comprised of continuous filamentary strands rather than discrete atoms. This simplified approach was taken because of the computational time required to calculate the full theoretical diffraction pattern for a "discrete atom" model. The strands mimicked the sugar-phosphate backbone of the molecule. Now because the phosphorous atoms have the highest atomic number among the D.N.A. atoms, they are responsible for a large part of the scattering. Therefore the scattering from a "filamentary strand" model is a good first approximation to that of the "discrete atom" model. The early results, using simplified models, showed that the S.B.S. was a viable alternative structure for D.N.A. (Rodley *et al.*, 1976; Bates *et al.*, 1977). The next step was to compare detailed models, which included all the atoms (except hydrogen - as their contribution to the scattering is insignificant), with the diffraction data. After modifying the models built by Rodley only slightly, good agreement was obtained. The remarkable thing here was that the models were built with very little regard to the diffraction pattern - chemical bonding "rules" and

intuition were the main factors in determining where an atom was positioned. Perhaps the results were fortuitous, but the fact remains. At this stage resistance to our work was becoming apparent. Three sequential papers were submitted to *Acta Crystallographica* in 1978. The referees rejected them, making reasonable and helpful comments. The papers were modified appropriately and resubmitted. Seven months later they were again rejected, this time for almost unbelievable reasons. It was felt that the referees comments were so outrageous that a research report (Bates *et al.*, 1978) was produced so that this situation would be recorded. The report contains the three papers concerned, a discussion of a relevant question neglected in the papers, the referees' comments, and a rebuttal of the referees' comments. Recent results concerning the S.B.S. model have been published in *Pramāna* (Bates *et al.*, 1980). The present scientific climate is becoming more receptive to alternative models of D.N.A. Recent studies of V-D.N.A. by Stettle *et al.* (1979) and of Z-D.N.A. by Wang *et al.* (1979) support the idea that D.N.A. can adopt more than one stable conformation. Another group, Sasisekharen *et al.* (1978), is also working on an S.B.S. model. The work on our S.B.S. model is still continuing at this University - mainly on optimising some of the bond lengths and angles between atoms in the bend region, which separates the left and right handed helical sections.

The people who are/have been directly involved here with this work are: G.A. Rodley, R.S. Scobie, G.B. Jameson, K.C. Ng, J.P. Day - Chemistry Department; R.H.T. Bates, R.M. Lewitt, C.H. Rowe, P.M. Andrea, R.P. Millane and myself - Electrical Engineering Department. My contribution has been in calculating axial Pattersons, and fibre diffraction patterns, from models consisting of discrete atoms. By comparing these with axial Pattersons calculated from the data, and actual diffraction patterns, one can determine how the model can be refined. The reader may well ask how Electrical Engineers can contribute to the determination of a molecular structure. The answer is that X-ray diffraction is very similar to electromagnetic and acoustic diffraction, which is among the major research preoccupations of Electrical Engineers nowadays.

The results contained in this chapter pertain to the S.B.S. model reported in *Pramāna* (Bates *et al.*, 1980).

2.2   D.N.A. DIFFRACTION

The scattering of X-rays by molecules can be described quite accurately by the Born approximation (Cowley, 1975; see also §1.6). Thus the diffraction pattern is related to the molecular structure via a Fourier transform relationship.

The Fourier transform, $F\{\cdot\}$, of a function $f(\underline{x})$ is defined as

$$F(\underline{u}) \equiv F\{f(\underline{x})\}$$

$$= \int_{-\infty}^{\infty} f(\underline{x}) \exp(j2\pi\underline{u}\cdot\underline{x}) \, d\underline{x} \quad . \tag{2.1}$$

The inverse transform, $F^{-1}\{\cdot\}$, is defined so that

$$f(\underline{x}) \equiv F^{-1}\{F\{f(\underline{x})\}\}$$

$$= \int_{-\infty}^{\infty} F(\underline{u}) \exp(-j2\pi\underline{u}\cdot\underline{x}) \, d\underline{u} \quad . \tag{2.2}$$

One interesting property of the Fourier transform is the relationship between multiplication and convolution. Denote convolution by the symbol $*$. That is

$$f(\underline{x}) * g(\underline{x}) = \int_{-\infty}^{\infty} f(\underline{x}) \, g(\underline{y}-\underline{x}) \, d\underline{x} \quad . \tag{2.3}$$

Now the Fourier transform of the convolution is

$$\iint_{\infty}^{\infty} f(\underline{x}) \, g(\underline{y}-\underline{x}) \, d\underline{x} \, \exp(j2\pi\underline{y}\cdot\underline{u}) \, d\underline{y} \tag{2.4}$$

$$= \int_{-\infty}^{\infty} f(\underline{x}) \, \exp(j2\pi\underline{x}\cdot\underline{u}) \, d\underline{x} \int_{-\infty}^{\infty} g(\underline{y}) \, \exp(j2\pi\underline{y}\cdot\underline{u}) \, d\underline{y} \tag{2.5}$$

$$= F(\underline{u}) \, G(\underline{u}) \quad . \tag{2.6}$$

((2.5) is obtained by redefining the variable $\underline{y}$ equal to $\underline{y}-\underline{x}$ in (2.4).) Thus

$$F\{f * g\} = F \, G \quad . \tag{2.7}$$

A similar result occurs for the correlation of two functions. Denote correlation by $\odot$, then

$$f(\underline{x}) \odot g(\underline{x}) = \int_{-\infty}^{\infty} f(\underline{x}) \; g(\underline{x} + \underline{y}) \; d\underline{x} \quad . \tag{2.8}$$

Using an argument similar to that used to obtain (2.7), it can be shown that

$$F\{f \odot g\} = F \; G^* \quad , \tag{2.9}$$

where G^* is the complex conjugate of G. A corollary of (2.9) is the relationship

$$F^{-1}\{|F|^2\} = f \odot f \quad . \tag{2.10}$$

This relationship says, in effect, that given the intensity of a function $(|F|^2)$, only the autocorrelation of the transform of that function can be obtained (these Fourier transform relationships are explained in Bracewell, 1965).

X-rays are scattered by the electrons within the molecules. The electric field component of the incident X-rays causes the electrons to oscillate. Each electron is therefore being accelerated, and so gives rise to electromagnetic radition. This radiation constitutes the "diffracted field" (Sherwood, 1976). In the analysis of the D.N.A. molecule presented in this chapter, only the scattering from the non-hydrogen atoms is considered, as the relative contribution of the hydrogen atoms is insignificant compared with the errors in the data.

Denote the electron density of the molecular structure by $\rho(r,\theta,z)$. Then the X-ray diffraction $F(r,\psi,\omega)$ is given by (Cowley, 1975)

$$F(R,\psi,\omega) = \int_{\infty}^{\infty} \int_{0}^{\infty} \int_{0}^{2\pi} \rho(r,\theta,z) \; \exp\left(j2\pi[Rr\cos(\theta-\psi)] + \omega z\right) \; rd\theta dr \, dz \quad , \tag{2.11}$$

where (r,θ,z) and (R,ψ,ω) represent cylindrical polar coordinates in the real and Fourier spaces respectively. Unfortunately it is not possible to measure the phases of X-rays - only intensity information can be obtained. Therefore only the modulus of the complex quantity F is available to estimate ρ . This is the so called "phase problem" (see for instance Bates, 1978). The only unambiguous quantity that can be

obtained from $|F|$ is the autocorrelation of ρ (see (2.10)). If the structure is very simple it is usually possible to unravel the autocorrelation using a meagre knowledge of chemical bonding requirements. For modest sized structures quasi-holographic techniques, such as "heavy atom replacement" (Sherwood, 1976), can be used. In this procedure an atom in the molecule is replaced by a much heavier atom. The heavy atom acts like a delta function. Write ρ as

$$\rho = \tilde{\rho} + \delta \quad , \tag{2.12}$$

where δ represents the heavy atom. The autocorrelation is

$$\rho \odot \rho = \tilde{\rho} \odot \delta + \tilde{\rho} \odot \tilde{\rho} + \delta \odot \delta + \delta \odot \tilde{\rho} \quad . \tag{2.13}$$

If the replacement atom is heavy enough it is responsible for most of the scattering, thus the $\tilde{\rho} \odot \tilde{\rho}$ term can be ignored, as a first approximation. The $\delta \odot \delta$ term is a large intensity point at the origin, and the remaining two terms provide the "real" and "inverted" images of $\tilde{\rho}$. When it comes to determining the structure of macromolecules, the ambiguity inherent in the autocorrelation is, however, more difficult to overcome.

The X-ray diffraction patterns show that D.N.A. is periodic along the longitudinal axis (hereafter called the z axis). Each period is known as a "repeat unit". Denote the electron density of a repeat unit by $e(r,\theta,z)$. Then

$$\rho(r,\theta,z) = \sum_{m=-\infty}^{\infty} e(r,\theta,z-mc) \quad , \tag{2.14}$$

where c is the length of the repeat unit, and e exists over the range $0 \leq z < c$. The diffraction of a structure periodic in one direction is observed only on layer planes, each separated by $1/c$ in Fourier space (this can be shown to be the case by considering the Fourier transform of a periodic structure, see Sherwood, 1976). The complex amplitude of the diffraction pattern on the ℓth layer plane $E_\ell(R,\psi)$ is given by

$$E_\ell(R,\psi) = \int_o^c \int_o^{\hat{r}} \int_o^{2\pi} e(r,\theta,z) \exp(j2\pi[Rr \cos(\theta-\psi) + \ell z/c]) \, r d\theta dr dz \quad , \tag{2.15}$$

where $\hat{r}$ is the radius of the circular cylinder which just encloses all the electron density of a repeat unit. It is very difficult to form good crystals of D.N.A. In many specimens the only ordering that can be achieved is the lining up of the axes of individual molecules. These specimens are known as fibre specimens and their measured diffraction patterns are called fibre patterns. The intensity of a fibre pattern is proportional to the rotational average of the diffracted intensity from a single molecule. The layer planes thus degenerate into layer lines. The normalised intensity of the ℓth layer line, $\Omega_\ell(R)$, is given by

$$\Omega_\ell(R) = 1/2\pi \int_0^{2\pi} |E_\ell(R,\psi)|^2 \, d\psi \quad . \tag{2.16}$$

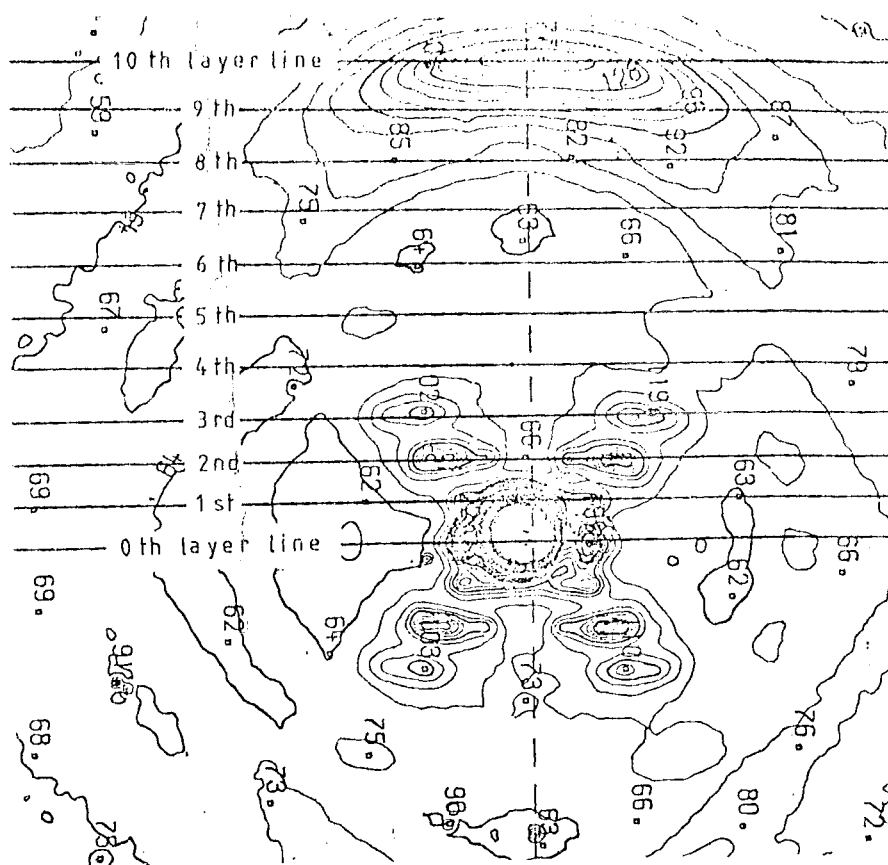Figure 2.3 is an isointensity densitometer plot of a D.N.A. diffraction pattern (Bram, 1977).

Fig. 2.3 Densitometer plot of a D.N.A. diffraction pattern (Bram, 1977).

The layer lines, spaced by 1/c, are indicated by the solid lines. After the third layer line they become difficult to identify. The pattern is symmetrical about the zeroth layer line, and also the vertical axis. The large intensity at the centre of the tenth layer line indicates that there is another strong periodic feature within the molecule. This "feature" is the regularly spaced base pairs - spaced at one-tenth of a repeat unit. The cross shaped region of high diffracted intensity, and the diamond shaped regions of low intensity, are characteristic of the diffraction from a helical structure. This sent Watson and Crick in search of a helical model; although it was realised at the time that there were many obstacles to uniquely interpreting the X-ray diffraction data (Wilkins *et al.*, 1953). Bates, in Bates *et al.* (1980), demonstrates that a wide class of structures can potentially give rise to these features in the diffraction pattern - the S.B.S. structure being one of them. In the following analysis, §§2.3 - 2.7, it is assumed that the intensity exists only along the layer lines. Thus the diffraction patterns are discussed in terms of the Ω_ℓ's. An infinite perfectly periodic structure would produce sharp layer lines. However the samples are in thermal motion, and probably only periodic for short lengths, hence the lines are blurred. This blurring, becoming more severe as one moves out from the origin, effects the resolution. The spacing of the layer lines is $1/34 \ \text{Å}^{-1}$. The diffraction pattern can be estimated quite well out to about the third layer line; between that and the tenth the diffraction data deteriorates. This means that good resolution of the molecule can be obtained down to about $11 \ \text{Å}$ $(34/3 \ \text{Å})$, and only mediocre resolution between $11 \ \text{Å}$ and $3.4 \ \text{Å}$. $3.4 \ \text{Å}$ is the resolution limit of the data. This is very marginal when it is realised that the repeat unit is only $34 \ \text{Å}$ long; bond distances are the order of $1.5 \ \text{Å}$; and there are 410 atoms (excluding hydrogens) per repeat unit, to determine the positions of. This is in addition to the theoretical ambiguity inherent in the Ω_ℓ's.

2.3 AXIAL PATTERSON

The obvious method of checking the suitability of a model is to compare the theoretical diffraction, as calculated from the model, with the actual diffraction pattern. However, there are two main reasons for looking at other methods of comparing the model with the data. Firstly, the Ω_ℓ's take a long time to calculate (about two hours on a PDP 11/70

computer), as effectively a three-dimensional Fourier transform of a 410 atom repeat unit has to be computed. And secondly, it is difficult (for most people) to understand how features in the Fourier transform relate to features in the molecule. This is because the effect of a particular point in "object space" is spread out over all of "Fourier space". Thus changing the positions of one or two atoms effects the whole diffraction pattern; and similarly every point on the diffraction pattern is due to the effect, in varying degrees, of every atom in the molecule. Mainly because of this latter reason the diffraction data is usually analysed in terms of its Fourier transform. As noted in §2.2 only the intensity of the diffraction pattern can be measured, thus the Fourier transform of the data is the autocorrelation of the molecule. In crystallography the specimens are periodic, therefore the auto-correlation is itself periodic. The crystallographic term for an autocorrelation of a periodic structure is the "Patterson" or "Patterson function" (Sherwood, 1976). It is generally easier for model builders to think in terms of the effects that features in their models have on the Patterson; rather than in terms of the effects that features have on the Fourier transform.

With fibrous D.N.A., the difficulty in interpreting the data is compounded by the random rotation of individual molecules in the specimens. Taking the Fourier transform of the Ω_ℓ's results in an "angularly averaged Patterson". Denote the Patterson by $P(r,\theta,z)$. Then the angularly averaged Patterson $\tilde{P}(r,z)$ is defined by

$$\tilde{P}(r,z) = 1/2\pi \int_0^{2\pi} P(r,\theta,z) \, d\theta \quad . \tag{2.17}$$

Notice that for $r = 0$, $\tilde{P} = P$. Thus the on axis form of the angularly averaged Patterson is equivalent to a straightforward autocorrelation with the shift being taken along the z axis of the molecule. This on axis form of the Patterson is called the "axial Patterson", P_{ax} (Bates et al., 1980).

The axial Patterson is calculated from the data as follows. P is given by Fourier transforming the diffraction pattern.

$$P(r,\theta,z) = \int_\infty^\infty \int_0^\infty \int_0^{2\pi} |F(R,\psi,\omega)|^2 \exp[j2\pi(Rr\cos(\theta-\psi) + \omega z] \, Rd\psi dRdz \quad . \tag{2.18}$$

Therefore, using (2.11) and (2.14) - (2.17) it can be shown that

$$P_{ax}(z) = \tilde{P}(o,z) \tag{2.19}$$

$$= \sum_{\ell=-\infty}^{\infty} \{ \int_{0}^{\infty} \Omega_\ell(R)\,RdR \} \exp(j2\pi\ell z/c) \tag{2.20}$$

where the term inside the brackets can be obtained from the intensity along the layer lines. The diffraction pattern is symmetrical, thus $\Omega_\ell(R) = \Omega_{-\ell}(R)$. Also the value for $\Omega_0(R)$ around the origin is unknown as it is obliterated by the incident X-ray beam (see Fig. 2.3). The effect of the Ω_0 term, in (2.20), is to shift the base line of P_{ax} up or down, but it does not alter the general "form" of P_{ax} (in electrical engineering parlance: it acts as a d.c. term). For this reason it can be omitted when calculating P_{ax} from the data, however on the following graphs it must be remembered that the base line is arbitrary. In addition to this, only those layer lines out to $\ell = 10$ are recorded. Therefore the P_{ax}'s, plotted in Fig. 2.4, are approximated to (2.20) by

$$P_{ax}(z) = \sum_{\ell=1}^{10} \{ \int_{0}^{\infty} \Omega_\ell(R)\,RdR \} \cos(2\pi\ell z/c) \quad . \tag{2.21}$$
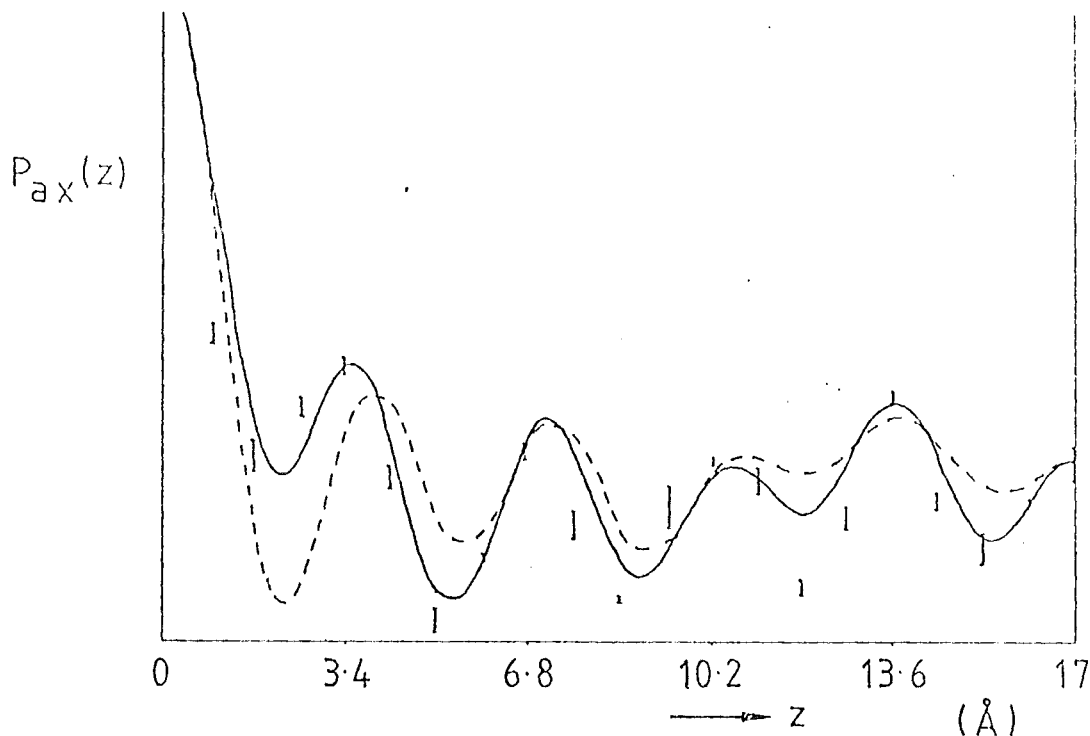

Fig. 2.4   Axial Pattersons calculated from the data: solid line - Feughelman et al., 1955; dashed line - Bram, 1977; error bars - Zimmerman and Pheiffer, 1979.

The axial Pattersons for fibrous B-DNA have a period of 34 $\overset{o}{A}$. However because they are even functions, they are shown here only from $\cdot 0$ to $17 \overset{o}{A}$ (see Fig. 2.4). Three sets of data were used in this study. The solid line (Fig. 2.4 again) was calculated using the Ω_ℓ's from Feughelman et al. (1955). The dashed line was produced from Bram's (1977) data. In this case the data were in the form of a densitometer plot (see Fig. 2.3). The positions of the layer lines were estimated, and the intensity values (corresponding to the Ω_ℓ's) were read directly off the plot. The remaining set of data, corresponding to the error bars, were obtained in the form of a photographic negative from Zimmerman and Pheiffer (1979). The positions of the layer lines were estimated, and a densitometer, scanning along these lines, was used to obtain the Ω_ℓ's. In this latter case it was difficult to be sure about the intensity on the ninth layer line - the error bars delimit the uncertainty. The integral in (2.21) was evaluated using trapezoidal integration. The P_{ax}'s for the three different sets of data were calculated, and then scaled so that the size of their peaks and troughs were roughly similar.

Notice that there is a peak every $3.4 \overset{o}{A}$. This corresponds to the separation distance of the base pairs; showing that the base pairs have a significant effect on the axial Patterson. The base pairs can be thought of as planar rectangular structures (approx. $11 \overset{o}{A} \times 5 \overset{o}{A} \times 1 \overset{o}{A}$, Arnott and Hukins, 1973). Now, in the case of B-DNA, the amplitudes of the peaks in the axial Patterson are largely due to the cross correlation between the base pairs. As the structure of the base pairs is similar to that of a long narrow rectangle, the correlation between two base pairs will be dependent on their relative angular orientation. Therefore a large amplitude peak in P_{ax}, at a point z_1 say, indicates that there are base pairs, separated by z_1, with a similar angular orientation. There is, of course, no way of distinguishing between two base pairs very closely orientated and several base pairs only partially orientated. D.N.A. consists of a whole series of planar base pairs stacked on top of each other. In terms of a model for D.N.A. the main question is, what is the stacking arrangement? Axial Patterson analysis is a useful means of investigating this question.

The axial Pattersons calculated for both the S.B.S. and D.H. models were obtained by autocorrelating the respective electron densities. The electron densities for each type of atom (denote by $h_j(r)$ for the jth atom) were assumed to be spherically symmetric.

They were determined by Fourier transforming the atomic scattering data (Lonsdale, 1968). The spherical symmetry approximation is satisfactory considering the resolution in the scattering data - and in any case it is not clear how to do any better. The electron density of a repeat unit is given by

$$e(r,\theta,z) = \sum_{j=1}^{410} h_j\big(A_j(r,\theta,z)\big) \quad , \tag{2.22}$$

where

$$A_j(r,\theta,z) = \{r^2 + r_j^2 - 2rr_j \cos(\theta-\theta_j) + (z-z_j)^2\}^{\frac{1}{2}} \quad , \tag{2.23}$$

and r_j,θ_j,z_j is the position within the model of the jth atom. The axial Patterson is given by

$$P_{ax}(z) = \int_0^{2\pi} \int_0^{\hat{r}} \int_0^{c} e(r,\theta,\alpha)\{e(r,\theta,\alpha+z) + e(r,\theta,\alpha-c+z)\}\, rd\theta dr d\alpha \quad , \tag{2.24}$$

for $0 \le z < c$. To evaluate this denote the cross-correlation between the jth and k^{th} atom by $C_{jk}(\beta)$, where β is the distance separating their centres. Therefore

$$P_{ax}(z) = \sum_{j=1}^{410} \sum_{k=1}^{410} \{C_{jk}(\beta_1) + C_{jk}(\beta_2)\} \quad , \tag{2.25}$$

where

$$\beta_1 = \{r_j^2 + r_k^2 - 2r_j r_k \cos(\theta_j-\theta_k) + (z_j-z_k+z)^2\}^{\frac{1}{2}} \quad , \tag{2.26}$$

and

$$\beta_2 = \{r_j^2 + r_k^2 - 2r_j r_k \cos(\theta_j-\theta_k) + (z_j-z_k+z-c)^2\}^{\frac{1}{2}} \quad . \tag{2.27}$$

The C_{jk} functions were evaluated and stored before the axial Pattersons were calculated. They were used in the form of "lookup tables". Because there are only four types of atoms (carbon, oxygen, nitrogen and phosphorus), there are only ten C_{jk} functions to store. The C_{jk} were obtained by inverse Fourier transforming the product of the scattering data from the jth and k^{th} atoms (refer to §2.2, (2.9) for this relationship).

The axial Pattersons calculated for the S.B.S. and D.H. models are shown in Fig. 2.5.
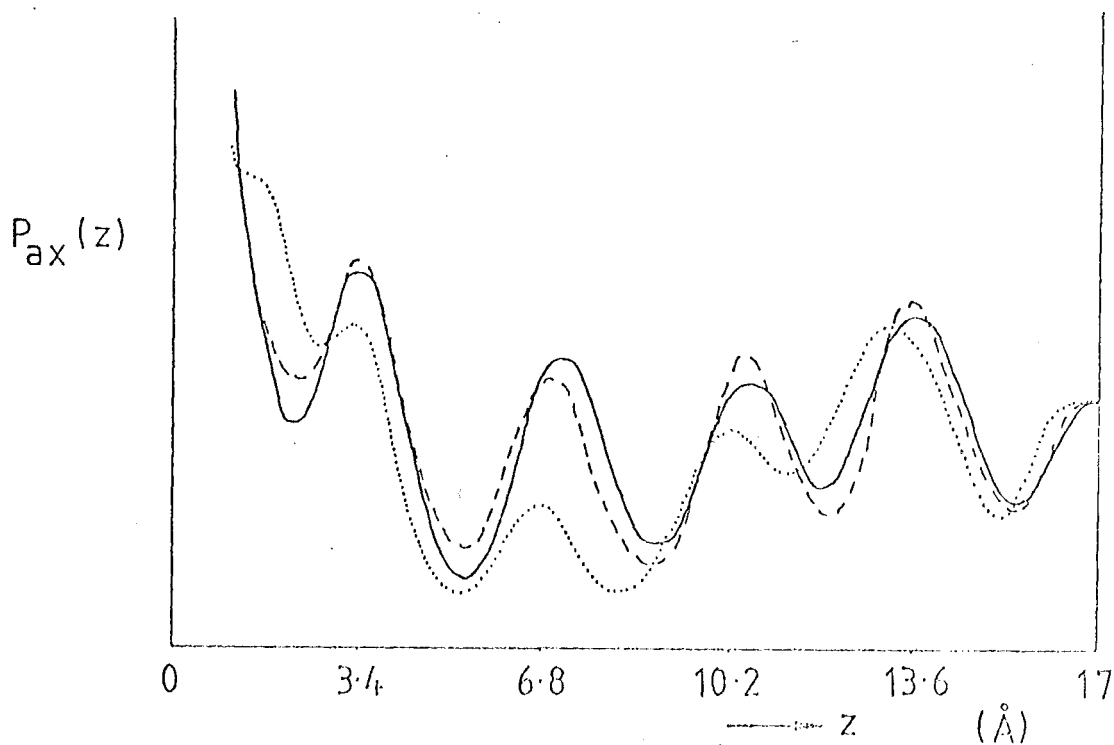

Fig. 2.5   The axial Pattersons of the two models compared with that calculated from the data; solid line - 'averaged' data; dotted line - D.H. model; dashed line - S.B.S. model.

The solid line is the average axial Patterson calculated from the data. The dotted line is that of the D.H. model; and the dashed line relates to the S.B.S. model. The scales of the axial Pattersons for both models have been adjusted to provide the best fit to the data curve. The curve for the S.B.S. model closely follows that of the data - the relative amplitudes of the peaks are almost the same as those of the data. The noticeable difference in the D.H. model curve, compared with the data, is the low amplitude peak at 6.8 Å. Even allowing for the variability in the data (see Fig. 2.4), this portion of the curve appears to be outside the limits.

The amplitudes of the peaks in P_{ax} are mainly due to the correlation between base pairs. Thus a high amplitude peak at 6.8 Å indicates that some of the base pairs separated by 6.8 Å are aligned at almost the same angle ($\pm 180°$). Consider the D.H. model. Adjacent base

pairs are separated by 36° (the diffraction data requires ten base pairs per repeat unit). The angular separation between base pairs displaced by 6.8 Å in the z direction is, therefore, 72°. The low peak at 6.8 Å (on the P_{ax} for the D.H. model) shows that the correlation at this angular separation is low. As the angular separation increases, the correlation reaches a maximum at 144° (z = 13.6 A). [The maximum is at 144°, rather than 180°, because the base pair, including the sugar group at each end, is actually crescent shaped (Arnott and Hukins, 1973).] The data require that models for D.N.A. produce more or less equal amplitude peaks in the axial Patterson at 6.8 Å, 10.2 Å and 13.6 Å. The easiest way of achieving this is by having alternating left and right handed helical sections in the model. With the S.B.S. model it is relatively simple to align the orientation angles of at least two base pairs at a given separation distance. However, with the D.H. model only one high amplitude peaks is possible (apart from at z = 0), because the angular separation between adjacent base pairs is fixed at 36°.

2.4 DIFFRACTION PATTERN ANALYSIS

The diffracted X-ray intensity, from specimens of fibrous B-DNA, lies along "layer lines". The most prominent layer lines are the zeroth through to third, and the tenth (see Fig. 2.3). In the graphs presented in this chapter showing the diffracted intensity along the layer lines, the different sets of diffraction data, and also the theoretical diffraction calculated from the models, are all normalised to have roughly equal amplitude at the highest intensity point on the second layer line.

The complex amplitude of the diffraction pattern of a D.N.A. molecule, in the ℓth layer plane, is

$$E_{\ell}(R,\psi) = \sum_{j=1}^{410} H_j [(R^2 + \ell^2/c^2)^{\frac{1}{2}}] \exp(j2\pi[Rr_j \cos(\theta_j - \psi) + \ell z_j/c]) \quad , \quad (2.28)$$

where (2.15) and (2.22) have been used. H_j is the scattering pattern due to the jth atom with electron density h_j (H_j is the Fourier transform of h_j). Again the effects of the hydrogen atoms are ignored. From (2.16) the fibre pattern along the ℓth layer line is found, after some manipulation, to be

$$\Omega_\ell(R) = \sum_{n=1}^{410} \sum_{m=n}^{410} \varepsilon_{n-m} H_n[(R^2 + \ell^2/c^2)^{\frac{1}{2}}] H_m[(R^2 + \ell^2/c^2)^{\frac{1}{2}}]$$

$$\times J_0[2\pi R(r_n^2 + r_m^2 - 2r_n r_m \cos(\theta_n - \theta_m))^{\frac{1}{2}}] \cos[2\pi\ell(z_n - z_m)/c] \quad , \quad (2.29)$$

where $J_0(\cdot)$ denotes the zeroth order Bessel function; and $\varepsilon_m = 1$ for $m = 0$, and $\varepsilon_m = 2$ for $m \neq 0$.

Fig. 2.6 shows the amplitude along the zeroth through to tenth layer lines of the fibre diffraction pattern of B-DNA measured by Feughelman et al. (1955) (solid line), Bram (1977) (crosses), Zimmerman and Pheiffer (1979) (circles); compared with those computed for the double helix model of Arnott and Hukins (1973) (dotted line), and the side-by-side model of Bates et al. (1980) (dashed line). The Ω_ℓ's for the two models were computed using (2.29). The most noticeable thing is the variation in the data. The discrepancy between both models and the data is appreciable on the eighth and ninth layer lines. [However, over this region of the actual diffraction patterns it is difficult to discern sharp layer lines.] On the zeroth, and the second through to seventh layer lines the differences between the data and models are of the same order. On the first layer line the D.H. model appears to be somewhat low in intensity while the S.B.S. model is tending towards the other extreme. The D.H. model also exhibits a discrepancy on the tenth layer line.

2.5 DIFFRACTION DIFFERENCES DUE TO BASE PAIRING

The calculated fibre diffraction patterns presented in §2.4 were produced assuming that all the base pairs were of the A-T variety. The alternative would have been to calculate diffraction patterns for various A-T G-C base pairing sequences and average the results. However this is unnecessary as is now shown. The basic "building block" of D.N.A., known as a dinucleotide, consists of a base pair with a sugar and phosphate group at each end. The manner in which these dinucleotides are stacked on top of each other determines the structure of the model. For instance, if the angular separation between adjacent dinucleotides is 36° the D.H. model is formed; and if the angular separation is firstly plus about 36° for five adjacent dinucleotides, and then minus about 36° for another five dinucleotides, a "rough S.B.S." model is formed.

$\Omega_l(R)$
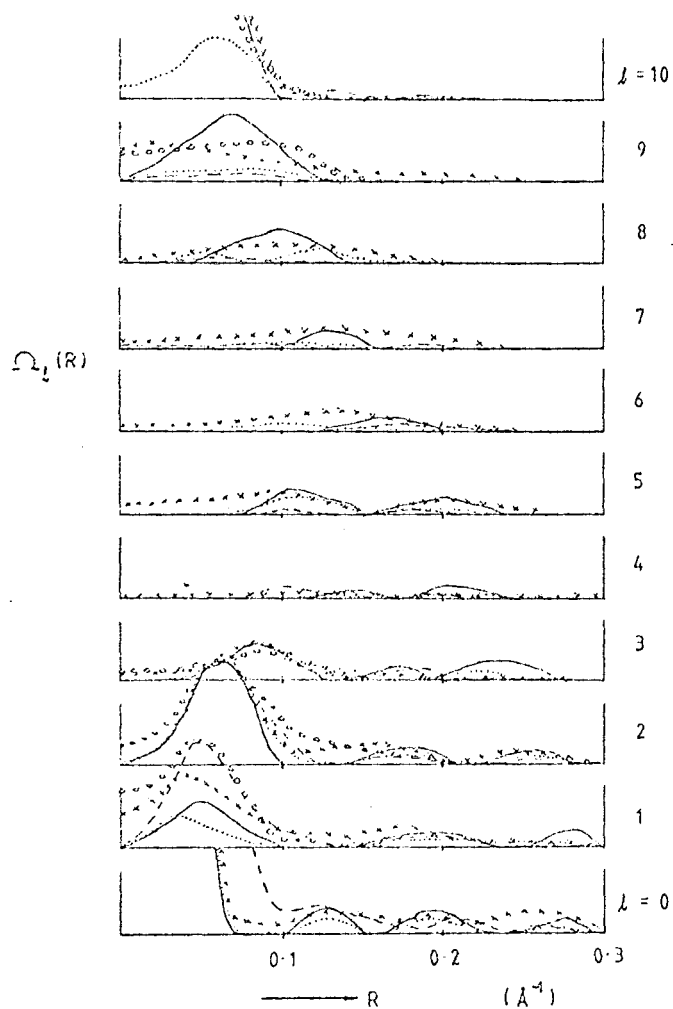

Fig. 2.6   The actual and theoretical diffraction intensity
            along the zeroth to tenth layer lines for fibrous
            B-DNA.

—————— Feughelman et al. (1955)

x x x   Bram (1971)

o o o   Zimmerman and Pheiffer (1979)

······   Double Helix Model (Arnott and Hukins, 1973)

— — — —   Side-by-Side Model (Bates et al., 1980)
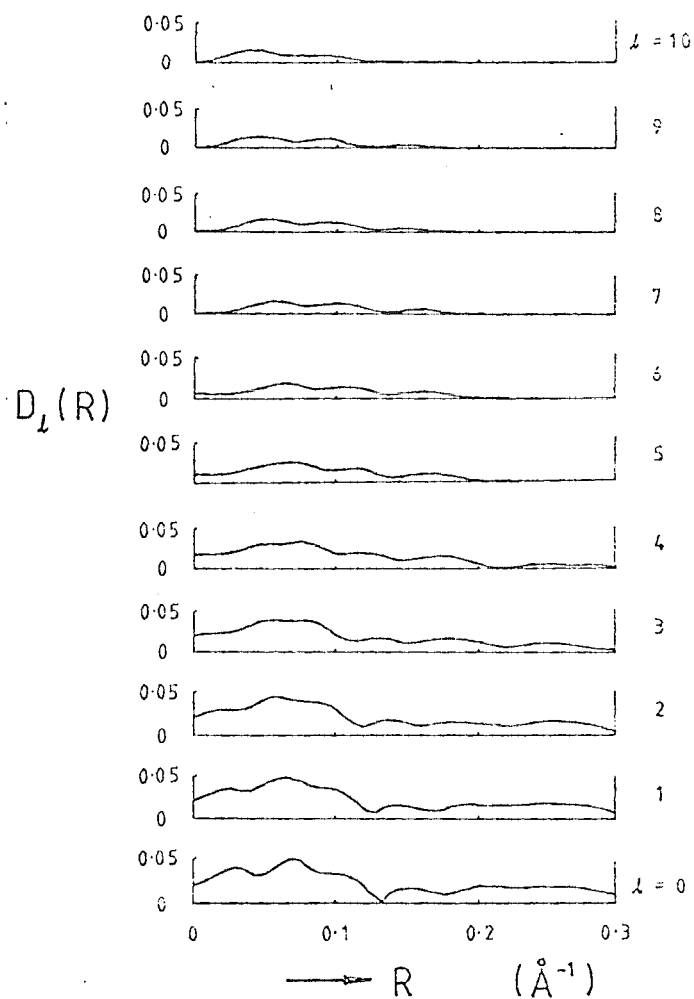
Fig. 2.7   Averaged normalised differences between diffracted
            intensities for a dinucleotide incorporating an
            A-T base pair and a dinucleotide incorporating a
            G-C base pair.

Now consider the calculated diffracted intensity for a single dinucleotide incorporating an A-T base pair, and from a single dinucleotide incorporating a G-C base pair. In the ℓth layer plane the diffracted intensity is denoted by $I_{A,\ell}(R,\psi)$ for the A-T base pair, and $I_{G,\ell}(R,\psi)$ for the G-C base pair. The layer planes are those that apply for B-DNA.

If the diffracted intensity for the two types of dinucleotide are almost the same in each layer plane, when averaged over three widely separated angles, then it seems that one is justified in assuming that the fibre patterns for B-DNA are virtually independent of the base pairing sequence. Define

$$\tilde{I}_{X,\ell}(R) = I_{X,\ell}(R,0) + I_{X,\ell}(R,\frac{\pi}{4}) + I_{X,\ell}(R,\frac{\pi}{2}) \quad , \tag{2.30}$$

where X stands for either A or G. Also define

$$\hat{I} = \tilde{I}_{A,2}(R_0) \quad , \tag{2.31}$$

where R_0 is the R value on the second layer line at which Feughelman et al. (1955) record the peak intensity. Finally the normalised intensity difference is defined by

$$D_\ell(R) = |\tilde{I}_{A,\ell}(R) - \tilde{I}_{G,\ell}(R)| / \hat{I} \quad . \tag{2.32}$$

$D_\ell(R)$ is shown in Fig. 2.7 for the zeroth to tenth layer lines.

The conclusion to be drawn from Fig. 2.7 is that the maximum effect due to different base pairing sequences on the diffraction data is about 5% of the peak intensity value on the second layer line. Such a change is insignificant compared with the variation in the available data. Therefore it is not likely to be a significant factor when it comes to deciding the conformation of D.N.A.

2.6   INTERPRETING THE AXIAL PATTERSON IN TERMS OF THE DIFFRACTION DATA

It seems strange that the widely varying sets of diffraction data should produce very similar axial Pattersons; and that the theoretical diffraction pattern for the D.H. model seems to be as good as that for the S.B.S. model yet the D.H. model exhibits a discrepancy with respect

to the axial Patterson. There are three questions here. Firstly, why are the data so varied? Secondly, why should varying sets of diffraction data produce similar axial Pattersons. And thirdly, what features of the diffraction pattern are responsible for the two almost equal amplitude peaks at 6.8 Å and 13.6 Å in the axial Patterson?

I am not really in a very good position to answer the first point, but my feelings are that each D.N.A. specimen used is slightly different in structure yet the overall conformation (be it S.B.S. or D.H.) is the same. This is probably due to the difficulty in preparing specimens of D.N.A. for analysis. The reason for the above comment is that Bram (1973) reports the results of an experiment to determine if the diffraction pattern is dependent on the ratio of A-T to G-C base pairs within the DNA specimens. The diffraction patterns obtained from supposedly identically prepared specimens of D.N.A. are in fact different. Although Bram attributes this to the A-T to G-C ratio, it can be shown (see §2.5, and Bates et al., 1980) that the effect of the different base pairs should be negligible (less than 5% of the maximum intensity on the 2nd layer line, whereas Bram's (1973) results show relative intensity varying by as much as a factor of 2). It seems that this variation is therefore likely to be due to the preparation of the specimens. Either that or the A-T:G-C ratio is changing the overall structure slightly. Although, considering the similarity of the base pairs, it is difficult to see why this latter possibility should be so.

This would also explain the second point. The axial Patterson tends to "average out" much of the detail in the diffraction data. Thus provided that the gross structural difference between molecules from different specimens is slight (even though there may be differences in detail) the axial Pattersons will be quite similar.

The answer to the third question is slightly more quantitative. The difference between the axial Patterson peaks at $z = 6.8$ A and $z = 13.6$ A is proportional to the quantity D defined by

$$D = (I_2 + I_3 + I_7 + I_8) - (I_1 + I_4 + I_6 + I_9) \quad . \qquad (2.33)$$

This is obtained using (2.21), where

$$I_n = \int_0^\infty \Omega_n(R) \, R \, dR \quad . \qquad (2.34)$$

Table 2.1 lists the I_n's for the sets of data and the models, plus their D values.

Table 2.1   The I_n and D values for the three
sets of data and the two models.

	I_n										D
Feughelman et al. (1955)	1.48	1.67	1.37	0.64	1.04	0.37	0.28	0.68	0.96	1.51	0.55
Bram (1979)	0.78	1.09	0.98	0.65	1.24	0.89	0.88	1.12	1.28	1.10	0.47
Zimmerman and Pheiffer (1979)	1.61	2.06	1.80	0.39	–	–	–	–	1.48	2.65	0.38
S.B.S.	2.23	2.00	1.00	0.75	0.92	0.62	0.48	0.46	0.31	1.23	0.03
Double Helix	1.09	2.94	1.72	0.55	0.80	0.59	0.27	0.67	0.37	1.05	3.00
n =	1	2	3	4	5	6	7	8	9	10	

There is very little correlation between the I_n's of the data and either model. However the relatively low I_1 value for the double helix model must be responsible for the low peak at 6.8 Å in the axial Patterson, as there is only a slight difference between the other I_n values for each model. Thus the ratio of I_1 to I_2 and I_3 appears to be the crucial factor in determining the relative amplitudes of the peaks at 6.8 and 13.6 Å in the axial Patterson.

2.7   DISCUSSION

There seems to be a *prima facie* case for a side-by-side model for D.N.A. However a definitive decision on the strength of the results reported in this chapter can not be made. The axial Patterson result, §2.3, is in favour of an S.B.S. model, but this is largely due to the diffracted intensity on the first layer line, §2.6. There is considerable variation in the data, and one can not be sure whether this variation is due to the differences between specimens of D.N.A., or errors in the data. The discrepancy exhibited by the D.H. model on the 10th layer line (see Fig. 2.6) is not significant in terms of choosing between the two models. The low intensity of $\Omega_{10}(R)$, for the D.H. model,

at small values of R, can be shown to be due to destructive interference between X-rays diffracted from each end of the dinucleotides. It could easily be overcome - for instance by adjusting the tilt of the dinucleotides.

There are two main points arising from this work. Firstly, in terms of the diffraction data, the S.B.S. model is certainly a viable alternative model for the D.N.A. molecule. And secondly, the present D.H. model for D.N.A. (Arnott and Hukins, 1973) requires further refinement. Considering that the D.H. model has been around for over 25 years, the agreement with some aspects of the diffraction data is less than satisfactory - even allowing for variability in the data.

3.   IMAGING THE BEATING HEART USING X-RAY COMPUTED TOMOGRAPHY

3.1   INTRODUCTION

The most impressive medical imaging devices are undoubtedly the X-ray computed tomography (C.T.) brain and body scanners. With these machines high quality cross sectional images can be obtained. Nevertheless present-day body scanners have some practical limitations. One of these is that they can not straightforwardly be used to produce images of moving organs, such as the heart.

Standard body scanners contain an X-ray source which is able to travel around the patient being imaged. A scan consists of taking X-ray views (known as "projections") through the patient at many different angular positions. [Projections are generally measured over 360° of X-ray source travel, however some scanners use a smaller angular range than this.] Now the minimum scan time is limited mainly by the technical difficulties associated with rapidly revolving an X-ray source about the patient. Typical minimum scan times are as small as one second. However, during even such a short time, the heart moves enough to blur the resulting image. It appears that scan times as low as 50 ms are required to capture "stationary" images of the heart. The fragile nature of X-ray tubes, and the problems associated with their power supply and cooling requirements, mean that it is unlikely that these 50 ms scan times will be attained with the conventional style of scanner. Research groups in various laboratories are investigating other methods of achieving very fast scans. At Mayo Clinic, Robb et al. (1979) are producing a scanner consisting of 28 X-ray sources, which enables multiple projections to be obtained simultaneously. Another approach, reported by Haimson (1979), envisages an X-ray tube large enough for its anode actually to surround the patient. By electronically scanning the electron beam around the anode, it is hoped that the resulting X-ray beam can be made to revolve around the patient - thus eliminating mechanical rotation.

In this chapter an alternative approach to imaging the heart is presented. Firstly, in §3.2, necessary preliminaries relating to X-ray computed tomography are introduced. Then, in §3.3, the characteristics

of heart motion are examined. The heart movement is reasonably periodic, therefore stroboscopic procedures can be employed. The timing signal for the "strobe" can be obtained from the electrocardiograph (E.C.G.) which, being an electrical signal produced by the muscular contractions within the heart, is closely related to heart movement. By measuring the E.C.G. throughout the scan, the X-ray projections which occur at the same phase of each heart cycle can be identified. Using only these projections, an image of the heart at that particular phase of its pumping cycle can be obtained. This is known as E.C.G. "gating"; as a small time "gate" at a particular phase of the E.C.G. is used to select the projections. The success of E.C.G. gating methods depends largely on the periodicity of the cardiac motion, and the stationarity of the rest of the material in the body. These two requirements typically hold true over only a small number of cardiac cycles. The number mainly depends on the period of time over which patients can comfortably hold their breath. Generally this is assumed to be not much more than about 12 heart beats.

Here two E.C.G. gating schemes for imaging the beating heart are presented. The first, §3.4, outlines a design philosophy for a new type of scanner – a cardiac scanner. The procedure involves taking a consecutive series of scans through the same "slice" of the patient (i.e. the X-ray source rotates through several revolutions). The scanner's rotation speed is synchronised in a particular manner with the E.C.G. Using such a device it should be possible to produce images of the heart, at each phase of its cycle, comparable in quality to those from conventional scans of stationary objects.

The second scheme, §3.5, is one that could be used with existing scanners. In conventional scanners the cables and pipes, supplying power and cooling oil to the X-ray tube, are simply "coiled up" as the tube rotates about the patient. This limits the number of revolutions the X-ray source can undergo in any direction. Generally only about 1½ revolutions are permitted before the system has to be "unwound". When only a single scan (i.e. 1 revolution) is possible, the number of projections available to reconstruct an image of the heart at each instant of its cycle is effectively equal to the number of heart beats per scan. [The "effectively" is explained in §3.5.] As noted above, this number is no more than about 12; which is small compared with the 720 or so projections normally used to reconstruct images of stationary

objects. Images of the heart reconstructed from only a few projections are low in quality. However in §3.5, a procedure for improving the image quality is presented. Basically the procedure involves recognising that the region external to the heart is stationary, and so can be estimated quite accurately using the full 720 projections. Then, using this estimate of the "stationary region", it is shown how the original projections can be modified to produce a more accurate image of the heart. Computer simulations demonstrating this procedure are presented in §3.6.

In §3.7 the significance of the ideas presented in this chapter are discussed, and also compared with previously reported E.C.G. gated C.T. schemes for heart imaging.

3.2 FUNDAMENTALS OF X-RAY COMPUTED TOMOGRAPHY

In X-ray computed tomography, the X-ray energy is treated as if it travelled in thin beams (or rays) which are subject only to attenuation effects. This apparent attenuation is mainly due to X-ray energy being scattered out of the beam. Suppose that the intensity of the X-ray beam is represented by I. Then it is found that the beam strength is attenuated according to the formula

$$\partial I/\partial \tau = -\lambda(\tau) \ I \quad ; \tag{3.1}$$

where τ is distance along the beam from the source position, and $\lambda(\tau)$ is the effective absorbtivity of the material in the path of the beam (see for example Stanton, 1969). In general λ is also dependent on the energy, or frequency, of the X-rays. For a wide bandwidth X-ray source a phenomena known as "beam hardening" occurs. What happens is that the low energy X-ray photons are absorbed at a greater rate than the higher energy photons. Therefore the ratio of high energy ("hard") photons to low energy ("soft") photons increases as the beam propagates. In the simplified analysis presented in this chapter it is assumed that the X-ray beam is mono-energetic. However in practical C.T. a correction for beam hardening is usually made (see for example Herman, 1979). Now consider an X-ray beam through an object (see Fig. 3.1). Suppose that the intensity of the beam at the source is I_o. Then, solving (3.1), the intensity at the detector, on the other side of the object,

$$I_d = I_o \exp\left\{-\int \lambda \, d\tau\right\}$$
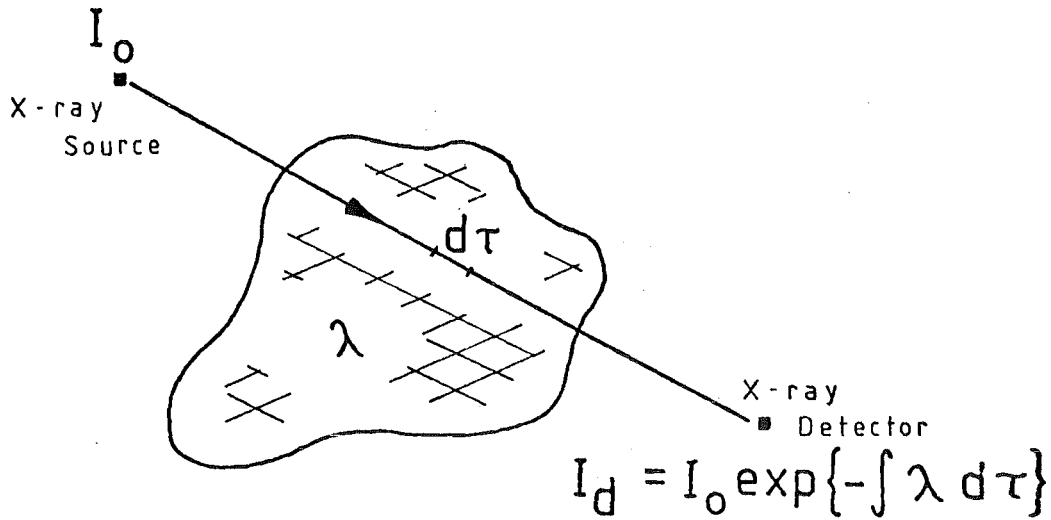
Fig. 3.1   In C.T., the propagation of X-ray energy is treated
in terms of rays subject to an apparent absorbtion.

is given by

$$I_d = I_o \exp\left(-\int_{\tau_o}^{\tau_d} \lambda(\tau) \, d\tau\right) \quad . \tag{3.2}$$

Provided I_o is known, the term $\int \lambda \, d\tau$ for a particular ray can be extracted from the measured intensity. The integral term is the projected X-ray absorbtivity along a ray. In most cases the rays all lie within a plane or "slice" of the object being imaged. The projected X-ray absorbtivity of a group of rays with some sort of geometrical relationship is known as a "projection". For example, rays orientated at the same angle and lying in the same plane produce a "parallel beam projection" (see Fig. 3.2). And rays emanating from the same source position and also lying in a plane produce a "fan beam projection" (see Fig. 3.3). In the following analysis the term "projection" refers to a parallel beam projection. There is no loss of generality in treating only the parallel beam case, as a complete set of fan beam projections can always be rearranged into parallel beam form.
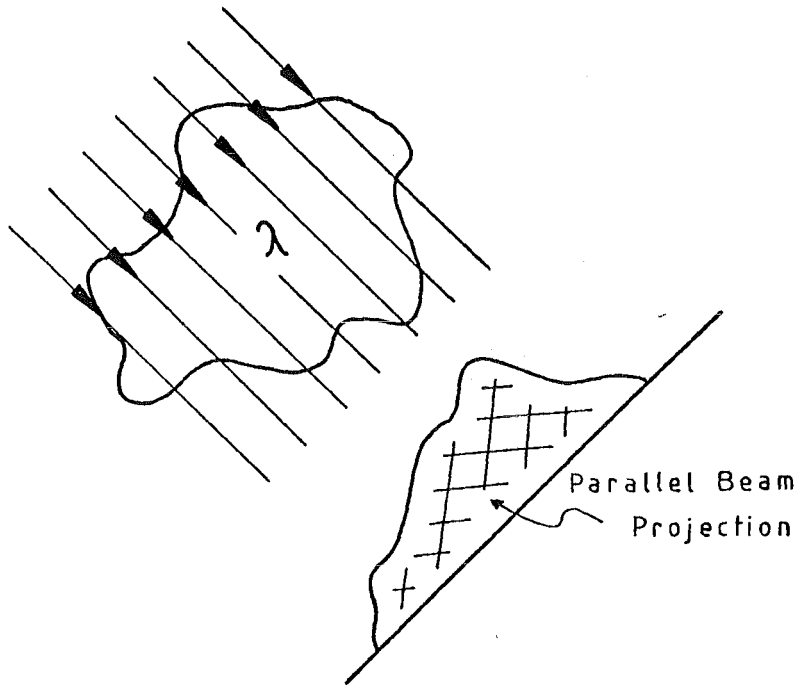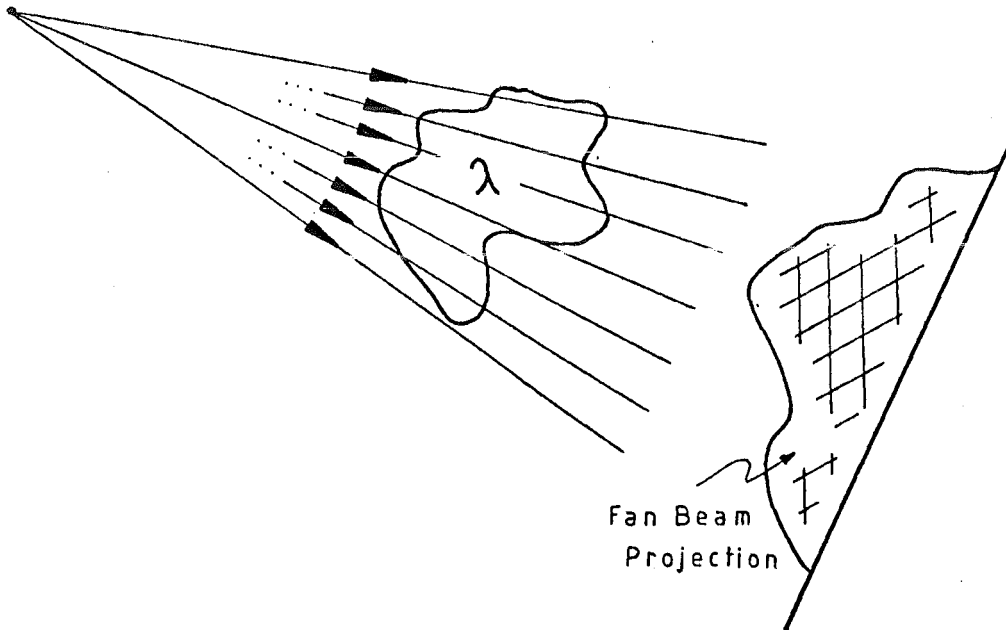
Fig. 3.2　Parallel beam projection.
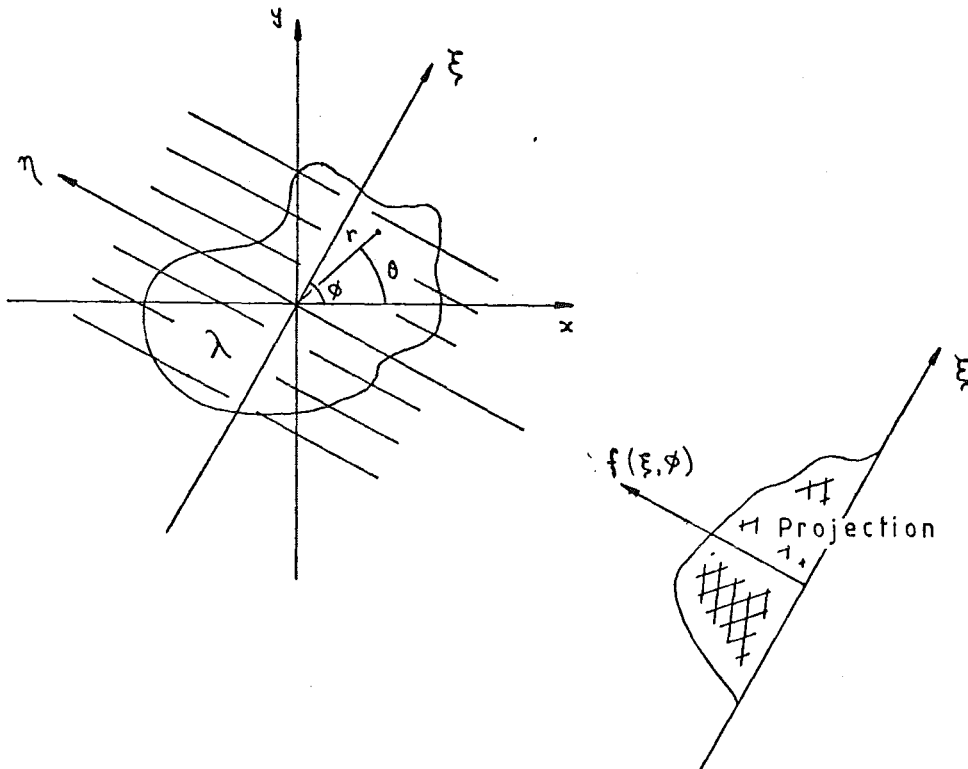

Fig. 3.3　Fan beam projection.

Fig. 3.4   The projection geometry. The (x,y) coordinates are fixed,
and the (ξ,η) coordinates rotate with projection angle ϕ .

Define an arbitrary datum within the X-ray absorbtivity
distribution λ (see Fig. 3.4). Also define three sets of coordinates -
stationary cartesian (x,y); stationary polar (r,θ); and revolving
cartesian (ξ,η) orientated at angle ϕ . The projection $f(\xi,\phi)$ is
obtained by integrating λ along lines parallel to the η axis. These
lines are specified by the (ξ,ϕ) coordinates. Notice that $f(\xi,\phi)$
$= f(-\xi,\phi+\pi)$. In §1.8 (chapter 1) it is shown how a two dimensional
function can be estimated from its projections. A line, passing through
the origin and orientated at angle ϕ say, in the two-dimension Fourier
transform of a function is equivalent to the one dimensional Fourier
transform of a projection, at angle ϕ, through that function. Therefore
the X-ray absorbtivity λ can be obtained from the projection $f(\xi,\phi)$ by
simply Fourier transforming the projections to obtain Λ - the two
dimensional Fourier transform of λ, and then inverse Fourier transforming
Λ . These operations can be written as

$$\lambda(r,\theta) = \int\limits_{-\infty}^{\infty} \int\limits_{0}^{\pi} \{ \int\limits_{-\infty}^{\infty} f(\xi,\phi)\exp(j2\pi\xi\rho)\,d\xi\} \ \exp[-j2\pi r\rho\cos(\theta-\phi)]\,|\rho| \ d\phi d\rho \quad . \quad (3.3)$$

Here (r,θ) are conventional cylindrical polar coordinates, but (ρ,ϕ) are defined to exist in the ranges $-\infty < \rho < \infty$ and $0 < \phi < \pi$. Rather than obtain λ by directly evaluating (3.3), it is usually calculated in a numerically more efficient manner using the "modified backprojection algorithm" (see for example Lewitt *et al.*, 1978). Define the "modified projection" $\tilde{f}(\xi,\phi)$ by

$$\int\limits_{-\infty}^{\infty} \tilde{f}(\xi,\phi) \ \exp(j2\pi\xi\rho) \ d\xi = |\rho| \int\limits_{-\infty}^{\infty} f(\xi,\phi) \ \exp(j2\pi\xi\rho) \ d\xi \quad . \quad (3.4)$$

Also define the one dimensional Fourier transforms of $\tilde{f}$ and f by $\tilde{F}(\rho,\phi)$ and $F(\rho,\phi)$ respectively. Then (3.4) is equivalent to

$$\tilde{F}(\rho,\phi) = |\rho| \ F(\rho,\phi) \quad . \quad (3.5)$$

Therefore $\tilde{f}$ is obtained by modifying the spatial frequency components of f by a weighting function which is proportional to frequency. [In practice, because of both the finite number of samples and the presence of noise in the measured $f(\xi,\phi)$, the weighting function $|\rho|$ is replaced by a truncated version. That is, the weighting function used equals $|\rho|$ up to some frequency component, ρ_c say, and thereafter is set equal to zero. In effect a "low pass filtered" $|\rho|$ weighting function is used. Also, as $f(\xi,\phi)$ is known at only a finite number of sample points, it is impossible (in general) to estimate $F(\rho,\phi)$ exactly. The errors in $F(\rho,\phi)$ are known as aliasing errors (see for example Brigham, 1974). To reduce the effects of these errors in the resulting image it is usual to multiply $F(\rho,\phi)$ by a "window function". Window functions are generally bell or gaussian shaped functions. That is, they are approximately equal to unity for small values of ρ, and tend smoothly to zero as ρ tends to ρ_c (in this case) (see Harris, 1978). Multiplying $F(\rho,\phi)$ by both the low pass filtered $|\rho|$ weighting function and the bell shaped window function is generally carried out in one operation.]

Now (3.3) can be rearranged, using (3.4), to give

$$\lambda(r,\theta) = \int\limits_{0}^{\pi} \int\limits_{-\infty}^{\infty} \tilde{f}(\xi,\phi) \ \{ \int\limits_{-\infty}^{\infty} \exp[j2\pi\rho(\xi - r\cos(\theta-\phi))]\,d\rho\} \ d\xi d\phi \quad . \quad (3.6)$$

The result of the ρ integration is the delta function $\delta\left(\xi - r\cos(\theta-\phi)\right)$. The ξ integration is then easily carried out, giving the equation

$$\lambda(r,\theta) = \int_0^\pi \tilde{f}\left(r\cos(\theta-\phi),\ \phi\right)\ d\phi \quad . \tag{3.7}$$

The term $r\cos(\theta-\phi)$ is, for a given projection angle ϕ, equivalent to the ξ coordinate (Fig. 3.4). Thus to obtain $\lambda(r,\theta)$ all the modified projections are spread back along ray paths, corresponding to a particular projection angle; and the value of the modified projection associated with each particular ray is added to those image points intersected by that ray. This part of the operation is known as back-projecting the modified projections. Hence the name "modified back projection" method.

3.3 CHARACTERISTICS OF CARDIAC MOTION

E.C.G. gating techniques for imaging the heart require that the motion be periodic or stationary. Any departure from these two states of motion reduces the spatial resolution. Here the expected movement and velocity of the heart muscle are discussed with respect to E.C.G. gated computed tomography.

The main cause of aperiodicity of the heart motion is respiration. This has been examined using cineangiograms (X-ray movies of the heart) (Bogren et al., 1977). As the heart sits on the diaphragm it is directly influenced by diaphragmatic movement. During normal breathing this movement can be as large as 2 cm, which produces severe blurring of the C.T. image. Bogren et al. (1977) also examine the effect of breath holding. With no diaphragmatic motion the heart movement was found to be periodic to at least within 1 mm. (The number of heart beats over which this periodicity was observed was not specified explicitly, but it would appear to be between four and six.) Abnormal beats destroy this periodicity, but they can usually be detected on the E.C.G. A sequence of regular heart beats is required for an E.C.G. gated scan.

Gating the respiratory as well as the cardiac cycle, to enable projection data to be collected over a large number of heart beats, has been suggested (Sagel et al. 1977). However, it is not known how

periodic the respiratory cycle is. In fact one study (Bogren *et al.*, 1977) has shown that it is difficult to repeatedly achieve the same level of inspiration - although the discrepancy with respect to diaphragmatic movement was typically only 5 mm. It seems that to obtain high resolution images, the total scan should be completed during the interval throughout which patients can comfortably hold their breath. In addition the body, excluding the heart, must remain still. It is doubtful that these requirements can generally be achieved for longer than 20 s. A more realistic figure might be 12 s.

The periodic motion of the heart muscle is now discussed. The valves, being so fine and moving quite rapidly, would be difficult to detect using X-ray C.T., except possibly when they are stationary. Movement of the left ventricular walls has been examined using cineangiograms (Gibson *et al.*, 1976). [The left ventricule is the main pumping chamber of the heart.] The maximum displacement of the wall is about 18 mm on a normal subject, with a maximum velocity of approximately 13 cm/s. These results have to be taken with a grain of salt, but they do allow some rough conjectures relating to scan timing and synchronisation to be made. The heart motion is composed of "relative calm" interspersed with "quite violent" movement. To demonstrate this, consider the heart cycle to be divided into 20 equal intervals of approximately 50 ms each. Within two of these intervals the maximum movement of the ventricular wall is about 6 mm. During two other intervals the ventricule is virtually stationary. And the maximum movement during each of the remaining sixteen intervals is about 2 mm.

The above characteristics of the heart movement are relevant to the multiple scanning E.C.G. gating scheme discussed in §3.4. There the projection data allocated to reconstruct a single image are collected over a segment of each E.C.G. cycle rather than at an instant. The duration of the segment, or rather the movement per segment, effects the resolution. It is difficult to explicitly relate resolution to the motion of the object being scanned, but the apparent blurring can be expected to be equal to the extent of the movement, as shown by Bates and Heffernan (1980). Thus the consequences of object motion during C.T. imaging are, to a reasonable approximation, similar to those in conventional photographic imaging. Suppose that the heart is assumed to be "quasi-stationary" during each of the 50 ms intervals. Then if an image is reconstructed using projections measured during a particular

interval, over a number of cycles, the spatial resolution can be
expected to be less than 6 mm and typically of the order of 2.mm.
These estimates do not take account of the spatial resolution limit
inherent in the scanner itself, which is generally of the order of
2 to 3 mm. The scanner resolution is effected by the apertures on the
X-ray source and detectors - in other words, the width of the X-ray
beam. By using "deconvolution" techniques though, it is possible to
achieve a resolution finer than the width of the X-ray beam (see
Bracewell, 1971, for a discussion on this).

3.4 CARDIAC C.T. SCANNER

In this section a scanning strategy is presented for cardiac
imaging using an X-ray body scanner. The method has been obliquely
suggested in a couple of papers (Harell et al., 1977; Berninger et al.,
1979). It requires a scanner with an X-ray source which is able to
travel through several revolutions in the same direction. Here it is
shown how the rotational speed can be optimised so that none of the
measured projections is redundant; thereby minimising the X-ray dosage
to the patient.

The heart cycle is divided into a number of equal time intervals.
The duration of each interval is chosen to be small enough that the
heart can be considered quasi-stationary within it (i.e. the amount of
heart movement per interval is less than the required resolution).
These quasi-stationary phases of the heart are identified with
particular segments of the E.C.G. (see Fig. 3.5). Each segment occurs
once each cycle. Now if X-ray projections and the E.C.G. are recorded
simultaneously as the scanner rotates, each projection can be identified
with a particular phase of the heart cycle via the segments of the E.C.G.
An image of the heart, for a particular phase, is obtained using only
those projections which are measured during the relevant segment. In
general, to obtain a reasonable image it is necessary that: (i) a
large number of projections are recorded; and (ii) the projections are
equally spaced in angle. To obtain a large number of projections,
equally spaced in angle, for each phase of the heart cycle, several
scans have to be performed. [Here the term "scan" implies one complete
revolution of the X-ray source about the patient.] If there is no
control over the projection angle, for a given heart phase, several
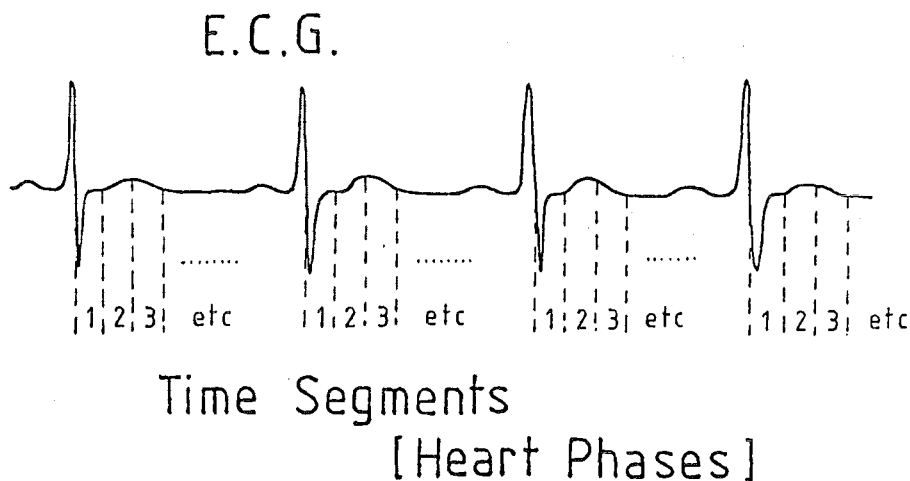
E.C.G.

Time Segments
[Heart Phases]

Fig. 3.5   The E.C.G. is divided into a number of time segments.

projections may be measured at certain angles and none at other angles.
Thus it may be necessary to take a "few more" scans to achieve an
evenly spaced set of projections.   This is basically the method used by
both Harell *et al.* (1977), and Berninger *et al.* (1979).   [However
instead of scanning through several revolutions in the same direction,
they had to perform one scan, reverse the X-ray source back to the
starting position, then scan again, followed by another reverse etc.]
This is rather a hit and miss affair, and a certain number of unnecessary
projections are likely to be measured.

By synchronising the projection angle with the E.C.G., the number
of scans required can be minimised.   Because a projection at angle ϕ
is equivalent to one at $\phi+\pi$, see §3.2, it is sufficient to measure
projections over $180°$ for each heart phase.   Suppose that the heart
cycle can be treated as having N phases.   Then at least N/2 complete
scans are required to collect all the projections.   Here a series of
consecutive scans is called a "multi-scan".   Further, remember that
to minimise the effects of patient movement and aperiodicities in the
cardiac cycle the multi-scan should be performed as quickly as possible.

An interesting feature of the projection collection strategy presented in this section is that, provided heart motion is periodic, it works largely independently of the number of E.C.G. cycles per multi-scan. Suppose that the heart can be treated as being periodic for M complete cycles. If the speed of rotation (ω) of the X-ray source is chosen to be π/M radians per segment, then, provided (N-M) is odd (unless M=1), none of the measured projections is redundant. To see how this works, consider the case where N = 4 and M = 3. Two complete rotations are required and ω = $\pi/3$ radians per segment, see Fig. 3.6.
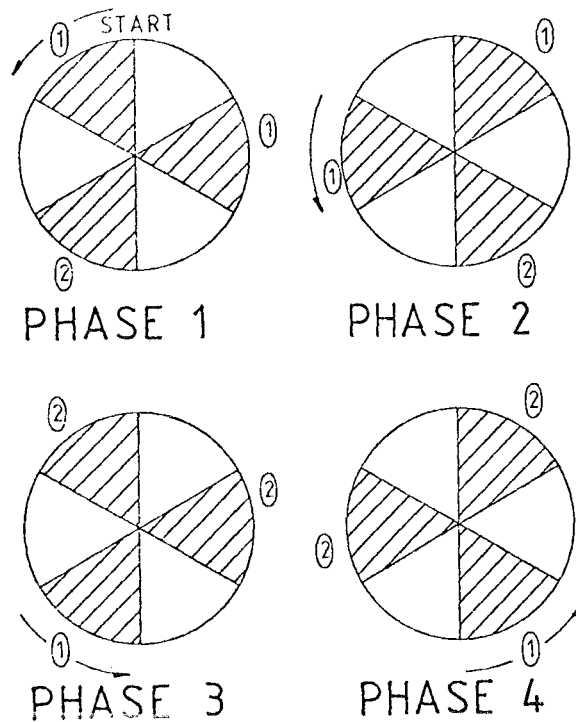

Fig. 3.6   Diagram showing that by synchronising the rotation
of the X-ray source, projections can be measured
over 180° for each heart phase without redundancy.

Each circle corresponds to a particular heart phase. The sectors correspond to the time segments, and the angles covered by the shaded sectors correspond to the angles at which projections are measured for that particular segment. The rotation is anti-clockwise, and the circled numbers indicate whether the projections are measured during the first or second revolution. As projections measured at angle ϕ are the same as those measured at $\phi + \pi$, Fig. 3.6 shows that a complete

set of projections for each phase is obtained. Thus, by synchronising the rotation with the E.C.G. in the prescribed manner, the scanning procedure is optimised. [If N-M is even the shaded sectors shown in Fig. 3.6 overlap each other rather than interleave.]

In §3.3 it is argued that to obtain reasonable resolution N has to be equal to about 20 (i.e. the time per segment is about 50 ms). Thus the scanner has to be capable of rotating continuously in the same direction for 10 revolutions. It is also argued in §3.3 that the maximum time for which the heart motion can be considered sufficiently periodic, is in the order of 12 s . Hence the scanner must be able to rotate at least once each second.

3.5   CARDIAC IMAGING USING FEW PROJECTIONS

Existing C.T. body scanners are not designed to perform multiple revolutions in the same direction. In this section a method for imaging the heart using conventional scanners is described. Again an E.C.G. gating scheme is used. However, when only one scan is possible (within the 12 s or so that the patient can remain still), relatively few projections are obtained for each phase of the heart cycle. It follows that image reconstruction errors are considerable. By recognising that the region surrounding the heart is more or less stationary, means can be devised for significantly reducing these errors.

To obtain a good image one requires a large number of consistent projections equally spaced in angle. A set of projections is consistent if all the projections are of the same X-ray absorbtivity distribution. In this case the projections all have to be measured during the same phase of the heart cycle. Now with an E.C.G. gated scan, the total number of projections available to reconstruct an image, at a particular phase, is equal to the number of E.C.G. cycles per scan. [In actual fact, due to the quasi-stationary period noted in §3.4, groups of projections are measured, and it is actually the number of groups that is equal to the number of E.C.G. cycles per scan. However, when only a single scan is possible, the reconstruction errors are still significant because of the large angular spacing between consistent groups of projections. The image reconstruction methods discussed here are all based on the very efficient modified backprojection algorithm (see §3.2). Therefore having several projections per group would improve the

signal to noise ratio. But, for simplicity, in the following analysis
it is assumed that only one projection per phase per E.C.G. cycle is
measured.] Provided that arguments developed in §3.3 are valid, the
maximum interval during which "consistent projections" can confidently
be obtained is about 12 s. Therefore, assuming a heart rate of 1 beat
per second, the number of projections measured per phase is 12. Thus
one has the problem of reconstructing an image of a particular heart
phase using about 12 projections. (Normal high resolution C.T. images
are reconstructed from as many as 720 different projections.)

To demonstrate the ideas presented in this section, a mathematical
"chest phantom", which includes a "moving heart", is used to generate
simulated X-ray projections. The model is more fully described in the
next section (§3.6). Here it is sufficient to be aware that the X-ray
absorbtivity values used in the model are typical of those found in
lung, muscle, spine, etc., of the body (Huang and Wu, 1979) - however
the X-ray absorbtivity of the "blood" within the heart has been modelled
as if a reasonable amount of contrast agent is present (about 5 ml/litre
of blood of sodium meglumine diatrizoate; see Kivisaari *et al.*, 1979).

Now two computational approaches are introduced:

> A: the straightforward approach
>
> B: the subtraction approach.

The straightforward approach, A, involves taking the small number
of projections, corresponding to a particular phase of the heart cycle,
and then reconstructing a complete section through the thorax.

At this point in the argument it is useful to show just how much
degradation of the reconstructed image there is when only a few
projections are available. Figs 3.7 and 3.8 are reconstructed images
for one phase, using 180 and 10, respectively, equally spaced consistent
projections per phase. It is seen that when using few projections the
errors in the image become significant. The most obvious are the spoke
like artifacts emanating from the dense (high X-ray absorbing) region
of the "spine" - lower right of image. These artifacts obliterate any
detail in the reconstruction of the "heart" region - upper centre of
image. One of the characteristics of images reconstructed from
projections is that features within the object being imaged cause errors
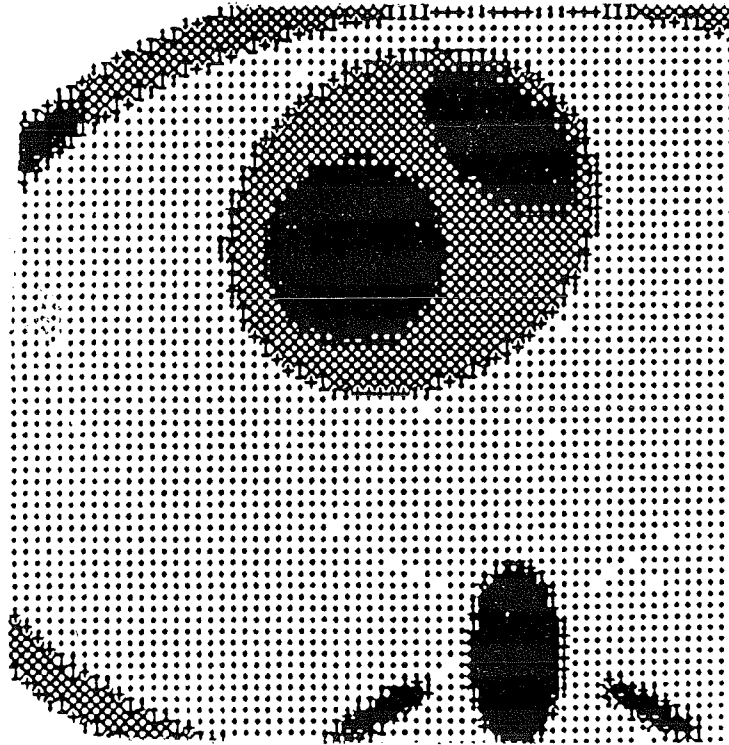over the whole of the image. Basically the absolute value of these
errors increases with:

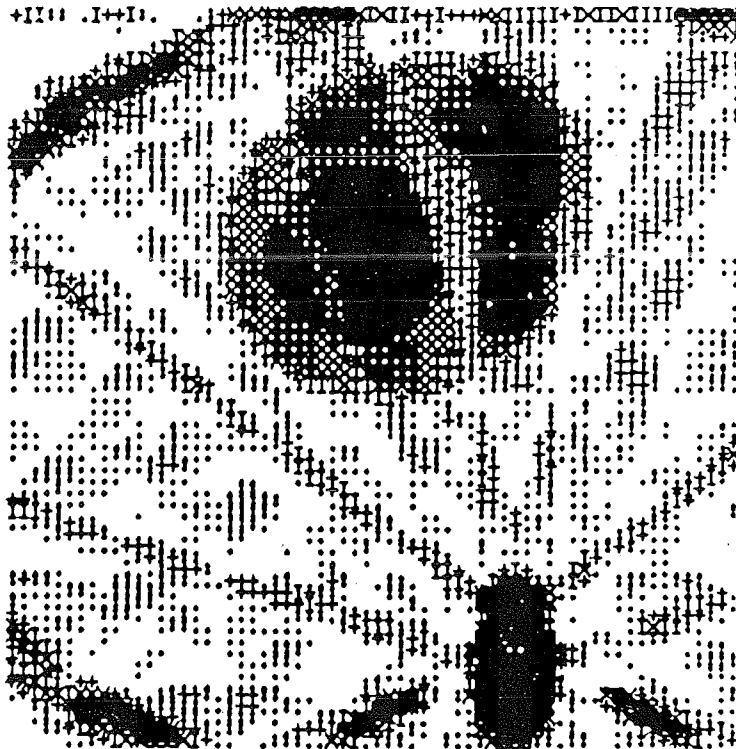Fig. 3.7   Image obtained using 180 consistent projections.

Fig. 3.8   Image obtained using only 10 consistent projections.

(i)   distance away from the feature within the image which is
      responsible for the particular errors, and

(ii) the difference in density between that of the feature responsible
      for the particular errors and that of the surrounding material.

Naturally these errors are reduced by increasing the number of
projections. Smith *et al.* (1973) show quantitatively how the
reconstruction errors depend upon the number of projections that are
used.

The main sources of the errors over the heart region then, are
the dense objects in the region surrounding the heart. Now it is worth
recognising that these dense objects are stationary and so can be
estimated from the full projection set - not just the 10 or so
projections available per heart phase. Consider the projections to be
composed of two components, one due to the stationary region and the
other due to the heart region. By invoking the "hollow projection"
methods of Lewitt and Bates (1978), the dense objects in the stationary
region can be reconstructed from the full set of projections - even in
the presence of heart motion. Therefore the contributions to the
projections of the stationary region can be estimated and subtracted
from the originally measured projections. The resulting "difference
projections" are in effect projections of the heart region only. This
is the essence of the subtraction approach B. Reconstructing images
of the heart region from these difference projections reduces the
amplitudes of the aforementioned artifacts and hence improves image
quality, because now only the heart region is being imaged, not the
whole chest cross section. The variation in X-ray absorbtivity levels
within the heart region is much less than within the stationary spinal
region - by a factor of about 10 (see next section). Therefore the
absolute errors in the image, obtained using the difference projections,
are significantly reduced, compared with the errors that occur when
the straightforward approach A is used.

3.6   RESULTS OF COMPUTER SIMULATION

To study the problems of imaging the beating heart using the
subtraction approach, B, (§3.5) a mathematical model of the chest was
developed. This model employs 13 ellipses to represent various X-ray

absorbing regions through a chest cross-section. Included in the model
are idealized "ribs", "sternum", "chest wall", "lung", "spine" and
"heart". The heart includes "heart muscle" and two "ventricles".
The X-ray absorbtivity assigned to each ellipse is representative of
that found in the corresponding regions of the body (Huang and Wu, 1976).
However, the value used in the ventricles has been increased over that
of blood to a level consistent with that obtained using a "normal"
intravenous injection of contrast material (Kivisaari et al., 1979).
X-ray absorbtivity values in C.T. are generally normalised so that air
has the value -1000 and water the value 0. On this scale the X-ray
absorbtivity of the model components are: "ribs' = 180, "sternum'= 180,
"chest.wall" = 20, "lung" = -200, "spine" = 200, "heart muscle" = 20,
and "heart ventricles" = 70 . The three ellipses used to represent the
heart are programmed to move in a continuous manner between the two
extremes shown in Figs 3.9 and 3.10. This simulates the beating heart.
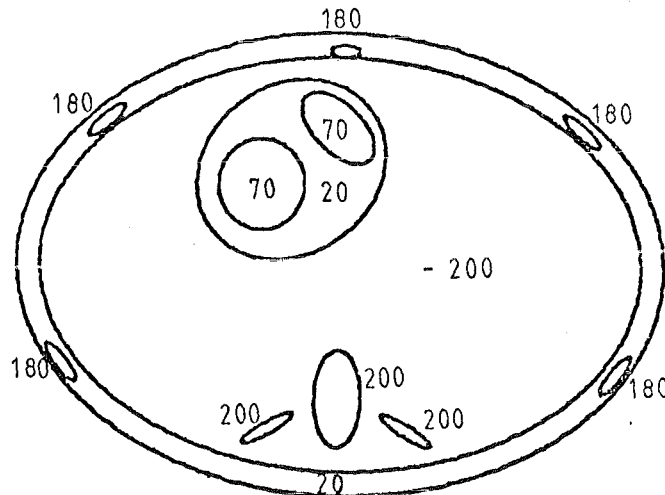

Fig. 3.9   The chest phantom when the "heart" is of maximum
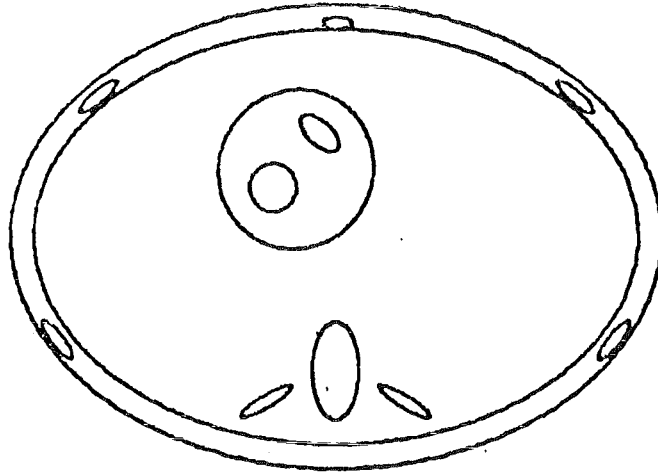   extent. Also shown are the relative absorbtivity
           levels.

Fig. 3.10   The chest phantom when the "heart"

if of minimum extent.

It is difficult to model the heart movement exactly as very little seems
to be known about it.   The extent of the motion of the simulated heart
is designed to be "a worst case".

The scan was simulated as follows: 180 projections, each
consisting of 255 points, were calculated - one projection per degree.
During the scan, 10 heart beats were simulated.   Thus 10 projections
per phase were obtained, and there were 18 different phases per heart
cycle.   To reconstruct the images the modified backprojection algorithm
was used (§3.2).   Refer to Figs 3.11 to 3.13, and 3.15 to 3.17, and note
that some of the images are presented at a different magnification to
permit clearer visualization of the structure.   The images were
calculated to a resolution of 256×256 pixels (picture elements).
However the display is only to a resolution of 64×64 pixels.

Reconstructing using all 180 projections results in a blurred
"heart" (see Fig. 3.11).   It is interesting, though, that much of the
detail in the stationary region is preserved - this is in accord with
actual C.T. chest scans.   The steps in the procedure for obtaining a
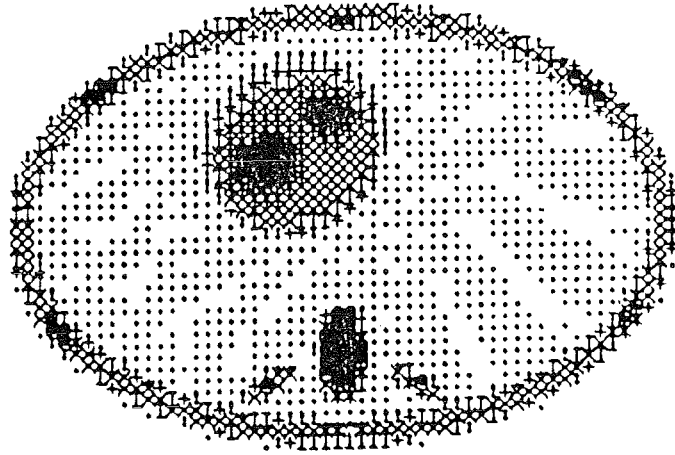useful image of any particular phase of the heart using only 10

Fig. 3.11   Reconstruction of the chest phantom containing
a "beating heart" using all projections.

projections are listed below:

(1)   Take the image obtained using the full 180 projections and
identify the heart region (see Fig. 3.12).

(2)   Remove the heart region from the image. This results in an
image representing the stationary region of the chest
(see Fig. 3.13).

(3)   Calculate projections through this image at angles corresponding
to those of the original projections.

(4)   Subtract these calculated projections from the originals to
obtain the difference projections. (The original projection,
the projection of the stationary region, and the difference
projection, for $\phi = -90^{\circ}$ are shown in Fig. 3.14(a), (b) and (c)
respectively.)

(5)   Use a set of 10 consistent difference projections to reconstruct
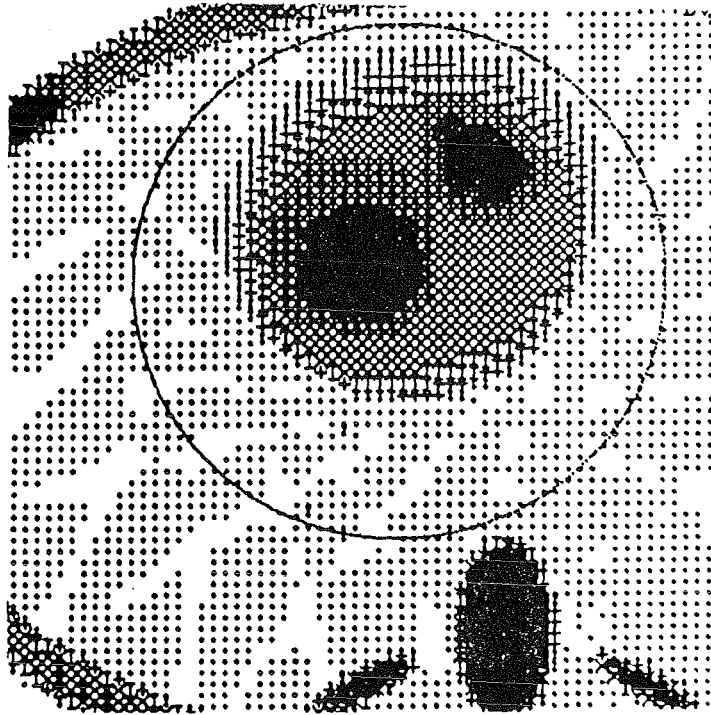an image of the heart (see Fig. 3.15).
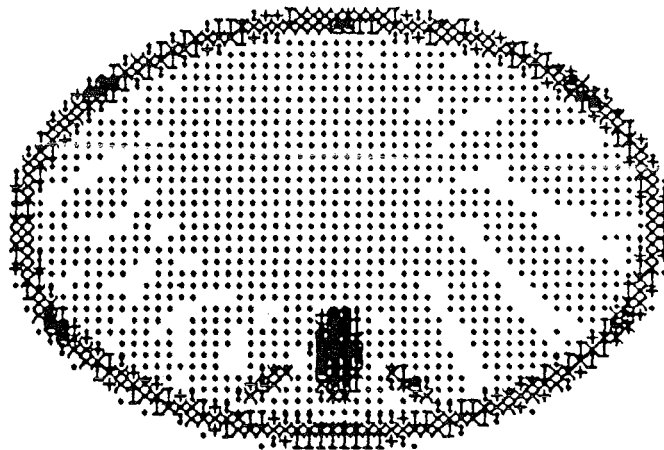
Fig. 3.12   A close up of the blurred "heart" region.

Fig. 3.13   Estimate of the stationary region. Obtained by
removing "heart" region from Fig. 3.11.

Original Projection

Estimated Projection
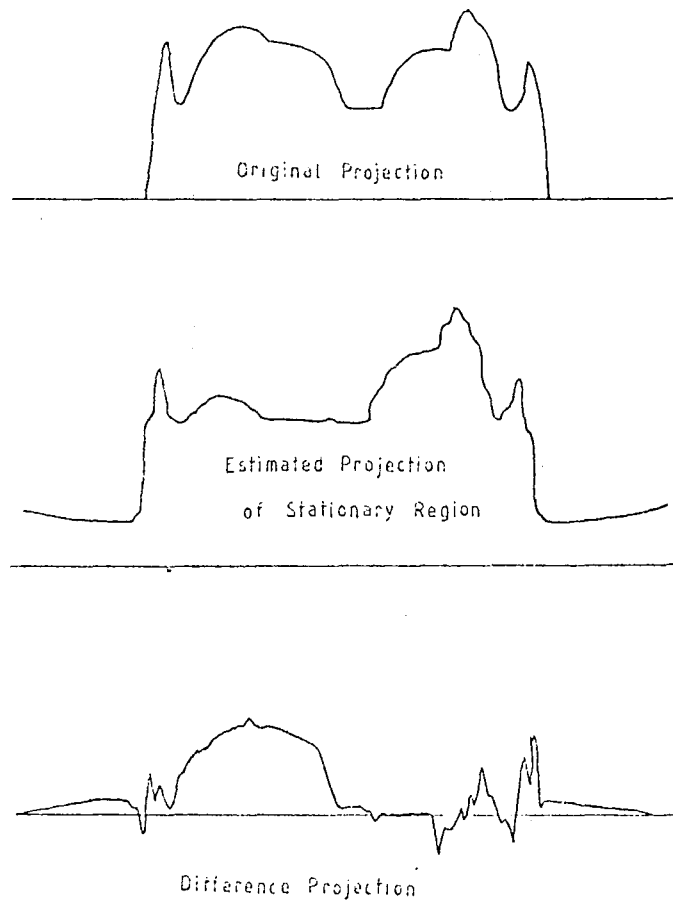of Stationary Region

Difference Projection

Fig. 3.14   The difference projections are obtained by subtracting
the projections through the estimated stationary
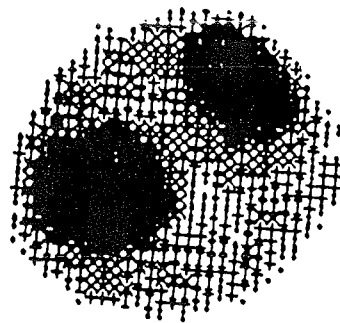region, from the original projections.

Fig. 3.15   Image of the "heart" obtained using
10 difference projections.

Compare the reconstructed image shown in Fig. 3.15 with the original object (Fig. 3.16), and the image obtained using 10 of the original projections (Fig. 3.17). Notice that errors over the heart region in Fig. 3.15 are significantly less than those in Fig. 3.17.
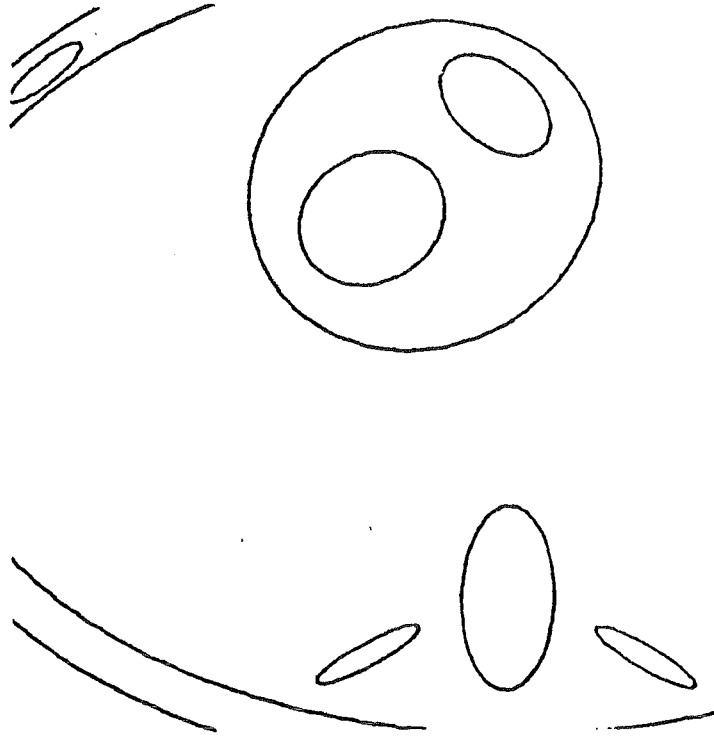

Fig. 3.16   The chest phantom as it appears during the "heart" phase corresponding to the reconstructed images of Figs 3.15 and 3.17.

Fig. 3.17   Image of the heart obtained using 10 of the original projections. (Same phase as Figs 3.15 and 3.16.)

Step (5) can be repeated for the other phases of the heart cycle to produce a time series of images depicting the heart throughout its cycle.

Steps (1) and (2) provide a "quick and nasty" method of estimating the stationary region. To obtain a more accurate estimate the hollow projection method of Lewitt and Bates (1978), mentioned in §3.5, should be used. The procedure is as follows:

(i)   Remove from the projections that section containing the "projection" of the heart region.

(ii) A set of hollow projections has now been obtained.

(iii) Using the consistent completion procedure, the hollow regions of the projections can be filled in, in such a manner as to provide a complete set of consistent projections.

(iv) These consistent projections are then used to reconstruct an estimate of the stationary region.

The consistent completion procedure (Lewitt and Bates, 1978) involves the following operations. The projections $f(\xi,\phi)$, as defined in §3.2, are decomposed into their angular Fourier series, i.e.

$$f(\xi,\phi) = \sum_{m=-\infty}^{\infty} f_m(\xi) \, \exp(im\phi) \quad . \tag{3.8}$$

Define the origin of coordinates such that the hollow regions of the projections lie in the range $|\xi| < a$, say, where a is a constant. Thus the $f_m(\xi)$ are unknown for $\xi < a$. In practice the object being imaged is of finite extent, therefore $f_m(\xi) = 0$, for $\xi > b$, say, where b is another constant. The method of Lewitt and Bates (1978) involves fitting a series of polynomials, which exist over the range $0 < \xi < b$, to the $f_m(\xi)$ (which are known for $a < \xi < b$). The "polynomials" belong to a special set which ensure that the $f_m(\xi)$ are consistent with respect to projections through a stationary object of finite extent. Their method is basically an analytic continuation process to estimate $f_m(\xi)$, in a consistent manner, over the region $0 < \xi < a$. Then, by invoking (3.7), a consistent set of $f(\xi,\phi)$ can be obtained.

However, with respect to heart imaging, consistent completion is probably unnecessary in practice. It has been found that the original projections usually behave as though they have been satisfactorily "completed". Even though the original projections are strictly inconsistent, due to heart movement, the effective extent

of the blurring seems always to be localised to the heart region. The stationary region of the chest is thus imaged quite sharply, implying that there is no need to invoke more sophisticated processing methods.

3.7   DISCUSSION

Two methods of high resolution cardiac imaging based on E.C.G. gated C.T. have been presented. The method described in §3.4 outlines the design for a new type of scanner. Although there are considerable practical difficulties in the way of producing a scanner capable of performing multiple revolutions, the approach suggested here would be very much more economical than that which uses multiple X-ray sources and detector banks (Robb et al., 1979).

The method outlined in §§3.5 and 3.6 is attractive, in the short term at least, because it can be implemented easily on existing scanners. The major question with this method is whether or not the differences in X-ray absorbtivities over the heart are sufficient to highlight the features of the organ. As there is a large amount of blood within the heart muscle, the heart muscle and chambers are almost homogeneous. It would seem that contrast material would have to be injected into the blood to enhance the absorbtivity differences between blood and muscle. But just how much would be required is unknown at this stage.

Others (Harell et al., 1977; Berninger et al., 1979; and Sagel et al., 1977) have reported E.C.G. gated C.T. schemes for imaging the heart. However, the images that are obtained are not particularly clear compared with C.T. images of stationary objects. It seems that there are two probable reasons for this:

(i)   The total scan time is too long, resulting in violation of the stationarity constraints mentioned in §3.3.

(ii)   The E.C.G. gating window (i.e. the interval over which the heart is assumed to be quasi-stationary) is also too long.

The previously reported schemes have attempted to obtain a full set of projections, for a particular phase of the heart, so that existing scanner software could be used to reconstruct the images. Because of this they have had to compromise on the consistency of the projections, which results in a poor image of the heart. The scheme presented in §§3.5 and 3.6 shows that by using algorithms specifically tailored

for heart imaging, it is possible to obtain reasonable results even from only a small number of projections.

4.   A LIMITATION ON ULTRASONIC TRANSMISSION TOMOGRAPHY

4.1   INTRODUCTION

Largely because of the success of X-ray computed tomography
(C.T.) (see for example Brooks and Di Chiro, 1976), increased interest
is being shown in ultrasonic transmission imaging (Dines and Kak, 1979;
Mueller et al., 1979; Glover and Sharp, 1977; Schomberg, 1978; Johnson
et al., 1975; Bates and Dunlop, 1977; Bates and McKinnon, 1979).
Unfortunately, however, ultrasonic C.T. is not enjoying the same success
as its X-ray counterpart. This is generally attributed to the lack of a
good solid theoretical foundation for the image reconstruction procedure.
The problem is that, even though both ultrasound and X-rays can be
treated in terms of rays, the propagation of ultrasonic rays is more
complicated than that of X-rays because, with ultrasound, reflection
and refraction effects are significant.

There are two general approaches to ultrasonic C.T. One uses
data obtained from attenuation measurements to form images of
absorbtivity distributions throughout body cross-sections. The other,
which images refractive index or velocity profiles, relies on
measurements of the propagation times of pulses travelling through
body cross-sections. These attenuation and propagation time measurements
can be converted into projections similar to those of X-ray C.T. (see
§3.2). That is, the line integrals of the absorbtivity or refractive
index distribution, along rays, can be derived from the measurements.
However, unlike X-rays which are straight, ultrasonic rays, being
subject to refraction effects, are generally curved. Thus the simple
and numerically efficient image reconstruction methods which contribute
much to the success of X-ray C.T., can not be applied exactly to
ultrasonic transmission tomography. Nevertheless "an image" can be
obtained by effectively treating the ultrasonic rays as being straight,
and using a standard X-ray C.T. reconstruction algorithm to process the
projections. Now, whether one is interested in absorbtivity or
refractive index distributions, the latter is required if ray curvature
is to be taken into account. This chapter deals only with the problem
of determining refractive index distributions from propagation time
measurements. Analytic solutions to this problem are known only for

very simple cases. For example, it is possible to determine circularly symmetric refractive index distributions from propagation time measurements - provided that the refractive index is monotonically increasing with distance from the centre of the distribution (see Bullen, 1963). Numerical "solutions" which allow for ray curvature have, so far, all been iterative in nature (Schomberg, 1978; Johnson et al., 1975; Bates and McKinnon, 1979). The initial input to the iterations has, in all cases, been the image derived assuming that the rays were straight.

While developing a new iterative method (see Appendix to this chapter), and also investigating the errors introduced into the image when the curved-rays are assumed to be straight, a particular limitation of ultrasonic transmission tomography became apparent. This is the main result reported in this chapter.

In §4.2 the fundamentals of an ultrasonic transmission imaging system, and other necessary preliminaries, are presented. In §4.3 it is shown that in most media there are regions which it is impossible to image when ray propagation is assumed. This is the limitation mentioned in the previous paragraph. Results of a computer simulation study, illustrating the effects of this limitation, are described in §4.4. Then, in §4.5, the practical significance of the results presented in this chapter are assessed, and related to previously reported theoretical and experimental results.

4.2   ULTRASONIC TRANSMISSION IMAGING PRELIMINARIES

The essentials of the envisaged ultrasonic transmission measurement system are shown in Fig. 4.1. Within a cross-section of the body being examined an arbitrary datum is chosen. Ultrasonic transmitter (Tx) and receiver (Rx) transducers move along lines, AA' and BB', respectively, which are parallel to the ξ axis. The coordinates (ξ, ϕ) specify the transmitter-receiver pair position. A set of "time-delay projections" is obtained by measuring the propagation time of a pulse travelling through the body from transmitter to receiver for many pairs of values of ξ and ϕ. The term "projections" is borrowed from standard X-ray C.T. jargon (see §3.2). It is known that the propagation of ultrasonic energy through many of the types of body of interest in biomedical applications can be usefully treated in terms
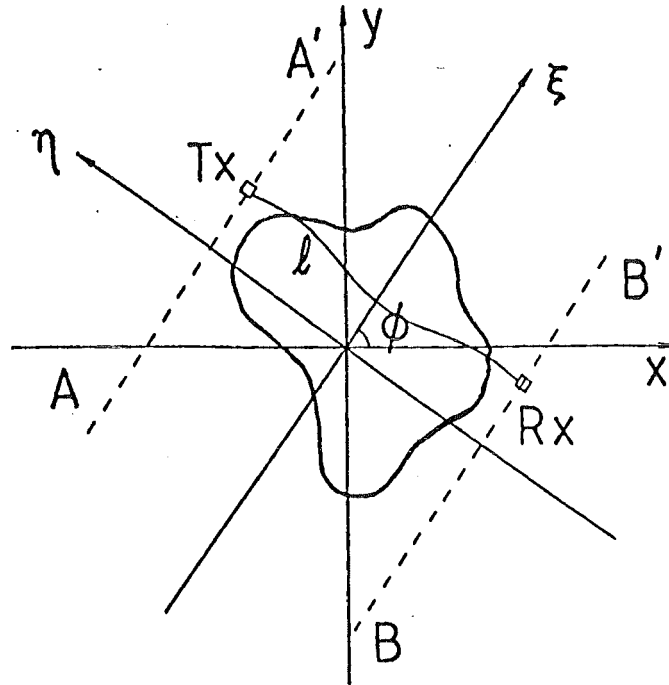
Fig. 4.1   Configuration for an ultrasonic transmission imaging system.

of rays which obey Fermat's Principle of minimum propagation time

(Friedlander, 1958; and §1.7). Thus each time-delay projection,

$f(\xi,\phi)$, is proportional to the integral of the ultrasonic refractive

index $v = v(x,y,z)$ evaluated along the ray path $\ell = \ell(\xi,\eta,\phi,z)$ for

which the propagation time between the transducers is a minimum.

The direction of the z axis is perpendicular to the plane of Fig. 4.1.

Since v varies in an initially unknown manner, the form of each ℓ is

unknown *a priori*. As the ray path is unknown, transmission tomography

is more complicated for ultrasound than for X-rays.

The limitation discussed in §4.3 applies whether or not v varies

with z . Since the analysis is much simplified if v is independent of

z , from now on it is assumed that $v = v(x,y)$, so that $\ell = \ell(\xi,\eta,\phi)$.

The simplest method of reconstructing an "image" of $V(x,y)$ from measured values of $f(\xi,\phi)$, is to assume that each ℓ is a straight line between the transducers $(\ell = \xi)$. The time-delay projections can then be interpreted exactly as in X-ray C.T. (§3.2). This is referred to here as a "straight-ray reconstruction". The literature contains several proposals for treating straight-ray reconstructions as initial inputs for iterative schemes (Schomberg, 1978; Johnson et al., 1975; Bates and McKinnon, 1979). These iterative schemes allow for each ℓ being a curved ray path and iterate, supposedly, to the actual image. The general nature of these "iterative correction schemes" is as follows:

(i)   Obtain an estimate of the image from the measured projections (by assuming the rays to be straight).

(ii) Calculate ray paths through the estimated image - between points corresponding to the positions of the transmitter and receiver during the propagation time measurements.

(iii) Calculate the propagation time along the ray paths found in (ii) (i.e. integrate the refractive index over the ray path). The calculated propagation times will, in general, differ from the measured values, as the estimated image is not necessarily the same image. Now, because "straight-ray reconstruction" algorithms, effectively, are used in the iteration cycle, estimates of the "straight-ray projections" are required. [Note that if the actual straight-ray projections could be obtained the object being imaged could be estimated exactly.] As the curved-ray projections denote the integral of refractive index over the minimum propagation time rays, they are necessarily less than or equal to the corresponding straight-ray projections. By obtaining the difference between the measured projections and the curved-ray projections calculated through the estimated image, and adding this difference to the projections used to obtain the estimated image, an estimate of the straight-ray projections is produced.

(iv) Use these estimated straight-ray projections to calculate another estimate of the image (using a straight-ray reconstruction algorithm).

(v)   The idea now is to repeat steps (ii) - (v), updating the estimated straight-ray projections, until the calculated propagation time for each ray is the same as the measured value.

The major computational difficulty with the earlier reported
iterative correction schemes (Johnson et al., 1975; Schomberg, 1978)
is that ray paths have to be calculated through each estimated image.
In an imaging system having a useful resolution (say 180 projections
at 100 points per projection) this amounts to calculating rays between
18,000 pairs of points. However a more efficient iterative algorithm
has been developed (Bates and McKinnon, 1979) in which the explicit
evaluation of the ray paths is not required. The algorithm has been
updated since the "Bates and McKinnon, 1979" paper - explanatory details
are set down in the Appendix to this chapter. One of the hazards of
using iterative correction schemes is that it is difficult to know
whether the image has converged to the true image or not. In §§4.3
and 4.4 it is shown that for certain refractive index distributions
there will be regions where iterative correction schemes will not
converge to the true image.

The limitation on ultrasonic C.T. with which this chapter is
concerned can be demonstrated by considering circularly symmetric
distributions of refractive index. Objects characterised, respectively
by piecewise constant and continuously varying distributions are
examined. For the piecewise constant case the ray paths are calculated
by applying Snell's law at each discontinuity in the refractive index.
Also, in the piecewise constant case, those rays which creep around
discontinuities have to be considered, as they may describe the path
with the shortest propagation time. For the continuously varying case
the integral formulas from Bullen (1963) are invoked. In some cases
there will be more than one ray path between two points. Whenever this
occurs the minimum propagation time ray path is chosen.

Snell's law can be developed from the eikonal equation (1.103)
and (1.114). The eikonal equation is

$$|\nabla s|^2 = \nu^2 \quad , \qquad\qquad\qquad (4.1)$$

where s is the "optical path length", and ν is the refractive index.
[s is equivalent to the function describing the propagation time of a
pulse travelling through a refracting media. The term "optical path
length" is borrowed from optics - that being the field where much of
the early work on refraction was carried out. "Optical path length" is
used here, in a general sense, to denote the integral of the refractive

index over the ray path.] Now because ∇s is in the direction of the ray (see §1.7), (4.1) can be rewritten as

$$\partial s/\partial \ell = \nu \quad , \tag{4.2}$$

where $\partial \ell$ is the differential element along the ray. Taking the grad of (4.2), and assuming that the order of the differentiation can be reversed, results in (see Born and Wolf, 1970)

$$\partial (\nabla s) / \partial \ell = \nabla \nu \quad . \tag{4.3}$$

Consider the case where the refractive index is varying in only one direction. That is $\nu = \nu(x)$ (see Fig. 4.2). Define $\hat{x}$ to be the unit vector in the x direction.
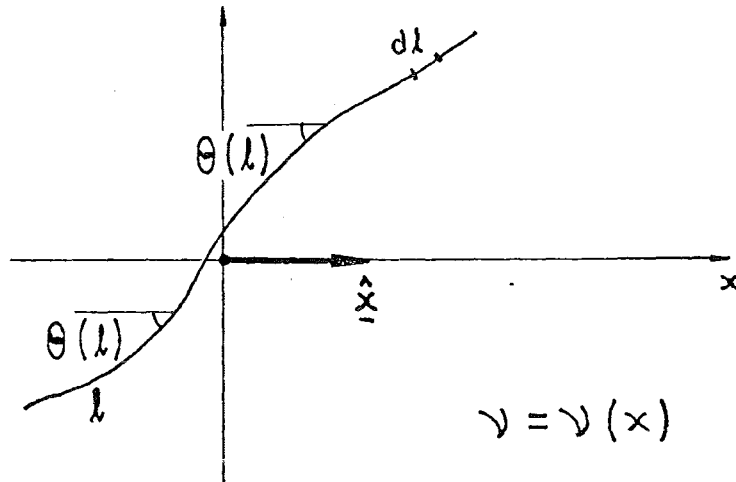

Fig. 4.2   A ray penetrating a medium in which the refractive index is varying in only one direction.

Now consider the variation of the vector $(\hat{x} \times \nabla s)$ along the ray ($\times$ denotes the vector cross product). That is

$$\partial (\hat{x} \times \nabla s)/\partial \ell = \hat{x} \times \partial (\nabla s)/\partial \ell \quad . \tag{4.4}$$

Since $\nabla\nu$ is in the x direction; using (4.3), the term on the R.H.S. of (4.4) can be shown to be zero. Therefore, with this type of refractive index distribution, $(\hat{x} \times \nabla s)$ is constant along each ray. Another way of writing this result, using (4.2) and Fig. 4.2, is

$$\nu \sin \theta = \text{const.} \quad , \tag{4.5}$$

which is Snell's law.

A similar result can be achieved for the case of a circularly symmetric refractive index distribution, $\nu = \nu(r)$ (see Fig. 4.3).
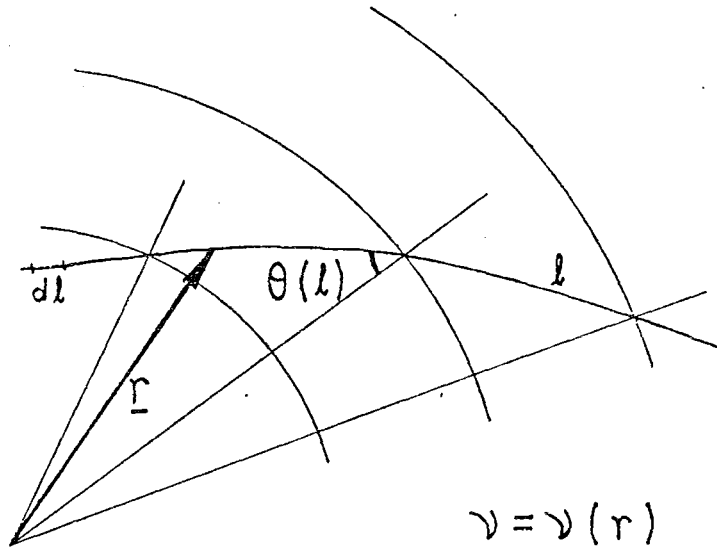

Fig. 4.3   A ray penetrating a circularly symmetric refractive index distribution.

Define the vector $\underline{r}$ to be the position vector of a point on a ray, and consider the variation of the vector $(\underline{r} \times \nabla s)$ along a ray. That is

$$\partial(\underline{r} \times \nabla s)/\partial\ell = \partial\underline{r}/\partial\ell \times \nabla s + \underline{r} \times \partial(\nabla s)/\partial\ell \quad . \tag{4.6}$$

Since $\partial\underline{r}/\partial\ell$ is in the same direction as ∇s, the first term on the R.H.S. is zero. And because $\nabla\nu$ is in the same direction as $\underline{r}$; using (4.3), the second term is also zero. This implies that, along a ray, $\underline{r} \times \nabla s$ is

constant. Thus, along each ray

$$\nu r \sin\theta = \text{const.} \quad ; \qquad\qquad (4.7)$$

where θ is the angle between the ray and $\underline{r}$.

In the case of continuously varying circularly symmetric refractive index distributions it is possible to derive analytic formulas for the propagation time of a pulse travelling along a ray path (Bullen, 1963). Now the constant in (4.7) is a parameter describing a particular ray. Denote the ray parameter by γ . Let P be any point of a ray whose ray parameter is γ . Let the polar coordinates of P be (r,ϕ) , as shown in Fig. 4.4, and denote the distance along the ray by ℓ .
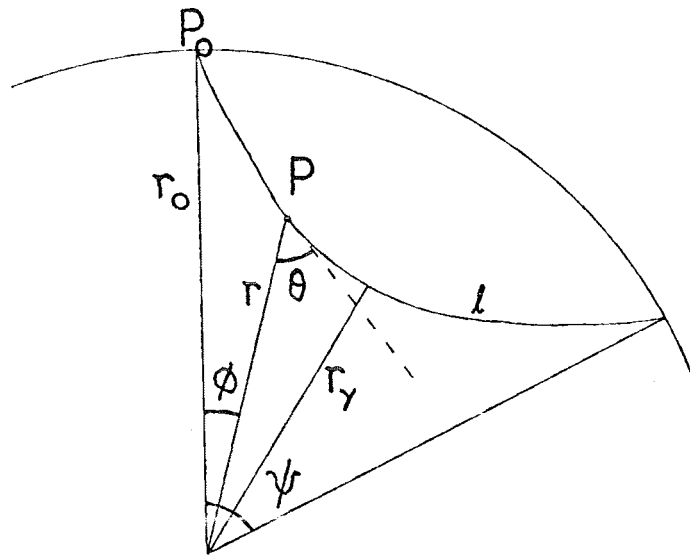

Fig. 4.4   A ray path between two points on the circumference of a circle. The refracting medium within the circle is circularly symmetric.

Note that the ray is between two points on the circumference of a circle. The angular separation of these two points is denoted by ψ . θ is the angle between the tangent to the ray, at P, and the line between P and

the origin. Using the Sine Law for triangles,

$$d\phi/d\ell = (\sin\theta)/r \quad . \tag{4.8}$$

Therefore (4.7) can be written as

$$r^2 \nu d\phi/d\ell = \gamma \quad . \tag{4.9}$$

Using the relationship

$$(d\ell)^2 = (dr)^2 + r^2(d\phi)^2 \quad , \tag{4.10}$$

and (4.9), one can obtain

$$r^4 \nu^2/\gamma^2 = (dr/d\phi)^2 + r^2 \quad . \tag{4.11}$$

That is

$$d\phi = \pm\gamma r^{-1}(\eta^2 - \gamma^2)^{-\frac{1}{2}} dr \quad , \tag{4.12}$$

where η is defined by

$$\eta(r) = r\nu(r) \quad . \tag{4.13}$$

Upon integrating (4.12) between P_o (the origin of the ray) and the deepest point of the ray, the expression for ψ is found to be

$$\psi = 2\gamma \int_{r_\gamma}^{r_o} r^{-1}(\eta^2 - \gamma^2)^{-\frac{1}{2}} dr \tag{4.14}$$

(where r_γ is the distance between the deepest point of the ray and the origin). If, this time, $d\phi$ is eliminated between (4.9) and (4.10), $d\ell$ can be shown, in a similar manner, to be

$$d\ell = \pm \eta (\eta^2 - \gamma^2)^{-\frac{1}{2}} dr \quad . \tag{4.15}$$

Since the propagation time s is equal to the integral of $\nu d\ell$ along the ray, using (4.15) a formula for s can be obtained, namely,

$$s = 2 \int_{r_\gamma}^{r_o} \eta^2 r^{-1}(\eta^2 - \gamma^2)^{-\frac{1}{2}} dr \quad . \tag{4.16}$$

Also, by elementary algebra

$$s = 2 \int_{r_\gamma}^{r_o} \{\gamma^2 r^{-1} (\eta^2 - \gamma^2)^{-\frac{1}{2}} + r^{-1} (\eta^2 - \gamma^2)^{\frac{1}{2}} \} \, dr \quad . \tag{4.17}$$

Hence, by (4.14),

$$s = \gamma\psi + 2 \int_{r_\gamma}^{r_o} r^{-1} (\eta^2 - \gamma^2)^{\frac{1}{2}} \, dr \quad . \tag{4.18}$$

Further, because $\sin\theta = 1$ at the deepest point of the ray, r_γ is given by the solution to

$$r_\gamma \, \nu(r_\gamma) = \gamma \quad . \tag{4.19}$$

[Note that the integral equations (4.14) and (4.16) have singularities at $r = r_\gamma$. The singularities are, however, integrable. To evaluate the integrals define a new variable $\beta = (\eta^2 - \gamma^2)^{\frac{1}{2}}$; thus $d\beta = (\eta^2 - \gamma^2)^{-\frac{1}{2}} \eta(\partial\eta/\partial r) dr$. By changing the integration variable to β , as defined here, the singularity "disappears".]

4.3   A LIMITATION ON THE RAY APPROACH TO ULTRASONIC C.T.

Suppose that the transducers shown in Fig. 4.1 are immersed in a fluid medium (refractive index normalised to unity), and that a circular cylinder of radius r and constant refractive index $\bar{\nu} > 1$ is inserted between the transducers. Refer to the two sets of ray paths shown in Fig. 4.5. The first set (dashed lines) are constructed so as to obey Snell's law at the cylinder boundary, and the second set creep around the edge of the cylinder. Since the ultrasound travels faster through the surrounding medium than through the cylinder, the creeping rays must be considered. For sufficiently large values of $\bar{\nu}$ none of the "shortest optical path length" rays penetrates the cylinder. For instance consider the transducers to be positioned symmetrically about the cylinder (i.e. $\xi = 0$ in Fig. 4.5). Further suppose that the transducer separation D is appreciably larger than r . The optical path length (or normalised propagation time) of the ray that creeps around the cylinder is then approximately $(D + 2r^2/D)$. The optical path length of the ray that passes straight through the cylinder is $(D + 2r(\bar{\nu} - 1))$.
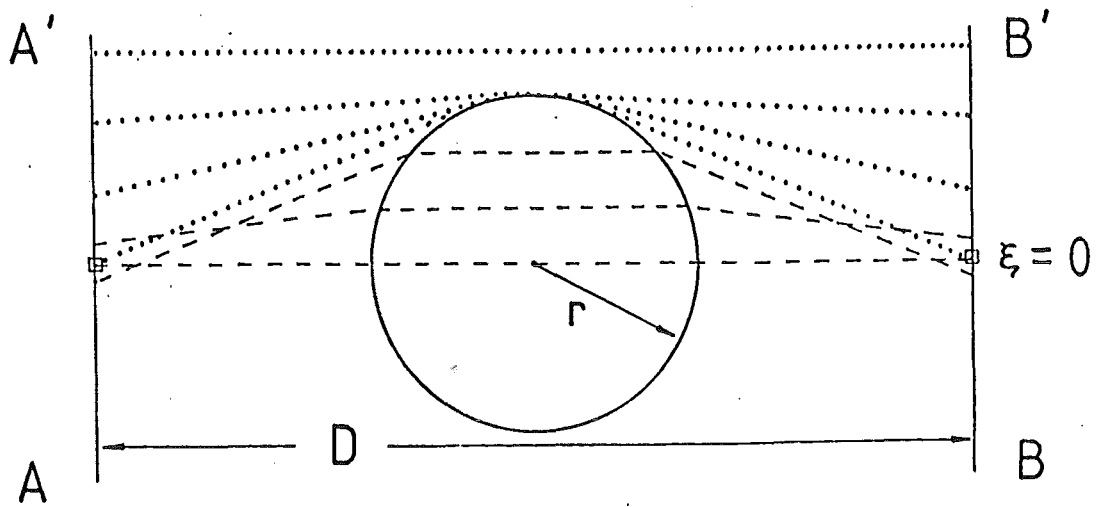
Fig. 4.5   Rays passing through a circular cylinder which
            has a refractive index greater than that of the
            surrounding medium.

So, if

$$\bar{v} > 1 + r/D \quad ,$$

(4.20)

none of the shortest optical path length rays passes through the
cylinder.

Suppose now that the object placed between the transducers has
a continuously varying refractive index $v = v(r)$. Suppose further that

$$\partial v/\partial r < 0 \quad ; \; r < r_a \quad .$$

(4.21)

Rays travelling through the part of the object for which $r < r_a$ all
curve around the origin (Bullen, 1963) (see Fig. 4.6), except for the
ray which travels straight through the centre. The optical path length
of this central ray is

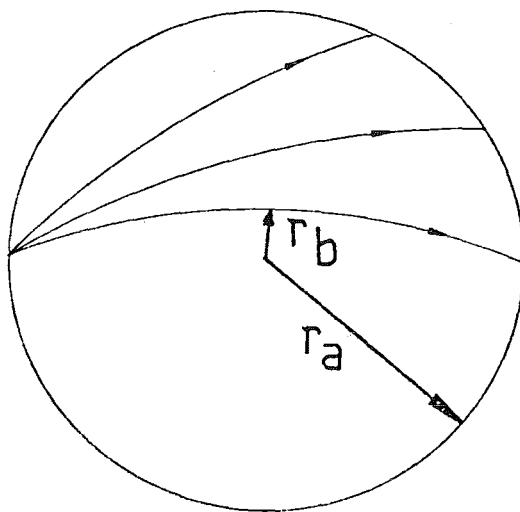$$T_s = 2 \int_0^{D/2} v \, dr \quad .$$

(4.22)

Fig. 4.6   When $\partial\nu/\partial r$ is negative the rays curve around the origin.

Now consider the curved ray travelling between transducers placed symmetrically about the object. As the curvature is towards the origin in the region $r < r_a$, this ray must pass through a minimum value of r, denoted here by $r_b > 0$. Write $\nu(r_b) = \nu_b$. The optical path length of this ray, from (4.18), is

$$T_c = r_b \nu_b \pi + 2 \int_{r_b}^{D/2} r^{-1}[(r\nu)^2 - (r_b\nu_b)^2]^{\frac{1}{2}} \, dr \quad . \tag{4.23}$$

When $T_s > T_c$, no minimum propagation time rays intersect that part of the object for which $r < r_b$. From (4.22) and (4.23) this occurs when

$$2 \int_0^{r_b} \nu \, dr > r_b \nu_b \pi - 2 \int_{r_b}^{D/2} \nu\{1 - [1 - (r_b \nu_b / r\nu)^2]^{\frac{1}{2}}\} \, dr \quad . \tag{4.24}$$

Two main points arise from the analysis in this section.

(i)   There are "forbidden regions" into which minimum propagation time rays never penetrate. It is, therefore, impossible to reconstruct the refractive index faithfully within these regions.

A forbidden region can surround any point where ν exhibits a localised maximum.

(ii) The sizes of the forbidden regions depend, in general, not only on the variation of ν but also on the transducers' separation. This separation should be as small as physical constraints permit in order to minimise the forbidden regions' sizes.

4.4   RECONSTRUCTIONS

In this section four circularly symmetric refracting objects are studied. The time-delay projection "data" used in this chapter are obtained from computer simulations which involve calculating ray paths through the objects. Both straight-ray reconstructions (using the modified back projection method - see §3.2) and iteratively corrected reconstructions (using the procedure described in the Appendix) are presented in this section. For each example, the amplitude of the refractive index is shown along a line passing through the centre of the distribution (full line represents the actual distribution, dashed line represents the straight ray reconstruction, and dotted line represents the iteratively corrected reconstruction).

Fig. 4.7 shows what happens when a cylinder of constant refractive index $\bar{\nu}$ is immersed in a medium of refractive index (normalised to) unity. When $\bar{\nu} < 1$ (Fig. 4.7a), so that there cannot be a forbidden region, both the straight-ray and the iteratively corrected reconstructions conform quite closely to the ideal distribution. Notice that the dotted line shows little improvement on the dashed line. This is due to the abrupt change in refractive index at the edge of the cylinder. When $\bar{\nu} > 1$ and D are large enough that there is a forbidden region (Fig. 4.7b), not only is the straight ray reconstruction less accurate than when $\bar{\nu} < 1$, but the iteratively corrected reconstruction shows signs of diverging from, rather than converging to, the actual distribution.

Figures 4.8 and 4.9 refer to the third and fourth examples respectively. The time-delay projections were calculated on the assumption that the transducers were placed on the circumferences of the circles shown in Figs 4.8a and 4.9a.
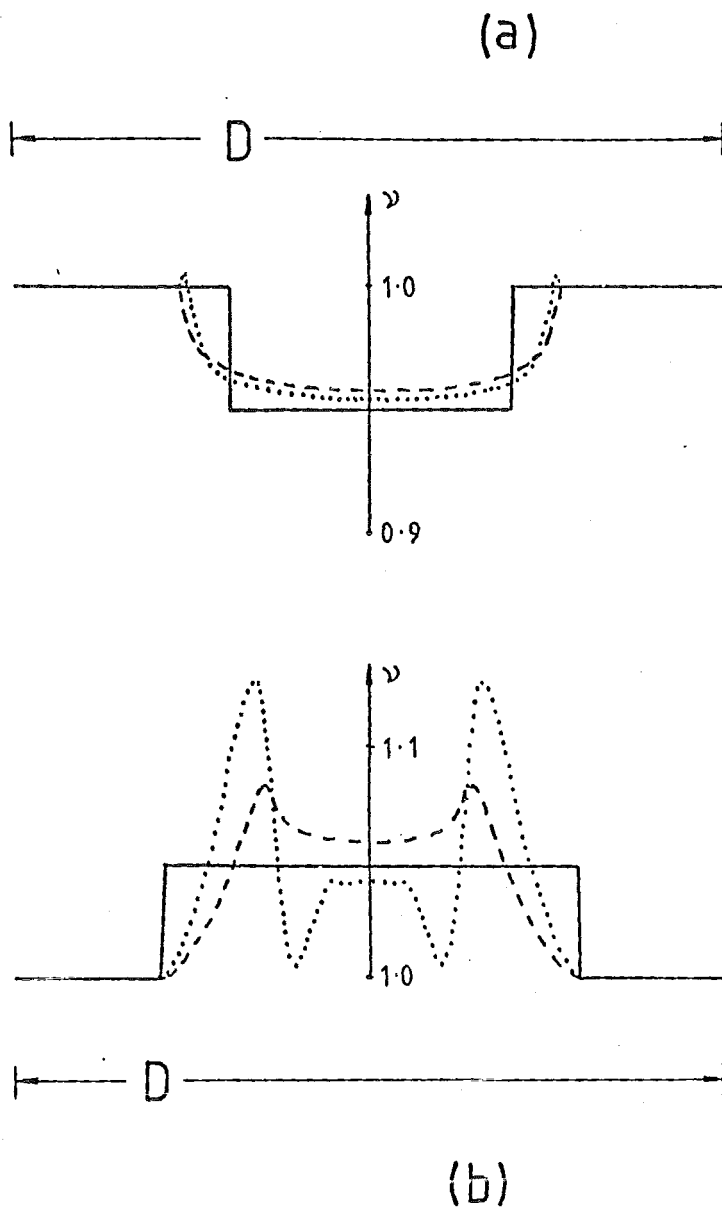
Fig. 4.7   Reconstruction of a cylinder immersed in a medium of
unit refractive index: (a) First example ($\bar{\nu}$ = 0.95);
(b) Second example ($\bar{\nu}$ = 1.05).

———————   actual object

— — —   straight ray reconstruction

. . . . . .   iteratively corrected reconstruction.
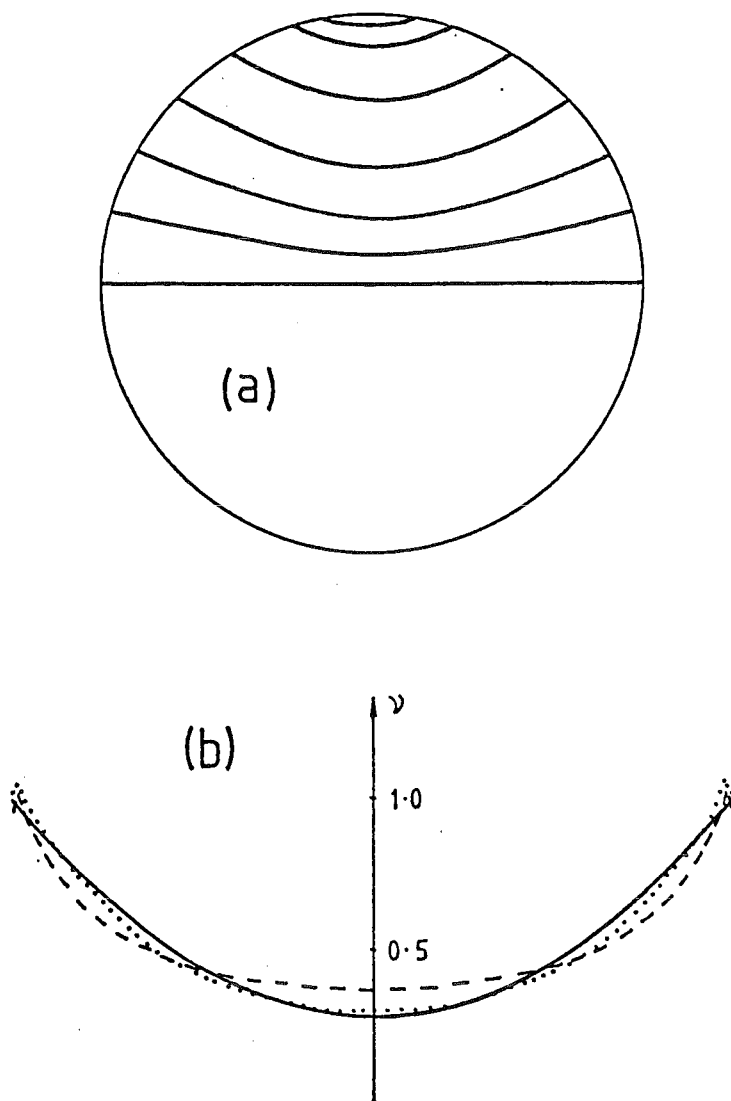
Fig. 4.8   Third example: (a) ray paths;   (b) reconstructions.

————————   actual distribution

— — —   straight ray reconstruction

. . . .   iteratively corrected reconstruction.

Fig. 4.9   Fourth example: (a) ray paths - note "forbidden

region".   (b) reconstructions.

—————   actual distribution

— — —   straight-ray reconstruction

. . . . .   iteratively corrected reconstruction.

There is no forbidden region for the third example and, not only is the straight-ray reconstruction an encouraging approximation to the actual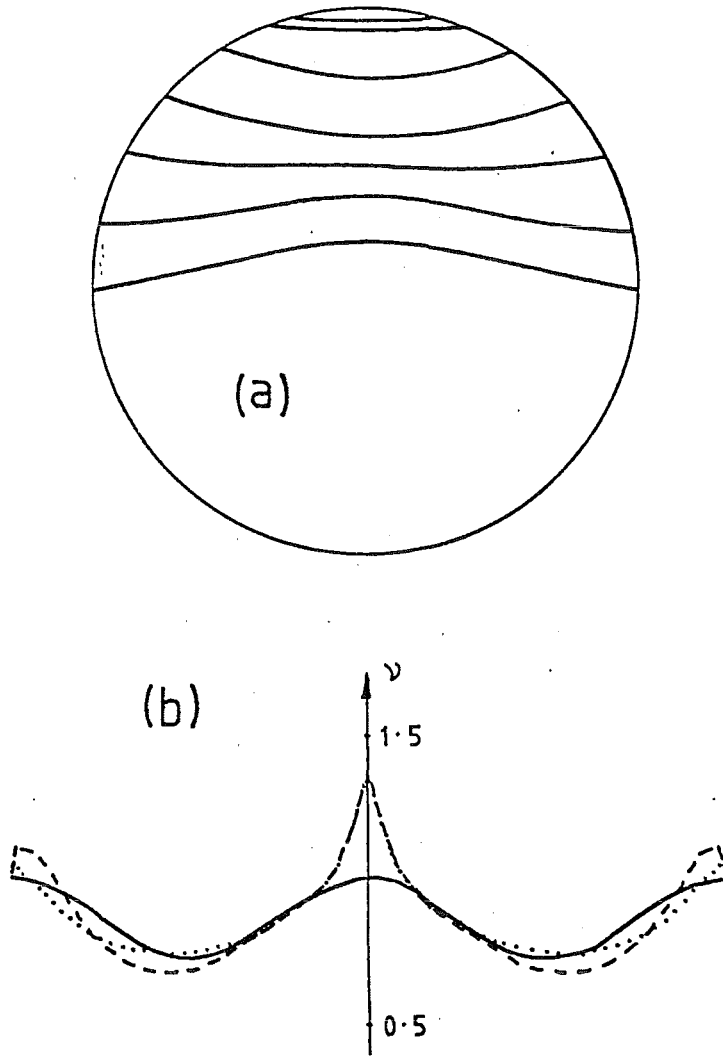 distribution, the iteratively corrected reconstruction shows definite improvement everywhere. There is a forbidden region for the fourth example - neither reconstruction is as close to the actual distribution as in the third example, and the iteratively corrected reconstruction only shows improvement outside the forbidden region. Note that if one did not know the form of the ideal distribution, the closeness of the dashed and dotted curves in Fig. 4.9b might suggest that true convergence is being manifested.

4.5   DISCUSSION

The problem of ray curvature in ultrasonic transmission tomography has been discussed by several researchers (Glover and Sharp, 1977; Schomberg, 1978; Johnson et al., 1975; Bates and Dunlop, 1977; Bates and McKinnon, 1979). The simplest method of treating it is to assume that the rays are straight. However, various schemes have been reported which attempt to improve on this (Schomberg, 1978; Johnson et al., 1975; Bates and McKinnon, 1979). Both Johnson et al. (1975) and Schomberg (1978) have presented images that have had refractive corrections applied to them, but it is difficult to notice any real improvement in image quality. One of the dangers of using iterative methods is that if an image seems to be "converging" there is a tendency to believe that it is approaching the true image. It has been shown in this chapter that this is not necessarily correct. The application of an iterative correction procedure may, however, compensate for certain distortions.

It is worth emphasising that many of the results reported in the literature for straight-ray reconstructions are of a potentially useful faithfulness (Dines and Kak, 1979; Glover and Sharp, 1977; Glover, 1979; Schomberg, 1978; Johnson et al., 1975; Bates and Dunlop, 1977), which suggests that ray curvature need not seriously degrade the quality of the reconstructed images. The results shown in §4.4 (and others obtained in the course of this work) indicate that straight ray reconstructions are adequate for many practical purposes if the refractive index never departs by more than about 10% from its mean value. Because of both this and the basic limitation on ultrasonic C.T. discussed in this chapter, it seems that it is probably unnecessary and

possibly misleading in practice to attempt to correct straight ray
reconstructions.

APPENDIX: A FAST ITERATIVE CORRECTION METHOD

In this appendix an iterative correction scheme for recovering
an image from "curved-ray" type projections is described. This approach
has a computational advantage over most of the other correcting schemes
in that ray paths are not explicitly calculated. This is an important
consideration as an ultrasonic transmission imaging system could make
use of about 15,000 "rays". Enormous computations would be required
to calculate such a large number of individual ray paths.

The method is based on estimating the difference between the
straight-line projections through the object, which are unknown, and
the measured time-delay projections. This difference is used to modify
the measured projections to obtain a "good estimate" of the straight-
line projections. Remember that if the latter could be found exactly,
then the true image could be accurately reconstructed by standard
procedures (see for example Lewitt *et al.*, 1978). The algorithm
consists of the following steps:

  (i)   Reconstruct an estimate of the image using the modified back
        projection method (see §3.2) on the measured time-delay
        projections.
  (ii) Calculate the difference between the straight-line and time-delay
        projections through this image.
  (iii) Add the difference to the measured projections to obtain a set
        of corrected projections.
  (iv) Reconstruct another estimate of the image using the modified
        back projection algorithm on the corrected projections.
  (v)   Repeat iteration (go to step ii) until the image converges,
        if in fact it does.

The major difficulty is estimating the time-delay projections
through the image. The two previously reported iterative correction
methods (Johnson *et al.*, 1975; Schomberg, 1978) effectively use
equation (4.3), that is

$$\partial (\nabla s) \: / \: \partial \ell = \nabla \nu \quad , \tag{A1}$$

to trace out a ray through the refracting medium ν. The ray is in the direction of ∇s (where s denotes normalised propagation time, or optical path length; and ℓ is length along the ray). The time-delay projection is found by integrating ν over the ray path. In the method described in this appendix, the time-delay projection is obtained in a rather different manner. This involves solving the eikonal equation (see §§1.7 and 4.2),

$$\left| \nabla s \right|^2 = \nu^2 \quad , \tag{A2}$$

on a square grid over the image ν. Refer to Fig. A1. Let Δ be the grid spacing, and ϕ the projection angle as in §3.2. Then to a first approximation (A2) can be written as

$$s(\eta + \Delta, \xi, \phi) = s(\eta - \Delta, \xi, \phi) + \{ [2\Delta\nu(\eta, \xi)]^2 - [s(\eta, \xi + \Delta, \phi) - s(\eta, \xi - \Delta, \phi)]^2 \}^{\frac{1}{2}} \quad .$$
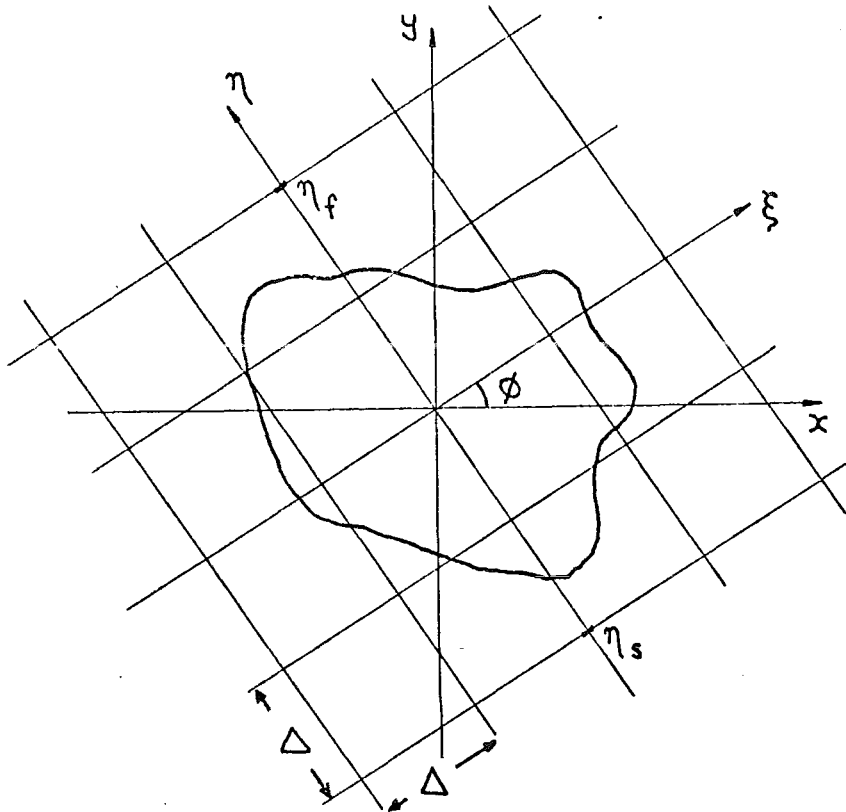
$$\tag{A3}$$

Fig. A1   The eikonal equation is solved on a grid
of points over the image of ν.

The initial conditions,

$$s(\eta_s - \Delta, \xi, \phi) = 0 \qquad\qquad (A4)$$

$$s(\eta_s, \xi, \phi) = \Delta\nu(\eta_s, \xi) \qquad\qquad (A5)$$

are specified and $s(\eta_s + \Delta, \xi, \phi)$ is calculated, using (A3), for all ξ. The variable η is incremented and the process is repeated until the image is traversed. The final value $s(\eta_f, \xi, \phi)$ is the estimate of the time-delay projection. The initial conditions given by (A4) and (A5) apply for an incident plane wave. However it was found that if the initial conditions were made to correspond to a disturbance emanating from a localised source, the quality of the reconstructed image was not significantly improved when ν did not depart by more than 10% from its mean value. By specifying the initial conditions as in (A4) and (A5) the time-delay projections, at a given angle ϕ, can be obtained for all values of ξ simultaneously - this is a considerable advantage in terms of computer time.

5. ASPECTS OF ELECTRICAL IMPEDANCE IMAGING

5.1 INTRODUCTION

Electrical impedance imaging is the process of determining volume conductivity distributions using low frequency electric fields. An electric potential is applied to the surface of an object, and the current flowing through the surface is measured (or vice versa). The goal is to reconstruct from these measurements the internal conductivity distribution of the object. For general conductivity distributions, exact solutions to this, the electrical impedance imaging problem, are not known. Consequently practical imaging schemes are approximate and largely ad hoc.

Electrical impedance imaging techniques have been used by geologists for many years (Keller and Frischnecht, 1966). However geological electrical impedance imaging methods still tend to be of an empirical nature; and generally they are based around the assumption that the structure of the earth is horizontally stratified (Inman *et al.*, 1973). Medical imaging is another field within which electrical impedance imaging procedures are applied. For instance, Henderson and Webster (1978) describe an "impedance camera" for imaging the thorax. Two grids of electrodes are placed on the thorax, one on the chest and the other on the back. By measuring the impedance between opposing electrodes on the front and back grids an "impedance image" is formed. From the impedance image it is possible, to a certain extent, to determine the locality of different conductivity regions within the chest - although the areas within the image corresponding to the different conductivity regions of the body are not sharply delineated. Perhaps the most interesting attempts at imaging conductivity distributions are those based on the reconstruction methods used in X-ray computed tomography (Price, 1979; Lytle and Dines, 1978; Schomberg, 1978). This imaging technique has been dubbed "impedance computed tomography" (I.C.T.) (Price, 1979). The material presented in this chapter arose out of an investigation into the theoretical aspects of I.C.T. The proponents of I.C.T. put forward plausible arguments, but they seem to have overlooked a basic physical limitation which prevents their methods from being successful in general.

In §5.2 various preliminaries relating to impedance imaging are collected. Then, in §5.3, the computed tomography approach to impedance imaging is discussed. §5.3 includes the arguments for why I.C.T. might work, and the limitation which prevents it from uniquely imaging conductivity distributions. Theoretical aspects of impedance imaging are examined in §5.4. Also, the types of independent measurements that can be made are outlined. In §5.5, a method for determining circularly symmetric conductivity distributions from impedance measurements, when the conductivity distribution is composed of a series of constant conductivity annular regions, is presented. The significance of the results contained in this chapter is assessed in §5.6.

The ideas presented here are of a preliminary nature. However further detailed work is being carried out by A.D. Seagar of the Electrical Engineering Department at this University. He is investigating solutions to the non-linear "imaging equations" discussed in §§5.4 and 5.5. He is hoping to apply his algorithms to the problem of "imaging" blood flow in the legs.

5.2 PRELIMINARIES

The equations describing the physical interaction between a static electric field $\underline{E}$ in a solid conducting object, with conductivity σ, are now introduced (see for example Stratton, 1941). The voltage (or potential) V is defined in terms of the electric field by

$$\underline{E} = - \nabla V \quad , \tag{5.1}$$

The current $\underline{J}$ in a conductor is related to $\underline{E}$ by

$$\underline{J} = \sigma \underline{E} \quad , \tag{5.2}$$

In the absence of free charges the current is conserved, i.e.

$$\nabla \cdot \underline{J} = 0 \quad . \tag{5.3}$$

These last three equations can be manipulated to give

$$\nabla^2 V = - \nabla V \cdot \nabla \tau \quad , \tag{5.4}$$

where $\tau = \nabla \ln \sigma$.

Now consider an impedance imaging system. For simplicity, the discussion in this chapter is confined to two-dimensional distributions of conductivity. However the results are easily extended to three dimensions. There is no fundamental difference between the analysis of static electric fields in two and three dimensions. Although, there are differences in detail - for instance, in the two-dimensional case, conformal transformations can be used within constant conductivity regions (see for example Churchill *et al.*, 1974). Refer to Fig. 5.1.
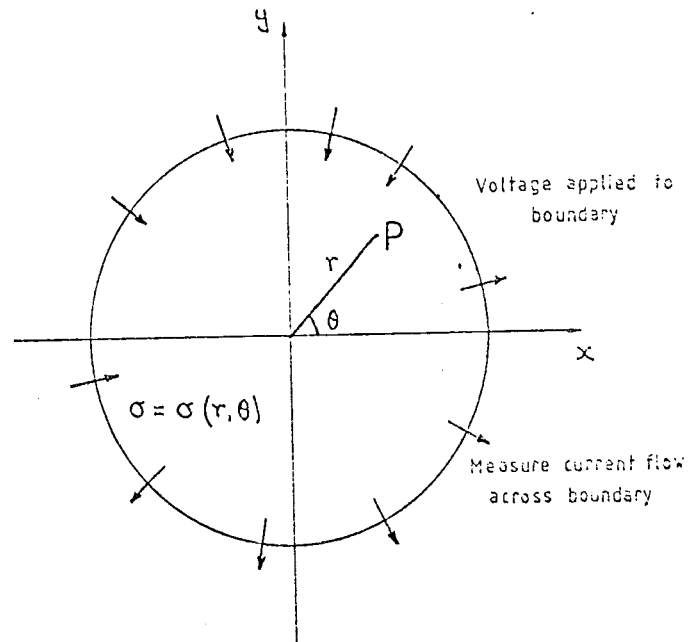
Fig. 5.1   Configuration of a simple impedance imaging system.

The object being imaged is confined to the unit circle. An arbitrary point P, within the unit circle, has polar coordinates (r, θ) and cartesian coordinates (x, y). The conductivity distribution of the object within the unit circle is $\sigma(r, \theta)$. Voltages are applied to the circumference of the unit circle, and currents flowing through the circumference are measured. Generally, to avoid polarisation effects at the electrodes, alternating voltages are used. However in all situations of biomedical interest, the frequency can be chosen low enough to make reactive effects negligible. Thus temporal variations of the voltage $V(r, \theta)$ and the current $\underline{J}(r, \theta)$ within the unit circle

can be neglected (see Plonsey and Fleming, 1969). In other words, (5.1) - (5.4), describing the interaction of static electric fields with a conducting body, can be used if the frequency is low enough. In this chapter it is assumed that the electric field is effectively static.

The problem of impedance imaging is to infer $\tau(r,\theta)$ from measurements of the current $J_c(\theta)$ flowing through the circumference of the unit circle, given the voltage $V_c(\theta)$ applied to the circumference. Note that

$$J_c(\theta) = -\sigma(1,\theta)\ \partial V(1,\theta)/\partial r \quad , \tag{5.5}$$

and

$$V_c(\theta) = V(1,\theta) \quad . \tag{5.6}$$

It is emphasised that it is not yet clear whether this problem has a unique solution in general.

5.3 COMPUTED TOMOGRAPHY APPROACH TO IMPEDANCE IMAGING

Several researchers (Price, 1979; Lytle and Dines, 1978; Schomberg, 1978), aware of the conceptual difficulty of interpreting J_c when σ can vary arbitrarily, have investigated an approach which has been called "impedance computed tomography" (Price, 1979). Their method involves applying a particular voltage distribution to the boundary of the object being imaged, such that current tends to flow along straight lines.

Consider two infinite parallel plates (see Fig. 5.2). The plate at $y = y_o$ has a potential V_o, and the potential on the plate at $y = y_o$ is $-V_o$. The two plates are separated by a medium with constant conductivity σ_o. Notice the current streamlines that are drawn on Fig. 5.2. The direction of these streamlines indicates the direction of current flow; and the separation of the streamlines is roughly inversely proportional to current density. The voltage distribution in the region $-y_o < y < y_o$, then, is given by

$$V = V_o\ y/y_o \quad . \tag{5.7}$$
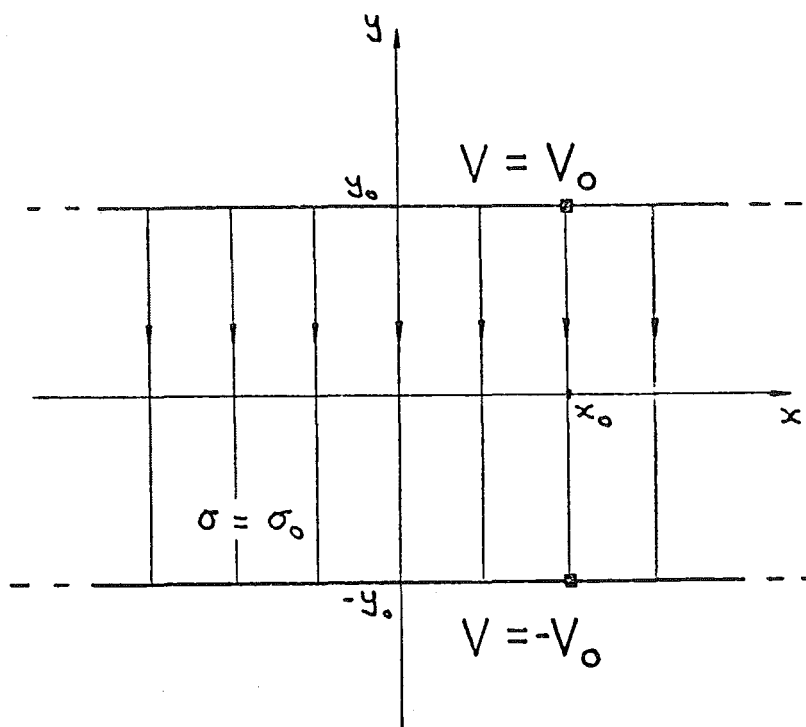
Fig. 5.2   Current distribution between two parallel plates
           separated by a constant conductivity medium.

Thus, from (5.1) and (5.2)

$$\underline{J} = - \sigma_o V_o \, \hat{\underline{y}}/y_o \quad , \tag{5.8}$$

where $\hat{\underline{y}}$ is the unit vector in the y direction. Now, in this case, the
current density is constant. Therefore an integral of the resistivity
ρ_o $(= 1/\sigma_o)$ between opposite points on respective plates, (y_o, x_o) and
$(-y_o, x_o)$ say, can be obtained straightforwardly. Denote the integral
of resistivity along a line at x_o, parallel to the y axis, by $R(x_o)$;
and the potential difference between the above two points by

$$\Delta V(x_o) = V(y_o, x_o) - V(-y_o, x_o) \tag{5.9}$$

$$= 2 V_o \quad . \tag{5.10}$$

Thus

$$R(x_o) = \Delta V(x_o)/(\underline{J}(y_o, x_o) \cdot \hat{\underline{y}}) \tag{5.11}$$

$$= 2 \, \rho_o y_o \quad . \tag{5.12}$$

Now suppose that an inhomogeneity is introduced into the region between the plates. The potentials of the plates are fixed, therefore ΔV is unchanged. But the current density becomes $\tilde{\underline{J}}$, say. If the current streamlines, within the current density distribution $\tilde{\underline{J}}$, remain parallel to each other; then R, defined by

$$R = \Delta V/\tilde{\underline{J}} \cdot \hat{\underline{y}} \quad , \tag{5.13}$$

represents an integral of the resistivity along the streamlines. However the introduction of an inhomogeneity effects the current density in a global manner. This means that, in general, the streamlines do not remain parallel to each other; and that the current density varies along a streamline. [Recall that the resistivity distribution could be determined using image reconstruction from projections algorithms (see §§1.8 and 3.2) if straight line integrals of the resistivity could be measured.] Now the proponents of impedance computed tomograpny (Price, 1979; Lytle and Dines, 1978; Schomberg, 1978) suggest that useful approximate images are obtained by treating R as being representative of straight line resistivity integrals. R is known as a "resistance projection". By treating the resistance projections in an analogous manner to the X-ray projections obtained with computerised tomography body scanners (§3.2), an "image" of resistivity (or conductivity) can be obtained. However the question is; are the curved streamlines "straight enough", and is the current density along streamlines "constant enough", that the resistance projections can be usefully treated as integrals of the resistivity along straight lines?

To investigate impedance computed tomography, consider the simple imaging configuration introduced in Fig. 5.1, §5.2. When $\sigma = \sigma_O$ is constant, parallel equally spaced current streamlines can be obtained by specifying V_c equal to

$$V_c = V_O \sin \theta \quad . \tag{5.14}$$

The current is then given by

$$\underline{J}(r,\theta) = -\sigma_O V_O \; (\hat{\underline{r}} \sin \theta + \hat{\underline{\theta}} \cos \theta) \quad , \tag{5.15}$$

where $\hat{\underline{r}}$ and $\hat{\underline{\theta}}$ are unit vectors in the r and θ directions respectively. The "resistance projection", in terms of the θ coordinate, using (5.13), is

$$\tilde{R}(\theta) = 2V_o \sin\theta/(\underline{J} \cdot \hat{\underline{y}}) \tag{5.16}$$

$$= 2\rho_o \sin\theta \quad . \tag{5.17}$$

Writing the resistance projection in terms of the x coordinate results in

$$R(x) = 2\rho_o (1 - x^2)^{\frac{1}{2}} \quad , \tag{5.18}$$

which is exactly equivalent to a straight line projection (see §1.8) through a circular cylinder of radius unity.

Now suppose that an inhomogeneity is introduced into the cylinder. To enable the results in this section to be presented in a simple analytic form, the "inhomogeneity" is taken to be a circular cylinder placed concentrically with respect to the first. Within the object being imaged, then, the conductivity is piecewise constant. Refer to Fig. 5.3.
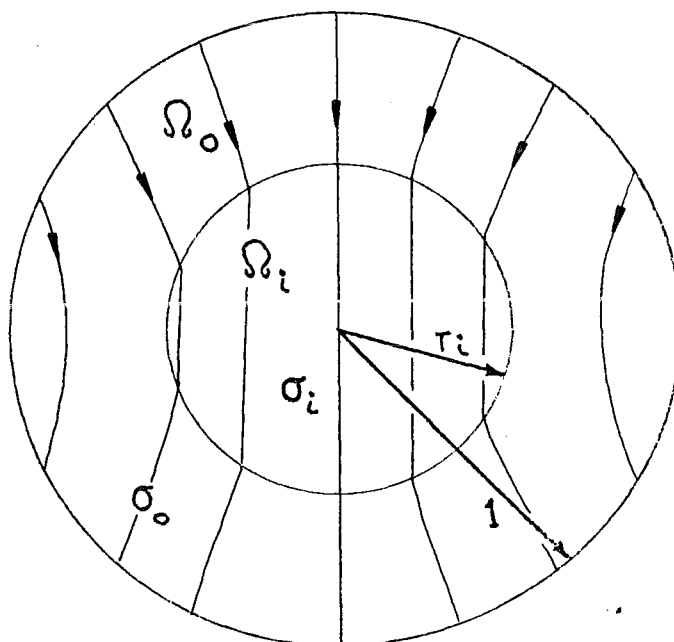

Fig. 5.3   Streamlines through two concentric
regions of different conductivity.

Denote by Ω_i the region $0 \leq r \leq r_i$, with conductivity σ_i; and the region $r_i < r \leq 1$, with conductivity σ_o, by Ω_o. Note that $\nabla \ln \tau = 0$ within Ω_o and Ω_i. Therefore, within Ω_o and Ω_i

$$\nabla^2 V = 0 \quad , \tag{5.19}$$

which is Laplace's equation. Solutions to (5.19), in two-dimensional polar coordinates, are $r^n \cos n\theta$, $r^n \sin n\theta$, $r^{-n} \cos n\theta$, $r^{-n} \sin n\theta$, $\ln r$, and $a = const.$; where n is any positive integer (see Morse and Feshback, 1953). Therefore the potential in Ω_o can be written in the general form

$$V_o(r,\theta) = \sum_{n=1}^{\infty} \{ a_n r^n \cos n\theta + b_n r^n \sin n\theta + c_n r^{-n} \cos n\theta$$

$$+ d_n r^{-n} \sin n\theta \} + e \ln r + f \quad . \tag{5.20}$$

However in Ω_i, because the potential must remain finite at $r = 0$, there are no $r^{-n} \cos n\theta$, $r^{-n} \sin n\theta$, and $\ln r$ terms. Therefore the potential in Ω_i is written in the particular form

$$V_i(r,\theta) = \sum_{n=1}^{\infty} \{ \alpha_n r^n \cos n\theta + \beta_n r^n \sin n\theta \} + \gamma \quad . \tag{5.21}$$

The unknown coefficients $a_n, b_n, c_n, d_n, e, f, \alpha_n, \beta_n$, and γ are determined from the boundary conditions. By satisfying (5.14) at $r = 1$, enforcing continuity of potential and current at $r = r_i$, and using the orthogonality of the sine and cosine terms, the potential is seen to satisfy

$$V_o(r,\theta) = V_o (r + \mu/r) \sin\theta/(1 + \mu) \quad , \tag{5.22}$$

where

$$\mu = r_i^2 (\sigma_o - \sigma_i)/(\sigma_o + \sigma_i) \quad . \tag{5.23}$$

The current density at the boundary is given by

$$\underline{J}(1,\theta) = -\sigma_o \nabla V(r,\theta) \Big|_{r=1} \tag{5.24}$$

$$= \{ \hat{\underline{r}}(1 - \mu) \sin\theta + \hat{\underline{\theta}}(1 + \mu) \cos\theta \} \sigma_o V_o/(1 + \mu) \quad . \tag{5.25}$$

Thus using (5.13), the resistance projection is

$$\tilde{R}(\theta) = \{2(1+\mu)\rho_o \sin\theta\}/\{1-\mu(\sin^2\theta-\cos^2\theta)\} \quad , \tag{5.26}$$

or

$$R(x) = \{2(1+\mu)\rho_o(1-x^2)^{\frac{1}{2}}\}/\{1-\mu(1-2x^2)\} \quad ; \tag{5.27}$$

where $\rho_o = 1/\sigma_o$. In order to produce useful images using impedance computed tomography it is necessary that the resistance projection (5.27) correspond reasonably closely to the "straight line projection" through the object being imaged. ["Straight line projection" refers to a projection obtained by integrating resistivity, in this case, along straight lines (see §1.8).] The straight line projection through the object shown in Fig. 5.3 is

$$S(x) = \begin{cases} 2\{\rho_o(1-x^2)^{\frac{1}{2}} + (\rho_i-\rho_o)(r_i^2-x^2)^{\frac{1}{2}}\} & ; \ |x| < r_i \\ \\ 2\{\rho_o(1-x^2)^{\frac{1}{2}}\} & ; \ r_i < |x| < 1 \quad , \end{cases} \tag{5.28}$$

where $\rho_i = 1/\sigma_i$. The significant difference between (5.27) and (5.28) is that the effect of the central cylinder is smeared over the whole resistance projection, whereas it is localised to $|x| < r_i$ within the straight line projection.

The effect of using resistance projections, defined by (5.27), to image resistivity distributions is now examined. Two reconstructed images are shown in Fig. 5.4. They were produced using the modified back projection algorithm (see for example Lewitt *et al.*, 1978; and §3.2). The resistivity values through the centre of both objects are shown as solid lines; and the resistivity values through the centre of both reconstructed images are shown as dashed lines. In Fig. 5.4a $\rho_o = 1$, $\rho_i = 0.25$ and $r_i = 0.5$; and in Fig. 5.4b $\rho_o = 1$, $\rho_i = 4$ and $r_i = 0.5$. Notice that there is no discernible feature in the images which might be suggestive of a boundary between Ω_o and Ω_i . The reason for this is fairly obvious when (5.23) and (5.27) are studied more closely. The information in the projection (5.27) regarding the conductivity and radius of the central cylinder is contained in only one parameter μ . And, from (5.23), μ can be constant for many different σ_i and r_i values. Thus there is inherent ambiguity in impedance computed tomography. All that can possibly be accomplished, without *a priori* information, is an estimate of the parameter μ .
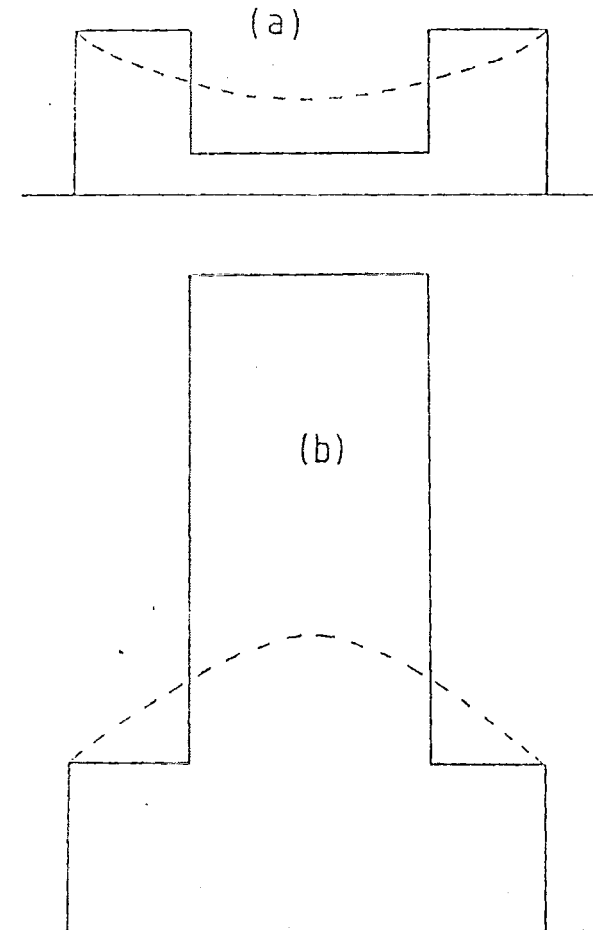
Fig. 5.4   Images obtained using resistance projections.

Solid line - actual object,

dashed line - reconstructed object.

5.4 SOME THEORETICAL ASPECTS OF IMPEDANCE IMAGING

In this section the types of independent measurements that can be made are discussed. Then the problem of determining the conductivity of an object from voltage and current measurements made around the boundary is examined.

Again consider the simple impedance imaging scheme shown in Fig. 5.1, and discussed in §§5.2 and 5.3. Notice that to obtain a resistance projection at angle ϕ, say (cf. §3.2), the potential distribution around the boundary is altered to

$$V_c(\theta, \phi) = V_o \sin(\theta - \phi) \quad . \tag{5.29}$$

Thus by varying ϕ, resistance projections at many different angles can be obtained. However for a given, general, conductivity distribution there are only two independent resistance projections, as is shown below.

$V_c(\theta, \phi)$ in (5.29) can be represented in terms of a linear combination of $\cos\theta$ and $\sin\theta$; i.e.

$$V_c(\theta, \phi) = V_o \cos\phi \sin\theta - V_o \sin\phi \cos\theta \tag{5.30}$$

Denote by $V(r, \theta, \phi)$ the voltage distribution within the unit circle when the potential around the boundary is given by (5.29). Now because of the principle of superposition and (5.30), $V(r, \theta, \phi)$ can be written as a linear combination of $V(r, \theta, 0)$ and $V(r, \theta, \pi/2)$. It follows that, in this case, there are only two independent current distributions. Hence only two independent resistance projections can be measured using boundary conditions of the form (5.29). It seems that in order to determine the conductivity distribution, many independent measurements are required. One way of obtaining independent voltage distributions throughout the unit circle is to specify V_c by

$$V_c = V_o \cos n\theta \tag{5.31}$$

and

$$V_c = V_o \sin n\theta \quad , \tag{5.32}$$

where n is any positive integer. All the different V_c defined by (5.31) and (5.32) are orthogonal around the circumference of the unit circle.

When V_c is defined by (5.31) and (5.32), it is useful to think of the voltage distribution existing within the unit circle as the "nth static mode".

Alternatively, it may be more practical to specify the current distribution J_c around the unit circle and then measure the voltage. To obtain independent measurements, J_c could be specified similarly to the V_c in (5.31) and (5.32). However it is considerably easier to use a current source with two point like electrodes - one injecting current and the other removing it (see Fig. 5.5).
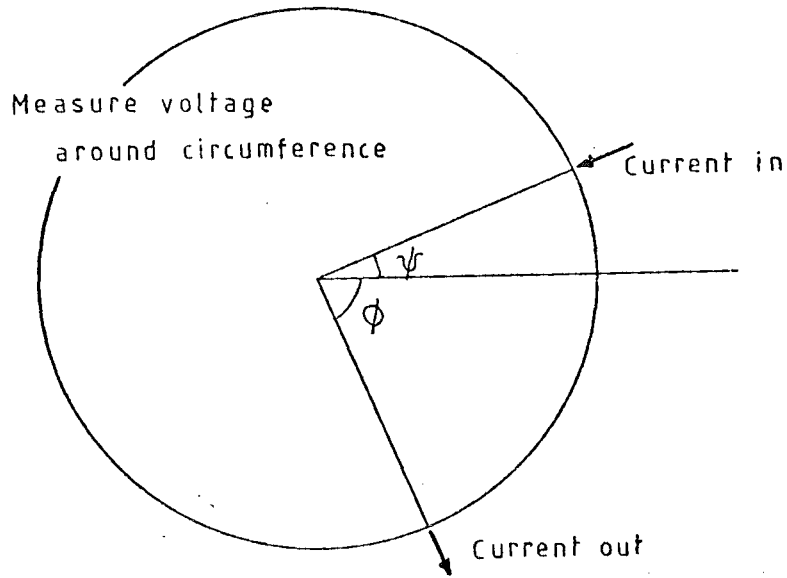
Fig. 5.5   Impedance imaging using a simple current source.

J_c can be represented, then, by two delta functions

$$J_c(\theta,\psi,\phi) = \delta(\theta-\psi) - \delta(\theta-\phi) \quad . \tag{5.33}$$

The positive delta function indicates that current is injected at $\theta = \psi$, and the negative delta function indicates the removal of current at $\theta = \phi$. Now consider the types of independent measurements that can be made using this simple current source. Using (5.33),

$$J_c(\theta,\psi,\phi) = \left(\delta(\theta-\psi) - \delta(\theta-\beta)\right) - \left(\delta(\theta-\phi) - \delta(\theta-\beta)\right) \tag{5.34}$$

$$= J_c(\theta,\psi,\beta) - J_c(\theta,\phi,\beta) \quad . \tag{5.35}$$

It can also be shown that

$$J_c(\theta,\psi,\phi) = J_c(\theta,\alpha,\phi) - J_c(\theta,\alpha,\psi) \quad . \tag{5.36}$$

Thus all the $J_c(\theta,\psi,\phi)$ are not independent. To obtain independent voltage distributions throughout the unit circle it is necessary to fix ψ, say, at some value ψ_o and vary ϕ over the range $0 < \phi < 2\pi$.

Now the problem of determining σ from J_c and V_c is examined. Suppose that τ is represented in terms of a series

$$\tau(r,\theta) = \sum_{n=-\infty}^{\infty} \sum_{m=0}^{\infty} a_{m,n} \; r^m \exp(jn\theta) \quad . \tag{5.37}$$

This series is sufficiently general to represent any continuously varying non zero conductivity distribution. Also represent the potential distribution within the unit circle in a similar manner, i.e.

$$V(r,\theta) = \sum_{p=-\infty}^{\infty} \sum_{q=0}^{\infty} b_{q,p} \; r^q \exp(jp\theta) \quad . \tag{5.38}$$

Insert these expressions for τ and V into (5.4). Then, using the orthogonality of the $\exp(jn\theta)$, and treating the terms associated with equal powers of r as being independent, a set of equations relating the $b_{q,p}$ to the $a_{m,n}$ is obtained. This is

$$(q^2 - p^2) \; b_{q,p} = \sum_{n=-\infty}^{\infty} \sum_{m=0}^{p} \{n(q-n) - m(p-m)\} \; b_{m,n} \; a_{p-m,q-n} \quad . \tag{5.39}$$

Now define V_{cp} and J_{cp} by

$$V_c(\theta) = \sum_{p=-\infty}^{\infty} V_{cp} \exp(jp\theta) \quad , \tag{5.40}$$

and

$$J_c(\theta)/\sigma(1,\theta) = \sum_{p=-\infty}^{\infty} J_{cp} \exp(jp\theta) \quad , \tag{5.41}$$

respectively. As $J_c(\theta)$ and $V_c(\theta)$ are known, the V_{cp} and J_{cp} can be determined (assume that $\sigma(1,\theta)$ is known). On the unit circle, $r = 1$,

$$V_{cp} = \sum_{q=0}^{\infty} b_{q,p} \quad , \tag{5.42}$$

and

$$J_{cp} = \sum_{q=0}^{\infty} q\, b_{q,p} \quad .$$

(5.43)

(These last two equations are obtained using the orthogonality of the exp(jnθ), and (5.5) and (5.6)). The imaging problem is "solved" by using (5.39), (5.42) and (5.43) to determine the $a_{m,n}$. Unfortunately because of (5.39) the equations are non-linear.

This approach to impedance imaging is not well understood, although it seems that iterative procedures are required. However it is not clear that a unique solution exists. This question, and problems associated with the numerical stability of any "solutions", have yet to be investigated.

5.5   IMPEDANCE IMAGING SOLUTION FOR CIRCULARLY SYMMETRIC CONDUCTIVITY DISTRIBUTIONS

Suppose that σ is circularly symmetric - i.e. $\sigma = \sigma(r)$ - but that it is "onion-like" in that it is constant within annular regions:

$$\sigma(r) = \sigma_m \quad , \quad \alpha_m < r < \alpha_{m+1} \quad , \quad 0 < m < M \quad .$$

(5.44)

where $\alpha_o = 0$ and $\alpha_M = 1$. Using an expression of the form (5.21) for the potential in the region $0 < r < a_1$, and expressions of the form (5.20) for the potential in the annular regions, and going through the procedure used to derive (5.22), it is found that

$$J_c = \begin{Bmatrix} \cos n\theta \\ \sin n\theta \end{Bmatrix} n\sigma_M [\Sigma_n(1,L) - \Sigma_n(L,2L)]/[\Sigma_n(1,L) - \Sigma_n(L,2L)] \quad ,$$

(5.45)

where the cos or sin is present depending on whether V_c is defined by (5.31) or (5.32) respectively. $L = 2^{M-1}$ and

$$\Sigma_n(j,k) = \sum_{\ell=j}^{k} \beta_{\ell n} A_\ell \quad .$$

(5.46)

The $\beta_{\ell n}$, which are all different, consist of products and quotients of the α_m. Each A_ℓ is of the form

$$A_\ell = \prod_{m=1}^{M} (\sigma_m \pm \sigma_{m-1}) \quad ,$$

(5.47)

where the combination of plus and minus signs is unique to the particular value of ℓ . By choosing an A_p and an A_q, say, that differ only with respect to the + or - sign in one position, A_p/A_q gives terms of the form σ_j/σ_k . If σ_M is known, it is clear that the remaining σ_m can be deduced from the A_ℓ . If the α_m are assumed known and if J_c is measured for V_c equal to 2L different $\cos n\,\theta$ and $\sin n\,\theta$ distributions, then (5.45) is a system of linear equations which are solvable for the A_ℓ .

Since (5.44), with $(\alpha_{m+1} - \alpha_m) = 1/(M+1)$, can usefully approximate virtually any circularly symmetric distribution of conductivity if M is large enough, it is clear from the above that the problem of impedance imaging has an effectively unique solution when σ is independent of θ . Notice that the unknowns in (5.45) are products of the σ_m . However by solving instead for the A_ℓ the problem is linearised. Although, unfortunately, this linearisation means that 2^M independent sets of measurements are required to obtain only M unknown σ_m.

5.6 DISCUSSION

It has been shown that it is impossible, using the straight-forward impedance computed tomography approach, to obtain unique images in general. The basic difficulty with impedance computed tomography is that it only attempts to excite a single "static mode" whereas arbitrary numbers of static modes are needed to characterise general distributions of conductivity.

It has also been demonstrated that any circularly symmetric distribution of conductivity can be unravelled if sufficient static modes are identified and measured. The simple analytic techniques invoked in §5.3 are inadequate when σ varies arbitrarily with r and θ . Even if it rurns out that the problem of impedance imaging has a unique solution in general, it seems that any algorithm for reconstructing $\sigma(r,\theta)$ will have to be iterative.

6.   CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

6.1   ON THE STRUCTURE OF D.N.A.

The most significant point to come out of the work discussed in Chapter 2 is that the available diffraction data is not sufficient to determine the structure of D.N.A. unequivocally. Producing good regular crystals of D.N.A. is exceedingly difficult. However until this is done, the credibility of any D.N.A. model, refined with respect to diffraction data, will remain questionable. Because of the large number of atoms in a repeat unit of D.N.A. it is very important to have high quality diffraction data.

Further, it seems that the theoretical considerations associated with the problem of determining molecular structure from diffraction measurements are not fully appreciated by many of the more applied crystallographic researchers. The task of determining the structure of large molecules uniquely from the diffraction data is no immense that success depends as much on analytic techniques as it does on skill in constructing crystals and making diffraction measurements.

6.2   IMAGING THE BEATING HEART WITH X-RAY C.T.

Two stroboscopic schemes are presented in Chapter 3. The "cardiac scanner", outlined in §3.3, should be investigated more fully from a practical point of view. It appears to be a rather straight-forward device, capable of taking high resolution images of the heart.

The heart imaging scheme which could be implemented on conventional C.T. scanners, §§3.4 and 3.5, requires clinical evaluation. Because the material within the heart has similar X-ray absorbtivity values, it is not known whether a dozen, or so, projections are sufficient to image the detailed structure of the heart. The simulations in §3.4 suggest that it would be possible to image the boundary between the ventricles and heart muscle, but this depends on the amount of contrast material in the blood.

As far as X-ray C.T. in general is concerned, the question of "required resolution" in a C.T. image should be reviewed. In the course

of this work the problem of image reconstruction using few projections
was briefly examined. Conventional scanners use 720 odd projections,
but quite reasonable images of the body can be obtained using 100 or so.
Having a "low resolution image" option on C.T. scanners would enable
images to be obtained at considerably reduced X-ray dosage levels to
the patient. Any artifacts in the image, caused by using few
projections, appear as straight lines. Now, provided that there is
enough dynamic range, these artifacts do not significantly reduce the
ability of the eyes and brain to interpret the image. In effect it is
like viewing a picture with shadows over certain regions. The brain can
easily take account of the "shadows". With sufficient dynamic range,
details under the "shadows" can be registered quite clearly.

6.3 ULTRASONIC TRANSMISSION TOMOGRAPHY

The lack of rigorous image reconstruction procedures has
handicapped ultrasonic transmission tomography. The results in Chapter
4 go some way towards alleviating this. It seems that in most cases
useful images could be obtained by simply assuming that the ultrasonic
rays are straight, and using standard X-ray C.T. image reconstruction
algorithms.

The main application of ultrasonic transmission tomography is
likely to be in mass screening for breast cancer. I feel that the next
step is to design and make a suitable device for performing this
function - the image reconstruction problem is "solved", the main
problems are in the mechanics of the device.

An interesting aspect of ultrasonic transmission tomography is
that, unlike echo-location procedures, a focused beam is not required.
Thus it is possible to make simultaneous propagation time measurements
in an analogous manner to the "fan beam" projections (see §3.2) in X-ray
C.T. This would considerably shorten the "scan time"; and does not
appear to have been exploited yet.

6.4 ELECTRICAL IMPEDANCE IMAGING

Electrical impedance imaging is much more complicated than the
imaging methods referred to in the previous three sections. There is
still much work to be done with regard to investigating theoretical

solutions to the imaging problem. Recent results achieved by
A.D. Seagar of the Electrical Engineering Department at this University,
show that the method is very insensitive to changes in conductivity.
For instance a large change in conductivity, over a small region towards
the centre of the object being imaged, has virtually no effect on the
current and voltage at the boundary. This means that measuring
equipment used has to be very sensitive, and accurately calibrated.

Still it is early days yet. Provided that the object is
relatively simple, and it is known a *priori* that this is the case,
one should be able to produce reasonably representative images using
electrical impedance measurements.

REFERENCES

Arnott S., and Hukins D.W.L. 1972. "Optimised Parameters for
    A-DNA and B-DNA", Biochemica Biophysica Research Comm.,
    Vol. 47, pp. 1504-1509.

Arnott S., and Hukins D.W.L. 1973. "Refinement of the structure
    of B-DNA and implications for the analysis of X-ray diffraction
    data from fibres of biopolymers", J. Mol. Biol., Vol. 81,
    pp. 93-105. [also, see above reference].

Bates R.H.T., and Peters T.M. 1971. "Towards improvements in
    tomography", New Zealand J. of Sci., Vol. 14, pp. 883-896.

Bates R.H.T. 1975. "Global solution to the scalar inverse scattering
    problem", J. Phys. A: Math. Gen., Vol. 8, No. 8, pp. L80-L82.

Bates R.H.T., Boerner W.M., and Dunlop R.G. 1976. "An extended Rytov
    approximation and its significance for remote sensing and inverse
    scattering", Opt. Commun., Vol. 18, pp. 421-423.

Bates R.H.T., and Wall D.J.N. 1976. "Chandrasekhar transformations
    improve convergence of computations of scattering from linearly
    stratified media", Trans. IEEE, Vol. AP-24, No. 2, pp. 251-253.

Bates R.H.T., and Dunlop G.R. 1977. "Inverse scattering and
    tomography", In: Ultrasonics International 1977 Conference
    Proc., (I.P.C. Science and Technology Press, Guilford, U.K.),
    pp. 104-110.

Bates R.H.T., Lewitt R.M., Rowe C.H., Day J.P., and Rodley G.A. 1977.
    "On the structure of DNA", J. Roy. Soc. New Zealand, Vol. 7,
    No. 3, pp. 273-301.

Bates R.H.T., and Wall D.J.N. 1977. "Null field approach to scalar
    diffraction", Phil. Trans. Roy. Soc. London, Vol. 287, No. 1339,
    pp. 45-114.

Bates R.H.T. 1978. "On phase problems: I and II", Optik, Vol. 51,
    No. 2, pp. 161-170, & No. 3, pp. 223-234.

Bates R.H.T., McKinnon G.C., and Millane R.P. 1978. "A new look at
        B-DNA diffraction data", Research Report, Electrical Engineering
        Dept, University of Canterbury, New Zealand.

Bates R.H.T., and McKinnon G.C. 1979. "Towards improving images in
        ultrasonic transmission tomography", Australasian Physical
        Sciences in Medicine, Vol. 2-3, pp. 134-140.

Bates R.H.T., and Heffernan P.B. 1980. "Image reconstruction from
        projections. V. Blurring due to object movement", Optik,
        Vol. 56, pp. 101-112.

Bates R.H.T., McKinnon G.C., Millane R.P., and Rodley G.A. 1980.
        "Revised interpretations of the available X-ray data for B-DNA".
        Pramana, Vol. 14, No. 4, pp. 233-252.

Berninger W.H., Redington R.W., Doherty P., Lipton M.J., and Carlsson E.
        1979. "Gated cardiac scanning: Canine studies", J. of Compt.
        Assist. Tomography, Vol. 3, No. 2, pp. 155-163.

Bloomfield V.A., Crothers D.M., and Tinoco I. 1974. "Physical
        chemistry of nucleic acids", Harper and Row, New York.

Bogren H.G., Lantz B.M.T., Miller R.R., and Mason D.T. 1977.
        "Effect of respiration on cardiac motion determined by
        cineangiography", Acta Radiologica Diagnosis, Vol. 18, No. 6,
        pp. 609-620.

Born M., and Wolf E. 1970. "Principles of optics", 4th Ed.,
        Pergamon Press, Oxford.

Bracewell R.N. 1965. "The Fourier transform and its applications",
        McGraw-Hill, New York.

Bracewell R.N. 1977. "Correction for collimator width (restoration)
        in reconstructive X-ray tomography", J. of Compt. Assist.
        Tomography, Vol. 1, No. 1, pp. 6-15.

Bram S. 1973. "Variation of type B-DNA X-ray fibre diagrams with
        base composition", Proc. Natl. Acad. Sci. USA, Vol. 70,
        pp. 2167-2171.

Bram S. 1977. Private Communications.

Brigham E.D. 1974. "The fast Fourier transform", Prentice-Hall, New Jersey.

Brooks R.A., and Di Chiro G. 1976. "Principles of computer assisted tomography (C.A.T.) in radiographic and radioisotopic imaging", Phys. Med. Biol., Vol. 21, pp. 689-732.

Bullen K.E. 1963. "An introduction to the theory of seismology", Cambridge University Press.

Churchill R.V., Brown J.W., and Verhey R.F. 1974. "Complex variables and applications", McGraw-Hill, Japan.

Cowley J.M. 1975. "Diffraction physics", North Holland/American Elsevier, Amsterdam.

Dines A.K., and Kak A.C. 1979. "Ultrasonic attenuation tomography of soft tissues", Ultrasonic Imaging, Vol. 1, pp. 16-33.

Dunlop R.G., Boerner W.M., and Bates R.H.T. 1976. "On the extended Rytov approximation and its comparison with the Born approximation", Univ. of Massachussetts at Amherst, IEEE Catalog No. 76 CH 1121-3AP.

Feughelman M., Langridge R., Seeds W.E., Stokes A.R., Wilson H.R., Hooper C.W., Wilkins M.H.F., Barclay R.K., and Hamilton L.D. 1955. "Molecular structure of deoxyribose nucleic acid and nucleoprotein", Nature, Vol. 175, pp. 834-838.

Friedlander F.G. 1958. "Sound pulses", Cambridge University Press.

Gelfand I.M., Graev M.I., and Vilenkin N.Ya. 1966. "Generalised functions: Vol. 5 Integral geometry and representation theory", Translated E. Saletan, Academic Press, New York/London.

Gibson D.G., Prewitt T.A., and Brown D.J. 1976. "Analysis of left ventricular wall movement during isovolumic relaxation and its relation to coronary artery disease", British Heart J., Vol. 38, pp. 1010-1019.

Glover G.H., and Sharp J.C. 1977. "Reconstruction of ultrasound propagation speed distributions in soft tissue: Time-of-flight tomography", Trans. IEEE, Vol. SU-24, pp. 229-234.

Górski F. 1975. "Critical review of the mechanisms separating the strands of DNA double helices during cell division", Folia Biologia (Krákow), Vol. 23, pp. 1-11.

Haimson J. 1979. "X-ray source without moving parts for ultra high speed tomography", Trans. IEEE, Vol. NS-26, pp. 2857-2861.

Harell G.S., Guthaner D.F., Breiman R.S., Morehouse C.C., Seppi E.J., Marshall W.H., and Wexler L. 1977. "Stop action cardiac computed tomography", Radiology, Vol. 123, pp. 515-517.

Harris F.J. 1978. "On the use of windows for harmonic analysis with the discrete Fourier transform", Proc. IEEE, Vol. 66, No. 1, pp. 51-83.

Henderson R.P., and Webster J.B. 1978. "An impedance camera for spatially specific measurements of the thorax", Trans. IEEE, Vol. BME-25, pp. 250-254.

Herman G.T. 1979. "Correction for beam hardening in computed tomography", Phys. Med. Biol., Vol. 24, No. 1, pp. 81-106.

Huang H.K., and Wu S.C. 1976. "Mass densities of the human body in vivo from C.T. scans", Compt. in Biol. and Med., Vol. 6, No. 4, pp. 337-343.

Inman J.R., Ryu J., and Ward S.H. 1973. "Resistivity inversion", Geophysics, Vol. 38, No. 6, pp. 1088-1108.

Johnston S.A., Greenleaf J.R., Samayoa W.A., Duck F.A., and Sjostrand J. 1975. "Reconstruction of three dimensional velocity fields and other parameters by acoustic ray tracing", In: 1975 Ultrasonics Symposium Proc. (IEEE Cat. #75 CHO 994-4SU), pp. 46-51.

Jones D.S. 1964. "The theory of electromagnetism", Pergamon Press, Oxford.

Jordan A.K., and Ahn S. 1979. "Inverse scattering theory and profile reconstruction", Proc. IEE, Vol. 126, pp. 945-950.

Kaveh M., Mueller R.K., and Iverson R.D. 1979. "Ultrasonic tomography based on perturbation solutions to the wave equation", Computer Graphics and Image Processing", Vol. 9, pp. 105-116.

Keller G.V., and Frischnecht F.C. 1966. "Electrical methods in geophysical prospecting", Pergamon Press, Oxford.

Kivisaari L., Kormano M., and Rantakokko V. 1979. "Contrast enhancement of the pancreas in computed tomography", J. Compt. Assist. Tomography, Vol. 3, No. 6, pp. 722-726.

Lewitt R.M., and Bates R.H.T. 1978. "Image reconstruction from projections. III: Projection completion methods (theory)", Optik, Vol. 50, pp. 268-728.

Lewitt R.M., Bates R.H.T., and Peters T.M. 1978. "Image reconstruction from projections. II: Modified back projection methods", Optik, Vol. 50, pp. 189-204.

Lonsdale K. 1968. "International tables for X-ray crystallography", Birmingham, U.K., Kynock Press, Vol. III, p. 201.

Lytle R.J., and Dines K.A. 1978. "An impedance camera: a system for determining the spatial variation of electrical conductivity", Lawrence Livermore Laboratory Report UCRL-52413.

Morse P.M., and Feshback H. 1953. "Methods of theoretical physics", McGraw-Hill, New York.

Mueller R.K., Kaveh M., and Wade G. 1979. "Reconstruction tomography and applications to ultrasonics", Proc. IEEE, Vol. 67, pp. 567-587.

Newton R.G. 1966. "Scattering theory of waves and particles", McGraw-Hill, New York.

Newton R.G. 1974. "The Gelfand-Levitan method in the inverse scattering problem", appears in Scattering Theory in Mathematical Physics - NATO advanced study series, Eds J.A. Lavita and J.-P. Marchand, D. Reidel, Holland, pp. 193-236.

Plonsey R., and Fleming D.G. 1969. "Biomedical phenomena",
    McGraw-Hill, New York.

Price.L.C. 1979. "Electrical impedance computed tomography (ICT):
    a new C.T. imaging technique", Trans. IEEE, Vol. NS-26,
    pp. 2736-2739.

Robb R.A., Ritman E.L., Gilbert B.K., Kinsey J.H., Harris L.N., and
    Wood E.H. 1979. "The DSR: a high speed three dimensional
    X-ray computed tomography system for dynamic spatial
    reconstruction of the heart and circulation", Trans. IEEE,
    Vol. NS-26, No. 2, pp. 2713-2717.

Rodley G.A., Scobie R.S., Bates R.H.T., and Lewitt R.M. 1976.
    "A possible conformation for the double stranded polynucleotides",
    Proc. Natl. Acad. Sci. U.S.A., Vol. 73, No. 9, pp. 2959-2963.

Sagel S.S., Weiss E.S., Gillard R.G., Hounsfield G.N., Jost R.G.T.,
    Stanley R.J., and Ter-Pogossian M.M. 1977. "Gated computed
    tomography of the human heart", Investigative Radiology,
    Vol. 12, No. 6, pp. 563-566.

Sasisekharan V., Pattabiraman N., and Gupta G. 1978. "Some
    implications of an alternative structure for DNA",
    Proc. Natl. Acad. Sci. USA, Vol. 75, pp. 4092-4096.

Schomberg H. 1978. An improved approach to reconstructive
    ultrasound tomography", J. Phys. D: Appl. Phys., Vol. 11,
    pp. L181-L185.

Schomberg H. 1978. "Reconstruction of spatial resistivity distribution
    of conducting objects from external resistance measurements",
    Philips GmbH Forschungslaboratorium, Hamburg, Manuskript NR:
    MS-H 1908V/78.

Sherwood D. 1976. "Crystals, X-rays and proteins", Longman, London.

Smith P.R., Peters T.M., and Bates R.H.T. 1973. "Image reconstruction
    from finite numbers of projections", J. Phys. A: Math, Nucl.
    Gen., Vol. 6, pp. 361-382.

Stanton L. 1969. "Basic medical radiation physics", Meredith
     Corporation, New York.

Stettler U.H., Weber H., Koller T., and Weissmann C. 1979.
     "Preparation and characterisation of form V DNA, the duplex
     DNA resulting from association of complementary circular
     single stranded DNA", J. Mol. Biol., Vol. 131, pp. 21-40.

Stratton J.A. 1941. "Electromagnetic theory", McGraw-Hill,
     New York.

Von Hippel A.R. (ed.) 1954. "Dielectric materials and applications",
     Technology Press of MIT/Wiley, New York.

Wade G., Mueller R.K., and Kaveh M. 1979. "A survey of techniques
     for ultrasonic tomography", In: Computer Aided Tomography
     in Ultrasonics and Medicine, eds J. Raviv, J.F. Greenleaf,
     and G.T. Herman, North Holland, Amsterdam, pp. 165-215.

Wang A.H.J., Quigley G.J., Kolpak F.J., Crawford J.L., Van Room J.H.,
     Van der Marel G., and Rich A. 1979. "Molecular structure of
     left handed double helical fragment at atomic resolution",
     Nature, Vol. 282, pp. 680-686.

Waterman P.C. 1965. "Matrix formulation of electromagnetic
     scattering", Proc. IEEE, Vol. 53, pp. 805-812.

Watson G.N. 1966. "A treatise on the theory of Bessel functions",
     Cambridge Univ. Press.

Watson J.D., and Crick F.H.C. 1953. "Molecular structure of nucleic
     acids: a structure for deoxyribonucleic acid", Nature, Vol. 171,
     pp. 737-738.

Wells P.N.T. 1977. "Biomedical Ultrasonics", Academic Press, London.

Wert C.A., and Thomson R.M. 1970. "Physics of solids", McGraw-Hill,
     Tokyo.

Wilkins M.H.F., Seeds W.E., Stokes A.R., and Wilson H.R. 1953.
     "Helical structure of crystalline deoxypentose nucleic acid",
     Nature, Vol. 172, pp. 759-762.

Wolf E. 1969. "Three dimensional structure determination of semi-transparent objects from holographic data", Optics Commun., Vol. 1, pp. 153-156.

Zimmerman S.B., and Pheiffer B.H. 1979. "Helical parameters of DNA do not change when DNA fibres are wetted: X-ray diffraction study", Proc. Natl. Acad. Sci. USA, Vol. 76, pp. 2703-2707.